

CRANFIELD UNIVERSITY

CHARLES MAZIVANHANGA

**FLOOD MODELLING APPROACHES FOR LARGE LOWLAND  
TROPICAL CATCHMENTS**

SCHOOL OF WATER, ENERGY AND ENVIRONMENT  
PhD in Water

PhD  
Academic Year: 2016 - 2019

Supervisor: Robert Grabowski  
Associate Supervisor: Tim Brewer  
October 2019



CRANFIELD UNIVERSITY

SCHOOL OF WATER, ENERGY AND ENVIRONMENT  
PhD in Water

PhD

Academic Year 2016 - 2019

CHARLES MAZIVANHANGA

FLOOD MODELLING APPROACHES FOR LARGE LOWLAND  
TROPICAL CATCHMENTS

Supervisor: Robert Grabowski  
Associate Supervisor: Tim Brewer  
October 2019

This thesis is submitted in partial fulfilment of the requirements for the  
degree of PhD

© Cranfield University 2019. All rights reserved. No part of this  
publication may be reproduced without the written permission of the  
copyright owner.



## ABSTRACT

Flooding is increasing in tropical regions, where millions of people are at risk, and challenges exist in providing reliable predictions and warnings. This research responds to this challenge by identifying and applying physics-based and data-based hydrological modelling approaches for large-scale flood modelling in lowland tropical regions. First, a distributed hydrological model was developed to accurately represent catchment conditions and processes in the model. Second, empirical data from nested catchments were analysed using statistical scaling relationships to complement the accuracy of peak discharge estimates. Finally, the effects of uncertainty propagation and interactions were quantified to increase the reliability of model results.

The research was conducted in the Grijalva catchment area (57 958 km<sup>2</sup>) southeast of Mexico. A large-scale model with a 2 x 2 km grid cell resolution was developed using the SHETRAN hydrological model and run enforced with 3-hour input rainfall data. Geostatistical techniques were used to quantify and reduce errors in input data, and all diverted flows were accounted for to optimise simulations. For the first time, the application of the Scaling theory of floods was applied in the study area to improve the estimation of peak discharge. A Monte Carlo technique was used to propagate and quantify rainfall and parameter uncertainties through a coupled hydrologic and hydraulic model and into model results.

Although the model under-predicted the magnitude of peak discharge, calibration results showed satisfactory model performance (NSCE = 0.72, CC = 0.74, Bias = -0.44% and RMSE 139.56 mm) and validation results were good (NSCE = 0.56, CC = 0.60, Bias = -6.3% and RMSE 62.59 mm). A statistical log-log relationship between intercepts ( $\alpha$ ) and peak discharge, from the smallest nested catchment, was used to complement the simulation of peak discharge magnitudes. It was observed that given rainfall uncertainties of  $\pm 71\%$ , ranging from 63 to 73%; the model generates discharge with uncertainties of  $\pm 46\%$ , ranging from 45 to 49% and errors of  $\pm 46\%$  ranging from 45 to 46%. The propagated uncertainties resulted in flood inundation extents of  $\pm 4.34$  km<sup>2</sup> varying from 1.66 to 7.02 km<sup>2</sup>.

Thus, flood modelling in large tropical regions can be achieved by optimally integrating several datasets with the best combination of the model parameter, input and output datasets based on uncertainty and error quantification and removal approaches.

**Keywords:**

Remote sensing, Flood risk, Scaling relationships, Scaling theory of floods, hydraulic modelling, Uncertainty propagation, Probabilistic flood inundation, Coupled hydrological models.

## **ACKNOWLEDGEMENTS**

First, I want to thank the Natural Environment Research Council (NERC) for funding this PhD research.

To my main supervisors Bob and Tim at Cranfield University, thanks for the opportunity to undertake this PhD study. Thanks for your unwavering support, persistence and trust, especially for your excellent guidance and leadership over the past three years.

I am also thankful to my co-supervisors, Steve and Greg at Newcastle University, for your continuous guidance for the opportunity to explore the mysteries of the SHETRAN hydrological model, for encouraging me to do that little extra work. To Chris, for your critical eye, your expert advice, and for asking the right questions and sharing your insights.

To Eunice and Victor at the Universidad Juárez Autónoma de Tabasco (UJAT), thanks for making my field trips to Villahermosa, Tabasco, Mexico an enriching and pleasant experience and taking the time to treat me more like a colleague than a student. Your support throughout and beyond this research is much appreciated.

I am also grateful to the members of my thesis committee, including my subject advisor, Dr Toby Waine, and to the independent chairman Professor. Naresh Magan for providing guidance, encouragement and suggestions in improving my ideas and solutions throughout this PhD study.

My most profound appreciation goes to my wife Colletta and daughter Tsungai, who stood by me through thick and thin - thanks for your unconditional love and support.

Finally, to my family and friends back in Zimbabwe, for your solid confidence in me and for reminding me that there is life outside the academic world. For helping me in so many ways, challenging me to think outside the box.





# TABLE OF CONTENTS

ABSTRACT.....	iii
ACKNOWLEDGEMENTS .....	v
LIST OF FIGURES.....	xii
LIST OF TABLES.....	xix
LIST OF EQUATIONS .....	xxiii
LIST OF ABBREVIATIONS.....	xxv
1 INTRODUCTION.....	29
1.1 Challenges in Tropical lowland regions.....	30
1.2 Large-scale, distributed hydrological modelling approaches .....	33
1.3 Scaling relationships.....	34
1.4 Uncertainty quantification.....	37
1.5 Knowledge gap and research needs.....	39
1.6 The study area.....	41
1.6.1 Villahermosa City and flood risk .....	42
1.6.2 La Sierra sub-catchment .....	43
2 AIMS AND OBJECTIVES.....	45
2.1 Thesis outline .....	45
3 LITERATURE REVIEW.....	47
3.1 Increased risk of flooding.....	47
3.1.1 The need for flood modelling.....	49
3.1.2 Installation of river gauge network and challenges .....	51
3.1.3 Challenges in tropical lowland regions .....	52
3.2 Opportunities for satellite data .....	57
3.2.1 Distributed hydrological models and satellite data.....	60
3.2.2 High-resolution modelling at a large-scale.....	67
3.2.3 Summary .....	69
3.3 The scaling theory of floods .....	71
3.3.1 The scaling theory of floods .....	71
3.3.2 Summary .....	75
3.4 Propagation of uncertainty through hydrological and hydraulic models ..	77
3.4.1 Uncertainties in flood modelling.....	77
3.4.2 Sources of uncertainty in flood modelling.....	78
3.4.3 Uncertainty propagation.....	79
3.4.4 Uncertainty analysis .....	80
3.4.5 Summary .....	83
4 METHODOLOGY.....	85
4.1 Study area .....	87
4.1.1 The main Grijalva catchment.....	87
4.1.2 La Sierra sub-catchment .....	92
4.2 The SHETRAN hydrological model and data requirements.....	94

4.2.1	Criteria for model selection.....	96
4.2.2	Data requirements and processing.....	99
4.2.3	Rainfall data processing.....	106
4.2.4	Geostatistical technique.....	115
4.2.5	Implementation of Geostatistical rainfall analysis.....	117
4.2.6	Rainfall correction.....	124
4.2.7	Thiessen polygons.....	125
4.2.8	Diverted flows.....	127
4.2.9	Assumptions relating to groundwater.....	129
4.3	SHETRAN hydrological modelling.....	131
4.3.1	Model runs and computing requirements.....	131
4.3.2	Sensitivity analysis.....	131
4.3.3	Model calibration.....	133
4.3.4	Capturing dam storage.....	137
4.3.5	Mass balance checks.....	138
4.3.6	Model validation.....	140
4.3.7	Model evaluation during calibration and validation.....	140
4.4	The Flood Modeller 2D model.....	143
4.4.1	The model selection criteria.....	143
4.4.2	Data requirements.....	144
4.5	Nested catchment modelling.....	147
4.5.1	Data availability.....	147
4.5.2	Defining nested catchments.....	148
4.5.3	The procedure for establishing scaling relationships.....	150
4.5.4	Hydrologic model simulations.....	152
4.5.5	Scaling relationships.....	153
4.5.6	Flood peak prediction framework.....	154
4.5.7	Measurement of uncertainty.....	156
4.5.8	Model selection.....	157
4.5.9	Estimation of parameters and peak discharge.....	158
4.6	Uncertainty propagation and analysis.....	159
4.6.1	The coupled or cascading mode setup.....	159
4.6.2	The meteorological level.....	160
4.6.3	The SHETRAN hydrological model and domain.....	163
4.6.4	The Flood Modeller 2D model.....	164
4.6.5	Model scenarios.....	168
4.6.6	Generation of parameter values.....	169
4.6.7	Historic flood levels and extents data.....	172
4.6.8	Uncertainty analysis.....	172
4.6.9	Measures of fit.....	173
4.7	Summary.....	176

5 RESULTS .....	177
5.1 Spatially distributed hydrological modelling.....	177
5.1.1 Rainfall error correction results.....	177
5.1.2 Sensitivity analysis results.....	185
5.1.3 Discharge output error correction.....	188
5.1.4 Calibration and validation results.....	191
5.2 Summary .....	196
5.3 Flood scaling relationships and estimation of peak discharge.....	197
5.3.1 Peak discharge scaling relationships .....	197
5.3.2 The effects of rainfall variability on scaling parameters .....	199
5.3.3 Flood scaling parameter estimation for La Sierra catchment .....	209
5.3.4 Intercept estimation from observed peak discharge .....	211
5.3.5 Exponent ( $\theta$ ) estimation using Intercept ( $\alpha$ ) values.....	213
5.3.6 Intercept ( $\alpha$ ) multi regressed with peak discharge .....	214
5.3.7 Exponents ( $\theta$ ) multi regressed with peak discharge .....	216
5.3.8 Significant statistical models developed .....	217
5.3.9 Model selection .....	217
5.3.10 Application of the results of the flood simulation framework .....	218
5.4 Summary .....	223
5.5 Uncertainty propagation and analysis results.....	224
5.5.1 Probability distribution fitting and rainfall ensemble generation .....	224
5.5.2 Calibration results of the flood inundation model .....	226
5.5.3 Rainfall uncertainty propagation results.....	229
5.5.4 Parameter uncertainty propagation results.....	234
5.5.5 Effects of high Strickler parameter values .....	234
5.5.6 Effect of low Strickler parameter values .....	239
5.5.7 Effect of high AET/PET parameter values .....	243
5.5.8 The effect of low AET/PET parameter values.....	247
5.5.9 Effect of high saturated conductivity parameter values.....	251
5.5.10 The effect of low saturated conductivity parameter values .....	255
5.5.11 The effect of high saturated water content parameter values .....	257
5.5.12 The effect of low saturated water content parameter values.....	261
5.6 Summary .....	264
6. DISCUSSION .....	265
6.1. Distributed hydrological modelling.....	265
6.1.1. High-resolution distributed modelling approaches .....	266
6.1.2. Full use of satellite datasets for the SHETRAN model .....	268
6.1.3. Application of SHETRAN distributed hydrological model.....	269
6.1.4. Rainfall error correction and distributed modelling .....	272
6.1.5. A greater understanding of the dynamics of tropical flood flows	274
6.1.6. Implications of dominant hydrological processes .....	275

6.2.	Scaling relationships in nested catchments .....	278
6.2.1.	Expanding the scaling theory of floods in tropical regions .....	278
6.2.2.	Evidence of power-law relationships .....	280
6.2.3.	Rainfall catchment coverage and flood parameters .....	281
6.2.4.	Flood parameter and peak discharge estimation.....	282
6.2.5.	Implications of scaling relationships .....	284
6.2.6.	Climate change adaptation .....	285
6.3.	Uncertainty propagation .....	286
6.3.1.	Explicit quantification of uncertainties .....	286
6.3.2.	Uncertainty quantification and distribution .....	287
6.4.	Discussion summary .....	292
6.5.	Recommendations for future research .....	294
6.5.1.	Large-scale, distributed hydrological modelling.....	294
6.5.2.	Scaling relationships .....	295
6.5.3.	Uncertainty quantification.....	296
6.6.	Conclusions .....	297
7.	REFERENCES .....	302
8.	APPENDICES.....	352
	Appendix A Rainfall input data .....	352
	Appendix B : Diverted flows calculations.....	361
	Appendix C : Simulated and observed discharge.....	381
	Appendix D : Nested catchment modelling .....	382
	Appendix E : Comparison of simulated and observed discharge .....	385



## LIST OF FIGURES

Figure 1.1: The projected probability of flood occurrence by 21st century for flood volumes related to flooding with a 100 year-return period.....	29
Figure 2.1: The thesis outline divides the thesis report into seven chapters. ...	46
Figure 3.1: (a) The number of flood events per income (bar) and (b) number of countries reporting flood events from 1960 to 2013 in the Emergency Events Database (EM-DAT).....	49
Figure 3.2: Global natural disasters between 1900 and 2010.....	50
Figure 3.3: Distributed hydrological model representation of catchment conditions, properties and processes in discrete grid cells. ....	61
Figure 3.4: Synergism between satellite data and distributed hydrological modelling for detailed representation of catchment conditions and process. ....	63
Figure 4.1: The modelling methodology, adapted from Anderson and Woessner (1992).....	86
Figure 4.2: The location of the main Grijalva river catchment and La Sierra sub-catchment areas in the southeast of Mexico. The main river network and major dams in the areas are shown.....	87
Figure 4.3: The mean monthly rainfall-runoff regime and seasons in the Grijalva River basin .....	89
Figure 4.4: The components of the SHETRAN hydrological model .....	95
Figure 4.5: Digital Elevation Model (DEM) for Grijalva catchment at 2 x 2 km grid resolution.....	103
Figure 4.6: Land cover types at 2 x 2 km grid map for the SHETRAN parameters .....	104
Figure 4.7: The soil grid map and categories used for setting up the SHETRAN hydrological model .....	105
Figure 4.8: The distribution of dams, rain gauge and river gauging stations and associated river network in the Grijalva catchment. ....	108
Figure 4.9: Time-averaged map of mean daily rainfall over 2013 to 2017 period at 0.25 degrees spatial scale (Approx. 25 x 25 km) [TRMM_3B42_Daily v7] in Grijalva, Mexico. ....	109
Figure 4.10: The mean monthly rain gauge rainfall distribution interpolated at 2 x 2 grid-scale (over 2013 to 2017 period) before merging. ....	112

Figure 4.11: The average monthly rain gauge rainfall distribution interpolated and upscaled at 2 x 2 grid-scale (over 2013 to 2017 period) before merging and compared to satellite rainfall below.....	113
Figure 4.12: The mean monthly satellite rainfall grid map (3B42V7) downscaled at 2 x 2 km grid-scale (over 2013 to 2017 period) before merging and compared to rain gauge rainfall above. The rainfall product underestimated rainfall amounts in the area and overestimate the Pacific coastal mean rainfall.....	113
Figure 4.13: Merging process of rain gauge and satellite rainfall datasets.....	114
Figure 4.14: Histogram of some of the rain gauge rainfall data after Log-normal and Box transformation. Data now has increased symmetry. Several rainfall maps in this study required Log-normal transformation, also called Log-normal Kriging. ....	117
Figure 4.15: A QQ plot showing the distribution of rain gauge rainfall data against the expected normal distribution where observations should lie approximately along a straight line. ....	118
Figure 4.16: A semi-variogram/Covariance model summarising the spatial variation of rainfall in the study area and how the variations change with the increasing separation distance between rain gauges .....	119
Figure 4.17: Summary of model diagnostics statistics on CoKriging prediction performance showing that the model can provide the best estimates of rainfall interpolation. ....	123
Figure 4.18: The lower Grijalva River showing a network of distributaries and tributaries and the location of gauging stations used in calculating diverted flows.....	127
Figure 4.19: Model calibration and sources of errors.....	133
Figure 4.20: A SWAT Calibration Helper v1.0 was modified and used to visualise and compare different plots of the SHETRAN water balance components and aid the calibration and validation of the model .....	139
Figure 4.21: Nested catchments and gauge network in the La Sierra catchment area for the application of scaling relationships between peak discharge and drainage areas. ....	149
Figure 4.22: Statistical flood scaling relationship analysis procedure to estimate scaling parameters and peak discharge in the La Sierra catchment. ....	150
Figure 5.1 Interpolated monthly total rain gauge rainfall distribution in Grijalva catchment highest totals occurring along northern slopes of La Sierra Mountains and covering most of Tabasco lowland areas while the central depressions of Chiapas have the lowest rainfall totals.....	179

Figure 5.2 Spatial distribution of mean monthly rainfall in the Grijalva catchment based on the Cokriging interpolation and merging of rain gauge rainfall, satellite rainfall and topographic datasets.....	180
Figure 5.3 Rainfall error distribution showing spatially varying error rainfall error fields in the Grijalva catchment area.....	181
Figure 5.4 The interpolated error fields derived from the Cokriging with external drift technique were integrated with SHETRAN Thiessen polygons.....	184
Figure 5.5: Percentage change in total discharge due to increases in parameter values.....	186
Figure 5.6 Spatially averaged flow totals over 2013 to 2017 simulation period for mass balance checks in La Grijalva catchment .....	187
Figure 5.7 Percentage fraction of seasonal and monthly-diverted flows in the lower reaches of the Grijalva River. ....	189
Figure 5.8: Calibration results flows-with diverted flows included .....	190
Figure 5.9: Calibration results using corrected flows- diverted flows removed	190
Figure 5.10: Comparison of simulated and observed discharge for the 2013 to 2015 calibration period .....	192
Figure 5.11: Comparison of simulated and observed discharge for the validation period 2016 to 2017.....	193
Figure 5.12: Peak discharge scaling relationships between peak discharge and nested catchment drainage areas experiencing complete rainfall coverage. Dotted lines show the upper and lower limit of the 95% confidence interval for estimated discharge values. ....	200
Figure 5.13: Peak discharge scaling relationships for partial rainfall coverage of catchments for rainfall events 2012-2015. Dotted lines show the upper and lower limit of the 95% confidence interval for all discharge values. ....	201
Figure 5.14: Peak discharge scaling relationship in catchments with both complete coverage and partial coverage rainfall.....	201
Figure 5.15: The scaling relationships between drainage areas and peak discharge from rainfall events that occurred in dry (a) and in wet (b) season. ....	204
Figure 5.16: Effect of rainfall location and direction on flood scaling parameters. Storms moving from NE have high exponent ( $\theta$ ) parameters and low intercepts ( $\alpha$ ) while NW storms have low exponents ( $\theta$ ) and high intercept ( $\alpha$ ) values.....	206
Figure 5.17: Results showing the relationships between simulated peak discharge and drainage areas based on the duration of rainfall events in La Sierra catchment area. ....	208



Figure 5.18: Scatter plot of the natural logarithm of the scaling intercept ( $\alpha$ ) against the natural logarithm of peak discharge observed at Puyacatengo Gauging Station. The 95% prediction interval estimates where the actual intercept ( $\alpha$ ) values are expected along the best fit line. ....	211
Figure 5.19: Scatter plot of the scaling exponent ( $\theta$ ) against the natural logarithm of the scaling intercept ( $\alpha$ ) .....	213
Figure 5.20: Comparison of observed and expected scaling Intercept ( $\alpha$ ) parameter values from 2012 to 2015 rainfall-runoff events. The grey dotted lines show the 95% prediction interval range where actual value can be predicted, and the orange line is the confidence interval showing precision of the regression model. ....	220
Figure 5.21: Comparison of estimated exponent ( $\theta$ ) and observed values from rainfall-runoff events occurred from 2012 to 2015. The grey dotted lines show the 95% prediction interval range where actual value can be predicted, and the orange line is the confidence interval showing precision of the regression model. ....	221
Figure 5.22: Comparison results between simulated discharge and observed discharge in La Teapa sub-catchment. ....	222
Figure 5.23 The outcome of the calibration process analysis showing changing flood extents to changes in Manning's coefficient values .....	226
Figure 5.24: The simulated flood extents (light green) and satellite-derived contours (blue) underline the adequate capacity of the calibrated model to replicate the actual observed flooded area. ....	228
Figure 5.25: Rainfall amounts have high variability in the 5% probability confidence level than for the upper (up to 95%) confidence limit. ....	230
Figure 5.26: The variability of discharge output ensembles within 5% and 95% confidence limits based on fixed/calibrated parameter scenario. ....	231
Figure 5.27: The 2015 hydrograph showing 5% and 95% confidence limits to represent uncertainty discharge due to fixed/calibrated parameter values. The red line represents the ensemble mean of discharge while the black is the observed discharge. ....	232
Figure 5.28: Probabilistic flood inundation outlines in the 5% and 95% confidence limits based on ensemble model runs based on fixed/calibrated parameter values.....	233
Figure 5.29: The uncertainty in the discharge output was relatively evenly distributed around the mean with large bands in less frequent probabilities than in the more frequent probabilities.....	235
Figure 5.30: The 2015 hydrograph showing 5% and 95% confidence limits in grey to represent uncertainty in discharge due to high Strickler parameter values	

for each vegetation type. The red line represents the discharge ensemble meanwhile the black is the observed discharge.....	237
Figure 5.31: Probabilistic flood inundation extent outlines in the 5% and 95% confidence limits based on ensemble model runs based on high Strickler parameter value scenario .....	238
Figure 5.32: The uncertainty in the discharge output was relatively evenly distributed around the mean and it has large bands in more frequent probabilities than in the less frequent probabilities.....	240
Figure 5.33: The 2015 hydrograph showing 5% and 95% confidence to represent uncertainty in discharge due to low Strickler parameter values for each vegetation category. The red line represents the ensemble mean of discharge while the black is the observed discharge. ....	241
Figure 5.34: Probabilistic flood inundation extent outlines in the 5% and 95% confidence limits based on ensemble model runs based on low Strickler parameter values for each vegetation type. ....	242
Figure 5.35: The uncertainty in the discharge output was not evenly distributed around the mean and it has large bands in more frequent probabilities than in the less frequent probabilities .....	243
Figure 5.36: The 2015 hydrograph showing the 5% and 95% confidence limits representing uncertainty in discharge due to high AET/PET parameter values for each vegetation type. The red line represents the mean ensemble of discharge while the black is the observed discharge. ....	244
Figure 5.37: Probabilistic flood inundation outlines in the 5% and 95% confidence limits based on ensemble model runs on high AET/PET parameter values for each vegetation type .....	246
Figure 5.38: The variability of discharge output due to low AET/PET parameter values was relatively evenly distributed, more frequent probabilities than the less frequent probabilities.....	247
Figure 5.39: The 2015 hydrograph showing 5% and 95% confidence limits (in grey) representing uncertainty in discharge due to low AET/PET parameters values for each vegetation type. The red line represents the ensemble mean, and the black is the observed discharge.....	248
Figure 5.40: Probabilistic flood inundation extents within the 5% and 95% confidence limits from ensemble model runs based on low AET/PET parameter values for each vegetation type. ....	250
Figure 5.41: The 2015 hydrograph showing 5% and 95% confidence limits to represent the uncertainty of discharge due to high saturated conductivity values for each soil type in the study area. ....	252

Figure 5.42: Probabilistic flood inundation outlines in the 5% and 95% confidence limits based on flow input into ensemble model runs based on high saturated conductivity values for each soil type in the study area.....254

Figure 5.43: The flood hydrography showing 5% and 95% confidence limits to representing the uncertainty of discharge due to high saturated water content parameter values. The red line represents the ensemble mean of discharge while the black is the observed discharge.....258

Figure 5.44: The 2015 hydrography showing 5% and 95% confidence limits to represent uncertainty discharge due to low saturated water content parameter values. The red line represents the ensemble mean of discharge while the black is the observed discharge.....262



## LIST OF TABLES

Table 3:1: Summary of ongoing and future hydrological modelling missions...	58
Table 4:1: Characteristics of nested catchment areas in La Sierra main catchment.....	93
Table 4:2 The standard questions and guidelines for model selection.....	97
Table 4:3 The primary data sources and types required for the SHETRAN hydrological modelling. ....	100
Table 4:4 The upper, middle and lower values of parameters used during calibration, including the original values, calibration/sensitivity range applied and final calibrated values as used.....	136
Table 4:5 Interpretation of statistical indicators for model calibration and validation .....	142
Table 4:6 Discharge data availability for the nested catchment areas. ....	148
Table 4:7 The primary variables used in Pearson's Correlation analysis .....	155
Table 4:8 Details of selected rainfall stations in La Sierra catchment .....	162
Table 4:9 The upper, lower and calibrated parameter values for each land cover and soil types used during uncertainty analysis and scenarios .....	170
Table 4:10 Recorded flow and levels for Las Gaviotas and nearby gauging stations on 27th October 2015.....	172
Table 5:1. The quantified errors between the merged rainfall dataset and original rain gauge rainfall data show that errors varied seasonally with high error fields prevalent during the wet season between (May to October) compared to the dry season.....	182
Table 5:2 Mean monthly rainfall errors per gauging station derived from the differences between merged rainfall dataset (rain gauge and TRMM rainfall) and original rain gauge rainfall in the Grijalva catchment. ....	183
Table 5:3 Percentage improvement in model simulations after removing the proportion of diverted flows.....	191
Table 5:4 The evaluation results for SHETRAN model performance ratings for the period 2013 to 2015 .....	192
Table 5:5 Studies that used SHETRAN for simulating river discharge and catchment size in comparison to the present study .....	195
Table 5:6 Results from scaling relationships between peak discharge from 59 rainfall events (that occurred in the La Sierra catchment) and nested	

catchment drainage areas showing scaling intercepts ( $\alpha$ ), exponents ( $\theta$ ) and coefficient of determination ( $R^2$ ) values. ....	198
Table 5:7 Summary of power-law relationships between the drainage areas and peak discharge from rainfall events that completely, partially and entirely cover catchments from 2012 to 2015.....	202
Table 5:8 Statistical significance of scaling relationships between complete, partial and both complete and partial rainfall coverage in La Sierra catchment .....	203
Table 5:9 Statistical significance of scaling relationship from complete catchment rainfall coverage and catchment areas in the wet (a) and dry (b) seasons. ....	203
Table 5:10 Pearson correlation analysis results showing significant and non-significant relationships between the exponents ( $\theta$ ) and intercept ( $\alpha$ ) (flood parameters) and other measurable catchment processes and conditions ( independent variables) in La Sierra catchment.....	210
Table 5:11 Results of the log-log relationship between Intercept ( $\alpha$ ) parameters and Puyacatengo peak discharge .....	212
Table 5:12 Results of the linear regression analysis between Exponents ( $\theta$ ) and Intercept ( $\alpha$ ) parameters .....	214
Table 5:13 Results from multilinear regression analysis of relationships between Intercept ( $\alpha$ ) parameters, Puyacatengo and Teapa peak discharge.....	215
Table 5:14 Results of the multilinear regression analysis between Puyacatengo, Gaviota peak discharge and exponent ( $\theta$ ) parameter. ....	216
Table 5:15 Statistical models developed for estimating scaling parameters. ..	217
Table 5:16 Model selected based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) estimators. ....	218
Table 5:17. Combined scores in comparing and selecting Probability distributions. ....	225
Table 5:18 Probability distribution parameter values for all rainfall gauging stations in La Sierra catchment. ....	225
Table 5:19 Rainfall input, discharge, flood extent and uncertainty quantification and comparison based on the calibrated model scenario. ....	229
Table 5:20 Rainfall, discharge and flood extents uncertainty and error quantification and comparison values. ....	236
Table 5:21 Rainfall input, discharge output and flood inundation extents, uncertainty and error quantification and comparison between calibrated and low Strickler parameter value values for each vegetation type.....	239

Table 5:22 Rainfall, discharge and flood inundation extents error and uncertainty quantification and comparisons due to high AET/PET parameter value values for each vegetation type. ....	244
Table 5:23 Rainfall, discharge, flood extents error and their uncertainty quantification including comparisons between calibrated parameter and low AET/PET parameter values for each vegetation type .....	248
Table 5:24 Rainfall, discharge and flood inundation extents uncertainty and errors quantification and comparisons between calibrated and high saturated conductivity values for each soil type in the study area.....	251
Table 5:25 Rainfall, discharge and flood extents error and uncertainty quantification comparisons between calibrated and low saturated conductivity parameter values for each soil type in the study area. ....	255
Table 5:26 Rainfall, discharge and flood inundation extents uncertainty and error quantification and comparisons between calibrated and high saturated water content parameter values for each soil type in the study area. ....	258
Table 5:27 Rainfall, discharge, flood extents uncertainty and error quantification and comparison between calibrated and low saturated water content parameter values for each soil type in the study area. ....	261





## LIST OF EQUATIONS

Equation 3.1 .....	71
Equation 4.1 .....	106
Equation 4.2 .....	115
Equation 4.3 .....	115
Equation 4.4 .....	116
Equation 4.5 .....	118
Equation 4.6 .....	120
Equation 4.7 .....	121
Equation 4.8 .....	121
Equation 4.9 .....	122
Equation 4.10 .....	122
Equation 4.11 .....	122
Equation 4.12 .....	124
Equation 4.13 .....	125
Equation 4.14 .....	126
Equation 4.15 .....	128
Equation 4.16 .....	128
Equation 4.17 .....	128
Equation 4.18 .....	132
Equation 4.19 .....	138
Equation 4.20 .....	140
Equation 4.21 .....	141
Equation 4.22 .....	141
Equation 4.23 .....	141
Equation 4.24 .....	154
Equation 4.25 .....	157
Equation 4.26 .....	157
Equation 4.27 .....	161

Equation 4.28.....	173
Equation 4.29.....	174
Equation 4.30.....	175
Equation 5.1.....	212
Equation 5.2.....	214
Equation 5.3.....	215
Equation 5.4.....	216

## LIST OF ABBREVIATIONS

AET	Actual Evapotranspiration
AIC	Akaike Information Criterion
BANDAS	National Surface Water Data Bank
BIC	Bayesian Information Criterion
CC	Correlation Coefficient
CEH	Centre for Ecology & Hydrology
CLICOM	Mexican National Climatological Database
CONAGUA	Mexican National Water Commission
CRED	Centre for Research on the Epidemiology of Disasters
DDS	Dimensioned Search Algorithm
DEM	The Digital Elevation Model
DTM	Digital Terrain Model
EM-DAT	Emergency Events Database
FAC	Flow Accumulation Grid
FDR	Flow Direction Grid
GASIR	Mexican Surface Water Management and Rivers Engineering
GBHM	Geomorphology-Based Hydrological Model
GCEW	Goodwin Creek Experimental Watershed
GDAS	Global Data Assimilation System
GIS	Geographical Information Systems
GPM	Global Precipitation Measurement
GPU	Graphics Processing Unit
ICHARM	International Centre for Water Hazard and Risk Management
INEGI	Mexican National Institute of Statistics and Geography
IPCC	Intergovernmental Panel on Climate Change
IT	Information Technology
ITCZ	Inter-Tropical Convergence Zone
JCGM	General Conference on Weights and Measures
JCGM	Joint Committee for Guides in Metrology
KED	Cokriging with External Drift
NSE	Nash-Sutcliffe
OFTAT	One Factor At a Time

OSTRICH	Optimisation Software Toolkit for Research Involving Computational Heuristics
PBIAS	Percentage Bias
PDF	Probability Density Function
PET	Potential Evapotranspiration
PIK	Potsdam Institute for Climate Impact Research
RMSE	Relative Mean Square Error
SHE	Systeme Hydrologique European
SMAP	Soil Moisture Active and Passive
SMOS,	Soil Moisture and Ocean Salinity
SRTM	Shuttle Radar Topography Mission
SRTM	Shuttle Radar Topography Mission
TM	Landsat Thematic Mapper
TMPA	TRMM Multi-satellite Precipitation Analysis
TRMM	Tropical Rainfall Measuring Mission
TVD	Total Variation Diminishing
VIC	Variable Infiltration Capacity
WHO	World Health Organisation



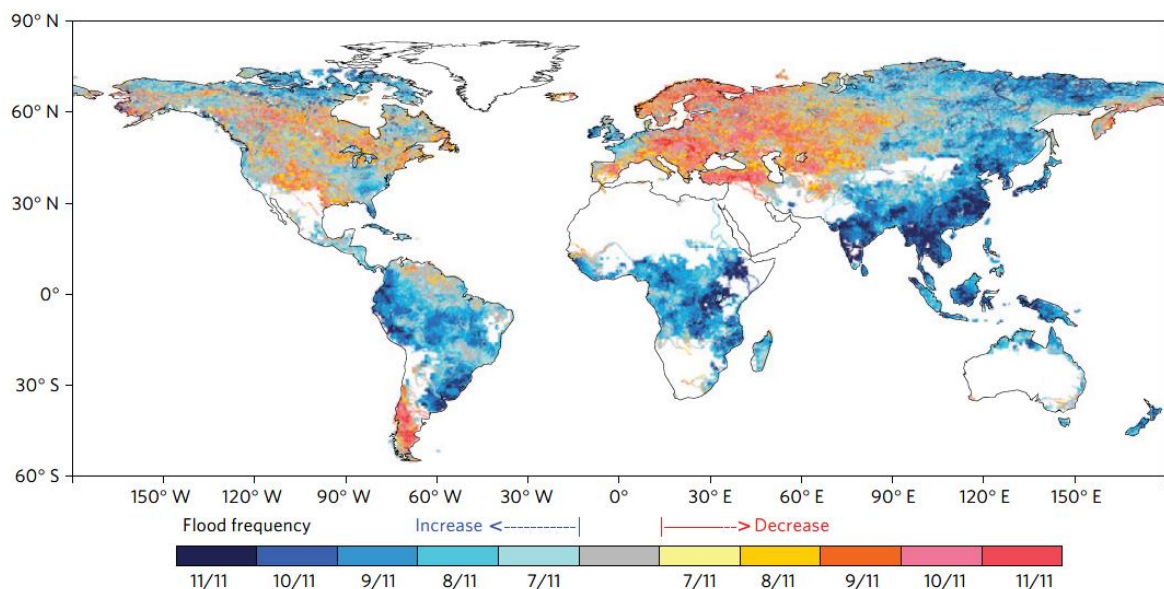


# 1 INTRODUCTION

*“Prediction is very difficult, especially if it is about the future”  
(Nils Bohr, Nobel laureate in Physics).*

*“Those who have knowledge, don't predict. Those who predict, don't have”  
knowledge (Lao Tzu, 6th Century BC, Chinese Poet).*

Flooding is the most devastating natural disaster in the world and its threats will increase in unprecedented magnitudes and frequency given climate change predictions (Chen and Wang, 2018; Khan et al., 2011). Recent scientific evidence based on outputs from eleven climate models and a global hydrological model coupled with an inundation scheme reveals increases in flood frequency in several highly populated tropical regions (Hirabayashi et al., 2013). The most significant increases are in tropical regions of Southeast Asia, Peninsular India, eastern and western Africa and central and southeast regions of South America (Figure 1.1).



**Figure 1.1: The projected probability of flood occurrence by 21<sup>st</sup> century for flood volumes related to flooding with a 100 year-return period (Hirabayashi et al., 2013).**

The number of people susceptible to flooding is estimated to double in Africa from 70 to 156 million, in Asia from 25 to 34 million and in South America from 6 million to 12 million (Willner et al., 2018). The estimated combined flood losses in tropical regions, particularly in Asia and Africa, will account for 95% of people globally affected by floods and 73% of the total with direct economic damage every year (Alfieri et al., 2017). Despite the deteriorating flood situation, most lowland tropical regions remain ungauged with inadequate in-situ measurements and hydrometeorological datasets to set up flood models and yet have significant populations at risk.

The climate change-induced flood risk demands accurate and reliable information on the potential of flood risk, particularly where large numbers of people are vulnerable (Henrik Sen et al., 2018). Flood models contribute to responding effectively to potential flooding and to the consideration of underlying catchment conditions and processes that help to develop flood risk mitigation strategies and measures (Ma et al., 2018; Wohl et al., 2012). Flood modelling is one of the critical steps towards climate change adaptation that provides flood risk information for sound decision-making in emergency flood response situations (WHO, 2012; Schumann et al., 2009). However, most tropical lowland catchments are large and trans-boundary with complex hydrological and hydraulic systems that present challenges for setting up hydrological models (Hidayat et al., 2017).

### **1.1 Challenges in Tropical lowland regions**

Flood modelling in large tropical lowland regions is a challenge due to the large scale and highly complex conditions and processes in catchments (Hidayat et al., 2017). The large catchments, which often cross geopolitical boundaries, present survey challenges for systematic data collection for hydrological modelling. It is a challenge to use conventional data collection techniques based on extensive field and cross-border surveys (Hossain et al., 2013). The scale of rivers is mostly large scale with dynamic hydrological links to peatlands, wetlands and groundwater, making tropical lowland regions a scientific challenge (Hidayat et al., 2017). The regions are distinguished by particular characteristics, such as extensive plains, high groundwater levels, high water holding capacity soils and close interactions between



river discharge, groundwater, wetlands however with unclear discharge-level relationships (Schmalz et al., 2008; Hidayat et al., 2017).

Besides, tropical lowlands are home to complex river systems that are complicated by the continually changing meanders, braided and deltaic river formations. As a result, simple hydrological techniques such as stage-discharge curves, level to level and flow to flow correlations, and rainfall-runoff models are difficult to apply without alternative approaches (Hidayat et al., 2017). Flood modelling in large tropical lowland regions suffers from lack of basin-wide data and challenges in deriving critical datasets for mapping large and highly variable catchment conditions and processes (Hossain et al., (2007).

As a result, flood extents, depths, velocities, levels, and volumes are not well known in tropical regions, as minimal flood modelling research has been conducted (Alsdorf et al., 2016). Catchment conditions and processes are often poorly defined in hydrological models due to costs, time availability and technological limitations in obtaining the required datasets (Coles et al., 1997). Successful hydrological modelling requires an ongoing and thorough understanding, building knowledge on hydrologic processes and the underlying parameters, boundary conditions, flow dynamics and the accurate representation of these in models (Wang et al., 2012). However, hydrometric records, previous hydrologic, hydraulic models and hydrodynamic information required for continuous flood modelling in tropical lowland tropical regions are usually not available (Coles et al., 1997; Hidayat et al., 2017).

Despite the challenging physical environment, lowland tropical regions coincide with highly populated areas of economic activity and growth (Hidayat, 2013). The regions have one of the world's fastest-growing populations and use of flood plains for agriculture due to availability of fertile alluvial soils that sustain crop production, drinking water and rivers that provide transport links into inland regions (Harding et al., 2017; Verhoeven and Setter, 2010). For example, over half the world's population lives in Asia's tropical lowlands, around one-quarter of the earth's land area and by 2050, half of the world's population is expected to live in tropical regions (Kabir et al., 2013).

River flooding in tropical regions is a critical issue due to impacts on large numbers of people living in flood risk areas without adequate flood protection measures (Alfieri et al., 2017). The most severe loss of life and property in tropical regions occurs due to flooding and affects many people every year, a result of developments in floodplains and alterations of natural environments. Rapid population growth linked to economic growth means that the impact on the natural environment increases the vulnerability to flooding of large numbers of people (Verhoeven and Setter, 2010). In several areas, population growth has caused significant loss of forest cover where agricultural and mining activities are widespread, resulting in changes in the natural systems through deforestation, urbanisation and overgrazing worsening the already deteriorating flood risk situation (Di Baldassarre et al., 2010).

As a result, when tropical lowland regions experience flooding under altered natural conditions, the social, economic and environmental impacts of flooding are catastrophic with loss of lives and properties. In response, investments by governments and flood risk management agencies in flood protection programmes typically resolve more of the consequences than the causes of the floods. Several tropical countries continue to face challenges in flood infrastructural development, and opportunities to improve are often restricted by limited human and institutional resources, including the capacity for good governance (Harding et al., 2017).

## **1.2 Large-scale distributed hydrological modelling approaches**

The rapidly increasing availability of satellite datasets coupled with increased computing power, Geographical Information Systems (GIS) and data analysis techniques provide new opportunities for large-scale distributed modelling approaches, particularly in data-poor tropical regions where millions of people are vulnerable to flood risk (Li et al., 2016). Satellite data offer extensive geographical coverage and can provide hydrologic models with vast quantities of spatially distributed data required for large-scale distributed hydrological modelling (Xu et al., 2014).

Despite recent progress in the availability of satellite data both in terms of quality and quantity, the often and commonly used lumped hydrological models cannot efficiently use this wealth of increasing satellite data for accurate flood modelling (Kauffeldt et al., 2016). Lumped models cannot make effective use of satellite data sets because they are based on vast areas of averaged parameter values over entire drainage areas (Sharad et al., 2018). As a result, satellite datasets have not been used as widely as anticipated in flood modelling, considering the growing number and variety of satellite platforms that have been sent into space. Practical approaches to increasing flood modelling accuracy and providing reliable flood simulations in large-scale lowland tropical regions need to be investigated, given the rapidly increasing availability of satellite data, computer power, cloud systems, and the advancement of the Geographic Information Systems (GIS) (McCabe et al., 2017).

The availability of improved hardware and software applications and digital terrain data has the potential to facilitate the application of distributed hydrological modelling approaches capable of providing a detailed representation of temporal and spatial variability of catchment conditions and processes (Chen and Wang, 2018). Distributed hydrological models divide the catchment into grid cells, with each grid cell having catchment properties uniquely represented in the model (Xu, 2002; Chen and Wang, 2018). These high definition models can integrate high-resolution satellite data because of their regular squared grids that are scalable and can capture individual variations/heterogeneity of catchment properties and processes at high resolutions (Andersen et al., 2001). Distributed hydrological modelling approach

provides opportunities to represent accurately the highly variable catchment properties, processes and state variables and can maximally use of available satellite data information (Chen and Wang, 2018).

Although distributed hydrological models have been successfully applied in some data-rich countries for flood forecasting purposes, they have been successfully applied in small-scale catchments. Few studies have applied distributed hydrological models using satellite data for predicting catchment outflow hydrographs in large tropical lowland regions. The focus of this thesis is therefore to explore the potential of fully distributed hydrological modelling approaches in large lowland tropical regions, given the rapidly increasing availability of satellite data, computer power, cloud systems, the advancement of the Geographic Information Systems (GIS).

### **1.3 Scaling relationships**

Flood prediction in ungauged regions is a global challenge, and most tropical regions with increasing flood risk do not have adequate hydrometric records required for setting up flood models. This research investigates the application of statistical scaling modelling approaches to complement the hydrological modelling undertaken in the catchment. In parallel to hydrological modelling, statistical models are essential to complement the accuracy of flood simulations, where there are limited records, especially in ungauged tropical regions (Gupta et al., 2007).

Floods result from complex rainfall interactions with various catchment properties, conditions and processes, including topography, land cover and moisture states. While climate change and human-made environmental changes will render flood prediction a very complex problem, the interactive catchment properties and processes can be statistically quantified, and the links can help to estimate the magnitude of peak discharge (Mandapaka et al., 2009).

The USA Geological Survey (USGS) pioneered a method of regional quartile equations by establishing regression relationships between catchment flood explanatory variables that relate to catchment conditions, processes and states. The Regional Flood Frequency (RFF) analysis is still used to predict annual peak flow quantiles, such as 100-year floods in ungauged basins establishing regression

relationships between catchment flood explanatory variables that relate to catchment conditions and processes. The Regional flood frequency equations are statistical descriptions that use historical streamflow records and the concept of homogeneous regions to complement the accuracy of flood quantile estimates because of limited record lengths (Mandapaka et al., 2009).

The equations, however, cannot include physical processes that produce rainfall-runoff flooding and are not suitable for short-term flood predictions. Required is a solution that considers short-term physical processes at a rainfall-runoff event scale in the generation of floods in drainage areas (Gupta et al., 2010). Thus, the challenge had been the prediction of flood parameters and peak discharge, from physical processes on a rainfall-runoff event scale instead of using the Regional flood frequency approaches based on annual scales and large homogenous regions.

In response to this challenge, the Regional flood frequency method subsequently underwent several refinements; however, drainage area continues to be identified as the most significant variable to predict peak discharge (Ayalew et al., 2018; Mandapaka et al., 2009). The approach later provided an opportunity to establish, among others, a new geophysical flood theory that establishes flood scaling relationships across a broad range of catchment scales to estimate flood parameters and peak discharge (Gupta et al., 2010). The theory of floods also referred to as the scaling theory put forward the hypothesis that at a single rainfall-runoff event there is a scale-invariant spatial structure of peak discharge with drainage areas that is defined by power-law relationships (Ayalew et al., 2018). The theory focused on river drainage areas and abandoned the concept of homogeneous regions and incorporated flood-producing physical processes at a rainfall-runoff scale (Gupta et al., 2010). It integrates flood-producing physical processes based on the understanding of scaling or power-law relationships between observed peak discharge and drainage areas. It provides a geophysical understanding of how flood parameters and peak discharge can be estimated from physical processes across spatial scales (Ayalew et al., 2018).

Several studies have so far provided the theoretical and empirical basis for the formulation of the theory of floods over the past twenty years (Gupta et al., 2004, 2010 & 2015; Menabde and Sivapalan, 2001; Ogden and Dawdy, 2003; Dawdy et al., 2012). The results confirm that peak discharge, from a single rainfall-runoff event in nested catchments, has a power-law relationship with nested catchment drainage areas and the exponents ( $\theta$ ) and intercepts ( $\alpha$ ) of the scaling equations change from one event to another (Gupta et al., 2010). To establish the physical processes governing flood parameters (exponents ( $\theta$ ) and intercepts ( $\alpha$ )) and their event to event variability or changes, most research has been focused on establishing and validating the scaling relationships between peak discharge and drainage areas including analysis of flood parameter changes related to rainfall variability and other measurable catchment conditions and processes (Ayalew et al., 2014).

However, most studies investigating the scaling relationships have been done in the United States and the United Kingdom (Ayalew et al., 2015; Wilkinson and Bathurst, 2018). Also, most of the studies were experimentally based on plots of less than 10 km<sup>2</sup>, mostly in the wet northern latitude catchments. Subsequently, other studies were done at catchment scales using realistic catchment conditions and processes. However, no similar research has been conducted in other climatic regions with different rainfall types and coverages.

Therefore, most flood scaling theory studies have been conducted in humid northern latitude catchments, and it is not known whether the scaling theory applies in other climatic regions. Scaling relationships in various climatic regions need to be investigated, considering different rainfall types and coverages, and contribute to the challenge of flood predictions in ungauged regions (Sivapalan et al., 2003; Wilkinson and Bathurst, 2018). The statistical scaling relationship studies should be expanded to larger tropical river basins where convective rainfall and isolated flood events often occur. An in-depth understanding of flood scaling relationships between peak discharge, drainage areas and other measurable physical catchment processes is essential and should develop statistical flood parameter models for peak discharge prediction.

This research extends the scaling theory of floods in tropical regions using La Sierra sub-catchment as a case study. The research analyses, statistical relationships between tropical rainfall variability, nested catchment data and other measurable catchment conditions and processes. Such an approach provides an opportunity to understand the event to event variability of flood parameters, statistical relationships with catchment conditions and processes in the process of flood generation in the study area (Ayalew et al., 2015). The derived flood parameter relationships and scaling equations will provide a complementary framework for estimating flood parameters and peak discharge and refining and verifying the magnitude of peak discharge produced by numerical models.

#### **1.4 Uncertainty quantification**

Proposed is a large-scale distributed hydrological modelling approach based on satellite data and limited in-situ data, but subject to errors and uncertainties from various sources. The estimation of rainfall input data, parameter values, the magnitude and timings of discharge, including inundation extents, are all uncertain. Reliable flood simulations are critical for successful flood modelling; however, discharge estimation relies on data and models prone to uncertainties. Because of several sources of uncertainties in flood modelling that compromise modelling results, there is a need to identify and quantify individual uncertainties to improve the reliability of model results.

Studies on uncertainties in flood inundation mapping have investigated the effect of different uncertainties on flood modelling results (Jung and Merwade, 2015). Several research results show that different sources of uncertainties of differing magnitudes affect flood modelling outputs and, in turn, interact and combine, resulting in different simulation results (Bates et al., 2014). However, most uncertainty studies have focused primarily on quantifying the combined or total uncertainty in model simulations without exploring the decomposition of individual uncertainty contributions to simulation results. Combined uncertainties from various sources; model structure, inputs and parameter values contribute to model output uncertainties. However, simulations are compromised by uncertainties from individual sources (Brown and Heuvelink, 2006). Uncertainties in hydrological models differ

based on different sources and are propagated individually through the model, although they interact and have a combined effect on model results.

A critical question that arises in modelling is how individual uncertainties associated with rainfall estimates propagate and interact with other uncertainties and affect flood simulations. The propagation of individual uncertainties and how they affect overall uncertainty in the model output is not well known due to many sources of uncertainties that interact and combine due to the propagation process (Jung and Merwade, 2015; Lim and Lee, 2018). Thus, the challenge that remains is how to disentangle the combined uncertainties and quantify each individual taking into account the complex interactions and effects in model outputs (Blöschl et al., 2019).

Therefore, explicit quantification of individual uncertainties associated with hydrological and hydraulic model simulations is proposed in this study to assess and comment critically on modelling weaknesses and how each uncertainty affects model results (Diaz-Ramirez et al., 2013). There is a need to quantify uncertainties as source-based individuals and determine the effect of propagating each within a cascading model chain while analysing the interactions and assessing the effect of each on discharge output and flood inundation extents.

Integrating specific, quantified uncertainties and interactions in model outputs help to convey complete flood risk information to decision-makers and people at risk (Zarzar et al., 2018). Model output users should be informed about the underlying conditions of the modelling process and the individual uncertainties and how each affects the final model output. Awareness of uncertainty characteristics in terms of specific types and magnitude is crucial for reliable model performance and accurate flood risk-based decision-making that enables resource targeting to particular catchment conditions or sources that contribute most to uncertainties and errors in flood simulations.



## 1.5 Knowledge gap and research needs

There is a need for accurate flood modelling to generate reliable simulations in data-poor tropical regions using the increasingly available satellite datasets, computer power and advancement of GIS tools and taking advantage of existing underutilised distributed hydrological models. Flooding is projected to increase in tropical regions given climate change predictions and is an urgent requirement in these regions where millions of people are at risk. A greater understanding of the dynamism and underlying mechanisms of tropical flood hydrology, including interactions and effects of floodwaters, is urgently required to implement effective flood risk mitigation strategies and measures (Wohl et al., 2012).

However, knowledge gaps still exist on the dynamics of flood flows, magnitudes, timing and impact of peak discharge in tropical regions (Wohl et al., 2012). Despite significant progress in satellite data availability, the frequently used lumped hydrological modelling approaches are not sufficiently efficient enough to use this wealth of accumulating satellite data for flood modelling (Kauffeldt et al., 2016). Instead, distributed hydrological modelling approaches can use efficiently the increasing high-resolution satellite data to represent accurately catchment conditions and processes in models and generate reliable simulations given the rapidly increasing computer power, cloud systems and the advancement of the Geographic Information Systems (GIS).

The high spatial resolution modelling over large-scale approach contributes to closing knowledge gaps that exist between the use of simplified, coarse resolution modelling approaches in large-scale catchments and use of detailed high-resolution modelling approaches in small-scale catchments due to limited data availability (Sampson et al., 2015). There is potential in using large-scale high-resolution distributed hydrological modelling approaches in increasing flood modelling accuracy and provision of reliable flood simulations in large tropical regions. Although considered to be a new trend, the use of distributed hydrological models for flood modelling is rarely used and their potential to accurately represent catchment properties, and processes are not fully realised (Chen and Wang, 2018; Kauffeldt et al., 2016). There is a need to contribute to meeting the challenge of representing accurately the highly

variable catchment conditions and processes and provide reliable flood simulations in large tropical regions where an increasing number of people are at risk of flooding.

To further improve the simulation of the magnitude of peak discharge, this study investigates the application of the flood scaling relationships between peak discharge, catchment properties and processes that control the generation of flooding in nested catchments of the La Sierra catchment. There is a need for more flood scaling studies with different storm types linking hill slope and headwater catchments to large catchment scales in various climatic regions (Wilkinson and Bathurst, 2018). Compared to wet temperate regions, few studies have been conducted in tropical regions to investigate the physical processes governing flood-generating parameters and estimation of peak discharge. The mechanism of flood generation in small catchments leading to flooding at larger catchments is still not fully quantified, including the universal application of flood scaling theory in other climatic regions remains unknown. There is a need to bridge the scaling gap by extending the scaling theory of floods from experimental plots and small catchments to large regional and continental river catchments in different climate regions.

Due to several sources of uncertainty in flood modelling that compromise modelling results, there is a need to identify and quantify individual uncertainties to improve the reliability of model results. The full range of sources of uncertainties in models is a challenge to analyse each source individually, avoiding confusion with those from different sources (Tscheikner-Gratl et al., 2018). The distribution of individual uncertainties and their effect on overall simulation results is necessary, however, their distributions are not widely understood (Jung and Merwade, 2015; Lim and Lee, 2018). The critical question that remains to be answered is how to disentangle combined uncertainties (from different sources) and quantify each uncertainty considering the complex interactions and effects on model outputs (Blöschl et al., 2019). Also, how uncertainties in input data propagate and interact to determine discharge output volumes and inundation extents. Further, what are the effects of individual uncertainties and their interactions with other uncertainties when propagated in a cascading modelling chain and the effects on discharge output and inundation extents (Rodríguez-Rincón et al., 2015)?

## 1.6 The study area

The Grijalva River catchment area was used as a case study for distributed hydrological modelling and is situated in the southeast of Mexico, covering large mountainous and lowland areas drained by one of Mexico's major rivers (Figure 1.2). The study area is a large trans-boundary catchment (57 958 km<sup>2</sup>), bounded by 89.6° and 94.5°W and 15.3° and 18.7°N, straddling much of the Mexican State of Tabasco in the north, Guatemala in the south, the State of Veracruz to the west, the state of Chiapas in central and south-eastern parts (Aparicio et al., 2009).

The Grijalva catchment is highly variable but can be divided into two broad physiographic regions, the low-lying Gulf coastal alluvial plains in the north (averaging 10 metres above sea level) and the southern highlands of the Chiapas and Guatemala comprising highly dissected mountain ranges, depressions and plateaus reaching elevations between 2 100 and 2 400 metres above sea level (INEGI, 2018). Although the basin experiences a hot, humid maritime tropical climate, it has highly variable climatic conditions closely following the two major physiographic regions. In the low-lying Tabasco plains, average temperatures reach 40°C and do not fall below 18°C in the coldest month, while in the Chiapas highlands average temperatures are 31°C and minimum are 14°C.

Rainfall is highest during the passage of the Inter-tropical Convergence Zone (ITCZ) that generates intense rainfall totals, and its movements control the seasonality of rainfall patterns in the catchment. However, rainfall occurs nearly all year-round (see Appendix A, Figure A-1), with pronounced wet and dry seasons, with the highest annual rainfall reaching approximately 4 000 mm and mean yearly of 2 550 mm (Gama et al., 2010; Arreguín-Cortés et al., 2014). The complex east-west mountainous regions (ranging to over 4 000 metres) are part of north-south Central America, La Sierra mountains, another factor in the high spatial variability of rainfall in the Grijalva catchment. The mountainous slopes have an orographic effect on moisture-laden north-east trade winds from Atlantic resulting in intense rainfall in the La Sierra sub-catchments, however, causing a drier leeward side mostly in the central depressions of Chiapas (Verdin et al., 2016).

The highly variable topographic and climatic conditions have resulted in contrasting patterns of soils and vegetation in the catchment. The Tabasco plains in the north are mainly composed of deep clayey soils which are hydromorphic, while others comprise azonal alluvial soils in flood plains, levees and terraces (Krasilnikov et al., 2013). In the highlands and upper plateau areas in the south, comprising prairie soils with associated coniferous vegetation (Arreguín-Cortés et al., 2014). There are highly variable vegetation types in the catchment with the predominant type, mainly medium-height evergreen forest, with different degrees of disturbance.

The Grijalva River is the major river in the study area originating in Guatemala in La Sierra de Cuchumatanes and then enters Mexico, crossing the central depression of Chiapas, where flows are regulated by four dams (Arreguín-Cortés et al., 2014). Further downstream, the river is divided, creating two of essential distributaries; the Carrizal and the Samaria Rivers, however, recovering its name (Grijalva River) after Carrizal distributary River that crosses Villahermosa city downstream to its end in the Gulf of Mexico (Ramos et al., 2009)

### **1.6.1 Villahermosa City and flood risk**

The key economic activities in the Grijalva catchment area are mainly oil production in the Tabasco state and small-scale mixed livestock farming in both states, including Chiapas, which are connected to rapid urbanisation and the deterioration of the surrounding natural vegetation (Ramos et al., 2009). Originally, Villahermosa city, the capital of Tabasco State, was surrounded by undisturbed rain forest and water bodies; however, rapid population growth changed the physiography of the city through deforestation, mixed agriculture and modified river courses (Ramos et al., 2009; Musálem-Castillejos et al., 2018). The depletion of vegetation cover led to the loss of soils and reduction in infiltration capacity of soils causing rapid surface runoff, hill-slope erosion, river siltation and flooding (López-Caloca et al., 2017).

Villahermosa is confined within the confluence of the Grijalva and La Sierra Rivers, and the floodplains often experience severe flooding, such as those that occurred in

2007, 2008 and 2009 (Arreguín-Cortés et al., 2014; Ramos et al., 2009). Also, the State of Tabasco straddles the delta of two major rivers in Mexico, the Grijalva and the Usumacinta Rivers and several of its districts and municipalities frequently experience severe flood events. During the intense rainy season (July to October), hurricanes and tropical waves occur in the study area, frequently leading to flooding, leaving populated centres isolated and without communication.

To protect the city from frequent flooding, some river flows were regulated through the construction of flood defence structures such as levees, barriers and dikes, roads and river diversions. Flood barriers were constructed along the Samara River, and dams were built to regulate river flows and reduce the risk of flooding in Villahermosa city (Ramos et al., 2009). In recent years, the construction of residential areas within the city has not only changed the landscape and resulted in the expansion of irregular settlements on the outskirts of the city and its rivers.

### **1.6.2 La Sierra sub-catchment**

La Sierra catchment is one of the sub-catchments in the main Grijalva catchment (Figure 1.2) and was selected to investigate the application of the flood scaling theory in nested catchments and for studying the propagation and interaction of uncertainties through coupled hydrologic and hydraulic models (Objective 1 and 2). The catchment is in the east of the Grijalva catchment covering an area of 6 743 km<sup>2</sup>, an important contributor of flood flows in the Grijalva River and further to the Gulf of Mexico (López-Caloca et al., 2017). La Sierra catchment is one of the unregulated sub-catchments in the Grijalva catchment with mostly natural flows. It is a major tributary of the main Grijalva River and one of the primary sources of flooding in Villahermosa city.

La Sierra sub-catchment is the wettest catchment in Mexico with an annual rainfall total of approximately 4 000 mm, annual averages of 2 550 mm occurring all year with very high totals from June to October compared to lower totals or drier months from November to May (INEGI, 2018). Because of high amounts of rainfall received in altered environments, the catchment generates significant amounts of runoff resulting in severe flooding in the downstream city of Villahermosa. The areas that

are frequently flooded in the city include downstream areas, between Las Gaviotas Gauging Station in the south and El Povernir Gauging Station in the north, mostly in the confluence areas of La Sierra and Grijalva River.

In La Sierra catchment, six smaller catchments were identified as nested catchments for a detailed study of the peak flood generation at various spatial scales. The nested catchment areas were defined from seven available flow gauging stations, each of which represents an outlet from each nested drainage area. The nested catchment areas and their sizes identified and named after their gauging stations were: Gaviotas (6 743 km<sup>2</sup>), El Pueblo (4 748 km<sup>2</sup>), El Puente (1 787 km<sup>2</sup>), Teapa (238 km<sup>2</sup>), El Puyacatengo (128.67 km<sup>2</sup>), Oxolotan (1 416 km<sup>2</sup>), and Tapijulapa (698 km<sup>2</sup>), sub-catchments. However, the El Puente and Tapijulapa gauging stations and their nested catchments were excluded from the study because of insufficient hydrometeorological records for the years under consideration.

## 2 AIMS AND OBJECTIVES

The research aim is to: identify and apply physics-based and data-based hydrological modelling approaches for large-scale flood modelling in lowland tropical regions using the Mexican State of Tabasco as a case study.

This aim is accomplished in the following objectives:

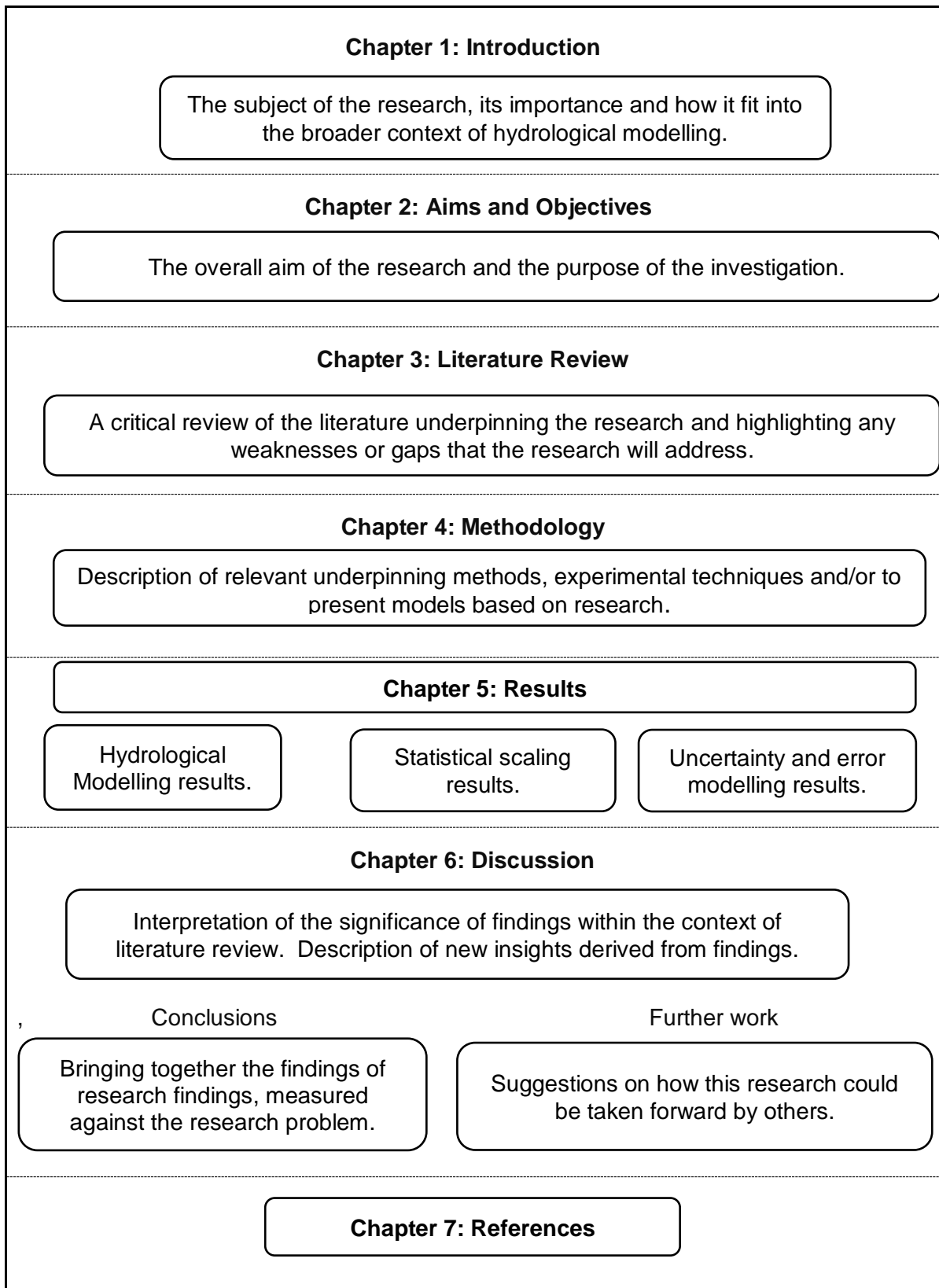
**Objective 1:** Integrate satellite datasets into the SHETRAN distributed hydrological model based on error quantification and correction to simulate the magnitude and timing of flood peak discharge in the Grijalva catchment area.

**Objective 2:** Investigate statistical scaling relationships of rainfall variability, flood peak discharge and nested catchment drainage areas to complement the estimation of peak discharge magnitudes.

**Objective 3:** Assess the effects of propagation and interaction of rainfall and parameter uncertainties on the magnitude of peak discharge and flood inundation extents in the study area.

### 2.1 Thesis outline

The thesis is written in monograph format, with discrete chapters included covering the different components required and cross-referenced to ensure coherence. The thesis is divided into seven parts, with the following summarising each chapter's content (Figure 2.1).



**Figure 2.1: The thesis outline divides the thesis report into seven chapters.**



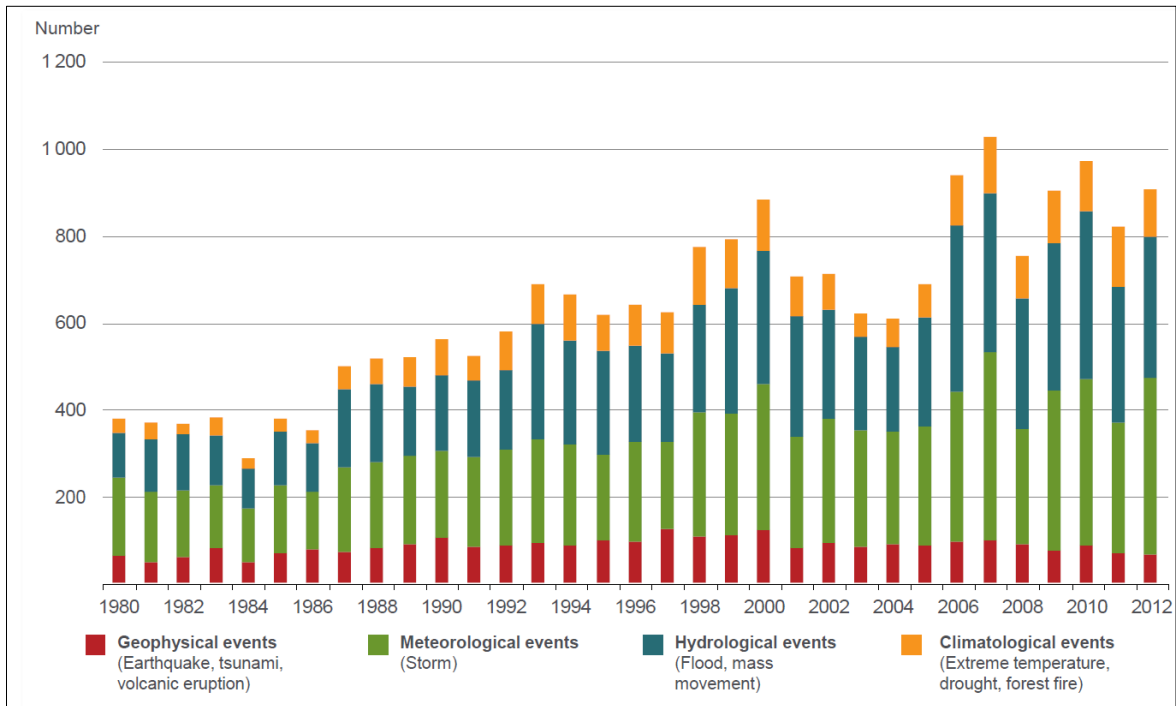
## 3 LITERATURE REVIEW

This section focuses on hydrological modelling approaches for large-scale flood modelling to improve flood modelling accuracy and provide accurate simulations of flood magnitude and extents in large data-scarce tropical regions. Using satellite data as a potential solution in overcoming challenges in the development of hydrological models in data-poor regions is reviewed. The chapter is divided into three main sections: (1) an overview of increased flood risk, satellite data opportunities, and distributed hydrological modelling approach (2) application of statistical scaling relationships to estimate flood parameters and peak discharge and (3) propagation and interaction of rainfall and parameter uncertainties on peak discharge and flood inundation extents.

### 3.1 Increased risk of flooding

Globally, floods are, by far the most prevalent natural hazards affecting human lives and infrastructure (Khan et al., 2011). The Centre for Research on the Epidemiology of Disasters (CRED) (2016) report that on average, 335 flood disasters occurred each year between 2005 and 2014, which is almost twice as high as in 1985-1994. Many of the observed changes in the global climate system have been unprecedented, and these observed changes include warming of the atmosphere and oceans, the rise of sea levels and a record increase of extreme rainfall events in several regions (IPCC, 2014).

More regions have seen increases rather than declines in the number of extreme weather events. The increasing trend in extreme rainfall is resulting in extreme river flows in several watersheds, implying a more significant regional-scale flood risk (IPCC, 2014). The trend shows an unprecedented increase in weather-related incidents over the last 36 years (Figure 3.1). For example, Munich Re, (2018) estimated global weather-related losses of \$330 billion in 2017, up from \$184 billion in 2016, nearly three times the average over the last ten years, to \$49 billion, and four times the average of \$35 billion over the last 30 years. In the past two decades, approximately half of the people killed due to natural disasters have been victims of flood-related incidents (Munich Re, 2018). Floods account for around one-third of the world's economic losses (WMO, 2013).



**Figure 3.1 Increase of weather-related incidents and compared to other hazards over the last years (Munich Re 2018).**

For example, historical flood reports from the Emergency Events Database (EM-DAT) (Guha-Sapir et al., 2009) show that the number of deaths recorded, and the number of flood losses increased between 1960 and 2013 (Figure 3.2). Reports show a 1.5% rate of increase in the total yearly number of fatalities and 6.3% of the annual estimated economic losses (Tanoue et al., 2016). It is shown that floods accounted for over 65,000 deaths, displaced over 1.1 billion people, and cost approximately \$280 billion in 2002 - 2011 alone (Wu et al., 2014).

Though in evaluating the potential impacts of climate change on global flood risk, there are some regional variations. Arnell and Gosling, (2016) estimated that, by 2050, the current one-hundred-year flood event will frequently occur, affecting nearly 450,000 million people and 430,000 km<sup>2</sup> of croplands, at least twice as often across tropical regions. Flood losses are projected to continue to increase because of climate change and the associated effects of land-use change, deforestation, sea-level rise and flood-prone developments, leading to a global increase of 2 billion by 2050 (Jain et al., 2017; Bogardi, 2004; ICHARM, 2009; Vogel et al., 2011). The deteriorating flood risk situation needs governments, communities, households and individuals to be aware of and take measures to mitigate current and potential flood risks.

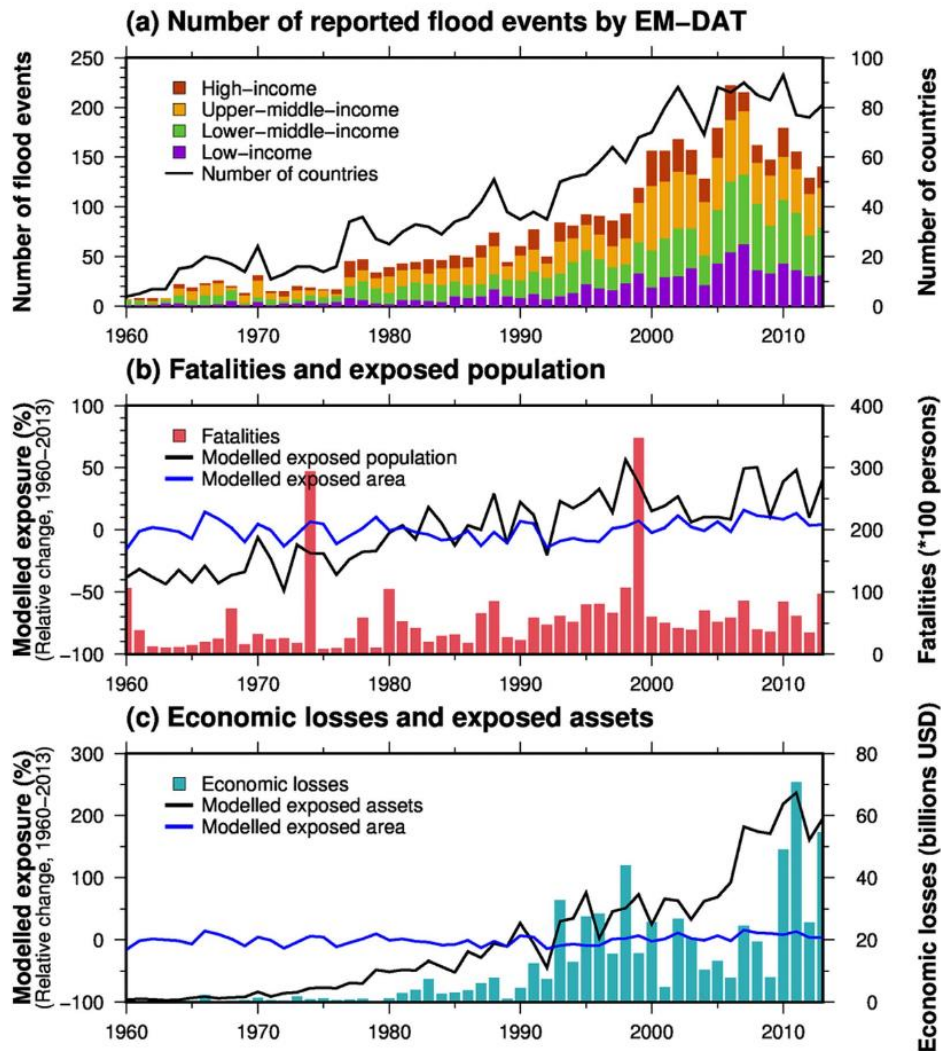
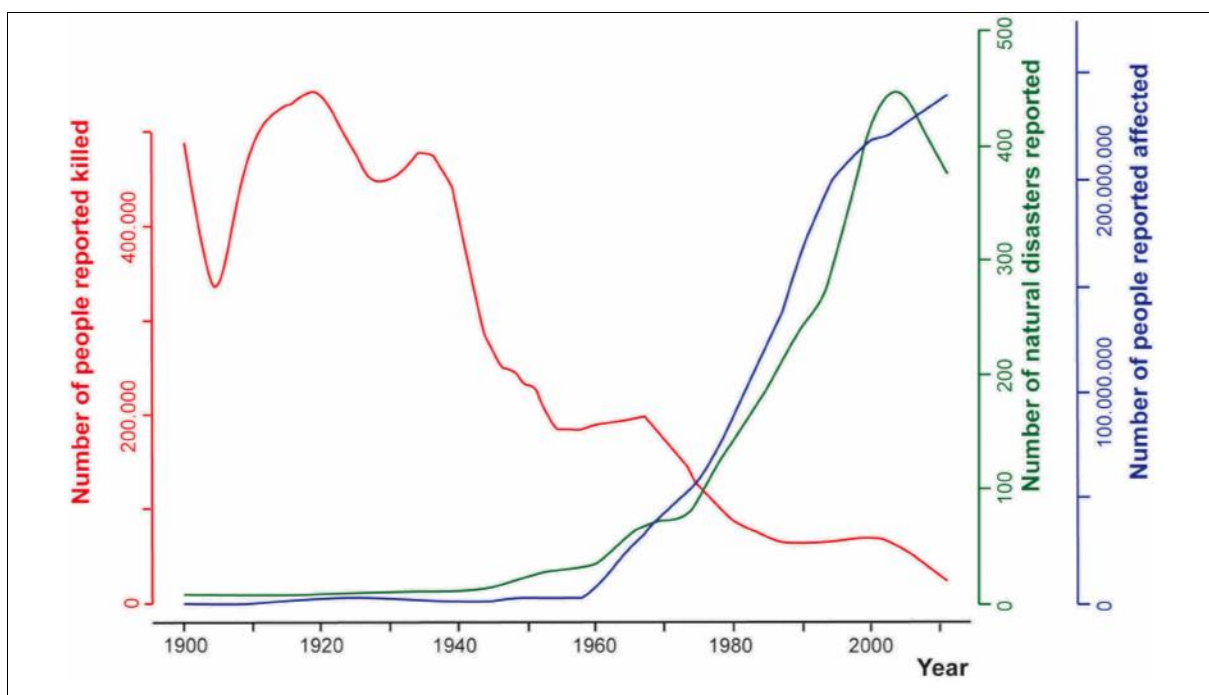


Figure 3.1: (a) The number of flood events per income (bar) and (b) number of countries reporting flood events from 1960 to 2013 in the Emergency Events Database (EM-DAT) (solid line) (Guha-Sapir et al., 2009).

### 3.1.1 The need for flood modelling

Flood modelling and warning systems are the most effective flood risk management technique, which can minimise the adverse effects of flooding. The systems are a valuable adaptive measure that generates flood warnings that are disseminated on a wide scale to the benefit of large numbers of people, including those who do not have protection from hard defences. The forecasts and warnings produced should be timely and reliable, and this is essential for sound decision-making in flood emergencies (Schumann et al., 2009).

The importance of flood modelling and warning systems is also acknowledged by WMO, (2012), which estimated that global flood-related economic losses increased by about 50 times over the last decade, but the loss of life and property decreased by 10 due to flood modelling and associated early warning systems linked to emergency preparedness and response. Evidence of this is reflected from studies of worldwide flood disaster incidences and impact data stored in the Emergency Events Database (EM-DAT) (Guha-Sapir 2009) which shows that the death toll has been declining while the number of people affected has increased since 1900 (Figure 3.3). This decrease is attributed, among other factors, to flood forecast systems that play a vital role in early flood warning, awareness and forecasting of pre-existing floods, helping people to take precautions to protect lives and property (Jain et al., 2018). The warning systems provide a vital service to alert communities when floods are developing and providing communities and organisations opportunities to prepare and mitigate the resulting impacts (WMO, 2013).



**Figure 3.2: Global natural disasters between 1900 and 2010 (includes the number of disasters, deaths and people affected). Source: EM-DAT: International Disaster Database (Guha-Sapir et al., 2009).**

Flood prediction and warning systems are different to conventional flood management approaches comprising structural flood defences such as flood reservoirs, floodwalls and embankments where the emphasis is on changing flood properties and minimising flood depth and inundation extents. Though these structural measures reduce flood risk, they cannot remove the risk altogether. Structural interventions are not possible in specific locations, especially in remote mountain areas, and are not useful for all inundation scenarios and can cause unintended environmental impacts (Sharad et al., 2018).

There is currently a fundamental change from flood defence to flood risk management, which involves enhanced approaches, such as greater advocacy for soft flood management approaches, allocation of responsibilities, and more focus on the responsibilities of private citizens. In this modern flood risk management approach, hydrological /hydraulic models are increasingly recognised as vital tools for understanding catchment systems, flood risk areas, risk assessments and encouraging the selection of appropriate mitigation measures (Revilla-Romero et al., 2015). Individual citizens should take greater responsibility for their state of readiness with an emphasis on improved awareness, preparedness and emergency response (Wilby and Keenan, 2012).

While flood forecasting is one of the hydrology's most difficult techniques, it contributes significantly to reducing life and infrastructure losses (Jain et al., 2017). Hard-defence systems are effective when combined with a range of measures that include flood forecasting and alert systems (Khatibi and Haywood, 2002). Therefore, the development of flood forecasting systems is a critical element in the ability of communities to adapt to the worsening flood risk situation.

### **3.1.2 Installation of river gauge network and challenges**

River gauge networks and their long-term data records are critical elements for developing robust flood forecasting systems. However, recent studies have shown that the number of existing gauge networks and the installations of new networks, including the quantity of hydrometeorological data collected in many parts of the world, has fallen significantly (Khan et al., 2011; Revilla-Romero et al., 2015). The number of gauging stations declined by 48% in the period 1985 - 2005 in the Aral Sea

basin, while from 1980 to 2004, over 2 000 gauging stations with 30 or more years of discharge record were discontinued in the USA and several regions in Europe, Latin America and Asia were also affected (WMO, 2013).

There are various explanations for this decline, and these include insufficient funding problems, inadequate institutional structures, insufficient understanding of long-term hydrological records and, sometimes, political instability. Current data collection networks have been reported to be insufficient to provide the required information to understand and clarify changes in the hydrological system (Mishra and Coulibaly, 2009). The loss in hydrometric stations is happening when global warming is worsening weather conditions, making it harder to track and predict flooding than it was 30 years ago. The loss is the most concerning issue threatening the foundation on which most hydrological processes are understood. Mishra and Coulibaly, (2009) warned of difficulties in validating remote sensing datasets and model outputs in many data-poor regions due to lack of long-term in-situ data required for accurate flood modelling work.

### **3.1.3 Challenges in tropical lowland regions**

Lowland tropical regions lie between 25° north and south of the equator and include areas in main river basins, downstream parts in the form of deltas. The lowland areas occupy one-fifth of the world's land and mostly occur below 300 metres above mean sea level. While tropical regions vary significantly, usually rainfall is higher than evaporation for at least 270 days annually (Wohl et al., 2012). The convergence of the surface winds in the humid tropics east and west causes strong convection and cyclonic vorticity, which typically results in tropical rainfall. The regions generate the highest amount of runoff and are subject to the greatest severity of alteration or disturbance of land cover. As a result, millions of people are affected by floods, as large numbers of people live in flood plains without sufficient flood protection. The need for flood warning systems is even more significant in these regions; however, the following are challenges that militate against setting up flood forecasting systems in these regions.

### ***Large/trans-boundary river basins***

The development of flood modelling and forecasting systems in tropical lowland river catchments is a challenge due to large, dynamic and complex water environments (Hidayat et al., 2017). Most tropical basins are large and trans-boundary, and this presents coordination challenges in obtaining the required modelling data. The size of the regions presents survey challenges and limitations for conducting routine field measurements of river dimensions, bathymetry, terrain properties and boundary conditions required for setting up hydrologic models. It is a challenge to use traditional data collection techniques that rely on extensive in-situ and cross-border field measurements or surveys (Hossain et al., 2013). The scale of the rivers, the complex geomorphology and the hydrological relationships with nearby peat moors and inland wetlands are all susceptible to fluvial flooding and render tropical lowlands a scientific research challenge (Hidayat et al., 2017).

### ***Complex hydrological and hydraulic environments***

Tropical lowland regions present complex hydrological and hydraulic environments that are further complicated by meandering and deltaic nature of river systems. Lowlands are characterised by distinctive features, including flat-topography with low hydraulic gradients, high water retention potential in peatlands, and strong river and wetland interactions that result in unclear stage-discharge relationships (Schmalz et al., 2008). Without new approaches, hydrological techniques such as the rating curve, level-to-level correlations and rainfall-runoff modelling can be a challenge to apply.

### ***Greater spatial and temporal variability***

In tropical regions, all hydrological components, including energy flows, water flows and interactions, are characterised by higher spatial and temporal variability, higher amplitudes, pronounced spatial gradients and the potential for rapid reaction due to human interventions (Wohl et al., 2012). Very few studies have been conducted to model complex hydrological and hydraulic dynamics and interactions, including spatial-temporal rainfall-runoff variability in large lowland tropical regions compared to humid lowland temperate regions (Hidayat et al., 2017). Tropical lowland regions are characterised by highly variable catchment conditions and processes covering

broad spatial and temporal scales (Wohl et al., 2012). The hydrology is influenced by large global circulations such as the Hadley cell, the decadal oscillations and multi-annual regional circulations. Also, tropical hydrology is controlled by the Global patterns of monsoon-like El Niño and La Niña: Southern Oscillations, mesoscale circulations, such as land-sea breezes, orographic flows and intense localised rainfall (Wohl et al., 2012). The results are highly variable catchment processes with broad inter-annual, sub-seasonal and daily variability, intense spatial gradients and water fluxes, particularly for rainfall-runoff and channel flow. The variability also involves local energy and land-water cycling, influenced by surface roughness, albedo, soil moisture variability, intensive rainforest convection, latent heating of large and small mountain ranges, and micro-scale cloud control.

### ***Increased flood risk***

Predicting the fate of tropical lowland river flows in response to existing flood risk challenges, including those resulting from climate change and anthropogenic origin, remains a challenge (Hidayat et al., 2017). Climate change is projected to affect disproportionately tropical areas, worsening existing flood risk management arrangements (Harding et al., 2017). Although natural disasters are rising worldwide in severity and scale, tropical lowland regions continue to experience a faster increase in hazardous flood events worsened by rapid urbanisation, deforestation and agriculture (Kabir et al., 2013). The life and properties of millions of people every year are threatened by significant flooding in the Southeast Asia lowlands, Peninsular India, East and West Africa and regions of South America. Flood risk management challenge is another dynamic theme that must be understood and managed in several lowland tropical regions.



### ***Limited data availability***

The development of flood forecasting systems in large lowland tropical regions suffers from lack of basin-wide data and challenges of deriving critical local datasets for mapping large and catchment conditions and processes (Hossain et al., 2007). Very few studies have studied the dynamics of major tropical rivers that have not been well studied compared to their counterparts in mid and high latitude regions (Hidayat et al., 2017). Successful hydrological modelling requires a thorough understanding of catchment conditions and processes and the underlying parameters, boundary conditions, flow dynamics and integration in hydrological models (Wang et al., 2012).

In tropical regions, however, flood depth, surface water and wetland stocks are not well known since only a few hydrological studies have been performed (Alsdorf et al., 2016). In most cases, small areas are often examined in isolation from other large-scale catchment processes, including global hydrological systems (Hidayat et al., 2017). As a result, hydrological processes are often poorly described, particularly for large-scale catchment modelling due to cost, time and technical limitations in conducting extensive field surveys (Coles et al., 1997). There is, therefore, limited knowledge of hydraulic processes and hydrodynamics of water flow necessary for hydrological modelling in several lowland tropical regions (Hidayat et al., 2017).

### ***Large populations vulnerable to flood risk***

Tropical lowland regions coincide with highly populated areas with concentrated economic activity and growth (Hidayat, 2013). The regions have one of the world's most rapidly growing populations due to availability of fertile soils, simple navigation and transport into inland regions (Harding et al., 2017). The major population centres are concentrated in fertile alluvial lands that sustain food production and availability of potable water and transport waterways along navigable rivers (Verhoeven and Setter, 2010). For example, over half of the world's population lives in Southeast Asia lowlands, accounting for about a quarter of the earth's land area (Kabir et al., 2013). Harding et al. (2017) study noted that by 2050, half the world's people would live in tropical regions.

Flooding affects most people in tropical lowland regions due to developments and settlements in flood-prone areas, lack of adequate flood risk management (Alfieri et al., 2017). As a result, floods account for the majority of victims each year due to vast flood-prone areas that are densely populated with altered natural conditions. About 95% of global floods are expected to result in combined flood losses in tropical regions, especially in Asia and Africa, with overall direct economic impact (Alfieri et al., 2017).

Rapid population growth, combined with economic development, means that environmental impacts will continue to increase the vulnerability to flooding of many people (Verhoeven and Setter, 2010). In some lowland regions, population increases have resulted in significant losses of forest cover, particularly where agricultural activities are prominent. For example, most agricultural lowland areas are heavily modified with ditches, drainage pipes, canals; river courses straightened, deepened and widened, including wetland areas reclaimed as landfill areas. When annual floods are experienced, the social, economic and ecological costs are staggering. As a response to flood disasters, most governments and communities in tropical regions invest more in flood protection infrastructure that addresses the effects rather than the causes of the flooding without long-term guarantee of protection.

Several countries in tropical regions face challenges in flood infrastructure development, but most times, they are limited by human and institutional resources and poor governance (Harding et al., 2017). What is clear and underlying is a challenging flood risk situation, however, there is lack of knowledge on flood dynamics, processes and extents of tropical lowland flooding that can be obtained from catchment-wide flood modelling and mapping. Although challenges and opportunities for infrastructure development exist, most times they are limited by human and institutional resources and poor governance that constrain progress (Cole et al., 2010; Harding et al., 2017).

## **3.2 Opportunities for satellite data**

The potential for satellite data in hydrological modelling has been recognised for some time. The earliest research on this potential took place in the early mid-seventies, where Landsat images for land use mapping were used for hydrological modelling (Andersen et al., 2001). However, very few satellite data sets were available for each water balance element to meet data requirements for hydrological modelling. Continual advances have been realised in the capability of remote sensing platforms or sensors to acquire datasets required to map hydrological processes, catchment properties and conditions such as rainfall, vegetation cover, surface water levels, soil moisture and other hydrological variables (McCabe et al., 2017).

Breakthroughs in sensor design and image processing algorithms have shown the potential of new platforms to represent hydrological cycle components in high definition (Chen and Wang, 2018). Many satellite missions have been launched, with some designed explicitly for hydrological research (Table 3.1). The Soil Moisture and Ocean Salinity Mission (SMOS), Soil Moisture Active and Passive (SMAP) and Global Precipitation Measurement (GPM) are rapidly providing useful datasets that can enhance flood modelling particularly in data-scarce regions (Chen and Wang, 2018). There is increasing satellite data production that has become too large and complex to handle using conventional data storage and processing methods. However, the satellite data increase in both quantity and quality is not matching the increase in the improvement in flood modelling in large data-poor regions. There is a gap between satellite data availability, improved quality and quantity, and its use in improving environmental risk management.

**Table 3:1: Summary of ongoing and future hydrological modelling missions (Source: Chen and Wang, 2018).**

Hydrological cycle component	Mission/sensor	Standard spatial resolution (km)	Standard temporal resolution	Launch year
Precipitation	GPM	5	0.125 days	2014
Evapotranspiration	Terra/MODIS	0.5	1 days	1999
	Aqua/MODIS			2002
	Landsat 8			2013
	Landsat 9			2023
Snow and ICE cover	ICESat-2	0.01	33 days	2018
	CryoSat-2	0.25	369 days	2010
Evapotranspiration	Terra/MODIS	0.5	1 days	1999
Soil moisture	SMOS	36	3 days	2011
	SMAP	36	3 days	2015
Streamflow	SWOT	0.1	11 days	2021
Groundwater	GRACE	220	30 days	2002
Water cycle	WCOM	2–5	–	2020

The increasing availability of satellite data in terms of quantity and quality is an opportunity to expand flood modelling and general environmental modelling in large data-poor trans-boundary catchments (Yan, 2015). Satellite data are potentially useful in developing large-scale hydrological models where modelling data is scarce and difficult to obtain (Sandholt et al., 1999). García-Pintado et al. (2015) note that satellite data has the potential and can enhance flood forecasting accuracy and reliability by accurately representing catchment conditions, processes and states in flood models.

Although there is an unprecedented wealth of satellite data, critical gaps remain in satellite data measurement capabilities. The estimates and mapping of evaporation levels, deep soil moisture, groundwater, snowfall and snow water equivalents remain limited. The unsaturated and saturated zone still cannot be represented in a physically based way, using satellite data (Andersen et al., 2001). There is a need for measurements and data retrieval algorithms that can provide continuous and aggregated water balance observations superior to the prevailing approaches (McCabe et al., 2017). Also, complementary approaches are required in remote sensing observations linked to in-situ measurements for data and model result validation to improve model accuracy. These approaches are required due to technological limitations of measuring instruments and data processing algorithms that result in low veracity and uncertainty in satellite datasets (McCabe et al., 2017).

Remote sensing measurements often result in data inconsistencies and discrepancies between the modelled and observed values of mapped catchment processes and conditions (Povey and Grainger, 2015). There are differences between the observed and modelled scales and these scale effects limit the accuracy of remote sensing datasets (Wu et al., 2014). The differences result in data errors, traditionally classified as random when the measurement error repeats and measurements differ unpredictably and, systematically if the error remains constant or varies predictably. Contrastingly, uncertainty describes the expected magnitude of the error that characterises its dispersion or distribution if the measurement was infinitely repeated (JCGM 200, 2012).

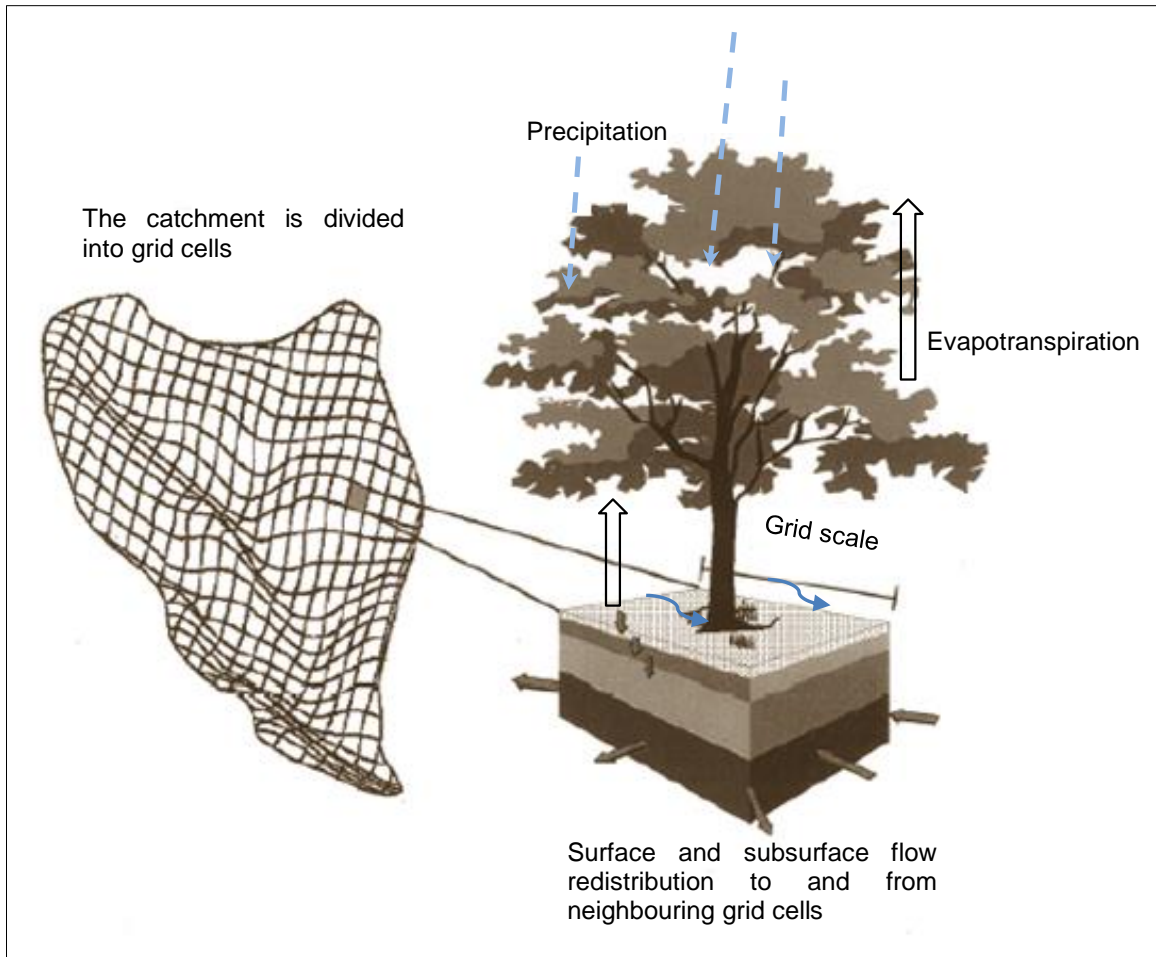
The design of most remote satellite observations involves many assumptions and methods that create unknown and quantifiable systemic errors (Kahl and Nachtnebel, 2008). Additional errors result from challenges of resolving scale variations in the measured phenomena (Povey and Grainger, 2015). For example, limitations of satellite datasets include data retrieval and interpretation algorithms which depend on measuring the spectral signature of reflected radiation where this information is used to retrieve a desired geophysical dataset (McCabe et al., 2017). Retrieval is achieved using complex models with many approximations and non-unique solutions to convert satellite data into a specific variable of interest.

In this respect, a quantified estimate of uncertainty and errors should be made so that satellite data sets are used correctly. However, the challenge is to quantify and characterise adequately individual uncertainties and errors and trace the uncertainty/error associated with each measured catchment condition and process (Chen and Wang, 2018). It is a challenge to assimilate satellite data successfully into models due to various uncertainties and errors that can be introduced, propagated and affect model simulations (Kahl and Nachtnebel, 2008). Although there are challenges in applying the available satellite datasets to flood modelling, the hydrological modelling process can integrate new paradigms such as data analysis and error quantification and use satellite data for accurate flood modelling (Chen and Wang, 2018). Recent data science analysis and interpretation technologies have been developed, including data analysis algorithms that can increase the number of satellite datasets that can be used in hydrological models.

### **3.2.1 Distributed hydrological models and satellite data**

The rapidly increasing availability of satellite data, computer power, cloud systems, the advancement of the Geographic Information Systems (GIS) and distributed hydrological models have the potential of increasing flood modelling accuracy and provision of reliable flood simulations in large data-scarce lowland tropical regions (Smith et al., 2012). The abundance of satellite datasets and their potential use in hydrological modelling have been noted; however, most of the studies have investigated the efficacy of one or two satellite datasets in limited selected small case studies and there is a growing need to realise their full potential (Mazzoleni et al., 2019). The development of satellite earth observations is speeding up the ability to collect data on highly variable catchment conditions and processes over extensive areas at a spatial resolution compatible with distributed hydrological models (Figure 3.4) (Abbott and Refsgaard, 2012).

The synergy between satellite data and distributed hydrological models offers opportunities for accurate mapping of temporal and spatial variations of catchment conditions and processes (Mancini et al., 2013; Chen and Wang, 2018). Distributed hydrological models divide the catchment terrain into grid cells (Figure 3.4) with the smallest catchment unit having unique terrain properties and process (Xuan et al., 2009). Hydrological processes are described on both the smallest and largest scales, allowing parameters to quantify hydrological conditions and processes that vary at each grid-scale up to catchment level. Also, when implemented spatially, from the hillslope to the continental scales, distributed models can integrate the spatial variations of catchment conditions and processes at a hillslope and catchment scale while also reflecting the subsurface heterogeneity in horizontal and vertical directions at each grid-scale (Fatichi et al., 2016). This capability enables the description of the heterogeneity of the smallest unit with terrain properties and processes up to the whole catchment scale (Chen et al., 2016).



**Figure 3.3: Distributed hydrological model representation of catchment conditions, properties and processes in discrete grid cells (Source: Vanshaar et al., 2002).**

The distributed feature of distributed hydrological models is fundamental in accurately simulating catchment conditions and processes at sufficient spatial resolution. It allows for comprehensive descriptions of hydrological conditions and state variables with varying heterogeneity and scales of different sub-catchment sizes (Tegegne et al., 2017). For example, distributed hydrological models can thoroughly define the highly spatial-temporal variability of rainfall across catchment areas and can define rain-runoff processes, particularly in tropical regions where highly variable and localised heavy rainfall often occur (Chen et al., 2016).

Although considered to be a new trend, the use of distributed hydrological modelling approaches for flood modelling is rarely applied, and their potential to represent variable catchment properties and processes accurately is not fully utilised (Chen et

al., 2016; Kauffeldt et al., 2016). This is because the scale and complexity of implementing the distributed hydrological model are totally out of proportion compared to the results that these models can achieve. Therefore, after two decades of modelling development and research, the use of distributed hydrological models is a fraction of their capacity (Abbott and Refsgaard, 2012).

Distributed hydrological models are considerably different from the widely used lump models in the representation of catchment conditions and processes. Lumped models consider that the entire catchment area has average values and uniform parameters that represent catchment conditions and processes (Abbott and Refsgaard, 2012). They represent the conditions and processes by averaging parameter values as unit values over the whole catchment. The terrain in the lumped model domain comprises average and constant values of parameters over the whole watershed, and this may not wholly represent the complex process in catchments studied (Chen and Wang, 2018). Therefore, lumped models, as the name suggests, lump catchment variables and physical parameters and they are easy to set up, requiring a shorter time to run as they have limited data requirements compared to distributed hydrological models. As a result, lumped models are widely applied for flood modelling as they require simplified catchment representations, fewer data requirements and less computational costs (Coron et al., 2017; Refsgaard et al., 1997; Chen and Wang, 2018).

However, distributed hydrological models represent, in detail, the observed conditions and phenomena. They are defined by their ability to incorporate the distributed watershed parameters and state variables comprehensively in the modelling framework. The distributed modelling approach is intended to make full use of available data, all relevant information documented in meteorology, land cover, soils, geology, river bathymetry and all other information relevant to the purpose of modelling (Singh, 2018). The grid-by-grid computations are based on a scalable Digital Elevation Model (DEM) and on the physical drainage properties that include location, slope, stream length and stream order. It is this detailed description of catchment properties and processes which enables process-based physically distributed models to offer significantly better accuracy and overall reliability of flood simulations (Abbott and Refsgaard, 2012).



Distributed hydrological approaches enhance flood modelling due to their detailed representation of catchment processes and conditions in models that allow the full use of available datasets (Chen and Wang, 2018). Satellite data utilises the raster data structure that fits well with the discrete digital grids of distributed hydrological models. The gridded nature of satellite data over broad areas, the digital acquisition of catchment processes and properties are linked to the gridded property of distributed hydrological models (Cole et al., 2010). This synergism between satellite data and distributed hydrological models coupled with the availability of the Geographical Information Systems (GIS) and increased computer power opens opportunities for maximum use of satellite datasets for large-scale distributed hydrological modelling (Figure 3.5).

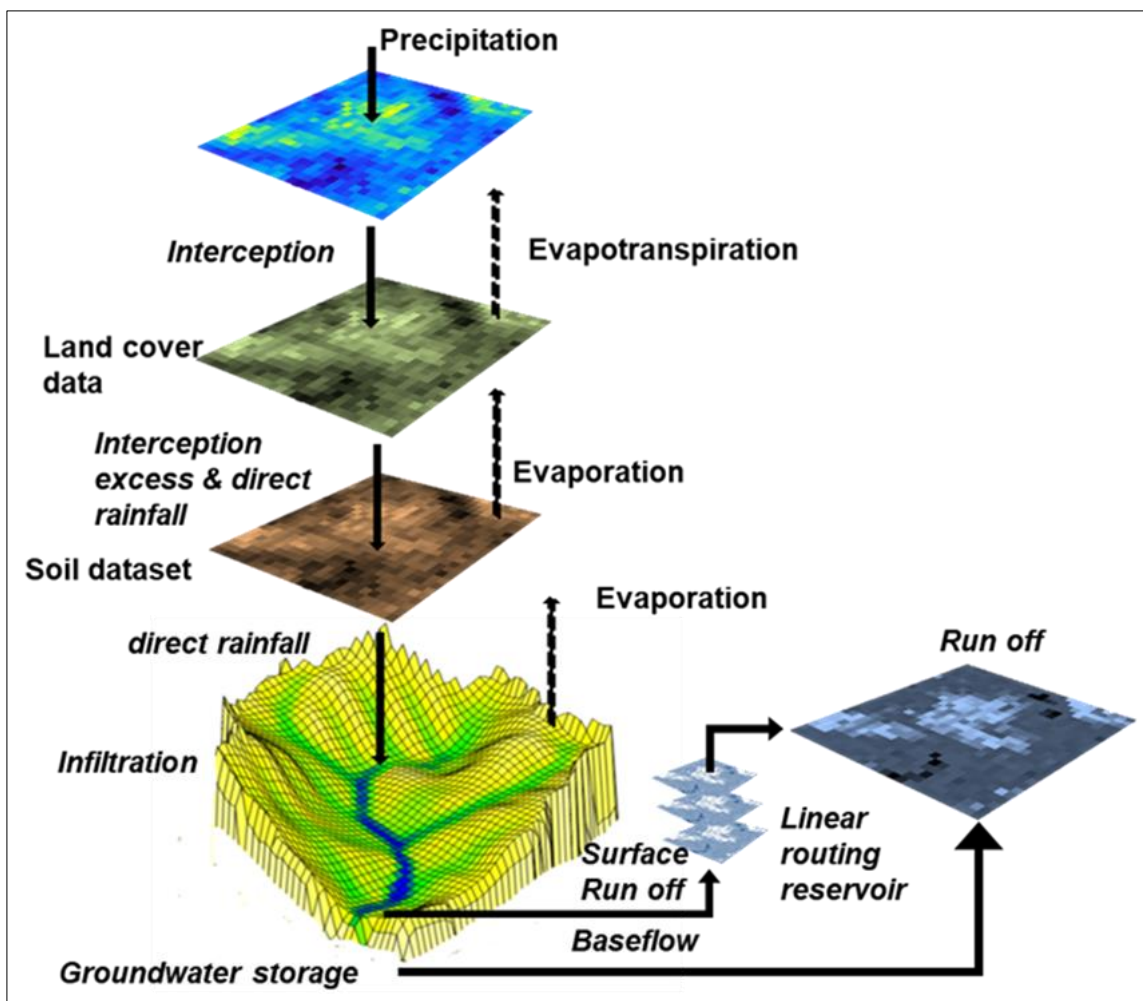


Figure 3.4: Synergism between satellite data and distributed hydrological modelling for detailed representation of catchment conditions and process. (Modified from Wigmosta et al., 1994).

Satellite datasets integrated with distributed hydrological models, therefore, provide a detailed description of catchment properties and processes that allows accurate representation of hydrological states, conditions and processes at grid-scale up to a watershed scale with different and variable parameters in each grid (Chen and Wang, 2018). However, lumped models cannot fully use satellite data sets because they are based on broad areas assumed to be homogenous over entire catchment areas (Sharad et al., 2018). The parameters in lumped models reflect a catchment with spatially averaged characteristics and are often not explicitly validated with field measurements. Lumped models typically use basic accounting methods to measure catchment processes and simulate time changes within a hydrological system (Yu et al., 2015). The advantage of the models is their easy conceptual parameterisation and less computation costs compared to distributed models. Simple lumped models have helped to model and understand large-scale patterns with self-similarity (scale invariance) features that can be mathematically explained using fractal theory and showing self-organisation of complex adaptive structures such as landscapes or flood quantiles (Fatichi et al., 2016).

There are, however, many situations in which thorough descriptions of complexity are needed to understand the dynamics of processes and the highly variable conditions and interacting processes in systems. Several semi-distributed (lumped) models have been developed in recent years with available digitally distributed and remotely sensed datasets (Yu et al., 2015). These semi-distributed models have been used widely for simulation of hydrological processes in climate, meteorological and hydrological studies, however, with the lumped characteristic of averaging catchment conditions and processes in sub-catchments units. Several semi-distributed models have been successfully implemented, including the Sacramento model developed by Burnash et al. (1995).

However, most of the conventional hydrological models do not fit well with satellite data, since they belong to the full lumped category of models which consider data inputs and state variables as mean values for whole catchments covering hundreds of km<sup>2</sup> in size (Sandholt et al., 1999; Andersen, 2001). Given the growing availability of satellite data and computer power, there is a need for accurate representation of

catchment characteristics and an improvement in flood modelling accuracy in data-poor regions where millions of people are at risk of flooding.

However, most studies point towards data availability still forming a significant challenge in implementing distributed models in large catchments (Beven, 2011). Full use of distributed models requires extensive data sets with accurate spatial and temporal details of natural parameters such as geology, terrain, soil, land cover and anthropogenic impacts (Abbott and Refsgaard, 2012). Distributed hydrological models require extensive datasets which may not be available, and this makes it difficult to set them up using traditionally surveyed data. High expectations have been placed on satellite data and remote sensing techniques for large-scale distributed hydrological modelling in ungauged regions. New satellite platforms have been launched in recent years, and it is now reasonable to expect the potential application of large distributed hydrological modelling in data-scarce regions (Abbott and Refsgaard, 2012).

The lack of reliable observed data for validating model data, results, and the uncertainty of representing parameters and catchment processes complicates the calibration of the distributed hydrological models. Distributed models can be challenging to develop and calibrate due to process complexity, interactions and scaling issues (Clark and Connolly, 2012). The expertise of the modeller is essential to use and configure the models as they require experience in calibration and requires a plausible control of model results. If a modeller understands and interprets the spatio-temporal distribution of parameters, these can be represented accurately in the model and generate reliable outputs.

Distributed hydrological models can be equipped with high-resolution data sets, can consider state variables and conditions to generate reliable flood simulations. The advancement of GIS tools and increased computer power facilitate the use of satellite data for distributed hydrological modelling (Andersen, 2001; Kherde et al., 2013). So far, the models have been successfully applied at small scales for flood forecasting purposes. However, challenges still exist in their application at large-scale due to reported issues of data availability and generation of errors (Chen and Wang, 2018). Also, the models have not been widely applied due to their prohibitively high computer

requirements, including time and expertise in calibrating the models (Lewis, 2016). With the growing availability of satellite imagery, increased computer power, cloud systems and the advancement of GIS tools, these models have the potential to use available datasets and generate improved simulations effectively. Distributed models are increasingly needed flood risk management tools and for effective representation of highly variable catchment conditions and processes (Abbott and Refsgaard, 2012).

One example of the application of a distributed hydrological modelling approach is the Grid-to-Grid model launched in England and Wales by the Environment Agency in partnership with the Flood Forecasting Centre, the Centre for Ecology & Hydrology (CEH) and Deltares in the Netherlands. The Grid-to-Grid model was required, as current and mostly implemented lumped hydrological models cannot maximise the use of high-resolution predictions of precipitation, and most are focused on large homogeneous areas (Clark and Connolly, 2012). To enhance flood forecasting, the Grid-to-Grid model was implemented and built on 1 km<sup>2</sup> high-resolution grid capable of capturing the variability of both hydrological and meteorological processes, enhancing the distribution of highly variable rainfall forecasts integrated with high-resolution global weather models (Cole et al., 2010). The Grid-to-Grid approach uses digital spatial topography, soil/geology and land cover details, and telemetry including Met Office rainfall forecasts for accurate flood predictions (Clark and Connolly, 2012). The model can calculate runoff from each grid redirected to the next downhill grid to the catchment source, where the model transforms rainfall into runoff and river flows.

Although distributed hydrological models were successfully applied for flood prediction in some data-rich countries, their implementation remains challenging in data-poor watersheds, particularly in tropical regions. Distributed hydrological models integrated with available satellite data are required to meet the demand for accuracy and reliability of flood prediction necessary for climate change adaptation and mitigation (Chen et al., 2016). Given the rapidly increasing availability of satellite data, computer power, cloud systems and the advancement of the Geographic Information Systems (GIS), this study sets out to explore the potential of increasing flood modelling accuracy in large data-scarce lowland tropical regions using distributed hydrological models.

### 3.2.2 High-resolution modelling at a large-scale

Regional and continental distributed hydrological models, most which are semi-distributed, have been implemented at low spatial resolution, with simplistic representations of catchment conditions and processes due to data scarcity resulting in simulation errors and uncertainties (De Paiva et al., 2013; Ibarra-Zavaleta et al., 2017; Yang et al., 2018). Most large-scale flood models simulate flood flows without detailed representation of highly variable catchment conditions and processes needed for reliable flood simulations (Thielen et al., 2009; Paiva et al., 2013, Alfieri et al., 2017; Pappenberger et al., 2012; Schumann et al., 2013; Sampson et al., 2015). The model domains are usually in the range of a few tens of square kilometres and are typically too coarse to reflect the complex catchment conditions and processes needed for local flood risk management (Yamazaki et al., 2011; Pappenberger et al., 2012; Schumann et al., 2013). For example, De Paiva et al., 2013) developed a large-scale Amazon River flood model using an in-channel hydrodynamic model equipped with a fill-in process to estimate storage volumes, and it did not have full floodplain hydraulics and could not dynamically reproduce flood dynamics in the study area.

While several high-resolution hydrological models with accurate representations of catchment conditions and processes have been successfully applied, they have mostly been applied at small scales (Khan et al., 2010; Dukic and Radic et al., 2016 and 2014; Zhang et al., 2015; Mourato et al., 2015; Birkinshaw et al., 2014; Op de Hipt et al., 2017). Compared to the SHETRAN model, developed in the Grijalva catchment area (57,960 km<sup>2</sup>) at 2 x 2 km grid spatial and 3-hour temporal resolution, examples of distributed hydrological models applied in tropical regions include: the CREST model that was applied to Nzoia sub-catchment (12 900 km<sup>2</sup>) in the Lake Victoria basin (Khan et al., 2010); the SHETRAN model applied to the Dano catchment (126 km<sup>2</sup>) Burkina Faso at 200 x 200 m<sup>2</sup> spatial and 1-hour temporal resolutions (Op de Hipt et al., 2017), the GSSHA model applied in the sub-tropical Guadalupe River basin (11,285 km<sup>2</sup>) in Texas at 150 m x 150 m grid and 15 minutes temporal resolutions (Chintalapudi et al., 2017) and the Geomorphology-Based Hydrological Model (GBHM) applied in Thailand (Wang et al., 2016).

Moreover, the majority of large-scale distributed hydrological models are primarily semi-distributed and cannot provide an accurate representation of hydrological variables and are susceptible to rainfall and parameter error due to their inability to completely use satellite data for parameter representation (Ibarra-Zavaleta et al., 2017; Paiva et al., 2013). However, the Variable Infiltration Capacity (VIC) model is a macroscale distributed model and is planned for extensive areas but has a relatively low spatial resolution (Yang et al., 2018).

Besides these research efforts, this research applies a large-scale distributed hydrological modelling approach that is sufficiently detailed to fit observed data with much finer spatial and temporal resolutions in a tropical lowland region (Schumann et al., 2013). The study applies a catchment-wide flood modelling approach with sufficient detail of local scale flood dynamics and extents. The research contributes to closing a gap that exists between simplified, coarse resolution large-scale modelling approaches on the one hand and detailed high resolution small-scale hydrologic modelling approaches on the other (Sampson et al., 2015). The study further contributes to meeting the challenge of accurate representation of highly variable catchment conditions and processes and the provision of reliable flood simulations in large data-scarce tropical regions where an increasing number of people are at risk of flooding (Schumann et al., 2013).

### **3.2.3 Summary**

Recognising the extreme flood risk situations that are projected to increase due to climate change, steps must be taken to achieve successful flood risk management in tropical regions. The development of accurate flood models is one of the first crucial steps towards climate change adaptation and successful mitigation based on early warning systems and provision of timely and reliable flood simulations required for sound decision-making in flood response situations. Despite the worsening flood situation, lowland tropical regions remain ungauged and lack in-situ measurements for setting up flood models and yet have significant populations prone to flooding. Flood modelling systems are critical in these regions and are urgently required to protect lives and property.

Although the development of flood models in lowland tropical regions is a challenge in data scarce tropical regions, satellite data can provide large-scale modelling data for large and highly variable catchments. A wealth of satellite platforms, some with solely hydrological observation missions, have been launched, leading to an increase in the availability of datasets for hydrological purposes. Given the increasing availability of high-resolution satellite data, computer power, cloud systems, improved GIS tools and available distributed hydrological models, opportunities exist to improve flood modelling in large lowland tropical regions. Although systematic and random errors associated with data and the modelling process present limitations on maximum use of satellite data, this could be resolved by employing error and uncertainty quantification approaches to improve the accuracy and reliability of model results. Techniques that account for model, parameter and input errors and uncertainties can contribute to the accurate incorporation of satellite data sets into models in a manner that is efficient, transparent and complete with the potential to improve the reliability of the forecasts produced.

### **Research gap**

Given the increasing flood risk, most lowland tropical regions remain ungauged, with limited in-situ measurements and related records that are required to set up flood models. Accurate flood modelling and reliable flood simulations are needed in these regions where a growing number of people are at risk of flooding. However, flood modelling in large tropical lowland regions is a challenge owing to the large scale of the basins, highly complex catchment conditions and processes, limited availability of data and associated rapid changes in the natural environment due to human alterations. The challenge is the accurate representation of highly variable catchment conditions and processes within the model and the provision of reliable flood simulations.

Given the increasing availability of high-resolution satellite data, computer power, cloud systems, improved GIS tools and available distributed hydrological models, there is a potential to increase flood modelling accuracy in large data-scarce tropical regions. Integrating high-resolution satellite data with distributed hydrological models presents opportunities for detailed representation of catchment processes and conditions in the model, maximum use of available datasets and the provision of reliable simulations.

Accuracy in flood modelling and the provision of reliable flood simulations in tropical regions can be made through optimally integrating satellite and limited local datasets using distributed modelling approaches with the best combination of the model parameter, input and output datasets based on error and uncertainty quantification and removal. There is a need to explore the potential for increasing flood modelling accuracy and providing reliable simulations in data-scarce tropical regions using increasingly available high-resolution satellite data and make use of available under-used distributed hydrological models.

Therefore, required is high-resolution modelling across large and highly variable domains to help close the knowledge gap that exists between simple, coarse resolution large-scale modelling approaches on one hand and high-resolution, small-scale hydrological modelling approaches on the other (Sampson et al., 2015). Very few studies have explored this issue, so it is worth exploring further.



### 3.3 The scaling theory of floods

This section provides a literature review on the application of the geophysical scaling theory of floods to improve the estimation of the magnitude of peak discharge in La Sierra sub-catchment area. It is proposed that statistical scaling relationships between peak discharge and physical catchment conditions and processes can be developed into statistical models for estimating flood parameters and the magnitude of peak discharge in the sub-catchment area. The objective is to investigate statistical scaling relationships of rainfall variability, flood peak discharge and nested catchment drainage areas to complement the estimation of peak discharge magnitudes across La Sierra catchment.

#### 3.3.1 The scaling theory of floods

The geophysical scaling theory of floods also referred to as the scaling theory, has a specific aim of connecting rainfall-runoff events to statistical power laws between peak discharge and drainage areas across space and time scales (Gupta, 2017). The core hypothesis is that a scale-invariant spatial peak discharge structure can be defined by a power-law relationship from each rainfall-runoff event (Ayalew et al., 2018). Several research results over the last 20 years have provided the theoretical and empirical basis for formulating the geophysical theory of floods in river catchments (Gupta et al., 1996, 2007, 2010, 2015; Robinson and Sivapalan, 1997; Menabde and Sivapalan, 2001; Ogden and Dawdy, 2003; Gupta., 2004; Dawdy et al., 2012). Significant research has been devoted to establishing relationships between peak discharge from nested catchments and individual rainfall events (Ayalew et al., 2014). It has been consistently found that peak discharge observed in nested catchments show a power-law variation with nested catchment drainage areas at a scale of a single rainfall-runoff event as:

$$Q_p = \alpha A^\theta \quad \text{Equation 3.1}$$

Where  $Q_p$  is peak discharge,  $A$  is nested catchment drainage area,  $\theta$  is the scaling exponent parameter, and  $\alpha$  is the scaling intercept parameter.

Gupta et al. (1996), first established the power-law relationships between flood peak discharge and drainage areas, but the results raised many questions, and many remain unanswered. Some findings were validated by several empirical-based studies mainly in the USA, namely by Ogden and Dawdy, (2003), Furey and Gupta, (2005), Gupta et al. (2010) and Ayalew et al. (2015) with some using computational rainfall-runoff model simulations in synthetic and natural river basins (Gupta et al., 1996, 2007; Gupta et al., 1998; Mantilla et al., 2007, 2011; Mandapaka et al., 2009; Ayalew et al., 2014). In a 21 km<sup>2</sup> Goodwin Creek Experimental Watershed (GCEW) in Mississippi (USA), Ogden et al. (2003) conducted and published the first empirical study on the relationship between peak discharge, drainage areas and rainfall. Subsequently, Furey and Gupta, (2005) examined Ogden and Dawdy (2003) results from the same Goodwin Creek Experimental Watershed. Gupta et al. (2010) analysed the effect of catastrophic floods in the Iowa River basin in June 2008 and found out that the power-law scaling relationship persists despite reasonable rainfall in the catchment.

However, most studies have been conducted in the United States, with few in the United Kingdom investigating the physical catchment conditions and processes that control flood scaling parameters using empirical data sets (Ayalew et al., 2015; Wilkinson and Bathurst, 2018). A recent UK study by Wilkinson and Bathurst, (2018) examined the scaling relationships between spatial rainfall variability and flood response across nested catchments. Most ground-based knowledge of the generation of floods has been experimentally based on scales less than 10 km<sup>2</sup> in wet northern latitude catchments in the United States. As a result, the non-linear geophysical relationships between peak discharge and drainage areas at different spatial scales do not provide a solid basis for large-scale extrapolation to other climatic regions.

The issue is a typical trend in flood scaling theory studies; they have been conducted in humid northern latitude regions, and it is not well known whether the theory applies to other climatic regions. The mechanism of producing floods from small headwater nested catchments into floods in larger catchments is not yet fully quantified, and the universal application to other climatic regions remains unknown (Wilkinson and Bathurst, 2018). Scaling relationships in different climatic regions need to be

investigated, considering different rainfall coverages, and contribute to the question of flood prediction in ungauged catchments (Wilkinson and Bathurst, 2018).

Tropical regions are characterised by large quantities of heavy convective rainfall, normally local with partial catchment coverage, generally resulting in flash floods. It has been reported from humid temperate studies in the USA that the flood scaling theory applies only to runoff events resulting from large-scale rainfall events with full catchment coverage (Ayalew et al., 2015). The studies showed that peak discharge could not be estimated if the studied catchment remains partially covered by rainfall and the peak discharge produced in such sub-catchment could not provide a scaling invariance with the drainage areas (Ayalew et al., 2014). Research by Ayalew et al. (2015) analysed the combined effect of rainfall intensity, runoff speed, channel flow speed and drainage network layout and quantified their relationships with flood scaling parameters. The research findings show that a scale break exists in partial rainfall-runoff processes on hillsides where surface runoff flows quickly in drainage systems.

However, Wilkinson and Bathurst, (2018) in a multi-nested catchment experiment in the UK considered all storm events, both partial and full catchment coverage, observed that all events chosen resulted in some scaling relationships. The results showed that observations of peak discharge events resulting from catchment-wide rainfall can still predict flooding, even if the entire catchment has limited rainfall coverage (Ayalew et al., 2015). Therefore, the flood theory needs to be further assessed to identify the spatial scale properties of peak discharge due to partial rainfall coverage of heavy and intense rainfall in tropical catchments.

Although theoretical progress in linking the flood scaling exponent ( $\theta$ ) to rainfall and catchment properties and conditions differs from event to event, limited progress has been made in testing the theory in large tropical catchments (Ayalew et al., 2014; Ayalew, 2015). Several studies have been conducted to investigate spatial variations of flood response in plots/hill slopes or small headwaters. Catchments studied were small, with event data limited to peak discharge scaling from hill slopes without linking to the broader catchment areas (Wilkinson and Bathurst, 2018). Ayalew et al. (2015) responded to this and extended the scaling relationships into the Iowa River Basin

(32,000 km<sup>2</sup>) for several types of flood events and noted that the temporal and spatial nature of rainfall affects the flood peak scaling relationships.

Although the relationship between exponent ( $\theta$ ) and intercepts ( $\alpha$ ) has been highlighted to result from excessive rainfall covering entire catchment areas, further studies are needed to establish whether these results are evident in tropical catchments. No study in the literature has shown whether the results from these studies hold in larger tropical watersheds where localised heavy and intense rainfall often occur. The actual process of generating floods from small headwater catchments into larger floods in large catchments is still poorly quantified (Wilkinson and Bathurst, 2018).

In this context, the flood scaling theory needs further exploration to include large tropical river basins, in which convective rainfall and isolated flood events occur often. Present knowledge on the variation of flood generation in tropical catchments as a function of the variability of tropical convective rainfall is lacking. Scale dependencies of flood peak scaling relationships in tropical regions need to be explored in all landscape scales, considering all rainfall coverages (Wilkinson and Bathurst, 2018). Besides, what is required is an evaluation of the effect of tropical convective rainfall variability on peak discharge to understand the spatial catchment flood response characteristics (Ayalew et al., 2015). The challenge that remains is bridging the scaling gap, by applying nested catchment studies that extend to other climatic regions with different rainfall coverage and types (Wilkinson and Bathurst, 2018).

Such experimental studies will involve defining rainfall-runoff-event variability and related flood parameters and establish relationships with observable catchment conditions and processes in drainage areas and estimate the overall flood parameters and peak discharge across spatial scales (Ayalew, 2015). The results are essential in solving the flood prediction problem in ungauged basins (Sivapalan et al., 2003). The approach will also provide a flood prediction technique that is not affected by climate change as it is not based on validation using observed discharge (Dawdy et al., 2012; Gupta, 2004 and 2010; Gupta et al., 2007). Further, it provides a framework to verify the magnitude of peak flow simulations provided by numerical rainfall-runoff models.

### **3.3.2 Summary**

The study investigates the power-law relationships between peak discharge, catchment properties and processes that control the generation of flooding in nested sub-catchments in La Sierra catchment to improve the estimation of the magnitude of peak discharge. There is need to explore the variations in the generation of floods in nested catchments as a result of spatial-temporal rainfall variability and the power-law relationships between drainage areas and peak discharge from tropical rainfall events that partially and entirely covers catchments. The study applies the recent flood scaling theory with the explicit objective of integrating catchment flood-producing processes through an understanding of scaling relationships between peak discharge, and drainage areas in tropical regions. There is a need for complementary mechanisms for estimating the magnitude of peak floods in data-scarce regions. The theoretical framework of statistical scaling relationships is investigated in the catchment area to check, track and improve the simulation of the magnitude of the peak discharge.

### **Research gap**

The literature review of the flood scaling theory has shown that there is still a lack of multi-scale nested catchment studies in different climate regions, involving scaling relationships that relate factors and processes involved in the generation of flooding from headwater catchments to broad or entire catchment areas. Few studies have been conducted to examine catchment conditions and processes that govern flood scaling parameters using empirical data sets, most of which have been conducted in the United States and the United Kingdom to date.

The mechanism by which floods are generated in small sub-catchments leading to flooding in large catchments in tropical regions is still not yet quantified. There is a lack of knowledge on tropical catchment flood response due to highly variable tropical rainfall. A gap still exists on how upstream flood response, contributes to flooding in downstream sub-catchments and more background data from nested catchment studies are required, particularly in data-poor tropical regions. The scaling relationships between peak discharge and nested catchment drainage areas in tropical climatic regions should be explored, considering the variability of rainfall in different catchment units. In this respect, this study expands the application of the scaling theory of floods to large tropical river basins, in which convective rainfall and localised flood events frequently occur.

### **3.4 Propagation of uncertainty through hydrological and hydraulic models**

This section presents a literature review on the propagation, interaction and effects of uncertainties on discharge simulations and flood extents through coupled hydrological and hydraulic models. Presented is an examination of literature which supports current knowledge of uncertainty propagation and interaction in coupled models and identifies deficiencies, discrepancies and knowledge gaps. The objective of the review is to examine the effects of propagation and interaction of rainfall and parameter uncertainties on the magnitude of peak discharge and flood inundation extents in the study area.

#### **3.4.1 Uncertainties in flood modelling**

Reliable flood simulation is critical for successful flood risk management; however, the estimation usually relies on data and models that are prone to uncertainties. The estimation of rainfall data, the model structure, parameters, discharge output, and flood extents are all uncertain. The accuracy of the flood volumes and the flood map derived from the modelling process depends on a variety of uncertainties from various sources. It is recognised that all models are imperfect and are a simple representation of catchment conditions and processes, so quantifying uncertainties resulting from this generalisation of the process is necessary (Rodríguez-Rincón et al., 2015).

However, a critical question that arises in coupled modelling for flood volumes and mapping is how uncertainties associated with rainfall estimates propagate and interact with other uncertainties and how each uncertainty affects flood discharge outputs and flood inundation extents. Due to several sources of uncertainties in the flood modelling process and data, the distribution of individual uncertainties and how they affect the overall uncertainty in the model results are unclear (Jung and Merwade, 2015; Lim and Lee, 2018). The challenge is how to disentangle the combined uncertainty in flood models and quantify individual uncertainties considering the complex interactions and effects on model outputs (Blöschl et al., 2019).

### 3.4.2 Sources of uncertainty in flood modelling

The source of uncertainties in hydrological modelling is the simplification of the modelling process, the inadequate definition of catchment conditions and processes in the model and the inability to identify accurate data for model development (Oubennaceur et al., 2018). As a result, hydrological modelling provides a simpler conceptual representation of complex catchment conditions and processes using simplified spatial and temporal data across broad and complex areas (Anees et al., 2017). Any process of modelling involves a reducible and inherent uncertainty of the system of interest, from data, model abstractions and natural heterogeneity (Merwade et al., 2008). Thus, flood modelling has several sources of uncertainties, including model input data, parameter data, model structure and topography (Di Baldassarre et al., 2010).

Model input uncertainties relate to input data, mostly direct rainfall time series that is often estimated in theory and selected model parameters values from the literature. Uncertainties manifest in different locations of the hydrological model, and the locations are the basis for uncertainty classification, that is according to where they originate. Similarly, a recent study by Refsgaard et al. (2013) divided sources of uncertainty into input data (such as external input data) model uncertainties and context uncertainties (comprising model boundaries). However, uncertainties from model structure, inputs and parameter sources can combine and contribute to model output uncertainty. Also, simulations can be compromised further through interactions with individual uncertain input and parameter values.

Although inaccuracies in the input and parameter data vary considerably based on unique sources and constrained by model algorithms, they are propagated individually through the model. In the process, they combine with other uncertainties from different sources to affect the magnitude of peak discharge and flood extents. Computing the full spectrum of sources of uncertainty that interact and combine is a challenge to tackle rigorously, isolating each source and avoiding confusion between them in the process (Tscheikner-Gratl et al., 2018).



### 3.4.3 Uncertainty propagation

The propagation of uncertainty occurs when input data and model uncertainties in coupled models lead to uncertain model outputs, and the initial uncertainties are said to have propagated through the model chain (Brown and Heuvelink, 2006). Studies on uncertainties in flood inundation mapping have investigated the impact of different uncertainties in flood inundation outputs and reported that individual uncertainties of varying magnitudes affect model outputs, but can interact or combine to produce substantially different simulation results (Jung and Merwade, 2015); Bates et al., 2014).

It has been widely reported that data input uncertainty, particularly rainfall, significantly affects the overall predictive uncertainty (Liu and Gupta, 2007; Jung and Merwade, 2015). However, the combination of uncertainties and interactions that exists in models can affect model outputs in various magnitudes (Bates et al., 2014). That means rainfall uncertainties can interact with parameter and model structure uncertainties as they are propagated through the model leading to greater uncertainties in model outputs.

Besides uncertainty interactions, parameter variables and their spatial and temporal dependencies could affect uncertainties as well. The dependencies can combine with increasing input and output uncertainties and can alter model parameter values and affect model outputs (Merwade et al., 2008). For example, the impact of rainfall uncertainty on the final flood map can be restricted if the hydraulic model is more sensitive to topographic data. Nonetheless, the contribution of a specific uncertain source in model simulations depends on the model's sensitivity to that uncertain variable and other related uncertainties (Brown and Heuvelink, 2006).

Therefore, quantifying and evaluating the contribution of different uncertainties to overall model output uncertainty is essential to identify sources that contribute significant amounts of uncertainty, enabling direct efforts to be directed to primary contributors. Considering and accounting for the sensitivity of individual uncertainties, their relative interactions, and dependencies with precise quantification of each are essential.

#### **3.4.4 Uncertainty analysis**

Knowledge of the magnitude of uncertainties in flood modelling is critical to understanding and interpreting information on flood risk (Oubennaceur et al., 2018). Uncertainty analysis should always be performed rigorously to help and improve modelling results, and any flood risk assessment should be conducted within a system of uncertainty quantification (Rodríguez-Rincón et al., 2015). Accurate quantification, characterisation and communication of uncertainties in modelling are essential for effective flood risk management and sound decision making during flood emergency response (Montanari, 2011). However, a study by Merwade et al. (2008) highlights that the propagation of uncertainties within a cascading model chain is still not understood, particularly the complex interactions between uncertainties from individual sources including effects on model results.

While uncertainties can be reduced, and model performance improved, uncertainties cannot be eliminated completely (Pappenberger et al., 2005). A key aspect is, therefore, a clear understanding of the uncertainties associated with many sources of flood modelling using probabilistic flood maps for decision-making. Confidence intervals are used to characterise ranges (5% and 95%) with a likelihood for the average estimate. Tolerance intervals may also be defined that include a certain proportion of unsure model calibration estimates considering all probable outcomes (Pechlivanidis et al., 2011; Rodríguez-Rincón et al., 2015).

The quantified uncertainty results can help to decide on less expensive solutions with a better understanding of other probable outcomes, enabling accurate decision-making and targeting resources towards uncertainty sources that contribute most to the overall uncertainty. Gaining knowledge of uncertainty characteristics in terms of type and magnitude is crucial for meaningful interpretation of model outputs (Warmink et al., 2011). Quantifying and identifying uncertainties is important for the generation of simulated results with quantified flood probability ranges (Thielen et al., 2009; Coccia and Todini, 2011). The quantified uncertainties improve flood risk management techniques focused on a transparent and truthful uncertainty analysis

aimed at the most probable outcome of the flooding, Probabilistic flood inundation maps.

Frequently, flood maps are interpreted as a single deterministic flood outline without considering uncertainties associated with uncertain input data, initial conditions, parameters and numerical model limitations (Zarzar et al., 2018; Merwade et al., 2008). In reality, estimates of rainfall data, model structure, model parameters and flood outlines are uncertain (Merwade et al., 2008). Typically, flood inundation models make no allowance for these uncertainties, yet significant uncertainty arises from the data and parameters that are used to build models (Rodríguez-Rincón 2015). Specific hydrological data sets, equations, simulation parameters and methods of processing have a relative uncertainty on flood maps (Lim, 2018). Thus, in flood modelling processes, various internal and external uncertainties influence the accuracy of flood simulations produced.

Despite uncertainties, however, flood maps are considered reliable, the reason why several studies suggest the inclusion of uncertainties in flood modelling results is to provide more flood information and improve reliability (Lim and Lee, 2018; Di Baldassarre et al., 2010). Without including uncertainties, determinist flood maps cannot convey to decision-makers and people at risk the full risk of flooding (Zarzar et al., 2018). Consumers must be made aware of leading modelling factors, their weakness and how they influence model outputs. Including uncertainty information in the decision-making process provides transparent communication which is considered to be critical and necessary.

Several studies have so far focused attention on investigating uncertainties in flood inundation model outputs derived from different methods of representing different initial meteorological conditions or differences in using a single model set versus a multi-model set (Pappenberger et al., 2008; Cloke and Pappenberge, 2009; Rodríguez-Rincón 2015). However, uncertainties have been identified as a challenge to quantify and remove in the final flood map, owing to limited financial resources, insufficient processing times, imperfect knowledge of flood science, and lack of an agreed coherent framework of quantifying uncertainties. Over the last four decades,

structural model uncertainty has been an issue in hydrology and is more elusive than input models and parameter uncertainties (Kirchner, 2019).

Analytical propagation methods rarely apply to every model because many have complex structures and because rainfall has a multivariate distribution with statistical characteristics that vary in space and time and its ensemble generation requires consideration of spatial-temporal characteristics of the process. The most popular and effective technique identified to evaluate rainfall input, uncertainty propagation and quantification has been the use of an ensemble approach.

Despite efforts, the question that remains is how model structural/parameter/input uncertainty can be disentangled and quantified in hydrological modelling simulations (Blöschl et al., 2019). No attention has been paid to the propagation and interaction of uncertainties in coupled hydrological and hydraulic models, including the effect of complex interactions on flood model outputs (Jung and Merwade, 2015; Lim and Lee, 2018). What remains unknown are the effects of individual propagated uncertainties, their interactions with other uncertainties within the modelling chain and the effects of each uncertainty on flood volumes and inundation extents (Rodríguez-Rincón 2015). Therefore, a critical issue that arises in applying a coupled modelling approach to flood simulation or mapping is how uncertainties that emerged in input data interact and propagate into discharge outputs and flood inundation extents.

### 3.4.5 Summary

Reliable flood simulations are critical for successful flood risk management; however, flood estimates are based on data and models that are prone to uncertainties. The estimation of rainfall data, model structure, model parameters and flood simulation is uncertain. It is also known that all models are imperfect and are a simple representation of catchment conditions and processes. Thus, flood modelling has several sources of uncertainty, namely model input data, structural uncertainty, and the different parameter ranges that match sparse calibration data and provide uncertain simulation results.

The full range of sources of uncertainty in flood modelling is a challenge to analyse rigorously, separating each from corresponding sources and avoiding confusion between (Tscheikner-Gratl et al., 2018). It is not well known how individual uncertainties propagate interact and affect the overall uncertainty in model results. The question is, how to disentangle the combined input/structural/parameter uncertainty, considering the complex interactions and effects on model outputs. A critical issue in flood modelling is how uncertainties that arise in rainfall and parameter data propagate and interact through coupled hydrological and hydraulic models into model results.

Given the wide range of sources of uncertainty in flood modelling, the distribution and interaction of individual uncertainties from the beginning of modelling to the endpoint is of great importance. Gaining knowledge of uncertainty characteristics in terms of type and magnitude is crucial for meaningful interpretation of model results. This knowledge is vital for producing simulated outputs with quantified flood probability ranges crucial to flood risk-based decision-making.

### **Research gap**

There is little research on how uncertainties in rainfall estimates propagate and interact with catchment level parameter uncertainties and the quantification of their effects on the magnitude of flood volumes and flood inundation extents. Current research concentrates on quantifying the overall uncertainty in model outputs without exploring the decomposition of individual uncertainty contributions from different sources to model output.

Due to several sources of uncertainty in flood modelling that compromise model simulations, there is a need to identify and quantify individual uncertainties in-order to improve the reliability of model results. It is essential to measure individual propagated uncertainties, evaluate their effect within the model chain, analyse the interactions with other parameters and determine the effects on discharge output and flood inundation extents. Clear and specific quantification of uncertainties related to the propagation of uncertainties in hydrological and hydraulic model simulations is essential for objective assessment of model merits and limitations (Diaz-Ramirez et al., 2013).

This study responds to the realisation that the deterministic representation of flood outputs and flood maps used in flood risk assessment and mapping is not the most proper representation of flood risk. The approach conveys absolute information despite uncertainties in the flood inundation map-generation process (Di Baldassarre et al., 2010; Lim and Lee 2018). The incorporating quantified uncertainties in model outputs helps to communicate the full risk of potential flooding to decision-makers and people at risk (Zarzar et al., 2018). Users of model simulation outputs must be informed of the conditions used, the constraints and the impact of the final model results in the modelling framework.

## 4 METHODOLOGY

This chapter describes the methodologies used to identify and apply physics-based and data-based hydrological modelling approaches for large-scale flood modelling in lowland tropical regions. Section one (4.1) presents the key characteristics of the study area, highlighting the large-scale, complex and highly variable catchment conditions and processes that influence modelling river flows. The hydrological modelling steps from model set up to simulation outputs follow the method outlined by Anderson and Woessner (1992) (Figure 4.1). The selection of the La Sierra catchment, a sub-catchment of the main Grijalva catchment, is presented for the application of the scaling theory of floods and the study of uncertainty propagation and quantification. The SHETRAN hydrological model and the Flood Modeller are described in sections (4.2 and 4.4) respectively, including their selection and suitability for producing various simulations and scenarios for analysis.

The data specification and development of each model are described, including the identification, source and definition of both satellite and in-situ datasets (Subsection 4.2.2 and 4.4.2). The methods used for data processing, quality checks and data error removal to enable the correct representation of catchment conditions and processes in the models are provided. The development of the hydrological models, including data processing, parameterisation, calibration, validation, and model performance evaluation, are discussed (Section 4.3 and subsection 4.6.3). Furthermore, the chapter presents a methodology adopted for developing statistical scaling relationships and the models developed to complement the estimation of the magnitude of peak discharge in La Sierra catchment (Section 4.5). To improve the reliability of model results, the chapter presents the methodology adopted for setting up coupled/cascading hydrological and hydraulic models for uncertainty propagation and interaction in the catchment (Section 4.6).

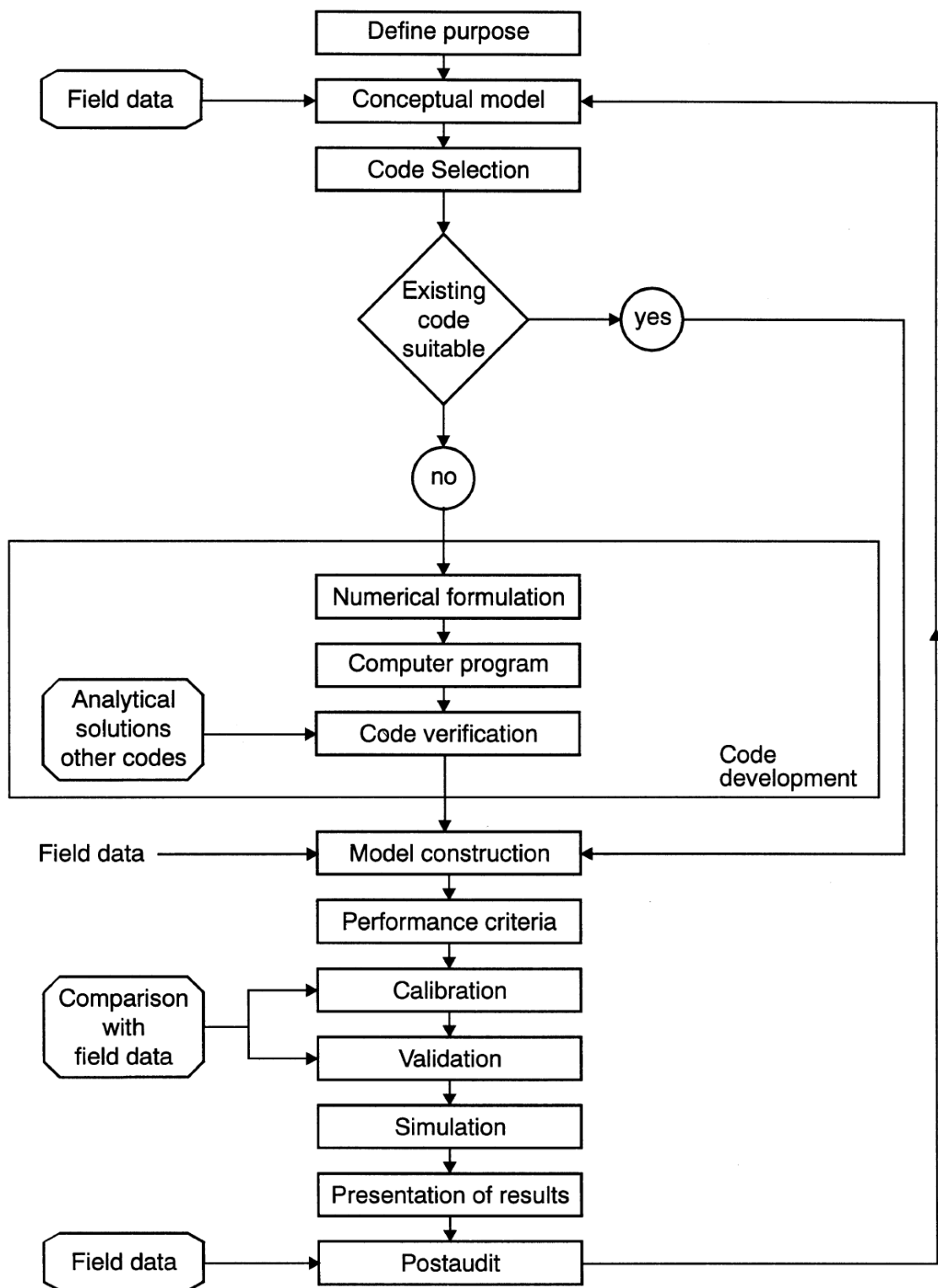


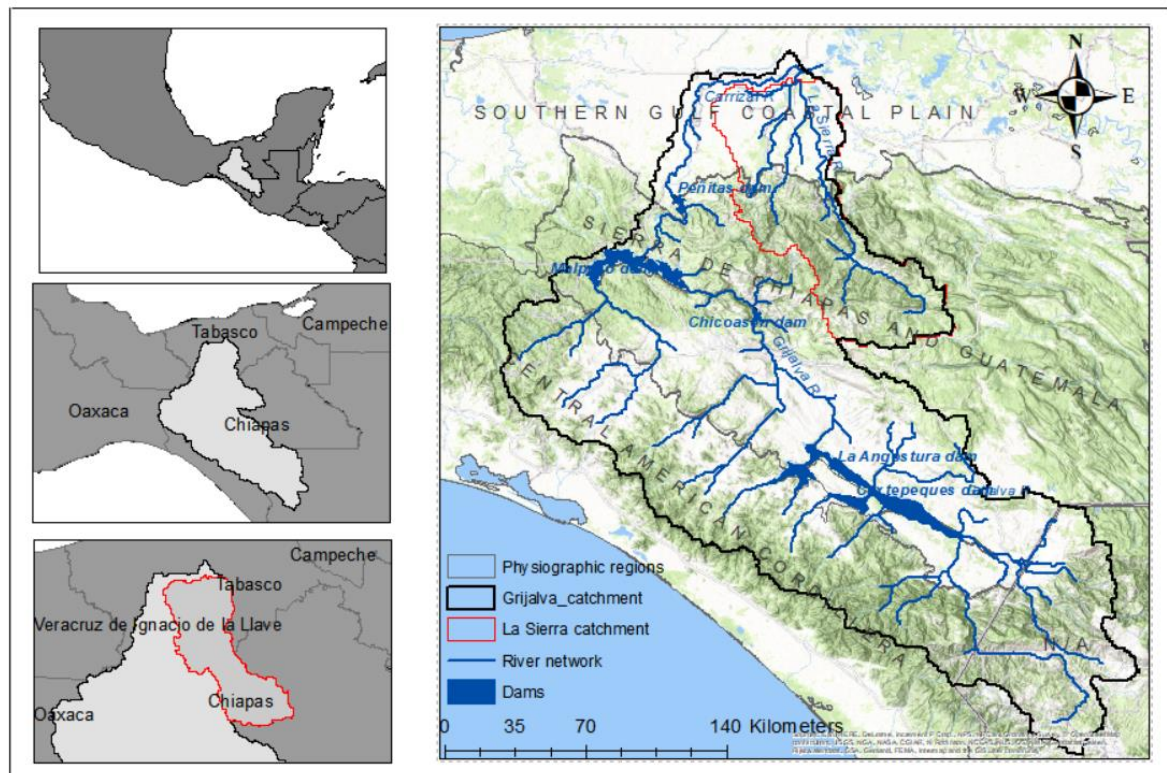
Figure 4.1: The modelling methodology, adapted from Anderson and Woessner (1992).



## 4.1 Study area

### 4.1.1 The main Grijalva catchment

The study area covers the tropical lowland areas of Tabasco and upland regions of Chiapas and Guatemala, southeast of Mexico, within the Grijalva River catchment (Figure 4.2). The catchment area is 57,958 km<sup>2</sup> in size, bounded by 89.6° and 94.5°W and 15.3° and 18.7°N occupied by lowland plains in Tabasco State in the north, mountainous highlands in Chiapas State and Guatemala in the south and Veracruz to the west (Aparicio et al., 2009). The Grijalva catchment area includes vast low-lying Gulf coastal alluvial plains in the north with an average altitude of 10 metres above sea level with some elevations not exceeding 30 metres. In the southern highlands of Chiapas and Guatemala region comprise highly dissected mountain ranges and plateaus reaching elevations of 2 100–2 400 metres above sea level (INEGI, 2018). Some mountain ranges are situated in the south of Grijalva basin and the most notable are Sierra Tapijulapa with 900 metres, La Pava 860 metres, and Sierra Puana 560 metres.



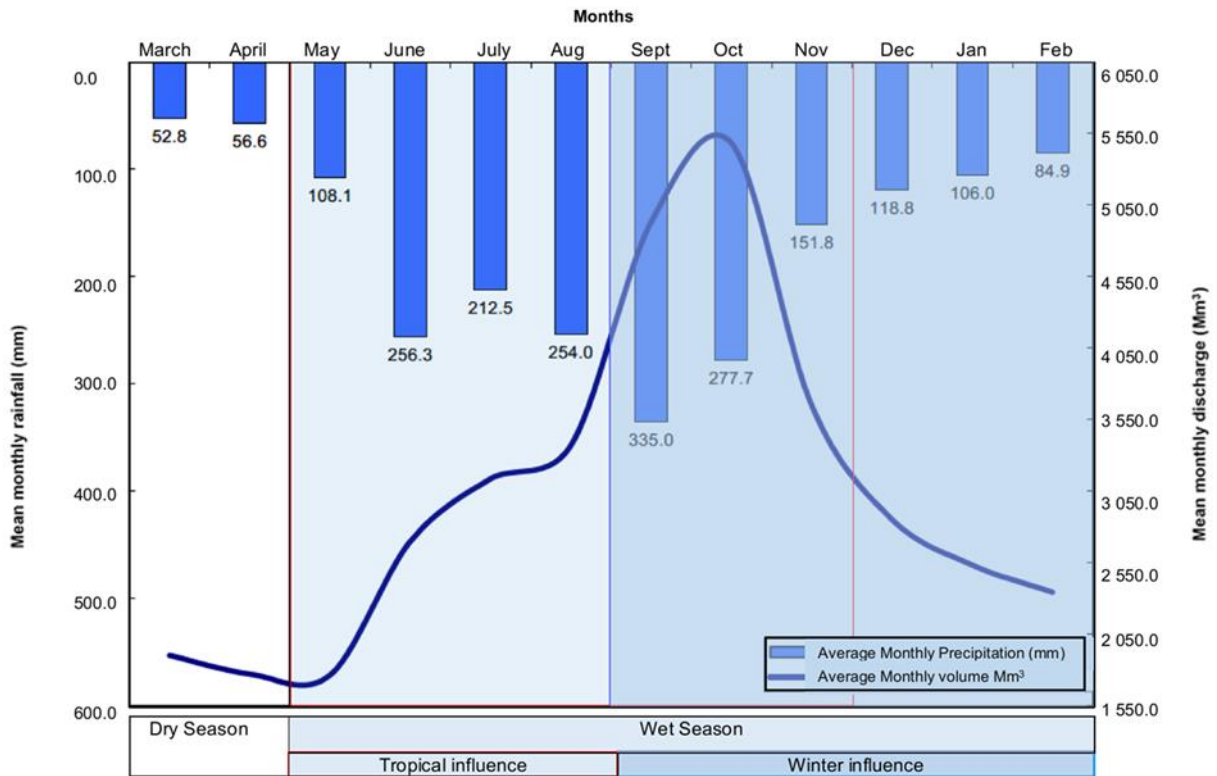
**Figure 4.2: The location of the main Grijalva river catchment and La Sierra sub-catchment areas in the southeast of Mexico. The main river network and major dams in the areas are shown.**

The Tabasco plains are mainly composed of sediments transported, deposited and then dissected by running surface water forming a complex and dynamic network of river channels, some with meandering, braiding and deltaic river channel formations (Arreguín-Cortés et al., 2014). The lowland plains comprise deep clayey soils which are hydromorphic, some of which resulted from gleisation processes while others comprise azonal alluvial soils, mostly found on flood plains, levees and terraces (Krasilnikov et al., 2013). The Tabasco lowlands form Mexico's most extensive alluvial coastal plain stretching from the Gulf of Mexico to Sierra de Chiapas and Guatemala mountainous region (INEGI, 2018). In the upper plateau areas in the south, there are local appearances of prairie soils, mostly laterite, humid forest soils with associated coniferous vegetation. However, the prevailing soil texture in the basin is clay, covering about 77%, followed by 14% sandy soil (Arreguín-Cortés et al., 2014).

The Grijalva basin experiences a hot, humid maritime tropical climate in the north, with temperatures not falling below 18°C in the coldest month. Over 95% of the Tabasco State experiences a hot and wet climate and hot and semi-humid climate in other regions of the far-north-eastern states (INAFED, 2010). Average annual temperatures are around 27° C, with high temperatures of 36 ° C in May and lows of 18.5 ° C in January. In the Chiapas highlands, temperatures are cooler with average temperatures of 20 °C but can reach as high as 40 °C and as low as 0 °C (Arreguín-Cortés et al., 2014).

There are three primary sources of rainfall in the catchment area: Tropical cyclones from the Caribbean Sea and the Atlantic Ocean, the Inter-Tropical Convergence Zone (ITCZ) which extends to higher latitudes in summer, affecting the upper Grijalva basin, and late summer Tropical waves which cause substantial rainfall in the northern part of the catchment area (Arreguín-Cortés et al., 2014). The rise of warm and humid air masses along the ITCZ generates intense rainfall, and its north and south movements control the seasonality of rainfall with a maximum between June and October and relative minimum rainfall during July and August. Rainfall occurs nearly all year-round, with a relatively dry season in February and May (Arreguín-Cortés et al., 2014). Total annual rainfall in the Grijalva catchment area can reach approximately 4 000 mm with an average of 2 550 mm, but from June to October it is

slightly wetter compared to the other months (Gama et al., 2010). The average monthly rainfall-runoff regime and seasons in the Grijalva River basin are illustrated in Figure 4.3 (see Appendix A, Figure A1).



**Figure 4.3: The mean monthly rainfall-runoff regime and seasons in the Grijalva River basin (Adapted from Arreguín-Cortés et al., 2014).**

The seasonal tropical rainfall regime significantly decreases in the upstream regions of Grijalva catchment, within the Guatemala headwater sub-catchments, however, in central parts a relatively intense rainfall regime continues from September through October to November (Arreguín-Cortés et al., 2014). Higher rainfall is experienced in Chiapas from June to September while in Tabasco, occurs from June to October (Gama et al., 2010). However, Tabasco State is the wettest state in Mexico with an annual rainfall of over 2 406 mm per year, followed by Chiapas with 1 969 mm per year (INEGI, 2018).

The Grijalva River originates in Guatemala in the Sierra de Cuchumatanes and enters Mexico, crossing the depression of Chiapas, where its flows are first regulated by four dams in the central depressions of Chiapas (Arreguín-Cortés et al., 2014). From the

highlands, the river enters the broad coastal plains of Tabasco, before discharging into the Gulf of Mexico. This last stretch is within an alluvial fan (delta) in which several branches have developed through time due to changes of the river course. The Grijalva and Mezcalapa Rivers form a deltaic mouth, comprising distributaries before coming into the sea. Also, the hydrology of the lowland areas is characterised by lakes, salt marshes and lagoons of shallow depths, which are interconnected by a considerable number of channels before draining into the Gulf of Mexico (Arreguín-Cortés et al., 2014).

The Tabasco State is frequently subjected to periodic and catastrophic flooding such as in 2007, 2008 and 2009 (Arreguín-Cortés et al., 2014). The high incidence of flooding in the State has been exacerbated by land subsidence, deforestation, oil and gas drilling and poor methods of farming which have led to excessive silting of river channels. Frequent flooding is a severe problem in downstream areas of the Grijalva River, particularly in Gaviotas flood plains, the eastern fringes of Villahermosa city. The city is located at the confluence of the Grijalva and Mezcalapa Rivers occupying an area of 61 km<sup>2</sup> with 353,577 inhabitants. According to the 2010 Mexican Census, Tabasco State with 17 municipalities has a population of 640,359 inhabitants and all municipalities taken together have 72 towns with over 1 000 inhabitants and 116 towns with over 500 inhabitants, all with varying levels of vulnerability to flooding.

Flood risk management structures comprise four dams in the upper and middle sections of the Grijalva River (Figure 4.2). The dams are Angostura, Chicoasen, Malpaso and Penitas dams, built for the production of hydropower and regulating river flows. The dams have a significant impact on the Grijalva River flows, where natural flows have been reduced. Despite the dams, Tabasco flood plains continue to be subjected to frequent flooding from several large rivers and complex networks of tributaries and wetlands fed with increasing high rainfall totals (Aparicio et al., 2009). During these extreme flood events, floodwaters interact with pollutants from offshore oil extraction and agricultural areas presenting a significant risk to people's lives, infrastructure and sensitive ecological habitats.

Worryingly, there are densely populated areas in most flood-prone areas, especially in Villahermosa, which often experience flooding with severe impacts on life and property. The city of Villahermosa occupies much of the floodplains of the La Sierra, Carrizal and Viejo Mezcalapa river network, which were dissected into broad flat plains with an average height of 10 m above sea level. In 2007, unprecedented rainfall fell in the Grijalva Usumacinta Basin, and produced large amounts of runoff and flooding around 70% of Tabasco plains with depths of 4m in some areas (Aparicio et al., 2009). About 20,000 people had to seek emergency shelter in Villahermosa, and according to the Mexican Ministry of Interior, in 2007 approximately 1.2 million people were affected.

Throughout decades, Villahermosa has been subjected to frequent flooding, primarily due to human activities that compromise the integrity of the surrounding environment. Activities that alter the water environment in the surrounding catchment areas include the draining and diversion of river flows, siltation, deforestation and mixed farming. Villahermosa remains vulnerable to severe flooding, despite substantial waterworks and flood protection measures implemented since the 1950s. The measures comprise four large dams, floodwalls, river diversion at Samara River and the El Macayo barrier along the Carrizal River (Audefroy, 2017).

Therefore, the State of Tabasco requires accurate flood modelling and provision of reliable flood forecasts to warn its inhabitants and assessing the risk of flooding. The availability of satellite data has the potential to enable flood modelling in the State and provide a flood forecasting service to manage flood risks. The outcomes of this project will contribute to the state's efforts in improving flood risk management in the area. It is expected that the research will deliver flood risk map outputs showing areas prone to flooding under different return periods.

#### **4.1.2 La Sierra sub-catchment**

La Sierra catchment is one of the sub-catchments of the Grijalva basin and was selected as a study area to investigate the application of the flood scaling theory and the propagation of rainfall and parameter uncertainties (Objective 1 and 2). The sub-catchment is located in the eastern part of the Grijalva catchment, covering an area of 6 743 km<sup>2</sup>. The uncertainty propagation and analysis study used a downstream section of the La Sierra River, approximately 8 km between Las Gaviotas Gauging Station (-92.89° N and 18.00°W) in the south and El Porvenir Gauging Station (-92.91° N and 17.96°W) in the north mostly comprising downstream flood plains areas of the La Sierra River.

La Sierra catchment is the wettest in Mexico with an annual rainfall total of approximately 4 000 mm with annual averages of 2 550 mm occurring all year (INEGI, 2018). As a result, settlements in La Sierra flood plains were heavily hit during the 2007, 2008 and 2009 flood events (Arreguín-Cortés et al., 2014). The flood-prone areas include permanent and informal settlements in the districts of Torno Largo, Coquitos, Valle Verde, Armenia, La Gaviotas Sur, Gaviotas Norte, La Manga I, II and III. There are other flood-prone areas on the left bank of the La Sierra River, comprising the districts of La Casa Blanca I and II and part of the downtown area of El Centro, which are usually affected by flooding (Audefroy, 2017).

#### ***La Sierra nested catchments***

To apply the flood scaling theory, six sub-catchments were identified as nested catchments that drain surface flows from smaller sub-catchments within the main La Sierra catchment area (Figure 4.21). The nested catchment areas were delineated from seven discharge gauging stations (Table 4:6) which completely link and direct the river network and recognise the distributional drainage areas in a nested chain of catchment areas with distributional junctions (Mercado et al., 2016). Each gauging station represented an outlet from each sub-catchment drainage area and taken as nested catchment pour points (Fig. 4.21).

The seven nested catchments (and their sizes) were identified and named after their gauging stations were: Gaviotas (6 743 km<sup>2</sup>), El Pueblo (4 748 km<sup>2</sup>), El Puente (1 787 km<sup>2</sup>), Teapa (238 km<sup>2</sup>), El Puyacatengo (128.67 km<sup>2</sup>), Oxolotan (1 416 km<sup>2</sup>), and Tapijulapa (698 km<sup>2</sup>), sub-catchments (Figure 4.4). However, the El Puente and Tapijulapa gauging stations and corresponding nested catchments were not included in the scaling analysis because of insufficient hydrometeorological records for the years under consideration.

The main Las Gaviotas nested catchment area includes parts of Tabasco and Chiapas States, and its main river system begins in the central highlands of Chiapas, covering a drainage area of 6 743 km<sup>2</sup> to its downstream outlet at the Gaviotas gauging station. The Oxolotan nested catchment was delineated from the Oxolotan gauge station in the southern upland areas, draining 1 416 km<sup>2</sup>. The river flows through the highest mountainous areas (2 216 metres above sea level) mainly covering mountainous areas of La Sierra (Table 4.1). Tapijulapa (698 km<sup>2</sup>), Teapa (238 km<sup>2</sup>) and El Puyacatengo (128.67 km<sup>2</sup>), are the smallest nested catchment areas and are the headwater catchments in the La Sierra catchment. The topography of these upper catchments comprises highly dissected mountainous areas with very steep slopes (4.0%) and narrow V-shaped river valleys, while downstream are areas are extensive and gentle (1.0%) lowland plains at 6 m above sea level. Appendix D, Figure D-1 describes other physiographic characteristics of each nested catchment area.

**Table 4:1: Characteristics of nested catchment areas in La Sierra main catchment.**

Catchment	Catchment area (km <sup>2</sup> )	Length of the main channel (km)	Minimum elevation of the main channel (m)	Maximum elevation of the main channel (m)	Average slope of the main channel (%)	Standard average annual rainfall	Standard runoff (%)
Gaviotas	6 743	173 615	6	2 216	1.27	1 730	20 -30
El Pueblo	4 748	157 442	9	2 177	1.38	1 730	20 -30
El Puente	1 787	89 219	7	902	1.00	2 101	20 -30
Oxolotan	1 416	101 064	109	2 216	2.08	1 602	>30
Tapijulapa	698	131 192	19	2 216	1.67	1 602	>30
Teapa	238	42 027	35	1 106	2.55	2 513	20-30
Puyacatengo	129	15 370	59	644	3.81	2 461	>30

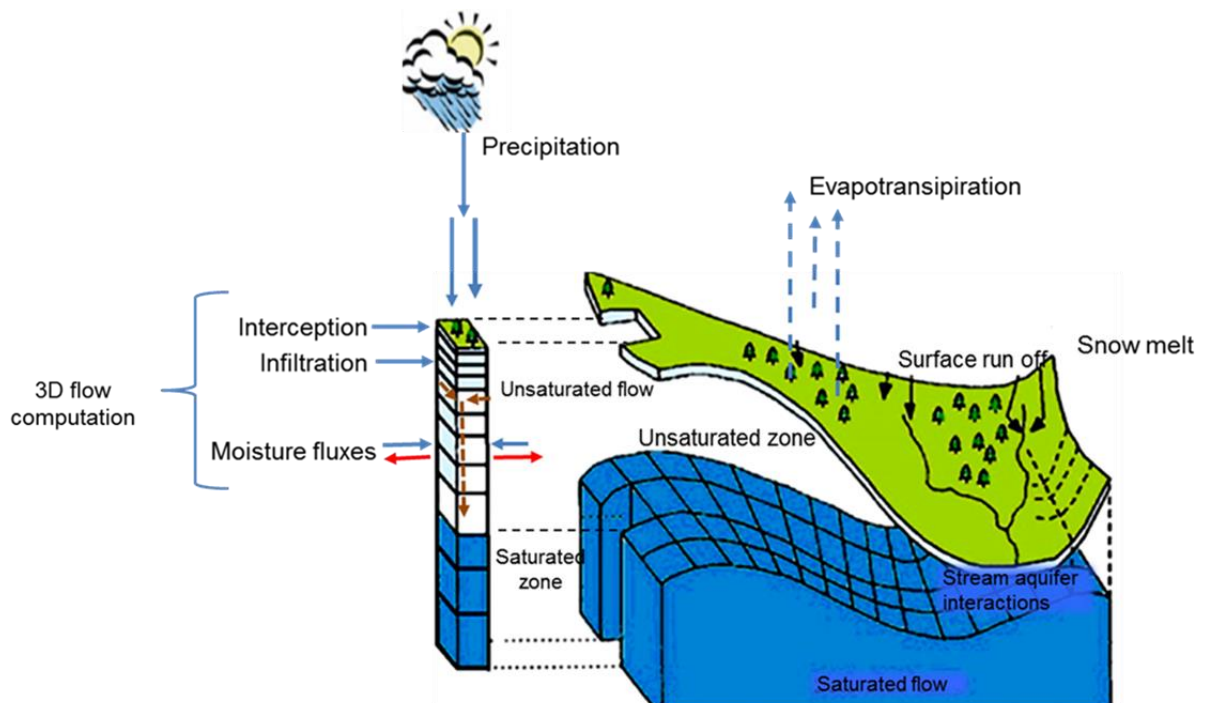
## **4.2 The SHETRAN hydrological model and data requirements**

The SHETRAN hydrological model used in this study is a physically-based fully distributed hydrological modelling system based on the Systeme Hydrologique European (SHE) model, developed by three European organisations, the British Hydrological Institute, the Danish Hydraulic Institute and the French firm SOGREAH (Abbott et al., 1986; Ewen et al., 2012). The UK Nirex Limited Company partially funded the development of the model to assess the health and safety of a proposed radioactive waste repository by studying the transport of radionuclides both on the surface and through shallow subsurface hydrological systems (Ewen et al., 2012).

The hydrological model was continuously improved and equipped with additional features, including the sediment component and a 3D subsurface water flow component (Wicks et al., 1996). The model was renamed SHETRAN by the School of Civil Engineering and Geo-sciences, Newcastle University, after integrating sediment and solvent transport modules. The Danish Hydraulic Institute (DHI) continuously invested resources into research and redeveloped and renamed the SHE model to MIKE-SHE (Abbott et al., 1986). Since then, further improvements have been made to the software. The most significant development was the implementation of a fully three-dimensional subsurface (VSS) component. These efforts led to the current form of SHETRAN model used to solve finite difference equations in describing the three-dimensional integrated surface/subsurface water flow and transport of sediments and pollutants (Ewen et al., 2012).

The SHETRAN hydrological model has a modular design with each element representing different hydrological cycle processes. The model is conceptualised as a sequence of columns and surface channel networks or river links running along column edges (Shrestha et al., 2017). Each column comprises a stack of computer cells with land cover type information at the top and sequential depths of soil horizons. The variation of catchment properties and processes and input data is interpreted horizontally in parallel to grid square networks by discretising the drainage system (Ewen et al., 2012). In this, the spatial distribution, variability of parameters and processes are represented in each grid column (Figure 4.4)





**Figure 4.4: The components of the SHETRAN hydrological model (Source: Ewen 1995).**

The Saint-Venant and diffusion approximation equations represent overland and channel flow in one and two dimensions, respectively. Potential evapotranspiration (PET) is calculated externally, with the Penman-Monteith equation (Equation 4.1) and added to the SHETRAN hydrology model. The model was designed for multi-resolution hydrological modelling and is suitable for simulations ranging from high to low-resolution modelling. It can provide flow simulations based on rainfall forecasts and data sets of potential evapotranspiration or remote sensing systems, including numerical model rainfall forecasts.

Given this context of development, it is not surprising that the strength of SHETRAN lies in its detailed grid-to-grid modelling of subsurface flow as a variably saturated porous subsurface directly coupled with the surface flow (Lewis, 2016). The model offers a broad range of applications such as assessments of groundwater resources, pollutant transport and flooding from multiple sources. The SHETRAN hydrological model can simulate the temporal and spatial variability of water flows and storage at standard grid cell resolution. This study focuses on the overland flow component for river flow modelling in a catchment containing highly variable catchment conditions

and processes and a channel system that feeds surfaces and subsurface processes with variable responses to rainfall. The utility of the SHETRAN model is demonstrated in this study by simulating river flows in a large tropical basin. However, the subsurface, sediment and solute transport capabilities are not fully employed, but the model developed can be further developed and improved to a model with such capabilities.

#### **4.2.1 Criteria for model selection**

The primary criterion for selecting the SHETRAN hydrological model was focused on understanding and defining the purpose of the model in this research (Wittwer, 2013). The selection considered data availability, catchment area size, topographical conditions and rainfall (Jakeman et al., 2006). It included how the catchment conditions and processes need to be represented in the model. The criteria required maintaining a balance between the purpose of the model, availability of model software, desired outputs, support and constraints due to computational costs and data availability. The following were taken into consideration (Cunderlik, 2003).

##### *Type of model*

The recommendations of the World Meteorological Organisation for selecting hydrological models were followed (Table 4.2) (Wittwer, 2013). The guidelines set out some basic questions for aiding model selection based on three primary model categories (fully distributed, semi-distributed and lumped models) (Scanlon et al., 2002). The selection is based on catchment size, the dominant rainfall that causes flooding, topographic characteristics and availability of gridded datasets (Wittwer, 2013). However, in selecting a suitable model for the large-scale hydrological modelling in the Grijalva catchment more attention was given to catchment conditions and processes, including the availability of datasets. The SHETRAN hydrological model was chosen to represent the highly variable catchment conditions and processes in the catchment area.

**Table 4:2 The standard questions and guidelines for model selection proposed by the WMO (Wittwer, 2013).**

Question 1	Catchment size?		
	Small (headwater)	Medium	Large
Routing	Mostly not needed	Hydraulic/hydrology	Hydraulic/hydrology Level to level correlations
Type of model	Lumped	Semi distributed	Distributed
Question 2	Catchment topography?		
	Flat/plain	Moderately/hilly	Pronounced/Mountainous
Type of model	Lumped	Semi distributed	Distributed
Question 3	What is the predominant flood causing rainfall?		
	Seasonal	Frontal/adveective	Convective
Recommended data resolution	Daily	Daily/hourly	Hourly/sub-hourly
Type of model	Lumped	Semi distributed	Distributed
Question 4	Are gridded data available?		
Constraints	No	Intermediate	Yes
Type of model	Lumped	Semi distributed	Distributed

### *Model availability*

Another criterion considered in the selection of the SHETRAN hydrological model was its accessibility as an open-source or upon request or by agreement. The model software is available for use both along with an active developer community with the possibilities of using the model for several purposes (Scanlon et al., 2002; Kauffeldt et al., 2016). The SHETRAN and LISFLOOD hydrological models were two options that were available for this research. Considering the criteria above, the LISFLOOD software was not open-source and not freely available, and this made it less suitable for this project. Therefore, the SHETRAN model was a suitable option and selected for hydrological modelling in the Grijalva catchment. Though the model had its advantages and disadvantages, it was envisaged that the chosen model would be subject to further analysis and minimisation of errors and uncertainties.

### *Existing support and user community*

The SHETRAN model was also chosen based on the presence of active users and as a model that is actively developed with its core developers defined to ensure sufficient support and for this research to recommend improvements. A requisite was a “user group” where users can ask, share problems, where they can get solutions quickly (Scanlon et al., 2002). The group was supposed to be large enough to provide adequate help without having to wait long times. Availability of technical support for model set up in the form of user guides and documents to guide potential users was also a criterion. (Kauffeldt et al., 2016).

### *Input data requirements*

Model selection criteria also focused on the availability of input datasets, where the availability of gridded data was limited, the lumped model would have been the only choice (Scanlon et al., 2002). Whereas distributed models require large datasets that must be available and extracted from existing databases and the datasets should be available both in sufficient quantities in terms of spatial and temporal resolutions and in-terms of variables required for the flood modelling undertaken. All input data required for modelling were supposed to be available within time and cost constraints of the project (Kauffeldt et al., 2016).

### *The flexibility of model grid structure*

The selection was also based on model scalability as per this study's requirement, which is a thorough representation of parameter variability and the model's ability to respond to short term variations in rainfall in different parts of the study area. The SHETRAN model uses a gridded structure where each grid/cell represents a computational unit. The flexible gridded structure uses the same spatial discretisation to describe the spatial variability of forcing data, geology, topography, soils and land cover to produce discharge simulations.

### *Flexibility in temporal resolution*

The model should provide the possibility of being run at a flexible temporal resolution that can be changed to represent the typically localised and variable rainfall experienced in the study area. If the flooding results from seasonal rainfall, then the daily data would be available, if it is caused by convective rainfall, then the sub-daily rainfall data would be required.

### *Size of the model domain*

This study was developed at a regional level, where hydrological components have interactive and complex relationships. A process-based and fully distributed model was suitable for detailed definition of highly variable catchment conditions and processes in the study area. The criterion relied on the regular squared grids that are scalable and can capture individual variations/heterogeneity of catchment properties and processes in the study area.

## **4.2.2 Data requirements and processing**

The SHETRAN hydrological model uses raster data and is closely coupled with GIS systems for processing its parameter maps (Table 4.3). GIS tools were used for data processing, which included spatial information analysis, map editing and presentation of findings, watershed description and interpretation using powerful distribution and location functions. The GIS hydrological tools were used to process the Digital Elevation Model (DEM) by extracting information on watershed boundaries, slope, flow path, Flow Accumulation and network configuration, providing a suitable framework for the SHETRAN hydrological modelling work. Raster maps for soil, vegetation, land use and water bodies were developed within the GIS environment. These included data layers showing the river channel network and rainfall gauge Thiessen polygons. All parameter layers developed were then converted into an ASCII file format compatible with SHETRAN data file requirements..

**Table 4:3 The primary data sources and types required for the SHETRAN hydrological modelling.**

<b>Data category</b>	<b>Data type</b>	<b>Dataset</b>	<b>Derivatives</b>	<b>Source</b>
<b>Satellite data</b>	Topographic	SRTM 90m Digital Elevation Model (DEM)	Flow Direction Grid (FDR) Flow Accumulation (FAC)	Shuttle Elevation Derivatives at multiple Scales (HydroSHEDS)
	Rainfall	TRMM_3B42; TRMM (TMPA) Rainfall Estimate L3 3-hour 0.250 x 0.250 V7	2013-2015 rainfall time series 2016-2017 rainfall time series	Geospatial Interactive Online Visualisation and Analysis database (Giovanni) data portal
	Evapotranspiration	Potential Evapotranspiration (PET)	2013-2018 Evapotranspiration time series	Famine Early Warning Systems Network (FEWS NET) global data portal
	Discharge	River discharge	1998-2020 Daily river discharge	Dartmouth Flood Observatory
	Soil	Parameter dataset	2 x 2 km gridded soil layer	Food and Agriculture Organisation (FAO) Harmonised World Soil Database (HWSD) database
	Land use	Parameter dataset	2 x 2 km gridded land-use layer	National Institute of Statistic and Geography (INEGI)
<b>In-situ data</b>	Rainfall	Observed daily rainfall datasets	Observed daily 2013-2015 and 2016-2017 rainfall time series dataset	Mexican National Climatological Database, CLimate COMputing (CLICOM)
	Discharge	Observed river discharge	Observed daily 2013-2015 and 2016-2017 discharge time series	Mexico surface Water Management and Rivers Engineering (GASIR) National Surface Water DataBank databases (BANDAS),
	Soil	Parameter dataset	Soil types classification grid	National Institute of Statistic and Geography (INEGI)
	Land use	Parameter dataset	Master map and land use grid	
	Vegetation	Parameter dataset	Vegetation grid	

### **Data quality checks**

All data sets were required to be of acceptable quality and accuracy before hydrological modelling. In particular, rainfall, potential evaporation and discharge data were checked for quality, continuity, reliability and pattern consistency. Incorrect data, anomalies and availability of data for calibration and validation periods were also checked. Then checks were done for time lag, time series rainfall and discharge patterns, including whether the data meet the predicted catchment trends.

The first thing was to convert data into standard units and into the same units and by plotting, for example, rainfall (mm/hr), potential evapotranspiration (mm/hr) and discharge data ( $m^3/s$ ) in adjacent excel columns. In the first column the common time format or function (*dd/mm/yyyy/hh:mm*) was used. To enable checking all datasets were converted into the same units, for instance, where discharge values originally in  $m^3/s$  were converted to mm/hr, using the *value \*3 600\*1 000/area in  $m^2$*  converter. Subsequently, all the time series data (now with same units), were plotted together to show up any spurious or missing values. Comparisons were made to track significant rainfall-runoff events in terms of time lag, recession limbs and peaks (Birkinshaw, 2012). The data were also checked for continuity and reliability using the Pivot table functions in excel. The functions summarised time series data for long-term bias estimates by plotting monthly values using the 'Month and Year' scales (Appendix A-1).

Data quality checks were carried out on time-series data with missing values; thus, incomplete series data were reviewed and validated with identical observed data. Prolonged dry periods of more than a month were established and excluded if comparable dry periods were not identified at nearby stations, including removal of duplicate rainfall totals in consecutive sub-daily scales. However, some time series data required filtering to eliminate systematic errors due to measurement instruments. Generic procedures were used to detect anomalies and outliers, some of which involved removing data from the upper and lower ranges identified for the observed average (Wilby et al., 2017). Recognising sources of data spikiness embedded in some measuring devices were considered, but spikes were not uncritically removed; cross-checking was performed against nearby gauge datasets for any unexpectedly high data values.

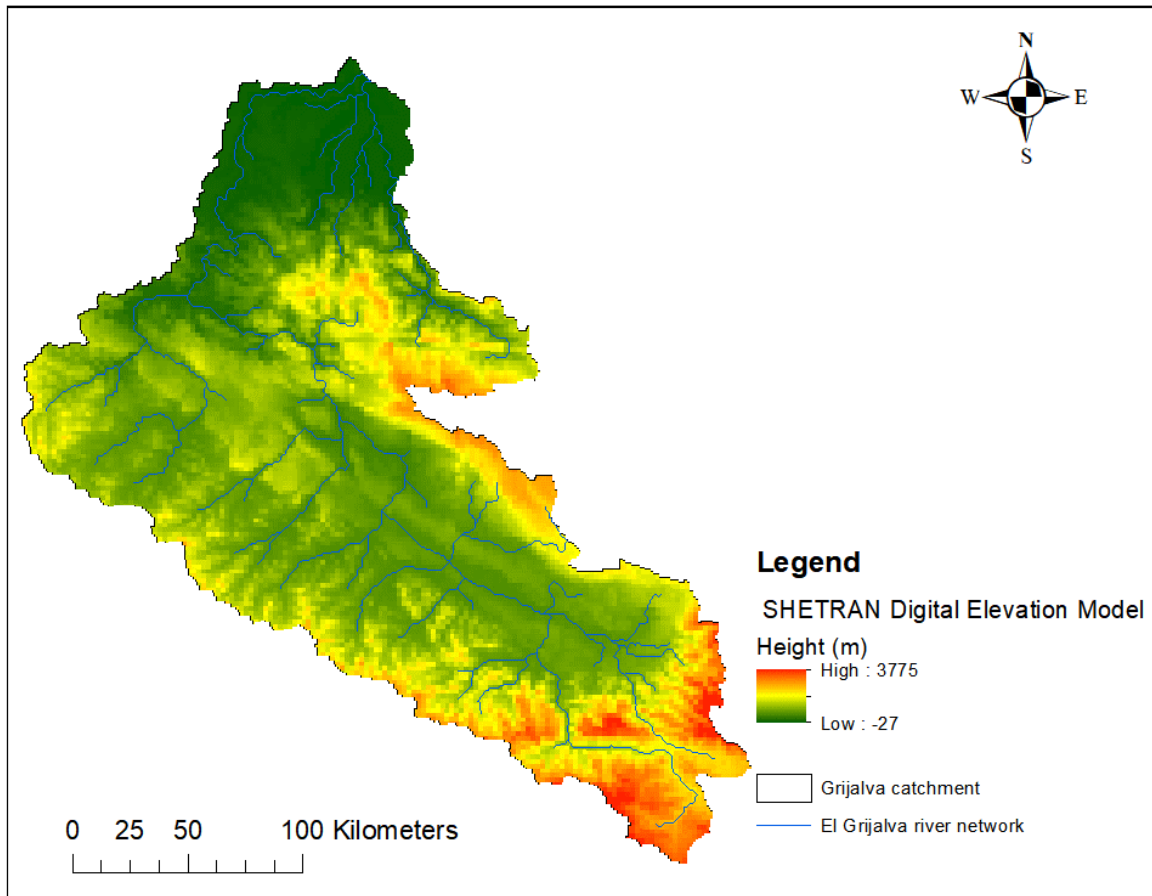
### ***Topographic data processing***

The Digital Elevation Model (DEM) was used as the principal data source representing watershed properties in the SHETRAN hydrological model. The elevation data were converted from latitude and longitude coordinates to the Mexican national grid projected coordinate system. Initially available at 3-arc resolution, the dataset was re-sampled to 2 x 2 km grid cell resolution (Figure 4.5) to fit the selected SHETRAN model grid size. The re-sampling of the DEM resulted in voids, sinks or pits that required re-processing of the dataset using an automated void filling procedure outlined in Jenson and Domingue (1988), and Clark et al. (2017).

The hydrologic terrain analysis tools in ArcGIS were used to create Flow Accumulation grids (FAC) that defined the number of upstream areas draining into each grid cell. The Flow Direction Grid (FDR) was also developed to determine the direction of flow from each grid to the steepest downstream neighbouring grid. A Grid Mask was produced for delineating the model domain. The location of the channel flow routes in SHETRAN was calculated by computing the number of upstream squares flowing in the downstream direction into each square. A river channel link was created when they reach a certain number. Channel heights were based on two adjacent grid squares, and sometimes heights were adjusted using the ArcGIS Fill tool so that the downstream flow path would always be present (Birkinshaw et al., 2010).

However, where significant tributaries and other sections of the existing river network were not showing, they were “burned” into the DEM to reduce elevation values of pixels in locations where the river channel is located (Clark et al., 2017). Lastly, all raster DEM layers developed were processed to make them compatible with SHETRAN regarding grid size and format with other raster layers, for example, soil type, vegetation type, water bodies and land use (Figure 4.5).

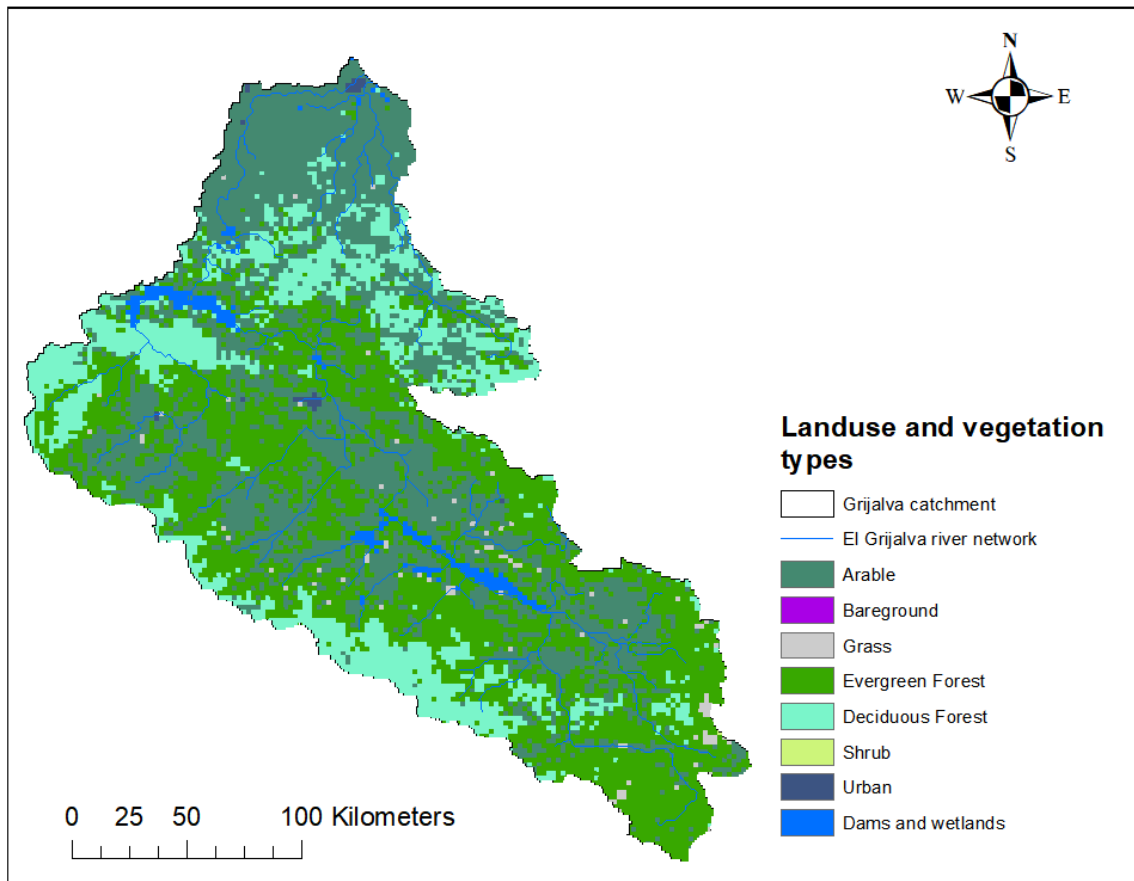




**Figure 4.5: Digital Elevation Model (DEM) for Grijalva catchment at 2 x 2 km grid resolution.**

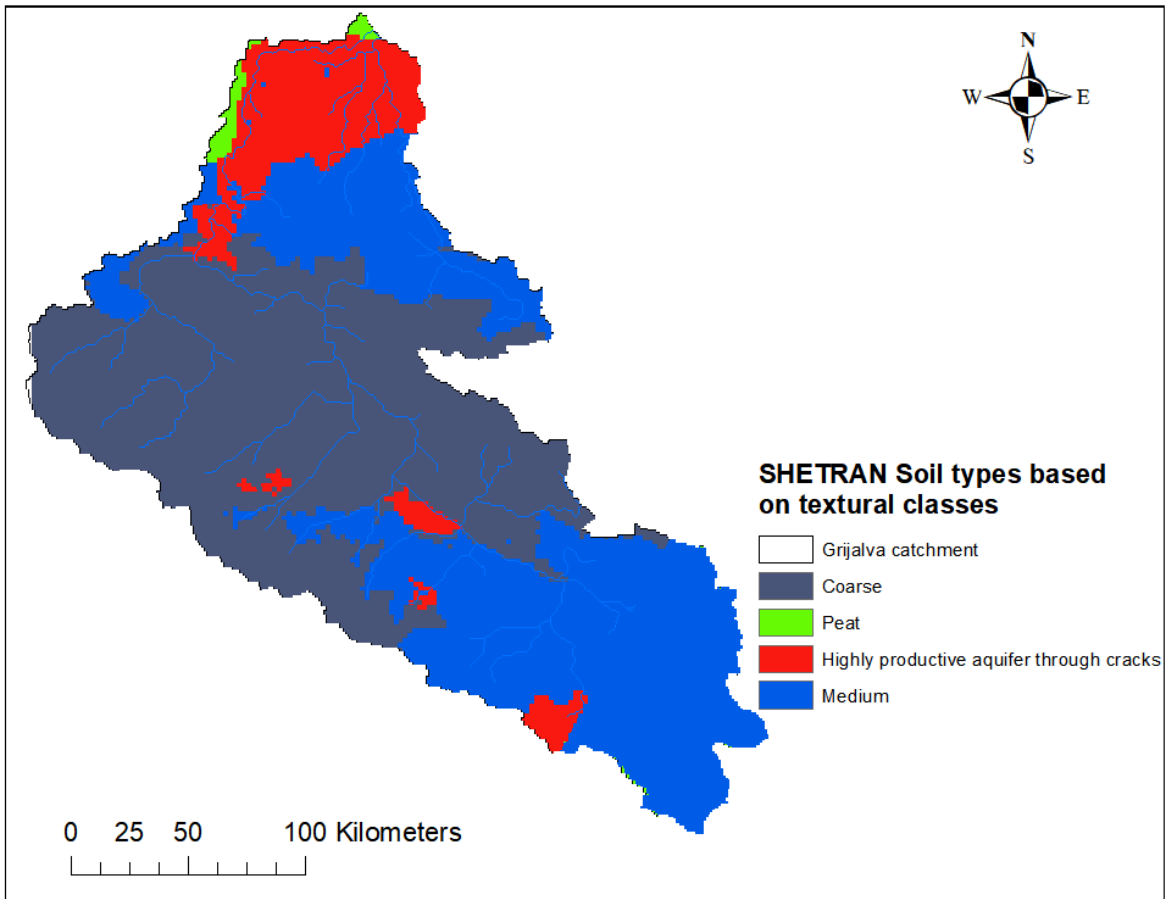
### ***Soil and land cover***

The SHETRAN hydrological model also required raster grids for soil and land cover types to specify the land surface properties, vegetation and land use that partition rainfall into infiltration and runoff. The soil and land cover data for each SHETRAN grid square were obtained from the 1km grid resolution the Harmonised World Soil Database (HWSD) database (FAO, 2018). Land cover types in the catchment were divided into seven groups, namely: arable, bare land grass, deciduous forest, evergreen forest, shrub and urban (Figure 4.6). This study used spatial patterns of satellite remote sensing data for parameter representation of spatial variations, which is a valid source of information for distributed hydrological modelling undertaken (Dembélé et al., 2020). Details of each land cover type, including parameter ranges were derived from the literature.



**Figure 4.6: Land cover types at 2 x 2 km grid map for the SHETRAN parameters.**

Soil maps were obtained from FAO, (2018) and INEGI, (2011, 2012) data portals. The topsoil and subsoil textures were used to assign the SHETRAN soil parameter values for saturated water content, soil conductivity, residual moisture content, saturated hydraulic conductivity and the van Genuchten parameters were specified from literature (Sreedevi and Eldho, 2019; Birkinshaw et al., 2017; Bathurst et al., 2004). The soil texture parameters were classified according to the British soil classification in which the distribution of the size of particles within the soil defines a soil type. Primary soil types considered were medium clay, coarse and highly productive aquifer through cracks. The spatial data were aggregated to 2 km SHETRAN grid squares with the soil profile chosen to be the most dominant in that square (Birkinshaw et al., 2017). Later, all gridded parameter layers were converted to ASCII file format and integrated as model files in SHETRAN to represent the spatially distributed and grid parameter values (Figure 4.7).



**Figure 4.7: The soil grid map and categories used for setting up the SHETRAN hydrological model.**

***Potential evapotranspiration***

The Potential evapotranspiration (PET) dataset represented the combined transpiration and evaporation from surfaces that are covered by a grass crop and whose water supply is unlimited. The Potential Evapotranspiration (PET) dataset was obtained from the global daily potential evapotranspiration database provided by the Famine Early Warning System Network, (FEWS NET) global data portal (FEWS NET 2018). The PET data set was calculated using net meteorological variables of radiation, temperature, humidity and wind speed using the Penman-Monteith equation as:

$$\lambda E = \frac{\Delta(H_{\text{net}} - G) + p_{\text{air}} \cdot C_p \cdot \frac{e_s - e_a}{r_a}}{\Delta + \gamma(1 + \frac{r_c}{r_a})}$$

**Equation 4.1**

Where  $\lambda E$  is latent heat flux density (MJ/m<sup>2</sup>/h),  $E$  is depth rate of evaporation (mm/h),  $\Delta$  is the slope of the saturation vapour pressure-temperature curve (kPa/°C),  $H_{\text{net}}$  net is radiation (MJ/m<sup>2</sup>/h), and  $G$  is soil heat flux (MJ/m<sup>2</sup>/h)

### **4.2.3 Rainfall data processing**

Rainfall was considered to be a critical hydrological variable and primary input data for the spatially distributed hydrological modelling undertaken (Chen and Wang, 2018). It was noted that accurate flood modelling and reliable simulations could not be achieved if the accuracy of input rainfall data were not high enough (Li et al., 2018). Also, no model, however well-founded in physical process theory or empirically supported by historical observations, can provide reliable flood simulations if enforced with inaccurate rainfall input data (Beven, 2008). In this regard, error quantification and correction of rainfall input data were considered necessary for the accuracy of the hydrological modelling undertaken and to ensure reliable simulations. Geostatistical methods (Kriging) were used for spatial interpolation of point rain gauge rainfall to continuous surfaces, merging interpolated rain gauge with satellite rainfall and topographical data and also for the quantification of errors used to improve the rainfall input data.

### ***Available rainfall datasets***

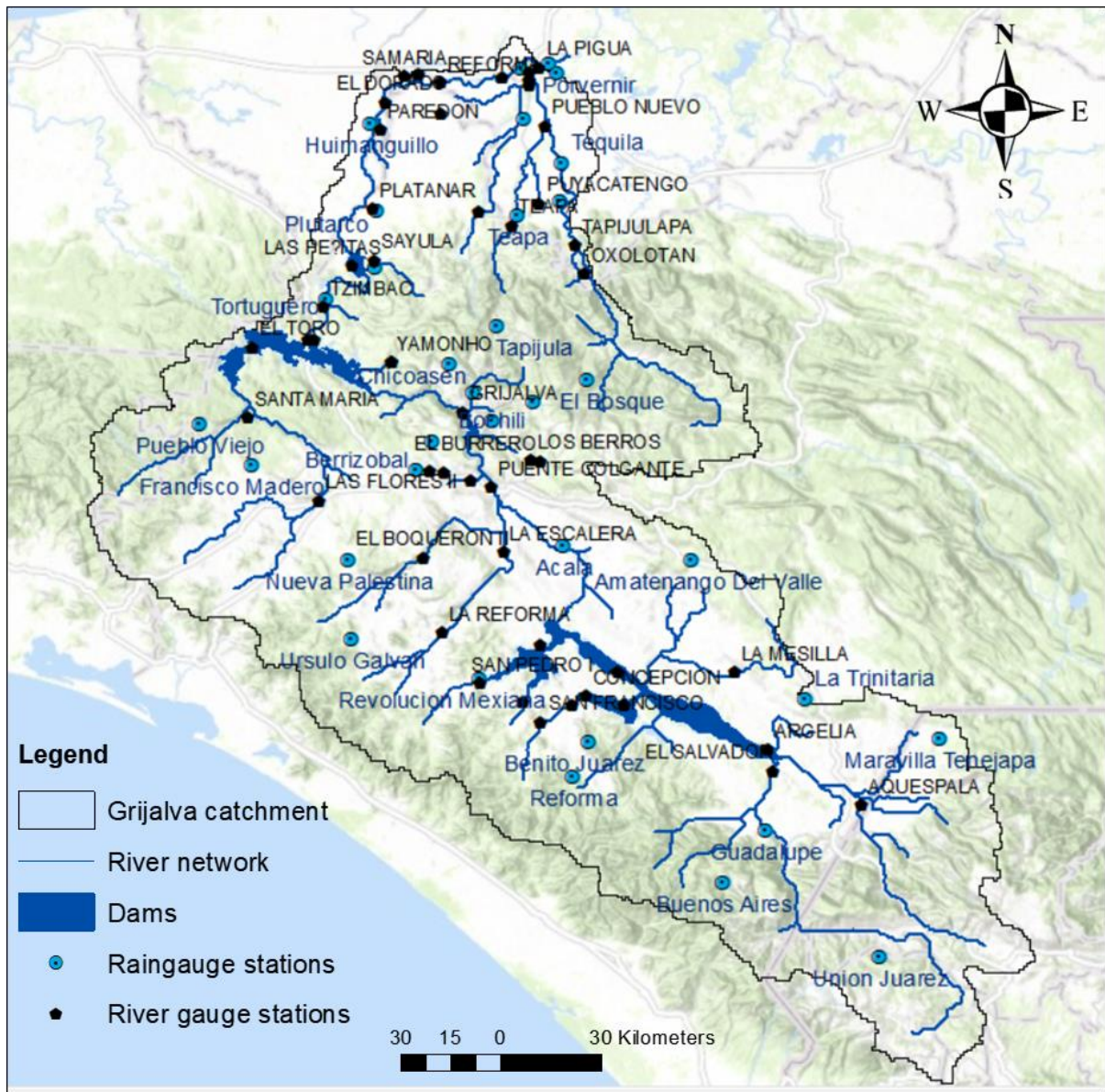
Rain gauge and satellite rainfall data were two key sources of rainfall data used in SHETRAN hydrological modelling and the datasets were merged, and errors quantified to improve rainfall input data for the model.

#### *Rain gauge rainfall data*

Daily rain gauge data from thirty-three rain gauge stations covering 57,958 km<sup>2</sup> of Grijalva catchment area were available for four years between January 2013 and December 2017. This translated to an average rain gauge network density of one rain gauge per 1 700 km<sup>2</sup> in the study area (Figure 4.8). The rainfall data were obtained from the Mexican rainfall database (CLIMCOM), maintained by the Mexican Meteorological Service (Servicio Meteorológico Nacional, SMN) and the Mexican Water Commission (Comisión Nacional de Agua).

Although the rainfall data were subjected to initial quality control before archiving it was not possible to obtain continuous temporally and spatial representative rainfall data for the study area. The spatial distribution of rainfall rain gauge stations in the Grijalva catchment was uneven with many gauges concentrated in highly populated areas, particularly in lowland areas of Tabasco and sparsely distributed in mountainous areas in Chiapas (Figure 4.8). In the southern areas of the study area, including in the southern headwater catchment areas in Guatemala, there were few stations such that each rain gauge represented an extensive area.

Nevertheless, the available rain gauge rainfall data provided direct point accurate rainfall estimates, though with limitations in describing its variability in the catchment (Pedersen et al., 2010; Sideris et al., 2014). Thus, the datasets were considered accurate point reference of rainfall in the study area collected directly at specific points and provided the best description of rainfall received but with limited spatial coverage. Thus, it was considered impossible to obtain accurate rain gauge rainfall estimates from rain gauge data alone, so satellite rainfall datasets were used to supplement rain gauge rainfall shortfalls in the study area.



**Figure 4.8: The distribution of dams, rain gauge and river gauging stations and associated river network in the Grijalva catchment.**

## Satellite rainfall datasets

Satellite rainfall data were available at 3-hour, daily (24 hours) and monthly temporal resolutions, however, the datasets were of the same spatial resolution, at  $0.25^\circ \times 0.25^\circ$  (approximately 25 x 25 km) (Figure 4.9) (Acker and Leptoukh, 2007; Huffman et al., 2010; Verdin et al., 2016). The latest Version 7 of the TRMM 3B42 of satellite rainfall products were freely available from the Giovanni online data portal, maintained by the NASA Goddard Earth Sciences Data and Information Service Centre (GES DISC) (<https://giovanni.gsfc.nasa.gov/giovanni/>). The satellite rainfall datasets provided the spatial description of rainfall information, particularly in areas where rain gauges were not available (Figure 4.9 and Appendix A, Figures A 4). The datasets were used to represent the overall spatial distribution of rainfall in the study area, although the data are less accurate in estimating rainfall intensity (Dinku et al., 2014).

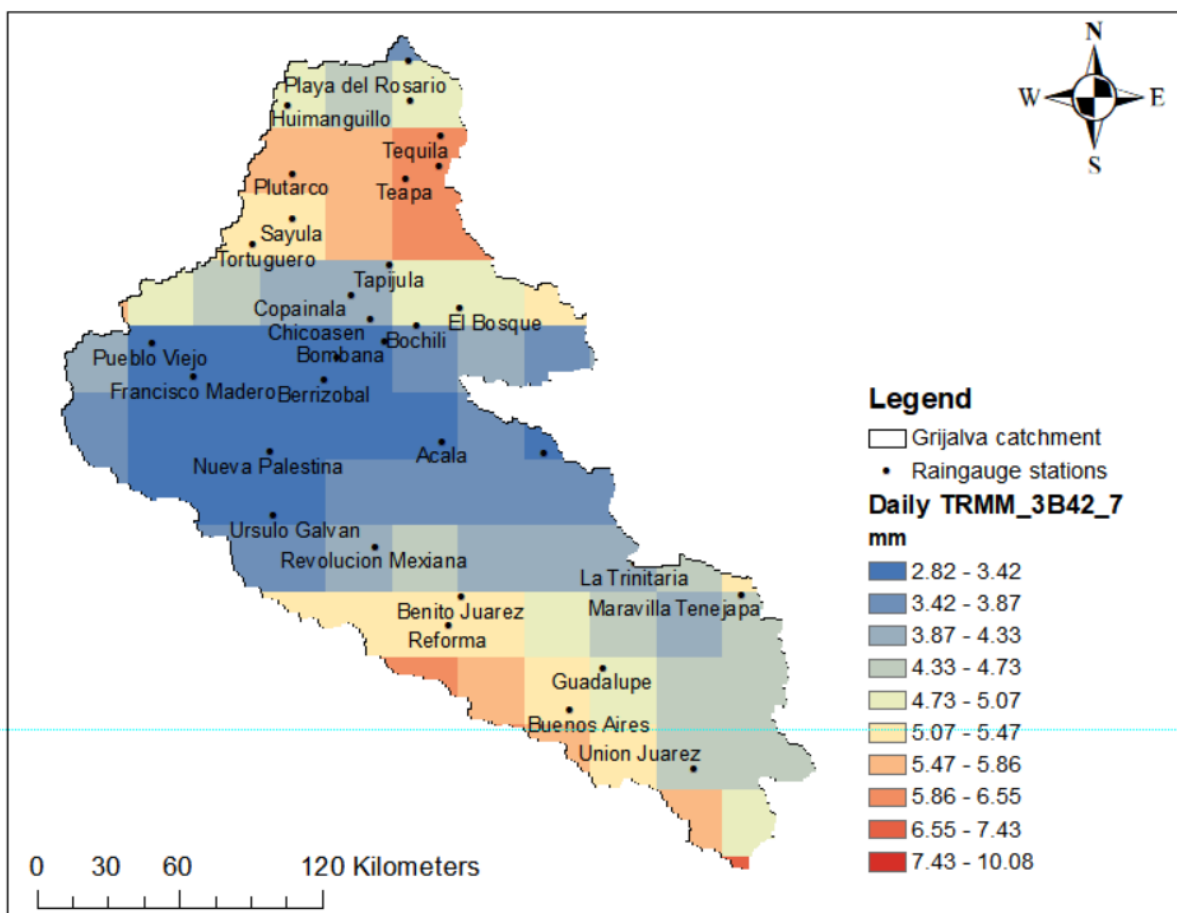


Figure 4.9: Time-averaged map of mean daily rainfall over 2013 to 2017 period at 0.25 degrees spatial scale (Approx. 25 x 25 km) [TRMM\_3B42\_Daily v7] in the Grijalva catchment area, Mexico.

### *Comparisons of rain gauge and satellite datasets*

The available rain gauge and satellite rainfall datasets were both considered to have limitations and merits. Each rainfall dataset had its advantages and disadvantages, and none of them seemed to be perfect.

Satellite datasets had the advantage of providing large and continuous spatial rainfall coverage, while rain gauge rainfall was considered accurate only at a few locations (Grimes et al., 1999; Dinku et al., 2014). Although satellite rainfall data provided opportunities to map rainfall at high temporal resolution, however, it had several inherent errors related to observation, instrument, and uncertain rainfall retrieval algorithms. For example, Chen et al., (2016) reported that the 3B42V7 satellite rainfall data underestimate low-intensity rainfall <24 mm and overestimate high-intensity rainfall (>100 mm) (see Appendix A, Figures A 2). Other critical shortcomings of satellite rainfall include the heterogeneous time series, a short period of observation and low accuracy, particularly at higher temporal and spatial resolutions (Grimes et al., 1999; Dinku et al., 2014). Also, satellite rainfall measurements were estimated from the physical characteristics of cloud top temperatures and hence provided an indirect measure of rainfall, based on temperature-rainfall correlations.

Although rain gauge rainfall data provided direct point accurate rainfall estimates, it was corrupted by random and systemic measurement errors. The systematic error was the most common, a likely result of losses due to the wind, wetting, evaporation and splashing. Furthermore, the rain gauge data had gaps in time series, incomplete values for significant events, others with spurious values, and pattern inconsistency. Also, the tipping bucket gauges used throughout the study area are known to underestimate rainfall for extreme events, compared to weighing bucket gauges.

Thus, to eliminate rainfall errors, it was necessary to merge both datasets to compensate deficiencies in each (Vila et al., 2009; Verdin et al., 2016; Chen and Wang, 2018). The merging of rain gauge and satellite rainfall datasets was envisaged to generate an improved rainfall product that has both high spatial and quantitative accuracy.



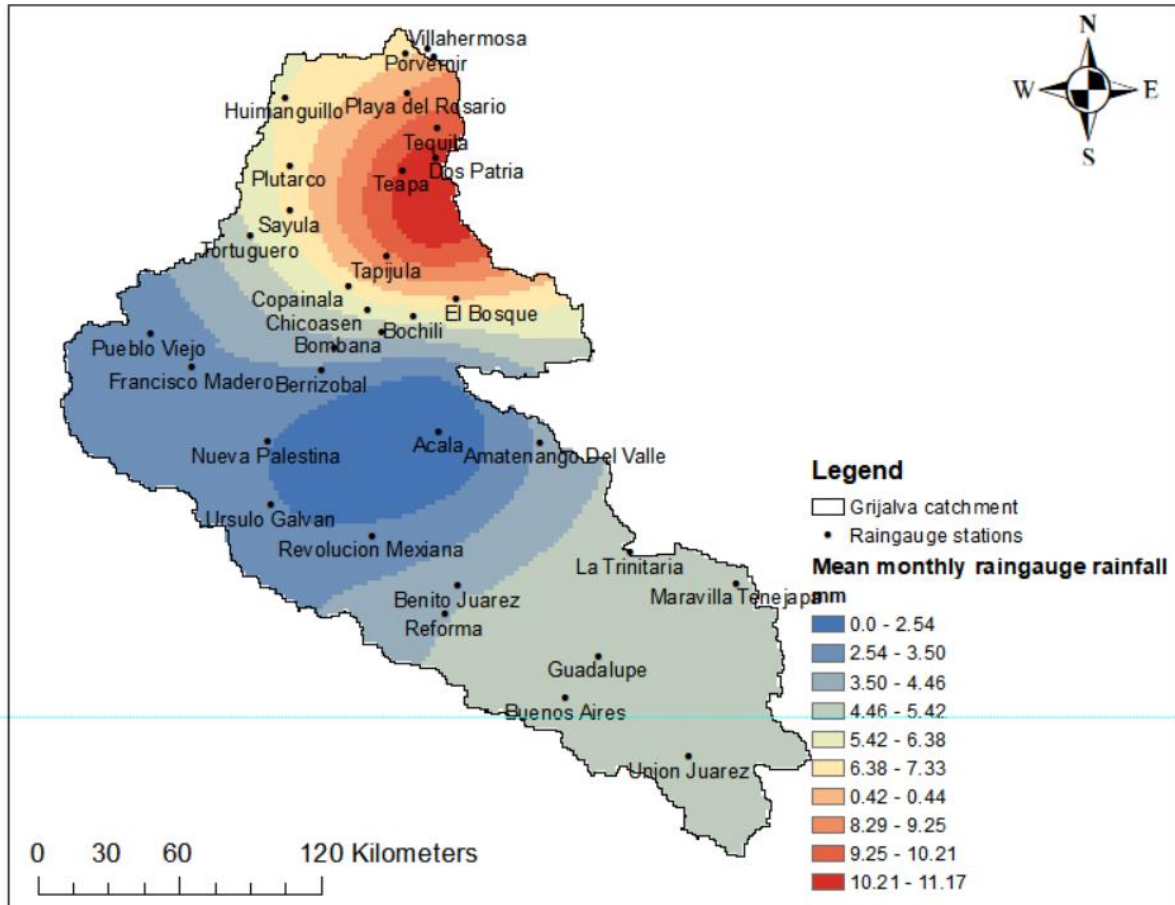
### ***Rainfall data scale issues and merging***

Rain gauge and satellite rainfall datasets were required to be merged to correct and improve rainfall information in the study area. The fact that each dataset provided rainfall estimates at different temporal and spatial scales complicated the integration of the two datasets (Goovaerts et al., 1997). Rain gauge rainfall data were available at daily resolutions while the satellite data were available at 3-hour, daily and monthly temporary scales.

Thus, daily rain gauge rainfall datasets were computed to the same monthly rainfall scale as satellite data to allow merging of the two (Huffman et al., 2010). The daily rain gauge and the 3-hour satellite rainfall resolutions were not used, as they were considered to be of high resolutions, which would provide a detailed variability of rainfall, not suited for a catchment with a sparse distribution of rain gauges. For long-term bias calculations, monthly rainfall resolution was used to better reflect the strong regional effects of topography and local circulation, which play a key role in the generation of rainfall in the area and so that the data can be easily smoothed over large scales (Vila et al., 2009).

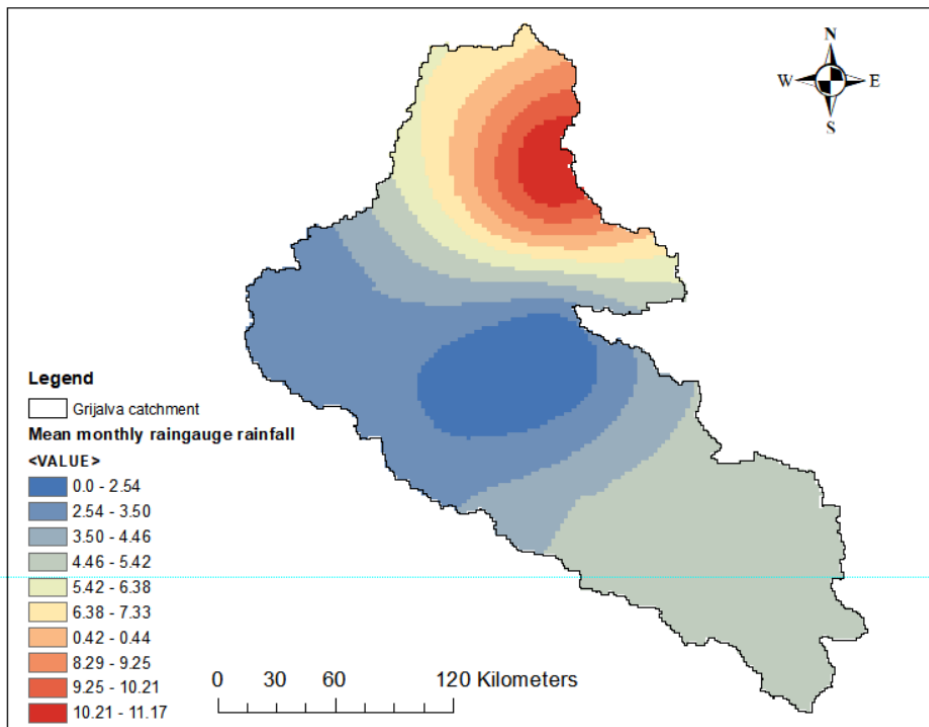
The daily rain gauge rainfall data were summarised into mean monthly values at each rain gauge station using Excel Pivot tables (Appendix A-1). The rainfall aggregation was carried out using the average summary functions of the Pivot tables to obtain average monthly rainfall for the study area. Subsequently, excel rainfall data tables were developed, comprising a list of all rainfall gauge names, corresponding coordinate locations, sub-catchment and average monthly rainfall values, and these were loaded into ArcMap and translated into shapefiles to store their geometric properties. A one-to-one relationship was maintained between individual rain gauge data attributes linked to record numbers, gauge name, monthly rainfall values and rain gauge positions as vector points.

Using the Kriging spatial model, point rain gauge rainfall datasets were interpolated and upscaled to 2 x 2 km<sup>2</sup> grid resolution before merging (Figure 4.10 and 4.11). The satellite rainfall datasets originally at 0.25 x 0.25 latitude/longitude grid-scale were downscaled to a fine spatial resolution of 2 x 2 km<sup>2</sup> grid (Figure 4.12) using the ArcGIS Resampling tool.

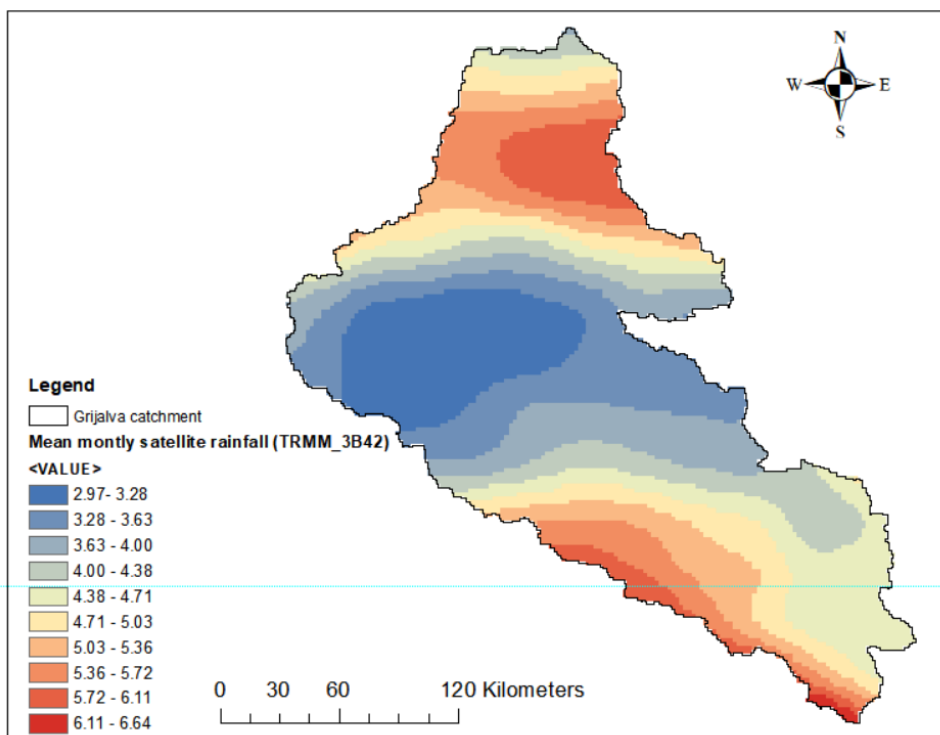


**Figure 4.10: The mean monthly rain gauge rainfall distribution interpolated at 2 x 2 grid-scale (over 2013 to 2017 period) before merging.**

The point rain gauge values were interpolated to areas with sparse rain gauge data at the same spatial resolution of 2 x 2 km<sup>2</sup> grid (Figure 4.11; Appendix A, Figures A 3) and subsequently combined with satellite data (Bringi et al., 2011). The Kriging spatial model was used to merge the point rain gauge and the satellite dataset to improve rainfall data for the study area (Grimes et al., 1999). The merging process of rain gauge and satellite rainfall datasets is summarised in Figure 4.13 and described in Section 4.2.4.



**Figure 4.11: The average monthly rain gauge rainfall distribution interpolated and upscaled at 2 x 2 grid-scale (over 2013 to 2017 period) before merging and compared to satellite rainfall below.**



**Figure 4.12: The mean monthly satellite rainfall grid map (3B42V7) downscaled at 2 x 2 km grid-scale (over 2013 to 2017 period) before merging and compared to rain gauge rainfall above. The rainfall product underestimated rainfall amounts in the area and overestimate the Pacific coastal mean rainfall.**

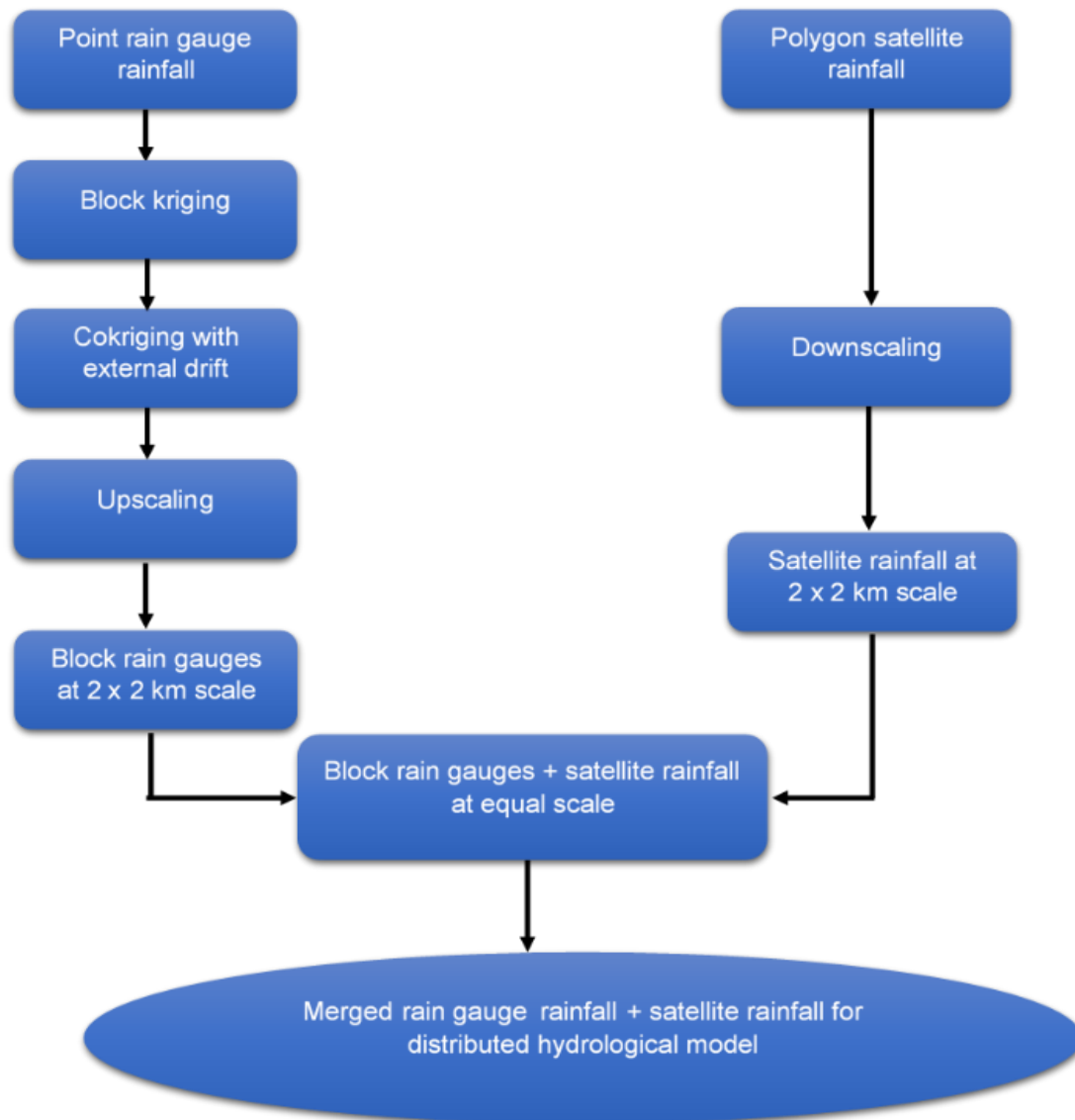


Figure 4.13: Merging process of rain gauge and satellite rainfall datasets.

#### 4.2.4 Geostatistical technique

Kriging, a geostatistical interpolation technique, was used as a weighted linear combination of available rainfall gauge data to estimate rainfall at ungauged locations and minimise expected squared errors (Verdin et al., 2016). The technique provided the best unbiased linear estimates of rain gauge point values to derive estimates of rainfall at ungauged locations based on a fitted theoretical variogram (Nerini et al., 2015). The estimates of rainfall values at ungauged locations were obtained as a weighted combination of all gauged rainfall values using this general formula:

$$r(s) = \sum_{i=1}^{N_s} \lambda_i r(s_i)$$

**Equation 4.2**

Where  $r(s_i)$  is the estimated value at the  $i$ th location,  $\lambda_i$  is an unknown weight for the measured value at the  $i$ th location;  $s_i$  is the prediction location, and  $N_s$  is equal to the number of measured values.

The block kriging method, an extension of Kriging technique, was used to upscale point rain gauge observations to blocks of 2 x 2 km<sup>2</sup> (Figure 4.10, 4.11 and 4.13) and geostatistical layers were converted into raster format using the GA Layer to Grid geoprocessing tool.

##### *Block kriging*

Block kriging works by predicting the average value of rainfall for several specified grids, and the value is averaged over the entire grid. The method uses a moving neighbourhood or block of given dimensions to estimate the average  $Z$  values over a surface (Teng et al., 2017). The average value of attribute  $Z$  over a block  $V$  centred at  $u$ , the block mean value  $Z_v(u)$  is defined as:

$$Z_v(u) = \frac{1}{N} \sum_{i=1}^N Z(u_i)$$

**Equation 4.3**

The block value  $Z_v(u)$  is a linear average of the  $N$  point estimators and has a minimum variance of estimation error.

The technique was used to estimate the average rainfall field from a point rain gauge location to up-scaled blocks of 2 x 2 km<sup>2</sup> covering the entire model domain (Figure 4.10 and 4.11). There are three Geostatistical model variants of the Kriging technique, the Ordinary Kriging (OK), the Ordinary Cokriging (OCK) and Kriging with an external drift (KED) (Verdin et al., 2016). The Ordinary Kriging OK technique was used with rain gauge information only, while the other two techniques incorporated secondary information to improve prediction in ungauged areas (Goovaerts 1997). In this study, the Cokriging with external drift (KED) technique was employed to incorporate satellite and topographic data to further improve the spatial interpolation of the rain gauge rainfall values in the study area (Grimes et al., 1999).

#### *Cokriging with external drift (KED) technique*

Cokriging with external drift (KED) is an extension of ordinary kriging that estimates rain gauge rainfall using secondary variables, satellite rainfall and topography datasets that were extensively available and known in the study area (Goovaerts 1997). The KED technique was used to integrate secondary variables (in this case satellite and topographic data) as they represented a trend surface to which the mean rainfall was linearly related to, expressed as::

$$m_1^*(\mathbf{u}) = a + bz_2(\mathbf{u})$$

**Equation 4.4**

Where  $m_1^*(\mathbf{u})$  is the mean rain gauge rainfall measured at location  $\mathbf{u}$ ;  $Z_2(\mathbf{u})$  is the value of the topographic variable at the same location;  $\mathbf{a}$  and  $\mathbf{b}$  are unknown coefficients.

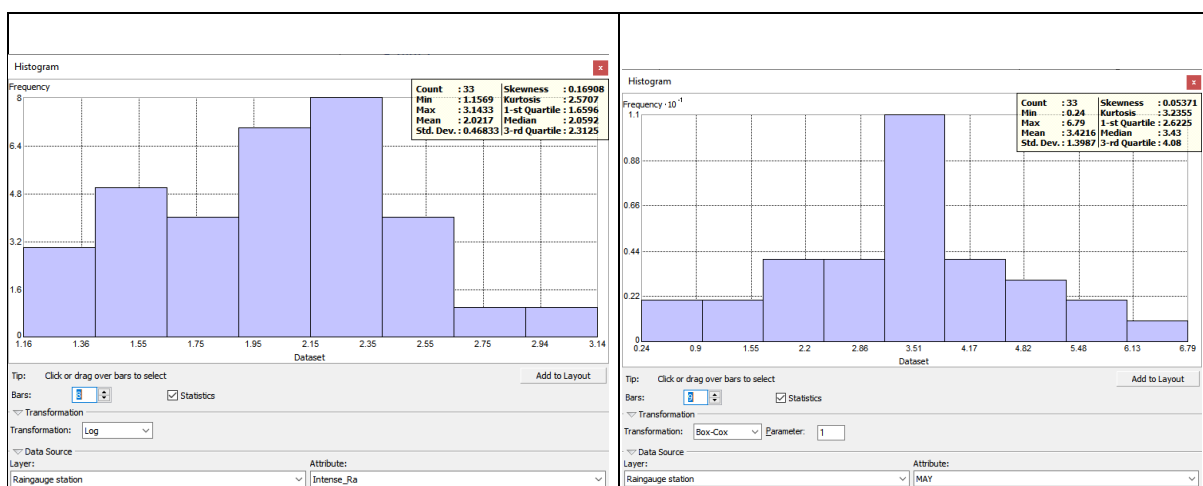
The basis for the Cokriging with external drift merging technique used in this study was the assumption that rainfall error estimation has an inherent spatial structure defined by location and elevation (Verdin et al., 2016). A Digital Elevation Model (DEM) was employed to incorporate topography as complementary information that can improve the interpolation of rainfall in the study area (Lloyd and Atkinson, 2004). The goal of using topographic information was to increase the accuracy of the representation of local rainfall variability in the Grijalva catchment area since elevation and rainfall co-vary locally. Thus, elevation data were included as a covariate due to its direct influence on rainfall occurrence in the area.

## 4.2.5 Implementation of Geostatistical rainfall analysis

The first step used in implementing the Geostatistical Rainfall Analysis was to explore rainfall data using the Exploratory Data Analysis tool to investigate the statistical and spatial characteristics of available rainfall data. The following three data features were verified: dependency, stationarity and distribution (Verdin et al., 2016). Three primary Geostatistical data tools were used to explore these features, and these were histogram exploration, semivariogram and trend analysis.

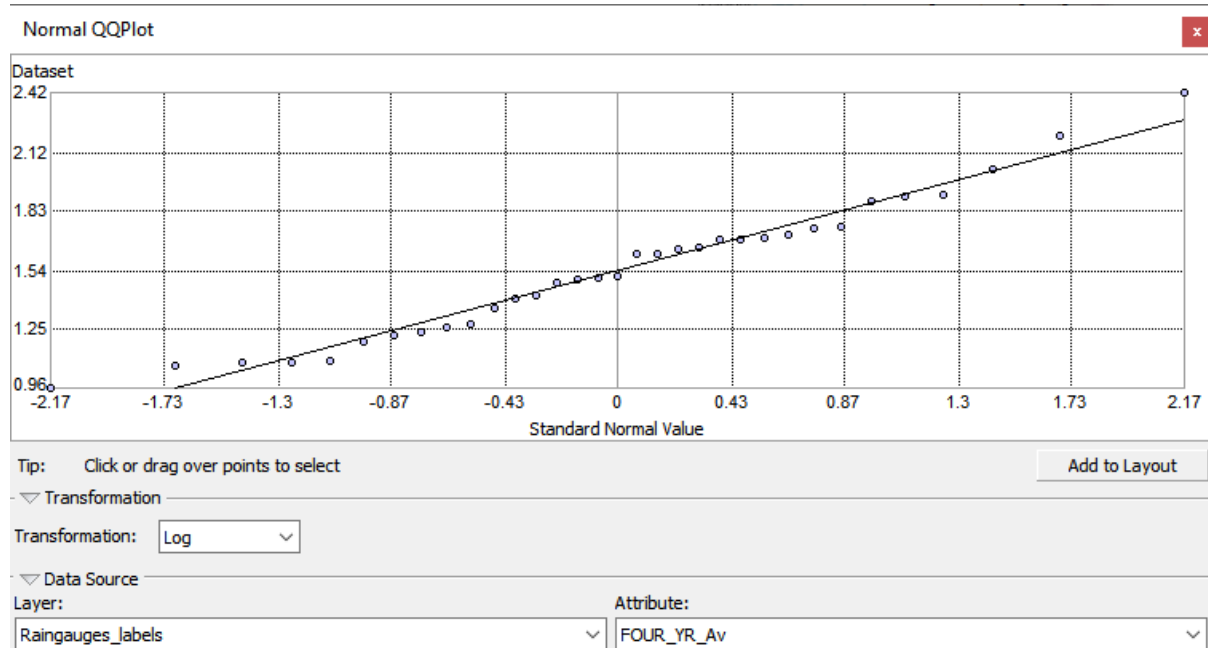
### *Histogram exploration of rainfall distribution*

A histogram exploration of the distribution of rainfall datasets was done to find out whether the data had a Gaussian distribution, or whether the transformation of the data was necessary to approximate a normal distribution for an effective Geostatistical analysis. The Kriging technique is driven by the Gaussian distribution to estimate the mean value of the rainfall, and the data needed to follow a normal distribution at least approximately. The data were checked for normal distribution by looking if it follows a bell-shaped or if the mean and median are similar or close together, the skewness near zero and kurtosis is near three (Figure 4.14). If the data is highly skewed, the data was transformed to make it normal. The Geostatistical Analyst provides functionalities for data transformation, namely: Log-normal, Box-Cox or Normal Score, so that the data became closer to a normal distribution (Figure 4.14) and this was essential to increase the accuracy of the Kriging predictions.



**Figure 4.14: Histogram of some of the rain gauge rainfall data after Log-normal and Box transformation. Data now has increased symmetry. Several rainfall maps in this study required Log-normal transformation, also called Log-normal Kriging.**

Besides the histogram, a QQ plot was used to compare the distribution of data to the standard normal distribution, providing a further measure of normality. The plot (Figure 4.15) showed the distribution of rain gauge rainfall data against the expected normal distribution, where observations should lie approximately along a straight line. If the data is not normal, the points form a curve deviates much from the straight line.



**Figure 4.15:** A QQ plot showing the distribution of rain gauge rainfall data against the expected normal distribution where observations should lie approximately along a straight line.

### ***Determining autocorrelation using the semivariogram***

Besides exploring data using a histogram and a QQ plot, the best rainfall semivariogram model in the study area was selected before the application of the Kriging technique. A variogram model summarises the spatial variation of rainfall and describes how its variation changes with the increasing separation distance between rain gauges (Chappell et al., 2013). The model  $\gamma(h)$  calculates half the average squared difference between paired data values and computed as:

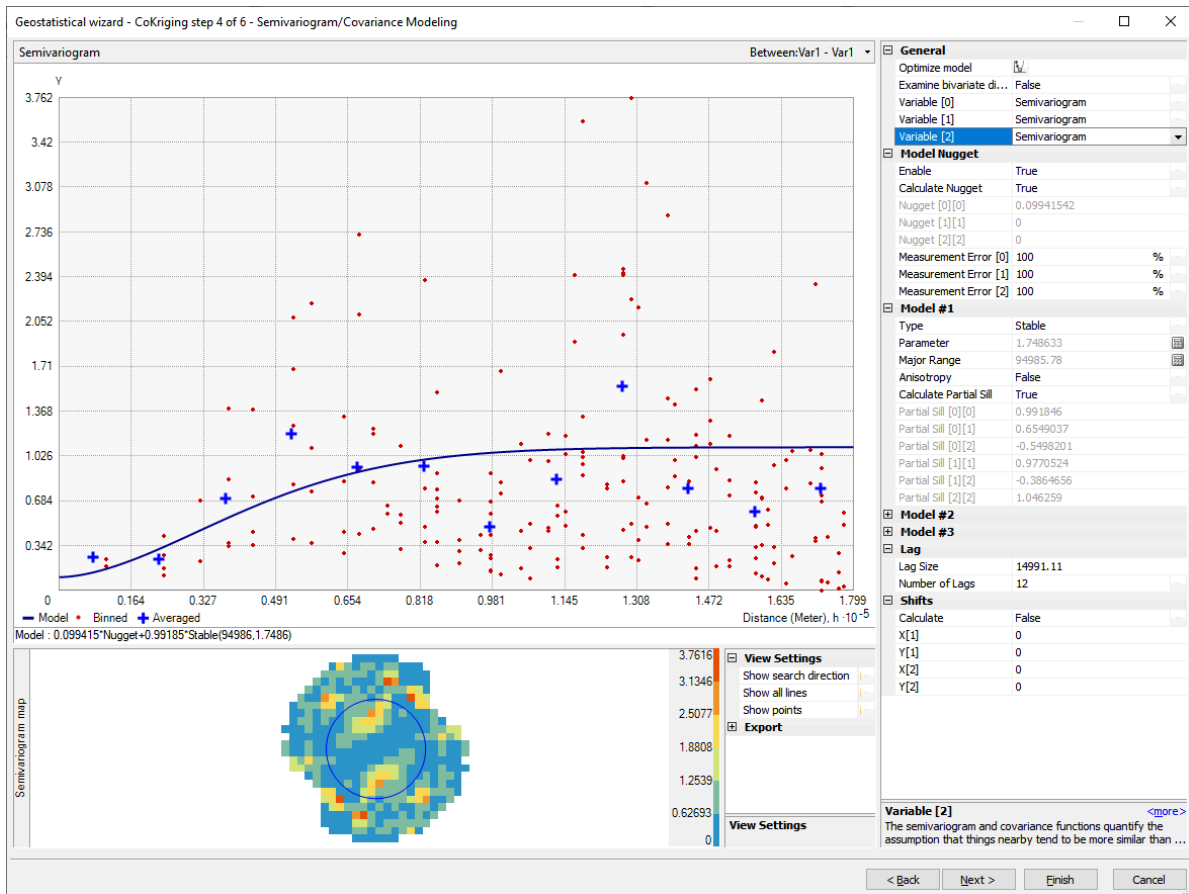
$$\tilde{\gamma}(h) = \frac{1}{2|N(h)|} \sum_{(i,j) \in N(h)} |r(s_i) - r(s_j)|^2$$

**Equation 4.5**

Where  $N(h)$  is the number of observations and  $|N(h)|$  is the observed number of point pairs  $s_i$  and  $s_j$  that are within a tolerance of a given lag  $|s_i - s_j| = h$ .



The formula calculates half the difference squared between rainfall values of paired rain gauge locations. To ensure correct Kriging interpolation and optimal merging of rain gauge and satellite rainfall datasets, several experimental semivariograms were tested/checked by analysing spatial dependencies in rainfall datasets (spatial correlation) (Figure 4.16)



**Figure 4.16: A semi-variogram/Covariance model summarising the spatial variation of rainfall in the study area and how the variations change with the increasing separation distance between rain gauges.**

## **Stationarity**

The Kriging technique also relies on the assumption of stationarity, requiring in part that all data values come from a distribution that has the same variability (Johnson et al., 2001). Stationarity proceeded distribution assumption and is a statistical property of rainfall that does not rely on exact locations requiring observations be replicated to estimate errors in the datasets (Verdin et al., 2016). The basic simple Kriging estimators under the assumption of weak stationarity were defined by Goovaerts et al., (1997) as:

$$\gamma(h) = \frac{1}{2} \text{Var}[r(s_i) - r(s_j)]$$

**Equation 4.6**

Where  $h = \|s_i - s_j\|$  is the Euclidean distance between observation locations  $s_i$  and  $s_j$ . It should be noted that the value of  $h$  has units of metres since latitude and longitude are translated to easting and northing (Verdin et al., 2016).

Stationarity was used to obtain the replication of observations and to estimate errors in the datasets. The mean of rainfall at one location should be equal to the mean at any other location and data variance constant and the correlation between any two-gauge locations should not have exact locations (Goovaerts et al., 1997).

## **Cross-validation**

The cross-validation process consisted of diagnostic tests done to show whether the model was appropriate and whether the model could provide the best rainfall estimates. The process also helped to make informed decisions about whether the model would provide the best predictions and use it for interpolation and merging. The statistical values obtained showed whether the interpolated values were realistic (Oliver et al., 2014)

Two different validation processes were performed to estimate the performance of the Kriging predictions. The leave-one-out cross-validation (L-1-OCV) was done by removing one rain gauge at a time from the analysis and recalculating its value from the remaining data. The process comprised withholding one or more rain gauge rainfall values in the study area and then use the model to predict the same values at a

particular location selected (Johnson et al., 2001). It was also performed by splitting the sample size randomly into a training (80%) and a test (20%) subset. The training subset was then used to predict the values of the test subset. In this way, comparisons were made on the predicted value to the observed value and useful information on the accuracy of the model was obtained. Geostatistical analysts provided several graphs and summaries of observed values, including predicted values, and a summary of model diagnostics statistics on Kriging prediction performance (Figure 4.17).

### ***Model performance metrics***

The Geostatistical tool provided five diagnostic algorithms, described below to assess whether the model developed can provide the best estimates of rainfall interpolation.

#### *Mean Error*

The mean error measures the average difference between the observed and predicted mean monthly rainfall values. The mean error is given as:

$$\frac{\sum_{i=1}^n (\hat{Z}(s_i) - z(s_i))}{n}$$

**Equation 4.7**

If the prediction of values is accurate, then mean error should be close to zero.

#### *Root Mean Square error (RMS)*

The Root Mean Square error (RMS) measures the prediction accuracy of the model by calculating the average error and is weighted according to the square error. It is computed as the square root of the average of the squared distances as given below:

$$\sqrt{\frac{\sum_{i=1}^n (\hat{Z}(s_i) - z(s_i))^2}{n}}$$

**Equation 4.8**

The smaller the RMS prediction error, the better or, the closer the predictions are to their actual values (Johnson et al., 2001).

Also, when the RMS is close to ASE, then the variability in prediction is correctly assessed since the two different estimates of prediction error give similar results.

#### *Average Standard Error (ASE)*

The Average Standard Error (ASE) provides the prediction standard errors as given as:

$$\sqrt{\frac{\sum_{i=1}^n \hat{\sigma}^2(\mathbf{s}_i)}{n}}$$

**Equation 4.9**

If the ASE were close to the RMS value, then the variability in prediction is correctly assessed. Also, if ASE value is greater than RMS value, it means the model is over predicting the variability of predictions (Johnson et al., 2001).

#### *Mean Standardised Error (MSE)*

The Mean Standardised Error (MSE) measures the average of the standardised errors, and the value should be close to 1.

$$\frac{\sum_{i=1}^n (\hat{Z}(\mathbf{s}_i) - z(\mathbf{s}_i)) / \hat{\sigma}(\mathbf{s}_i)}{n}$$

**Equation 4.10**

#### *Root Mean Square Standardised Error (RMSS)*

The Root Mean Square Standardised Error (RMSS) error reflects the accuracy of the model by considering the variability of the rainfall data. The RMSS prediction error should be close to 1.

$$\sqrt{\frac{\sum_{i=1}^n [(\hat{Z}(\mathbf{s}_i) - z(\mathbf{s}_i)) / \hat{\sigma}(\mathbf{s}_i)]^2}{n}}$$

**Equation 4.11**

### CoKriging model performance

The cross-validation process provided diagnostic measures that show whether the model can provide the best estimates of rainfall interpolation and merging (Figure 4.17). The smallest RMS prediction error was 1.65, showing that the model developed was an optimal model for rainfall interpolation and merging in the study area. Multiple tests were conducted to ensure that the RMS error was close to the ASE, and the nearest was 1.99 for RMS and 1.83 for ASE, the smallest difference was 0.16. The Mean Standardised Error was close to zero at 0.03, and the Root Mean Square Standardised Error was close to 1, at 1.65. With these results, the model was considered optimal for rainfall interpolation and merging in the study area.

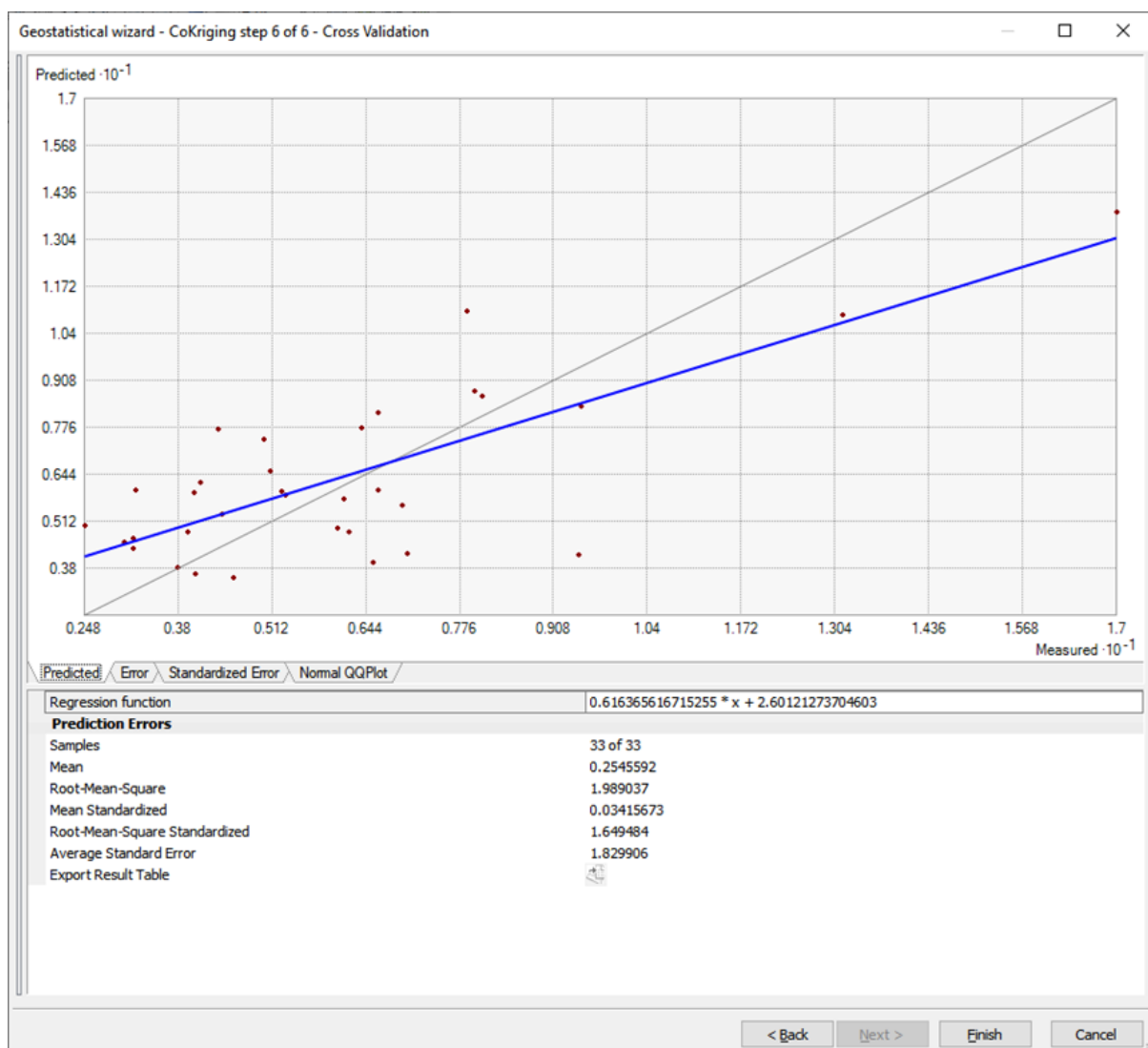


Figure 4.17: Summary of model diagnostics statistics on CoKriging prediction performance showing that the model can provide the best estimates of rainfall interpolation.

#### 4.2.6 Rainfall correction

The correction of rainfall input data for the SHETRAN hydrological model was performed following Vila et al. (2009) rainfall correction procedure that uses an additive correction scheme. The rainfall scheme used in this study included the calculation of the difference between the observed monthly rain gauge rainfall and the interpolated and merged rainfall data at each rain gauge station to obtain rainfall bias for each station (Appendix A, Figures A5). Monthly rain gauge rainfall values were higher compared to merged rainfall data due to satellite rainfall, which usually underestimates the amount of rainfall (Appendix A, Figures A2). As a result, subtracting merged rainfall from monthly rain gauge rainfall did not yield negative rainfall values.

The rainfall error quantification and correction were done in two stages: 1) Finding the long-term monthly difference between observed monthly rain gauge rainfall and the merged rainfall dataset derived from Kriging interpolation. 2) Adding the long-term difference (bias) to rainfall input data at each time interval to improve model input rainfall accuracy. The additive bias correction was defined using (Equation 4.12).

$$Corr_{rainfall} = Rain_s + \overline{(Rain_g^i - Rain_m^i)}$$

**Equation 4.12**

Where:  $Corr_{rainfall}$  represented the corrected rainfall,  $Rain_s$  satellite rainfall,  $Rain_g$  the gauge rainfall and  $Rain_m$  the merged rain gauge rainfall data and the bar  $\overline{(Rain_g^i - Rain_m^i)}$  represents the bias between observed rain gauge rainfall and merged rainfall dataset at each gauging station (denoted by the superscript  $i$ ).

The correction of rainfall input data for the model was then carried out by adding the mean rainfall bias obtained to each rainfall time series step at each rain gauge station (Equation 4.12). However, Vila et al. (2009) reported that the additive bias correction equation is not suitable when rain gauge rainfall values are low, particularly for sub-daily rainfall observations. Such situations are typically the case with short-term rain gauge rainfall (sub-daily and daily rainfall) which is much more variable, often with zero or low values compared to merged rainfall combined with regional effects such as topography which further smooths rainfall details. However, when dealing with highly variable daily rainfall data with low rainfall values, another component of the

additive formula available was the multiplicative bias correction (Equation 4.13) which suggests a ratio between the gauge and merged rainfall datasets. The ratio bias correction (RAT) is defined as:

$$rr^* = rr_{sat} \times \left( \frac{rr_{obs}^i}{rr_{sat}^i} \right)$$

**Equation 4.13**

Where the same conventions as in the additive scheme apply, the multiplicative rainfall correction formula is not suitable when substantial differences exist between the observed and estimated values (Vila et al., 2009). However, the two schemes can be combined into a single method to overcome the limitations of both schemes when used separately. In this study, no significant monthly discrepancies were found, and there were no low monthly rain gauge rainfall values in the Grijalva basin. Monthly mean satellite rainfall values were lower relative to rain gauge values. Therefore, it was appropriate to use the additive correction technique on its own. The multiplicative approach was not used in this study.

#### **4.2.7 Thiessen polygons**

Rainfall input data are assimilated into the SHETRAN hydrological model as time-series data from each rain gauge station enclosed in Thiessen polygons. Rainfall values were approximated from the closest rain gauge location and spatially averaged within large polygonal areas (Williams and Berlamont, 2002). The time series values of the closest rain gauge were assigned to each grid within a polygon, showing characteristics of the polygonal zones of influence around each rain gauge. In this way, rainfall input data divided the model domain into polygons with perpendicular bisectors between rain gauge locations in each and all points nearer to each rain gauge location enclosed in polygonal areas (Canli et al., 2018).

The Thiessen polygons were constructed in ArcMap using the Create Thiessen Polygon tool that converts rain gauge points to an output feature class with Thiessen proximal polygons around each rain gauge station (ArcGIS Desktop 2011). The points were automatically connected in GIS by drawing straight lines between the stations, and these lines were perpendicularly bisected and meet to form the polygons.

The areas of each polygon were calculated and expressed as a fraction of the total area using the Create Thiessen Polygon tool, which calculates grid-scale rainfall values based on the area-weighted sum of rain gauge data as shown in Equation 4.14 below:

$$V_g = \sum_{i=1}^n w_i V_{Si}$$

**Equation 4.14**

Where  $V_g$  is the aggregated grid-scale rainfall,  $n$  is the number of intersected Thiessen polygons within a grid,  $w_i$  is the percentage of area for intersected Thiessen polygon  $i$  in the grid and  $V_{Si}$  is the precipitation value of the intersected Thiessen polygon  $i$ .

However, polygon sizes were unevenly distributed, and the method did not accurately account for localised rainfall events. Hence, the Thiessen polygons result in errors in the representation of rainfall, particularly for short-term rainfall periods, which usually occur in the study area where the rainfall intensity varies over short distances. Also, the Thiessen polygons describe spatially distributed rainfall in the study area as abrupt transitions between arbitrary polygon boundaries that are unrealistically straight-lined (Canli et al., 2018). It would be impossible to obtain accurate rainfall input data from rain gauge data averaged in polygons, as the method does not reflect the spatial variability of rainfall in the Grijalva catchment area. The inaccuracy resulting from the use of Thiessen polygons was reduced by computing errors at each rain gauge location and correct each polygonal rainfall station.

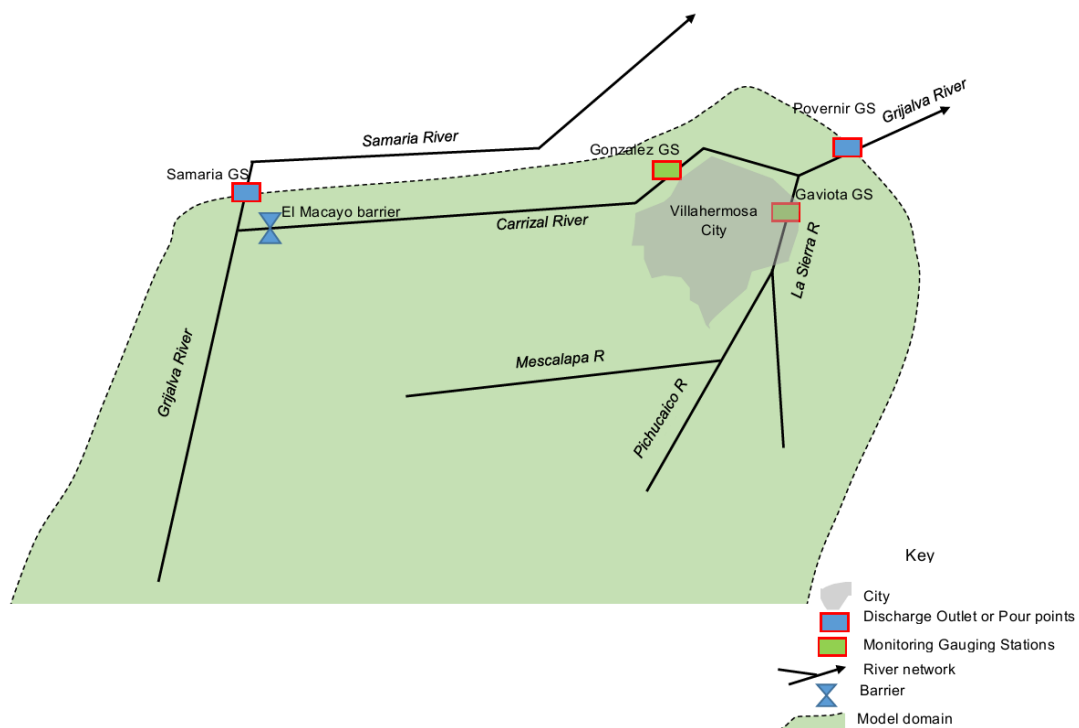
Therefore, the Kriging model was applied to rain gauge and satellite rainfall datasets to obtain merged rainfall estimates with correct spatial dimensions and magnitudes of rainfall in the Grijalva catchment area. The analysis of spatial-temporal patterns of the two rainfall datasets was conducted to identify estimated errors between the merged rainfall product and the original observed rain gauge rainfall at each station. The differences between the observed and merged estimates of datasets were taken as the rainfall error, and this was applied at each rain gauge to correct rainfall input data for the SHETRAN hydrological modelling.



## 4.2.8 Diverted flows

### *Flows from the model domain*

At the lower reaches of the Grijalva River, flows are diverted away from the Grijalva model domain into a river channel locally named Samaria River. The Grijalva River is thus split into two deltaic watercourses, the Samara River, which flows northwest, east, and the Carrizal River, which flows east along the northwest fringes of Villahermosa Region. Figure 4.18 depicts schematically the lower Grijalva River and its distributaries (Appendix B, Figures B-1). However, during the model setup, both the Flow direction and Flow Accumulation base layers of the SHETRAN hydrological model could not detect the Samaria river channel. Several SHETRAN grid squares flowing into subsequent downstream squares could not create a link between the main Grijalva River channel link with the Samaria River (Appendix B, Figures B-12).



**Figure 4.18: The lower Grijalva River showing a network of distributaries and tributaries and the location of gauging stations used in calculating diverted flows.**

As a result, the diverted flows in the model were not represented correctly as they were part of the observed flow at Povernir outlet, about 10 km downstream. To correct the model, the diverted flows were defined and all flows in the catchment area were correctly represented. Thus, all flows were accounted for, and those that were diverted were removed from the simulated flows at Povernir outlet.

### ***Diverted flows calculation***

To account for diverted flows, a percentage fraction of flows in each channel was determined based on the overall flows of the Grijalva River (Appendix B, Figures B-16 to 20). The combined total volume of flows through Samaria and Gonzalez gauging stations was determined using the Area under the hydrograph curve method (Hoyt and Grover, 2016). The flow was measured as a volume passing through a cross-section ( $V_Q$ ) of each gauge station over a given duration ( $t$ ), computing the area under the curve of hydrographs and plotting instantaneous discharge ( $Q$ ) against time ( $t$ ). Then the total volume ( $m^3/s$ ) was found by multiplying the time for the entire year by discharge in cubic metres per second. At any given time ( $t$ ) the observed discharge ( $Q$ ) was paired according to their time-series position and discharge ( $V_Q$ ) written as:

$$V_Q = (t_1 - t_2) \left( \frac{Q_2 + Q_1}{2} \right) + (t_3 - t_2) \left( \frac{Q_3 + Q_1}{2} \right) + \dots + (t_N - t_{N-1}) \left( \frac{Q_N + Q_{N-1}}{2} \right)$$

**Equation 4.15**

Where  $V_Q$  is discharged at a gauging station,  $Q$  is instantaneous discharge over a specified duration time  $t$  and  $N$  is the last observation in the time series. The equation can be generalised by denoting the location of the subscription  $i$  so that the term of each subscript is given as in Equation 4.16.

$$V_Q = (t_{i+1} - t_i) \left( \frac{Q_{i+1} + Q_i}{2} \right)$$

**Equation 4.16**

The equation can be written in summation form as (Equation 4.17):

$$V_Q = \sum_{i=1}^{N-1} (t_{i+1} - t_i) \left( \frac{Q_{i+1} + Q_i}{2} \right)$$

**Equation 4.17**

Where: operations on the right of the summation symbol were performed for each time series step from  $i=1$  to  $i= N-1$ , and the results for each were summed to obtain the total annual discharge.

The result was the cumulative annual total flow volume of the Grijalva River and its major downstream tributaries and distributaries. Similarly, the annual total discharge passing through Samaria and Gonzalez gauging stations was calculated and used to get a percentage share of each as a proportion of the total flows of the Grijalva River. The percentage share of the Samara River represented the diverted flow component. This share was used to correct the over-predicted simulated flows by removing them from the Povernir discharge output at each time step.

#### **4.2.9 Assumptions relating to groundwater.**

Modelling the Grijalva river flows assumed a default no-flow boundary condition that does not allow flow but results in permeability loss at its boundary. It was assumed that the Grijalva catchment area has a watertight subsurface boundary with a groundwater divide that coincided with the surface water divide. The groundwater basin was modelled as an impervious bed with no inter-basin movement of groundwater (Sharad et al., 2018). The lower boundary was assumed to be at a depth where no significant flow of deeper regional aquifers occurs during the timescale of the simulations. Also, the catchment subsurface was assumed to be isolated apart from the downstream (outflow) boundary condition at the model outlet or pour point (Lewis et al., 2018).

The flow boundary condition was defined at the downstream end of the river only, no-flow boundary conditions were used for all aquifer boundary conditions (Birkinshaw et al., 2008). All fluxes between sub-catchments were transferred as a channel flow. The net gain or loss of groundwater through the riverbed was also assumed to be zero, with no evidence to the contrary.

The subsurface material was assumed to comprise a variably saturated porous medium with a depth from the bottom layer (m) varying within  $\geq 1.2$  and  $\leq 6.2$  (> layer above and < layer below). The porous media has two layers: a highly permeable A-

horizon depth ranging between 1 and 1.2 m and a less permeable B horizon at a depth of 5 to 6.2 m. Also, the aquifer porous material properties were assumed to be characterised by the hydraulic conductivity parameter, a measure of speed of water flow through the porous media. Other assumptions of the porous media include infiltration that only occurs in a vertical direction; however, shrinking and expansion of the porous media was negligible, and no incrustation on the media surface occurs during rainfall events. The study assumed that the runoff in the study region was controlled by the surface water flow rather than the subsurface flow.

## **4.3 SHETRAN hydrological modelling**

This subsection presents the methodology adopted to develop a large distributed hydrological model using the SHETRAN hydrological model. Presented first are the computing requirements, preliminary model runs, and analysis. The second part provides data on model input, calibration and validation, and parameterisation. The performance evaluation criteria used to assess SHETRAN performance is also provided.

### **4.3.1 Model runs and computing requirements**

The fully distributed modelling approach applied at a large catchment scale using high-resolution datasets had high computational power requirements. Large memory was required for each grid cell to store information on catchment conditions and processes. The model was developed using three-hour rainfall time-series input data at 2 x 2 km<sup>2</sup> spatial grid. The SHETRAN model ran over a five-year simulation period from 2013 to 2017. Preliminary model runs were performed using baseline (average) values for all parameters to investigate how the outputs differed from observed discharge. There was no knowledge of which parameters were sensitive in the catchment area. Sensitivity analysis was, therefore, conducted to determine the relevant and sensitive parameters for model calibration.

### **4.3.2 Sensitivity analysis**

A sensitivity analysis was conducted to assess the effect of changes in model parameters on model results (Pechlivanidis et al., 2011). The analysis was done to assess model performance response to model parameter changes (Moriasi et al., 2007). The calibration process required knowledge on the sensitivity of discharge output results to model parameter changes i) to better assess the model response, ii) to reduce the number of parameters for calibration and iii) to define the parameter uncertainty (Ewen et al., 2007; Op de Hipt et al., 2017).

Besides, the sensitivity analysis helped to identify parameters with negligible effect on discharge output. A higher sensitivity test value suggested a highly relevant parameter

for selection and use during calibration. The sensitivity analysis tests were therefore done to minimise the number of parameters to be adjusted when calibrating the model, required were critical parameters to focus on based on their sensitivity (Arnold et al., 2012). A local sensitivity analysis was done using the traditional “one factor at a time” (OFTAT) method, also known as one-variable-at-a-time. The impact of varying parameter values provided insight into the most sensitive ones (Op de Hipt et al., 2017).

The sensitivity was determined for each parameter within the range of allowable literature parameter values for each. A traditional approach was used of increasing each parameter's value initially by 10% to test the change in discharge output, and this was followed by a decrease of each parameter by -10% to control the output. The impact of parameter changes on the total variance of model output was computed, which means the total maximum discharge was used for comparisons. Equation 4.18 was employed to quantify the sensitivity of each parameter analysed (Op de Hipt et al., 2017).

$$SI_{90} = \frac{|O_{90} - O_{-90}|}{O_0}$$

**Equation 4.18**

Where:  $SI_{90}$  is the index of sensitivity,  $O_{90}$  and  $O_{-90}$  are the model outputs resulting from increasing or decreasing the value of the parameter by 90%, and  $O_0$  is the model output from the baseline run.

However, some parameters had an insignificant impact on discharge and had their baseline average values maintained during calibration (Arnold et al., 2012). Although sensitivity analysis of single parameters ignores interactions between parameters, it provided information on how sensitive discharge output was in the catchment to changes of parameter values.

### 4.3.3 Model calibration

Calibration is a process of changing parameter values of the model to replicate observed historical data within a range of performance criteria (Abbott et al., 1986). Calibration attempts to improve the parameterisation of a model to field conditions and reduce forecast error (Figure 4.19) (Arnold et al., 2012). Although parameters of physically-based models, theoretically do not need calibration, adjustments were required to overcome unrealistic representation of catchment properties and processes in the model (Beven, 2011; Op de Hipt et al., 2017).

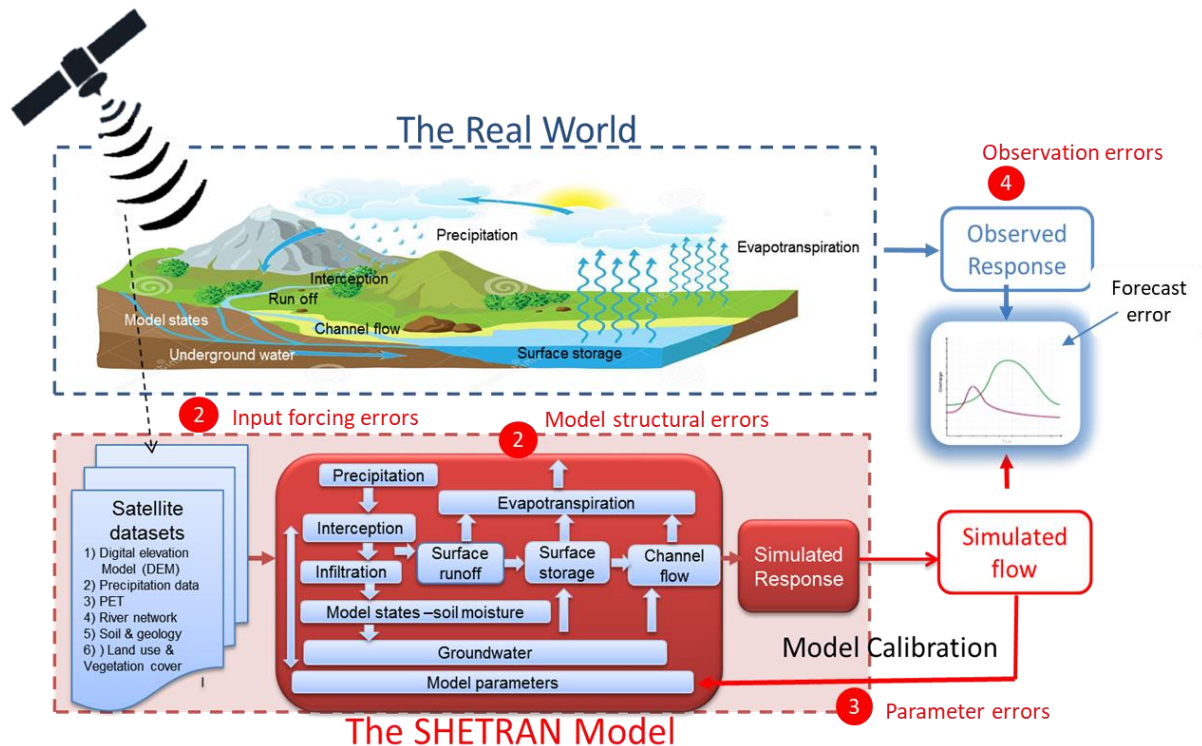


Figure 4.19: Model calibration and sources of errors (Adapted from Ryu, 2014).

The SHETRAN hydrological model is sensitive to initial conditions, and two options were available to manage this sensitivity. The first was the use of a Hot Start file that allows a new model to run starting with the previous soil water potential, allowing the model to run with the same initial conditions as the previous. The second option available was to discard the run-in period up to the first weeks of any simulation. The second option was used because after each model was run, the Hot Start file approach produced large files, taking a lot of computer storage.

Due to limited calibration and validation data, a temporal split-sample calibration-validation strategy was used to calibrate the SHETRAN hydrological model (Daggupati et al., 2015). In this strategy, observed data were split into two periods, that is, for the 2013 to 2017 sample period (Appendix C, Figures C-1) was divided into 2013 to 2015 and 2016 to 2017 sample periods. The 2013 to 2015 sample was used to calibrate the model, and they used 2015 to 2017 for model validation.

The SHETRAN hydrological model was calibrated by adjusting the selected (sensitive) parameters so that the model simulations best match the observed discharge (Duan et al., 1992; Duan et al., 1994). Considerable efforts were made to find parameter values in literature and manually use the try-and-error method to find optimum values that result in model simulations matching as closely as possible observed discharge (Ajami et al. 2007). Several model runs were conducted after each parameter change with different combinations of parameter values to give the near-optimal solution with reduced simulation error.

### ***Calibrated parameters***

The parameters that were calibrated for discharge simulations were the Strickler overland coefficient, the saturated hydraulic conductivity for the unsaturated zone, and the AET/PET at field capacity (Table 4.4). The parameters were adjusted to obtain envelopes of discharge (upper and lower) and vertical variation of the base flow of the simulation period. The lower values of AET/PET and hydraulic conductivity increased discharge output and were used to establish the upper boundary envelope of the hydrograph. The lower boundary envelope was determined by higher values of AET/PET, hydraulic conductivities, and the upper values of hydraulic conductivity and Strickler overland coefficient parameters that cause a decrease in discharge output. Increasing saturated hydraulic conductivity and Strickler overland parameter coefficient values contributed to increases in discharge output. Conversely, an increase in parameters values for saturated hydraulic conductivity, the Strickler overland coefficient and the actual/potential evapotranspiration ratio result in a decrease in total discharge.



Also, the overland flow was generated by upward saturation of the soil or by an excess of rainfall over soil infiltration rate. The groundwater table dynamics were adjusted by varying the physical properties of subsurface material by calibrating soil depth, hydraulic conductivity, and porosity parameter values (Shrestha et al., 2017). Overland flow inputs determine channel flow by interacting with the saturated zone. The parameters were varied within permissible bounds until some calibration criterion was met using a visual match of simulated and observed trends. Also, a more formal criterion was applied, which comprised minimising the percentage error in the predicted peak, minimising the Root Mean Square value of the difference between the simulated and measured discharges throughout the calibration period.

The resulting baseline values, the upper, middle and lower values of parameters used during calibration including the original values, calibration/sensitivity range applied, and final calibrated values of the critical parameters are shown in Table 4.4. These are the best estimates of the parameter values, and their bounds were used for uncertainty analysis in Section 4.6. However, there is a need to measure parameter values at the appropriate spatial scale for representing at SHETRAN grid-scale of 2 x 2 km<sup>2</sup>.

**Table 4:4 The upper, middle and lower values of parameters used during calibration, including the original values, calibration/sensitivity range applied and final calibrated values as used.**

Sensitivity	Parameters	Description	Value range	Original values	Final values	Unit
Highly sensitive	K <sub>1</sub>	Strickler overland (Arable)		2.55	9.48	m <sup>1/3</sup>
	K <sub>2</sub>	Strickler overland	≥0.3 ≤ 9.9	1.5	6.70	m <sup>1/3</sup>
	K <sub>3</sub>	Strickler overland		1.0	7.48	m <sup>1/3</sup>
	K <sub>s1</sub>	Saturated cond. (Coarse)		27.0	1.0	m/da
	K <sub>s2</sub>	Saturated cond. (highly)	≥0	56.25	0.20	m/da
	K <sub>s3</sub>	Saturated cond. (Medium)		12.0	0.30	m/da
	AET/P	AET/PET ratio (Arable)		0.05	0.35	(-)
	AET/P	AET/PET ratio (Deciduous)	≥0.01 ≤1.99	0.15	0.70	(-)
	AET/P	AET/PET ratio (Evergreen)		0.1	0.84	(-)
		h <sub>1</sub>	Topsoil depth (Coarse soil)		1.0	1.0
	h <sub>2</sub>	Topsoil depth, highly Prod.	≥ 0.3 ≤20.0	1.2	1.2	m
	h <sub>3</sub>	Topsoil depth (Medium soil)		1.2	1.2	m
Low sensitivity	n <sub>1</sub>	van Genuchten n (Coarse)		1.38	1.38	(-)
	n <sub>2</sub>	van Genuchten n highly	≥ 1.2	6.0	6.0	(-)
	n <sub>1</sub>	van Genuchten n (Medium)		1.18	1.18	(-)
	α <sub>1</sub>	van Genuchten α (Coarse)		0.04	0.04	cm <sup>-1</sup>
	α <sub>2</sub>	van Genuchten α highly	>0.0	0.01	0.01	cm <sup>-1</sup>
	α <sub>1</sub>	van Genuchten α (Medium)		0.03	0.03	cm <sup>-1</sup>
	θ <sub>r1</sub>	Residual water content	≥ 0 ≤0.8	0.02	0.02	m <sup>3</sup> /m
	θ <sub>r2</sub>	Residual water content		0.02	0.02	m <sup>3</sup> /m
	θ <sub>r3</sub>	Residual water content		0.01	0.03	m <sup>3</sup> /m
	θ <sub>s1</sub>	Saturated water content		0.40	0.81	m <sup>3</sup> /m <sup>3</sup>
θ <sub>s2</sub>	Saturated water content	≥ 0.05 ≤1.0	0.3	0.30	m <sup>3</sup> /m <sup>3</sup>	
θ <sub>s3</sub>	Saturated water content		0.439	0.44	m <sup>3</sup> /m <sup>3</sup>	

#### 4.3.4 Capturing dam storage

The hydrological modelling in the Grijalva catchment area considered dams to have a major impact on the hydrological flows in the area, so all primary water holding bodies in the study area, including reservoirs or dams, lakes and wetlands were represented in the model. The SHETRAN hydrological model was designed to allow input of a Lake mask layer to represent the location of lakes and dams in the model and capture the effect of storage to represent the effect of water bodies on flows in the study area.

The Lake mask layer was configured within the SHETRAN XML (library) file, `<LakeMap>grijalva_lakes.txt</LakeMap>` to represent the location and effect of water storage in the model. The layer sets the Strickler coefficient value for river channels where dams and lakes are located to a value of 3 instead of the default value of 20 for normal river flow. All river channels were assigned a Strickler default parameter value of 20 for uncontrolled river flow, and SHETRAN adjusted this parameter using the Lake mask layer developed, so it corresponds to 3, a Strickler value set for lakes, dams or reservoirs (Birkinshaw, 2012). If the Lake mask grid cell intersects with a river link, it is treated as an open water body, reducing the default Strickler coefficient value from 20 to 3, in this way controlling or varying river flows. When the value of the Strickler coefficient is low in the river channel, and heavy rainfall with large river inflows occurs in the channel, the water level increases and overflows into nearby grid squares of the model domain (Lewis et al., 2018).

Therefore, the low Strickler coefficient value has an effect of slowing river flows and increase storage of water in the river channel. In this way, the Lake mask layer and the Strickler coefficient value were calibrated to attenuate river flows similarly when a river flows into a lake or a dam. The flows were assessed based on a series of annual water balances for the Grijalva catchment, where individual input and output components of the water cycle were estimated for water mass accounting of the catchment.

### 4.3.5 Mass balance checks

To aid calibration, mass balance checks were conducted to ensure the model was not gaining, storing and losing inappropriate amounts of flows. The initial calibration of the model was done to get the overall annual mass balance to acceptable levels to facilitate calibration for the correct timing, prediction of peaks, and base flows. The mass balance output values were extracted from the SHETRAN mass balance file; *output CATCHMENT mb.txt* file (Birkinshaw, 2012). After each model run, the model records catchment variables, and these were compared and checked using the mass balance equation (Equation 4.23). Recorded values, including simulated discharge and other simulated output variables, were checked and compared with observed values (Birkinshaw, 2012). The simulated and observed discharge values were plotted in Excel using the Month and Year functions on Pivot tables, and the accumulated total is obtained to populate the Mass balance equation below.

$$\text{Inflow} = \text{Outflow} \pm \text{change in storage } (\Delta V)$$

**Equation 4.19**

Inflow (Cum Precipitation) = Outflow - (Cum. Canopy Evap. + Cum. Soil Evap + Cum. Discharge + Cum. Transport)  $\pm$  Change in storage (Subsurface Storage + Land Surface Storage + Channel Storage + Canopy Storage).

When the simulated outflows were high, more storage was required, the land surface storage was increased by reducing the value of the Strickler coefficient to store more water in the system. Another important parameter that was calibrated included the AET/PET ratio for each vegetation type that contained useful information to constrain parameter sets and to set discharge at correct levels. For example, to increase evaporation, the AET/PET ratio was increased to allow more water to leave the system. Besides discharge observations, the mass balance data contained useful information for constraining parameter sets. A SWAT based Calibration Helper v1.0, a customised VBA coded Microsoft Excel dashboard file was modified and used to provide options for controlling, evaluating and calibrating water balance components of the SHETRAN hydrological model (Figure 4.20). Also, the spreadsheet dashboard enabled visualisation and comparison of different plots of water balance components (Shrestha et al., 2017).

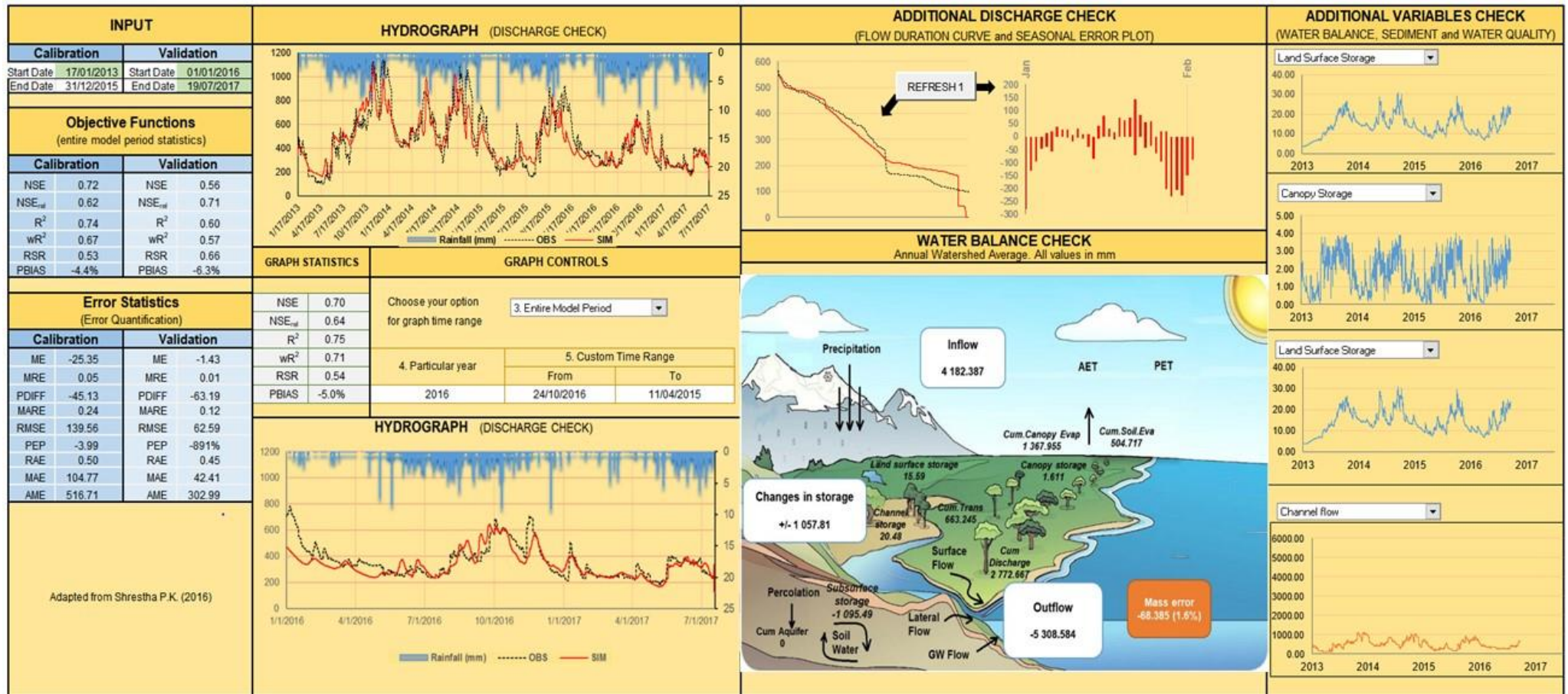


Figure 4.20: A SWAT Calibration Helper v1.0 was modified and used to visualise and compare different plots of the SHETRAN water balance components and aid the calibration and validation of the model (Shrestha et al., 2017).

#### 4.3.6 Model validation

Model validation was performed to ascertain if the calibrated flows mirrored realistic flow conditions in the study area. The model was run with other related rainfall events of a different time that occurred in the study area. The observed data between 2016 and 2017 period was used for the validation, to test the prediction ability of the model in different rainfall conditions and verify if it can generate sufficiently accurate simulations (Refsgaard, 1997 in Arnold et al., 2012). The model was also validated using the same parameters selected in the sensitivity analysis and the discharge results were compared with those observed in 2016 to 2017 (Daggupati et al., 2015).

#### 4.3.7 Model evaluation during calibration and validation

The SHETRAN's hydrological model performance during calibration and validation was assessed using Correlation Coefficient (CC), Nash-Sutcliffe (NSE), Percentage Bias (PBIAS) and the Root Mean Square Error (RMSE). A combination of graphic methods and dimension error-index statistics for model evaluation was adopted.

The NSE (Equation 4.19) was used to show the statistical strength of the simulated and observed discharge (Nash & Sutcliffe, 1970):

$$NSE = 1 - \frac{\sum(Q_{i,o} - Q_{i,c})^2}{\sum(Q_{i,o} - \bar{Q}_o)^2}$$

**Equation 4.20**

Where:  $Q_{i,o}$  is the mean observed discharge,  $Q_{i,c}$  is the simulated discharge and  $Q_o$  is the observed discharge output values. If  $NSE \geq 0.5$ , the model simulation was judged to be satisfactory (Table 4.5).

The Pearson Correlation Coefficient (CC) was used to evaluate the relationship between simulated and observed discharge (Equation 4.20).

$$cc = \frac{\sum(Q_{i,o} - \bar{Q}_o)(Q_{i,c} - \bar{Q}_c)}{\sqrt{\sum(Q_{i,o} - \bar{Q}_o)^2(Q_{i,c} - \bar{Q}_c)^2}}$$

**Equation 4.21**

Where:  $Q_c$  is the average simulated discharge values and other values as above. When  $CC \geq 0.5$ , the model output was considered to be satisfactory (Table 4.5).

The Percentage Bias (PBIAS) was used to measure the prevailing inclination of simulated values greater or less than the observed data used. The ratio evaluated the systemic bias of the simulated discharge (Equation 4.21) and was given as:

$$BIAS = 100 \times \frac{\sum_{t=1}^T Q_o^t - \sum_{t=1}^T Q_m^t}{\sum_{t=1}^T Q_o^t}$$

**Equation 4.22**

A percentage BIAS of  $\leq 25\%$  was considered optimal for river discharge, low values indicate accurate model simulation. Positive values show model overestimation error and negative values show a model overestimation error (Moriasi et al., 2007).

The Root Mean Square Error (RMSE) is the quadratic scoring rule which measures the average magnitude of the error. It was used to estimate the difference between the observed and simulated discharge values with 0 showing a perfect fit and was given as below (Equation 4.22).

$$RMSE = \left( \frac{1}{T} \sum_{t=1}^T (Q_o^t - Q_m^t)^2 \right)^{1/2}$$

**Equation 4.23**

If  $RMSE \leq 0.70$  is considered satisfactory and model simulation is judged as adequate (Table 4.5).

**Table 4:5 Interpretation of statistical indicators for model calibration and validation (Moriasi et al., 2007).**

<b>PBIAS (%)</b>	<b>NSE</b>	<b>CC</b>	<b>RMSE (mm)</b>	<b>Interpretation</b>
PBIAS < 10	0.75 < NSE < 1.0	0.75 < CC < 1.0	0.00 < RMSE < 0.50	Very good
10 < PBIAS < 15	0.65 < NSE < 0.75	0.65 < CC < 0.75	0.50 < RMSE < 0.60	Good
15 < PBIAS < 25	0.50 < NSE < 0.65	0.50 < CC < 0.65	0.60 < RMSE < 0.70	Satisfactory
PBIAS > 25	NSE < 0.50	CC < 0.50	RMSE > 0.70	Unsatisfactory

*Notes: PBIAS, Percent bias; NSE, Nash-Sutcliffe efficiency; RMSE-Root Mean Square Error*



## **4.4 The Flood Modeller 2D model**

This study used the Flood Modeller, a 2D/1D inundation model, to simulate flood extents between Gaviotas and Povernir gauge stations along the lower reaches of the La Sierra River. The 2D modelling application was coupled with the SHETRAN hydrological model to study rainfall and parameter uncertainty propagation to assess and quantify the effects of propagation and interactions of rainfall and parameter uncertainties on the magnitude of peak discharge and flood inundation extents. The two models were linked in such a way that outputs from the SHETRAN hydrological model feed into Flood Modeller as input data.

### **4.4.1 The model selection criteria**

The Flood Modeller application was selected to simulate flood inundation extents in the Las Gaviotas flood plains, downstream of the La Sierra River. The application is for simulation of river levels and flood extents in river channels and floodplains using one-dimensional (1D) and two-dimensional (2D) hydraulic computational solvers. The 1D solver provides stable and dynamic 1D solvers and three 2D solvers designed to simulate river, overland estuarine and coastal flow situations where variations in rapid flow are absent (Banks et al., 2013). The Flood Modeller also comprises a wide range of tools for drawing flood maps for flood forecasting, flood management, catchment techniques and many other applications including low-flow modelling, sedimentation and water quality.

The package was chosen primarily because it is an efficient 1D/2D modelling software with fast simulations up to 10 times faster than similar applications (Néelz and Pender, 2013). The flood modelling tool speeds up flood risk assessments using simplified hydraulics, for example, the 2D FAST Solver produces results in seconds or minutes and is up to 1 000 times faster than conventional 2D models which take days to complete modelling tasks. The Total Variation Diminishing (TVD) solver can depict rapid shifts in the profile of water surfaces but can take longer run times. This study used the steady-state ADI Solver to simulate river flooding for both channel and floodplain flooding using the power of modern graphics cards to solve quickly shallow water equations. The Flood Modeller 2D Graphics Processing Unit (GPU) uses a

similar solution strategy that is ten times faster compared to the solving shallow water equations using the standard computer Central Processing Unit (CPU).

Also, the Flood Modeller was chosen because it has been peer-reviewed and extensively validated by several organisations for a long period (Néelz and Pender, 2013). There is a sufficient pool of people in industry and academia who are experienced in using the selected software. Support and training were available to address any bugs or issues and to obtain expert guidance on rigorous or unusual model schematisations. The software was considered not to restrict the potential use and advancement of the model to other particular goals, either because the software is in house to a particular organisation or because of limited skills. Also considered were licensing options and costs for model re-runs or updates. The free licence used was able to meet the project objectives and provided required workflows, support and model runs.

#### **4.4.2 Data requirements**

The data required for the 2D flood inundation modelling fall into four distinct categories, (a) hydrological data (b) terrain data of the river channel and its floodplain (c) flow data provided by model input boundary conditions, (d) roughness coefficient values for the channel and floodplain.

##### ***Discharge data***

The SHETRAN hydrological model was linked to the Flood Modeller application in such a way that its outputs provided input data and the required initial boundary conditions for the Flood Modeller inundation model. The La Sierra catchment is relatively well gauged with all rainfall and flow gauging stations maintained and operated by CONAGUA and these comprise river level and flow records. Observed discharge data were available at Povernir and Las Gaviotas Gauging Stations. Additional flow data were obtained from the Mexican Surface Water Management and Rivers Engineering (GASIR) National Surface Water Data Bank database (BANDAS). Rainfall data for the SHETRAN model were available from 11 rain gauge stations spaced approximately 10–15 km apart in the La Sierra catchment sub-catchment (Figure 4.21).

### ***High-resolution LiDAR Digital Terrain Model (DTM)***

Topographic data were used to determine water surface heights, baseline data for flood elevations and flood extent outlines (NRC, 2009; Mason et al., 2011). The High-Resolution LiDAR Digital Terrain Model (DTM) with ground surface objects removed was obtained from the Mexican Institute of Statistics and Geography (INEGI), (2012) and used as the topographical base data for the inundation modelling. The dataset provided continuous data of "bare earth" topography across the study area at a horizontal grid resolution of 5 metres.

### ***Channel bathymetry***

In-channel bathymetry was obtained from a boat survey undertaken in 2018 by CONAGUA in La Sierra River (Priego-Hernández et al., 2018). The Bathymetry data comprised a series of channel depth cross-sections obtained using vessel-mounted sonar equipment. The surveyed areas covered the river channel, and the downstream sections of the river and the data provided information on channel geometry, which includes the depth and width of the river sections.

### ***Aerial photography***

Aerial photography was also used as an information source for the model roughness parameter. Since hydraulic roughness is a land use and vegetation cover classification characteristic, hydraulic grid maps were derived from several sources. Although errors were present, aerial photographs, land use/cover maps and vegetation cover, including satellite data, a source of spatial roughness parameter values (Schumann et al., 2007). Therefore, Manning's 'n' coefficient of roughness values was determined primarily using aerial photographs, land use/cover maps and satellite vegetation maps. The roughness values in Las Gaviotas flood plains were based on four types of land use: Arable, Grass Forest, Deciduous/evergreen and Urban.

### ***Historic flood level and extent data***

Historic flood levels, photographs and satellite imagery were used to provide information for calibrating the model (Stephens et al., 2012). The flood event modelled in this study occurred on the 27 October 2015 at 3 pm and flows and levels observed

at Las Gaviotas Gauging Station and the extents reached were used for model calibrating. During the 27 October 2015 flood event, the Las Gaviotas flow and level gauging station recorded flow levels that reached 4.86 m above sea level and maximum river discharge of 744.75 m<sup>3</sup>/s. The Landsat 7 ETM+ satellite imagery was useful in reflecting the extents of the flood envelope in the study area.

## **4.5 Nested catchment modelling**

A study was conducted to investigate scaling relationships between rainfall variability and peak discharge in nested catchment drainage areas in the La Sierra catchment. An examination of the effect of temporal rainfall variation and peak discharge on the flood scaling structure was conducted. The rainfall variability properties that were taken into consideration were rainfall depths, direction, location and duration. A study of the spatial relationship between peak discharge and drainage areas for individual rainfall events was included (Gupta et al., 1996). The analysis also included peak discharge simulations using a hydrological model that was fed with rainfall inputs of various temporal resolutions (Mandapaka et al., 2009). The results were translated into statistical models for estimating scaling parameters and peak discharge in the La Sierra sub-catchment (Wilkinson and Bathurst, 2018).

### **4.5.1 Data availability**

Discharge data for La Sierra catchment were available for some nested gauging stations, mostly for the years 2012-2015 inclusive (Appendix D, Figure D-2-5). Most of the hydrological discharge data for the nested hydrometric network were obtained from CONAGUA, and the rest were obtained from the Mexico Surface Water Management and Rivers Engineering (GASIR) and the National Surface Water Data Bank (BANDAS) databases. Rainfall data were available from twelve rain gauge stations covering the 2012 to 2015 years covering La Sierra catchment area (Figure 4.21). However, discharge data for El Puente and Tapijulapa nested catchment areas were limited, with missing values or showing some gaps in their time series (Table 4.6).

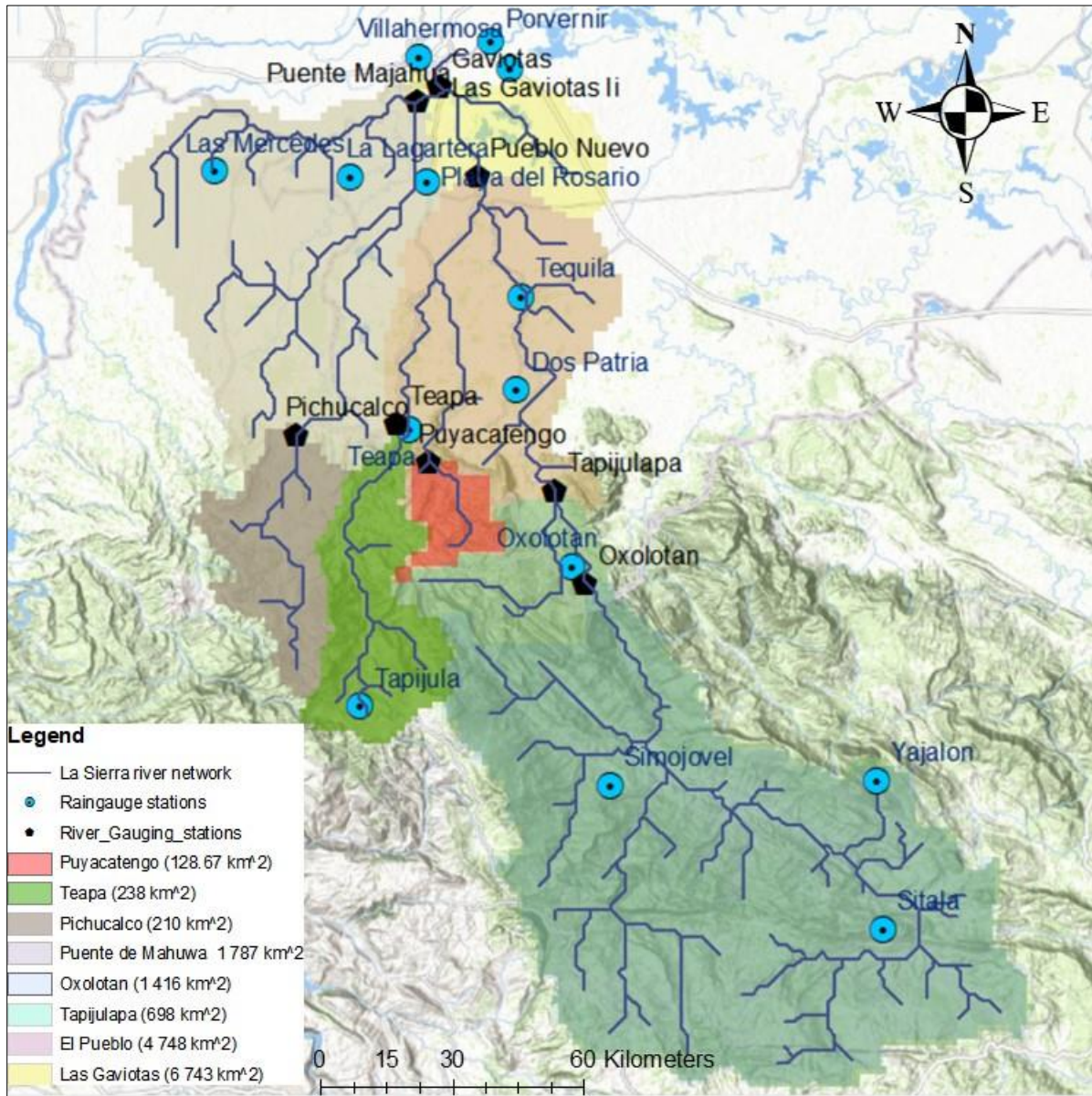
**Table 4:6 Discharge data availability for the nested catchment areas.**

Gauging Stations	Years with data								
	1940	1949	1959	1968	1978	1987	1996	2006	2015
Gaviotas	■	■							■
Pueblo Nuevo			■	■	■	■	■	■	■
El Puente									
Oxolotan				■	■	■	■	■	■
Tapijulapa				■	■	■		■	
Teapa		■	■	■	■	■	■	■	■
Puyacatengo			■	■	■	■	■	■	■

#### 4.5.2 Defining nested catchments

The La Sierra catchment was divided into six nested sub-catchments to investigate the flood scaling structure across various spatial scales defined by nested catchments (Figure 4.21). Flow gauging stations in the catchment were used as model pour points or outlets in delineating the sub-catchments. The gauged catchments were unit drainage areas of the larger basin (La Sierra catchment), allowing analysis of flood generation at different scales. The nested catchments were delineated and extracted using the 30 arc-second-resolution Digital Elevation Model (DEM) from the Shuttle Radar Topography Mission (SRTM) (Rabus et al., 2003).

The following sub-catchments were delineated and named after their outlet gauging stations: Gaviotas (6 743 km<sup>2</sup>), El Pueblo (4 748 km<sup>2</sup>), El Puente (1 787 km<sup>2</sup>), Teapa (238 km<sup>2</sup>), El Puyacatengo (128.67 km<sup>2</sup>), Oxolotan (1 416 km<sup>2</sup>), and Tapijulapa (698 km<sup>2</sup>) sub-catchments (Figure 4.21). However, the El Puente and Tapijulapa gauging stations and their nested catchments were not used in the analysis because of insufficient hydrometeorological records for the years under consideration. ArcGIS zonal statistics were used to calculate the drainage area of each nested catchment.



**Figure 4.21: Nested catchments and gauge network in the La Sierra catchment area for the application of scaling relationships between peak discharge and drainage areas.**

### 4.5.3 The procedure for establishing scaling relationships

The following procedure (Figure 4.22) was followed to examine the relationships between rainfall variability and flood scaling parameters in nested catchment areas (Ayalew et al., 2018).

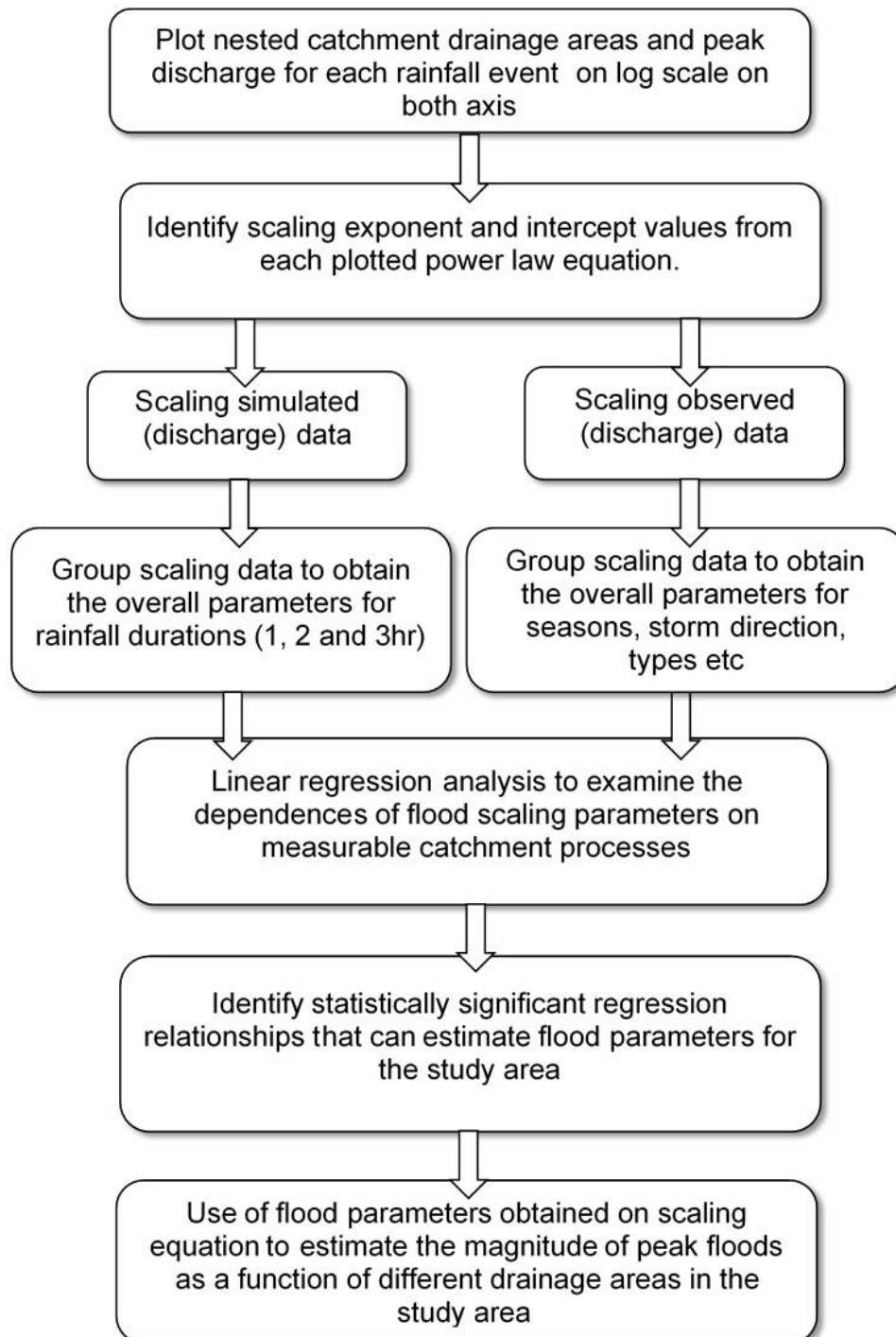


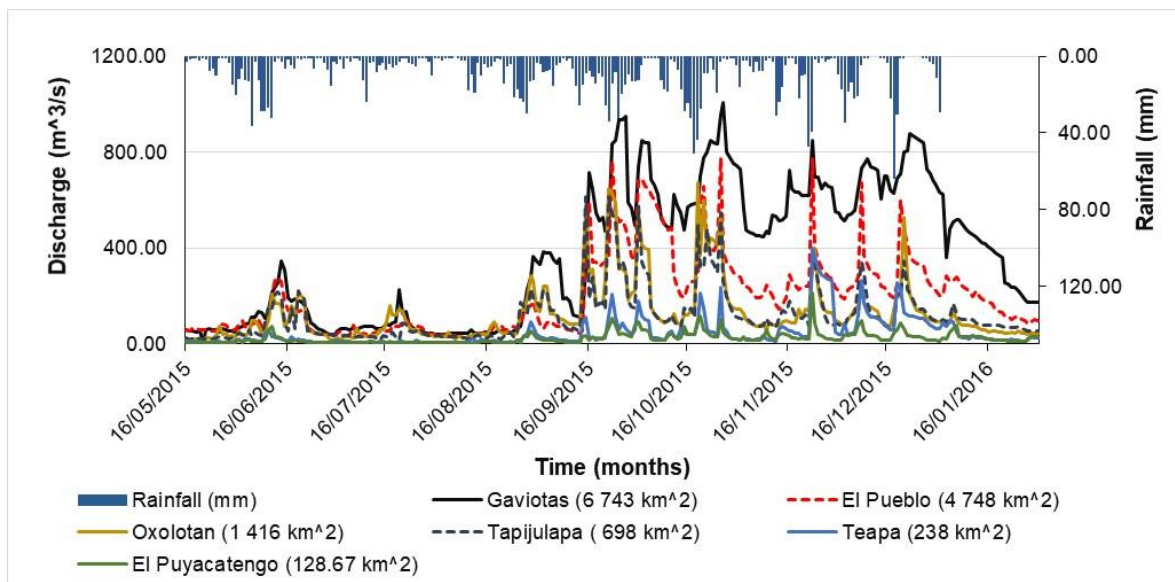
Figure 4.22: Statistical flood scaling relationship analysis procedure to estimate scaling parameters and peak discharge in the La Sierra catchment.



### **Selection of observed peak discharge events**

In selecting rainfall-runoff events that resulted in peak discharge, the first step was to identify a window that separates individual independent peak discharge events at each nested catchment outlet. Rainfall-runoff events associated with peak discharges were identified using the time of concentration of flows in the La Sierra catchment, that is, the period needed for a flood wave to travel downstream from the furthest headwater catchment point to the catchment outlet. This minimum time window size separates individual rainfall-runoff-peak events (Ayalew et al., 2015). The estimated time of concentration used for the La Sierra catchment was approximately three days.

Rainfall data relating to peak discharge were selected and analysed, assuming this happens over a time window that begins three days before the first peak discharge is observed in a catchment (Ayalew et al., 2015). All peaks were selected from the observed discharge time series, which show stream flow response within the three days until the peak discharge was observed at the outlet (Figure 4.23). A total of 59 storms that occurred between 2012 and 2015 were identified for analysis.



**Figure 4.23: Peak discharge was selected from the observed discharge time series of each nested catchment, which shows stream flow response due to individual rainfall events.**

However, only peak discharge above a two-year-return period was selected for analysis, and this threshold was a strict criterion set out to define only significant rainfall-runoff events in the catchment were studied. Compared to other studies that analysed floods resulting from basin-wide rainfall alone, this study considered all types of storms with a maximum and partial rainfall coverage of the La Sierra catchment area. The approach considered the typical localised convective tropical rainfall that occurs in the study area.

#### **4.5.4 Hydrologic model simulations**

Six nested hydrological models were developed using the SHETRAN hydrological model, and these were used to conduct simulation experiments to investigate the effect of rainfall of various time scales on the structure of scaling relationships. All nested models developed were calibrated using the OSTRICH (Optimisation Software Toolkit for Research Involving Computational Heuristics) (v17.12.19) (Matott, 2005) tool, a model-independent program that allows automatic calibration. The Dynamically Dimensioned Search (DDS) algorithm was employed to calibrate the several SHETRAN parameters and was suited for the computationally demanding calibration of the distributed hydrological modelling undertaken (Tolson and Shoemaker, 2007). Appendix E, Figure E-1 to 4 shows a comparison of observed and simulated discharge of each nested catchment.

The MuDRain software package, a multivariate disaggregation tool, was used to disaggregate the rainfall input data into finer resolutions (Koutsoyiannis et al., 2003). The disaggregation comprised two steps: first, the generation of hourly rainfall values without reference to daily totals and, second, the application of an adjustment procedure to preserve the daily totals observed. The adjustment corrected the hourly values generated in such a way that their daily totals over the entire duration of disaggregation were equal to the observed daily totals without affecting the first and second-order properties of the stochastic process. The hourly rain gauge rainfall data were later summarised into 2 and 3 hourly resolutions at each rain gauge station to reproduce rainfall dynamics with variable time intervals in the study area.

The simulated flow outputs were analysed using a similar analysis procedure applied in the scaling analysis of observed peak discharge. The 2015 flood events were simulated in all six nested catchments using merged satellite and rain gauge rainfall as input rainfall.

Each calibrated nested hydrological model was forced with rainfall scenarios with variable time intervals of 1, 2 and 3 hours while keeping the remaining rainfall properties constant. When rainfall intensity decreased as the time duration increased or when it increased as the time duration decreased, the rainfall volume remained constant. The simulation experiments examined the impact of rainfall resolutions only while other variables are maintained fixed at their respective values (Ayalew et al., 2014). In all simulation scenarios, it was assumed that surface runoff exists only with negligible evaporation, constant subsurface flow and rapid runoff into the river channel (Mandapaka et al., 2009).

#### **4.5.5 Scaling relationships**

Peak discharge scaling relationships were established for each of the selected storms that produced peak flows in all six sub-catchments. For each independent peak flow, a statistical power regression model was used to explore the scaling relationship between peak discharge and drainage areas. The log-log relationships for all individual peak flows were established where both peak discharge and drainage areas were plotted in 2 dimensions, with drainage areas on the x-axis and peak discharge on the y-axis. When peak discharge changes with the power of drainage area, a log-log graph displays the relationship as a straight line and each relationship result in the  $r^2$  value and a log-log equation with two scaling parameters, the exponent ( $\theta$ ) and intercept ( $\alpha$ ) values.

The following procedure was adopted to approximate the expected value of peak floods across the different spatial scales of the catchment area: 1) Determine the sub-catchment drainage area where rainfall-runoff occurred, 2) Identify peak discharge from the rainfall-runoff events 3) Use the power-law equation to establish scaling relationships between peak discharge and the drainage areas 4) The exponents ( $\theta$ ) and intercepts ( $\alpha$ ) for each rainfall event were derived from the power-law equation.

### **Peak scaling analysis**

The relationships between drainage areas, peak flow, intercept ( $\alpha$ ), and exponent ( $\theta$ ) values were investigated further to infer scaling relationships and to determine the mean scaling parameter values for the La Sierra catchment. The analysis also identified the overall log-log relationships of the nested catchment drainage areas for all 59 rainfall-runoff events, intercept ( $\alpha$ ), and exponent ( $\theta$ ) values grouped to obtain the overall scaling parameters. The analysis looked at the influence of rainfall variability by establishing the relationship between peak discharge and drainage areas and group the scaling data by (a) seasons; (b) storm type (convective or synoptic), (c) storm direction and (d) storm location. Overage scaling parameters were then calculated for each group (Wilkinson and Bathurst, 2018).

#### **4.5.6 Flood peak prediction framework**

A flood prediction framework was employed to assess if peak discharge could be estimated in the study area using the estimated scaling parameters and other measurable catchment processes. The estimated exponent ( $\theta$ ) and intercept ( $\alpha$ ) values (flood parameters) for each nested catchment drainage area were used in the power-law equation (Equation 4.24) to predict the expected magnitude peak discharge at any location in the study area:

$$Q_P = \alpha A^\theta$$

**Equation 4.24**

Where:  $Q_p$  is peak discharge,  $A$  is drainage area,  $\theta$  is the scaling exponent (slope of the best fit line) and  $\alpha$  is the scaling intercept parameter.

Pearson's correlation analysis was used to assess the intensity and direction of linear relationships between catchment properties and processes with scaling parameters using the IBM SPSS Statistics for Windows, Version 25. A data table was created for Pearson's correlation analysis (Table 4.7).

**Table 4:7 The primary variables used in Pearson's Correlation analysis**

<b>Scaling parameters</b>	<b>Peak discharge from nested catchments</b>	<b>Rainfall properties</b>	<b>Other catchment properties</b>
Scaling intercepts from the observed discharge	Observed Peak discharge	Accumulated rainfall amounts (1, 2. and 3 days)	River levels
Scaling exponents ( $\theta$ ) from observed discharge	Simulated Peak discharge (1hr, 2hr and 3hr)	Rainfall duration	Surface soil moisture
Simulated intercepts ( $\alpha$ ) from simulated flows		Rainfall direction	
Scaling exponents ( $\theta$ ) from simulated flows		Rainfall location	

The linear regression model was employed to analyse individual flood scaling parameter dependence on characteristics of rainfall variability, river levels, moisture and peak discharge (Statistics, 2018). The log-log relationships between flood scaling parameters were regressed against the natural logarithms of peak discharge and rainfall properties (Ayalew et al., 2015). The analysis allowed the description of rainfall variability in terms of rainfall accumulation, direction, location and time scales.

The regression model established the relationship between scaling parameters (as dependent variables) and rainfall variability, and other measurable catchment properties and processes. Also, the linear regression equations generated were used to estimate the catchment scaling parameters, and these were used with scaling equations to predict the expected peak flow as a function of the various nested catchment drainage areas. A multilinear regression analysis was also used to test the interdependence of flood scaling parameters on rainfall properties and peak discharge and other measurable catchment properties and processes.

The following procedure was used to establish significant relationships that can estimate exponents ( $\theta$ ) and intercepts ( $\alpha$ ) for the catchment.

1. Use the power-law equation (Equation 4.24) to establish the scaling relationships between peak discharge from individual storms and drainage areas of nested catchments to derive estimates of exponent ( $\theta$ ) and intercept ( $\alpha$ ) values.
2. Use the linear regression model to find the relationships between exponents ( $\theta$ ) and intercepts ( $\alpha$ ) with individual properties of rainfall variability, peak discharge and other catchment measurable catchment properties. A series of diagnostic tests were used to confirm the validity of the regression models developed using the goodness fit  $R^2$ , the normal distribution of errors, the t-statistic, collinearity statistics and the Dublin Watson and Shapiro-Wilk tests on the residuals (Ayalew et al., 2015).
3. The best linear regression model presented in item 2 above was used to estimate the corresponding scaling exponent ( $\theta$ ) and the scaling intercept ( $\alpha$ ) for various scales of the nested catchments.
4. Application of the power-law equation (Equation 4.24) and the estimated exponents ( $\theta$ ) and intercepts ( $\alpha$ ) from number 3 to estimate peak discharge across nested catchment drainage areas.

#### **4.5.7 Measurement of uncertainty**

Prediction and confidence intervals were used to show the uncertainty in the estimated flood scaling parameters and peak discharge values. The prediction interval was taken as the range likely to include the mean value of the estimated flood parameter and peak discharge values given the specific values of independent variables. On the other hand, confidence interval shows the most well-known range of values, showing the natural variability bounds of the estimated values and thus express uncertainty in the estimates. The uncertainty was presented using 5% and 95% confidence interval bounds, which are the limits of the range of the individual discharge and parameter

values frequently vary (NIST, 2012; Wadsworth H.M, 1990). Prediction intervals were calculated using equation 4.25.

$$\hat{y} \pm (t_{\alpha/2})s\sqrt{1 + \frac{1}{n} + \frac{(x_p - \bar{x})^2}{SS_{xx}}}$$

**Equation 4.25**

Confidence intervals deal with regression line estimation accuracy and help to compare the accuracy of various estimates. The interval establishes the natural variability bounds for flood parameter and peak discharge estimates. The intervals were calculated using a polynomial and power regression equation 4.26 below (Helsel and Hirsch, 2002).

$$\Delta X_i = Y_i \pm t_{\alpha} SE \sqrt{\frac{1}{DF} + \frac{(X_i - X_m)^2}{SS_{xx}}}$$

**Equation 4.26**

Where  $\Delta X_i$  is the confidence interval of the predictor,  $Y_i$  is the response variable,  $t_{\alpha}$  is the t-value at  $\alpha = 0.05$  significance level, SE is the standard error, DF is the degrees of freedom,  $X_m$  is the mean of X and  $SS_{xx}$  is the sum of squared differences.

#### **4.5.8 Model selection**

Model selection was made using the XLSTAT version 2019.3.2 software, which fit data with a variety of applications, including selecting the best models to predict dependent variables as predictor variables. The software enabled the selection of 2 best statistical models that can predict flood exponents ( $\theta$ ) and the other for predicting intercepts ( $\alpha$ ) parameters for the catchment based on the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). The selection was made to avoid model selection that would over fit the data and not perform well in new contexts.

The AIC and BIC metrics are objective measures for model estimation error and are used for comparing models and for selecting the optimal model (Ding et al., 2018). The primary tenant of AIC is to penalise the integration of more variables in a model and introduces a penalty to increase of errors in the event of additional values (Heinze

et al., 2018). The BIC is an AIC version with a higher penalty when additional variables are made. The lower the values of the metrics, the better the model that is desirable when comparing several candidate models, identifying models that are better for prediction (Shmueli et al., 2010; Brewer and Cooksley, 2016). Thus, it was feasible to determine the relative strength of one model relative to another based on the BIC and AIC values. The best models selected were then used to estimate flood parameters and for the estimation of discharge using the power-law equation.

#### **4.5.9 Estimation of parameters and peak discharge**

A framework for flood parameter prediction comprising the log-linear relationship between the intercept ( $\alpha$ ) and the exponent ( $\theta$ ) was used to estimate the scaling exponent ( $\theta$ ) values, and the log-log relationship between the intercept ( $\alpha$ ) parameter and the peak discharge observed from the smallest nested catchments was used to estimate the intercept ( $\alpha$ ) values for La Sierra catchment. The following three steps used by Ayalew et al. (2014) were applied to predict the expected value of peak discharge in the study area (1) a regression model was employed to calculate the scaling intercept ( $\alpha$ ) value from peak discharge observed from the smallest nested catchment. 2) A regression model was applied to estimate the corresponding scaling exponent ( $\theta$ ) from the natural logarithm of the scaling intercept ( $\alpha$ ) estimated in step one. Finally, 3) the estimated flood parameters (exponents ( $\theta$ ) and intercepts ( $\alpha$ )) were used to determine the expected magnitude value of peak discharge across spatial scales using the flood scaling equation (Equation 4.24). The steps presented provided an estimate of the expected value of peak discharges that exhibit variability around their expected value. The uncertainty (natural variability) of the predicted peak discharges was quantified using a Helsel and Hirsch, (2002) method for quantifying uncertainty.



## **4.6 Uncertainty propagation and analysis**

This section presents the method adopted to study rainfall and parameter uncertainty propagation through a coupled hydrological and inundation model to assess the effects on flood extents in the La Sierra catchment, downstream of the Las Gaviotas gauging station. A combined model chain comprising an ensemble rainfall model rainfall-runoff model and an inundation model is provided. Also, presented in this section is a description of each model in the cascading chain, followed by an overview of each model setup process, the hydraulic model's sensitivity and calibration. Presented is a method of estimating uncertainty propagation through the coupled model chain based on statistical methods to enable quantification of uncertainty in flood volumes and extents due to the uncertain rainfall input and variation of parameter values.

### **4.6.1 The coupled or cascading mode setup**

The coupled or cascading model set up comprised three levels: (a) the meteorological level b) the hydrological model and c) the hydraulic or inundation model linked together (Rodríguez-Rincón 2015; Tanaka et al., 2018). It linked the models in a nested modelling approach where outputs from one model are fed into the other model after the first model is enforced with rainfall input data (Di Baldassarre et al., 2010). The Monte Carlo simulator represented the first model at the meteorological level for generating rainfall ensemble data.

A hydrological model represented the second model; the SHETRAN distributed hydrological model enforced with rainfall ensembles generated by the first model (Loveridge et al., 2013). The model allowed research into how uncertainties originating from rainfall are propagated and interact with catchment level parameters (Tanaka et al., 2018). The last phase composed of the hydraulic model; the Flood Modeller inundation model enforced with each discharge assemble from the SHETRAN hydrological model. The Flood Modeller generated flood inundation extents for each input flow ensemble member (Rodríguez-Rincón 2015). The three coupled or cascading models were run to produce scenarios based on fixed and variable parameter values and generated several ensembles within 5% and 95% confidence limits.

#### **4.6.2 The meteorological level**

The GoldSim Monte Carlo simulation software was used for uncertainty propagation and the generation of ensemble rainfall time series (Loveridge et al., 2013; GoldSim Technology Group 2017; Zarzar et al., 2018). The software used a Monte Carlo simulation technique to produce random ensemble members required for one year (2015) covering variable rainfall circumstances (Rodríguez-Rincón 2015). The rainfall ensemble members were selected within the 5% and 95% confidence limits and enforced into a hydrological model to generate flow ensemble outputs that acted as inputs into a hydraulic model.

##### ***Rainfall ensemble generation***

The Monte Carlo simulation technique was used to generate separate and independent members of rainfall ensembles, each showing possible rainfall variability in the La Sierra catchment area (Rubinstein and Kroese, 2007). The probability density functions (PDFs) were defined for each station in the catchment area and each rainfall dataset was associated with a specific PDF, each of which was fitted with the most appropriate probability distribution function to derive the statistical distribution parameters.

The EasyFit software was used to select the most fitting probability distributions for rainfall data in the study area and to identify PDF parameters (Schittkowski, K., 2002). Several PDFs identified by the application were fitted using the MLE (Maximum Likelihood Estimates) method and calculated the  $p$ -values based on the Kolmogorov-Smirnov and Chi-Squared goodness-of-fit statistics. This is because the rainfall variables in the model input are random rather than fixed (Mamnoon and Rahman, 2019).

The Monte Carlo simulation technique was applied following these steps:

- 1 Identification and description of input rainfall variables and parameters from Probability Distributions Functions.
- 2 Generate random values for each input variable from their respective probability distributions.

- 3 Run the model with each set of the input variables produced and generate a model output for the set of parameters specified.
- 4 Analyse the distribution of model outputs to determine the cumulative distribution function and other statistical characteristics (for example, the mean and standard deviations).

The simulation technique allowed quantification of the uncertainty in rainfall input data resulting from its spatial and temporal variability expressed in the form of confidence limits on probable rainfall estimates. It was possible to examine the effects of many combinations of rainfall input variability on the magnitude of flood volumes and extents (Mamnoon and Rahman, 2019).

### ***Selection of ensemble members from non-symmetrical distributions***

Several rainfall datasets analysed maintained a consistent right or positively skewed distribution, where the mean was found to be larger than the median and the data containing very few large values. The analysis assumed a normal distribution of uncertainties from these non-symmetrical distributions. To make the distributions normal, non-symmetrical distributions were transformed by taking the square root of the observation logarithm using the Box-Cox method. A Box-Cox power transformation was used to make the data normal, where all data was converted into a particular exponent, shown by a Lambda value that makes the variance more constant and normalises the rainfall data. The Box-Cox normality plot which approximately normalises the data is defined as:

$$T(Y) = (Y^\lambda - 1)/\lambda$$

**Equation 4.27**

Where Y is the dependent variable, and  $\lambda$  is the transformation parameter. For  $\lambda = 0$ , the natural log of the data was taken instead of using the above formula.

The transformation of the data set yielded the data to follow approximately the normal distribution, which increased the applicability and usefulness of the statistical techniques based on the normality assumption. Computing the correlation coefficient of a normal probability plot helped define a measure of the normality of the resulting transformation (NIST, 2012). Twelve PDFs were examined in selecting the best-fit

probability distribution for the annual rainfall time series in La Sierra catchment (Table 4.8).

**Table 4:8 Details of selected rainfall stations in La Sierra catchment.**

Station name	Period of data	Record length (years)	Longitude	Latitude
Dos Montes	1973 - 2016	43	92.8W	18.0W
Plaza del Rosario	1983 - 2016	33	92.0W	17.9N
La lagarte	1986 - 2016	30	92.6W	17.4N
Las Mercedes	1984 - 2015	31	92.6W	17.4N
Tequita	1970 - 2016	46	92.8W	17.7N
Dos Patria	1960 - 2016	56	92.8W	17.6N
Teapa	1960 - 2015	55	93.0W	17.0W
Tapijula	1960 - 2015	55	93.2W	17.2W
Simojovel	1965 - 2016	51	92.3W	17.2N
Yayalon	1942 - 2016	74	92.3W	17.0N
Sitala	1984 - 2015	31	92.8W	17.0W
Oxolotan	1974 - 2016	42	92.8W	17.4N

### ***The goodness-of-fit tests***

The Goodness of fitness tests were conducted to select the best PDFs that could describe possible rainfall values, and probabilities for each rainfall station would vary within the La Sierra catchment area. The fitness test was used to determine whether a distribution fits the rainfall time series data set adequately and calculate statistics of goodness-of-fit that helps to rank the fitted distributions according to the quality of fit concerning the observed data (Mamnoon and Rahman, 2019). The best fit distribution was selected from a list of candidate probability distributions for the year 2015 rainfall time series data from twelve rain gauge stations in the study area.

For each PDF, the following three fitness measures at a significance level of 0.05 were used: (i) Kolmogorov-Smirnov; (ii) Anderson-Darling; and (iii) Chi-squared (Haddad and Rahman 2011). The selected twelve probability distributions were ranked on a scale of 1 to 5 for all three tests separately, with rank one representing the best fit distribution, and rank 2 being the second-best, and so on. The final selection was made based on cumulative test scores obtained from adding all three fitness measures, a distribution that fit the highest numbers of selected stations. A maximum

score of 5 was given to rank one probability distribution, and 4 to rank two distributions, and so forth (Mamnoon and Rahman, 2019).

To identify the best-fit probability distribution overall, a relative scoring method was adopted based on the results given by the three goodness-of-fit tests. To determine the best fit probability distribution that corresponds to the highest number of stations selected, a relative scoring method based on the results of the three fitness tests was adopted (Mamnoon and Rahman, 2019).

### ***Monte Carlo simulation***

The Monte Carlo simulation was used to sample each PDF for each rainfall gauge station data and produced a defined number of possible ensemble members for each gauge with confidence limits showed (Rubinstein and Kroese, 2007). Historical rainfall time series were fed into the Monte Carlo simulation module as external time histories with a known distribution before the simulation (GoldSim Technology Group 2017). The technique computed its output repeatedly, with input values that were randomly sampled from pdf parameters. A random sampling method of PDF parameters was performed to incorporate uncertainty in the rainfall input data that was then propagated into the flood inundation model (Kalyanapu et al., 2012). However, the process ignores uncertainty in model parameters and model structure, but these were later included in the analysis.

### **4.6.3 The SHETRAN hydrological model and domain**

The SHETRAN was applied because of its distributed nature of rainfall-runoff generation through grid-to-grid routing and gridded representation of catchment properties and processes. The detailed representation of parameter variability and the ability of the model to respond to differences in rainfall variability made the model suitable to investigation parameter uncertainty and effect on flow outputs. The hydrological model was applied to La Sierra catchment to investigate the propagation of rainfall and parameter uncertainties and effect on discharge.

La Sierra catchment is one of the wettest catchments in Mexico with an annual rainfall total of approximately 4 000 mm, with annual averages of 2 550 mm occurring all year. However, the river flows in the catchment have minimal influence from regulatory

artificial river control structures like dams, barriers and river diversions compared to the main Grijalva River. As a result, most flood-prone areas of the La Sierra floodplains, especially in the heavily populated areas of Villahermosa, often experience floods with severe impacts on life and property. The heavily populated areas were seriously affected during the 2007, 2008 and 2009 flood events and the flood-prone areas include permanent and informal settlements.

#### **4.6.4 The Flood Modeller 2D model**

The Flood Modeller application was used to simulate the October 2015 flood event and was able to provide an approximation of flood depths, velocity and extents. The hydraulic model was used at a relatively broad scale (15-metre resolution) to assess the impact of uncertainties on the inundation of flooding extents rapidly through a coupled hydrological/hydraulic model. A detailed description of the Flood Modeller application is given in section 4.4. The Flood Modeller simulations were run on a 64-bit operating system computer with processor type Intel (R) Core (TM) i7–6700HQ CPU operating at a speed of up to 3.41 GHz. The 16.0 GB RAM and the NVIDIA Quadro K2200 graphics card with 640 CUDA streaming processors at 4 GB of GDDR5 GPU memory with fast bandwidth that enables the creation of large and complex models. The Flood Modeller was run using a spatial resolution of 15 m.

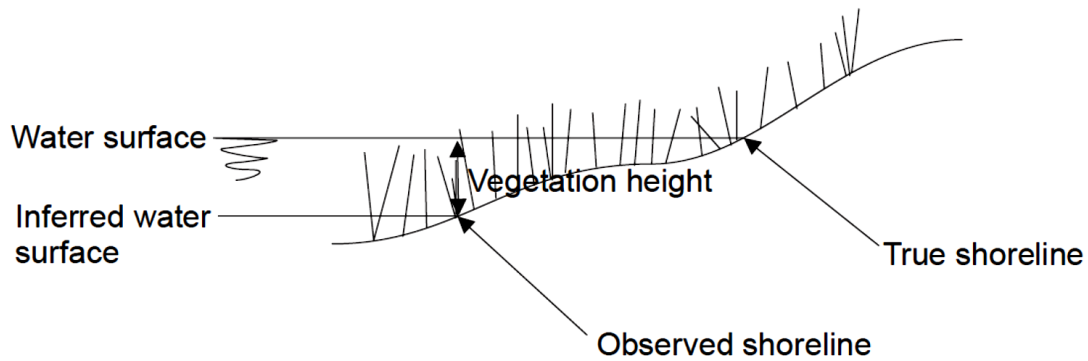
#### ***Model sensitivity and calibration***

The Flood Modeller's sensitivity to Manning coefficient values for the channel and floodplain was investigated to calibrate the flood model. The model was run using several similar simulations based on variations of the coefficient values of the Manning varying from 0.04 to 0.055 m, describing a winding river with ponds and reefs in the lower levels and inefficient slopes and sections (Chow, 2008). Within this range, the roughness coefficient values were adjusted to minimise the disparity between the model simulations and the observed flood outlines. The variation of the values was applied in incremental steps of 0.001 until the best fit value matched the flood outlines. To achieve the best fit, the simulated and observed flood extents when compared should have a roughness value that minimised the differences between the 2.

Calibration of the inundation model comprised finding the optimal Manning coefficient roughness values for the channel and floodplain since these have considerable influence on simulated flood extents (Van Wesemael et al., 2016; 2011). Ideally, the friction values should reflect the spatial variability present in the channel and floodplains and be estimated explicitly from physical or biological variables present in the catchment area (Mason et al., 2006). However, as an alternative to the representation of spatial variability, a conventional technique of using two different global static coefficients, one for channel and the other for floodplain was considered. Satellite images, land use maps and aerial photographs were used to obtain the spatially distributed friction values that reflected the floodplain and for the channel roughness values for the study area (Mason et al., 2006).

The traditional technique of calibrating the output of inundation models was to match simulated flood extents with observed extents from satellite images (Stephens et al., 2014). The evaluation of the efficiency of flood inundation models has often used satellite information to measure how well the output produced matches realistic flood extents (Lim and Lee, 2018). However, errors and uncertainties exist in the measurement of flood extents extracted from remote sensing images caused by inappropriate wavelength selection and application of incorrect algorithms (Stephens et al., 2014). In particular, the delineation of flood extents is complicated by emerging vegetation and buildings, water surface roughing and water surface rain distortions that cause multiple reflections that cause dry land on the edges of flooded areas to appear similar to flooded areas (Mason et al., 2009).

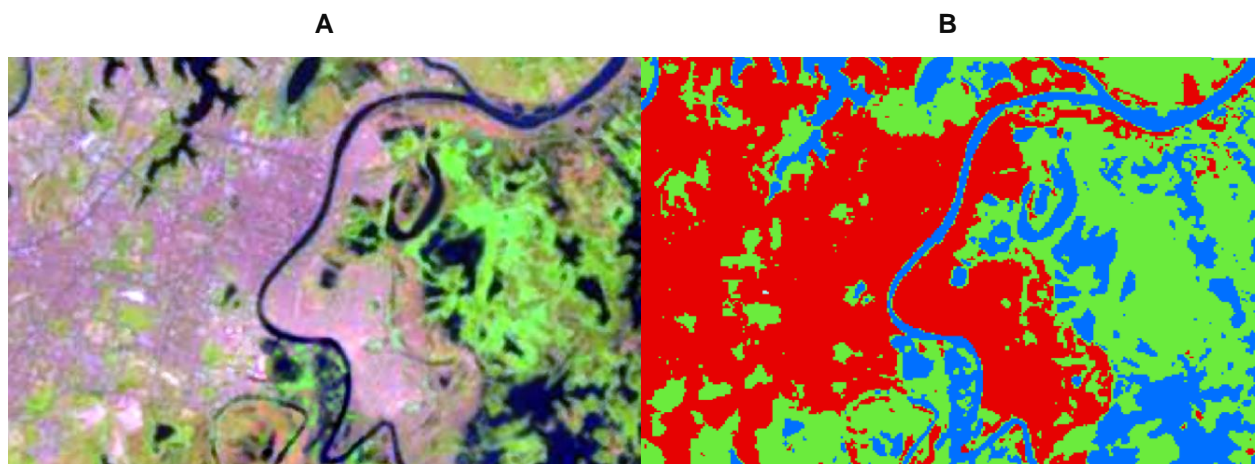
The waterline was shifted towards the river by the effects of emergent vegetation, causing an underestimation of water levels (Fig. 4.24). The levels were corrected by adding the vegetation depth at the waterline found in the satellite imagery (Horritt et al., 2003). Using combined datasets comprising gauged water elevations, contour-based LiDAR DEM and remote sensing images were used to generate reasonably observed flood extents (Stephens et al., 2014; Zhang et al., 2018).



**Figure 4.24: The effect of vegetation on flood outlines where the water level is underestimated (Horritt et al., 2003).**

***The extraction of observed flood extent outlines***

The extraction of flood outlines was performed using the ISODATA (Iterative Self-Organising Data Analysis) unsupervised image classification method (Jung et al., 2014). Landsat 5 Thematic Mapper (TM) images were used to show the level of flooding in the study area during the flood event on 27 October 2015. The ISODATA was applied to the Landsat spectral bands 1, 4 and 7, which were classified into 20 classes (Song et al., 2001). The spectral difference between water bodies and other surrounding land became more apparent when the spectral bands combine into Bands 1, 4 and 7 (blue depicted by Band 1, near-infrared by Band 4, short wave infrared by Band 7) (Jung et al., 2014) Figure 4.25.



**Figure 4.25: Comparison of the band variations in the downstream of Las Gaviota gauging station on 27 October 2015. (a) The original Landsat imagery on 27 October 2015. (b) The combination of Bands 1, 4 and 7 (blue: Band 1, near-infrared: Band 4, short wave infrared: Band 7).**



However, the boundary between dry and flooded land appeared vague and discontinuous with some having vague flood boundary outlines. The accuracy of the Landsat satellite image was improved using the combined use of over one dataset, and these included observed gauged water levels and contour-based LiDAR DEM data (Stephens et al., 2014). The observed water levels with discharge data and related rating curves across the entire year of 2015, covering the 17 October flood event were provided by CONANGUA.

To obtain contour-based LiDAR DEM data of the study area, LIDAR data was converted to contour lines at 0.5 m interval using GIS spatial tools. The 1D flood calculator tool in Flood Modeller was used to calculate a static flood extent from the location of the observed/gauged water levels. A constant water level layer from this static flood extent was calculated to represent a horizontal water surface that spread across the entire area with defined depth, water extents and flood shoreline (Stephens et al., 2012).

The static flood extent layer was overlaid onto the contour-based LiDAR DEM layer, and the flooded margin that matched with a contour line on the LiDAR topographic data defined the observed flood extents margin (Stephens et al., 2012). The combined dataset technique was used to reduce errors and uncertainties in the flood delineation process. It enhanced the accuracy of flood extent delineation based on the use of combined and complementary datasets rather than the more frequent use of satellite imagery only to measure the degree to which the observed and simulated flood extents agree.

### ***The hydraulic model set up and model domain***

The model was set up to simulate flood extents along an 8 km reach between Gaviotas and Povernir gauging stations. The set up comprised upscaling the 5 m resolution LiDAR DTM to 15 m so that the 2D modelling was within the 10 000 cells, the maximum number of cells permissible for the free Flood Modeller application. The Flood Modeller Resample tool was used to generate an upscale DTM that calculates an average value based on the source grid cells. As a result, the up-scaling process resulted in spatial averaging of values in each grid and removed some low-resolution topographic features such as embankments.

All flood embankments were presented in a separate layer in which values were kept in the model. The model set up also required, separate layers for buildings, vegetation and features present on the ground surface used for determining the distributed floodplain roughness coefficient values.

Since the LiDAR DTM did not show riverbed elevations, but only the water surface elevation, it was merged with surveyed bathymetry data to obtain an approximation of the real channel depths. Several water depths between 2.50 m and 4 m were tested during model calibration and in combination with the main channel roughness coefficient values. Also typical in the La Sierra River, there is a deposition of sediments that leads to siltation during high flows. Riverbed dredging is commonly practised in the La Sierra River that results in variable channel depths. In this way, channel depth served as one of the calibration parameters that represented a variable river depth.

#### **4.6.5 Model scenarios**

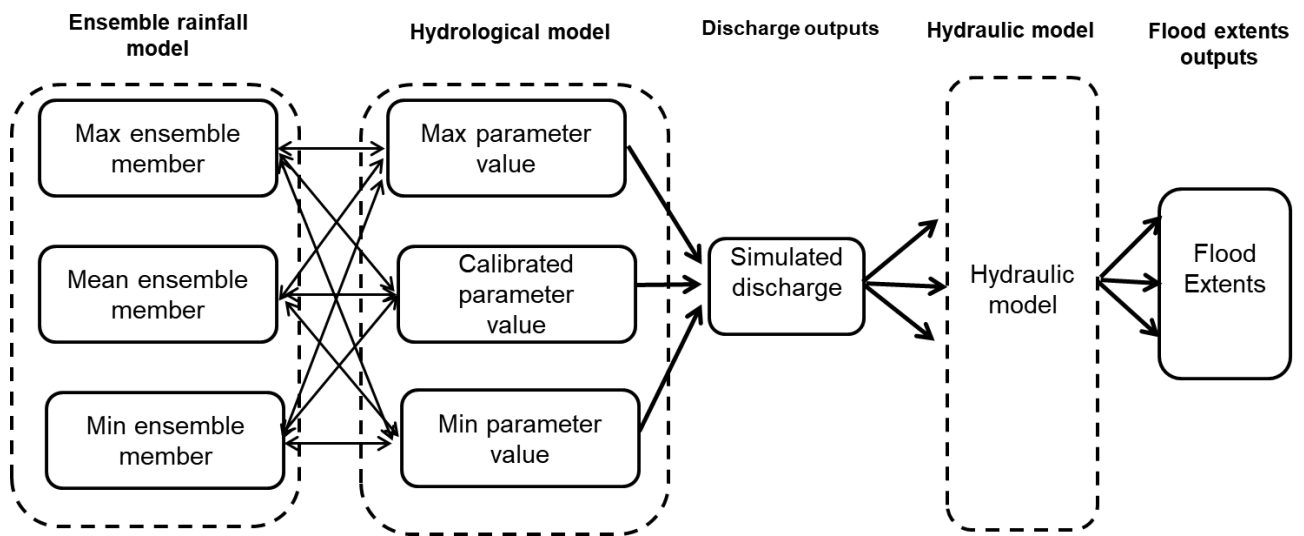
The ensemble simulations proceeded under two scenarios: (i) the calibrated 59 model runs based on fixed calibrated parameters to investigate the variability and propagation of rainfall uncertainties, and (ii) the same 59 model runs based on variable or individual parameter value changes this time to investigate variability and propagation of parameter uncertainties.

##### ***Scenarios without changes to the parameters***

Initially, the hydrological model was run enforced with each of the 59 rainfall ensembles based on fixed calibrated parameter values, (which were constant values, maintaining values obtained during calibration) to investigate the variability and propagation of rainfall uncertainties. In this way, the model chain was used to generate control simulations of ensemble discharge and flood inundation maps and enabling an understanding of rainfall uncertainties and the behaviour of the model in simulating ensemble discharge output (Zarzar et al., 2018).

### Scenarios based on the variable parameter model runs

Also, the model was run to generate 59 flow ensemble outputs based on parameter changes (Zarzar et al., 2018). One parameter at the time was selected for each set of 59 model runs and its value changed between maximum and minimum values within acceptable literature values for each separate set of 59 model runs (Figure 4.26). As a result, 59 flood inundation outline maps were obtained by enforcing the model with rainfall ensemble input members based on the selected parameter value.



**Figure 4.26: Rainfall and parameter uncertainty propagation through a coupled hydrological and inundation model based on enforcing the model with multiple ensemble members run with varying parameter values.**

#### 4.6.6 Generation of parameter values

The most sensitive parameters that most affect the discharge simulations in the study area (selected in section 4.4.2) were used for uncertainty estimation of parameters. The parameters selected were, the Strickler overland flow resistance coefficient (K), Saturated water content ( $\theta_s$ ), AET/PET and the Saturated conductivity (Ks). The parameter values were selected by defining parameter ranges for each of the critical land-use cover and soil types in the catchment area. The range of parameter value used for the uncertainty analysis was taken from the reference manual of the SHETRAN hydrological model. The model parameters, along with the parameter ranges considered in uncertainty analysis, are presented in Table 4.9. The major land cover types considered in the uncertainty analysis were deciduous forest, grass,

arable, and evergreen forests. Primary soil types considered in the uncertainty analysis included Medium soil (18%, clay, 35% and, 15% sand or 18%, clay and 15%, sand, 65%), Coarse soil (18% clay and 65%) and highly productive aquifer through cracks soils.

**Table 4:9 The upper, lower and calibrated parameter values for each land cover and soil types used during uncertainty analysis and scenarios.**

Parameters	Description	Range	Low	High	Calibrated	Unit
K <sub>1</sub>	Strickler overland (Arable)		0.5	9.9	9.48	m <sup>1/3</sup> s <sup>-1</sup>
K <sub>2</sub>	Strickler overland (Deciduous)	≥0.3 ≤ 9.9	0.1	7.8	6.70	m <sup>1/3</sup> s <sup>-1</sup>
K <sub>3</sub>	Strickler overland (Evergreen)		0.05	8.9	7.48	m <sup>1/3</sup> s <sup>-1</sup>
K <sub>s1</sub>	Saturated cond. (Coarse soil)		0.23	2.22	1.0	m/day
K <sub>s2</sub>	Saturated cond. (Highly)	≥0	0.11	1.19	0.20	m/day
K <sub>s3</sub>	Saturated cond. (Medium soil)		0.31	0.93	0.30	m/day
AET/PET <sub>FC1</sub>	AET/PET ratio (Arable)		0.2	1.9	0.35	(-)
AET/PET <sub>FC2</sub>	AET/PET ratio (Deciduous)	≥0.01≤1.99	0.4	1.6	0.70	(-)
AET/PET <sub>FC3</sub>	AET/PET ratio (Evergreen)		0.5	1.7	0.84	(-)
θ <sub>s1</sub>	Saturated water content (Coarse)		0.45	1.0	0.81	m <sup>3</sup> /m <sup>3</sup>
θ <sub>s2</sub>	Saturated water content (highly)	≥ 0.05 ≤1.0	0.42	0.9	0.30	m <sup>3</sup> /m <sup>3</sup>
θ <sub>s3</sub>	Saturated water content (Medium)		0.06	0.7	0.44	m <sup>3</sup> /m <sup>3</sup>

A pseudo-random number generator was used to randomly select parameters from the highest and lowest allowable values for each land cover type (Lewis and Orav, 1989). The Random Number Generator v1.4 application based on Mersenne Twister was used to generate random parameter values (Matsumoto and Nishimura 1998). The application generated parameter values by drawing samples from each selected parameter range (Figure 4.26 and 4.27). A list of random numbers was generated between the lower and upper parameter values for each land cover and soil types. The generated outputs were then sorted from highest to minimum, facilitating the selection of highest and lowest parameter values for use in model scenarios.

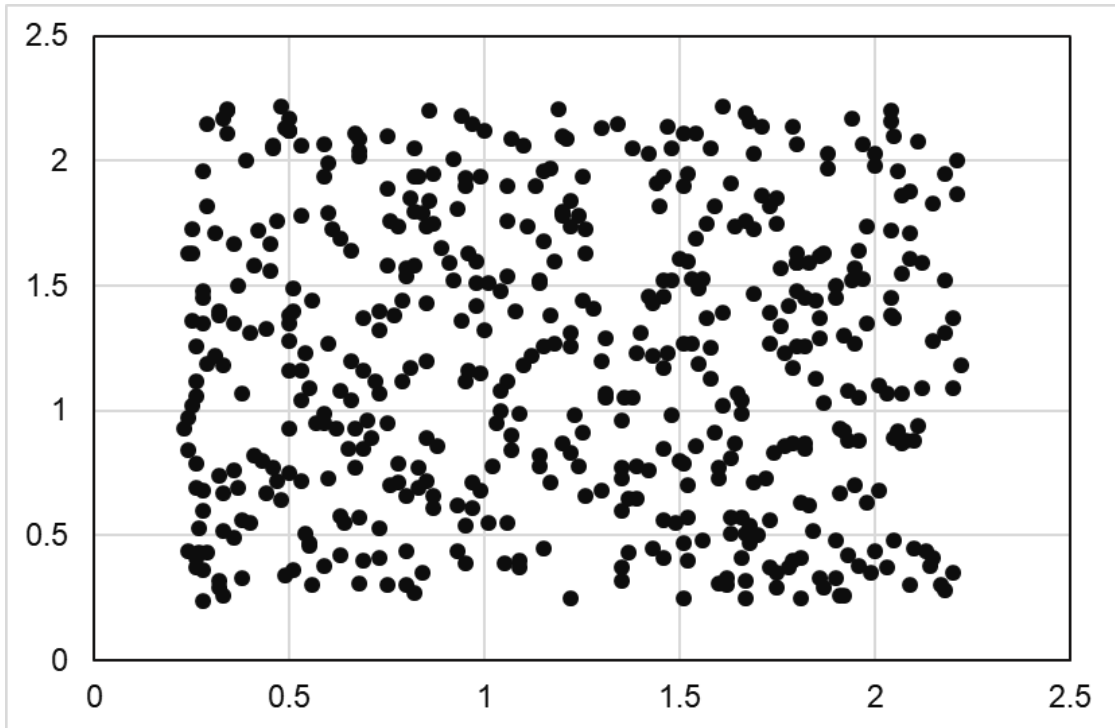


Figure 4.26: A random number scatter showing parameter values (500) randomly sampled from saturated conductivity parameter range (0.23 to 2.22) for coarse soil.

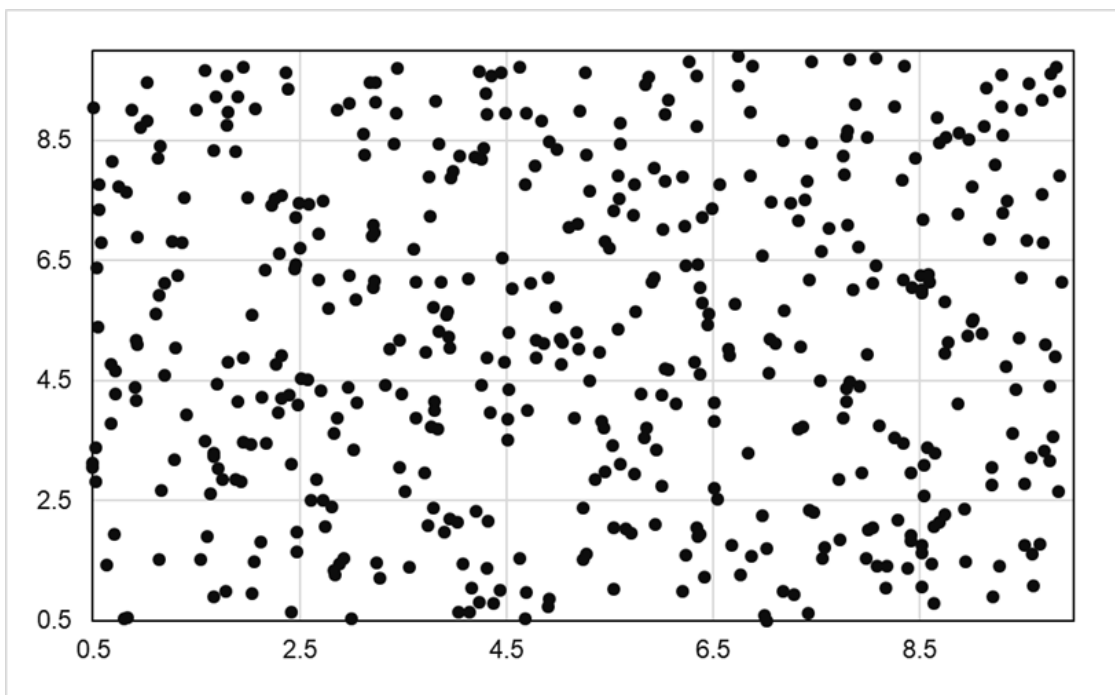


Figure 4.27: A random number scatter showing the parameter values (500) that were randomly sampled from the Strickler parameter range (0.5-9.9) for Arable land.

#### 4.6.7 Historic flood levels and extents data

From October 25 to 28, 2015, exceptional rainfall fell in the upper and middle Grijalva-Usumacinta River basin, generating runoff, which caused flooding in Tabasco and some neighbouring states. A system of low pressure resulted in several parts of the Grijalva catchment experiencing heavy rainfall and flooding since 26 October 2015. Late on the 25 October 2015, Tropical Cyclone Patricia formed in the eastern Pacific Ocean in southern Mexico and approached the coast of Michoacán and Colima States in Mexico as a hurricane. There are reports of over 254 mm of rainfall falling in 36 hours, causing flooding up to 2 and 75 cm deep and several areas were flooded. Although some minor flooding took place in Villahermosa city, severe rainfall and flooding took place in the nearby state of Quintana Roo, where a state of emergency was declared for its seven municipalities.

The flood levels and photographs from the 27<sup>th</sup> October 2015 flood were used to provide the required information for calibrating the hydraulic model (Stephens et al., 2012). In this flood event, Las Gaviotas flow and level gauging station recorded flow levels reaching 4.86 m above sea level and discharge reaching 744.75 m<sup>3</sup>/s (Table 4.10). The flood extents were visible from Landsat TM/ETM+ image, and these were used to represent the extent of the flood envelope used for model calibration.

**Table 4:10 Recorded flow and levels for Las Gaviotas and nearby gauging stations on 27th October 2015.**

Date	Level (metres above sea level)	Discharge (m <sup>3</sup> /s)	Gauging station
27/10/2015	3.9	822.09	El Povenir
27/10/2015	4.86	744.75	Las Gaviotas
27/10/2015	6.62	385.49	El Pueblo

#### 4.6.8 Uncertainty analysis

##### ***Statistical uncertainty quantification***

Statistical methods were used to quantify uncertainty in flood extents due to the uncertain rainfall input into the model (Roelofs et al., 2018). Rainfall input data were selected for analysis because it was considered the primary source of uncertainty in hydrological models that propagates into model outputs (Lobligeois et al., 2014). A more in-depth analysis of uncertainty required statistical analysis of individual input

and outputs variables for the model runs and to quantify uncertainty propagation across the modelling chain. The analysis was used to quantify the rainfall uncertainty at each level of the modelling chain defined by coupling meteorological, hydrological and hydraulic models.

Before model runs, rainfall uncertainty was represented using Probability rainfall Density Functions (PDF), a typical uncertainty characterisation and quantification approach. In this sense, statistical parameters were determined using a measure of central tendency theory (Loveridge et al., 2013). In this study, the standard deviation was used as a measure of uncertainty in rainfall input data. A small standard deviation showed that rainfall data was close to the mean often recognised as the expected value, whereas a high standard deviation suggested that the data would typically be distributed across a broader range of values. In principle, the small spread implies less uncertainty.

#### **4.6.9 Measures of fit**

##### ***Index of fit measure (F)***

Simulated flooding extents were compared to observed flood extent maps using an index of fit measure  $F$  (Komi et al., 2017; Di Baldassarre et al., 2010). The measure of fit,  $F$ , was given as:

$$F = \frac{A-B}{A+B+C}$$

**Equation 4.28**

Where:  $A$  is the wet area accurately estimated by the model (the observed inundated area),  $B$  is the inundated area estimated by the model (area estimated to be wet, but observed dry, that is, over-predicted) and  $C$  is the wet area not estimated by model (under-predicted) (Di Baldassarre et al., 2010; Komi et al., 2017). Each water depth map was converted to maximum flood extent polygon shapefiles using built-in Flood Modeller tools while the area of each flood extent polygon was calculated using ArcGIS geometry tools. Each flood extent map produced was compared to the observed flood extents, to obtain  $A$ ,  $B$  and  $C$ .

The  $F$  value ranges from  $-1$  to  $+1$ , where  $1$  is the best fit for good flood modelling results and enables comparison with observed outlines at different Manning coefficient value ranges (Horritt et al., 2007; Di Baldassarre et al., 2010; Jung and Merwade, 2015). In the measure of fit  $F$  equation, the term  $B$  is used to penalise if the model over-predict since satellite images show a composite of all flood types and related processes such as surface water flooding, backwater effects and groundwater flooding yet flood inundation models simulate only river discharge that overflows the river channel into flood plains (Schumann et al., 2009). The calibration process comprised comparing results from each model run to the observed flood extent map in terms of the measure of fit. After running several calibration simulations, a global coefficient value of  $0.04$  gave the best fit and was selected to represent Manning's coefficient roughness parameters for the reach studied.

### ***Coefficient of variation***

The coefficient of variation (CV) (Equation 4.29) is a measure of dispersion or variability of the probability of distribution about the mean of a population. It is also shown as a percentage of the ratio of the standard deviation to the mean (Yu et al., 2016).

$$CV = \frac{\sigma}{\mu} * 100\%$$

**Equation 4.29**

Where:  $\sigma$  and  $\mu$  are the standard deviation and the mean value respectively, of dispersion of a probability distribution. In this research, the coefficient of variance was used to assess the level of dispersion of 59 rainfall ensemble members around the ensemble mean. The higher the coefficient of variation, the higher the dispersion level around the mean and the measure is expressed as a percentage (Nakakita et al., 2016).



### ***Relative Interval Length (RIL)***

An adapted measure called Relative Interval Length (RIL) was used to estimate and summarise the uncertainty reported by the model (Jin et al., 2010). The RIL measures the resolution of the predictive distributions (Li et al., 2011). It is a standardised measure and is calculated as a function of lower and upper limits and median outlines of flood extent and useful for comparing flood inundation map scenarios (Ahmadisharaf et al., 2018). The RIL was computed using equation 4.30.

$$\mathbf{RIL} = \frac{X_U - X_L}{\dot{X}}$$

**Equation 4.30**

Where:  $X_L$  and  $X_U$  are the lower and upper confidence limits, respectively, and  $\dot{X}$  is the median value. The ARIL values should be as small as possible, meaning a smaller value represents better performance with less uncertainty (Jin et al., 2010; Li et al., 2011).

### ***Confidence interval and limits***

In this study, the 5% and 95% confidence intervals were considered as the lower and upper limits of the confidence interval with plausible ensemble member values in line with other flood inundation research where a similar measure was used to quantify uncertainty in inundation maps (Pappenberger et al., 2005; Loveridge et al., 2013; Ahmadisharaf et al., 2018). All ensemble input data discharge and flood extent outputs were classified into probability intervals between 5% and 95% confidence limits, which are the upper and lower ends of the rainfall ensemble values where their actual values lie with some certainty around an ensemble mean.

A confidence level is a probability that the uncertainties lie in the interval. Confidence intervals were used to describe ranges (5% and 95%) within which the average estimate occurs with defined probability. In this way, the 95% confidence interval has a 0.95 probability of containing the ensemble value; while 5% has a less frequent probability of 0.05 (Lane, 2015).

## 4.7 Summary

This section has presented the methodology to study rainfall and parameter uncertainty propagation through a coupled model chain to assess the effects on flood inundation extents in the La Sierra floodplain. The methodology comprises the meteorological model represented by the Monte Carlo simulator for generating rainfall ensembles. The model is coupled to a second model represented by the hydrological model comprising the SHETRAN hydrological model that is enforced with rainfall ensembles generated by the first model (Loveridge et al., 2013). The last is the inundation model enforced with flow ensembles produced by the hydrological model for a generation of flood inundation extents. The research explored how uncertainties originating in rainfall ensembles interacted with catchment level parameters and propagated to determine flood inundation extents. The estimation of uncertainty propagation through the coupled model chain was provided based on statistical methods that quantified uncertainty in flood extents due to uncertain in rainfall input and parameters.

## 5 RESULTS

This chapter present results of three main sections; (i) the calibration and validation of distributed hydrological modelling for the Grijalva catchment using a combination of satellite and in-situ datasets (section 5.1; Objective 1), (ii) the nested catchment modelling to investigate scaling relationships between peak discharge, nested catchment areas and rainfall variability in the La Sierra catchment (section 5.2; Objective 2) and (iii) the coupled hydrological-hydraulic modelling to investigate rainfall and parameter uncertainty, propagation and effects on predicted flood inundation extents (section 5.4; Objective 3).

### 5.1 Spatially distributed hydrological modelling

The development of a distributed hydrological model, which forms the foundation for this research, is presented in this section. The results are structured in the same order as outlined in the modelling protocol (Figure 4.1) comprising sensitivity analysis, calibration, validation, error quantification and correction.

#### 5.1.1 Rainfall error correction results

This section present results of rainfall interpolation, merging and error correction and the temporal and spatial rainfall error distribution in the Grijalva catchment. Also presented are calculation results of diverted discharge from the model domain, including the corrected simulated flows obtained and the percentage improvement results. Finally, the results of calibration and validation that show the performance of the SHETRAN hydrological model and its ability to simulate the magnitude of the observed discharge are presented.

#### *Geostatistical tool and analysis results*

Results of the Geostatistical exploration of rain gauge rainfall datasets in the study area show rainfall is mainly approximated by the Logistic, Normal and Gumbel Probability Distributions Functions. Generally, the Logistic distributions fit most rainfall datasets and were distinctly skewed to the right with heavier tails accommodating lower limits of rainfall values. The right tail of the histograms showed a relatively small

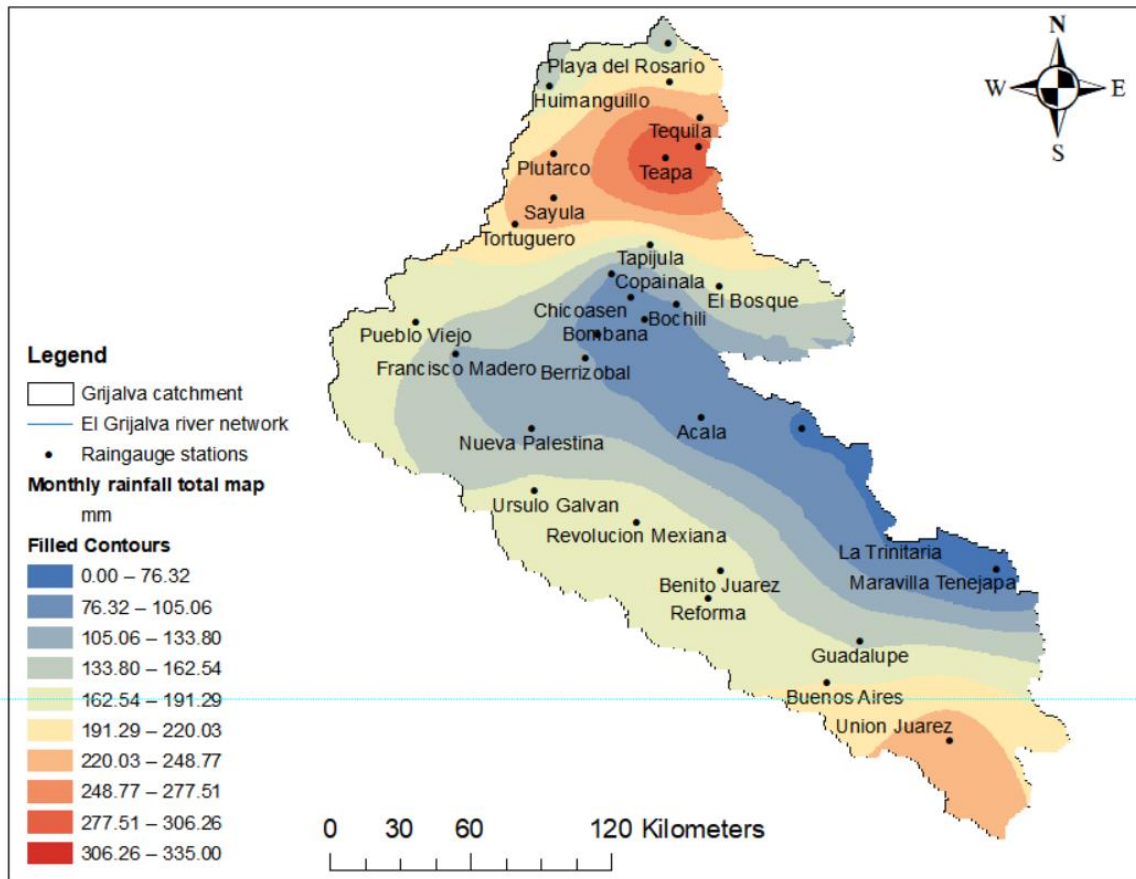
number of rain gauge rainfall values, however with large amounts of rainfall accumulations on the right. For several datasets analysed, rainfall maintained a consistent right or positively skewed distribution, where the mean was found to be larger than the median and the data containing very few large values. The distribution had values close to zero, and the data distribution skewed to the right. The fitting of the other distributions gradually decreased from Logistic, Normal to Gumbel Distribution functions. A Box-Cox power transformation was used to make the data normal, to make the variance more constant and normalise the rainfall data.

Also, cross-validations were performed to check the predictive capability of the Kriging model in assessing how well it will perform (section 4.2.4). Results show that the spread of points was close around the best fit (1:1) line (Fig. 4.17). Diagnostic statistics on the performance of the Kriging model prediction were given to further assess the model performance. The results showed the smallest Root Mean Square (2.44), a small Root Mean Square Standardised prediction error (1.42) which was the closest to 1. Also, the results had a small Mean Standardised Prediction Error (-0.13) and the lowest Average Standard Error (1.57). The results in the scatterplot for the predicted and observed data showed that the Kriging model performed well during the cross-validation process.

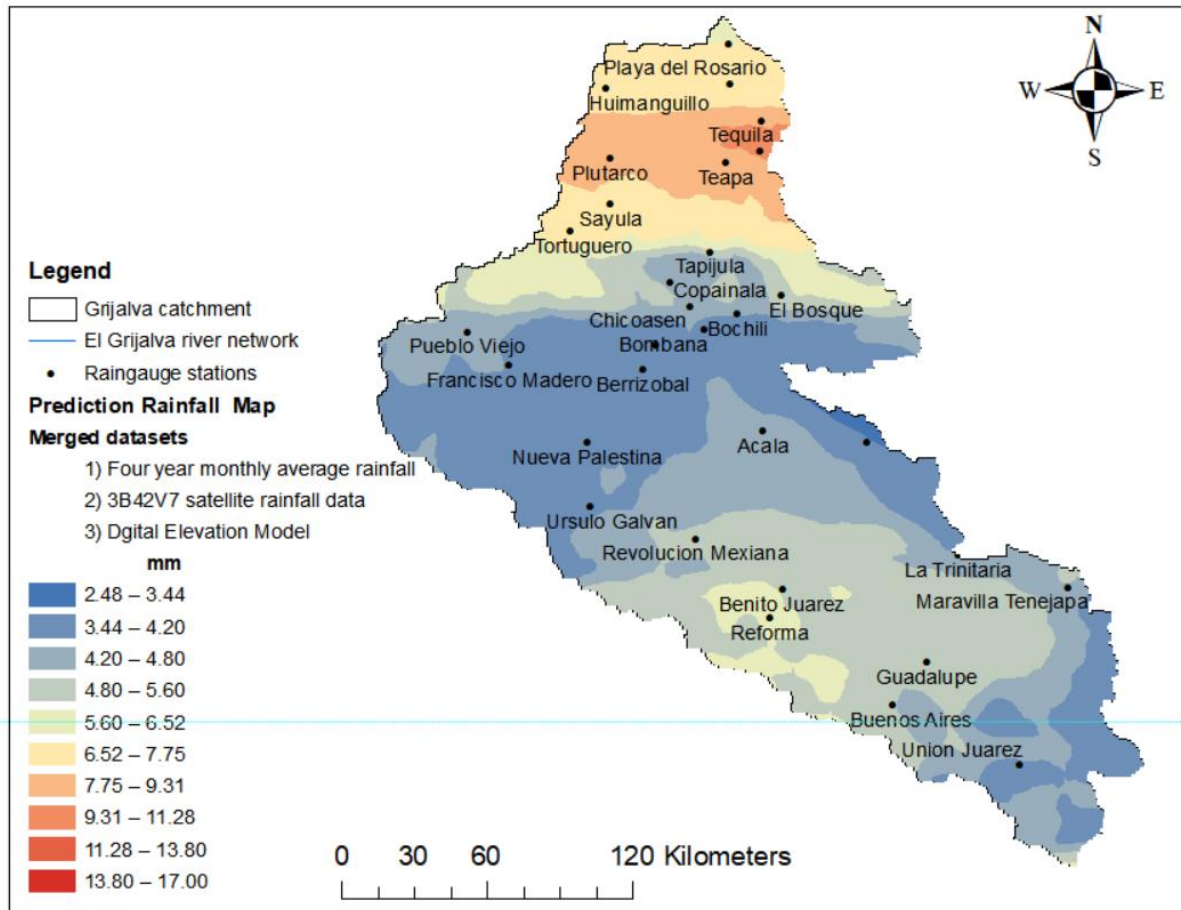
### ***Rainfall distribution***

Results of rainfall error correction are presented in this section, including the temporal and spatial rainfall error distributions in the Grijalva catchment. The results show that rainfall is highly spatially variable in the catchment and is primarily influenced by topography. The spatial distribution of rainfall reflects the orographic effect on rainfall with the western-eastern orientation of the mountainous regions of La Sierra having a direct effect on the north-east trade winds, which collect moisture across the Gulf of Mexico and then deposit large amounts as rain as they reach land. The highest mean monthly rainfall occurs along the windward northern slopes of the La Sierra Mountain, and parts of the lowland regions of Tabasco in the north. The most striking observation from rain gauge rainfall alone is that the upper and the middle sections of the Grijalva catchment have the highest rainfall totals in Mexico reaching monthly rainfall accumulations of over 300 mm (Figure 5.1 and 5.2).

However, results show that rainfall decreases gradually towards the south with the lowest rainfall occurring in the interior depressions of Chiapas with a long-term monthly rainfall total below 130 mm but increasing again on the western highlands along with the Pacific Ocean.



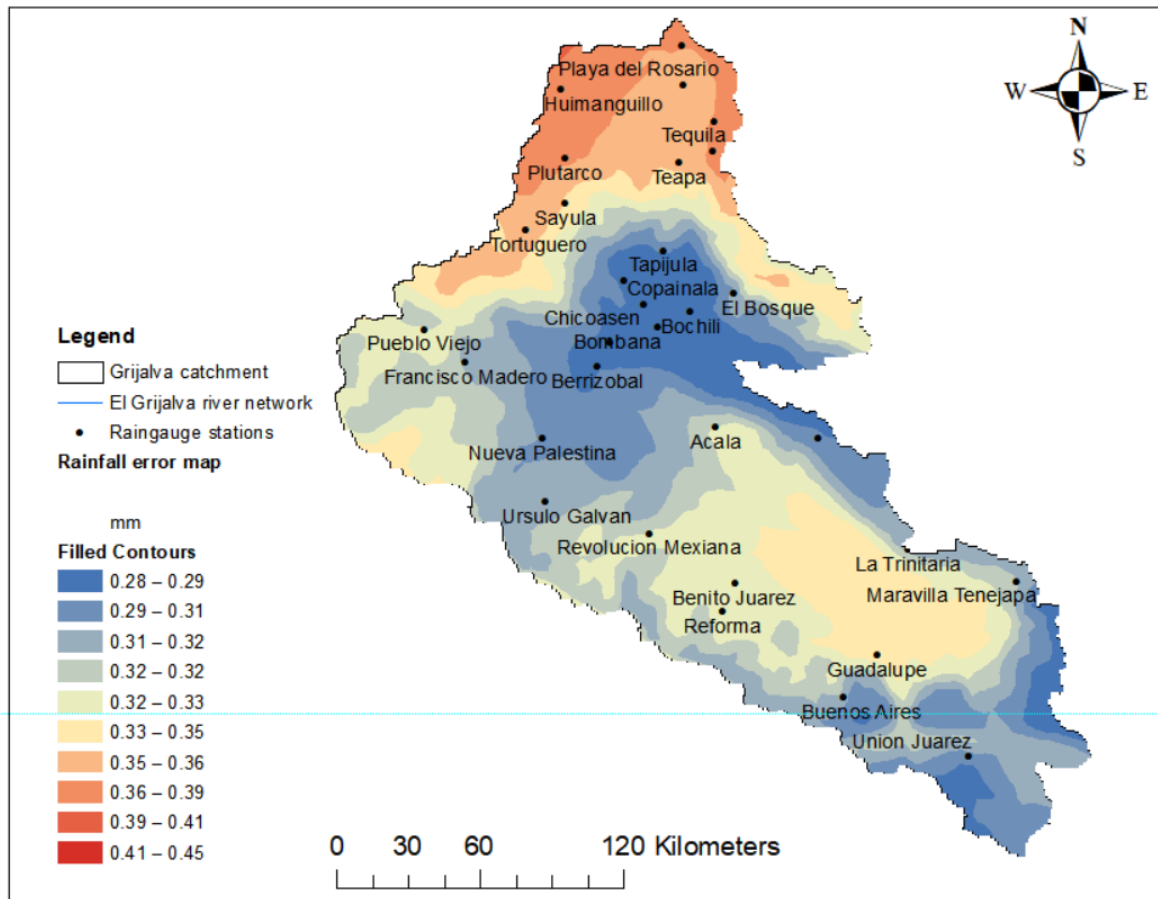
**Figure 5.1** Interpolated monthly total rain gauge rainfall distribution in Grijalva catchment highest totals occurring along northern slopes of La Sierra Mountains and covering most of Tabasco lowland areas while the central depressions of Chiapas have the lowest rainfall totals.



**Figure 5.2** Spatial distribution of mean monthly rainfall in the Grijalva catchment based on the Cokriging interpolation and merging of rain gauge rainfall, satellite rainfall and topographic datasets.

### ***Rainfall error distribution***

The spatial variability of rainfall errors in the catchment (Table 5.1) shows highly variable rainfall patterns in the study area (Figure 5.3). Results show spatial error variations following topography with low error fields (0.28 to 0.32 mm) in the southern slopes of Sierra mountainous regions covering the central depressions of Chiapas (Figure 5.3). Results show substantial errors ranging between 0.29 to 0.45 mm along the north-facing slopes of La Sierra mountainous regions, including the low-lying Tabasco plains. However, pockets of high rainfall errors ranging from 0.32 to 0.36 mm occur varyingly around La Angostura dam in the central depressions of Chiapas State.



**Figure 5.3 Rainfall error distribution showing spatially varying error rainfall error fields in the Grijalva catchment area.**

The temporal variability of rainfall errors in the catchment (Table 5.1) shows some of the highest monthly rainfall error distributions in the catchment with an average error of 6.48 mm, ranging between 5.27 mm in December and 7.39 mm in September. Most places receive almost all rainfall between May and October with a monthly average of 3.91 mm, while November to May are dry months with a long-term monthly average of 2.10 mm. Results show that low errors are in the dry season between November and May for nearly all rain gauge stations compared to the wet season. For example, large seasonal variability of errors in the dry season with average rainfall errors of 5.93 mm while in the wet season, it was 7.02 mm (Table 5.1). Conversely, the occurrence of the Inter-Tropical Convergence Zone (ITCZ) in wet months promotes highly variable and localised events, thus deteriorating the spatial correlation of the error field.

**Table 5:1. The quantified errors between the merged rainfall dataset and original rain gauge rainfall data show that errors varied seasonally with high error fields prevalent during the wet season between (May to October) compared to the dry season.**

Months	Mean monthly satellite rainfall	Mean monthly obs. RG rainfall	Mean monthly rainfall	% bias	Root mean-square error (RMSE)	Mean error	Average Standard Error (ASE)	Mean squared error
Jan	0.08	4.06	2.07	-98.11	5.94	3.99	1.99	15.89
Feb	0.06	1.49	0.77	-96.04	6.02	1.43	0.72	2.04
Mar	0.07	1.66	0.86	-95.84	6.27	1.59	0.79	2.53
Apr	0.07	1.87	0.97	-96.31	6.56	1.80	0.90	3.25
May	0.24	4.41	2.33	-94.57	6.89	4.17	2.09	17.43
Jun	0.41	8.42	4.42	-95.11	7.16	8.01	4.01	64.19
Jul	0.28	5.02	2.65	-94.44	7.03	4.74	2.37	22.50
Aug	0.34	7.43	3.89	-95.39	7.34	7.09	3.54	50.21
Sep	0.44	11.12	5.78	-96.08	7.39	10.68	5.34	114.10
Oct	0.34	8.52	4.43	-96.03	6.30	8.18	4.09	66.93
Nov	0.21	6.24	3.23	-96.59	5.54	6.03	3.01	36.36
Dec	0.12	5.42	2.77	-97.70	5.27	5.30	2.65	28.04
Long-term mean	0.22	5.47	2.85	-96.02	6.48	5.25	2.63	35.29

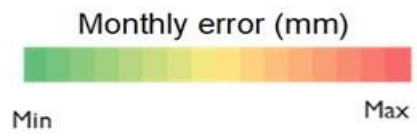
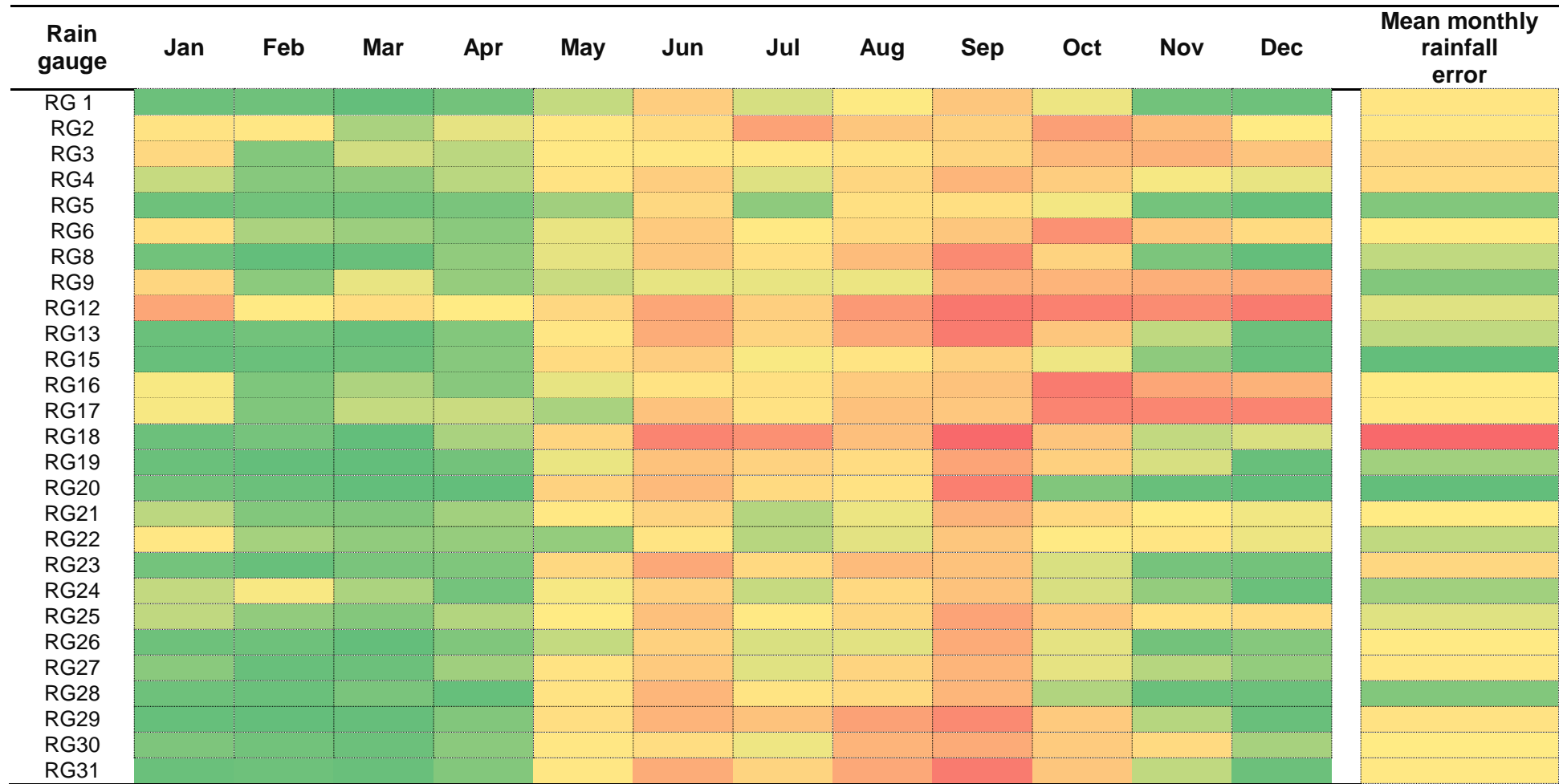
### ***Rainfall error correction***

Rainfall input data were corrected using the mean monthly rainfall errors per each gauging station derived from the differences between merged rainfall dataset and observed rain gauge rainfall in the Grijalva catchment. Results show that the average error at each rain gauge station in the Grijalva catchment is 0.3 mm. The lowest is 0.13 mm at Dos Montes and the highest 0.83 mm at Huimanguillo rain gauge station (Table 5.2) (Appendix A, Figure A.5). The errors at each gauging station were used to correct rainfall input data at each station.

Although Grijalva catchment is a poorly gauged with one rain gauge station per 1 700 km<sup>2</sup>, the Geostatistical merging technique was able to deliver satisfactory results because of the strong correlation between monthly rainfall observations, satellite rainfall data in ungauged areas and introducing the external drift as topographic data. The results show that the relative performances of the rainfall, interpolation and merging methods are strongly influenced by the rain gauge density and topography of the catchment.

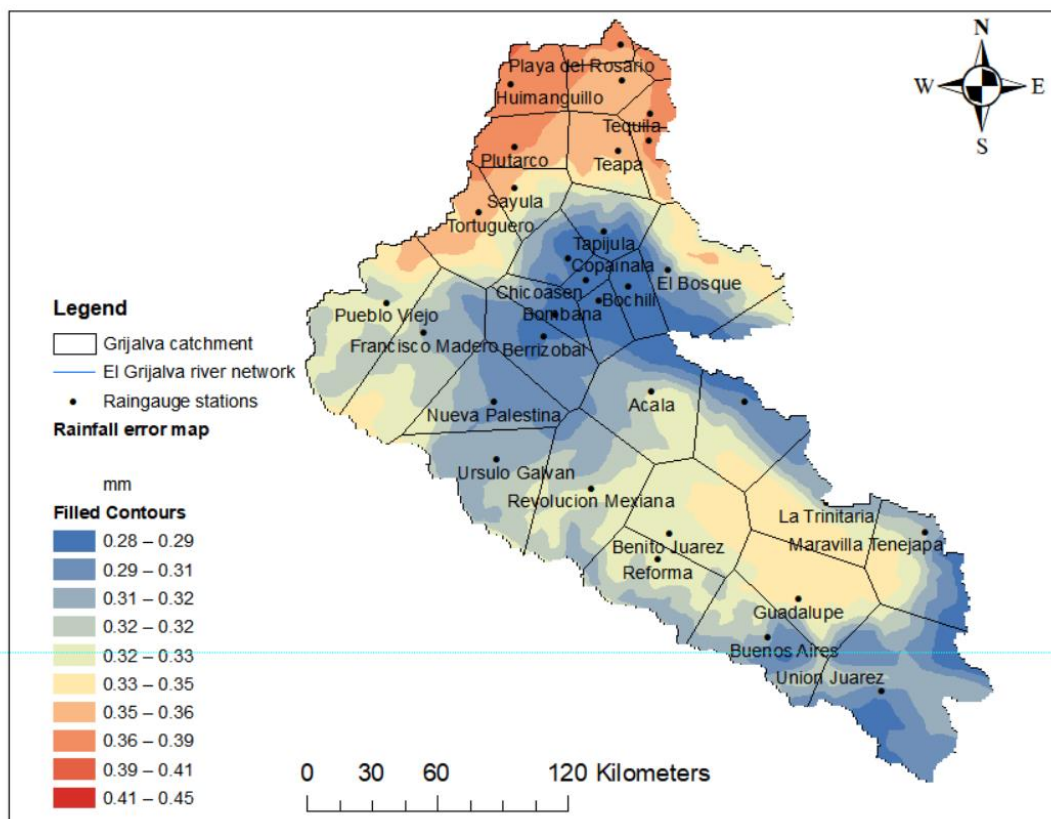


**Table 5:2 Mean monthly rainfall errors per gauging station derived from the differences between merged rainfall dataset (rain gauge and TRMM rainfall) and original rain gauge rainfall in the Grijalva catchment.**



### ***SHETRAN Thiessen polygons integrated with error fields***

The interpolated error fields derived from the Cokriging with external drift technique were integrated within the SHETRAN Thiessen polygons, as shown in Figure 5.4. The Thiessen polygon rainfall input data for SHETRAN were corrected using quantified errors at each rain gauge, and this enabled a bias-corrected input rainfall for each polygon. The inaccuracy resulting from the use of Thiessen polygons for rainfall representation in the SHETRAN hydrological model was reduced by computing errors at each rain gauge location, resulting in the correction of input rainfall in each polygon (Table 5.2).



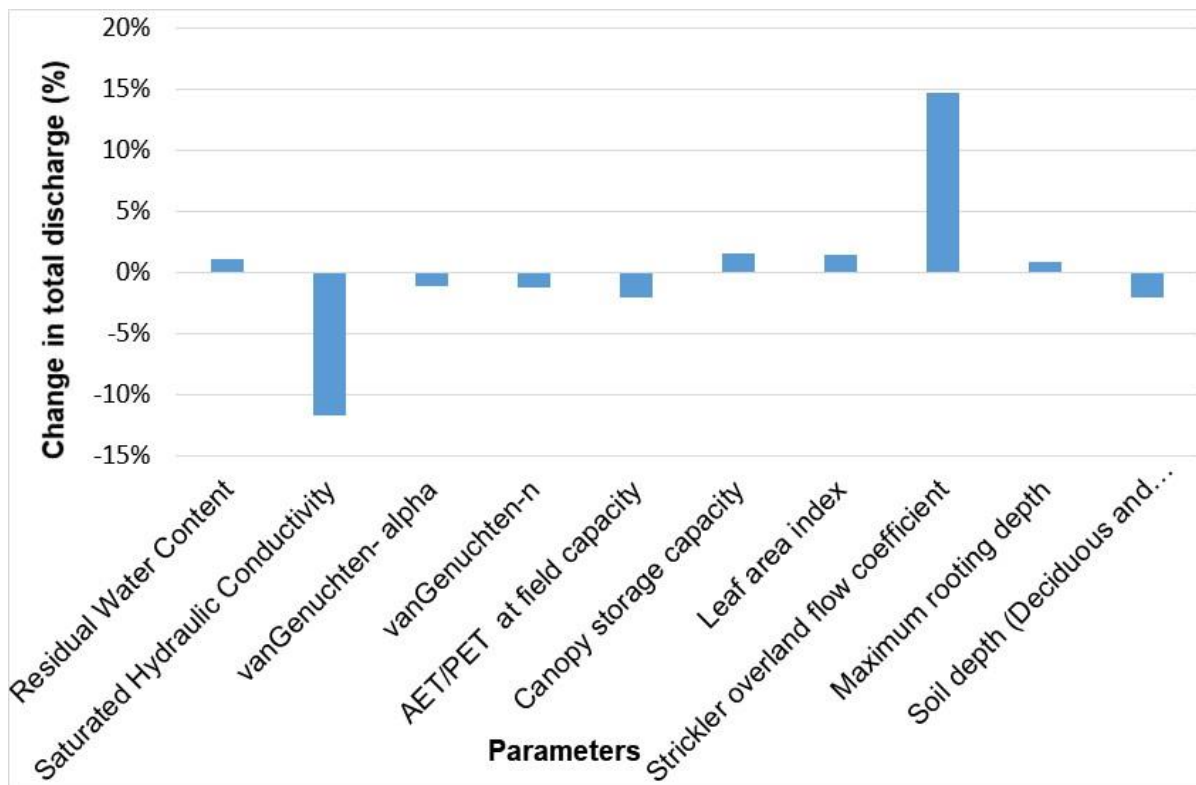
**Figure 5.4** The interpolated error fields derived from the Cokriging with external drift technique were integrated with SHETRAN Thiessen polygons.

### 5.1.2 Sensitivity analysis results

Diagnostic assessment of parameter sensitivity analysis was based on comparisons with total discharge as variable response output. The contribution of each parameter change to the variance in total discharge was investigated (Guse et al., 2006). A local sensitivity analysis of the type “one-at-a-time” (OAT) technique was applied by analysing the effect of one parameter change on discharge output at a time, keeping the other parameters fixed. Preliminary model results using baseline (average) parameter values derived from the literature (Zhang et al., 2013; Zhang et al., 2015) showed that average values do not represent actual catchment conditions and processes in the study area. There was a need to determine how different values of individual parameters affect model simulations.

Results from sensitivity analysis indicate that surface roughness (Strickler overland flow coefficient) ( $K$ ) and Saturated conductivity ( $K_s$ ) parameters caused the most considerable change in the total discharge output (Figure 5.5). Results based on variations of parameter values showed that increasing the Strickler overland coefficient value by 1.5 for each land use type (Deciduous forest, Grass, Evergreen forest, Arable, Bare ground, Shrub and Urban land use) resulted in the most significant change, reaching 15% of total discharge (Figure 5.5). A higher Strickler coefficient value represents lower roughness of the land surface, a smoother surface results in more runoff into the channel (Xevi et al., 1997). The higher speed of surface runoff increases the total discharge in the river channel.

Saturated conductivity ( $K_s$ ) parameter had the second-highest effect on total discharge. The Saturated conductivity ( $K_s$ ) parameter values were increased by 1.5 and resulted in a reduction in the total discharge output by 12%. Results show that high saturated hydraulic conductivity values result in higher soil infiltration rates resulting in a lower runoff with corresponding decreases in peak discharge. All other tested parameters caused substantially less change, and this was followed by Saturated water content ( $\theta_s$ ) AET/PET and soil depth.



**Figure 5.5: Percentage change in total discharge due to increases in parameter values (Adapted from Sreedevi and Eldho, 2019).**

Although the three parameters were relatively less sensitive, they were complementary parameters that were used to refine calibration results from the two most sensitive parameters. Therefore, the sensitivity analysis showed that changes in surface roughness (Strickler overland flow coefficient) (K) were observed to have a more pronounced effect on total discharge. Among soil parameters, saturated hydraulic conductivity had the most effect on total peak discharge. Saturated water content ( $\theta_s$ ) and AET/PET were less sensitive, however, they complemented in calibrating the model. The rest of the parameters were less sensitive, and their values remained fixed during calibration.

### Mass balance checks

As part of the sensitivity analysis, mass balance checks were done to ensure the model was not gaining or losing inappropriate amounts of discharge volumes. Hydrological models gain or lose discharge due to numerical rounding errors, model instability and convergence (Ackers et al., 2010). This gain or loss of volume is referred to as 'mass error.' Knowing the sum of the mass error and comparing it to other components of the water balance in the hydrological analysis makes it easier to understand uncertainties in the results of the model. The spatially averaged flow totals over 2013 to 2017 simulation period were calculated to summarise the overall water balance required for mass balance values (Figure 5.6). Check results show mass errors with a deficit of 1.6% in the Grijalva catchment. The deficit was acceptable as it was not over 5%. If greater than 5%, it would suggest that model schematisation or model type (software) is not robust and needs to be challenged (Ackers et al., 2010).

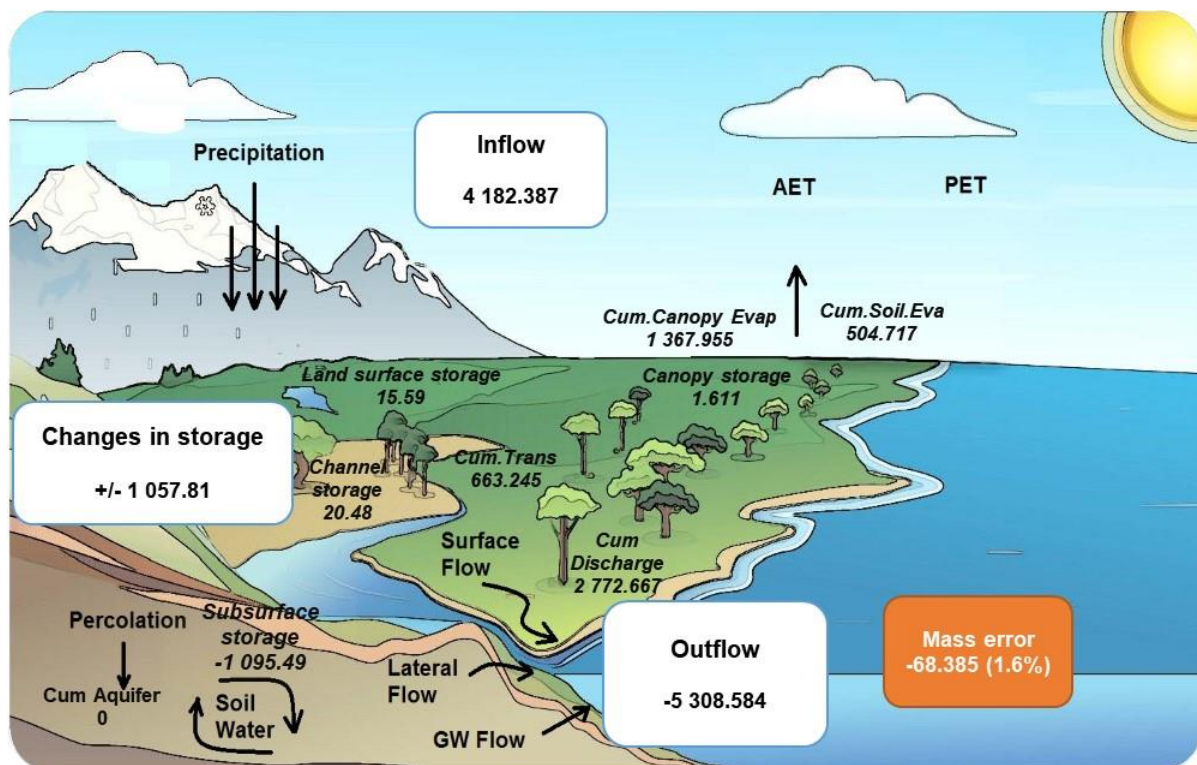


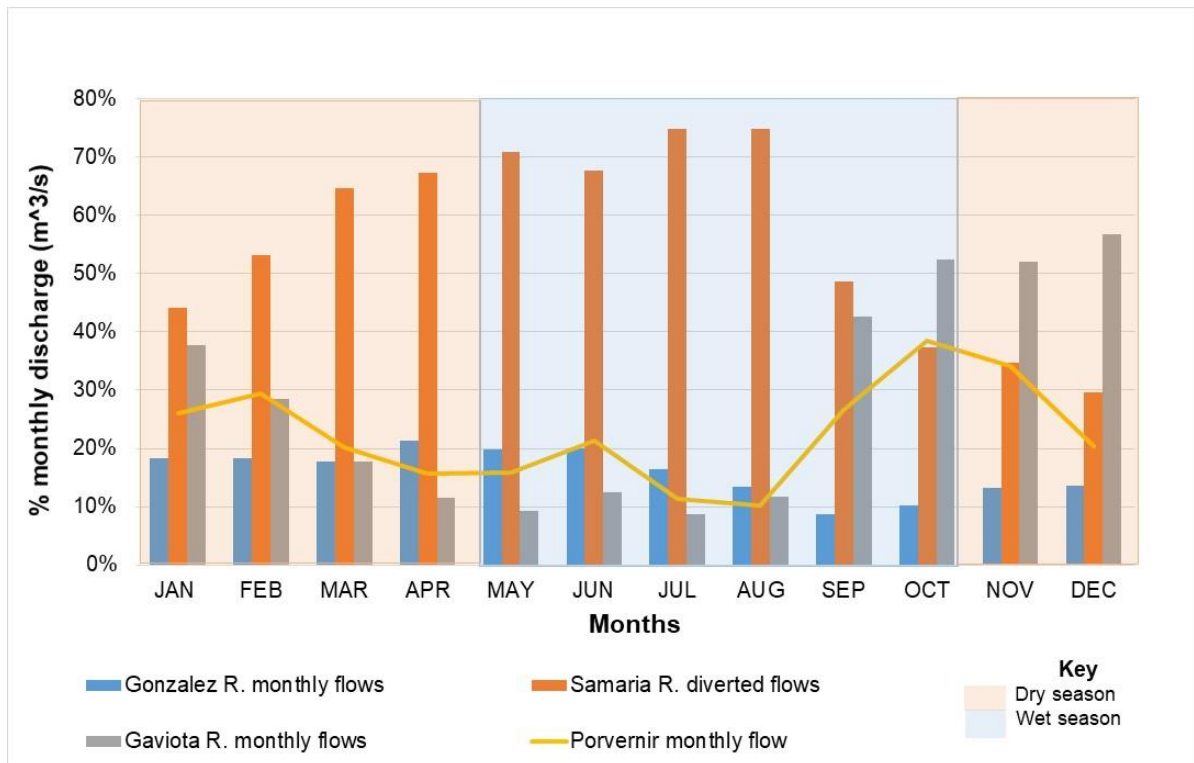
Figure 5.6 Spatially averaged flow totals over 2013 to 2017 simulation period for mass balance checks in La Grijalva catchment (Background picture Met Office 2009).

### 5.1.3 Discharge output error correction

The analysis of river flow records revealed that substantial discharge is diverted through Samaria River, a distributary of the main Grijalva River. Results showed that 53% of Grijalva flows are diverted from the model domain through this Samaria River, leaving 16% in the Carrizal River and 31% in the La Sierra River. That means 46.61% of discharge (combined flows for Carrizal 16% and La Sierra River 31%) passes through the model domain outlet at Povenir gauging station, located northeast of Villahermosa. Most of the Grijalva River flows are diverted into the Samaria River (53%), and the remaining flows in the Carrizal and La Sierra Rivers comprise the remaining 46.61% that pass through the domain outlet.

Based on calculations without flow contribution from La Sierra River (31%) but taking into account flows from Samaria and the Carrizal Rivers only, 77% of the Grijalva River flows were found to be diverted through Samaria River, leaving 23% as Carrizal River flows. This remaining 23% is the proportion of total discharge flowing towards Villahermosa city and is consistent with the maximum operational flow threshold for Carrizal River used at the El Macayo flood control barrier (Arreguín-Cortés et al., 2014). Also, results reveal seasonal variations of diverted discharge with 57% of flows diverted in the rainy season, while 54% are diverted in the dry season (Figure 5.7). The analysis of hydrological records shows that monthly flow variations for July and August have the highest diverted flows (75%) while December has the lowest (30%) (Figure 5.7).

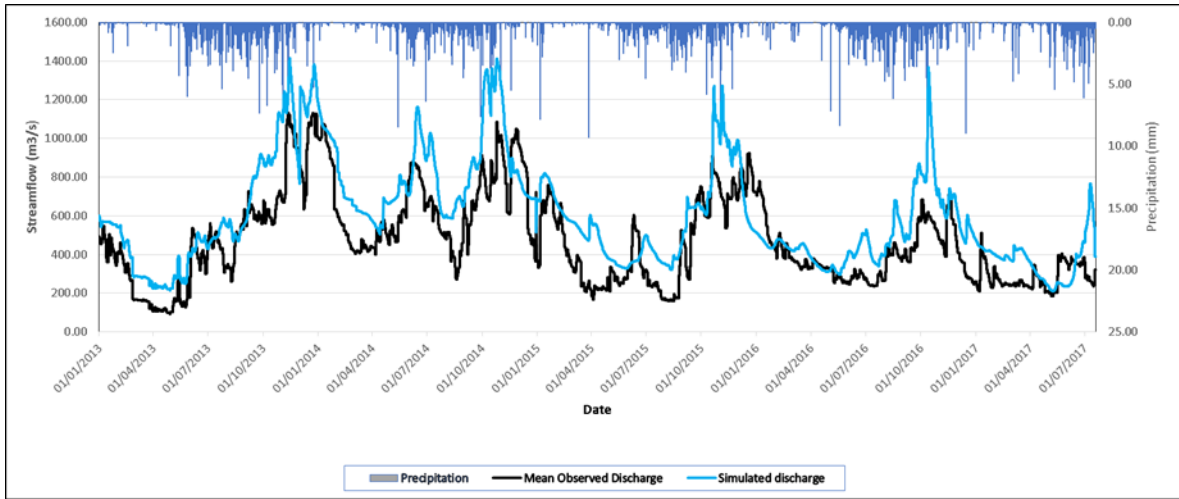
Povernir discharge is highest in October, the last month of the wet season, reaching 38% of Grijalva flows, while the lowest discharge (10%) is in August, mid-wet season (Figure 5.7). Also, flow balance checks showed that the annual flow surpluses at Povernir and Samaria Gauging Stations were approximately 149 008.35 m<sup>3</sup>, while at Gaviota and Gonzalez Gauging Stations there was a deficit of 149 008.35 m<sup>3</sup> each. Flow surpluses at the Samaria and Povernir outlets were able to balance inflows from the Gonzalez and Gaviotas Rivers. The flow in the Samaria River represented the proportion of flows (53%) that are diverted from the model domain. This proportion of flows was used to correct simulated discharge at Povernir outlet).



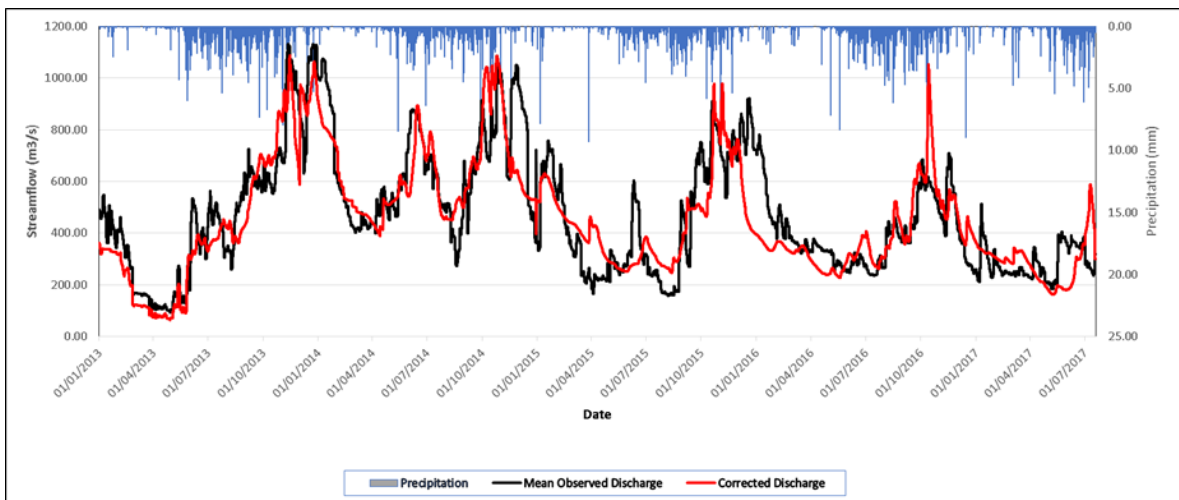
**Figure 5.7 Percentage fraction of seasonal and monthly-diverted flows in the lower reaches of the Grijalva River.**

***Effect of river diversion reproduced in simulated time series.***

When diverted flows were incorporated into output discharge flows and reproduced in simulated time series, they had an over-predictive effect (Figure 5.8). The model generates high surface flow than reality when diverted flows are included with overestimation of high peaks and low level or smaller peaks. However, the removal of diverted flows from the model domain resulted in a close match or agreement between the observed and the simulated discharge (Figure 5.9). Results show simulations were improved after the removal of diverted flows (CC improved by 33.39%, NSE 3.67%, PBIAS with 33.99 and RMSE by 27.36%) (Table 5.3). Significant improvements were realised in the percentage BIAS, and the NSE had the least of improvement.



**Figure 5.8: Calibration results flows with diverted flows included.**



**Figure 5.9: Calibration results using corrected flows- diverted flows removed.**



**Table 5:3 Percentage improvement in model simulations after removing the proportion of diverted flows.**

<b>Evaluation indices</b>	<b>Scale/Units</b>	<b>Original</b>	<b>Revised</b>	<b>% Improvement</b>
<b>Correlation Coefficient (CC)</b>	0 < CC < 1	0.52	0.69	33.39
<b>Nash–Sutcliffe coefficient (NSE)</b>	0 < NSE < 1	0.68	0.71	3.67
<b>Percentage BIAS</b>	(%)	-8.18	-5.41	33.99
<b>Root Mean Square Error (RMSE)</b>	(mm)	192.13	139.13	27.36

#### **5.1.4 Calibration and validation results**

This section provides the results of the calibration and validation of the SHETRAN hydrological model following the correction of rainfall inputs and discharge data. The results of model input and output data, error quantification and removal are also presented, including improvements made from the detailed representation of catchment conditions and processes using high-resolution satellite data. The model took 3 hours and 45 minutes to run the calibration (2013 - 2015 period) and 2 hours 10 minutes to run the validation (2015 - 2017 period), on a 2 x 2 km grid resolution domain of 57,958 km<sup>2</sup> running on a computer with i7–6700HQ CPU operating at a speed of up to 3.50GHz. The model took 5 hours and 30 minutes to run the entire calibration and validation period (2013-2017) on the same model domain with the same grid resolution using the same computer.

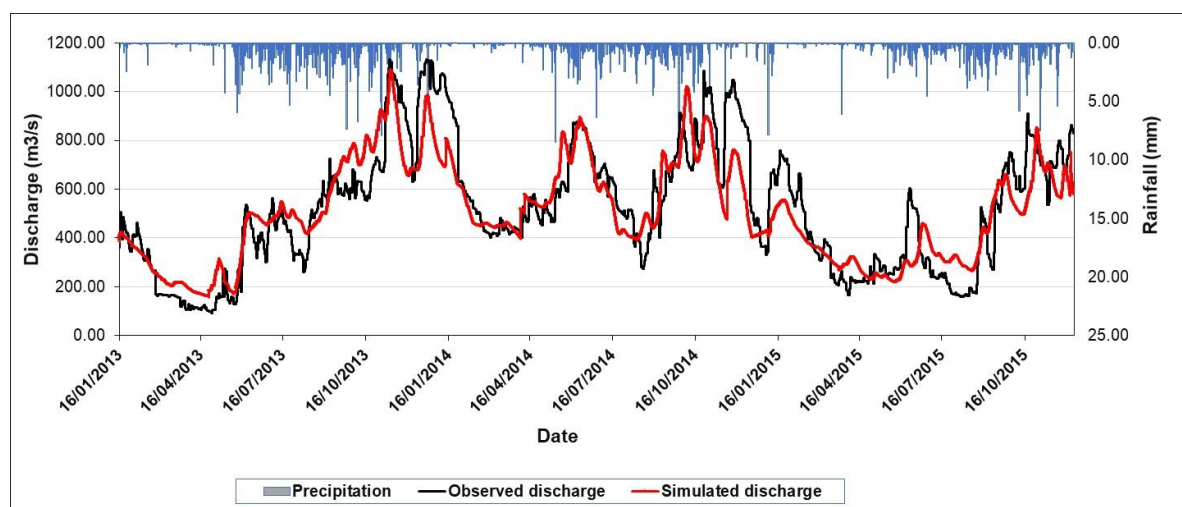
##### ***Calibration results***

Calibration results show that the SHETRAN hydrological model was able to simulate river discharge with similar seasonal trends to the observed discharge (Figure 5.10). Table 5.4 shows satisfactory results obtained from the assessment of the model's performance in simulating discharge in the Grijalva catchment area (NSE = 0.72, CC = 0.74, Bias = -0.44% and RMSE 139.56). The results show the Pearson Correlation Coefficient (CC) of (0.74) showing the strength of the correlation between the simulated and observed discharge, which should be above 0.5 (CC ≥ 0.5) for good performance. The PBIAS values ranged from -8.4 to 20.3%, but the final value (- 4.4%)

was within a very good calibration range ( $PBIAS < \pm 10$ ) (Table 5.4). Apart from the PBIAS model performance indication, the SHETRAN discharge simulation showed a "good" performance of NSE trends (0.72) within a good statistical fit of simulated flows and was within  $0.5 < NSE < 0.65$  thresholds (Moriasi et al., 2007; Ritter and Muñoz-Carpena, 2013). However, the RMSE of 139.56 mm was not satisfactory, showed substantial errors when simulated and observed discharge was compared (Table 5.4). These results show the need for new approaches and complementary tools to reduce errors and improve the reliability of model results.

**Table 5:4 The evaluation results for SHETRAN model performance ratings for the period 2013 to 2015 (description classes from Moriasi et al. (2007)).**

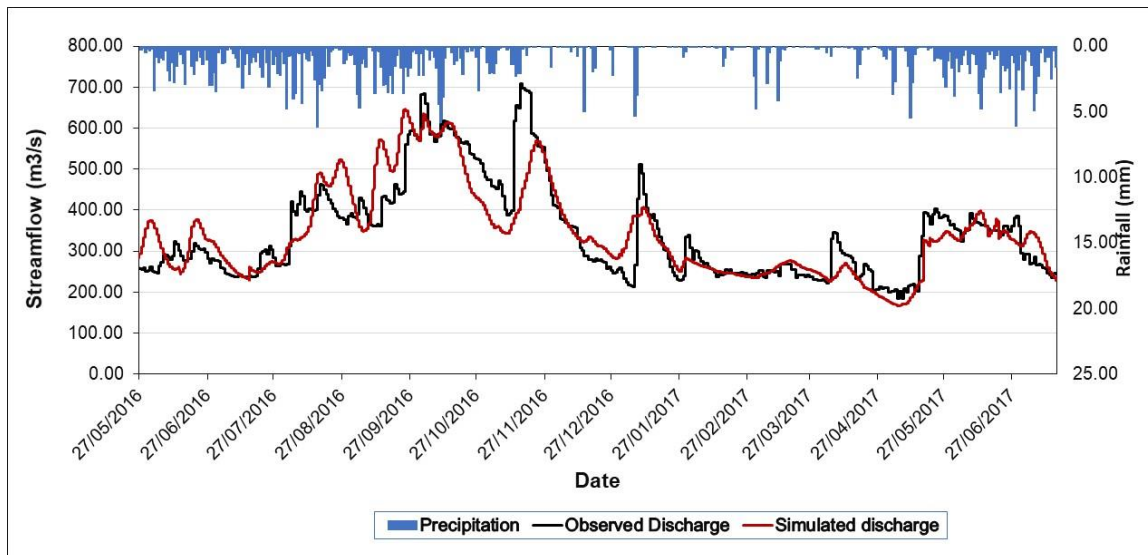
Process	Statistic	Value	Description
Calibration	NSE	0.72	Good
	CORRELATION	0.74	Good
	PBIAS (%)	-4.44%	Very Good
	RMSE (mm)	139.56	Unsatisfactory
Validation	NSE	0.56	Good
	CORRELATION	0.60	Good
	PBIAS (%)	-6.3%	Very Good
	RMSE (mm)	62.59	Satisfactory



**Figure 5.10: Comparison of simulated and observed discharge for the 2013 to 2015 calibration period.**

## Validation results

Model validation results based on the 2016 -2017 period showed good results (Figure 5.11, Table 5.3). The results show that the observed and simulated discharge values were relatively closely aligned with NSCE of 0.56 which is within the good statistical goodness-of-fit range of  $0.5 < NSE < 0.65$  (Moriasi et al., 2007; Ritter and Muñoz-Carpena, 2013). The validation results showed a Correlation Coefficient (CC) of Pearson of 0.60, showing a good association between simulated and observed discharge, which should be above 0.5 ( $CC \geq 0.5$ ). The PBIAS values were -13.3% to 24.3% during validation, but the ultimate magnitude (-6.3%) was classified as a very good validation result ( $PBIAS < \pm 10$ ).



**Figure 5.11: Comparison of simulated and observed discharge for the validation period 2016 to 2017.**

The RMSE result of 62.59 mm was satisfactory in terms of the average magnitude of the errors between the estimated and the observed discharge. The simulation results of seasonal river flow mirrored the observed water levels with the highest in the wet season and low in the dry season. Like the calibration, the validation results did not accurately reflect some sharp river peaks, especially at the end of each rainy season. Generally, model results showed an underestimation of high peaks and overestimation for low level or smaller peaks. However, results show that model efficiency during calibration exceeds the validation period with a small variance in performance.

The SHETRAN performance measure of NSE was within the range of 0.7 and 0.79 for calibration and between 0.72 and 0.56 for validation, which is comparable to other studies (Table 5.5) that used SHETRAN (Birkinshaw et al., 2014; Dukic and Radic, 2016, Dukic and Radic, 2014; Mourato et al., 2015; Naseela et al., 2015; Tripkovic, 2014; Zhang, 2015). The performance of the model was reasonable in other similar measures and in the range of other studies that used SHETRAN (Table 5.5). In these studies, NSE values above 0.5 are frequently reported. For example, by using the hourly resolution rainfall data, daily PET data and 200 m spatial resolution Op de Hipt et al., 2017 obtained NSE of 0.65 and 0.7 at the basin outlet for the SHETRAN simulations of the Dano catchment in Burkina Faso. Naseela et al., (2015) calibrated and validated the SHETRAN model in a tropical basin of 69 430 km<sup>2</sup> in India using daily and obtained NSE of 0.80 and 0.90 at the basin outlet.

**Table 5:5 Studies that used SHETRAN for simulating river discharge and catchment size in comparison to the present study (Op de Hipt et al., 2017).**

Study	Location	Spatial/Temporal Resolution	Catchment/Plot size	Performance
				Discharge
Present study	Mexico	2 km/3 hr and 1km/1-3 hr continuous	6,740 km <sup>2</sup> , 57,960 km <sup>2</sup>	0.7 ≤ NSE ≤ 0.6
Op de Hipt et al., 2017	Burkina Faso	200 m/h (continuous)	126 km <sup>2</sup>	0.65 ≤ NSE ≤ 0.7
Janes et al., 2018	UK	1 km <sup>2</sup> /hourly continuous	2,400 km <sup>2</sup>	0.9 ≤ NSE ≤ 0.6
Đukic and Radic et al., 2016 and 2014	Serbia	25 m/h (event)	114 km <sup>2</sup>	0.8 ≤ R <sup>2</sup> ≤ 0.9
Zhang et al., 2015	Portugal	2 km/h (event)	705 km <sup>2</sup>	0.7 ≤ NSE ≤ 0.8
Mourato et al., 2015	Portugal	-/daily (continuous)	61–83 km <sup>2</sup>	0.5 ≤ NSE ≤ 0.7
Naseela et al., 2015	India	-/daily (continuous)	69,430 km <sup>2</sup>	0.8 ≤ R <sup>2</sup> ≤ 0.9
Birkinshaw et al., 2014	UK	50 m/h (continuous)	1.5 km <sup>2</sup>	0.8 ≤ NSE ≤ 0.9
Shrestha et al. 2013	Kathmandu valley	700m x700m/hourly	600 km <sup>2</sup>	0.8 ≤ NSE ≤ 0.7
Tripkovic et al. 2014	UK	10 m, 100 m/h (continuous, event)	0.09 km <sup>2</sup> , 9.2 km <sup>2</sup>	0.5 ≤ NSE ≤ 0.9
Elliott et al., 2012	New Zealand	20 m/15 min (event)	1.46–167 km <sup>2</sup>	0.6 ≤ NSE ≤ 0.9
Bathurst et al., 2011	Middle/South America	50–500 m/h, daily (continuous)	0.35–131 km <sup>2</sup>	0.8 ≤ NSE ≤ 0.9
Birkinshaw et al., 2010	Chile	50 m/h (continuous)	0.35 km <sup>2</sup>	0.8 ≤ NSE ≤ 0.9
de Figueiredo and Bathurst et al., 2007	Brazil	5 m–2 km/daily–monthly (continuous)	100 m <sup>2</sup> –137 km <sup>2</sup>	0.3 ≤ R <sup>2</sup> ≤ 0.9
Adams et al., 2004	New Zealand	0.5 m/min (event)	970 m <sup>2</sup>	-
Norouzi Banis et al., 2004	UK	20 m/5 min (continuous)	0.03, 0.05 km <sup>2</sup>	-
Anderton et al., 2002	Spain	100 m/20 min (continuous)	0.56 km <sup>2</sup> , 4.17 km <sup>2</sup>	0.5 ≤ NSE ≤ 0.9
Lukey et al., 2000	France	50 m/daily (continuous)	0.86 km <sup>2</sup>	0.03 ≤ R <sup>2</sup> ≤ 0.4

## 5.2 Summary

The development of the large-scale hydrological modelling in the Grijalva catchment area involved data processing involving quality checks, error quantification and removal and data merging for correct/accurate data input and representation of catchment conditions and processes in the model. A high-resolution, fully distributed hydrological model integrated with high spatial (< 2 km) and temporal (3 hr) resolution satellite datasets were developed to improve the simulation of the timing, magnitude and extents of flooding in a large lowland tropical catchment.

The sensitivity analysis done showed Strickler overland flow coefficient and Saturated hydraulic conductivity as the most sensitive parameters in the catchment. Other less sensitive parameters were the AET/PET field capacity and soil depth parameters. The four parameters were selected for model calibration and validation. Considerable efforts were made to find proper estimates of parameter values that result in model simulations matching the observed discharge carefully and consistently (Ajami et al., 2007).

Model performance was further enhanced by removing diverted flows from simulations, and the model outputs were satisfactory, capable of simulating trends and reflecting the observed flow conditions in the study area. Results show that improvements in the simulations were obtained by optimally integrating several datasets with the best combination of the model parameter, input and output datasets based on input and output data error quantification and removal.

## 5.3 Flood scaling relationships and estimation of peak discharge

This section presents results of power-law relationships between peak discharge and catchment properties and processes governing the generation of peak floods in the La Sierra catchment area. First presented is proof or evidence of power-law relationships between drainage areas and peak discharge in the catchment. Also presented are statistical analysis results describing the relationships and the extent to which rainfall variability control the scaling intercept ( $\alpha$ ) and scaling exponent ( $\theta$ ) (flood parameters). The section also comprises results from the application of power-law relationships to estimate flood scaling parameters and the magnitude of peak discharge in the La Sierra catchment.

### 5.3.1 Peak discharge scaling relationships

In this section, presented are result from the analysis of peak discharge data based on hydrometric records of peak flows from 59 rainfall events that occurred in La Sierra catchment between 2012 and 2015. The results were obtained from power-law relationships between peak discharge from individual rainfall events, and the corresponding catchment drainage size plotted on a logarithm scale. The scaling intercepts ( $\alpha$ ), exponents ( $\theta$ ) (flood parameters) and  $R^2$  value were derived from the equation of the regression line (Ayalew et al., 2015).

Results show that the average coefficient of determination ( $R^2$ ) of the relationships between individual peak discharge and corresponding drainage areas for all the 59 rainfall-runoff events was 0.86 with 80% of events having an  $R^2$  value of greater than 0.8 (Table 5.6). Also, the results show that the flood scaling exponent ( $\theta$ ) values vary from 0.3 to 0.8 for rainfall events that cover the entire catchment area and from 0.2 to 0.8 if partial rainfall coverage of the catchment is included.

**Table 5:6 Results from scaling relationships between peak discharge from 59 rainfall events (that occurred in the La Sierra catchment) and nested catchment drainage areas showing scaling intercepts ( $\alpha$ ), exponents ( $\theta$ ) and coefficient of determination ( $R^2$ ) values.**

Storm No.	Peaking date	$\alpha$	$\theta$	$R^2$	Storm No.	Peaking date	$\alpha$	$\theta$	$R^2$
1	01/01/2012	6.72	0.46	0.76	31	24/05/2014	29.07	0.31	0.66
2	08/01/2012	3.98	0.51	0.97	32	28/05/2014	5.89	0.81	0.90
3	30/01/2012	12.32	0.34	0.88	33	03/06/2014	5.11	0.61	0.82
4	17/04/2012	37.38	0.14	0.51	34	05/07/2014	3.39	0.49	0.78
5	14/05/2012	30.40	0.22	0.52	35	14/07/2014	5.84	0.50	0.73
6	21/05/2012	80.85	0.20	0.53	36	03/08/2014	5.67	0.70	0.75
7	26/06/2012	10.76	0.36	0.59	37	15/08/2014	5.04	0.45	0.52
8	12/08/2012	50.40	0.22	0.57	38	26/08/2014	5.03	0.56	0.53
9	27/09/2015	17.46	0.35	0.59	39	31/08/2014	4.75	0.58	0.85
10	20/12/2012	10.76	0.36	0.59	40	12/09/2014	5.09	0.54	0.78
11	20/01/2013	33.01	0.24	0.56	41	19/09/2014	4.39	0.66	0.83
12	08/06/2013	5.41	0.84	0.58	42	23/09/2014	5.78	0.73	0.79
13	06/07/2013	5.25	0.59	0.81	43	14/10/2014	16.22	0.35	0.58
14	18/08/2013	28.08	0.24	0.60	44	26/10/2014	10.87	0.39	0.50
15	27/08/2013	11.57	0.38	0.59	45	29/10/2014	51.13	0.47	0.55
16	07/09/2013	10.75	0.44	0.53	46	08/11/2014	28.60	0.24	0.57
17	14/09/2013	48.45	0.17	0.59	47	13/11/2014	55.09	0.20	0.64
18	26/09/2013	46.15	0.18	0.51	48	27/11/2014	15.09	0.37	0.60
19	14/10/2013	12.14	0.41	0.71	49	13/06/2015	10.80	0.40	0.85
20	28/10/2013	1.72	0.73	0.79	50	29/08/2015	7.73	0.45	0.92
21	16/11/2013	12.32	0.44	0.87	51	15/09/2015	5.39	0.81	0.82
22	27/11/2013	1.15	0.77	0.86	52	23/09/2015	9.33	0.60	0.82
23	13/12/2013	0.61	0.95	0.80	53	01/10/2015	6.62	0.63	0.94
24	21/12/13	19.91	0.39	0.89	54	20/10/2015	9.72	0.61	0.66
25	13/03/2014	2.30	0.61	0.97	55	26/10/2015	17.52	0.45	0.91
26	08/04/2014	10.12	0.88	0.57	56	14/11/2015	5.41	0.56	0.70
27	16/04/2014	27.30	0.25	0.50	57	23/11/2015	82.32	0.26	0.70
28	20/04/2014	5.60	0.69	0.88	58	08/12/2015	19.70	0.42	0.75
29	05/05/2014	5.37	0.70	0.78	59	20/12/2015	12.24	0.50	0.92
30	14/05/2014	4.64	0.60	0.76					



An important outcome from the analysis is that peak discharge in La Sierra nested catchments, exhibit power-law relationships with drainage areas, and this is similar to findings in catchments studied in humid temperate regions. The scale invariance between the peak discharge and drainage areas is valid at a rainfall-runoff event scale under conditions of either catchment-wide or partial rainfall coverage (Ayalew et al., 2015).

Besides, results show that flood scaling parameters, the intercepts ( $\alpha$ ) and exponents ( $\theta$ ) changes from one rainfall event to another. This study investigated these event-to-event changes and establish relationships with rainfall variability and other catchment processes and properties that control the spatial generation of peak discharge in La Sierra catchment. Results show that flood scaling intercept ( $\alpha$ ) and exponent values change from event to event depends on rainfall variability in combination with other catchment properties and processes. Relationships between these physical processes that control variations in flood parameters have been quantified to use the estimated parameters to predict peak discharge in the La Sierra sub-catchment area.

### **5.3.2 The effects of rainfall variability on scaling parameters**

The scaling relationships between rainfall variability and scaling parameters in nested catchments (in the La Sierra catchment area) were statistically analysed and quantified. Properties of rainfall variability that were considered in the analysis include rainfall coverage, seasonal variations, direction, type, rainfall duration and accumulation or depth (Table 4:7). The resulting flood parameters were calculated by grouping rainfall-runoff events by 1) coverage 2) by seasons 3) storm direction 4) rainfall type 5) rainfall accumulations (Wilkinson and Bathurst, 2018). The natural variability of estimates obtained was presented using the 95% confidence interval showing the range over which discharge naturally vary expressing confidence in peak discharge estimates.

### Effect of rainfall coverage on exponent ( $\theta$ ) values

Following the power-law relationships established between scaling parameters, peak discharge and drainage areas, results show that rainfall events that cover the whole catchment (excluding partial rainfall coverage events) have high exponent ( $\theta$ ) values (0.70) and low intercepts ( $\alpha$ ) (6.70) with a model fit ( $R^2 = 0.75$ ) (Figure 5.12). However, when peak discharge is generated from individual rainfall events that both completely and partially cover nested catchments, results showed low exponent ( $\theta$ ) values (0.46) and slightly increased intercepts ( $\alpha$ ) (6.72) values with an equally good model fit ( $R^2 = 0.76$ ) (Figure 5.13). The 95% and 5% intervals in Figures 5.12 and 5.13 show bounds where the peak discharge estimates lie confidently along the best fit line (Helsel and Hirsch, 2002).

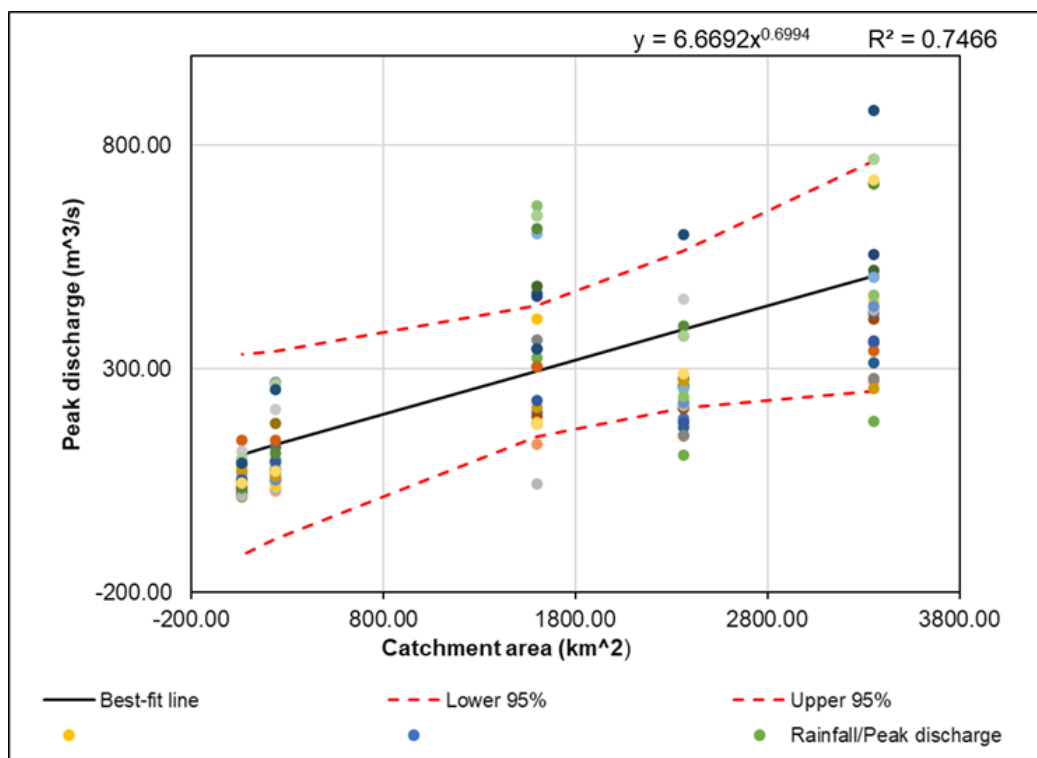


Figure 5.12: Peak discharge scaling relationships between peak discharge and nested catchment drainage areas experiencing complete rainfall coverage. Dotted lines show the upper and lower limit of the 95% confidence interval for estimated discharge values.

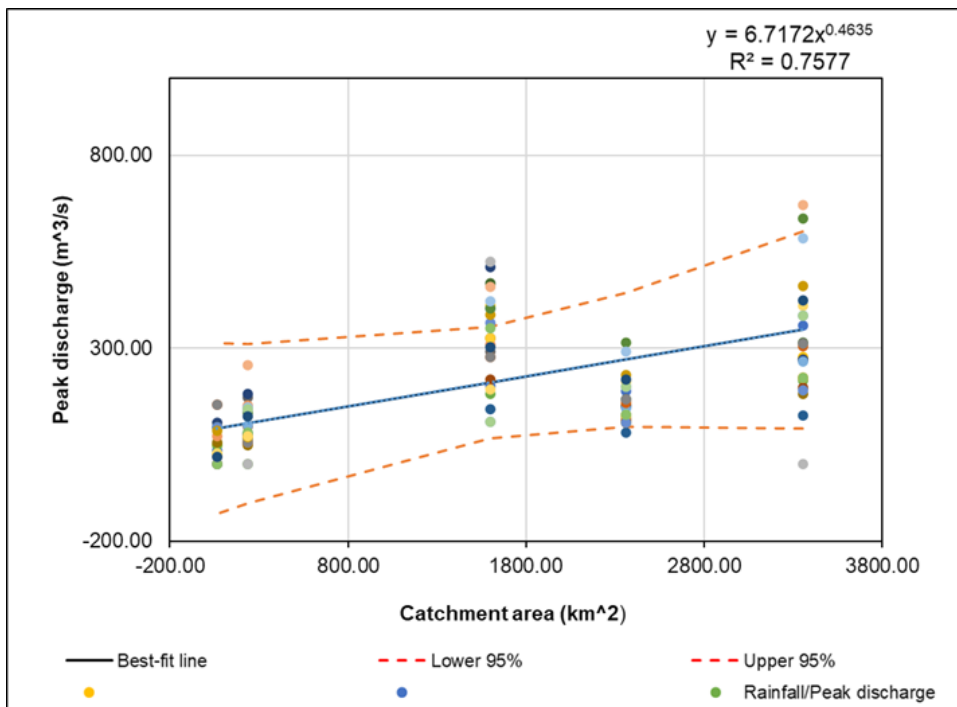


Figure 5.13: Peak discharge scaling relationships for partial rainfall coverage of catchments for rainfall events 2012-2015. Dotted lines show the upper and lower limit of the 95% confidence interval for all discharge values.

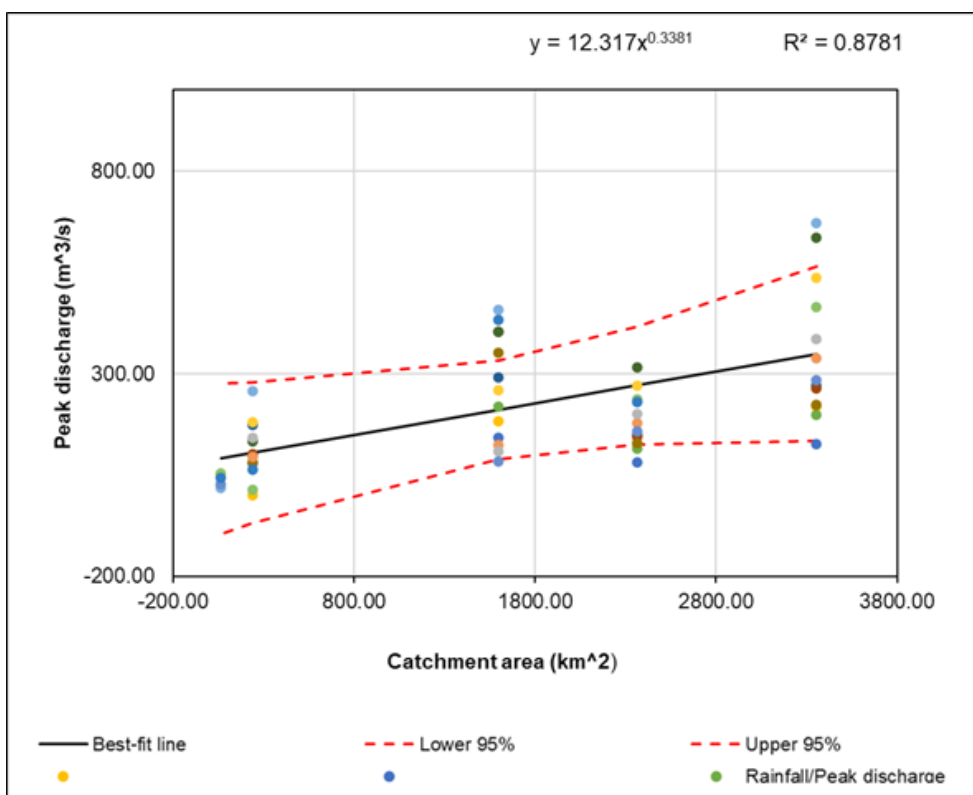


Figure 5.14: Peak discharge scaling relationship in catchments with both complete coverage and partial coverage rainfall.

However, the lowest exponent ( $\theta$ ) value (0.34), an increased intercept ( $\alpha$ ) value (12.32) and a good coefficient of determination ( $R^2$  of 0.88) were obtained from individual rainfall events that partially covers nested catchments (Figure 5.14). The 95% and 5% interval in Figure 5.14 shows the range of individual discharge values where they frequently vary with confidence (Helsel and Hirsch, 2002).

Results from the power-law relationships established between drainage area and peak discharge from all individual rainfall events that partially and entirely cover catchments are shown in Table 5.7. It can be seen that the scaling exponent ( $\theta$ ) values decrease with increasing scaling intercept ( $\alpha$ ) values as the rainfall catchment coverage decreases. When the catchment experience rainfall-runoff events that cover only part of the basin, results show that the intercept ( $\alpha$ ) increases while exponents ( $\theta$ ) decrease (Table 5.7). The results confirm findings from Ayalew et al. (2015) study that reveals that when scaling intercept ( $\alpha$ ) values increases, the scaling exponent ( $\theta$ ) decreases as runoff generation decreases with the decrease of rainfall coverage in nested catchments.

**Table 5:7 Summary of power-law relationships between the drainage areas and peak discharge from rainfall events that completely, partially and entirely cover catchments from 2012 to 2015.**

Rainfall coverage in catchments	Average values		$R^2$
	Intercept ( $\alpha$ )	Exponent ( $\theta$ )	
Complete coverage	0.67	0.70	0.75
Complete coverage and partial coverage	6.32	0.46	0.76
Partial coverage	12.32	0.34	0.88

Results show that the event to event changes of flood scaling intercept and exponent ( $\theta$ ) values due to differences in rainfall catchment coverage are significant ( $P < 0.05$ ). A student's  $t$ -test on the coefficients on each rainfall coverage type shows the scaling relationships are each statistically significant at a 95% confidence interval, as shown in Table 5.8.

**Table 5:8 Statistical significance of scaling relationships between complete, partial and both complete and partial rainfall coverage in La Sierra catchment.**

Rainfall coverage	Value	Standard error	t	Pr >  t	Lower bound (95%)	Upper bound (95%)
Complete	0.161	0.026	6.237	<b>0.003</b>	0.089	0.232
Partial	0.112	0.022	5.167	<b>0.007</b>	0.052	0.172
Complete and Partial coverage	0.112	0.024	4.596	<b>0.010</b>	0.044	0.179

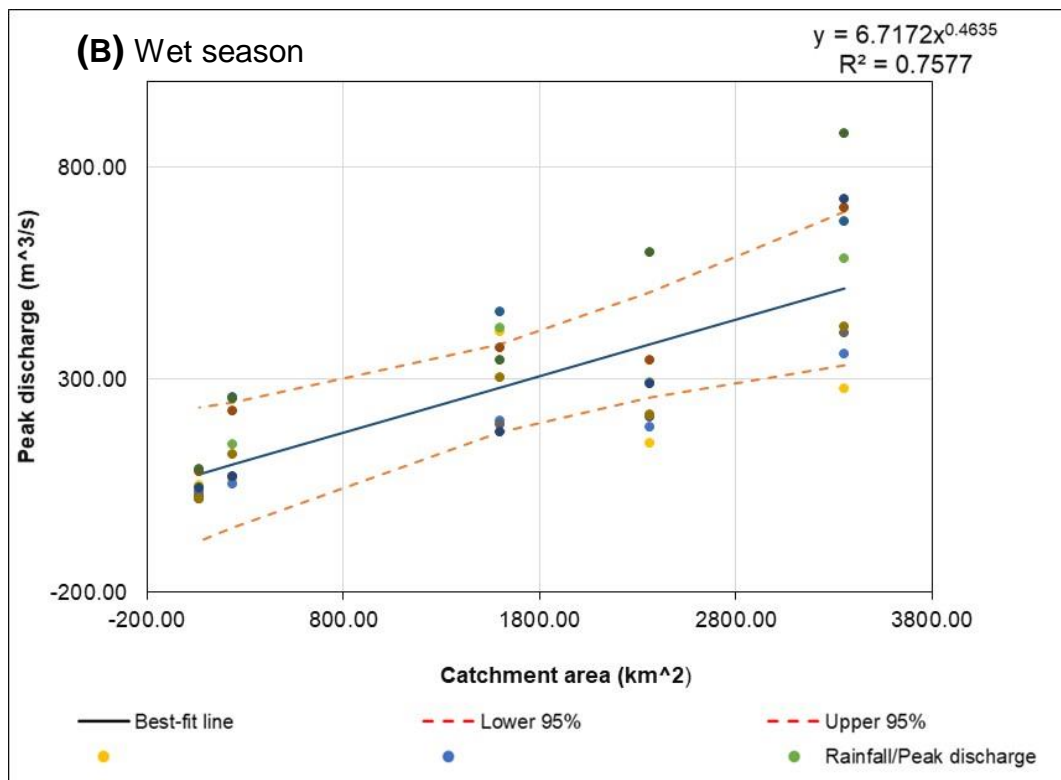
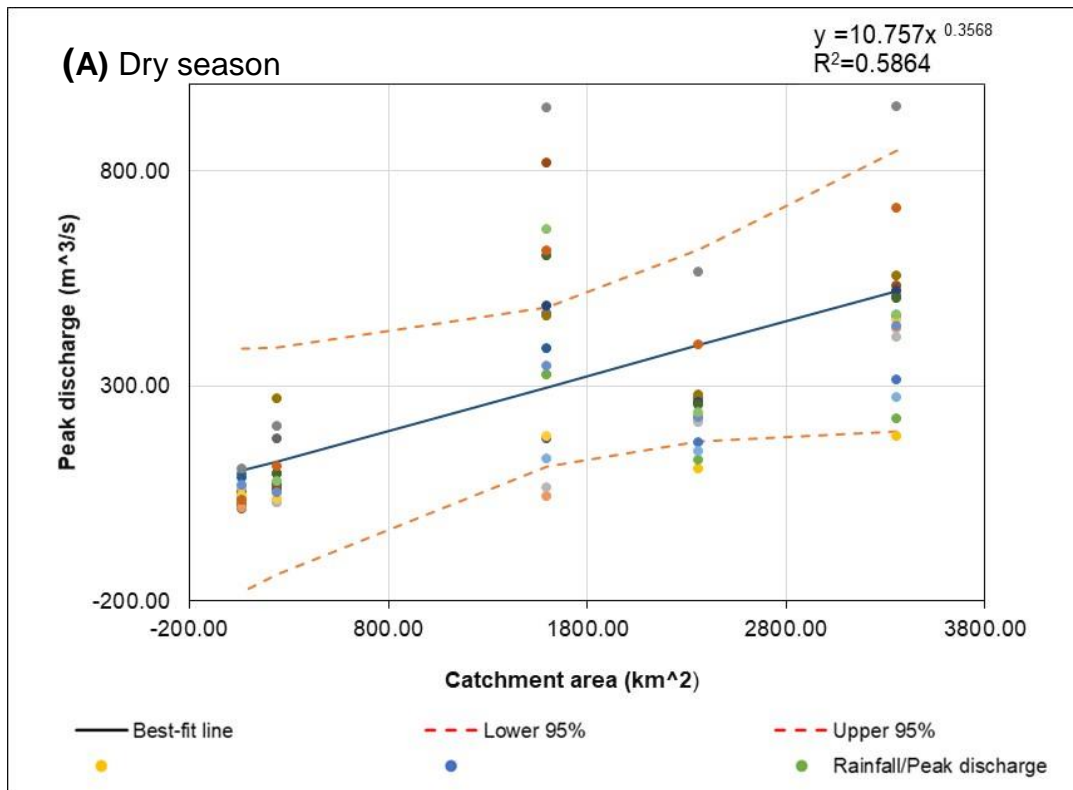
***Effect of seasonal rainfall variations***

Analysis of scaling relationships in wet and dry seasons shows that seasonal variations affect scaling parameters in the La Sierra catchment. Results show that changing from dry to the wet season has an effect of increasing exponent ( $\theta$ ) values and decreases intercept ( $\alpha$ ) values. However, changing from wet season to dry season has an opposite effect of decreasing exponent ( $\theta$ ) values. Results show that the dry season has low exponent ( $\theta$ ) parameter values (0.36) with a good model fit ( $R^2 = 0.57$ ) compared to relatively higher exponent ( $\theta$ ) values (0.46) with  $R^2$  of 0.76 in the wet season (Figure 5.15 A and B).

Statistical results summarised in Table 5.9 show that changing from dry to wet season has a significant effect on flood parameter values ( $P < 0.05$ ). A *t*-test shows there are statistically significant differences in scaling relationships between the two seasons. The resulting 95% confidence interval ranging from 0.47 to 1.41 in the dry season and 0.64 to 1.30 in the wet season shows some confidence in wet season estimates. A narrow confidence interval in wet season increases confidence in the result, while a broad interval in the dry season shows the existence of large uncertainties and hence low confidence in the estimates shown.

**Table 5:9 Statistical significance of scaling relationship from complete catchment rainfall coverage and catchment areas in the wet (a) and dry (b) seasons.**

Source	Value	Standard error	t	Pr >  t	Lower bound (95%)	Upper bound (95%)
Wet season	0.972	0.119	8.197	<b>0.001</b>	0.642	1.301
Dry season	0.940	0.170	5.522	<b>0.005</b>	0.468	1.413



**Figure 5.15: The scaling relationships between drainage areas and peak discharge from rainfall events that occurred in dry (a) and in wet (b) season. The dotted red lines are the upper and lower 95% confidence intervals. The narrow interval in the wet season shows increased confidence in the estimates compared to the dry season.**

### ***Rainfall type effect***


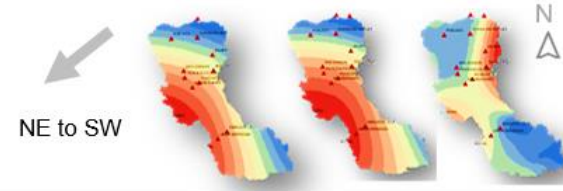
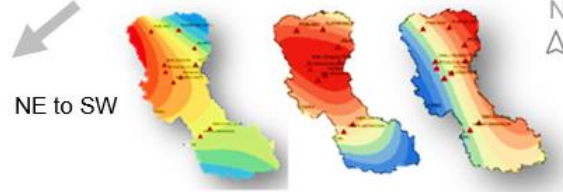
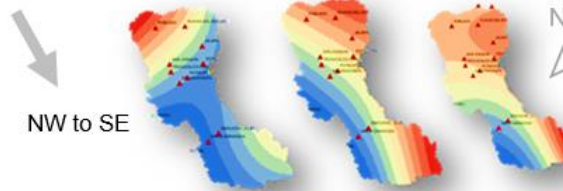
Rainfall types vary from place to place with respect to coverage of drainage areas and influence the magnitude of flood parameter values in the study area. Convective rainfall was assumed to be localised rainfall and result in partial coverage of drainage areas while frontal rainfall is widespread with complete coverage of drainage areas. Results show differences in the exponent ( $\theta$ ) and intercept ( $\alpha$ ) values related to whether rainfall is convective and frontal rainfall.

Results show that widespread rainfall events, particularly from tropical waves that usually cover the whole catchment for more prolonged duration results in higher exponent ( $\theta$ ) values (0.63), while localised convective storms that occur as localised patchy rainfall have low exponent ( $\theta$ ) values (0.40). A statistical *t*-test shows that the differences in the power-law relationship for the groupings for convective and frontal rainfall types are significant ( $P < 0.05$ ). Therefore, results show that widespread rainfall type of tropical wave origin results in high exponent ( $\theta$ ) values compared to the localised convective type of rainfall.

### ***Rainfall direction effect on exponent ( $\theta$ ) parameters***

Results show that the effect of rainfall direction on scaling parameters across all spatial scales in La Sierra catchment is significant. Analysis results of peak discharges for 2015 rainfall events only (Fig 5.16) showed that rainstorms moving from NE to SW (Storm number 22 and 23) crossing much of the catchment have high exponents ( $\theta$ ) values (0.60 and 0.63) both with a good model fit  $R^2 = 0.82$  and  $0.94$  respectively,

However, results show that storms that move from North-West to South-East crossing the catchment result in low exponents ( $\theta$ ) values (0.40 and 0.50) with a good model fit of  $R^2 = 0.85$  and  $0.92$  respectively, (Storm 19 and 28) (Figure 5.16). For example, on the 2nd of October 2015, storm number 22 moved from NW to SE crossing the central areas of the catchment resulted in some flood peaks at Tapijuluka, El Pueblo and Gaviota gauging station and this resulted in a relatively high ( $\theta=0.50$ ) scaling exponent ( $\theta$ ) parameter values. It is evident that in La Sierra catchment, exponents ( $\theta$ ) values are higher for rainfall events from NE than rainfall moving from NW. Therefore, results show that storm direction has some effect and contributes to the event to event changes of scaling parameters.

Storm number	Date of storm occurrence	Storm location and direction	Intercept ( $\alpha$ )	Exponent ( $\theta$ )	R <sup>2</sup>
19	15/06/2015	 NW to SE	10.80	0.40	0.85
22	25/09/2015	 NE to SW	9.33	0.60	0.82
23	02/10/2015	 NE to SW	6.62	0.63	0.94
28	22/12/2015	 NW to SE	12.24	0.50	0.92

**Key**

(Rainfall Hourly) mm

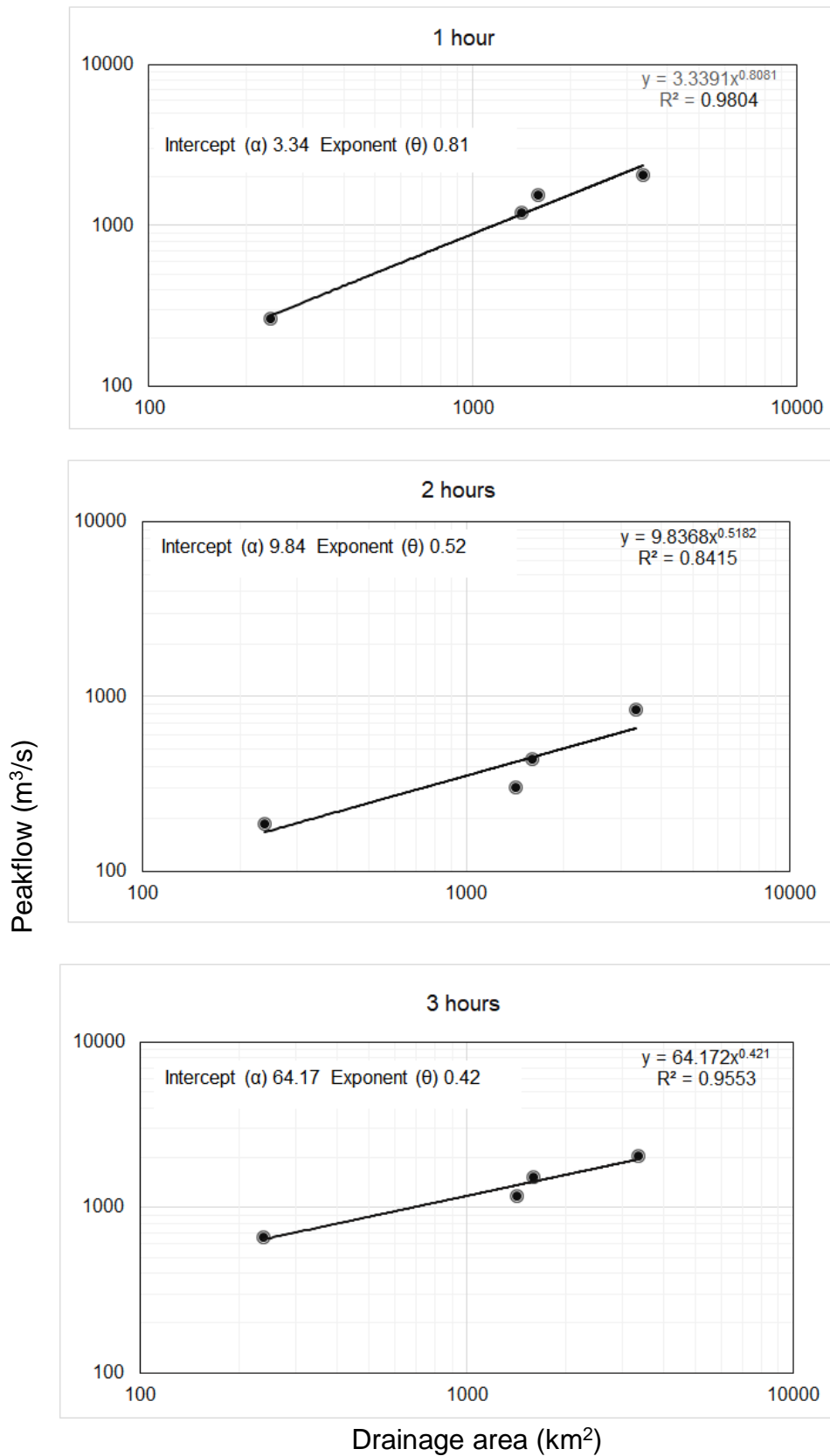
- 2.48 – 3.44
- 3.44 – 4.20
- 4.20 – 4.80
- 4.80 – 5.60
- 5.60 – 6.52
- 6.52 – 7.75
- 7.75 – 9.31
- 9.31 – 11.28
- 11.28 – 13.80
- 13.80 – 17.00

Figure 5.16: Effect of rainfall location and direction on flood scaling parameters. Storms moving from NE have high exponent ( $\theta$ ) parameters and low intercepts ( $\alpha$ ) while NW storms have low exponents ( $\theta$ ) and high intercept ( $\alpha$ ) values.



### ***Rainfall duration effect***

Simulations of peak discharge based on temporally varying synthetic rainfall of fixed amounts were enforced in a distributed hydrological model using four nested catchments of different sizes. The simulations were done to understand the effects of rainfall of varying durations on flood scaling parameters in the study area. Results from the experiment showed that rainfall duration affects scaling relationships, causing flood parameters to change from one rainfall event to another as rainfall duration changes (Figure 5.17). It was observed that peak discharge generated from one-hour duration rainfall has high exponent ( $\theta$ ) values (0.8) compared to 0.5 and 0.4 for two and three-hour rainfall duration, respectively. The results in figure 5.17 show that the exponent ( $\theta$ ) parameter values decrease with increasing rainfall duration, while intercepts ( $\alpha$ ) increase with increasing duration. Decreasing rainfall duration is related to decreases in intercept ( $\alpha$ ) values and increases exponent ( $\theta$ ) parameter values. The results confirm Gupta and Waymire (1998) findings that showed rainfall duration controls scaling parameters and determine the amount of overland flow that leads to peak floods.



**Figure 5.17: Results showing the relationships between simulated peak discharge and drainage areas based on the duration of rainfall events in La Sierra catchment area.**

### 5.3.3 Flood scaling parameter estimation for La Sierra catchment

The previous section found that spatial and temporal rainfall variability in nested catchments controls flood scaling parameters. Rainfall variability is related to event to event changes of the flood parameters, however in interaction with other catchment conditions and processes. The relationships between these physical processes can be used to estimate representative flood parameters and then the magnitude of peak discharge in the La Sierra catchment. Rainfall variability alone would not be adequate to provide flood estimates, and there is need to find significant relationships between scaling parameters and other measurable catchment processes and conditions and develop statistical models for estimating first, the flood scaling parameters, (the exponents ( $\theta$ ) and intercept ( $\alpha$ )) and then the magnitude of peak discharges in sub-catchments using the flood scaling equation (Equation 4.24). This section presents the significant relationships identified that could be used to find two representative flood parameters for the catchment. Also presented are the results of using the estimated flood parameters in the flood scaling equation (Equation 4.24) to estimate the peak discharge across nested catchment areas.

A 95% predictive interval is given with estimates to show where the actual parameter values are expected, while confidence intervals show the precision of the regression model in its estimation (Helsel and Hirsch, 2002). Prediction and confidence intervals provide information on the natural variability bounds for the estimates of flood parameter and peak discharge and therefore express the uncertainty in model results. A narrow prediction interval increases confidence in the result, while a broad interval shows the existence of large uncertainties and hence low confidence in the estimates shown.

The Pearson correlation analysis was employed to investigate the strength of linear associations between scaling exponents ( $\theta$ ) and intercept ( $\alpha$ ) parameters and catchment processes and conditions in the La Sierra catchment. Results in Table 5.10 show some strong relationships used with significance at  $p= 0.005$ . The significant statistical relationships identified can potentially be used to predict the flood parameter values for the catchment and further be used to estimate the magnitude of peak discharge using the scaling equation (equation 4.11).

**Table 5:10 Pearson correlation analysis results showing significant and non-significant relationships between the exponents ( $\theta$ ) and intercept ( $\alpha$ ) (flood parameters) and other measurable catchment processes and conditions ( independent variables) in La Sierra catchment.**

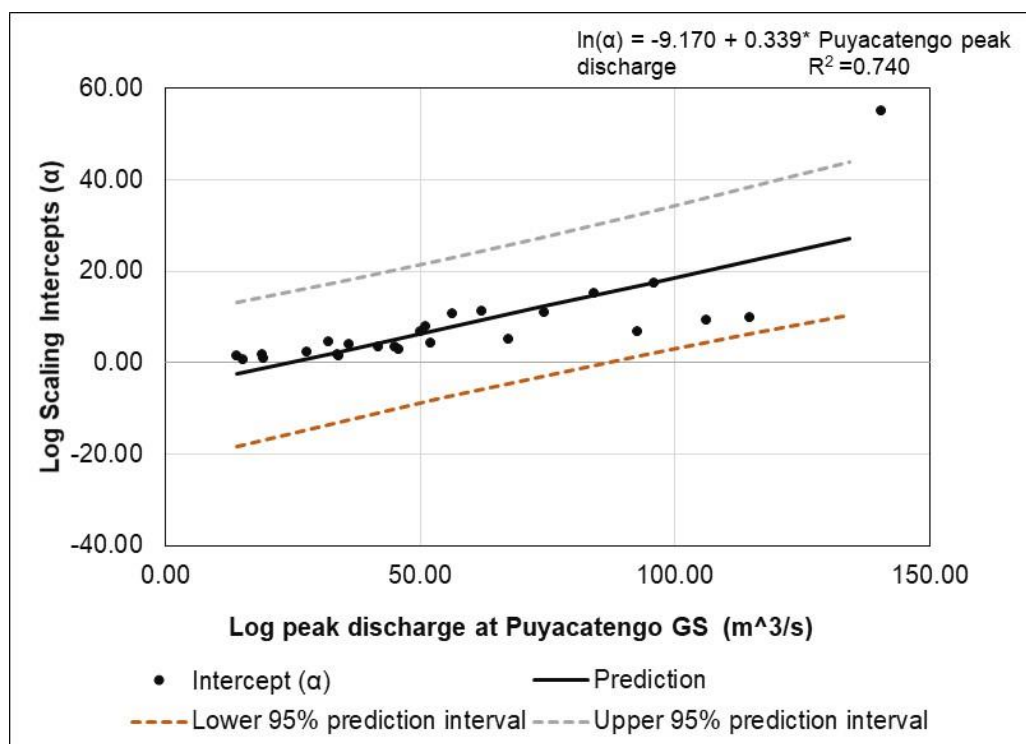
Correlations		Gaviota peakflow	El Pueblo peakflow	Teapa peakflow	Puyacaten go peakflow	Intercept ( $\alpha$ )	Exponent ( $\theta$ )	Gaviota average rainfall	Pueblo average rainfall	Oxolotan average rainfall	Gaviota rainfall acc.	Pueblo rainfall acc.	Teapa rainfall acc.	2hr oxolotan peakflow	Puyacaten go Stage (m)	Storm location
Gaviota peakflow	Pearson Correlation	1.00	.858**	.556**	0.14	-0.17	.430**	.587**	.587**	.473**	0.26	0.26	.332*	0.22	0.61	0.13
	Sig. (2-tailed)		0.00	0.00	0.37	0.21	0.00	0.00	0.00	0.00	0.09	0.09	0.03	0.63	0.06	0.32
El Pueblo peakflow	Pearson Correlation	.858**	1.00	.504**	0.21	-0.12	.319*	.414**	.414**	.314*	0.12	0.12	0.18	0.05	0.30	.266*
	Sig. (2-tailed)	0.00		0.00	0.15	0.38	0.01	0.00	0.00	0.04	0.43	0.43	0.24	0.91	0.37	0.04
Teapa peakflow	Pearson Correlation	.556**	.504**	1.00	.309*	.350**	-0.04	.531**	.531**	0.29	0.08	0.08	0.15	-0.40	.900**	.309*
	Sig. (2-tailed)	0.00	0.00		0.04	0.01	0.76	0.00	0.00	0.05	0.58	0.58	0.31	0.33	0.00	0.02
Puyacaten go peakflow	Pearson Correlation	0.14	0.21	.309*	1.00	.861**	-.709**	0.15	0.15	-0.02	-0.15	-0.15	-0.16	-0.51	.770**	.796**
	Sig. (2-tailed)	0.37	0.15	0.04		0.00	0.00	0.35	0.35	0.91	0.37	0.37	0.34	0.20	0.01	0.00
Intercept ( $\alpha$ )	Pearson Correlation	-0.17	-0.12	.350**	.861**	1.00	-.735**	0.08	0.08	-0.12	-0.24	-0.24	-0.25	-0.49	.635*	.524**
	Sig. (2-tailed)	0.21	0.38	0.01	0.00		0.00	0.61	0.61	0.43	0.11	0.11	0.10	0.22	0.04	0.00
Exponent ( $\theta$ )	Pearson Correlation	.430**	.319*	-0.04	-.709**	-.735**	1.00	.306*	.306*	.438**	.313*	.313*	.335*	.791*	-0.43	-.543**
	Sig. (2-tailed)	0.00	0.01	0.76	0.00	0.00		0.04	0.04	0.00	0.03	0.03	0.02	0.02	0.19	0.00
Gaviota average rainfall	Pearson Correlation	.587**	.414**	.531**	0.15	0.08	.306*	1.00	1.000**	.844**	.384**	.384**	.409**	-0.32	.744*	0.04
	Sig. (2-tailed)	0.00	0.00	0.00	0.35	0.61	0.04		0.00	0.00	0.01	0.01	0.00	0.53	0.02	0.78
Pueblo average rainfall	Pearson Correlation	.587**	.414**	.531**	0.15	0.08	.306*	1.00	1.00	.844**	.384**	.384**	.409**	-0.32	.744*	0.04
	Sig. (2-tailed)	0.00	0.00	0.00	0.35	0.61	0.04	0.00		0.00	0.01	0.01	0.00	0.53	0.02	0.78
Oxolotan average rainfall	Pearson Correlation	.473**	.314*	0.29	-0.02	-0.12	.438**	.844**	.844**	1.00	.359*	.359*	.353*	-0.23	0.65	-0.06
	Sig. (2-tailed)	0.00	0.04	0.05	0.91	0.43	0.00	0.00	0.00		0.01	0.01	0.02	0.65	0.06	0.69
Gaviota 3 day rainfall acc.	Pearson Correlation	0.26	0.12	0.08	-0.15	-0.24	.313*	.384**	.384**	.359*	1.00	1.000**	.958**	-0.13	0.07	-0.08
	Sig. (2-tailed)	0.09	0.43	0.58	0.37	0.11	0.03	0.01	0.01	0.01		0.00	0.00	0.81	0.85	0.60
Pueblo 3 rainfall acc.	Pearson Correlation	0.26	0.12	0.08	-0.15	-0.24	.313*	.384**	.384**	.359*	1.000**	1.00	.958**	-0.13	0.07	-0.08
	Sig. (2-tailed)	0.09	0.43	0.58	0.37	0.11	0.03	0.01	0.01	0.01	0.00		0.00	0.81	0.85	0.60
Teapa 3 rainfall acc.	Pearson Correlation	.332*	0.18	0.15	-0.16	-0.25	.335*	.409**	.409**	.353*	.958**	.958**	1.00	-0.23	0.12	-0.09
	Sig. (2-tailed)	0.03	0.24	0.31	0.34	0.10	0.02	0.00	0.00	0.02	0.00	0.00		0.66	0.76	0.57
2hr duration oxolotan peak	Pearson Correlation	0.22	0.05	-0.40	-0.51	-0.49	.791*	-0.32	-0.32	-0.23	-0.13	-0.13	-0.23	1.00	-0.67	-0.54
	Pearson Correlation	0.63	0.91	0.33	0.20	0.22	0.02	0.53	0.53	0.65	0.81	0.81	0.66		0.07	0.17
Puyacaten go Stage (m)	Pearson Correlation	0.61	0.30	.900**	.770**	.635*	-0.43	.744*	.744*	0.65	0.07	0.07	0.12	-0.67	1.00	0.42
	Sig. (2-tailed)	0.06	0.37	0.00	0.01	0.04	0.19	0.02	0.02	0.06	0.85	0.85	0.76	0.07		0.20
Storm location	Pearson Correlation	0.13	.266*	.309*	.796**	.524**	-.543**	0.04	0.04	-0.06	-0.08	-0.08	-0.09	-0.54	0.42	1.00
	Sig. (2-tailed)	0.32	0.04	0.02	0.00	0.00	0.00	0.78	0.78	0.69	0.60	0.60	0.57	0.17	0.20	

\*\* . Correlation is significant at the 0.01 level (2-tailed).

\* . Correlation is significant at the 0.05 level (2-tailed).

### 5.3.4 Intercept estimation from observed peak discharge

All scaling intercept ( $\alpha$ ) values derived from the power-law relationships between peak discharge (from 59 rainfall-runoff events occurred between 2012 and 2015) and nested catchment drainage areas were correlated with other measurable catchment processes that occurred during the same period. Results show a strong ( $r = 0.86$ ) positive relationship between the scaling intercept ( $\alpha$ ) values and peak discharge observed at Puyacatengo Gauging Station (Table 5.10). Due to this strong association, it was possible to use the relationship identified to predict the scaling intercept ( $\alpha$ ) parameter values for the catchment. To establish a predictive relationship, the natural logarithm of the scaling intercept ( $\alpha$ ) values was regressed on the natural logarithm of the peak discharge observed at Puyacatengo Gauging Station (Figure 5.18). Results show that, given  $R^2 = 0.74$  by the model, around 74% of the variability of the predicted intercept ( $\alpha$ ) values could be explained by the model, using observed peak discharge at Puyacatengo GS as a predictor variable.



**Figure 5.18: Scatter plot of the natural logarithm of the scaling intercept ( $\alpha$ ) against the natural logarithm of peak discharge observed at Puyacatengo Gauging Station. The 95% prediction interval estimates where the actual intercept ( $\alpha$ ) values are expected along the best fit line.**

The log-log relationship between scaling intercept ( $\alpha$ ) and observed peak discharge from Puyacatengo Gauging Station (the smallest nested catchment in the La Sierra catchment) (Table 5.11) and all coefficients are significant ( $P < 0.05$ ). Results show that peak discharge observed at this station can significantly predict intercept ( $\alpha$ ) parameter values for the study area using the log-log equation with an intercept ( $\alpha$ ) value of -9.17 ranging from -14.00 to -4.34 and the estimated mean slope value of 0.34 varying from 0.25 to 0.42 at 95% confidence level (Table 5.11).

**Table 5:11 Results of the log-log relationship between Intercept ( $\alpha$ ) parameters and Puyacatengo peak discharge.**

*Model parameters (Intercept ( $\alpha$ )):*

Source	Value	Standard error	t	Pr >  t	Lower bound (95%)	Upper bound (95%)
ln(Intercept ( $\alpha$ ))	-9.170	2.399	-3.822	< 0.003	-14.002	-4.338
ln(Puyacatengo)	0.339	0.042	7.995	< 0.001	0.254	0.424

The regression result shows that peak discharge observed at Puyacatengo Gauging Station can estimate intercept ( $\alpha$ ) parameter values for La Sierra catchment using the log-log regression model as summarised in equation 5.1.

$$\ln \text{Intercept } (\alpha) = -9.170 + 0.340 * \ln \text{Puyacatengo peak flows}$$

**Equation 5.1**

A student's t-test on the coefficients of the equation shows that all coefficients are statistically significant at a 95% confidence interval. Checks of the Q-Q plot of the residuals between intercepts ( $\alpha$ ) and observed discharge at Puyacatengo Gauging Station and the Shapiro-Wilk test for normality  $W = 0.96$  and  $p = 0.0067$ , show the residuals are normally distributed. Therefore, the results confirm the model's robustness and show that the scaling intercept ( $\alpha$ ) for the La Sierra catchment can be calculated using the peak discharge observed from the smallest nested catchment.

### 5.3.5 Exponent ( $\theta$ ) estimation using Intercept ( $\alpha$ ) values

The correlation results show a strong linear negative relationship ( $r = -0.74$ ) between scaling intercepts ( $\alpha$ ) and exponents ( $\theta$ ) values showing that 74% decrease in the exponent ( $\theta$ ) parameter values could be explained by increases in intercept ( $\alpha$ ) values (Table 5.10). The scaling exponent ( $\theta$ ) values were regressed on the natural logarithm of the scaling intercept ( $\alpha$ ) values (Figure 5.19). Results show that, given the regression model results with  $R^2 = 0.54$ , 54% of the variability of the dependent scaling exponent ( $\theta$ ) variables can be predicted using the explanatory intercept ( $\alpha$ ) variables on a log-linear relationship.

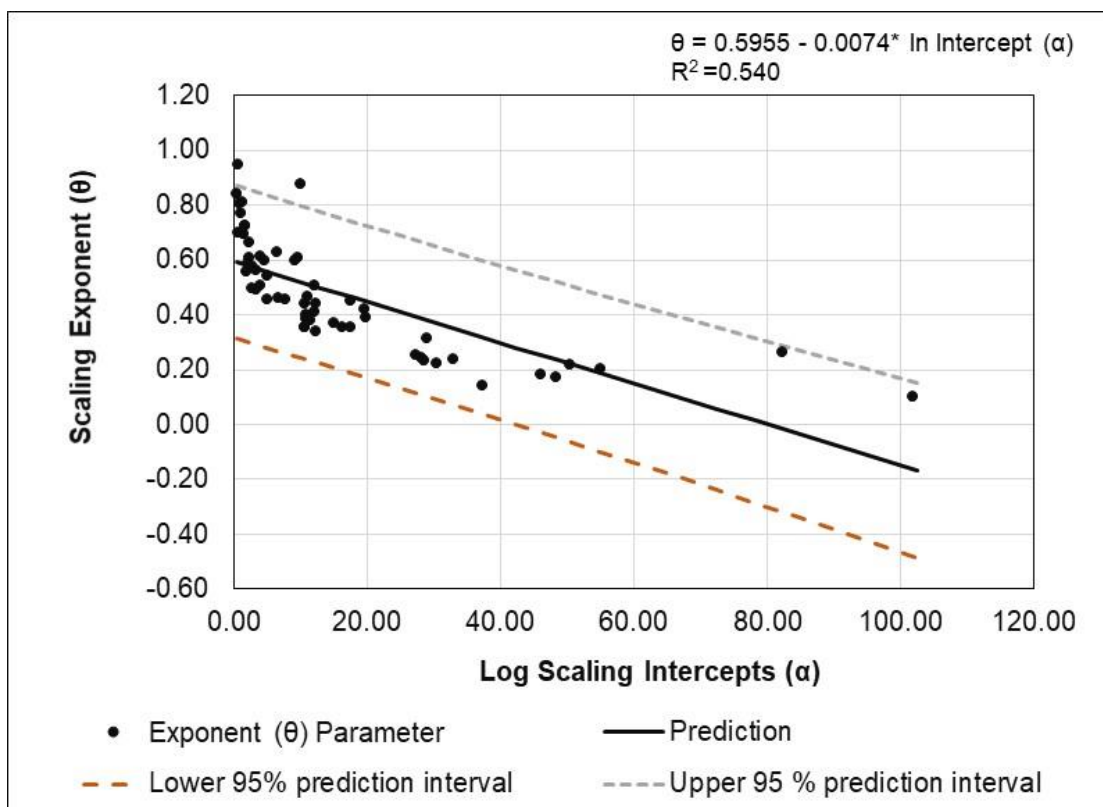


Figure 5.19: Scatter plot of the scaling exponent ( $\theta$ ) against the natural logarithm of the scaling intercept ( $\alpha$ ) (Adapted from Ayalew et al., 2018).

The results in Table 5.12 shows the statistical significance of the log-linear relationship between the scaling exponents ( $\theta$ ) and intercept ( $\alpha$ ) parameters with all coefficients being significant ( $P < 0.05$ ). It is shown that intercept ( $\alpha$ ) parameters can significantly predict scaling exponents ( $\theta$ ) parameter values for the study area using the log-linear

equation with an intercept ( $\alpha$ ) value of 0.60 ranging from 0.55 to 0.64 and the estimated mean slope value of -0.01 varying from -0.01 to 0.06 at 95% confidence level (Table 5.12). The results confirm the significance of the relationship between exponents ( $\theta$ ) and intercept ( $\alpha$ ) parameters as reported by Ayalew et al. (2015) that it can be used to predict exponents ( $\theta$ ) values across nested catchment areas.

**Table 5:12 Results of the linear regression analysis between Exponents ( $\theta$ ) and Intercept ( $\alpha$ ) parameters.**

Source	Value	Standard error	t	Pr >  t	Lower bound (95%)	Upper bound (95%)
Exponent ( $\theta$ )	0.596	0.023	25.947	< 0.0012	0.550	0.641
Intercept ( $\alpha$ )	-0.007	0.001	-8.184	< 0.0014	-0.009	-0.006

The resulting overall log-linear regression model is summarised in the following equation.

$$\text{Exponent } (\theta) = 0.596 - 7.426E - 03 * \ln \text{Intercept } (\alpha)$$

**Equation 5.2**

A student's t-test on the coefficients of the equation 5.2 shows that at 95% confidence interval all coefficients of the equation are statistically significant and the Shapiro-Wilk test for normality  $W=0.89$  and  $P =0.0063$  shows that the residuals are normally distributed. The relationships are significant ( $P < 0.05$ ), and it is possible to use the log-linear relationship to predict exponent ( $\theta$ ) values using intercept ( $\alpha$ ) values obtained in section 15.3.4. An examination of residuals between the intercepts ( $\alpha$ ) and exponent ( $\theta$ ) parameters was found to be normally distributed on Q-Q plot. Therefore, this study has developed a valid log-linear relationship between the scaling exponent ( $\theta$ ) and the intercept ( $\alpha$ ) that can be used to estimate the scaling exponent ( $\theta$ ) for the La Sierra catchment.

**5.3.6 Intercept ( $\alpha$ ) multi regressed with peak discharge**

Besides using two variables to predict flood parameters, a multiple regression analysis was used to estimate the intercept ( $\alpha$ ) and exponent ( $\theta$ ) parameters, using several



exploratory variables in the La Sierra catchment area. It is a further analysis to find a range of alternative significant relationships/models from which to select the best model for estimating intercept ( $\alpha$ ) parameters. Results show the intercept ( $\alpha$ ) parameter values have a strong negative correlation with the observed discharge at Puyacatengo GS (-0.77) and a slightly strong (0.35) with discharge at Teapa GS (Table 5.13).

A stepwise forward elimination procedure adopted shows that all coefficients are significant ( $P < 0.05$ ). Results show a significant relationship exists when the natural logarithm of the intercept ( $\alpha$ ) parameter values are multi regressed on the natural logarithm of discharge observed at Puyacatengo and Teapa Gauging Stations. The coefficient of the log-linear relationship was significant ( $P < 0.05$ ) in predicting the intercept ( $\alpha$ ) parameter for the catchment (Table 5.13).

**Table 5:13 Results from multilinear regression analysis of relationships between Intercept ( $\alpha$ ) parameters, Puyacatengo and Teapa peak discharge.**

Model	Unstandardised Coefficients		Standardised Coefficients		t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta				Zero-order	Partial	Part	Tolerance	VIF
ln(Constant)	0.667	0.04			17.73	0.004					
ln(Puyac-)	-0.003	0.00	-0.81		-8.03	0.003	-0.73	-0.77	-0.77	0.91	1.11
ln(Teapa)	0.0001	0.00	0.27		2.65	0.01	0.02	0.38	0.26	0.91	1.11

The resulting log-multilinear regression model is summarised as:

$$\ln \text{Intercept} = 0.667 - 0.003 \ln \text{Puyaca peak flows} + 0.0001 \ln \text{Teapa flows}$$

**Equation 5.3**

Checks on residuals were found to be normally distributed on the Q-Q plot and the regression line and were found to be approximately normally distributed. A Student's  $t$ -test on the coefficients of the equation 5.3 shows that Puyacatengo and Teapa peak discharge values can significantly ( $P < 0.05$ ) predict the intercept ( $\alpha$ ) parameter values for the catchment. It was therefore concluded that the relation between the observed peak discharges at the two stations could be used to predict intercept ( $\alpha$ ) parameters. Checks for multicollinearity in the model using collinearity statistics shows that it is 1.1, which is within the tolerance of  $> 0.1$  or  $VIF < 10$ .

### 5.3.7 Exponents ( $\theta$ ) multi regressed with peak discharge

This section presents the analysis results of significant relationships between exponent ( $\theta$ ) parameter values and other observable catchment processes and conditions. A strong positive correlation was found between exponent ( $\theta$ ) and Puyacatengo flows (0.74) and a moderate correlation with Gaviota flows (0.48) (Table 5.10). The natural logarithm of the scaling exponent ( $\theta$ ) values was multi regressed on the natural logarithm of peak discharge observed at Puyacatengo and Gaviotas Gauging Stations. Results from a stepwise forward elimination procedure on the multi-regression model show that the coefficients of the model were each statistically significant at ( $p < 0.05$ ) (Table 5.14).

**Table 5:14 Results of the multilinear regression analysis between Puyacatengo, Gaviota peak discharge and exponent ( $\theta$ ) parameter.**

Model	Unstandardised Coefficients		Standardised Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	0.493	0.03		16.34	0.002		
Puyacatengo Q	-0.003	0.05	-0.79	-12.60	0.001	0.98	1.02
Gaviota Q	0.0001	0.07	0.54	9.36	0.006	0.98	1.02

The equation 5.4 below summarises the resulting log-multilinear regression model:

$$\ln Exponent = 0.493 - 0.003 \ln Puyaca\ peakflow + 0.0001 * Gaviota\ peakflows$$

**Equation 5.4**

Results suggest that the log-log relationship between observed discharge at Puyacatengo and Gaviota GS can be used to approximate exponent ( $\theta$ ) parameter values for the study area. All coefficients in equation 5.4 are statistically significant ( $P < 0.05$ ). Test results show that the overall exponent ( $\theta$ ) parameter value can be significantly predicted from observed discharge at Puyacatengo and Gaviota GS. Checks for multicollinearity in the model using collinearity statistics shows that it is at 1.02 with a tolerance of 0.98, with  $> 0.1$  or  $VIF < 10$ .

### 5.3.8 Significant statistical models developed

Six statistical models have been developed to describe the significant relationships between intercepts ( $\alpha$ ) and exponents ( $\theta$ ) and catchment properties and processes and establish statistical models for estimating flood parameters for La Sierra catchment (Table 5.15). From the six models identified, four are for exponent ( $\theta$ ) parameter estimation with an average coefficient of determination,  $R^2 = 0.57$ , about 57% of the variability of exponents ( $\theta$ ) parameter values can be predicted by the model. Model 4 and 5 were identified as valid models for estimating intercept ( $\alpha$ ) parameter values with a good fitness of fit coefficient ( $R^2 = 0.62$ ). Given the number of models identified, there was a need to select the best models and prevent the use of models that overfit data and that do not work well in new contexts.

**Table 5:15 Statistical models developed for estimating scaling parameters.**

Model number	Statistical models	MSE	R <sup>2</sup>	Adjusted R <sup>2</sup>
1	Exponent ( $\theta$ ) = 0.595530584044186-7.42631001955593E-03*ln Intercept ( $\alpha$ )	0.019	0.540	0.532
2	Exponent ( $\theta$ ) = 0.698987330308686-3.00574514224434E-03*lnPuyacatengo peak discharge	0.017	0.503	0.492
3	Exponent ( $\theta$ ) = 0.31727972290143+3.21006324253182E-04*ln2hr Oxolotan peak discharge	0.011	0.625	0.563
4	ln Intercept ( $\alpha$ ) = 0.667 – 0.03ln(Puyaca peak discharge+0.001*lnTeapa peak discharge	79.648	0.690	0.676
5	ln Intercept ( $\alpha$ ) = -9.30434442210789+0.340025349111949*lnPuyacatengo peak discharge	78.188	0.542	0.536
6	Exponent ( $\theta$ ) = 0.493 – 0.03lnPuyaca peak discharge +0.0001*Gaviota peak discharge	81.188	0.545	0.555

### 5.3.9 Model selection

Model selection was done using XLSTAT 2019.1 software equipped with the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) estimators (Addinsoft, 2019). The software was used for selecting the best two statistical models for estimating flood exponents ( $\theta$ ) and intercepts ( $\alpha$ ) for the catchment. Model selection results show that model number 1 has the lowest BIC (-228.14) and AIC (- 232.30) and was selected as the best log-linear regression model for estimating exponents ( $\theta$ ) parameter values La Sierra catchment (Table 5.16). The best model for estimating intercept ( $\alpha$ ) parameter values showed that model number 5 is the best with the lowest

BIC (113.79) and AIC (111.36). Overall, results show that linear models have the lowest BIC and AIC values and were better than multi-variant models developed. Therefore, the two models (the log-log and the log-linear) were selected for parameter and discharge estimation for the catchment, with the added advantage of their simplicity.

**Table 5:16 Model selected based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) estimators.**

No	Overall models	Model selection metrics			
		MSE	R <sup>2</sup>	AIC	BIC
1	Exponent ( $\theta$ ) = 0.595530584044186-7.42631001955593E-03*ln Intercept ( $\alpha$ )	0.02	0.54	-232.30	-228.14
2	Exponent ( $\theta$ ) = 0.698987330308686-3.00574514224434E- 03*lnPuyacatengo peak discharge	0.02	0.50	-192.54	-188.80
3	Exponent ( $\theta$ ) = 0.31727972290143+3.21006324253182E-04*ln2hr Oxolotan peak discharge	0.01	0.63	-34.16	-34.00
4	ln Intercept ( $\alpha$ ) = 0.667 – 0.03ln(Puyaca peak discharge+0.001*lnTeapa peak discharge	79.65	0.69	211.20	214.94
5	ln Intercept ( $\alpha$ ) = 9.30434442210789+0.340025349111949 *lnPuyacatengo peak discharge	78.19	0.54	111.36	113.79
6	lnExponent = 0.493 – 0.03ln(Puyaca peak discharge +0.0001*Gaviota peak discharge	81.19	0.55	215.11	274.98

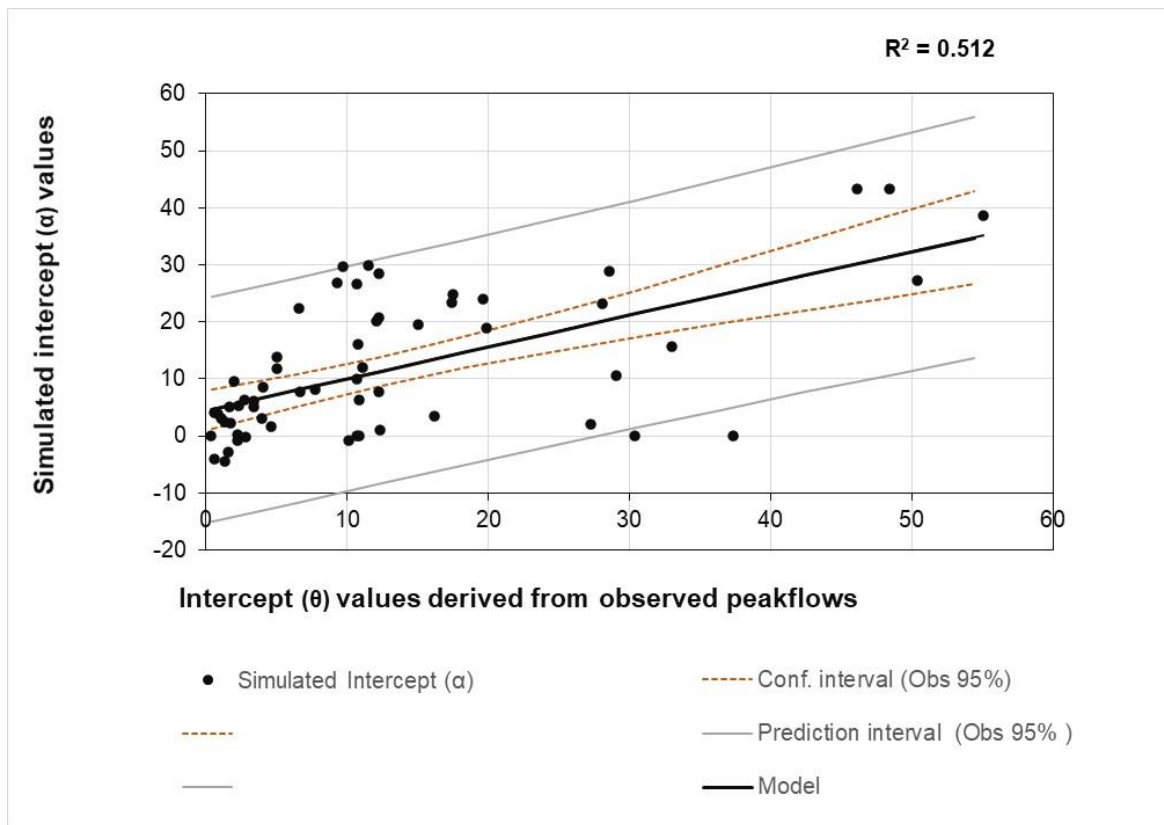
### 5.3.10 Application of the results of the flood simulation framework

A flood prediction framework comprising statistical models developed for estimating flood parameters and peak discharge in any part of La Sierra catchment is presented. The following method used by Ayalew et al. (2014) was followed to predict the expected value of flood parameters and peak discharge in a selected nested drainage area. A log-log regression model presented in equation 5.1 (Selected model # 5) was used to calculate the scaling intercept ( $\alpha$ ) from the observed peak discharge at Puyacatengo gauging station. A log-linear regression model presented in equation 5.2 was used to estimate the corresponding scaling exponent ( $\theta$ ) from the natural logarithm of the scaling intercept ( $\alpha$ ) estimated in step 2. Finally, the estimated values for the flood exponents ( $\theta$ ) and intercepts ( $\alpha$ ) and Teapa drainage area were incorporated into the flood scaling equation (Equation 4.24) to estimate the expected magnitude value of peak discharge for Teapa catchment.

The steps presented provide the estimates of the expected values of flood parameters and peak discharge across a range of spatial scales (nested catchments) in La Sierra catchment. Teapa sub-catchment was taken as an example in the implementation of the flood prediction framework. The section presents a comparison of results of modelled/simulated estimates and observed parameter values derived from observed peak discharge from 59 rainfall-runoff events. The results are presented using prediction and confidence intervals to express the natural variability bounds for the estimates of flood parameter and peak discharge. The 95% prediction intervals provide the spectrum of individual discharge and parameter values in which they often vary with confidence (Helsel and Hirsch, 2002). Also, it shows the probability range where the accurate linear regression estimates lie with confidence along the regression line.

### ***Results from the application of the log-log model for estimating intercept ( $\alpha$ ) parameter values***

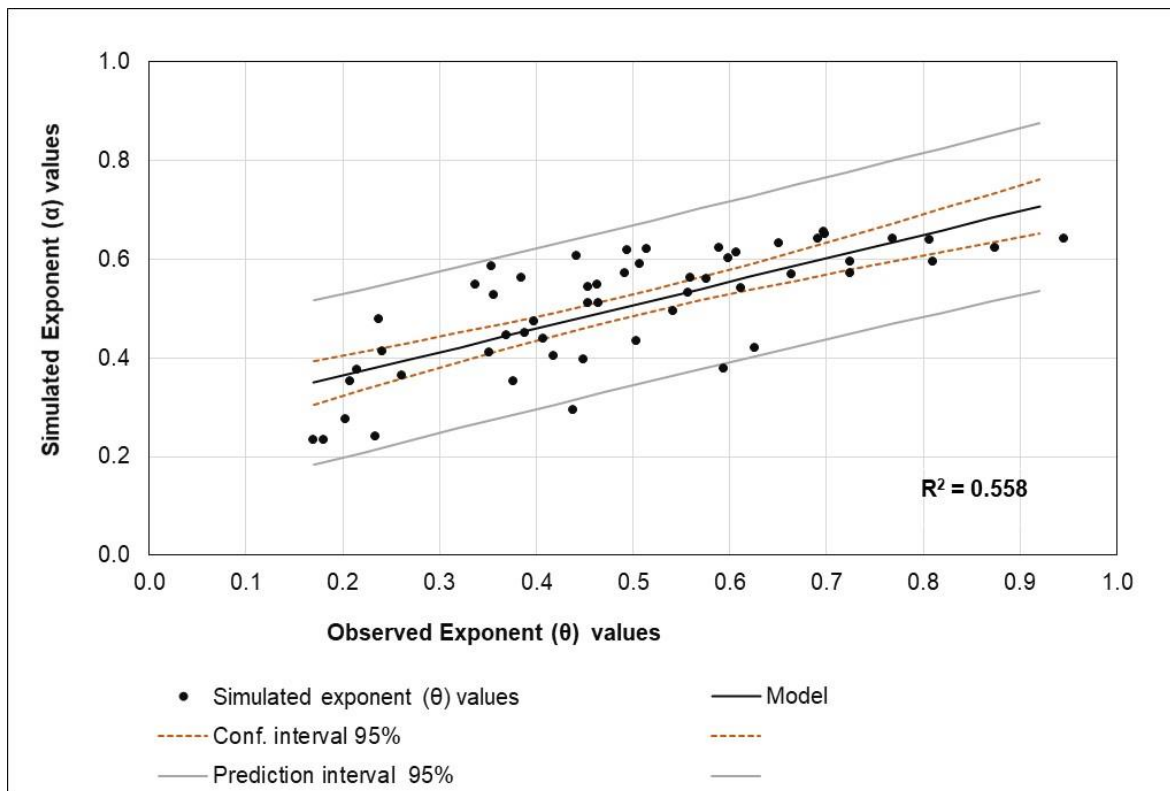
To demonstrate the application of the log-log model, the log-log relationship between the intercept ( $\alpha$ ) and the observed peak discharge from Puyacatengo Gauging Station was used to estimate the intercept ( $\alpha$ ) parameter value for a selected nested catchment area. The Teapa nested catchment was selected and used as an example, and its drainage area was included in the flood scaling equation (Equation 4.24). Comparison results between estimated intercepts ( $\alpha$ ) and observed intercepts ( $\alpha$ ) values derived from observed peak discharge from 2012 to 2015 rainfall-runoff events are shown in Figure 5.20. It can be seen that estimated parameter values can mirror changes in the observed parameter values with a good model fit of  $R^2 = 0.51$  showing that the model can reasonably predict 51% of the variability of intercepts ( $\alpha$ ) parameter values for the catchment. However, the model shows large errors in estimating intercept ( $\alpha$ ) values with a Root Mean Square Error (RMSE) of 11.36 mm, a Mean Square Error (MSE) of 129.13 mm and Percentage Bias (PBIAS) of -10.39% showing underestimation in its predictions.



**Figure 5.20: Comparison of observed and expected scaling Intercept ( $\alpha$ ) parameter values from 2012 to 2015 rainfall-runoff events. The grey dotted lines show the 95% prediction interval range where actual value can be predicted, and the orange line is the confidence interval showing precision of the regression model.**

### ***Results from the application of the log-linear model to estimation exponent ( $\theta$ ) parameter value***

To estimate exponent ( $\theta$ ) parameters for the catchment, the selected log-linear regression model presented in equation 5.2 was employed. The exponent ( $\theta$ ) values estimated were compared to the observed exponent ( $\theta$ ) values derived from observed peak discharges from 2012 to 2015 rainfall-runoff events (Figure 5.21). Results show a good model fit ( $R^2 = 0.56$ ), showing the model capacity to predict 56% of the variability of exponent ( $\theta$ ) parameter values. That means about 56% of the dependent variable, the exponent ( $\theta$ ) values can be predicted by the independent observed variables used. Overall, the results show the skill of the model (equation 5.2) in estimating the value of flood exponents ( $\theta$ ) in La Sierra catchment. However, the model shows large errors in estimating intercept ( $\alpha$ ) values with a Root Mean Square Error (RMSE) of 12.57 mm, a Mean Square Error (MSE) of 158.09 mm, and Percentage Bias (PBIAS) of 2.08% showing overestimation in its predictions.

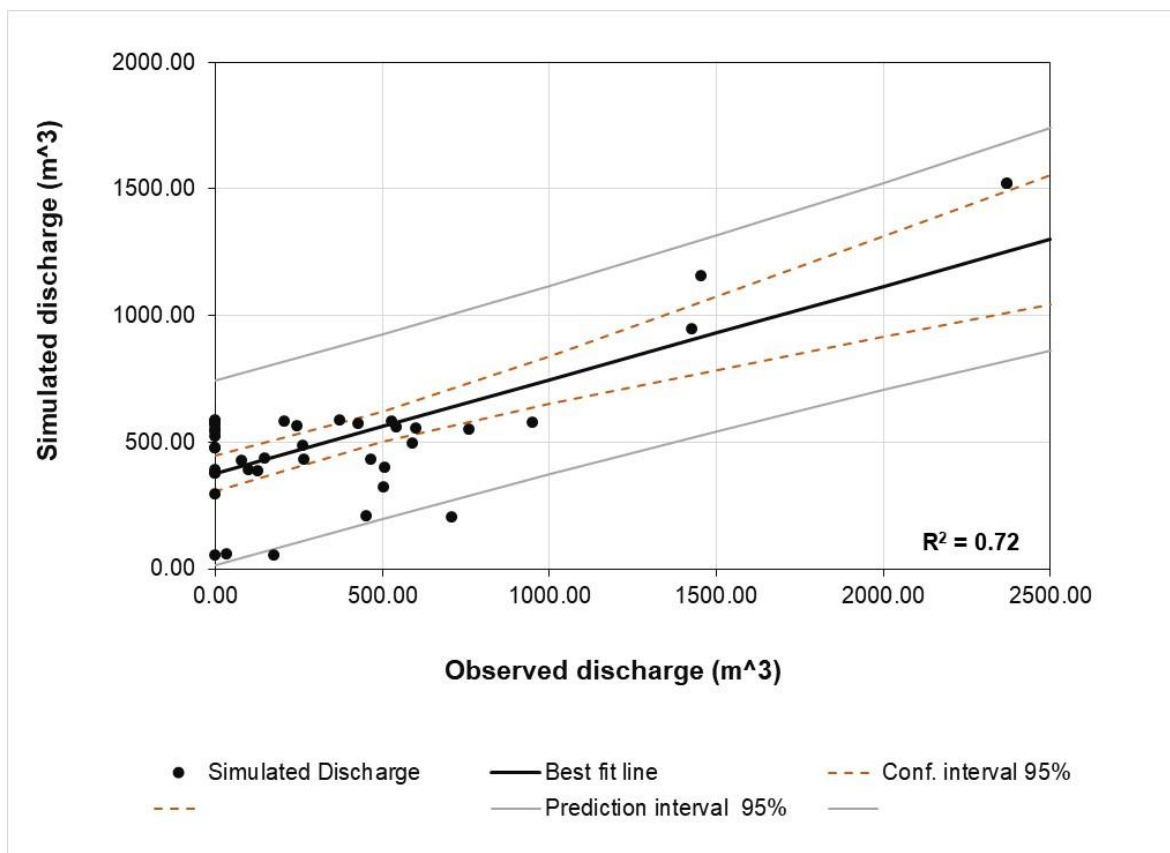


**Figure 5.21: Comparison of estimated exponent ( $\theta$ ) and observed values from rainfall-runoff events occurred from 2012 to 2015. The grey dotted lines show the 95% prediction interval range where actual value can be predicted, and the orange line is the confidence interval showing precision of the regression model.**

***Results from the application of the flood scaling equation to estimate the magnitude of peak discharge***

The estimated values for the flood exponents ( $\theta$ ) and intercepts ( $\alpha$ ) and Teapa drainage area were incorporated into the flood scaling equation (Equation 4.24) to estimate the expected magnitude value of peak discharge in the catchment. Figure 5.22 shows comparison results of the observed and predicted peak discharges in Teapa sub-catchment with a good correlation coefficient (0.79), good Nash-Sutcliffe Efficiency (0.60) however with a large Root Mean Square Error RMSE of 9.04 mm. The results between observed Teapa and simulated peak discharge shown on a scatter plot depicted a relatively strong relationship ( $R^2 = 0.72$ ) (Figure 5.22). The results show that peak discharges in La Sierra catchments can be reasonably be estimated from discharges observed in small headwater catchments, however, with some errors.

The results presented in Figure 5.22 show that the estimated peak discharges exhibit variability around their expected value, shown as the best fit line. The uncertainty (natural variability) of the predicted peak discharges was quantified using a Helsel and Hirsch, (2002) method for quantifying uncertainty. The prediction interval was used to estimate the variability (uncertainty) of peak discharges around their expected value. The intervals (grey line in Figure 5.22) set the natural variability bounds for peak discharge predictions. Also, the 95% confidence interval over which discharge usually varies in the study area is displayed using the orange dotted line and was determined. The confidence interval shows higher accuracy of the model in estimating low-level discharge magnitudes relative to larger quantities. The 95% prediction interval show a range of predictive interval that is relatively wide, showing large errors or uncertainties in model results. Though the model presents large uncertainties in estimating large flood magnitudes, there is high confidence in its ability to estimate small-scale magnitudes of peak discharge.



**Figure 5.22: Comparison results between simulated discharge and observed discharge in La Teapa sub-catchment.**



## 5.4 Summary

The study revealed that peak discharge in La Sierra catchment exhibits power-law relationships with drainage areas similar to findings from humid temperate regions. Results further show that the power-law relationship between discharge and drainage is valid even in conditions of partial rainfall coverage. This study statistically quantified the effect of rainfall variability on flood scaling parameters (the exponents ( $\theta$ ) and intercepts ( $\alpha$ )). The results showed the nature of the scaling relationships between flood parameters and spatial and temporal variability of rainfall in a tropical catchment. The effect of tropical rainfall variability on flood scaling parameters and flood peak generation was quantified and found to be significant.

Also, the results showed the nature of the dependence of flood parameters on the spatial and temporal variability of other catchment processes and conditions (rainfall included) and established statistical relationships that can be used to estimate flood parameter values and the magnitude of peak discharges in the La Sierra catchment area. Results have shown that the log-linear relationship between the intercept ( $\alpha$ ) and the exponent ( $\theta$ ) can reasonably estimate flood parameters and the log-log relationship between the intercept ( $\alpha$ ) parameter and the peak discharge observed from the smallest nested catchments can be used for estimating intercept ( $\alpha$ ) values for the study area. Peak discharges in the catchment areas can reasonably be estimated from discharge observed in small headwater catchments with short-term rainfall responses. The study showed the potential application of the estimated scaling parameters factored in the power-law equation along with the catchment drainage areas to estimate the magnitude of peak discharge in nested catchments.

## **5.5 Uncertainty propagation and analysis results**

This section presents results of rainfall and parameter uncertainty propagation and quantification through coupled hydrological and hydraulic models and the effects on discharge and flood inundation extents (Yu et al., 2016). First presented are the results from probability distribution fitting of rainfall time series and the generation of rainfall ensembles. The results for the hydraulic model calibration to find the optimal Manning coefficients of roughness values are presented. Also, presented are the results from ensemble model run scenarios based on fixed or calibrated parameters and the variations of parameter values.

The ensemble model runs produced different results depicting distinct uncertainties with a range of expected discharge outputs and flood extents. Uncertainties and errors were quantified at every stage of the cascading model chain, and their effects on flood volumes and inundation extents were quantified. The uncertainties in rainfall input data and the variability of discharge and flood extent output were compared to similar outputs from the fixed or calibrated parameter model run scenarios (Yu et al., 2016).

### **5.5.1 Probability distribution fitting and rainfall ensemble generation**

Twelve probability distributions were examined to select the best fit distribution for the annual rainfall time series data in La Sierra catchment. The probability distributions identified in the catchment area were six Logistic, three Normal, two Beta, one Gen Pareto, and one Gumbel distribution. To identify the best-fit probability distribution with the highest number of stations selected, a relative scoring method was adopted based on the results of the three goodness-of-fit tests (Mamnoon and Rahman, 2019). The goodness-of-fit test results showed that the combined test scores for all three tests are presented in Table 5.17. The test results showed that the Logistic probability distributions had a score of 90 and were the best fit distribution. Overall, the Logistic distribution appeared in the top 5 ranked distributions of the highest number of stations, accounting for 58% of the selected stations. The Normal distribution was second, accounting for 25% of the stations, followed by Beta (17%), General extreme value (8%) and Gumbel (8%).

**Table 5:17. Combined scores in comparing and selecting Probability distributions.**

Distributions	Score
Logistic	90
Normal	60
Beta	60
Three-Parameter Weibull	36
Generalised Extreme Value	24
Generalised Pareto	15

The most selected probability distribution (considering ranks 1 to 3) is the Logistic distribution, in this case, it was selected for the uncertainty analysis. However, using a Box-Cox power transformation, non-symmetric distributions were transformed into normal distributions or closer to normal distributions. The fitted distribution was then used to estimate design rainfall ensembles using the Monte Carlo application (Lee et al., 2010). The application has a random number generator which was used to generate parameter values from selected probability distribution functions (Table 5:18) (Kalyanapu et al., 2012). The randomly generated input parameters were then used to generate randomly the required rainfall ensemble members.

**Table 5:18 Probability distribution parameter values for all rainfall gauging stations in La Sierra catchment.**

Gauging station	Distribution	Parameters		
		$\kappa$	$\sigma$	$\mu$
Dos Montes	Normal		8.716	2.5885
Playa de Rosario	Logistic		6.1324	3.3787
Tapijulapa	GenPareto	0.73598	1.178	-0.40989
Sitala	Logistic		7.8111	6.0714
Dos Patria	General extreme value	0.68378	2.1556	0.75606
Tequita	Logistic		10.01	7.1512
Las Mercedes	Logistic		6.0641	7.8058
La lagarte	Logistic		8.3752	6.703
Yayalon	Gumbel		10.787	-1.2313
Oxolotan	Logistic		8.8974	5.7021
Simojovel	Logistic		9.3684	5.3534
Teapa	Normal		9.3223	2.7439

### 5.5.2 Calibration results of the flood inundation model

The flood inundation model was subjected to a multi-response calibration process to find optimal Manning coefficient values for La Sierra River and its floodplain (Di Baldassarre et al., 2010). The results of each model run were compared to the satellite flood map, the Landsat 7 ETM+ images using the index of fit measure,  $F$  (Jung et al., 2014). Running the hydraulic model with different values of the Manning's  $n$  ranging from 0.025 to 0.05  $s/(m^{1/3})$  in the floodplains resulted in the of flood extent areas varying between 1.33 to 1.79  $km^2$  (Figure 5.19).

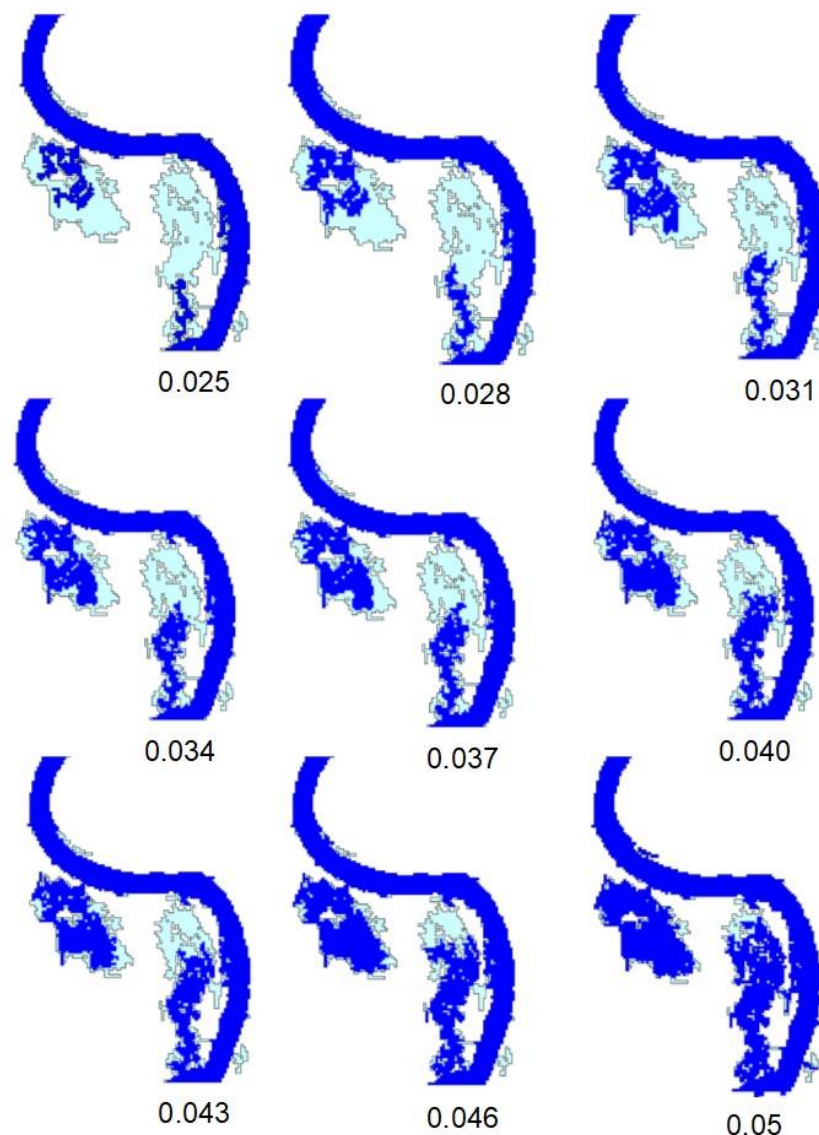


Figure 5.23 The outcome of the calibration process analysis showing changing flood extents to changes in Manning's coefficient values ( $s/[m^{1/3}]$ ).

The results showed that channel roughness values had an impact on flood extents. Any minor changes to roughness value resulted in visible changes in flood extent outlines. The flood extent sizes were sensitive to changes to the roughness values. Although variations in the flood area were not as high as those caused by variability in input data. Variations due to roughness were visible enough to alter the size of the flooded area. A global coefficient value of  $0.05 \text{ s/ (m}^{1/3}$  was selected after several calibration simulations, coming up with the best fit flood outlines and the value was selected to represent the Manning's coefficient of roughness parameter for the studied reach. The selected Manning coefficient value produced the best fit model ( $F= 0.92$ ) when the simulated flood extent outline was compared to the observed satellite flood extent outlines (Di Baldassarre et al., 2010). The results of the F-statistics show the spatial fitness between the simulated and observed flood outlines derived from Landsat images.

The results of the flood simulations from both scenarios showed that the Guayabal flood washlands south of the study region were the first to be flooded, while the rest of the channel barely experienced flooding. During floods, most of the floodwaters and river flow between the Guayabal and Laguna washlands in the La Sierra River were largely limited to the river channel within the floodwalls on either side of the river banks (Figure 5.24). However, the floodwaters were designed to spread in upstream and downstream areas where floodwaters are required to flood the Col. Los Pinos and La Laguna washlands in the north and the Los Guayabal washlands in the south.



**Figure 5.24: The simulated flood extents (light green) and satellite-derived contours (blue) underline the adequate capacity of the calibrated model to replicate the actual observed flooded area.**

### ***Scenario modelling***

Ensemble simulations were performed on the following two scenarios: (i) the calibrated model was run based on fixed calibrated parameter simulations to investigate the variability and propagation of rainfall uncertainties and (ii) the same model was run based on variable or individual parameter value changes this time to investigate variability and propagation of parameter uncertainties. The two scenarios enabled quantification and comparisons of the effect of parameter changes on uncertainties in discharge outputs and size of flood extent areas. All model runs were enforced with rainfall ensembles as input data and these generated discharge ensembles, each run based on parameter variations and fixed/calibrated parameters scenarios.

### 5.5.3 Rainfall uncertainty propagation results

The uncertainty in the rainfall input data was calculated using Probability Distributions Functions (Lee et al., 2010) and it was determined that rainfall in La Sierra catchment is mainly approximated by three distributions; the Logistic, Normal, Beta, Gen Pareto, and Gumbel probability distributions (Table 5.17). The Coefficient of Variation (CV), the mean and the Relative Interval Length (RIL) metrics (Li et al., 2011; Ahmadisharaf et al., 2018) were used for uncertainty quantification (Table 5.19).

The rainfall input ensembles were classified into probability intervals between 5% and 95% confidence limits (Loveridge et al., 2013; Ahmadisharaf et al., 2018). The limits were the upper and lower ends of the rainfall ensemble values where their actual values lie with some degree of certainty. The intervals were calculated based on the standard deviations of each ensemble member relative to the ensemble mean of all rainfall ensembles.

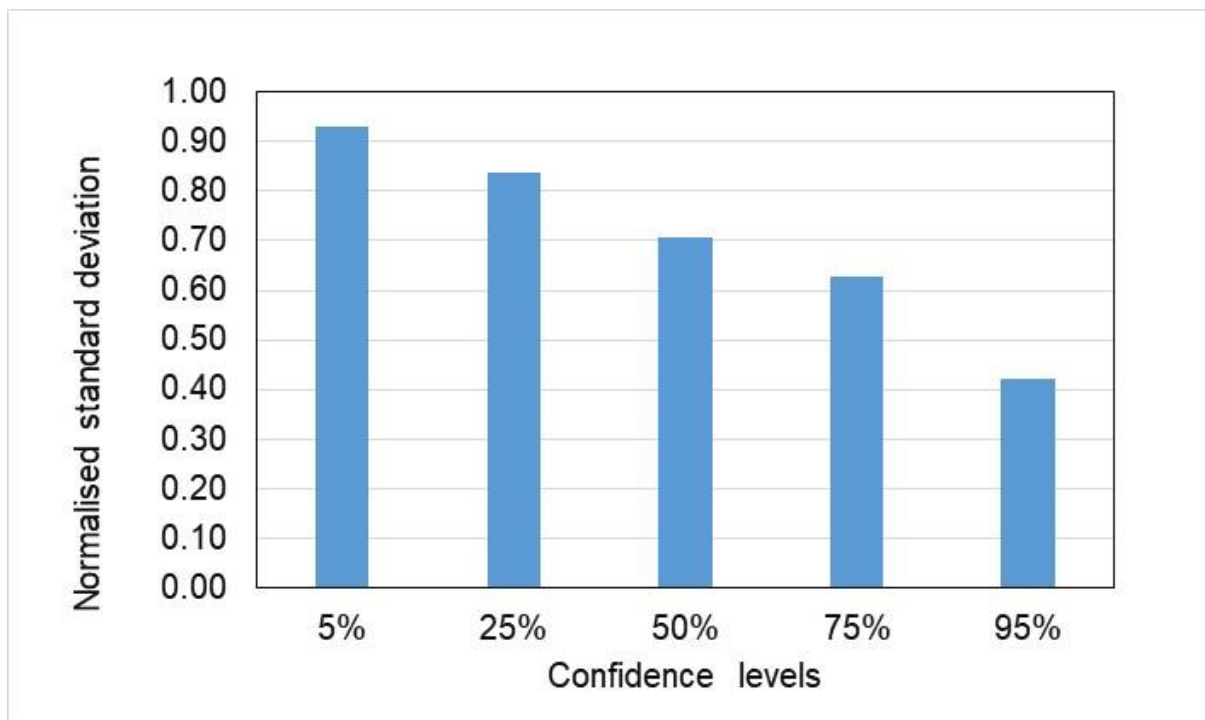
The confidence level is the probability that uncertainties lie within the interval. Confidence intervals were used in this study to describe ranges between 5% and 95% within which the average estimate occurs with defined probability. The 95% confidence interval has much frequent probability of 0.95 of containing the mean ensemble, while the 5% has a less frequent probability of 0.05 (Lane, 2015). Rainfall analysis results show that 92% of the ensemble members fell between 5% and 95% uncertainty limits, while the remaining 8% of members were approximately 10% of rainfall total.

**Table 5:19 Rainfall input, discharge, flood extent and uncertainty quantification and comparison based on the calibrated model scenario.**

Controlled run on fixed/ calibrated parameters		UNCERTAINTY						ERROR			
		95%		CV (%)		5%		RIL		RMSE (mm & m <sup>3</sup> /s)	
		mean	median	mean	median	mean	median	95%	mean	median	5%
Meteorological	Rainfall	63	70	71	73	0.49	0.47	0.45	0.46	0.46	0.46
Hydrological model	Discharge	45	45	45	49	0.45	0.46	29.9	57.8	54.31	92.1
Hydraulical model	Statistical level	Discharge (m <sup>3</sup> /s)		Flood extents (km <sup>2</sup> )		RIL					
	5%	744.75		7.02							
	95%	466		1.66		0.45					
	median	622		4.34							
	mean	624		4.36							

### **Variability in rainfall input**

The Coefficient of Variation (CV) for ensemble input rainfall calculated before rainfall was enforced in the hydrological model and was found to have 5% and 95% confidence limits of 73% and 6%, respectively, with a median variability of  $\pm 71\%$  and a mean of 70% (Table 5.19). Results show that rainfall amounts have higher variability in the 5% probability confidence level than in the upper (95%) level, particularly in less frequent probabilities for maximum rainfall totals. Rainfall uncertainties are higher in the low-frequency probabilities (5% confidence level) than in the high-frequency probability limits (95% confidence level) as shown in Figure 5.25.



**Figure 5.25: Rainfall amounts have high variability in the 5% probability confidence level than for the upper (up to 95%) confidence limit. (The 95% confidence interval has a 0.95 probability of containing the ensemble mean while 5% have a less frequent probability of 0.05).**

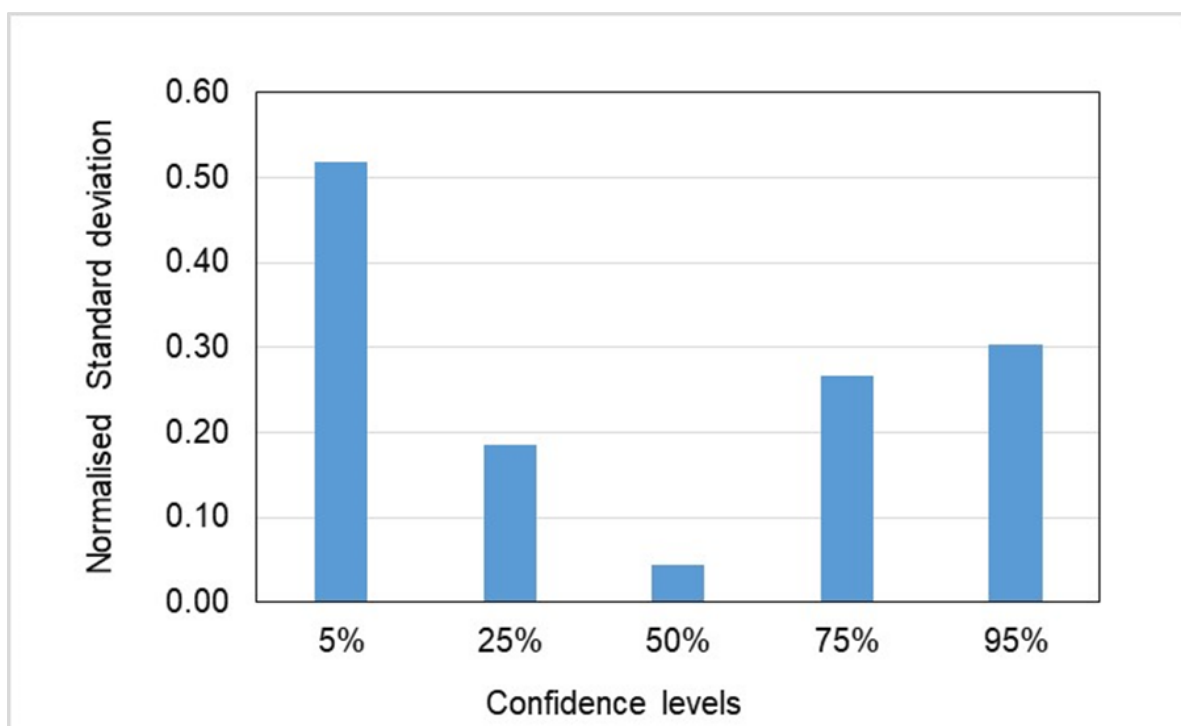
The Relative Interval Length (RIL) was used in combination with the Coefficient of Variation (CV) to measure the resolution of ensemble member predictive distributions (Ahmadisharaf et al., 2018) or the dispersion of the rainfall or discharge probability of distribution about the ensemble mean (Yu et al., 2016). A Relative Interval Length (RIL) with a median of  $\pm 47$  and a mean of 0.49 was found in the ensemble rainfall input data



between 5% and 95% confidence limits, showing a slightly lower average uncertainty. Here, the mean is the ensemble average estimated across all individual probable ensemble members. Each ensemble rainfall input data were compared to the observed rainfall and the Relative Mean Square Error (RMSE) metric was used to quantify errors. RMSE results show that the median error measured across all rainfall ensembles, between 5% and 95% was  $\pm 0.46$  mm, ranging from 0.45 to 0.48 mm, respectively (Table 5.19).

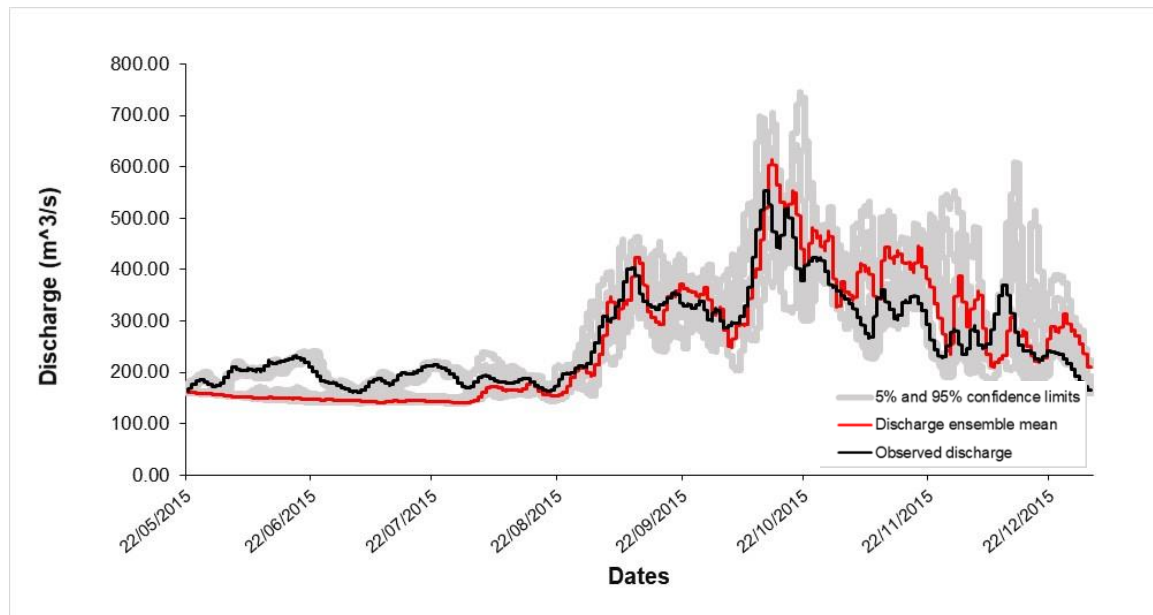
### ***Variability in discharge output***

Results from discharge ensemble simulations based on fixed/calibrated parameters scenario had a median variability of  $\pm 46\%$ , ranging from 44 to 49% and a mean of 45% in the 5% and 95% confidence limits (Table 5.19). The variability of discharge volumes was relatively evenly distributed with flood volume variability lower across all levels. However, there is higher variability in the lower probability limit than the upper limit and a much lower discharge uncertainty in the 50% confidence level (Figure 5.26).



**Figure 5.26: The variability of discharge output ensembles within 5% and 95% confidence limits based on fixed/calibrated parameter scenario.**

It is apparent from Figure 5.26 that there is less variability of ensembles in the more frequent probabilities (95% confidence limit) showing around 30% compared to the less frequent probability (5% confidence limit) around 51%. Therefore, the variability of discharge was found to be lower (RIL = 0.46) than the variability in rainfall input data that had a median Relative Interval Length (RIL) of 0.47 (Figure 5.27). The result shows a lower variability of discharge compared to rainfall input data when the model is run based on calibrated parameters.



**Figure 5.27: The 2015 hydrograph showing 5% and 95% confidence limits to represent uncertainty discharge due to fixed/calibrated parameter values. The red line represents the ensemble mean of discharge while the black is the observed discharge.**

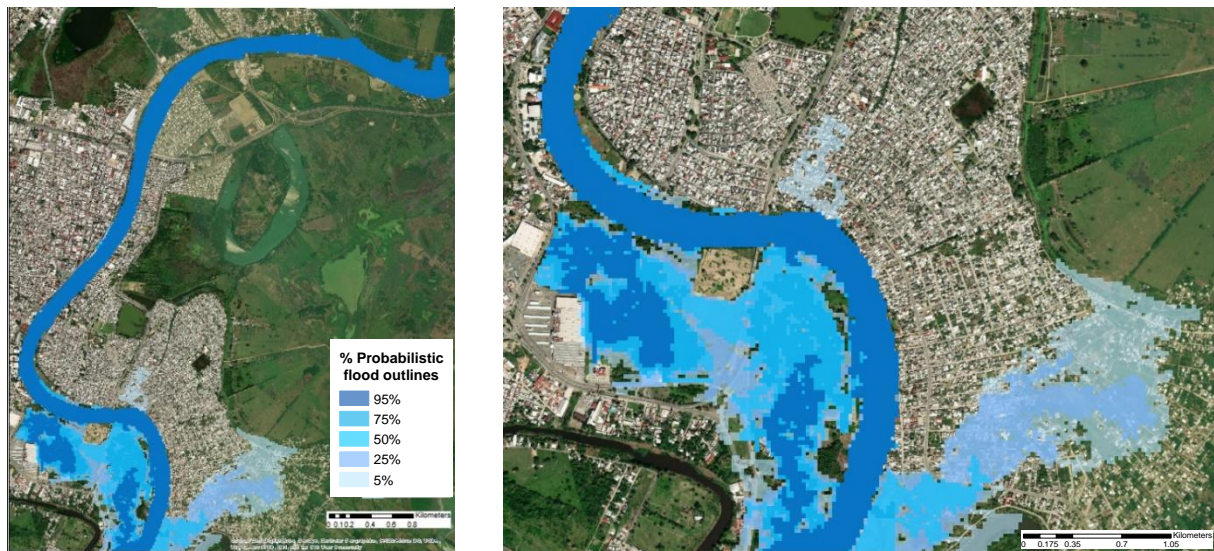
### ***Discharge error output***

Comparison results between observed and mean discharge ensemble based on fixed/calibrated parameter scenario show that discharge errors in the 5% and 95% confidence limits had a median variability of about  $\pm 56.68 \text{ m}^3/\text{s}$  ranging from  $31.9 \text{ m}^3/\text{s}$  to  $82.7 \text{ m}^3/\text{s}$  and a mean variability of  $56.64 \text{ m}^3/\text{s}$  (Table 5.19). As with rainfall errors, the variability of discharge errors was fairly evenly distributed across the 95% and 5% limits; discharge errors had a much wider band within a lower limit (confidence limit of 5%) compared to the upper limit (confidence limit of 95%). The variability of discharge error reduces with more frequent probabilities and increase with less probabilities.

### **Variability in flood inundation extents**

Results from simulated flood extents from the hydraulic model runs enforced with flow ensemble input data from hydrological model runs based on fixed/calibrated parameters had a 5% and 95% confidence limit of 7.02 and 1.66 km<sup>2</sup>, respectively (Table 5.19). Results show a median rainfall variability of  $\pm 71\%$  related to discharge variability of  $\pm 46\%$  and flood extents of  $\pm 4.34$  km<sup>2</sup>. In the 5% probability limit, there is 73% uncertainty in rainfall, which is related to 49% uncertainty in discharge and 7.02 km<sup>2</sup> uncertainty in the size of flood extent areas. While at the 95% probability limit, there was 63% uncertainty in rainfall was related to 45% uncertainty in discharge and 1.66 km<sup>2</sup> uncertainties in flood extents.

Therefore, under the calibrated parameter scenario, rainfall input data uncertainties between 63% and 73%, can cause discharge volumes to vary from 744.8 to 624 m<sup>3</sup>/s and the inundation extents from 1.66 to 7.02 m<sup>3</sup>/s (Figure 5.28). A Relative Interval Length (RIL) index was calculated to summarise the variability in the flood range and results showed RIL of 0.45, which is low compared to rainfall (RIL=49) input data, but equal to RIL=0.45 for discharge output data. The RIL measures the resolution of the ensemble distributions for each data for the cascading model (Li et al., 2011).



**Figure 5.28: Probabilistic flood inundation outlines in the 5% and 95% confidence limits based on ensemble model runs based on fixed/calibrated parameter values.**

## 5.5.4 Parameter uncertainty propagation results

This section presents results of parameter uncertainty propagation and interaction through a coupled hydrologic and hydraulic model based on variations of the following parameter values: Strickler overland flow resistance coefficient (K), Saturated water content ( $\theta_s$ ), AET/PET and the Saturated conductivity (KS). The parameter values were selected by defining parameter ranges for each land-use cover and soil types in the catchment area. The major land cover types considered in the uncertainty analysis were deciduous forest, grass, arable, and evergreen forests. Primary soil types considered included medium soil (18% clay, 35% and 15% sand or 18% clay and 15% sand or 65%), coarse soil (18% clay and 65%) and highly productive aquifer through cracks soils.

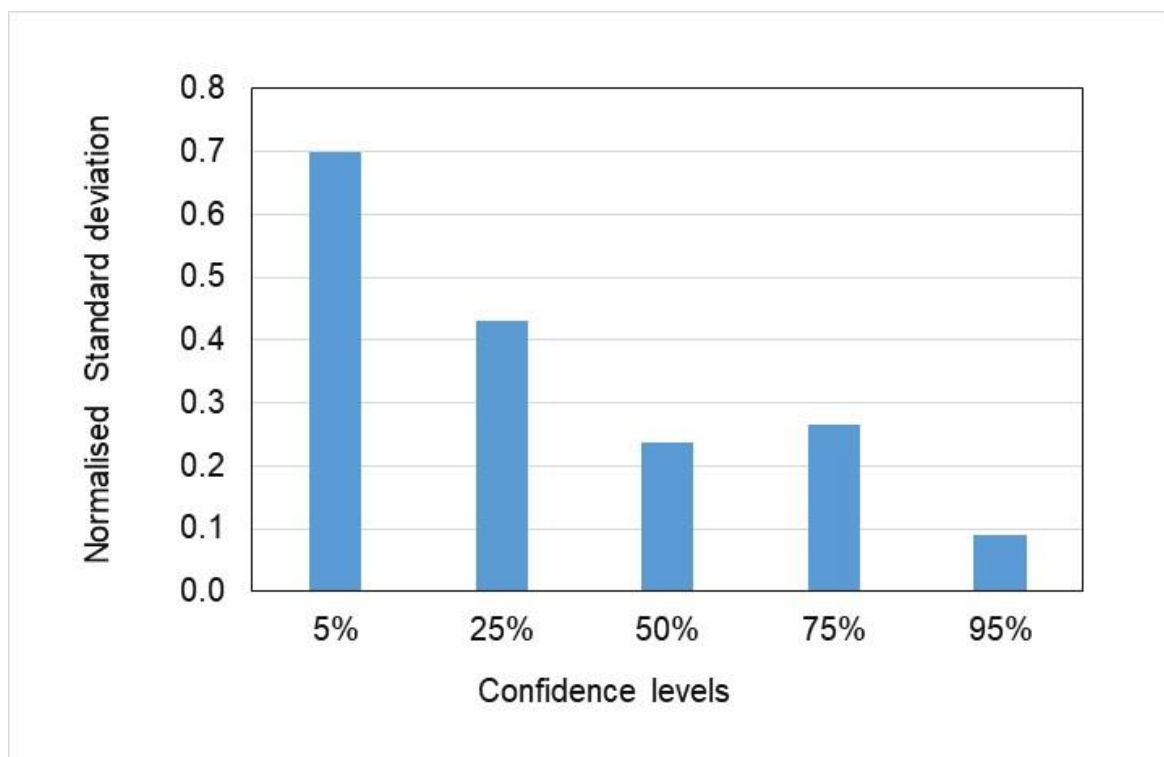
The propagation also includes results of quantified uncertainties at each stage of the coupled model chain, and related effects on the magnitude of peak discharge and flood extents (Rodríguez-Rincón et al., 2015; Yu et al., 2016; Tanaka et al., 2018). Like ensemble rainfall input, the discharge output ensemble results were classified into probability interval classes between 5% and 95% confidence limits. The ensemble discharge results generated were expressed using the statistical variance or measures of central tendency in terms of the mean, range, standard deviation and median. These measures were expressed related to each other as the Coefficient of Variation, the ratio of the mean to standard deviation. Propagated uncertainties were quantified and effects on discharge and flood extent areas were quantified using the Relative Interval Length (RIL). The errors between simulated and observed rainfall and discharge were calculated and quantified using the Root Mean Square Error (RMSE) metric.

### 5.5.5 Effects of high Strickler parameter values

Besides running the hydrological model enforced with rainfall ensembles based on fixed/calibrated parameter values, the model was also run based on randomly generated variations of the Strickler parameter values between high and low values for each vegetation category, but within permissible literature values. Each ensemble model scenario was run based on either high or low Strickler parameter values and each discharge output was compared to equivalent model runs outputs based on the calibrated parameters.

### ***Discharge uncertainty***

Results from the probabilistic model runs based on high Strickler parameter values generated variations of discharge in the 5% and 95% confidence limits with a median variability by  $\pm 45\%$ , ranging from 38 to 46% and a mean of 44% (Table 5.20). The uncertainty in the discharge output was relatively evenly distributed around the mean with considerable variations in more frequent probabilities than in the less frequent probabilities (Figure 5.29). The Coefficient of Variability (CV) of discharge output for the 5% confidence limit decreased by 6%, and for the 95% confidence limit, decreased by 16%, which is a decrease of discharge uncertainty in both less and frequent probabilities. A low Relative Interval Length (RIL) value (0.18) with a mean of 0.36 shows fewer uncertainties across the 5% and 95% confidence interval. The result shows a decrease of 61% of variability when compared to discharge output from the ensemble model runs based on calibrated parameter scenario (Table 5.20).



**Figure 5.29:** The uncertainty in the discharge output was relatively evenly distributed around the mean with large bands in less frequent probabilities than in the more frequent probabilities.

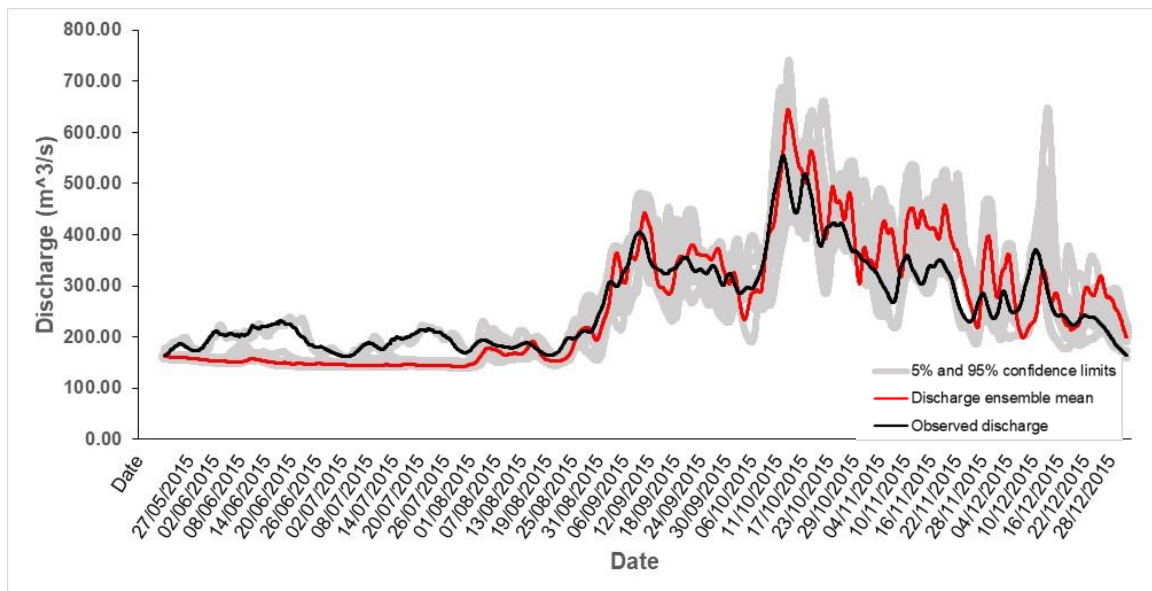
**Table 5:20 Rainfall, discharge and flood extents uncertainty and error quantification and comparison values.**

High Strickler parameter		UNCERTAINTY				ERROR					
		95%	mean	CV (%) median	5%	mean	RIL median	RMSE (mm & m <sup>3</sup> /s)			
Meteorological	Rainfall	63	70	71	73	0.49	0.47	0.45	0.46	0.46	0.47
Hydrological model	Discharge	38	44	45	46	0.36	0.18	34.7	53	50.47	54.3
	% change	-16%	-2%	0%	-6%	-20%	-61%	16%	-8%	-7%	-41%
Hydraulical model	Statistical level	Discharge (m <sup>3</sup> /s)		Flood extents (km <sup>2</sup> )		RIL					
	5%	741.34	-0.46%	2.82	-60%						
	95%	589.38	26.48%	1.91	15%	0.89					
	media	632	1.61%	1.99	-54%						
	mean	175.9	-71.81%	1.23	-72%						

### ***Discharge error variations***

The simulated discharge based on high Strickler parameter values were each compared to the observed discharge to estimate discharge error across ensembles within 5% and 95% confidence limits. The comparison results show that discharge errors across the confidence interval have an estimated median variability of  $\pm 50.47$  m<sup>3</sup>/s ranging from 34.71 to 54.29 m<sup>3</sup>/s and a mean of 52.95 m<sup>3</sup>/s (Table 5.20).

The variability of discharge errors was relatively evenly distributed around the mean, with the variability of discharge volumes having large bands in the lower probability limits than upper frequent probabilities in Figure 5.30. The model runs based on high Strickler values caused the variability of discharge errors to increase by 16% for high frequent probabilities and decrease by 41% for less frequent probabilities compared to the calibrated model run scenario.



**Figure 5.30: The 2015 hydrograph showing 5% and 95% confidence limits in grey to represent uncertainty in discharge due to high Strickler parameter values for each vegetation type. The red line represents the discharge ensemble mean while the black is the observed discharge.**

### ***Flood inundation extents***

The results of the variability of flood extent areas due to high Strickler parameter values had a 5% and 95% confidence limits of 2.82 km<sup>2</sup> and 1.91 km<sup>2</sup>, respectively. The median variability was found at around  $\pm 1.99$  km<sup>2</sup>, (a decrease of 54%) and a mean of 1.23 km<sup>2</sup> (another decrease of 72%) if compared to the calibrated parameter scenario. Results show that a relationship exists between discharge volumes and flood extents due to high Strickler parameter value. In the 95% probability limit, rainfall uncertainty of 63% is related to uncertainty in the discharge of 38% and flood extents reaching 1.91 km<sup>2</sup>. While in the 5% confidence limit, rainfall variability of 73% is linked to discharge with a variability of 46% and flood extent reaching 2.82 km<sup>2</sup>. Results show the highest Relative Interval Length (RIL=0.89) compared to rainfall input and discharge output data.

### ***Effect on flood extents***

The uncertainty analysis results show that high Strickler parameter values resulted in a general decrease in flood extent areas for less frequent probabilities (5%) and increases in more frequent probabilities (95%). Flood extent areas decreased for less frequent probabilities while they increase in more frequent probabilities. Compared to calibrated parameter model runs scenario high Strickler parameter values had an effect of

decreasing uncertainty in flood volumes by 22% for less frequent probabilities and a decrease in the size of flood inundation areas by 60%. In real terms, it was a decrease from 7.02 to 2.82 km<sup>2</sup> of flooded areas.

However, results show that in more frequent probabilities, an increase in uncertainty of 38% of flood volumes resulted in a 15 % increase in inundated areas from 1.66 km<sup>2</sup> to 1.91 km<sup>2</sup>. Also, the parameter variations resulted in a decrease of 60% in flood extent areas in the 5% confidence limit and an increase of 15% in the 95% confidence limit. Thus, under this scenario, the size of the areas flooded varied from 1.91 to 2.82 km<sup>2</sup> (Figure 5.31).



**Figure 5.31: Probabilistic flood inundation extent outlines in the 5% and 95% confidence limits based on ensemble model runs based on high Strickler parameter value scenario.**

### ***The overall effect of the high Strickler parameter values***

Uncertainty analysis results showed that in the 5% and 95% confidence limits, the median variability of rainfall was around  $\pm 71\%$ , ranging from 63 to 73%. This uncertainty was propagated through the coupled models (with a hydrological model run based on the highest Strickler parameter value) generated discharge output with an uncertainty of  $\pm 45\%$ , ranging from 38% to 46% with RMSE of  $\pm 50.47$  m<sup>3</sup>/s ranging from 35 to 54 m<sup>3</sup>/s. The propagated uncertainties resulted in some shifts in peak discharge of around  $\pm 632$  m<sup>3</sup>/s ranging from 589 to 741 m<sup>3</sup>/s and flood extents varying around  $\pm 2.0$  km<sup>2</sup> ranging from 1.9 to 2.8 km<sup>2</sup> (Figure 5.31).



## 5.5.6 Effect of low Strickler parameter values

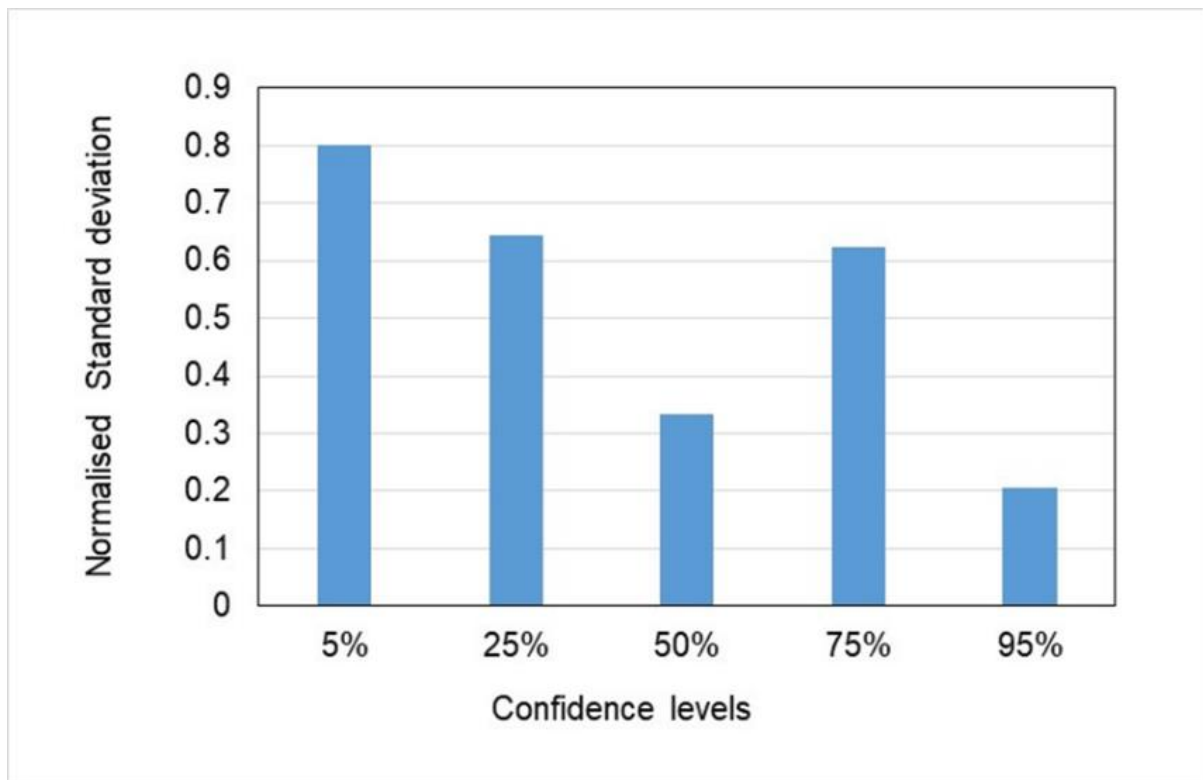
### *Variability in discharge output*

The results of simulated discharge output volumes from the probabilistic model run based on low Strickler parameter value scenario had a median variability in the 5% and 95% confidence limit of  $\pm 43\%$ , ranging from 37 to 45% and a mean of 44% (Table 5.21). Compared to discharge volumes generated from model runs based on fixed/calibrated parameter scenario, the variability for both 5% confidence limit and the 95% confidence limit decreased by 8% and 18% respectively.

**Table 5:21 Rainfall input, discharge output and flood inundation extents, uncertainty and error quantification and comparison between calibrated and low Strickler parameter value values for each vegetation type.**

Low Strickler parameter		UNCERTAINTY				ERROR					
		CV (%)				RIL		RMSE (mm & m <sup>3</sup> /s)			
		95%	mean	median	5%	mean	median	95%	mean	median	5%
Meteorological	Rainfall depth	63	70	71	73	0.49	0.47	0.45	0.46	0.46	0.47
Hydrological model	Discharge	37	44	43	45	0.38	0.21	31.9	53.5	53.63	94.3
	% change	-18%	-2%	-4%	-8%	-16%	-54%	7%	-7%	-1%	2%
Hydraulical model		Statistical level		Discharge (m <sup>3</sup> /s)		Flood extents (km <sup>2</sup> )		RIL			
		5%	651.70	-12%	2.12	-70%					
		95%	493.24	6%	1.35	-18%	0.27				
		media	589.74	-5%	1.65	-62%					
	mean	591.03	-5%	1.72	-60%						

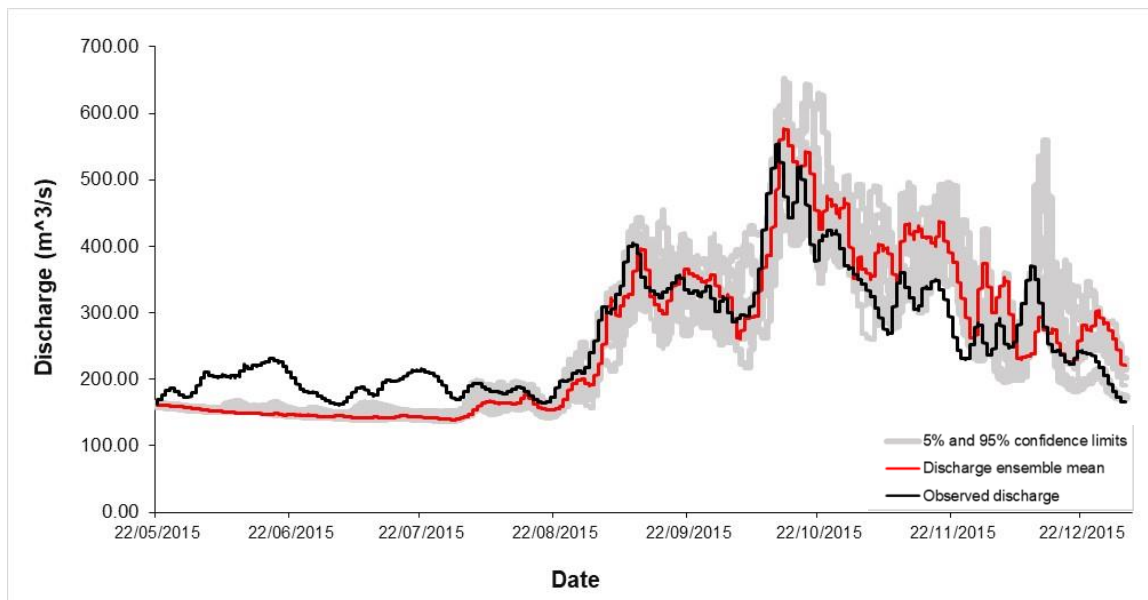
Analysis results showed that running the SHETRAN hydrological model based on low Strickler parameter value had an effect of decreasing discharge volumes across the 5% and 95% confidence limits. The uncertainty was summarised using the Relative Interval Length (RIL) and found to be 0.21, a mean of 0.38 showing that low Strickler parameter-based model run results in low uncertainties, which is a decrease of 54% when compared to the calibrated parameter scenario. The variability of discharge output was relatively evenly distributed across all confidence levels with the highest (80%) variation in less frequent probabilities than in the more frequent probabilities (Figure 5.32). However, with an increased spread in the 75% confidence level and the lowest in the 50% level.



**Figure 5.32:** The uncertainty in the discharge output was relatively evenly distributed around the mean and it has large bands in more frequent probabilities than in the less frequent probabilities.

### ***Discharge error variations***

To estimate the Relative Mean Square Error (RMSE) across the 5% and 95% confidence interval, each ensemble discharge output based on a low Strickler parameter value was compared to the observed discharge (Figure 5.33). Results showed that discharge errors across confidence intervals (5% to 95%) had a median error of  $\pm 53.63$   $\text{m}^3/\text{s}$  ranging from 31.89 to 94.33  $\text{m}^3/\text{s}$  and a mean of 53.5  $\text{m}^3/\text{s}$ . Compared to probability discharge errors from the calibrated parameter scenario, the result showed a 7% increase in errors in the high-frequency probabilities and an increase of 2% in low frequent probabilities.



**Figure 5.33: The 2015 hydrograph showing 5% and 95% confidence to represent uncertainty in discharge due to low Strickler parameter values for each vegetation category. The red line represents the ensemble mean of discharge while the black is the observed discharge.**

### ***Variability in flood inundation extents***

Results from simulations of flood extent areas due to low Strickler parameter values show a 5% and 95% confidence limits of 2.12 and 1.35 km<sup>2</sup>, respectively. Compared to the calibrated parameter scenario, the flood extent areas decreased by 70% in the 5% confidence limit and 18% in the 95% confidence limit. The variability in the size of flooded areas due to low Strickler parameter values showed a median variability of  $\pm 1.65 \text{ km}^2$ , which is a 62% decrease and a mean of 1.72 km<sup>2</sup>, a 60% decrease compared to calibrate parameter scenario run. Results show a link between discharge and the size of flooded areas, where there is a decrease of 8% uncertainty in maximum discharge, there is an equivalent decrease of 70% in the uncertainty in the size of flooded areas. Also, a decrease of 18% in minimum discharge uncertainty is linked with a related decrease of 18% in flood area variability.

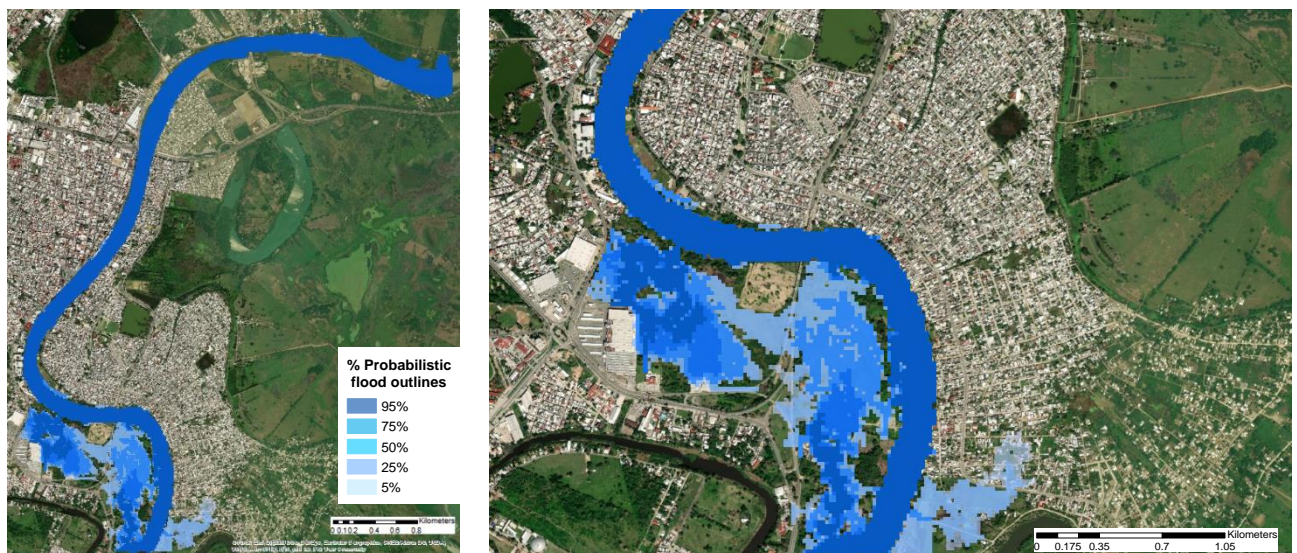
### ***Effect on flood extents***

Results show that low Strickler parameter values had an effect of consistently decreasing the size of flood extent areas in both less frequent (5%) and more frequent probabilities (95%). If compared to the calibrated parameter scenario, the low roughness values were linked with a decrease of 18% uncertainty in discharge in more frequent probabilities (95% confidence limit) and this was related to a decrease of 70% in the

variability in the size of flooded areas. For example, in the 95% confidence limit, flood extents decreased from 1.66 km<sup>2</sup> to 1.35 km<sup>2</sup> and for the 5% confidence limit showed a decrease from 7.02 km<sup>2</sup> to 2.12 km<sup>2</sup> in flood extents. The low parameter values resulted in a 70% decrease in flood extent areas in the 5% confidence limit and an 18% decrease in the 95% confidence limit, ranging from 1.35 to 2.12 km<sup>2</sup>. The uncertainty in the flood extents was summarised using the Relative Interval Length (RIL) index and results show a RIL of 0.27, showing slightly increased uncertainties compared to discharge output (RIL=0.21).

### ***The overall effect of low Strickler parameter***

The uncertainty analysis results showed that in the 5% and 95% confidence limits, the variability of rainfall had a median variability of  $\pm 71\%$ , ranging from 63 to 73% and if the uncertainty is propagated through the hydrological model based on lowest Strickler parameter value generates discharge output with a median variability of  $\pm 43\%$  ranging from 37% to 45% with median variability errors of  $\pm 53.63$  m<sup>3</sup>/s ranging from 31.89 to 94.33 m<sup>3</sup>/s. The peak discharge produced had a median variability of  $\pm 590$  m<sup>3</sup>/s ranging from 493 to 652 m<sup>3</sup>/s resulting in flood extents of  $\pm 1.6$  km<sup>2</sup> ranging from 1.4 to 2.1 km<sup>2</sup> (Figure 5.34).

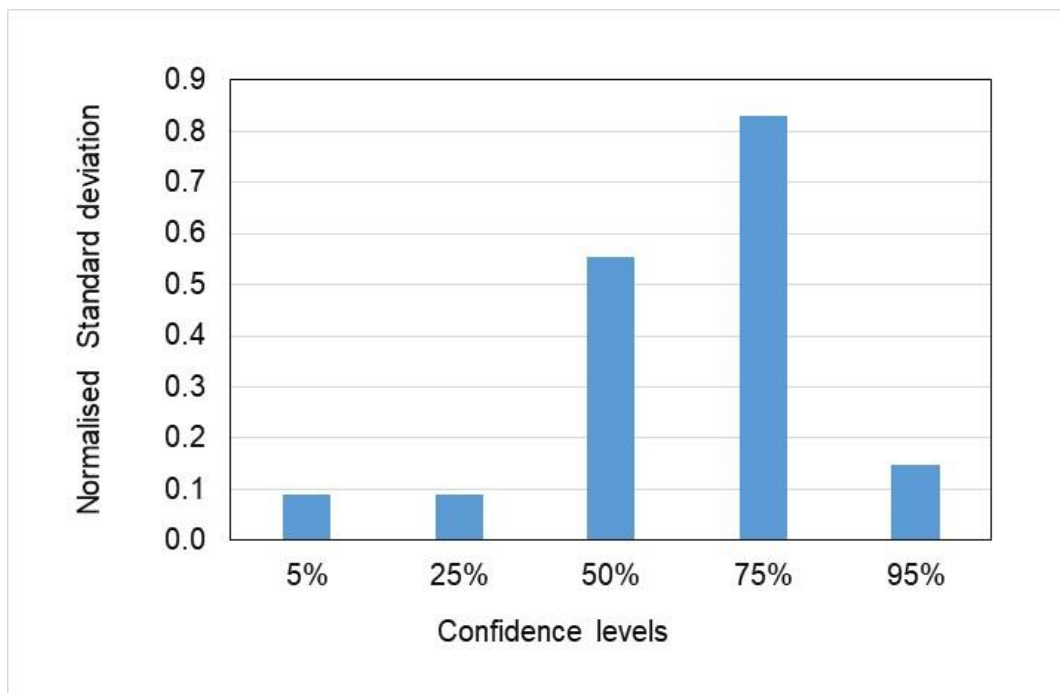


**Figure 5.34: Probabilistic flood inundation extent outlines in the 5% and 95% confidence limits based on ensemble model runs based on low Strickler parameter values for each vegetation type.**

### 5.5.7 Effect of high AET/PET parameter values

#### *Variability in discharge output volumes*

The results from the ensemble model runs based on highest AET/PET parameter values for each vegetation type influenced the variability of discharge output volumes in the 5% and 95% confidence limits. Results show a median variability of  $\pm 42\%$ , ranging from 35 to 44% and a mean of 43% (Table 5:22) The variability of discharge was relatively evenly distributed mainly for 75% confidence level, having highly uncertain discharge volumes with a wider band than the lower and upper levels (Figure 5.35).



**Figure 5.35: The uncertainty in the discharge output was not evenly distributed around the mean and it has large bands in more frequent probabilities than in the less frequent probabilities.**

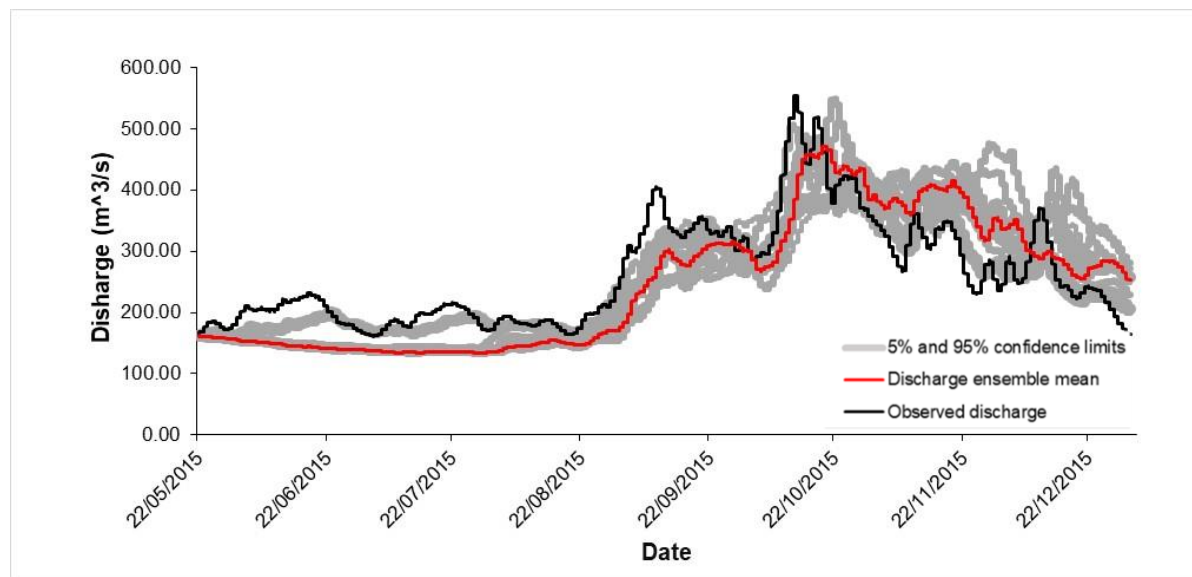
Compared to discharge volumes generated from the ensemble model runs based on calibrated parameters, the variability for discharge showed that discharge volumes were consistently decreased in all probabilities by 10% in the 5% confidence limit and 22% in the 95% confidence limit when compared to discharge output for the calibrated parameter scenario. However, the median Relative Interval Length (RIL) of 0.48 and a mean of 0.49 was found in the high AET/PET scenario, showing an increase of 9% and 4% respectively when compared to discharge output from the calibrated parameter model run scenario (Table 5.22).

**Table 5:22 Rainfall, discharge and flood inundation extents error and uncertainty quantification and comparisons due to high AET/PET parameter value values for each vegetation type.**

High AE/PE		UNCERTAINTY						ERROR				
		95%	mean	CV (%)		5%	mean	RIL	RMSE (mm & m <sup>3</sup> /s)			
				median			median	95%	mean	median	5%	
Meteorological	Rainfall depth	63	70	71	73		0.49	0.47	0.45	0.46	0.46	0.47
Hydrological model	Discharge	35	43	42	44		0.49	0.48	39.81	63.39	64.34	75.16
	% change	-22%	-4%	-7%	-10%		9%	4%	33%	10%	18%	-18%
		Statistical level		Discharge (m <sup>3</sup> /s)		Flood extents (km <sup>2</sup> )		RIL				
		5%		554.36		-26%		1.49		-79%		
		95%		411.20		-12%		1.33		-20%		0.30
Hydraulical model		median		483.59		-22%		1.35		-69%		
		mean		483.94		-22%		1.37		-69%		

### Discharge error variations

The simulated discharge outputs based on high AET/PET parameter values were compared to the 2015 observed discharge (Figure 5.36) to estimate the Relative Mean Square Error (RMSE) within the 5% and 95% confidence limits. The discharge error across the confidence intervals had a median variability of  $\pm 64.34$  m<sup>3</sup>/s ranging from 39.81 to 75.16 m<sup>3</sup>/s and a mean of 63.39 m<sup>3</sup>/s. Compared to the calibrated parameter discharge output, the analysis results showed an error increase of 33% for frequent high probabilities and a decrease of 18% in low frequent probabilities.



**Figure 5.36: The 2015 hydrograph showing the 5% and 95% confidence limits representing uncertainty in discharge due to high AET/PET parameter values for each vegetation type. The red line represents the mean ensemble of discharge while the black is the observed discharge.**

### ***Variability in flood inundation extents***

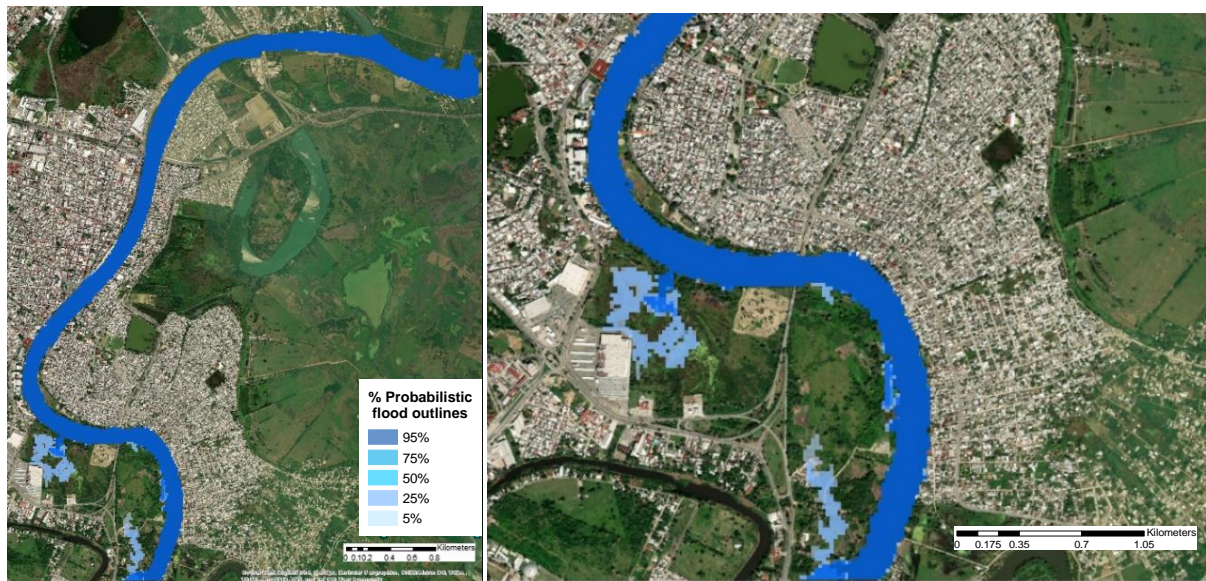
The results of the variability of flood inundation extents due to high AET/PET parameter values had a 5% and 95% confidence limits of 1.49 and 1.33 km<sup>2</sup>, respectively. Unlike flood inundation extents based on calibrated parameter scenario, high AET/PET parameter values resulted in decreases in the size of flooded areas by 79% in the 5% confidence limit however with an increase of 20% in the 95% confidence limit. The variability of the sizes of flooded areas from the ensemble model runs based on high AET/PET parameters had a median of  $\pm 1.35$  km<sup>2</sup> a decrease of 69% and a mean of 1.37 km<sup>2</sup> another decrease of 69% when compared to the calibrated parameter scenario (Table 5.22). Besides, the result shows a decrease of 10% in uncertainties in maximum flows is related to a decrease of 79% in uncertainties in inundation areas when the model is run based on high AET/PET parameter values. Also, a decrease of 22% in more frequent probabilities (5% confidence limit) was related to a decrease of 20% in the aerial extent of inundation areas.

### ***Effects on flood extents***

Results show that the variability of flood inundation areas due to high AET/PET parameter values resulted in a significant decrease in the sizes of flooded areas for both less frequent probabilities (5%) and more frequent probabilities (95%). The high AET/PET parameter values resulted in a 10% decrease of uncertainty in the maximum discharge of less (5%) frequent probabilities with a consequent 79% decrease in inundation areas from 7.02 to 1.49 km<sup>2</sup> (Figure 5.37).

However, for more frequent probabilities, a 22% decrease in discharge is linked to a 20% decrease in the size of flooded areas. In real terms, it is a decrease from 1.66 km<sup>2</sup> to 1.33 km<sup>2</sup> in flood extents. Unlike the calibrated parameter scenario, the discharge output based on high AET/PET parameter values was consistently reduced. The high AET/PET parameter values caused a 79% decrease in flood extents in the 5% confidence limit and a 20% decrease in the 95% confidence limit and the extents varying from 1.33 km<sup>2</sup> to 1.49 km<sup>2</sup> showing high AET/PET parameter values for each vegetation category in the study area have a significant effect on flood extents in the La Sierra catchment.

To summarise the relative distribution of uncertainty reported by the cascading model at each stage, the Relative Interval Length of 0.30 was found, which is the lowest RIL compared to rainfall input (RIL = 47) and discharge data (RIL = 48) in this scenario.



**Figure 5.37: Probabilistic flood inundation outlines in the 5% and 95% confidence limits based on ensemble model runs on high AET/PET parameter values for each vegetation type.**

### ***The overall effect of high AET/PET parameter***

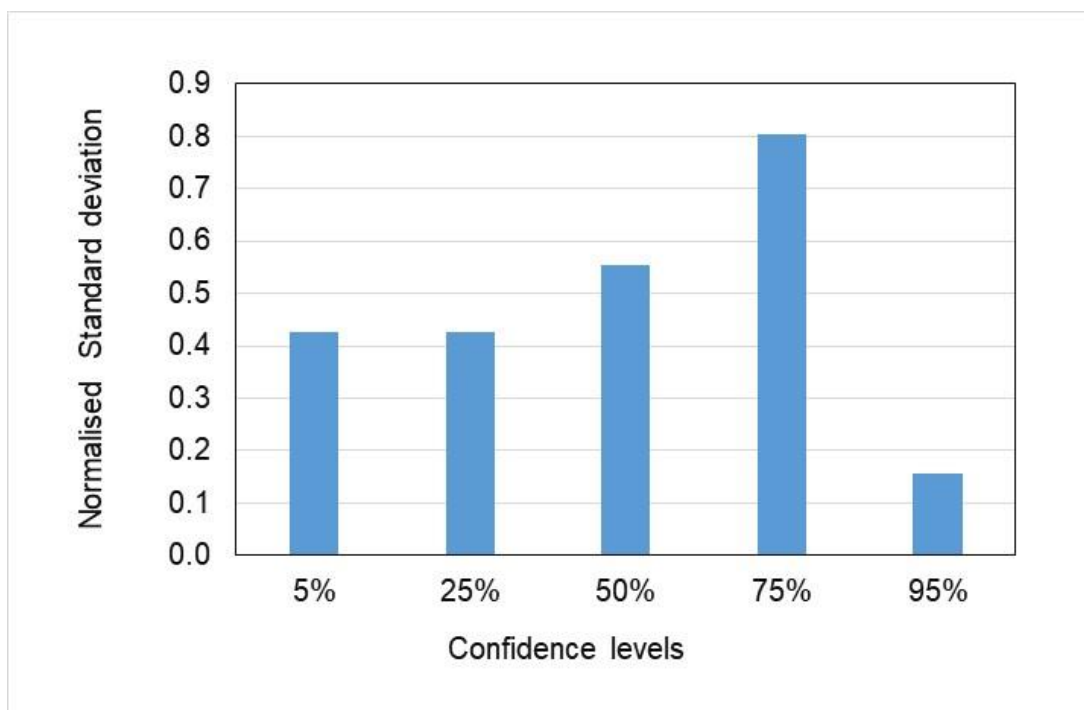
Overall, uncertainty analysis results have shown that in the 5% and 95% confidence limits, the ensemble input rainfall with a median variability of  $\pm 71\%$ , ranging from 63 to 73% when enforced and propagated through a hydrological model run on high AET/PET parameter values it generates a discharge output with a median variability of  $\pm 42\%$  ranging from 35% to 44% and errors with a median variability of  $\pm 64\%$  ranging from 40% to 75%. The propagated uncertainty resulted in shifts in peak discharge with a median variability of  $\pm 484 \text{ m}^3/\text{s}$  ranging from 411 to 554  $\text{m}^3/\text{s}$  and flood extents of  $\pm 1.3 \text{ km}^2$  varying from 1.3 to 1.5  $\text{km}^2$ .



## 5.5.8 The effect of low AET/PET parameter values

### *Variability in discharge output volumes*

Analysis results showed that the median variability of discharge output ensembles due to model runs based on low AET/PET parameter values for 5% and 95% confidence thresholds was  $\pm 44\%$ , ranging from 43 to 45% and a mean of 45% (Table 5.23). Compared to discharge volumes generated from model runs based on the calibrated parameter model run scenario, the variability for the 5% confidence limit decreased by 8% while the 95% limit decreased by 4%. The results show a consistent decrease of uncertainties in discharge across all confidence interval limits (Figure 5.38). A Relative Interval Length (RIL) of 0.47 was found showing an increase of 2% and an average increase of 0.49 and an increase of 9% compared to the calibrated parameter model run scenario.



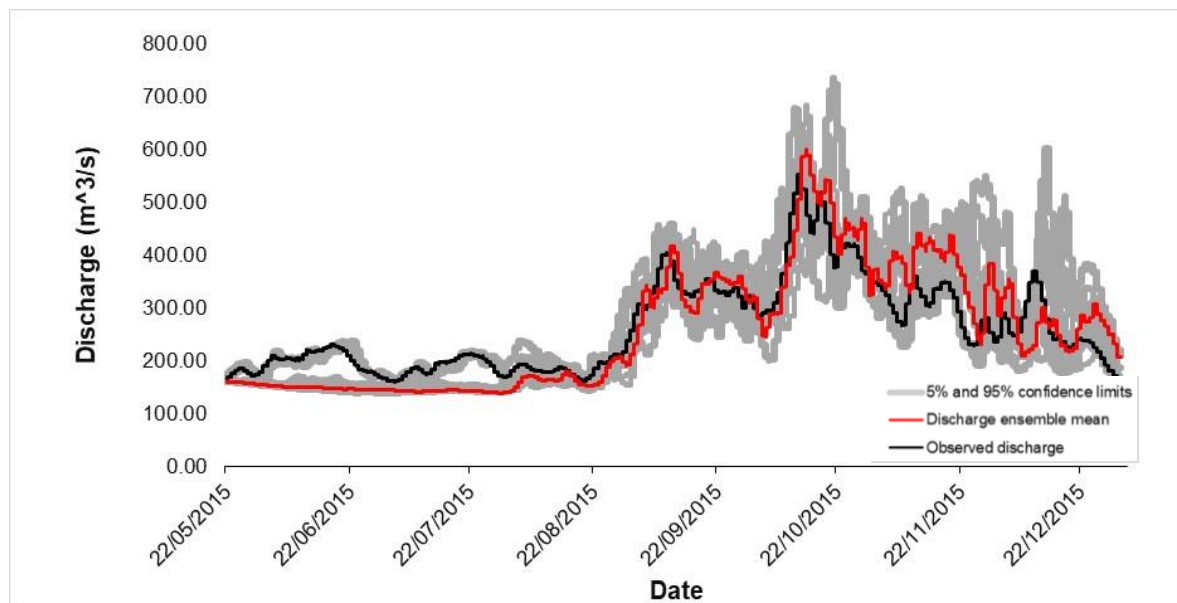
**Figure 5.38:** The variability of discharge output due to low AET/PET parameter values was relatively evenly distributed, more frequent probabilities than the less frequent probabilities.

**Table 5:23 Rainfall, discharge, flood extents error and their uncertainty quantification including comparisons between calibrated parameter and low AET/PET parameter values for each vegetation type.**

Low AE/PE		UNCERTAINTY				ERROR						
		CV (%)				mean	RIL		RMSE (mm & m <sup>3</sup> /s)			
		95%	mean	median	5%		median	95%	mean	median	5%	
Meteorological	Rainfall depth	63	70	71	73	0.49	0.47	0.45	0.46	0.46	0.47	
Hydrological model	Discharge	43	45	44	45	0.49	0.47	48.4	60.8	56.96	86.1	
	% change	-4%	0%	-2%	-8%	9%	2%	62%	5%	5%	-7%	
Hydraulical model	Statistical level		Discharge (m <sup>3</sup> /s)		Flood extents (km <sup>2</sup> )		RIL					
	5%	734.82	-1%	2.76	-61%	0.41						
	95%	486.54	4%	1.35	-19%							
	media	612.69	-1%	1.82	-58%							
	mean	615.98	-1%	1.90	-56%							

### Discharge error

Discharge outputs from model runs based on low AET/PET parameter values were compared to the observed discharge to estimate relative errors in discharge ensembles across 5% and 95% confidence limits (Figure 5.39).



**Figure 5.39: The 2015 hydrograph showing 5% and 95% confidence limits (in grey) representing uncertainty in discharge due to low AET/PET parameters values for each vegetation type. The red line represents the ensemble mean, and the black is the observed discharge.**

The low AET/PET parameter values generated discharge errors that varied across the 5% and 95% confidence limits with a median variability of around  $\pm 56.96 \text{ m}^3/\text{s}$  ranging from 48.38 to 86.07  $\text{m}^3/\text{s}$  and a mean of 60.81  $\text{m}^3/\text{s}$ . Compared to calibrated parameter discharge, this result showed a 62% increase in discharge errors for the low-frequency probabilities and a decrease of 7% for low frequent probabilities.

### ***Variability in flood inundation extents***

The results of the variability of flood extent areas due to low AET/PET parameter values show a 5% and 95% confidence limits with 2.76 and 1.35  $\text{km}^2$  respectively (Table 5.23). Compared to flood extents generated from the calibrated parameter model runs scenarios the low AET/PET parameter value resulted in decreases of 61% in flood extents for the 5% confidence limit and another reduction of 19% for the 95% confidence limit. The median variability of flood extents due to low AET/PET parameter was  $\pm 1.82 \text{ km}^2$ , a decrease of 58% and a mean of 1.90  $\text{km}^2$ , which is a decrease of 56% when compared to the calibrated parameter scenario. The result reveals some links between the variability of discharge and the size of areas flooded, where a decrease of 8% uncertainty is related to a decrease of 61% in inundation areas. Also, a decrease of 4% uncertainty in more frequent flows is related to decreases of 19% uncertainty in flood extents.

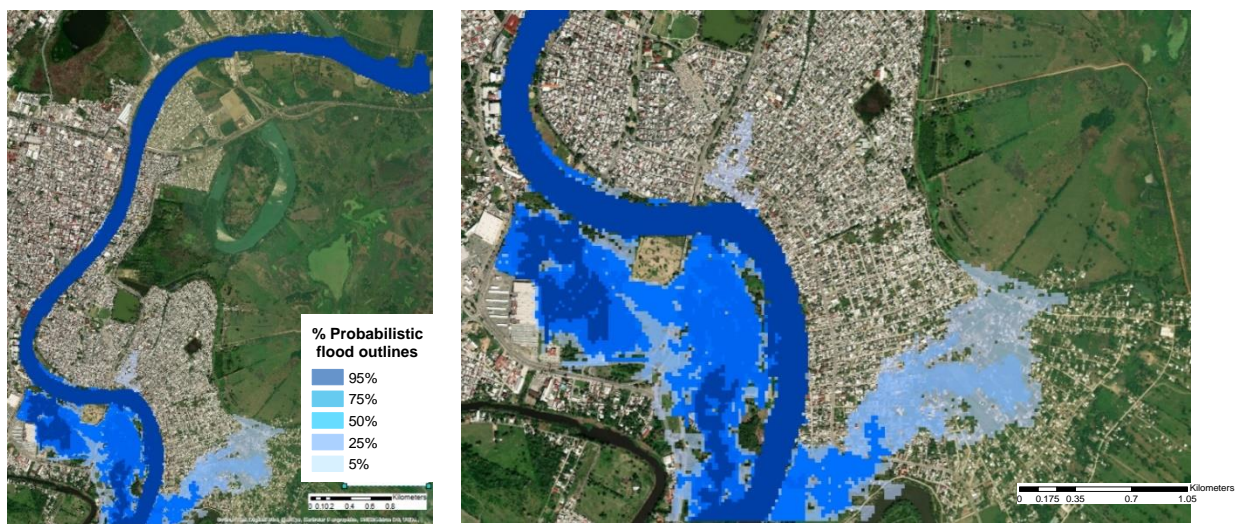
### ***Effect on flood extents***

Uncertainty analysis results show that low AET/PET parameter values for each vegetation type in the study area had an effect of decreasing the variability of flood extent areas in both less frequent probabilities (5% confidence limit) and more frequent probabilities (95% confidence limit). Low parameter values resulted in a decrease of 8% in discharge uncertainty, corresponding to a decrease of 61% in the size of flooded areas, from 7.02  $\text{km}^2$  to 2.76  $\text{km}^2$  of flood extent areas. However, for more frequent probabilities (95% confidence limit) comprising minimum flows, there was a decrease of 4%, which is related to a decrease in flooded areas by 19%. In real terms, it is a reduction from 1.66 to 1.35  $\text{km}^2$  in flooded areas. Also, results show that a decrease of 8% uncertainty of maximum flows for more frequent probabilities was related to a decrease of 61% uncertainty of inundation areas.

To summarise the relative distribution of uncertainties reported by the cascading model using the Relative Interval Length, it was found to be  $RIL=0.41$ . The RIL obtained is the lowest reported by the model chain, compared to rainfall input data ( $RIL=47$ ) and discharge data out ( $RIL=47$ ). The RIL measures the resolution of uncertainties' in the ensemble distributions by normalising the uncertainties in each confidence interval (Jin et al., 2010; Li et al., 2011).

### ***The overall effect of low AET/PET parameter values***

Analysis results of the effect of low AET / PET parameter showed that within the 5% and 95% confidence limits, rainfall input data had a median variability of  $\pm 71\%$ , ranging from 63 to 73% when enforced and propagated through a hydrological model generates ensemble discharge with a median variability of  $\pm 44\%$  ranging from 43% to 45% and relative errors of  $\pm 57\%$  ranging from 48% to 86%. The propagated uncertainties cause some shifts in discharge ensembles with a median variability of  $\pm 613\text{m}^3/\text{s}$  ranging from 487 to 735  $\text{m}^3/\text{s}$  resulting in flood extents of  $\pm 1.8 \text{ km}^2$  ranging from 1.3 to 2.8  $\text{km}^2$  (Figure 5.40).



**Figure 5.40: Probabilistic flood inundation extents within the 5% and 95% confidence limits from ensemble model runs based on low AET/PET parameter values for each vegetation type.**

## 5.5.9 Effect of high saturated conductivity parameter values

### *Variability in discharge output volumes*

Results from the analysis of variability of discharge from ensemble model runs based on high saturated conductivity values for each soil type in the 5% and 95% confidence limits had a median variability of  $\pm 46\%$  ranging from 38 to 47% with a mean of 45% (Table 5.24). Considering the mean 45% within the range of 38 to 47% the variability of discharge was relatively evenly distributed through flood volumes had a much broader band in lower limit than for the upper limit, particularly for less frequent probabilities.

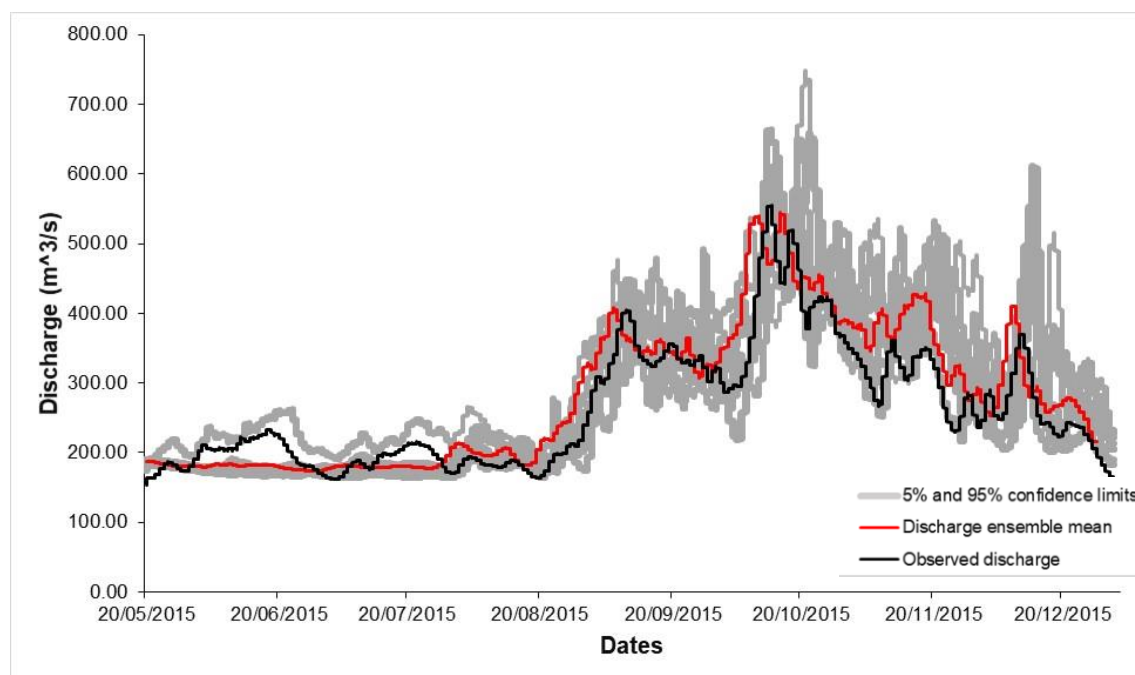
Compared to discharge volumes generated from the ensemble model runs based on the calibrated parameter scenario, the variability for the 5% confidence limit decreased by 4% whereas the 95% confidence limit the variability decreased by 16%. The result showed that the variation of discharge decreased across all probabilities when the ensemble model runs are based on high saturated conductivity values for each soil type in the study area. Results show a median Relative Interval Length (RIL) of 0.48 and a mean of 0.49, which is an increase of 9% and 4% respectively.

**Table 5:24 Rainfall, discharge and flood inundation extents uncertainty and errors quantification and comparisons between calibrated and high saturated conductivity values for each soil type in the study area.**

High saturated conductivity		UNCERTAINTY						ERROR			
		CV (%)				RIL		RMSE (mm & m <sup>3</sup> /s)			
		95%	mean	median	5%	mean	median	95%	mean	median	5%
Meteorological	Rainfall depth	63	70	71	73	0.49	0.47	0.45	0.46	0.46	0.47
Hydrological model	Discharge	38	45	46	47	0.49	0.48	31.9	59.2	59.51	83.7
	% change	-16%	0%	2%	-4%	9%	4%	7%	2%	10%	-9%
Hydraulical model	Statistical level	Discharge (m <sup>3</sup> /s)		Flood extents (km <sup>2</sup> )		RIL					
	5%	748.19	0.46%	4.02	-43%						
	95%	212.80	-54%	1.28	-82%	0.88					
	media	609.04	-2%	2.21	-68%						
	mean	585.39	-6%	2.45	-65%						

## Discharge error

To estimate discharge errors within 5% and 95% confidence limits, each discharge output ensemble based on the high saturated conductivity parameter values was compared with the observed discharge data (Figure 5.41). The high parameter values resulted in discharge output with errors of median variability of around  $\pm 59.51 \text{ m}^3/\text{s}$  ranging from  $31.88$  to  $83.74 \text{ m}^3/\text{s}$  and a mean error of  $59.19 \text{ m}^3/\text{s}$ . Compared to the model output based on the calibrated parameter scenario, the result showed a 7% increase in discharge in the high-frequency probabilities and a 9% decrease in low-frequency probabilities.



**Figure 5.41:** The 2015 hydrograph showing 5% and 95% confidence limits to represent the uncertainty of discharge due to high saturated conductivity values for each soil type in the study area.

## Variability in flood inundation extents

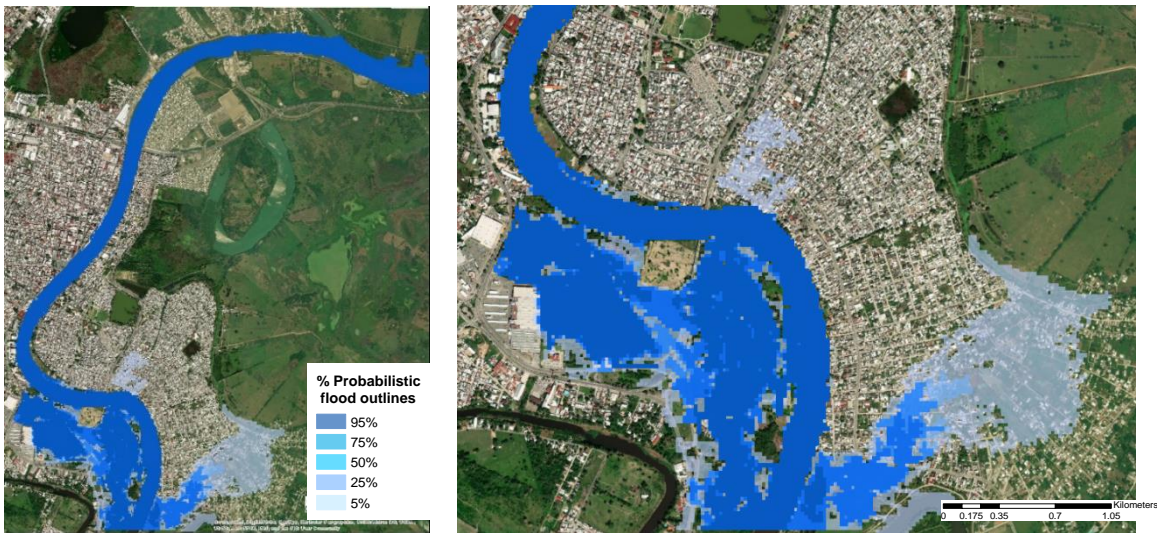
Results from the analysis of variability of flood extents due to high saturated conductivity values had a 5% and 95% confidence limit of  $4.02$  and  $1.28 \text{ km}^2$ , respectively. Compared to inundation extents based on calibrated parameter scenario, flooded areas decreased by 43% for the 5% confidence limit and 82% for the 95% confidence limit. The ensemble of flood inundation extent areas had a median variability of  $\pm 2.21 \text{ km}^2$ , which is a decrease of 68% and a mean of  $2.45 \text{ km}^2$ , another decrease of 65% compared to the calibrated model scenario.

The result reveals a relationship between discharge volumes and flood extents when the ensemble models are run based on high saturated conductivity values for each soil type in the study area. A decrease of 4% uncertainty of discharge in the less frequent probabilities is related to a decrease of 43% uncertainty in flood extents. Also, a decrease of 16% uncertainty in the more frequent probabilities is related to decreases of 82% in flooded areas (Table 5.24). The variability of flood extents was further estimated and summarised using the mean Relative Interval Length and was found to be higher (RIL=0.88) than rainfall input (RIL=47) and discharge output (RIL=48).

### ***Effect on flood extents***

Running the ensemble models based on the high saturated conductivity parameter reveals that flood extents decrease in both less and more frequently probabilities compared to the calibrated parameter scenario. The high saturated conductivity values had an effect of decreasing uncertainties by 43% discharge output, particularly in less frequent probabilities (5% confidence limit) which is a decrease of 7.02 to 4.02 km<sup>2</sup>.

However, for more frequent probabilities in the 95% confidence limit, discharge decreased by 16%, and this corresponds to decreases of 82% in flooded areas. In real terms, it was a decrease from 1.66 to 1.35 km<sup>2</sup> of inundated areas. Analysis results show that a 16% decrease in discharge output is associated with a 82 % decrease in flood areas. High saturated conductivity parameter values resulted in decreases of flooded areas by 43% in the 5% confidence limit and 82% in the 95% confidence limit valid within 1.28 km<sup>2</sup> to 4.02 km<sup>2</sup> of flooded areas (Figure 5.42).



**Figure 5.42: Probabilistic flood inundation outlines in the 5% and 95% confidence limits based on flow input into ensemble model runs based on high saturated conductivity values for each soil type in the study area.**

***The overall effect of high saturated conductivity parameter***

Analysis results have shown that in the 5% and 95% confidence limits, ensemble rainfall input data with a median variability of  $\pm 71\%$ , ranging from 63 to 73% when enforced and propagated through a cascading model chain based on high saturated conductivity parameters generates discharge output with a median variability of  $\pm 46\%$  ranging from 38% to 47% and with relative errors of  $\pm 60\%$  ranging from 32% to 84%. The propagation of uncertainties resulted in some shifts in peak discharge, having variabilities of  $\pm 609$  m<sup>3</sup>/s ranging from 213 to 748 m<sup>3</sup>/s and flood extents with a variability of  $\pm 2.2$  km<sup>2</sup> ranging from 1.3 to 4.0 km<sup>2</sup> of inundation areas (Figure 5.42).



### 5.5.10 The effect of low saturated conductivity parameter values

Results from ensemble model runs based on low saturated conductivity parameter values for each soil type in the study area show some shifts in the variability of discharge volumes in the 5% and 95% confidence limits. The discharge output uncertainties had a median variability of  $\pm 48\%$ , ranging from 40 to 49% and a mean of 50% (Table 5.25). It was observed that the variability of discharge was evenly dispersed around the mean with flood volumes having a much broader band in lower limit than in the upper limit, particularly for the less frequent probabilities. However, compared to discharge volumes generated from probability model runs based on calibrated parameters, the 5% confidence limit discharge volumes did not change (0%) while in the 95% confidence limit discharge volumes decreased by 11%.

Thus, results from the probability model runs based on low saturated conductivity parameter values showed that uncertainty in discharge decreased in high-frequency probabilities only and maintained the same variability in the low probabilities. Besides, a median Relative Interval Length (RIL) of 0.41 with a mean 0.46 was found from this scenario, showing a decrease of 11% when compared to a probabilistic calibrated model run scenario.

**Table 5:25 Rainfall, discharge and flood extents error and uncertainty quantification comparisons between calibrated and low saturated conductivity parameter values for each soil type in the study area.**

Low saturated conductivity		UNCERTAINTY						ERROR			
		CV (%)		RIL		RMSE (mm & m <sup>3</sup> /s)					
		95%	mean	median	5%	mean	median	95%	mean	median	5%
Meteorological	Rainfall depth	63	70	71	73	0.49	0.47	0.45	0.46	0.46	0.47
Hydrological model	Discharge	40	50	48	49	0.46	0.41	34.9	79.6	65.5	67.1
	% change	-11%	11%	7%	0%	2%	-11%	17%	38%	21%	-27%
Hydraulical model	Statistical level	Discharge (m <sup>3</sup> /s)		Flood extents (km <sup>2</sup> )		RIL					
	5%	1096.00	47%	13.91	98%						
	95%	490.36	5%	1.82	10%	0.94					
	media	645.56	4%	5.05	16%						
	mean	689.79	11%	5.83	34%						

### ***Discharge error***

Discharge errors were calculated from comparisons between the discharge ensemble mean and observed discharge recorded at the Las Gaviotas Gauging Station. Uncertainty analysis results of discharge ensembles across the 5% and 95% confidence interval had a median variability of  $\pm 65.5$  m<sup>3</sup>/s ranging from 34.89 to 67.10 m<sup>3</sup>/s and a mean of 79.63 m<sup>3</sup>/s. The low saturated conductivity values did not increase uncertainties in discharge, particularly for the less (5%) frequent probabilities, and this is related to an increase of 98% in flood extent areas. In real terms, it is an increase from 7.02 km<sup>2</sup> to 13.91 km<sup>2</sup>. Results show that saturated conductivity parameter values caused some shifts in the variability of discharge errors. Unlike discharge from the calibrated parameter model run scenario, the shift resulted in error increases of 17% for high frequent probabilities and decreases by 27% for low frequent probabilities.

### ***Variability in flood inundation extents***

The simulated flood extents from ensemble model runs based on low saturated conductivity parameter values had a 5% and 95% confidence limits of 13.91 and 1.82 km<sup>2</sup>, respectively. Compared to the probability model runs based on calibrated parameters reflecting current catchment conditions, the size of the flooded area increased by 98% in the 5% confidence limit and 10% in the 95% confidence limit. The flooded areas had a median variability of  $\pm 5.05$  km<sup>2</sup>, an increase of 16% with a mean of 5.83 km<sup>2</sup> again, an increase of 34% compared to the calibrated probabilistic model scenario.

The results revealed a link between the variability of discharge volumes and the variability in inundated flood extents. Analysis results show that 0% in the variability of maximum discharges is related to increases of 98% in flood extents. The absence of variability in maximum flows is associated with a 98% increase in flood areas when the hydrological model is run based on low saturated conductivity values for each soil type in the study area. The relative distribution of uncertainty reported by the coupled models at each stage, given by the Relative Interval Length, was found to be 0.94. The RIL value obtained in this scenario was the highest (RIL = 0.94) compared to rainfall input (RIL = 0.47) and discharge data (RIL = 0.48). The RIL summarizes the uncertainty reported by the model by normalising uncertainty between interval sizes at each stage of the cascading model.

### ***Effect on flood extents***

Analysis results show that when the ensemble model is run on low saturated conductivity parameter values, flood extent areas are reduced, particularly in the more frequent probabilities (95%). A decrease of 11% in discharge is related to a 10% increase in inundated areas. In real terms, it is an increase from 1.66 km<sup>2</sup> to 1.82 km<sup>2</sup>. However, for less frequent probabilities in the 5% confidence limit, discharge volumes did not change but remained similar to model runs on the calibrated parameter; however, linked to increases of 98% in the flooded areas. This means decreases of 11% uncertainty of minimum flows is related to decreases of 10% uncertainty of inundation areas. In comparison, no shifts in less frequent probabilities resulted in increases (98%) in flooded areas.

### ***The overall effect of low saturated conductivity parameter***

The analysis of the effect of low saturated conductivity parameter values has shown that in the 5% and 95% confidence limits, ensemble rainfall input data with a variability of  $\pm 71\%$ , ranging from 63 to 73% when enforced and propagated through a cascading model generates discharge of  $\pm 48\%$  ranging from 40% to 49% with relative errors of  $\pm 66\%$  ranging from 35% to 67%. The propagation causes some shifts on peak discharge uncertainty to  $\pm 646$  m<sup>3</sup>/s ranging from 490 to 1 096 m<sup>3</sup>/s and variability in flood extents to  $\pm 5.1$  km<sup>2</sup> ranging from 1.8 to 13.9 km<sup>2</sup>.

#### **5.5.11 The effect of high saturated water content parameter values**

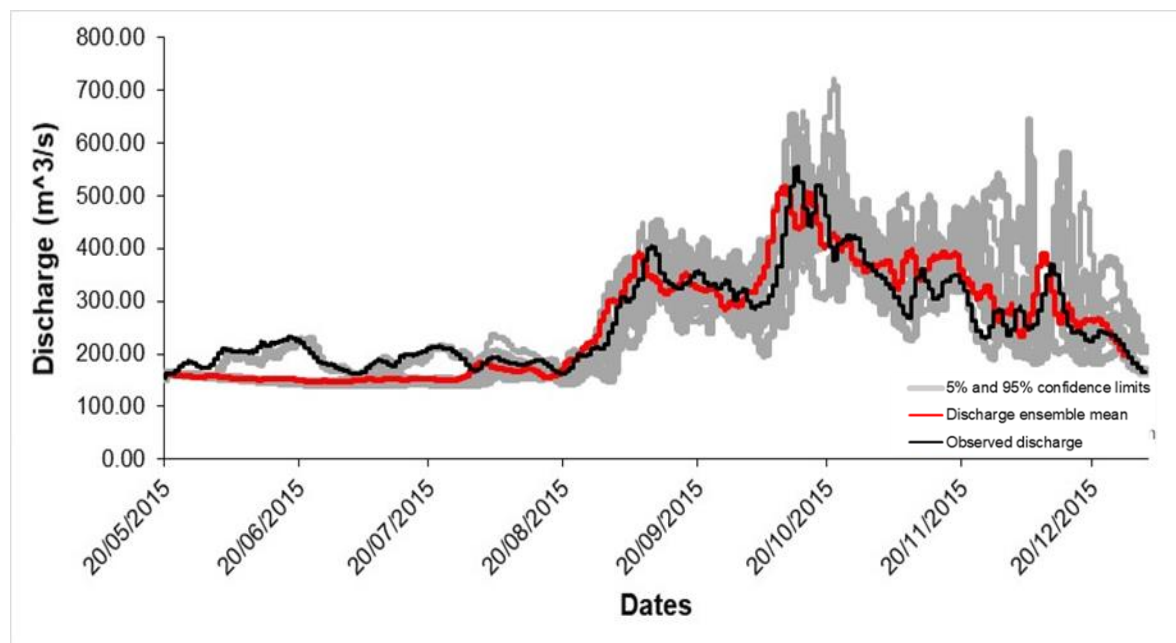
The discharge output results from ensemble model runs based on high saturated water content parameter values show a median variability of  $\pm 43\%$ , ranging from 38 to 45% and a mean of 44% within the 5% and 95% confidence limits. The result shows a decrease in the variability of discharge across the confidence limits when the ensemble model is run based on high saturated water content parameters values. Compared to the discharge output from the calibrated parameter scenario, the variability of discharge in the 5% confidence limits decreased by 8%, while in the 95% confidence limit decreased by 18%. Results show a median Relative Interval Length (RIL) of 0.48 and a mean of 0.5, which is an increase of 11% when compared to the calibrated parameter model run scenario (Table 5.26)

**Table 5:26 Rainfall, discharge and flood inundation extents uncertainty and error quantification and comparisons between calibrated and high saturated water content parameter values for each soil type in the study area.**

High saturated water content		UNCERTAINTY				ERROR					
		95%	mean	CV (%) median	5%	mean	RIL median	RMSE (mm & m <sup>3</sup> /s)			
Meteorological	Rainfall depth	63	70	71	73	0.49	0.47	0.45	0.46	0.46	0.47
Hydrological model	Discharge	38	44	43	45	0.5	0.48	31.9	58.5	58.5	83.7
	% change	-16%	-2%	-4%	-8%	11%	4%	7%	1%	8%	-9%
Hydraulic model	Statistical level	Discharge (m <sup>3</sup> /s)		Flood extents (km <sup>2</sup> )		RIL					
	5%	720.12	-3%	0.94	-87%						
	95%	476.44	2%	0.04	-97%	0.42					
	media	584.30	-6%	0.40	-91%						
	mean	597.04	-4%	0.48	-89%						

### Discharge error

Discharge outputs resulting from model runs based on high saturated water content parameter values were compared to observed discharge values to estimate the relative errors across 5% and 95% confidence limits (Figure 5.43). Results show that discharge errors across the confidence intervals had a median variability of  $\pm 58.5$  m<sup>3</sup>/s ranging from 31.88 to 83.74 m<sup>3</sup>/s and a mean of 58.52 m<sup>3</sup>/s. Compared to the calibrated parameter scenario, the result shows an error increase of 7% in the high-frequency probabilities and a decrease of 9% in low frequent probabilities.



**Figure 5.43: The flood hydrography showing 5% and 95% confidence limits to representing the uncertainty of discharge due to high saturated water content parameter values. The red line represents the ensemble mean of discharge while the black is the observed discharge.**

### ***Variability in flood inundation extents***

The simulated ensembles of flood extent outlines due to high saturated water content values had a 5% and 95% confidence limits of 0.94 and 0.04 km<sup>2</sup>, respectively. Results show that high saturated water content parameter decreased by 87% for the 5% confidence limit and 97% for the 95% limit when compared to flood inundation levels based on the calibrated parameter scenario. Also, the flood inundation extent areas due to high saturated water content values showed a median variability of  $\pm 0.40$  km<sup>2</sup>, which is a decrease of 91% and a mean of 0.48 km<sup>2</sup> compared to the calibrated parameter scenario.

The result shows relationships between discharge and inundation areas where a decrease of 8% uncertainty in discharge for the 5% confidence limit is related to a decrease of 87% uncertainty in the inundation areas. A decrease of 16% uncertainty for more frequent probabilities with minimum discharge volumes is related to the most substantial decrease of 97% uncertainty in inundation areas.

To summarise uncertainty reported by the cascading model, the Relative Interval Length was found to be 0.42. The RIL value obtained for flood extents was low compared to that of rainfall input data (RIL=47) and discharge output data (RIL=48).

### ***Effect on flood extents area***

Results show that high saturated water content parameter values generate consistent decreases in flooded extent areas for both less frequent probabilities (5%) and more frequent probabilities (95%). However, for more frequent probabilities discharge, decreases of 16% are equivalent to decreases of 97% of flood extents. In real terms, these are decreases from 1.66 km<sup>2</sup> to 0.94 km<sup>2</sup>. Also, a decrease of 8% uncertainty in maximum flows is related to decreases of 87% uncertainty of inundation areas. These are decreases from 7.02 km<sup>2</sup> to 2.76 km<sup>2</sup> of flooded areas.

Unlike the calibrated parameter output scenario, results show that the discharge output based on the high saturated water content parameter values was consistently reduced across 5% and 95% confidence limits. It was observed that a decrease of 9% in flood extent areas in the 5% confidence limit is linked to increases of 7% in flooded areas in the 95% confidence thresholds ranging from 0.04 to 0.94 km<sup>2</sup>.

### ***The overall effect of the high saturated water content parameter values***

The uncertainty analysis results have shown that given the 5% and 95% confidence limits, an ensemble rainfall input data with median variability  $\pm 71\%$ , ranging from 63 to 73% when propagated through a hydrological model based on high AET/PET parameters it generates discharge volumes with a median variability of  $\pm 43\%$  ranging from 38% to 45% and errors of  $\pm 59\%$  ranging from 32% to 59%. In real terms, the propagated uncertainties made shifts on discharge output with a median variability of  $\pm 584 \text{ m}^3/\text{s}$  ranging from 476 to 720  $\text{m}^3/\text{s}$  and flood extents of  $\pm 0.4 \text{ km}^2$  ranging from 0.04 to 0.94  $\text{km}^2$ .

### 5.5.12 The effect of low saturated water content parameter values

Results from the ensemble model runs based on the lowest saturated water content parameter values for each soil type in the study area showed some variability of discharge volumes across the 5% and 95% confidence limits. The discharge output had a median variability of  $\pm 44\%$ , ranging from 38 to 46% and a mean of 45% (Table 5.27). The result shows that the variability of discharge was relatively evenly distributed around the ensemble mean with a much broader band in lower limit than in the upper limit, especially at lower frequent probabilities (Figure 5.38).

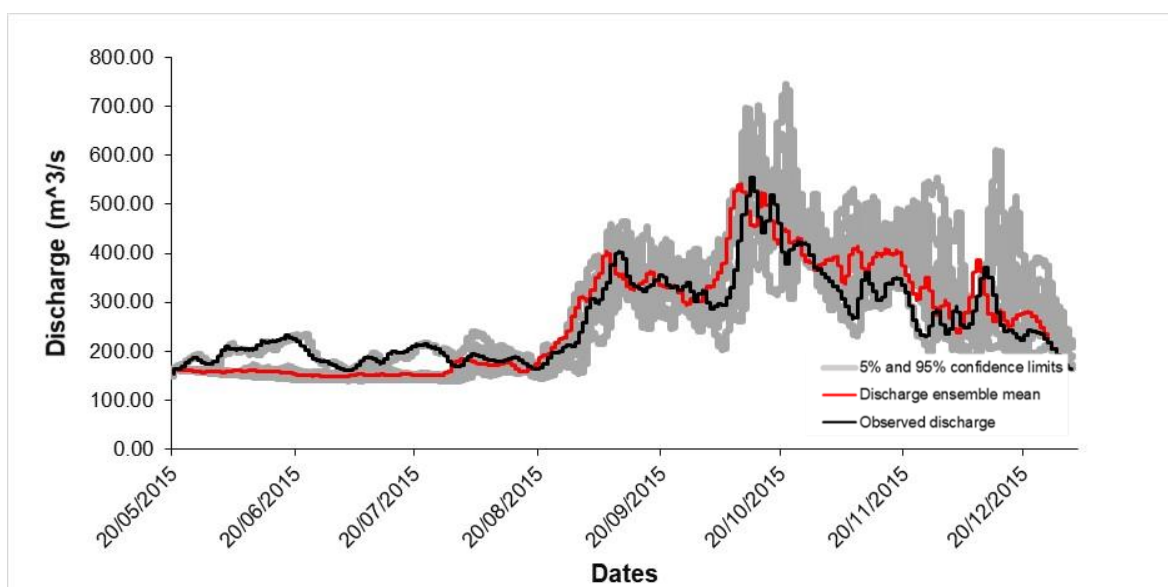
Unlike the variability of discharge output volumes generated from model runs based on the calibrated parameter scenario, the variability for the 5% confidence limit decreased by 6% and for the 95% confidence limit, also decreased by 16%. The result showed that uncertainty in discharge decreased across the confidence interval. A median Relative Interval Length (RIL) of 0.48 and a mean of 0.49 was found in this scenario, showing an increase of 4% compared to ensemble model runs based on calibrated parameters.

**Table 5:27 Rainfall, discharge, flood extents uncertainty and error quantification and comparison between calibrated and low saturated water content parameter values for each soil type in the study area.**

Low saturated water content		UNCERTAINTY				ERROR					
		CV (%)				RIL		RMSE (mm & m <sup>3</sup> /s)			
		95%	mean	median	5%	mean	median	95%	mean	median	5%
Meteorological	Rainfall depth	63	70	71	73	0.49	0.47	0.45	0.46	0.46	0.47
Hydrological model	Discharge	38	45	44	46	0.49	0.48	31.9	60.6	59.21	87.7
	% change	-16%	0%	-2%	-6%	9%	4%	7%	5%	9%	-5%
Hydraulical model	Statistical level	Discharge (m <sup>3</sup> /s)		Flood extents (km <sup>2</sup> )		RIL					
	5%	704.75	-5.37%	6.82	-3%						
	95%	493.24	5.85%	1.85	11%	0.31					
	media	620.11	-0.30%	4.47	3%						
	mean	625.95	0.31%	4.58	5%						

### **Discharge error**

Discharge outputs resulting from model runs based on low saturated water content parameter values for each soil type in the study area were compared to the observed discharge to estimate relative errors of each ensemble member within the 5% and 95% confidence limits (Figure 5.44). Results show that the low saturated water content parameter conditions resulted in some shifts in discharge errors. The discharge errors had a median variability of  $\pm 59.21 \text{ m}^3/\text{s}$  and a mean of  $60.57 \text{ m}^3/\text{s}$  ranging from  $31.89$  to  $87.74 \text{ m}^3/\text{s}$ . Compared to the probability model runs based on the calibrated parameter scenario, it is a shift of 7% increase in the high-frequency probabilities and is a decrease of 5% in low frequent probabilities.



**Figure 5.44: The 2015 hydrography showing 5% and 95% confidence limits to represent uncertainty discharge due to low saturated water content parameter values. The red line represents the ensemble mean of discharge while the black is the observed discharge.**

### **Variability in flood inundation extents**

Results show that the simulated flood extent outlines due to ensemble model runs on low saturated water content parameters had a 5% and 95% confidence limits of  $6.82$  and  $1.85 \text{ km}^2$ , respectively. Compared to inundation extents based on the calibrated parameter scenario, the flood inundation extent areas decrease by 3% for the 5% limit and increased by 11% for the 95% limit. The spatial extent of inundation areas had a median variability of  $\pm 4.47 \text{ km}^2$ , which is a decrease of 3% compared to the calibrated parameter scenario with a mean of  $4.58 \text{ km}^2$ .



The result showed some relationships between variability of discharge output and the variability of flood extent areas in which a decrease of 6% uncertainty in maximum discharge is linked to a decrease of 16% uncertainty in inundation areas. Also, a decrease of 6% uncertainty in maximum flows in the less frequent probabilities is related to decreases of 3% uncertainty in inundation areas. Results showed a Relative Interval Length of 0.31 for the flood extents, which is low compared to rainfall input (RIL = 47) and discharge data (RIL = 48) uncertainties.

### ***Effect on flood extents***

Results show some shifts in the variability of discharge due to low saturated water content parameter characterised by a reduction in uncertainty in flooded inundation areas, particularly in less frequent probabilities (5%) however with increases in more frequent probabilities (95%). The low saturated values resulted in decreases of 5% in uncertainty in maximum discharge in the less (5%) with a similar decrease of 3% in flood inundation areas, and it is a decrease from 7.02 km<sup>2</sup> to 6.82 km<sup>2</sup>.

This means a decrease of 6% uncertainty in maximum discharge is related to decreases of 3% in inundation areas. It seems possible that a decrease from 7.02 km<sup>2</sup> to 6.82 km<sup>2</sup> occurs when low saturated water content conditions characterise catchment conditions. In more frequent probabilities (95% confidence limits) with minimum flows, low saturated water content parameter decreases of 16% are related to an increase of 11% in inundated areas. In terms of coverage, it is an increase from 1.66 km<sup>2</sup> to 1.85 km<sup>2</sup> compared to the calibrated ensemble run scenario.

### ***The overall effect of low saturated water content parameter***

The uncertainty analysis results have shown that in the 5% and 95% confidence limits, rainfall ensemble input with median variability of  $\pm 71\%$ , ranging from 63 to 73% when enforced and propagated through a coupled hydrological and hydraulic model run based on low saturated water content parameter it generates a discharge with a median variability of  $\pm 44\%$  ranging from 38% to 46% and relative errors of  $\pm 59\%$  ranging from 32% to 88%. The result shows that some shifts in discharge output with a median variability of  $\pm 620 \text{ m}^3/\text{s}$  ranging from 493 to 705  $\text{m}^3/\text{s}$  and variability of flood extents of  $\pm 4.5 \text{ km}^2$  varying from 1.9 to 6.8 km<sup>2</sup>.

## 5.6 Summary

This study used a Monte Carlo simulation technique to propagate rainfall and parameter uncertainties through coupled hydrological and hydraulic models and quantified the magnitude of uncertainties and effects on flood volumes and inundation extents. The coupled model chain comprised the SHETRAN hydrological model and the Flood Modeller, a 1D/2D flood inundation model. The coupled models were enforced with rainfall and discharge ensembles based on two scenarios comprising fixed or calibrated parameter and the other on the variation of parameter values for each vegetation and soil category in the La Sierra catchment area.

The propagation of uncertainty showed the propagation and quantification of uncertainties in discharge output and flood inundation extents and including the probabilistic flooding rather than using a single deterministic flood outline. Results showed that given rainfall uncertainties with a median variability of  $\pm 71\%$ , ranging from 63 to 73% if enforced into a hydrological model and propagated generates discharge volumes with uncertainties of  $\pm 46\%$ , ranging from 45% to 49% with relative errors of  $\pm 46\%$  ranging from 45% to 46%. The propagated uncertainties resulted in some shifts in peak discharge with a median variability of  $\pm 622 \text{ m}^3/\text{s}$  ranging from 466 to 744.75  $\text{m}^3/\text{s}$  and inundation extents of  $\pm 4.34 \text{ km}^2$  varying from 1.66 to 7.02  $\text{km}^2$ .

Also, the results showed that the most significant amount of uncertainty existed in rainfall input data and these uncertainties were propagated, disaggregated and accumulated in different amounts and probabilities across 5% and 95% confidence interval. However, uncertainties in rainfall were not translated into more substantial uncertainties in flood volume outputs and flood extents. Uncertainty propagation due to parameter changes did not occur in equal magnitudes across all probabilities; instead, uncertainties were translated differently across the probability distribution limits given. The results showed that high AET/PET parameter values had a significant effect on flood extents, with the most significant decrease in flood volumes and extents, while low saturated hydraulic conductivity has had the most extensive flood ranges.

## **6. DISCUSSION**

This chapter discusses the identification and application of physics-based and data-based hydrological modelling approaches for large-scale flood modelling in data-scarce tropical regions. The discussion is organised into three subsections drawing on the results presented from Chapter 5. First, the results of integrating satellite datasets into the SHETRAN distributed hydrological model based on error quantification and correction approaches to improve the simulations of the magnitude and timing of flood flows are presented (section 6.1). Next presented is a discussion on the findings of the investigation of the power-law relationship between peak discharge and physical catchment conditions and processes in nested catchments in the La Sierra catchment area (section 6.2). Finally, the results of rainfall and parameter uncertainty propagation through a coupled hydrological and hydraulic modelling approach and effects on the magnitude of peak discharge and flood inundation in a selected high flood risk area in the study area are discussed (section 6.3).

### **6.1. Distributed hydrological modelling**

The first research objective investigated the integration of satellite datasets into the SHETRAN distributed hydrological model based on error quantification and correction techniques to improve the magnitude and timing of flood peak discharge in the Grijalva catchment. In this objective, the study responded to the challenge of providing reliable simulations of river flow dynamics, magnitudes and timing, contributing to flood risk management in large data-poor lowland tropical regions.

Given the rapidly increasing availability of satellite datasets and the power of computers, this study applied a fully distributed hydrological modelling approach integrated with high-resolution satellite datasets to improve the simulation of river flows in a large data-poor lowland tropical catchment. In this way, the present study overall contributes to solving the current problem of flood prediction in data-poor regions (Sivapalan et al., 2003).

The following reflects on the main results from integrating satellite datasets into the SHETRAN distributed hydrological modelling, based on error quantification and correction to improve flood discharge estimation: 1) the high-resolution distributed modelling approaches over large domains, 2) the maximum use of satellite datasets in large-scale modelling, 3) rainfall error correction and distributed modelling approach 4) performance of large-scale distributed hydrological models over large domains and 5) increased understanding of the dynamics of tropical flood flows.

### **6.1.1. High-resolution distributed modelling approaches over large domains**

The study has shown that the use of large-scale and high-resolution distributed hydrological modelling approaches in tropical lowland regions has the potential to improve flood simulations (Figure 5.10 and 5.11; Appendix C Figure C 1). The results show that accurate representation of large-scale catchment conditions and processes using high-resolution, fully distributed hydrological models integrated with high spatial (< 2 km) and temporal (<3 hr) resolution satellite datasets can provide reasonable levels of flood simulation accuracy in large lowland tropical regions.

The high-resolution distributed modelling approach over large domains pursued in this study contributes to closing the knowledge gap that exists between simplified, coarse resolution, large-scale modelling approaches on one hand and the detailed high-resolution, small-scale hydrological modelling approaches on the other (Sampson et al., 2015). Most large-scale distributed hydrological models are applied at low spatial resolutions using simplified representations of catchment conditions and processes. While the approach is suitable where data is scarce, it results in simulation errors and uncertainties in model results (Paiva et al., 2013; Ibarra-Zavaleta et al., 2017; Yang et al., 2018). Conversely, the high-resolution distributed hydrological modelling approach with accurate representations of catchment conditions and processes has been successfully applied, but mostly only at small scales (Khan et al., 2010; Birkinshaw et al., 2014; Zhang et al., 2015; Mourato et al., 2015; Dukic and Radic et al., 2016 and 2014; Op de Hipt et al., 2017). The application of high-resolution distributed modelling approaches on a large-scale is very rare due to limited data availability, computer processing time limitations and cost-time limitations. The results show that the application of distributed modelling approach in large tropical lowland regions is feasible

by making use of the growing available satellite data complemented by locally available in-situ datasets (Table 4.3). The results are comparable to a recent example of the application of a large-scale semi-distributed hydrological model developed by Paiva et al. (2013) on the Amazon River, enforced with daily rainfall input data at 5 km spatial resolution. Although the model provided predictive solutions for its intended purpose, it generated simulations with low accuracy of spatial-temporal variabilities of catchment conditions and processes (Schumann et al., 2013). However, high-resolution distributed hydrological models have been successfully applied in some data-rich countries in humid temperate regions for flood forecasting purposes; however, most have been developed at small scales (Clark and Connolly, 2012). <

There are few examples of fully distributed hydrological models applied in tropical regions that can be compared to the size of the current model developed in the Grijalva catchment area (57,960 km<sup>2</sup>) and at a fine 2 x 2 km<sup>2</sup> grid spatial and 3-hour temporal resolutions. Most distributed modelling approaches are applied at medium to small scales in tropical regions, examples include the CREST model applied to Nzoia sub-catchment (12 900 km<sup>2</sup>) in the Lake Victoria basin, enforced with daily rainfall, at 1 km<sup>2</sup> grid resolution (Khan et al., 2010). The SHETRAN model applied to the Dano catchment (126 km<sup>2</sup>) Burkina Faso at 200 x 200 m<sup>2</sup> spatial and 1 hr temporal resolutions (Op de Hipt et al., 2017) and the GSSHA model applied in the sub-tropical Guadalupe River basin (11,285 km<sup>2</sup>) in Texas at 150 m x 150 m grid and 15-min temporal resolutions (Chintalapudi et al., 2017).

Most of these high-resolution models were applied on a relatively small scale, and those implemented on a large scale were low-resolution and reported to be susceptible to rainfall and parameter errors due to lack of full use of the distributed data to estimate parameter values (Ibarra-Zavaleta et al., 2017; Paiva et al., 2013). The application of the distributed hydrological modelling approach requires a range of ground observation datasets; however, the availability of sufficiently high-resolution in-situ data sets is a challenge (Jiang et al., 2019). The poor performance of distributed hydrological models is partly due to a lack of high-resolution in-situ data for an accurate representation of catchment conditions and processes in the model (Aadhar et al., 2017).

This research identifies the scarcity of in-situ data as an obstacle that hinders the implementation of large-scale distributed hydrological modelling approaches in data-

scarce tropical regions (Smith et al., 2012; 2004). However, given the availability of high-resolution satellite datasets and computing power, the performance results of the distributed hydrological models can be enhanced by merging satellite datasets with a few available in-situ datasets and other available complementary data information.

### **6.1.2. Full use of satellite datasets for the SHETRAN model**

Despite recent advances in the availability of satellite data and computational power, the frequently used lumped hydrological modelling approaches are not efficient in using the wealth of increasing satellite data for flood modelling. As a result, satellite datasets have not been used as widely as expected, despite the increasing number of such datasets and the variety of satellite sensors that have been sent into space. This study shows that current satellite data availability can meet essential data requirements for large-scale distributed modelling (Table 4.1). The publicly available satellite datasets could complement the scarce in-situ datasets and provided continuous hydrometeorological data for distributed hydrological modelling in data-poor regions. Unlike previous small-scale hydrological modelling methods using one or two satellite datasets, this study used a fully distributed model and met most of its data requirements using satellite datasets, complemented by few locally available in-situ datasets (Mazzoleni et al., 2019).

However, it was observed that the SHETRAN hydrological model does not assimilate all its data as spatially gridded datasets. Instead, some data are presented in the model as time series and others with lumped parameter values. Lumped parameter values are averaged spatial characteristics of catchment properties, and processes covering the entire catchment or sub-catchments (Brirhet and Benaabidate, 2016). For example, in the SHETRAN hydrological model, rainfall and Potential Evapotranspiration (PET) data are represented as time series while soil moisture is represented as lumped values making integration of gridded satellite dataset less effective. Although high-resolution soil moisture satellite datasets were reported to improve model states and simulations, they are not represented as spatially gridded data in the SHETRAN model structure (Brocca et al., 2014; Petropoulos et al., 2015; Zhang et al., 2015; MacLeod et al., 2016; Zhuo, Han & Dai 2016; Li et al., 2016). Soil moisture data values are represented in the SHETRAN hydrological model using the two van Genuchten parameters  $\alpha$  and  $n$ , the soil water content ( $\theta$ ), the saturated soil moisture content and the residual soil moisture

content parameters. Therefore, there is a need to modify the previously developed distributed hydrological models so that they can efficiently integrate new forms of remote sensing data, particularly the gridded satellite data sets for soil moisture, rainfall and parameter data.

### **6.1.3. Application of SHETRAN distributed hydrological model**

This research developed a process-based fully distributed hydrological model, integrated with high - resolution satellite datasets, based on error quantification and correction in a large lowland tropical catchment area. The results of the study contribute to the ongoing debate on the strengths and weaknesses of the physical spatially distributed hydrological models for large-scale flood modelling. Many arguments have been put forward in the literature concerning the theoretical underpinnings of physical models, which appear to relate to questions of scale, complexity, data requirements, catchment size and parameter estimation. The results show that the potential applications, resulting from the implementation of the distributed hydrological approach in large tropical regions are significant given the increasing availability of satellite data, computing power, cloud systems and the enhancement of the Geographic Information Systems (GIS).

The results show that distributed hydrological models (represented by the SHETRAN hydrological model) have a satisfactory performance when applied on a large-scale and the results compare well with similar studies, although applied at small scales (Table 5.4) (Tripkovic et al., 2014; Zhang et al., 2013; Birkinshaw et al., 2014; Mourato et al., 2015) Naseela et al., 2015; Mourato et al., 2015; Zhang et al., 2015; Đukic and Radic et al., 2016 and 2014; Lewis, 2016; Op de Hipt et al., 2017; Janes et al., 2018; Shrestha et al., 2017; Sreedevi and Eldho, 2019). Results show that the SHETRAN hydrological model can reasonably simulate Grijalva River seasonal flow patterns, and the simulated outputs mirrored the observed flow dynamics (Figure 5.10 and 5.11; Appendix C Figure C 1). The model's ability to simulate seasonal flow patterns shows that the model concept and parameterisation are sufficiently representative and can simulate temporal and spatial flow dynamics in the study area. The SHETRAN model's potential ability to simulate medium to long-term river flow dynamic trend shows that the model is most likely to be useful for climate studies, but more work is needed to compare predictive differences with other models. Coupling the model with a Regional Weather Model

generator could provide some insights and produce a set of simulations with a range of scenarios for analysis and evaluation.

Although the SHETRAN hydrological model was able to simulate seasonal flow patterns, it had difficulties simulating certain high discharge events. The simulated flood peaks compared well in terms of timings with the observed discharge, except that the model could not accurately capture the magnitude of some large peak discharge. This is evident at the end of each rainfall season where simulated peaks are not sufficiently flashy compared to the observed (Figures 5.7, 5.10 and 5.11; Appendix C, Figure C 1). The observed end of season sharp rises and the 'big steps down' flows regularly experienced at the end of each wet season are likely a result of releases of excess stored water from upstream reservoirs and dams, also due to end of season flow or flood regulation at the El Mecayo barrier (Appendix C, Figure B 13). However, due to the sensitivity of dam releases operations that are associated with recurrent flooding problems in downstream communities, all release information was listed as sensitive information in the Tabasco State and all attempts to check on this assumption (of periodic releases) were unsuccessful.

The general underestimation of the magnitude of the peak discharge could also be attributed to the spatial segmentation of rainfall fields using Thiessen polygons. Using Thiessen polygons in SHETRAN may cause under-representation of short - term rainfall intervals that generate peak discharge. The approach may not be sufficient to account for localised rainfall events. The Thiessen polygon method cannot accurately represent the high spatial variation of short - term rainfall, particularly in mountainous regions where topography accentuates rainfall intensity, making it highly variable over short distances. Thiessen polygons characterise spatially distributed rainfall in the study area as abrupt transitions between arbitrary polygon boundaries that are unrealistically straight-lined. It is, thus, not possible to obtain accurate short - term rainfall input data from rain gauge data averaged in polygons as it does not represent the highly spatial variations of rainfall-runoff process in the catchment and results in an underestimation of the magnitude of discharge. Interpolation methods such as Kriging may also be inadequate to reflect local storms as they smooth rainfall intensities over larger spatial ranges of low rain gauge densities.



Nonetheless, a combination of calibration, correction of input data error and accounting for diverted flows has resulted in some improvement in the simulation of some peak discharge (Table 5.3). Results show that simulation improvements could not be attributed solely to rainfall input data errors, but also parameters and discharge outputs. The combination of rainfall input, error quantification and removal/correction, the accounting of diverted flows and the calibration of the most sensitive parameters was critical.

However, calibration processes could not inspect the full spectrum of parameter values and address simulation errors in their entirety. The calibration results could show a data error compensation problem where data errors are concealed during calibration (Ajami et al., 2004). The model has several parameters that allow the model to run under different validation conditions with significant uncertainties. It was also found that model parameters could be over-calibrated to compensate for errors, which could cause deleterious effects if the calibrated model were to be used in different scenarios. However, input data errors could be fully compensated for with some propagated into model results (Xu et al., 2017). This is consistent with section 4.1 observation that parameters could be over-calibrated and could account for errors from other sources.

Nevertheless, results show that distributed hydrological models (represented by the SHETRAN hydrological model) have a satisfactory performance for large-scale modelling using satellite data and the performance compares well to other studies that use the same model, though applied at a small-scale. This result provides insight into the potential application of large-scale distributed models in large domains, given the availability of computer resources and the increasing quality and quantity of satellite datasets. The potential application of these detailed models contributes to meeting the challenge of accurate representation of highly variable catchment conditions and processes in models and the provision of reliable flood simulations in large data from scarce tropical regions where an increasing number of people are at risk of flooding. Results of this study have shown that accuracy in flood modelling and provision of reliable flood simulations in data-scarce tropical regions can be realised through optimally integrating several datasets using distributed modelling approaches however with the best combination of the model parameter, input and output datasets based on error and uncertainty quantification and removal.

#### **6.1.4. Rainfall error correction and distributed modelling**

Rainfall is a critical data input in hydrological models, and it is impossible to produce accurate discharge simulations if the model is enforced with rainfall data that is not bias or error corrected (Zhang et al., 2018; McMillan et al., 2011). The results are consistent with observations suggesting that rainfall is a significant cause of errors in distributed hydrological modelling in which precise estimates of spatially and temporally distributed rainfall are required (Heistermann et al., 2011; Thiemig et al., 2011; Thiemig et al., 2013; Chen and Wang, 2018; Deng et al., 2013). Results showed that irrespective of how much calibration goes into the model, enforcing the model with inaccurate rainfall input data results in errors that get propagated into model simulation results.

The preliminary model runs of this study using rainfall input data that was not error corrected resulted in the model over predicting with a PBIAS of over 20% (section 5.3). However, after running the model using bias-corrected rainfall, simulations with a substantial reduction of errors were generated such that the model was under predicting (-8.18%). Results also show the same under-prediction due to the use of satellite rainfall data alone as input data and this confirms findings by Dinku et al. (2014) who reported that satellite rainfall data underestimate rainfall when compared to rainfall data (Grimes et al., 1999). Likewise, the results show that it is impossible to obtain accurate simulations either from rain gauge data alone or from satellite data.

Thus, each type of rainfall data was seen to have distinct advantages and disadvantages. Satellite rainfall data underestimate the intensity of convection rainfall compared to rain gauges and considered to be less reliable, although it had the advantage of providing large and continuous rainfall coverage (Dinku et al., 2014; Grimes et al., 1999). While rain gauge rainfall data were precise at point location, it had few point locations in the study area and less coverage. The combination of the two types of rainfall measurements (rain gauge and satellite) was able to take the maximum advantage of the two datasets. As a result, the available rain gauge and satellite rainfall datasets were merged to enhance their rainfall information.

The sparse rain gauge distribution in the catchment area required the use of large-scale satellite rainfall data to supplement the scarcely available in-situ datasets requiring a closer look at data merging techniques to make better use of available datasets. To

enhance rainfall information, geostatistical techniques were employed to interpolate, quantify and reduce errors and improve rainfall input data for the model (section 4.3.1). Although spatial coverage of rain gauges was uneven, results showed that merging rain gauge rainfall and satellite rainfall data significantly improved the distribution of rainfall in the catchment (section 5.3).

Besides, the present study adds additional evidence that the lack of sufficiently high spatial and temporal resolution data for locally occurring short-term rainfall is one of the significant sources of error in distributed hydrological modelling results (Niemczynowicz, 1988; Notaro et al., 2013; Cristiano et al., 2017). Although the temporal rainfall dimension was not tested, the results show that uneven spatial coverage or limited distribution of rain gauges in several parts of the Grijalva catchment area was associated with high levels of rainfall error compared to areas with a dense rain gauge distribution (Figure 5.2). Flood modelling accuracy can, therefore, be achieved by optimally integrating remote sensing and in-situ datasets to obtain the best combination of model data based on error quantification and correction. Accuracy and reliable flood simulations can be achieved through the optimal and efficient integration of high-resolution satellite data and available in-situ data using distributed modelling approaches.

### **6.1.5. A greater understanding of the dynamics of tropical flood flows**

Compared to mid and high latitude rivers, few studies have been done to model the spatial-temporal variability of hydrologic and hydraulic dynamics and interactions of river flood flow in large tropical regions (Hidayat et al., 2017; Wohl et al., 2012). In contrast, all components of hydrology in tropical regions, including energy fluxes, soil moisture exchanges, water flows and interactions, are characterised by greater spatial and temporal fluctuations, higher magnitudes, marked by steep spatial gradients and the ability to respond quickly due to human alterations (Wohl et al., 2012).

Given the gaps in knowledge on flood flows and their impact over large data-sparse tropical lowland regions, this research has contributed to providing insights into the nature of tropical flood flows and the underlying flood parameters using Grijalva catchment as a case study (Figures 5.10 and 5.11; Table 4:4; Appendix C Figure C 1). The results of the study show that the rapidly increasing availability of satellite data, computer power, cloud platforms and Geospatial Systems (GIS) enables mapping of highly variable catchment conditions and processes in large tropical lowland catchments. The results show the feasibility of applying large-scale distributed modelling approaches integrated with high-resolution satellite data to represent accurately highly variable catchment conditions.

However, the results have provided simulations of the dynamics of flood flows that are limited to river flooding only without other components that contribute to flooding in the Grijalva catchment. Specific components of the lowland hydrology include groundwater, wetlands and dams that are strongly linked to flooding but remain less explored in this study. Increased efforts are required to quantify human impacts on all dimensions of catchment hydrology and comparative severity of interactions of different reservoirs in the study area. Human activities in the catchment area, dam operations, water abstraction, oil drilling, river flow diversions and river flow regulation, have altered natural river flows in the area. Such activities should be further explored and fully considered in any future distributed hydrological modelling in the study area.

### **6.1.6. Implications of dominant hydrological processes**

The results show that the dominant hydrological processes in the Grijalva catchment area include the highly variable and extreme rainfall, vegetation cover with high Strickler overland flow coefficients and low saturated hydraulic conductivity soils that are conducive to rapid surface runoff and frequent flooding. The catchment has higher Strickler coefficient values, which vary depending on the land cover types. Values were found to be highest in urban areas, bare land and arable land types located in several parts of the catchment area, while low values were found in evergreen and deciduous forest land categories. The largest land use category with high Strickler values were mostly small-scale mixed-farm areas that made up much of the catchment area. The higher Strickler values, which characterise a significant part of the catchment area, increase the speed of surface runoff and increase the amount of discharge.

The result is that surface runoff in mixed agricultural and urban areas is higher, and the more natural environments that are altered and converted into urban and arable land types contribute to even increased surface runoff and increased risk of flooding in the study region. The rapid influx of people into the Tabasco State due to increased oil production would lead to rapid urbanisation and expansion of small-scale mixed farming practises in the surrounding areas to meet increasing food demand. However, agricultural practises in the catchment area are associated with alterations of natural vegetation cover, creating conditions that promote rapid runoff and flooding. Also, the rapid urbanisation of several towns in the catchment including Villahermosa city results in increased depletion of vegetation cover, rapid surface runoff, soil erosion, reduced infiltration capacity of soils, increased river siltation and consequently increased flooding.

Villahermosa city is the hub of business and administration for Mexico's oil industry and is the largest city in the south-eastern region that has recently witnessed an increase in oil production in Mexico and the associated increase in urbanisation. The leading causes of increases in flood frequency and magnitude in the city are related to changes in land use in the sub-catchments located upstream and expansion of the city itself. For example, in October 2007, Tabasco State was hit by an intense flooding event resulting from heavy rainfall that generated intense runoff throughout the Tabasco plains, flooding

about 70% of the state. The capital, Villahermosa city, was severely hit, suffering economic losses of over 3 billion dollars.

Thus, the removal of vegetation for agricultural purposes in the catchment appears to lead to a loss of soils and a reduction in infiltration capacity, causing higher volumes of surface runoff, hillslope erosion and siltation of rivers. Also, the rapid transformation of land cover widely reported in the catchment is leading to significant loss of vegetation cover that, together with an increase in the frequency of extreme weather phenomena, is increasing flood risk throughout the catchment. The results confirm the general hypothesis that the conversion of natural vegetation to cropland, cattle ranching, and settlements has adverse effects on catchment hydrology and flooding. The analyses show that land use has statistically significant effects on runoff and overflow of river channels.

Another dominant hydrological process in the catchment is low hydraulic conductivity values of the soil that also vary depending on the land use type. The conversion of savannah vegetation into mixed farming that is rapidly taking place in Grijalva catchment, particularly in Tabasco State, leads to decreases in infiltration capacity of soils and rapid overland flow during high-intensity rainfall. Rapid urbanisation and expansion of agricultural activities promote increases in runoff and reduces the conductivity of the soil in the catchment. Nevertheless, in undisturbed forested areas of the catchment, the higher hydraulic conductivity values in these areas increase infiltration capacity of soils. Agricultural activities and widespread deforestation in the catchment area result in lower conductivity values in most land use categories, and this decreases the base flow at the catchment outlet, and available underground water storage and so increases the peak flows.

These results validate the initial assumption made in this study, that surface runoff in Grijalva catchment is dominated by surface runoff than subsurface flow. The simulated results, the catchment simulated water balances, and inferred model results show that this assumption is valid. The results show that the dominant hydrological process in the catchment includes high-intensity rainfall-runoff, reduced canopy interception, less infiltration, higher surface runoff and lower groundwater recharge and river flooding.

Therefore, changes in land cover from forest to arable lands generate rapid surface runoff, transporting large quantities of sediments due to increased erosion rates in the catchment, and the siltation of rivers are considered to be the significant processes leading to frequent flooding in the catchment. However, the exposure of the population to river flooding has dramatically increased in recent years due to hurricanes in the coastal areas of the Pacific and Atlantic oceans, which produce unprecedented heavy rainfall and flooding. The combined effect of climate change, alteration of natural forests and rapid urbanisation are the main hydrological processes responsible for increased flood rates in the Grijalva catchment.

## 6.2. Scaling relationships in nested catchments

This section provides a discussion of results on the application of statistical flood scaling relationships that were identified, analysed and developed into equations for estimating flood parameter values (the intercepts ( $\alpha$ ) and the exponents ( $\theta$ )) and magnitude of peak discharge in the La Sierra catchment area. The discussion focuses on the application of scaling relationships, their characteristics, and the development of scaling relationships between peak discharge and nested catchment drainage areas for estimation of flood parameters and peak discharge from each rainfall event. Besides, a discussion is presented on the development of a framework for estimating flood parameters and the magnitude of peak discharge in the La Sierra catchment area.

The following are the interpretations and implications relevant to results presented in chapter 4, section 4.5 linked to the second research objective on investigating the scaling relationships between peak discharge and measurable catchment conditions and processes in nested catchment areas to provide a complementary framework for peak discharge estimation in La Sierra catchment.

### 6.2.1. Expanding the scaling theory of floods in tropical regions

To the knowledge of the author, this is the first study to be carried out to apply the scaling theory of floods in tropical regions and investigate the relationships between rainfall variability and the scaling structure of peak discharge considering all rainfall coverages. The available literature shows little on whether the scaling theory of floods holds in larger tropical watersheds (Wilkinson and Bathurst, 2018; Gupta, 2017). This study applied the scaling theory in a tropical catchment and investigated the physical processes that influence the flood scaling parameters using empirical data sets and considering the differences in rainfall coverages. The study is thus one of the few studies to investigate the physical processes that influence flood scaling parameters in a tropical setting using empirical data sets, most of the related studies have so far been studied in the United States and the United Kingdom (Ayalew et al., 2015; Gupta 2017; Wilkinson and Bathurst, 2018).

Although theoretical studies have been done in establishing relationships between flood parameters, the intercept ( $\alpha$ ) and exponent ( $\theta$ ) and the physical catchment properties



and processes in temperate regions, little progress has been made to test these relationships using observations from other climatic regions, (Ayalew et al., 2014 and Ayalew et al., 2015). Thus, the study has contributed to "bridging the scaling gap" by extending the scaling theory of flood from the smallest nested catchments to large regional and continental river catchments in different climate regions. The results contribute to the universal application of the scaling theory of floods.

Besides results show that the physical processes that govern the generation of floods in La Sierra catchment vary in space. The spatial variability of flood parameters controls the generation of floods in the catchment, a result of differences in catchment properties and processes. Capturing the variability of these conditions and processes in the whole Grijalva catchment area requires the spatial decomposition of Grijalva catchment into nested catchments with natural river network and flows. Also, extending scaling relationships to the larger Grijalva catchment area would be possible, if sufficient hydrometric data records are available for each sub-catchment.

However, several subs-catchments in the Grijalva catchment area are ungauged, and many are sparsely gauged with one streamflow gauge station per every 2 700 km<sup>2</sup>. Also, the river flows in the broader Grijalva catchment area are heavily modified by human activities through damming, river diversions and some major river channels have barriers for regulating river flows. Dams in the study area have a major effect on hydrological flows, and natural flows are decreased, and some are diverted. The presence of water holding bodies in the model domain has a significant effect on regulating natural river flows in the study area.

Conversely, La Sierra sub-catchment has an unregulated river system of mostly natural flows. The sub-catchment has the highest rainfall and flow gauge density in the Grijalva catchment area, as it is the wettest catchment with the highest flood risk in Mexico. As a result, it was found suitable to apply the scaling theory of floods and establish scaling relationships that link the physical processes in the catchment with statistical power-law parameters. The scaling equations were established in the La Sierra catchment, as most of its catchment conditions and river processes were unregulated and had natural river flows. The scaling relationships between peak discharge and drainage areas are suitable in catchments that are more natural and less altered by human activities.

### 6.2.2. Evidence of power-law relationships

An important finding from the application and analysis of the scaling relationships in a tropical setting is that peak discharge exhibits a power-law relationship with drainage areas, similar to findings from humid temperate regions (Fig 5.12, 5.13 and 5.14). Results from the analysis of scaling relationships studied in La Sierra sub-catchment confirmed that the flood parameters (the intercepts ( $\alpha$ ) and the exponents ( $\theta$ )) in the scaling equations identified changes from one rainfall event to another. The results confirm findings from other U.S. and UK studies that found out that peak discharge has a power-law relationship with drainage areas at each rainfall event (Gupta et al., 1996, 2007; Ogden and Dawdy, 2003; Furey and Gupta, 2005; Gupta, 2004; Gupta, 2010; Mantilla et al., 2007, Mantilla et al., 2011; Ayalew et al., 2014a, 2014b, Ayalew et al., 2015).

However, most of the studies were experimental studies and investigated spatial variations in flood response at either plot/hill slope scale or in small headwater catchments (Wilkinson and Bathurst, 2018). The results from this study show that the scaling theory of floods is valid in larger tropical watersheds, comparable to findings in humid temperate climatic regions. The results are consistent with similar studies conducted in actual catchments with empirical data, notably by Ayalew et al. (2015) and Wilkinson et al. (2018), which confirmed the existence of scaling relationships between peak discharge and catchment scaling parameters in the USA and UK, respectively. Therefore, this study confirms the same relationships between the flood parameters, they are valid even in tropical regions and can estimate flood parameter values for discharge estimation in the study area (Tables 5.15 and 5.16).

### 6.2.3. Rainfall catchment coverage and flood parameters

This work contributes to the understanding of variations in the generation of floods in nested catchments as a result of spatial-temporal rainfall variability. The results provide new insights into the power-law relationships between drainage areas and peak discharge, from rainfall events that partially and entirely covers catchments (Figure 5.8). The findings add to the growing evidence that widespread rainfall events covering the whole catchments have higher exponent ( $\theta$ ) values, while localised convective storms (with patchy runoff) have low exponent ( $\theta$ ) values.

The study results show that generally, the exponent ( $\theta$ ) values in La Sierra catchment decrease when rainfall events partially cover catchments (Figure 5.7). Results show that high exponent ( $\theta$ ) values result from catchment-wide runoff that generates increases in downstream flows in the catchment. It has been observed that with increasing scaling intercept ( $\alpha$ ), the scaling exponent ( $\theta$ ) decreases as the coverage of rainfall in catchments decreases (Table 5.5). The results confirm research findings in the mid temperate regions that showed some differences in the exponent ( $\theta$ ) values from rainfall events that entirely and partially cover catchments studied (Figure 5.8) (Gupta and Waymire, 1998; Grimaldi et al., 2012; Ayalew et al., 2014).

However, this analysis considered both storm events, which partially and wholly covered catchments, avoiding limitations in Ayalew et al. (2015) recent large-scale study, which considered only catchment coverage events. When the catchment receives a rainfall-runoff event that covers only part of the basin, results show that the intercept ( $\alpha$ ) increases while exponents ( $\theta$ ) decrease (Table 5.5). It was observed that when scaling intercept ( $\alpha$ ) increases, the scaling exponent ( $\theta$ ) decreases as runoff generation decreases with the decrease of rainfall catchment coverage.

Results show that high exponent ( $\theta$ ) values result from catchment-wide runoff that generates higher rates of downstream increase in runoff from large storms occurring in the catchment. Also, the study shows that exponent ( $\theta$ ) values in La Sierra catchment decrease when rainfall events partially cover the catchments under study. Contrastingly, intercept ( $\alpha$ ) values increase with low rates of downstream runoff. It seems the variation of parameter values is a function of coverage and downstream rates of runoff in

catchments. Thus, the result enhances understanding of the inverse variation of flood parameters due to rainfall coverage, catchment responses and flood generation.

However, a limitation of this study is the lack of long records of hydrometric datasets. This study analysed 2012 to 2015 flood events only, and caution must be exercised in applying results over time. While this may have limited temporal validity, the scaling relationships established for the catchment were spatially valid and significant (Table 5.10). This research provides an understanding of factors that governs the variations of flood parameters and the generation of flooding in tropical catchments. Among other factors, this research shows that the event to event changes of flood parameters are mainly attributed to the spatial variation of rainfall coverage and other catchment conditions and processes.

#### **6.2.4. Flood parameter and peak discharge estimation in La Sierra catchment**

The results obtained contribute to answering one of the twenty-three unsolved questions in hydrology, that is, Question 6: '*What are the hydrologic laws at the catchment scale and how do they change with scale?*' (Blöschl et al., 2019 p.1148). The results contribute to establishing how flood scaling relationships from a rainfall event scale can be related to broader catchment-scale processes. The results show how point scale processes in the smallest nested catchment; for example, observed discharge is related to broader catchment-wide flooding. This result is critical to the estimation of peak discharge and flood monitoring in flood-prone nested catchments.

The study confirms the existence of log-linear statistical relationships between the intercept ( $\alpha$ ) and exponent ( $\theta$ ) values (Equation 5.2) and the log-log power-law relationships between intercept ( $\alpha$ ) parameters and the peak discharge from the smallest headwater nested catchments (Table 5.12) can estimate the magnitude of flood scaling parameters and peak discharge in flood-prone nested catchments (Equation 5.1). The results show all significant scaling relationships in the catchment that can estimate flood parameter and peak discharge values based on the scaling equation across spatial scales in the nested catchments (Table 5.11).

The results demonstrate the potential application of scaling relationships between flood scaling parameters and drainage properties and processes to estimate the magnitude

of peak floods in La Sierra catchment (section 5.3.10). The results are consistent with recent findings that have also shown that peak discharges observed in the smallest headwater catchments can be used to estimate intercept ( $\alpha$ ) parameter values, and the exponent ( $\theta$ ) values can be estimated from the intercept ( $\alpha$ ) parameter values obtained (Ayalew et al., 2015). This finding has important implications for developing scaling relationship equations as complementary tools for refining or correcting the magnitude of peak discharge simulations from numerical models.

However, the results presented in Figure 5.22 show that the estimated peak discharges show significant variability around their expected value, particularly the high magnitude flood events. The results suggest there is greater confidence in scaling model developed in estimating low-level discharge magnitude values relative to its ability to predict large extreme flood events. The model presents significant uncertainties in the estimation of large flood events, although there is high confidence in its ability to estimate small-scale events.

The uncertainties in the scaling relationships established are partly due to lack of long records of hydrometric datasets. The study examined several low to medium level flood events that occurred over a four-year period in which very few major or severe floods occurred. The more frequent low-level floods were more representative of the data analysed than the less frequent extreme and high magnitude flood events that occur less regularly. Long hydrometric dataset records are necessary to reflect accurately the large flood events that occur less frequently. Hydrometric records should be long enough derived from spatially dense networks of streamflow gauges that continuously monitor stream flows in nested basins. While the model might have limited temporal validity, the scaling relationships established are spatially valid and significant (Table 5.10).

### **6.2.5. Implications of scaling relationships**

The process by which floods are generated from small sub-catchments in multi-scale nested catchments building up into large floods is still not quantified in many climatic regions (Wilkinson and Bathurst, 2018). This study has gone some way in quantifying the processes by which floods are generated from small headwater catchments to larger catchments and apply the relationships obtained to estimate flood parameters and peak discharge across the study area (Table 5.15). The scaling relationships established have significant implications in developing complementary tools for operational flood estimation and monitoring across nested catchments.

The findings show the potential for scaling relationships to estimate the severity of peak floods based on the inherent relationship between peak discharge and geophysical catchment properties and processes (section 5.3.4). The relationships obtained provides insights into the broader catchment flood response and flood generation linked to headwater flood peaks (Gupta et al., 2017). An implication of this result is the potential formulation of a flood risk strategy that promotes a network of flow gauging stations in smaller headwater streams for peak discharge monitoring and flood estimation in larger downstream catchment areas. The implication is particularly crucial for the La Sierra basin, which drains floodwaters from smaller mountainous catchments to downstream floodplains in the Las Gaviotas districts, densely populated areas that are frequently flooded in Villahermosa (Arreguín-Cortés et al., 2014). Also, the catchment is the wettest in Mexico with an annual rainfall total of approximately 4 000 mm, which results in frequent flooding in the downstream flood-prone areas.

Although this research has established valid scaling relationships between small headwater catchments and flood peak generation in larger downstream catchment areas, not all flooding processes were included in formulating the scaling relationships established. Although the inclusion of other physical processes had a less significant effect, several other catchment properties and processes that directly or indirectly control flooding were not considered. For example, the findings did not consider channel geometry and roughness due to lack of relevant data, key factors that strongly influence channel flow and peak discharge. Where data are available, further research on these factors is recommended.

Also, the research has assumed that peak discharge generation in the catchment area studied takes place under conditions of excess infiltration when soil moisture conditions are saturated. The study was not designed to assess the alteration of catchment cover by human activities, though their activities have a significant impact on land cover, alteration of the drainage network and river flows. However, these were not considered since scaling laws only consider physical catchment conditions and process.

While some other flood-generating variables were not considered, the findings are statistically relevant and validated by comparing to historical data (Figure 5.22). The result shows that the estimated flood parameters could reduce discrepancies between the observed and simulated peak discharge and can be used to refine or double-check simulated peak discharge magnitudes from numerical models.

#### **6.2.6. Climate change adaptation**

The study has gone some way to show the potential use of statistical models to simulate the magnitude of peak discharges without using hydrological models, and this is essential and relevant under current climate change conditions. The estimation process provides a flood estimation technique that is not affected by climate change and does not require re-calibration against historical observations to forecast peak discharge (Gupta, 2004; Gupta et al., 2007; Gupta et al., 2010; Dawdy et al., 2012).

In changing climatic conditions, the accuracy of historical data will be unreliable, and the calibration of parameters using historical data does not reflect future conditions (Gupta et al., 2017). The self-similarity of river networks and the flood scaling parameters are not expected to change as climate changes. The scaling laws and statistical models developed would remain unchanged for some time (Gupta, 2017). Therefore, there is potential for long-term use of statistical models for flood monitoring and prediction, including as complementary tools that can be used alongside hydrological models in times of climate change. The research provides new research direction to make flood predictions at multiple space and time scales under a changing global hydroclimate. However, there is a need to investigate the role of climate variability and change using the scaling framework for floods.

## **6.3. Uncertainty propagation**

Accurate flood modelling is a challenge due to the dynamic and unpredictable nature of rainfall, runoff processes and uncertainty in flood modelling processes (Volkman et al., 2010; Liu et al., 2012; Tscheikner-Gratl et al., 2018). Despite uncertainties, flood inundation results are usually communicated unreliably using a single deterministic flood extent for discharge and as a basis for significant investments and detailed flood risk management projects (Tscheikner-Gratl et al., 2018). Thus, due to several sources of uncertainty in flood modelling that compromise modelling results, this study identified and quantified individual rainfall and parameter uncertainties to improve the reliability of model results. The study quantified and assessed the effects of propagation and interactions of rainfall and parameter uncertainties on the magnitude of peak discharge and flood inundation extents in the study area.

### **6.3.1. Explicit quantification of uncertainties**

Current uncertainty studies focus mainly on quantifying the overall uncertainty in model outputs without investigating the composition of uncertainty contributions from individual sources (Willems, 2012, Yang et al., 2018; Tscheikner-Gratl et al., 2019). The challenge is to precisely disaggregate uncertainties by sources within the cascading model chain, which is complicated by uncertainty interactions between different sources. There is a need for a clear understanding of the propagation and interaction of individual uncertainties in coupled hydrological and hydraulic models.

The study has demonstrated that the combined model structural/parameter/input uncertainty could be disaggregated into separate uncertainties through individual treatment of each uncertainty by source, while other sources remain fixed. This research has identified rainfall input and parameter uncertainties and quantified their magnitudes at each stage when propagated through cascading models, including their interactions at catchment level. The significant outcome of this analysis is the explicit quantification of uncertainties associated with dynamically interlinked meteorological, hydrological and hydraulic models and the extent of errors and uncertainties, including the effects of flood volumes on flood extents (Tables 5.19 to 5.27).



Although it is challenging to distinguish uncertainties by source due to the underlying number of data categories and elements that need to be analysed (Aronica et al., 2002), an attempt has been made to measure each propagated individual uncertainty by source. Separating uncertainties by source is subject to epistemic and spontaneous errors that can arise from different sources. Though there are interdependences of uncertainties, the analysis could determine different magnitudes, sensitivities and effects of individual uncertainties on model results when perturbations or variations were applied on individual input rainfall ensembles and parameter values. In this study, rainfall and parameter uncertainties were explicitly quantified, and the propagation of each uncertainty and effects of variations on both peak discharge estimates and flood inundation extents was assessed and quantified.

Thus, the study has demonstrated a new uncertainty quantification framework in which uncertainty from rainfall and parameter sources could be disaggregated through individual treatment or perturbation by source, while other sources remain fixed. The research contributes to the formulation of an agreed framework of quantifying propagated and interactions of uncertainties and effects on discharge volumes and flood inundation extents.

### **6.3.2. Uncertainty quantification and distribution**

Results show that variability in rainfall input data contributes to most uncertainties in river discharge estimation, which causes uncertainties in flood extents (Tables 5.19 to 5.27). This finding is consistent with similar research that found that the most substantial uncertainties are in rainfall input data rather than in model structure, parameters and outputs. The results show that uncertainties are propagated from different sources, some are disaggregated, some get concealed in parameter spaces, and some compensate for parameter errors, and the rest are propagated into model results (Komma et al., 2007; Xuan et al., 2009; Nakakita et al., 2016). Other studies report that rainfall uncertainties are amplified when propagated through coupled models resulting in more considerable uncertainties in simulation results (Vila et al., 2009).

However, in this study, rainfall uncertainties were not amplified into larger values when propagated through the coupled models into simulation results. Results show that individual uncertainties propagated through the models decrease, contrary to some research reports showing that uncertainties are amplified as they propagate through the

models. The reduction of uncertainties could be explained by the fact that there are many parameters in the SHETRAN hydrological model with large parameter ranges/spaces that can have different values with negligible effect on the simulated flow (equifinality). It indicates the possible existence of uncertainty compensation in these parameter spaces where propagated uncertainties can accumulate and get concealed without affecting model performance (Anderton et al., 2002).

Results reveal that uncertainty in the input rainfall data when propagated lead to uncertainty in the estimate of river discharge output, which leads to uncertainty in flood levels and extents. However, uncertainties are propagated in different magnitudes across confidence limits and along different stages of the cascading model chain. Generally, uncertainty propagation due to parameter changes occurred in relatively equal magnitudes across 5% and 95% probabilities. Results are inconsistent with other research in that uncertainties were not translated and amplified in model outputs. The catchment flood response shows that higher Strickler coefficient values and low saturated hydraulic conductivity results in widespread flooding. However, high AET/PET results in the most significant decrease in flood volumes and extents in the study area.

### **6.3.3. Implications of uncertainty quantification in flood risk management**

Flood maps are usually represented as a single deterministic flood map shown without considering imperfect forcing uncertainties, varying initial conditions, model parameterisations and numerical limitations (Merwade et al., 2008; Zarzar et al., 2018). Flood inundation simulations based on deterministic model output cannot reflect the full range of flood risk information, as they do not convey the uncertainties associated with the modelling process and data used. Uncertainties should be included in flood risk maps to reflect the full range of flood risk information. Assessing uncertainty, quantification, and visualisation are essential aspects of flood modelling and risk communication. By not adding uncertainties, deterministic flood inundation maps are not conveying the full risk of potential flooding to decision-makers and people at risk. Including uncertainties in the planning and decision-making process and its clear communication is increasingly critical and necessary in flood risk management (Lim and Lee, 2018).

The ensemble approach implemented in this study provides a workable framework for augmenting flood inundation estimates with the underlying uncertainty information. The results have shown a flood simulation approach comprising cascading components of meteorological, hydrologic and inundation models that are all liable to considerable uncertainties in input data, parameters and the model itself. The approach has provided a clear understanding of the propagation and interaction of individual uncertainties in the cascading model chain. The research contributes to the ongoing formulation of a framework for quantifying the propagation and interaction of uncertainties and effects on discharge volumes and flood extents.

The uncertainty estimates are useful in engineering conceptual designs and applications, flood insurance assessments, flood risk management, land use planning and other flood risk measures. Quantifying and defining uncertainties is essential for generating simulated outcomes with quantified flood probability ranges, and this is critical for flood response decision-making in flood risk management (Thielen et al., 2009; Coccia and Todini, 2011; Pappenberger et al., 2012; Weerts et al., 2011; Liu et al., 2012). Given the complex interactions of uncertainties from several sources in the flood inundation modelling process, an integrated modelling system that connects input uncertainties, their interactions with parameter uncertainties and model outputs is vital (Rougier et al., 2010; Merwade et al., 2008).

However, some uncertainty sources were not included as part of this uncertainty propagation study. For example, the role of model uncertainty was not included due to time limitations. The study of uncertainty should be performed systematically to account for all uncertainty. Flood risk assessments should be carried out within the context of maximum uncertainty accounting that considers all sources of uncertainty. All uncertainties should be tracked to give both the relative importance of various uncertainty sources in the system, but also the total uncertainty from the combination of each component in flood simulations. Considering uncertainties in all sources is essential in risk-based decision-making for probabilistic flood forecasting and warnings and operation of flood risk management structures (Pappenberger et al., 2005).

However, considering the full range of sources of uncertainty in flood modelling is a challenge requiring rigorous analysis of all uncertainties, separating each from corresponding sources and avoiding confusion between them (Tscheikner-Gratl et al.,

2018). It is a challenge to separating uncertainties due to the underlying sheer number of data categories and elements that need analysis. Separating uncertainties by source is subject to epistemological and random errors that can occur from various sources. As a result, common operating flood prediction systems and associated research shy away from full uncertainty analysis due to these and other factors, including the high computational demand that the probabilistic approach would entail. However, the computational burden can be reduced by using High-Performance Computing Facilities (HPCF) and increasingly available cloud storages. However, questions remain how to disentangle the combined structural/parameter/input uncertainty, considering complex interactions and effects on model outputs. The model uncertainty is elusive and a challenge to disentangle.

Nevertheless, efforts towards quantification and description of uncertainties are required for generating probabilistic estimates of peak discharge and inundation extents critical for flood risk-based decision-making process (Thielen et al., 2009; Pappenberger et al., 2012; Weerts et al., 2011; Coccia and Todini, 2011; Liu et al., 2012). Knowledge of the spectrum of uncertainties in flood mapping is important for understanding and interpreting information on flood risk. Information on quantified uncertainties can be used to focus on less costly alternatives for a better understanding of other potential choices, allowing a variety of decision-making options and applying resources to sources that contribute most to uncertainty.

Also, knowledge of quantified uncertainties improves flood risk management techniques and develops a robust and clear uncertainty analysis aimed at the most possible outcome of flooding, considering all probable outcomes. The results enhance the quality of flood management strategies and promote an open communication of uncertainties, enabling the most likely outcome to be planned and mitigated (Rodríguez-Rincón et al., 2015). The probabilistic estimation of flood extents represents a step forward towards improvement in flood risk communication to flood-prone communities and households. Furthermore, quantified uncertainties in model results indicate limitations on the effectiveness and reliability of flood simulations when used in decision-making, providing details on constraints and opportunities for decision-making. Quantified uncertainty results can be used to define cost-effective options with a clearer

understanding of other possible outcomes, allowing effective decision-making and allocation of resources to sources that contribute most to uncertainties.

Thus, effective quantification and description of uncertainties are essential for generating simulated results with accurate and actionable guidance on predictive uncertainty enabling flood risk-based decision-making. Consideration of the whole spectrum of uncertainties from all sources is important in risk-based decision-making for issuing flood warnings, evacuation and management of flood risk structures. Besides, adequate definition and reporting of quantified uncertainties are important for the successful implementation of flood risk management strategies and accurate decision-making in the event of a flood emergency response (Montanari, 2011).

## 6.4. Discussion summary

The study responded to the challenge of providing reliable simulations of river flow dynamics, magnitudes and timing in large data-poor tropical regions and contributed to new insights and approaches into the long-standing issue of flood prediction in data-scarce regions. A complementary framework for estimating peak discharge for refining and verifying the magnitude of peak discharge was developed. To enhance the reliability of the results, the study has contributed to the development of a framework for quantifying propagated uncertainties and their effects on discharge volumes and flood extents.

The study shows the feasibility of applying large-scale and fully distributed hydrological modelling approaches integrated with high-resolution satellite data to represent the highly variable catchment conditions and processes in a large lowland tropical regions given the rapidly increasing availability of satellite data, computer power, cloud systems, the advancement of the Geographic Information Systems (GIS) and distributed hydrological models. Model results show that fully distributed hydrological modelling approaches integrated with high-resolution satellite datasets can generate reliable simulations of peak discharge in large catchments that have so far been applied at small scales. The study contributes to meeting the challenge of accurate representation of highly variable catchment conditions and processes and the provision of reliable flood simulations in large data-scarce tropical regions where an increasing number of people are at risk of flooding. Results show that improvements in simulations can be realised through optimally integrating satellite and in-situ datasets with the best combination of parameter, input and output datasets based on error and uncertainty quantification and removal.

The study investigated the application of statistical flood scaling approaches in tropical regions to complement the simulation of the magnitude of peak discharge. In this way, the study is the first to apply the scaling theory of floods in tropical regions to estimate flood parameters and the magnitude of peak discharge. The research results facilitate the universal application of the scaling theory of floods, contributing to “bridging the scaling gap” by extending the scaling theory of floods from experimental hillside plots, small headwater nested catchments to larger continental river basins. It

shows that peak discharge from the smallest headwater catchments can be used to estimate flood scaling parameters and peak discharge for any location in the catchment. The result contributes to quantifying the mechanism by which floods are generated from small headwaters to larger catchments. Also, the results provide a framework for estimating flood parameters and peak discharge values, which are important for monitoring the downstream generation of floods in nested catchments, as well as complementary tools for the verification and refinement of peak discharge simulations generated by the hydrological models.

Due to several sources of uncertainty in flood modelling that compromise modelling results, the study identified and quantified individual rainfall and parameter uncertainties to improve the reliability of model results. To enhance the reliability of flood simulation results, the effects of propagation and interactions of rainfall and parameter uncertainties on the magnitude of peak discharge and flood inundation in the study area were quantified and assessed. The study represents a new uncertainty quantification framework in which uncertainties that arise from input data are propagated and interact within the coupled model chain are quantified, including the effects of uncertainties on flood volumes and flood extents. An interesting outcome of this study is that uncertainties in rainfall input data are reduced when propagated through the hydrological system, a result of uncertainty compensation in parameter spaces where they accumulate and get concealed.

The general conclusion emerging from this modelling study is that it is feasible to implement large scale distributed hydrological modelling approaches integrated with satellite data, based on data merging, error quantification and removal including the incorporation of quantified uncertainties in the final model results. Also, the result contributes to quantifying the process by which floods are generated in catchments, providing a close insight into the headwater flood response and its links with catchment-wide flooding. To enhance the reliability of simulations, the study contributes to the development of a framework for quantifying the propagation and interactions of uncertainties and effects on discharge volumes and flood inundation extents.

## **6.5. Recommendations for future research**

This section provides recommendations for future research that address the limitations and proposes further research in response to the results discussed in this study. The recommendations involve building on from results of this study, addressing research shortcomings/limitations and re-evaluating and extending theories discussed. The recommendations are presented and organised into three subsections drawing on from the discussions made on 1) large-scale distributed hydrological modelling approaches, 2) the application of the scaling theory of floods in nested catchments and 3) uncertainty propagation, interaction and quantification through coupled hydrological and hydraulic models.

### **6.5.1. Large-scale, distributed hydrological modelling**

This research demonstrates the use of high-resolution datasets to represent the highly variable catchment conditions and processes in a large tropical catchment using distributed modelling approaches. It establishes and maps interactions between rainfall variability, the heterogeneity of catchment conditions and flood response at different scales. Recommended are the following three research aspects: (1) the development of automated satellite data processing platforms, (2) coupled meteorological, hydrological and hydraulic modelling and integrated modelling of surface water–groundwater for groundwater simulation.

First, the main obstacles to integrating datasets into large-scale distributed hydrological models are the complex and slow data processing methods for large-scale datasets. Distributed hydrological models require vast amounts of datasets that need processing and integrating to represent varied catchment conditions and processes. The availability of large amounts of data, including extensive spatial data and information on geology, topography, soil, river and hydrological system, vegetation and human activity parameters, is a prerequisite for the full implementation of a distributed modelling approach. There is a need for automated satellite data processing algorithms and related platforms similar to Google Earth Engine that speeds up data processing for large-scale distributed hydrological modelling. For example, the extraction of watershed



boundaries, processing Digital Elevation Model (DEM) data, Flow Direction and Flow Accumulation grids and configuration of the stream network, can be done in a suitable framework and platform for rapid setting up large-scale distributed hydrological models.

Second, further research is needed to understand the interconnected complexity of flood dynamics across various spatial and temporal scales. There is a need to link local flood responses to real-time regional and global meteorological processes and changes using integrated modelling techniques based on coupling meteorological, hydrological and hydraulic models. The coupled meteorological hydrological and hydraulic modelling approaches integrated with groundwater interactions are recommended for realistic flood simulations in the Grijalva catchment area.

Specific components of the Grijalva catchment such as soil moisture and groundwater are linked to flooding but remain less explored in this study. The river flows in the catchment have strong groundwater interactions, particularly when the river leaves the mountainous uplands; the flows take underground routes and then emerge in the broad alluvial plains as surface flow before flowing into the sea. Groundwater flooding is one of the elusive issues in Tabasco lowland areas and requires research. Increased efforts are recommended for integrated modelling approaches, which explores the human influence on catchment hydrology and the comparative extents of the interplay between the various water storages.

### **6.5.2. Scaling relationships**

In several catchments, the process by which floods are generated from small headwater catchments to larger downstream catchments is still not quantified. The challenge is to "bridge the scaling gap" by extending the scaling theory of floods (scaling relationships) from the smallest rivers on the mainland to large regional and continental rivers in different climate regions (Wilkinson and Bathurst, 2018). There is still a shortage of scaling relationship studies in multi-scale catchments in several climatic regions apart from temperate climatic regions. Therefore, there is a need for further studies that apply the scaling theory of floods and establish patterns of scaling relationships across regional and continental river catchments. There is a need to generalise the scaling theory of floods to medium and large river basins spanning different climatic regions.

### **6.5.3. Uncertainty quantification**

Given the complex interactions of uncertainties from several sources in the flood modelling process, an integrated modelling approach linking these sources is needed. Consideration of one or two sources of uncertainty leads to partial treatment of uncertainty and inadequate estimation of uncertainties in model outputs. This study recommends the consideration of all primary sources of uncertainties comprehensively and explicitly in the final model results. There is a need for further research that can fully disentangle, quantify all uncertainties, particularly those from the model structure (Blöschl et al., 2019). Therefore, explicit quantification of uncertainties associated with simulations of hydrological models is required to evaluate and report inherent uncertainties in models in an integrated and objective manner. Further efforts are needed to develop an agreed framework for quantifying the propagation and interaction of uncertainties and effects on discharge volumes and flood extents. Also, research is required to investigate how these uncertainties can be effectively communicated to decision-makers and the public.

## 6.6. Conclusions

This research set out to identify and apply physics-based and data-based hydrological modelling approaches for large-scale flood modelling in lowland tropical regions based on satellite data, error and uncertainty quantification. The study is composed of three distinct research components. The first explored the integration of satellite and in-situ datasets into the SHETRAN distributed hydrological model based on error quantification and correction to simulate the magnitude and timing of river discharge (Objective 1). An investigation of the statistical scaling relationships of rainfall variability and flood peak discharge in nested catchments was done to provide complementary estimates of the magnitude of discharge in La Sierra catchment (Objective 2). The study also examined the propagation and quantification of input data uncertainties and assessed effects on the magnitude of peak discharge and flood inundation extents (Objective 3).

First, the application of distributed hydrological modelling in large data-poor tropical regions was investigated using the SHETRAN hydrological model in the Grijalva River catchment as a case study. Given the rapidly increasing availability of satellite data, computer power, cloud systems and the advancement of the Geographic Information Systems (GIS), the study developed a process-based fully distributed hydrological model integrated with high-resolution satellite and in-situ datasets based on error quantification and correction in a large data-scarce lowland tropical catchment. The study provides new insights into the feasibility of applying large-scale and fully distributed hydrological modelling approaches, integrated with high-resolution satellite datasets to represent accurately the highly variable catchment conditions and processes in large tropical lowland regions, which have so far been studied on a small-scale using distributed models.

Also, the study contributes to providing an understanding of the performance of large-scale distributed models in large domains given the availability of computer resources and increasing quality and quantity of satellite datasets. Results show that distributed hydrological models represented by the SHETRAN hydrological model have satisfactory performance for large-scale modelling that compares well to other studies that use the same model, although applied in small catchments. The study notes that the commonly used lumped modelling approaches cannot fully use alternative forms of datasets (satellite data sets) because they are based on broad areas with average parameter

values across entire drainage areas. The grid nature of satellite data across large areas and the digital acquisition of catchment processes and properties are well linked to distributed hydrological modelling approaches.

The study has shown that large-scale flood modelling in tropical regions can be achieved by optimally integrating several datasets both satellite and in-situ with the best combination of the model parameter, input and output datasets based on error and uncertainty quantification and removal. In this way, the study contributes to the challenge of providing flood simulations of river flow dynamics, magnitudes and timing in large data-poor tropical regions. Also, it contributes to providing new insights into solving the long-standing flood prediction issue in data-scarce regions. The high spatial resolution modelling over large domains approach that is pursued in this study contributes to closing the knowledge gap that exists between simplified, coarse resolution large-scale modelling on the one hand and detailed high-resolution small-scale hydrologic modelling approaches on the other.

Second, empirical approaches to predict flooding were investigated in the La Sierra catchment, an unregulated sub-catchment of the main Grijalva basin and one of the primary sources of flooding in Villahermosa city. This study is the first to test the scaling theory of floods in a tropical catchment and investigate the physical processes that control flood scaling parameters using empirical data sets and considering different rainfall coverages. Most related studies have so far been studied in humid temperate regions, mainly in the United States and the United Kingdom. As a result, this study has contributed to "bridging the scaling gap" by expanding the scaling theory of floods from humid temperate regions to tropical regions, facilitating the potential extension to other climatic regions and ultimately the universal application of the flood scaling theory. To verify its application, the scaling theory of floods needs to be applied to continental and global basins that reflect various hydro-climatic regions.

The study of scaling relationships in tropical regions, using La Sierra catchment as a case study, has shown that peak discharges and nested catchment areas have power-law relationships even in catchments with partial rainfall coverage. The results confirm that flood parameters and peak discharge could be estimated using a log-linear relationship between intercepts ( $\alpha$ ) and exponents ( $\theta$ ) and the log-log relationship between intercepts ( $\alpha$ ) and peak discharge observed from the smallest nested

catchment. These findings substantiate the largely ignored log-linear relationship identified by Ayalew et al., (2015) between the exponent ( $\theta$ ) and the intercept ( $\alpha$ ) of the power-law relationships. Consideration of this log-linear relationship reduces the number of flood scaling parameters that need to be inferred from scaling relationships in nested catchments.

Also, the scaling relationships identified show that peak floods in large tropical catchment areas can reasonably be predicted from discharge observed in small headwater catchments that experience short-term rainfall responses. The results establish how point scale catchment processes in the smallest nested catchment; for example, observed discharge is related to broader catchment-wide flooding. The result contributes to quantifying the process by which floods are generated from small headwater to larger catchments. This provides a framework for estimating flood parameters and peak discharge values and monitoring the downstream generation of flooding in the entire catchment. Thus, the scaling models are essential complementary tools for verifying and refining the magnitude of the peak discharge simulations from hydrological models.

Third, the propagation and interactions of uncertainties through coupled hydrological and hydraulic models and effects on simulations of discharge and flood inundation extents was investigated. The study applied a coupled hydrological and hydraulic modelling approach to flood mapping and quantified how uncertainties that emerged in input data propagate and interact with parameter uncertainties into discharge and flood inundation extent simulations.

The results show a new uncertainty quantification approach in which uncertainty could be disaggregated through individual treatment of uncertainties by source, while other sources remain fixed. It is a framework of uncertainty propagation, interaction and quantification for a large cascading hydrodynamics system that has not been constructed yet. In this way, the research contributes to ongoing research efforts to develop a yet to be agreed framework for quantifying the propagation and interaction of individual uncertainties and effects on discharge volumes and flood inundation extents.

The study contributes to meeting the challenge on how to disentangle the combined structure/parameter/input uncertainty and quantify individual parameter uncertainties,

considering the complex interactions and effects on model results. Though there is still a need to fully disentangle and quantify all individual uncertainties, particularly those from the model structure, this study has demonstrated an approach in quantifying rainfall and parameter uncertainties and their effects on discharge and flood extent simulations.

Also, the study has provided new insights into the propagation of individual uncertainties in input data and the effects of interactions, including the quantified effects on discharge output and flood extents. Gaining insight on individual uncertainty characteristics in terms of source, propagation, interaction, and magnitude is vital for meaningful interpretation of flood simulations and is essential for the generation of simulated outputs with quantified flood probability ranges. The approach is essential for a sound flood risk management strategy that focuses on efforts and resources on specific sources that contribute most to uncertainties in flood risk.

Therefore, the study offers new insights into the feasibility of applying large-scale and fully distributed hydrological modelling approaches integrated with satellite data. The flood scaling theory provides complementary approaches that relate rainfall events to broader catchment scale flooding. The results provide a framework for estimating flood parameters and peak discharge values essential for monitoring the downstream generation of flooding in catchments and as complementary tools for verifying and refining the magnitude of the peak discharge simulations provided by numerical models. The research on uncertainty analysis contributes to ongoing efforts to formulate an agreed framework for quantifying the propagation and interaction of uncertainties and effects on discharge volumes and flood inundation extents simulations. Reliable flood simulations with quantified uncertainties are critical for sound decision-making in emergency flood risk response and the development of flood risk adaptation and mitigation strategies. Such results are essential for flood risk management in tropical regions where millions of people are exposed to flooding.



## 7. REFERENCES

- Aadhar, S. and Mishra, V., (2017). High-resolution near real-time drought monitoring in South Asia. *Scientific data*, 4, p.170145. Available from: <https://doi.org/10.1038/sdata.2017.145>
- Acker, J.G. and Leptoukh, G., (2007). Online analysis enhances use of NASA earth science data. *Eos, Transactions American Geophysical Union*, 88(2), pp.14-17. Available from: <https://doi.org/10.1029/2007EO020003>
- Ackers, J.C., Rickard, C.J. and Gill, D.S., (2010). *Fluvial design guide*. Available from: <http://evidence.environment-agency.gov.uk/FCERM/en/FluvialDesignGuide.aspx> (Accessed: 26 April 2019).
- Abbott, M.B. and Refsgaard, J.C. eds., (2012). *Distributed hydrological modelling*. (Vol. 22). Springer Science & Business Media. Available from: <https://www.springer.com/gp/book/9780792340423> (Accessed: 25 March 2019).
- Abbott, M.B.; Bathurst, J.C.; Cunge, J.A.; O'Connell, P.E.; Rasmussen, J. (1986). "An introduction to the European Hydrological System - Systeme Hydrologique European, "SHE", 1: History and philosophy of a physically-based, distributed modelling system". *Journal of Hydrology*. 87 (1–2): 45–59. [https://doi.org/10.1016/0022-1694\(86\)90114-9](https://doi.org/10.1016/0022-1694(86)90114-9).
- Adams, R., Parkin, G., Elliott, S. and Rutherford, K. (2004). Modelling of hillslope erosion from New Zealand pasture using a rainfall simulator. In *Proceedings of the British Hydrological Society International Conference*, London, UK (pp. 415-420). Available from: <https://doi.org/10.1002/hyp.5670>
- Addinsoft, S.A.R.L. (2019). *XLSTAT software, version (2019).3.2*. Addinsoft, Paris, France. Available from: <https://www.xlstat.com/en/news/xlstat-version-2019-1> (Accessed: 23 March 2019).



- Ahmadisharaf, E., Kalyanapu, A.J. and Bates, P.D., (2018). A probabilistic framework for floodplain mapping using hydrological modelling and unsteady hydraulic modelling. *Hydrological sciences journal*, 63(12), pp.1759-1775. Available from: <https://doi.org/10.1080/02626667.2018.1525615>
- Aparicio, J., Martínez-Austria, P.F., Güitrón, A. and Ramirez, A.I., (2009). Floods in Tabasco, Mexico: a diagnosis and proposal for courses of action. *Journal of Flood Risk Management*, 2(2), pp.132-138. Available from: <https://doi.org/10.1111/j.1753-318X.2009.01026.x>
- Ajami N.K., Gupta H., Wagener T. and Sorooshian S., (2004), Calibration of a semi-distributed hydrologic model for streamflow estimation along a river system, *Journal of Hydrology*, 298(1-4), 112-135. Available from: <https://doi.org/10.1016/j.jhydrol.2004.03.033>
- Ajami, N.K., Duan, Q. and Sorooshian, S., (2007). An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water resources research*, 43(1). Available from: <https://doi.org/10.1029/2005WR004745>
- Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., Wyser, K. and Feyen, L. (2017). Global projections of river flood risk in a warmer world. *Earth's Future*, 5(2), pp.171-182. Available from: <https://doi.org/doi/full/10.1002/2016EF000485>
- Alsdorf, D., Beighley, E., Laraque, A., Lee, H., Tshimanga, R., Loughlin, F. O., Mahe, G., Dinga, B., Moukandi, G. and Spencer, R. G. M. (2016), Opportunities for hydrologic research in the Congo Basin, *Rev. Geophys.*, 54, 378–409, <https://doi.org/10.1002/2016RG000517>.

- Andersen, J., Refsgaard, J.C. and Jensen, K.H., (2001). Distributed hydrological modelling of the Senegal River Basin—model construction and validation. *Journal of Hydrology*, 247(3-4), pp.200-214. Available from: [https://doi.org/10.1016/S0022-1694\(01\)00384-5](https://doi.org/10.1016/S0022-1694(01)00384-5)
- Anderson, M.P. and Woessner, W.W., (1992). The role of the post-audit in model validation. *Advances in Water Resources*, 15(3), pp.167-173. Available from: [https://doi.org/10.1016/0309-1708\(92\)90021-S](https://doi.org/10.1016/0309-1708(92)90021-S)
- Anderton, S.P., Latron, J., White, S.M., Llorens, P., Gallart, F., Salvany, C. and O'Connell, P.E. (2002). Internal evaluation of a physically-based distributed model using data from a Mediterranean mountain catchment. *Hydrology and Earth System Sciences Discussions*, 6(1), pp.67-84. Available from: <https://doi.org/10.5194/hess-6-67-2002>
- Anees, M.T., Abdullah, K., Nordin, M.N.M., Ab Rahman, N.N.N., Syakir, M.I. and Kadir, M.O.A., (2017). One-and Two-Dimensional Hydrological Modelling and Their Uncertainties. *Flood Risk Management*, p.221. Available from: <https://doi.org/10.5772/intechopen.68924>
- Arnell, N.W. and Gosling, S.N., (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, 134(3), pp.387-401. Available from: <https://doi.org/10.1007/s10584-014-1084-5>
- Arnold, J.G., Moriasi, D.N., Gassman, P.W., Abbaspour, K.C., White, M.J., Srinivasan, R., Santhi, C., Harmel, R.D., Van Griensven, A., Van Liew, M.W. and Kannan, N., (2012). SWAT: Model use, calibration, and validation. *Transactions of the ASABE*, 55(4), pp.1491-1508. Available from: <https://elibrary.asabe.org/abstract.asp??JID=3&AID=42256> (Accessed: 20 March 2018).

- Aronica, G., Bates, P.D. and Horritt, M.S., (2002). Assessing the uncertainty in distributed model predictions using observed binary pattern information within GLUE. *Hydrological Processes*, 16(10), pp.2001-2016. Available from: <https://doi.org/10.1002/hyp.398>
- Arreguín-Cortés, F.I., Rubio-Gutiérrez, H., Dominguez-Mora, R. and Luna-Cruz, F.D., (2014). Análisis de las inundaciones en la planicie tabasqueña en el periodo 1995-(2010). *Tecnología y ciencias del agua*, 5(3), pp.05-32. Available from: [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S2007-24222014000300001](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-24222014000300001) (Accessed: 23 March 2018).
- Audefroy, J.F., (2017). Risk and Unsustainability in Villahermosa, Mexico. *WIT Transactions on Ecology and the Environment*, 226, pp.17-27. Available from: <https://doi.org/10.2495/SDP170021>
- Ayalew, T.B., Krajewski, W.F. and Mantilla, R., (2014a). Connecting the power-law scaling structure of peak-discharge to spatially variable rainfall and catchment physical properties. *Advances in Water Resources*, 71, pp.32-43. Available from: <https://doi.org/10.1016/j.advwatres.2014.05.009>
- Ayalew, T.B., Krajewski, W.F., Mantilla, R. and Small, S.J., (2014b). Exploring the effects of hillslope-channel link dynamics and excess rainfall properties on the scaling structure of peak-discharge. *Advances in Water Resources*, 64, pp.9-20. Available from: <https://doi.org/10.1016/j.advwatres.2013.11.010>
- Ayalew, T.B., Krajewski, W.F., Mantilla, R. and Zimmerman, D.L., (2018). Can floods in large river basins be predicted from floods observed in small subbasins?. *Journal of Flood Risk Management*, 11(3), pp.331-338. Available from: <https://doi.org/10.1111/jfr3.12327>

Ayalew, T. B., W. F. Krajewski and R. Mantilla (2015), Insights into expected changes in regulated flood frequencies due to the spatial configuration of flood retention ponds, *Journal of Hydrology*, . Eng., 04015010. Available from: [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001173](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001173)

Ayalew, T.B., Krajewski, W.F., Mantilla, R. and Small, S.J., (2014). Exploring the effects of hillslope-channel link dynamics and excess rainfall properties on the scaling structure of peak-discharge. *Advances in Water Resources*, 64, pp.9-20. Available from: <https://doi.org/10.1016/j.advwatres.2013.11.010>

Ayalew, T.B., Krajewski, W.F. and Mantilla, R., (2015). Analyzing the effects of excess rainfall properties on the scaling structure of peak discharge: Insights from a mesoscale river basin. *Water Res Research*, 51(6), pp.3900-3921. Available from: <https://doi.org/10.1002/2014WR016258>

Banks, J. C.; Camp, J. V.; Abkowitz, M. D. (2013). "Adaptation planning for floods: a review of available tools". *Natural Hazards*. 70 (2): 1327–1337. <https://doi.org/10.1007/s11069-013-0876-7>.

Bates, P.D., Pappenberger, F. and Romanowicz, R.J., (2014). Uncertainty in flood inundation modelling. In *Applied uncertainty analysis for Flood Risk Management* (pp. 232-269). Available from: [https://doi.org/10.1142/9781848162716\\_0010](https://doi.org/10.1142/9781848162716_0010)

Bathurst, J.C., Iroumé, A., Cisneros, F., Fallas, J., Iturraspe, R., Novillo, M.G., Urciuolo, A., de Bièvre, B., Borges, V.G., Coello, C. and Cisneros, P., (2011). Forest impact on floods due to extreme rainfall and snowmelt in 4 Latin American environments 1: Field data analysis. *Journal of Hydrology*, 400(3-4), pp.281-291. Available from: <https://doi.org/10.1016/j.jhydrol.2010.11.044>

Bathurst JC, Ewen J, Parkin G, O'Connell PE, Cooper JD (2004) Validation of catchment models for predicting land-use and climate change impacts. 3. Blind validation for internal and outlet responses. *Journal of Hydrology* 287:74-94  
Available from: <https://doi.org/10.1016/j.jhydrol.2003.09.021>

Beven, K., Leedal, D., McCarthy, S., Lamb, R., Hunter, N., Keef, C., Bates, P., Neal, J. and Wicks, J., (2011). *Framework for assessing uncertainty in fluvial flood risk mapping*. FRMRC Research Rep. SWP1, 7. Available from: [https://www.ciria.org/Resources/Free\\_publications/fluvial\\_flood\\_risk\\_mapping.aspx](https://www.ciria.org/Resources/Free_publications/fluvial_flood_risk_mapping.aspx) (Accessed: 20 March 2019).

Beven, K.J., 2008. *Environmental Modelling: An Uncertain Future?* Routledge, London.  
Available from: <http://www.uncertain-future.org.uk/> (Accessed: 19 March 2019).

Birkinshaw, S.J., Guerreiro, S.B., Nicholson, A., Liang, Q., Quinn, P., Zhang, L., He, B., Yin, J., Fowler, H.J., (2017). Climate change impacts on Yangtze River discharge at the Three Gorges Dam. *Hydrology and Earth System Sciences* 21, 1911–1927. Available from: <https://doi.org/10.5194/hess-21-1911-2017>.

Birkinshaw, S.J., (2012) *Producing a good hydrological simulation*. Available at. <http://research.ncl.ac.uk/shetran/Producing%20a%20good%20hydrological%20simulation.pdf>. (Accessed: 01 June 2019).

Birkinshaw, S.J., Bathurst, J.C., Iroumé, A. and Palacios, H., (2010). The effect of forest cover on peak flow and sediment discharge an integrated field and modelling study in central-southern Chile. *Hydrological Processes*, 25(8), pp.1284-1297. Available from: <https://doi.org/10.1002/hyp.7900>

Birkinshaw, S.J., Bathurst, J.C. and Robinson, M., (2014). 45 years of non-stationary hydrology over a forest plantation growth cycle, Coalburn catchment, Northern England. *Journal of Hydrology*, 519, pp.559-573. Available from: <https://doi.org/10.1016/j.jhydrol.2014.07.050>

- Birkinshaw, S.J., Parkin, G. and Rao, Z., (2008). A hybrid neural networks and numerical models approach for predicting groundwater abstraction impacts. *Journal of Hydroinformatics*, 10(2), pp.127-137. Available from: <https://doi.org/10.2166/hydro.2008.014>.
- Blöschl, G., Bierkens, M.F., Chambel, A., Cudennec, C., Destouni, G., Fiori, A., Kirchner, J.W., McDonnell, J.J., Savenije, H.H., Sivapalan, M. and Stump, C., (2019). Twenty-three unsolved problems in hydrology (UPH)—a community perspective. *Hydrological Sciences Journal*, 64(10), pp.1141-1158. Available from: <https://doi.org/10.1080/02626667.2019.1620507> (Accessed: 29 March 2019).
- Bohr, N., (1909). *Selected Scientific Works*. Vol. I, Papers, 1925, p.149. Available from: <https://iopscience.iop.org/article/10.1070/PU1985v028n10ABEH003949/meta> (Accessed: 09 March 2018).
- Bogardi, J. (2004). *Update.unu.edu*. The newsletter of the United Nations University and its international network of research and training centres/programmes. Issue 32. Available from: <http://update.unu.edu/archive/issue32.htm> (Accessed: 29 March 2018).
- Brewer, M.J., Butler, A. and Cooksley, S.L., (2016). The relative performance of AIC, AICC and BIC in the presence of unobserved heterogeneity. *Methods in Ecology and Evolution*, 7(6), pp.679-692. Available from: <https://doi.org/10.1111/2041-210X.12541>
- Bringi, V. N., L. Tolstoy, M. Thurai, and W. A. Petersen, (2015): Estimation of Spatial Correlation of Drop Size Distribution Parameters and Rain Rate Using NASA's S-Band Polarimetric Radar and 2D Video Disdrometer Network: Two Case Studies from MC3E. *J. Hydrometeor.*, 16, 1207–1221, Available from: <https://doi.org/10.1175/JHM-D-14-0204.1>.

- Brirhet, H. and Benaabidate, L., (2016). Comparison of Two Hydrological Models (Lumped and Distributed) Over A Pilot Area of the Issen Watershed In The Souss Basin, Morocco. *European Scientific Journal*, 12(18). Available from: <http://dx.doi.org/10.19044/esj.2016.v12n18p347>
- Brocca, L., Ciabatta, L., Massari, C., Moramarco, T., Hahn, S., Hasenauer, S., Kidd, R., Dorigo, W., Wagner, W. and Levizzani, V., (2014). Soil as a natural rain gauge: Estimating global rainfall from satellite soil moisture data. *Journal of Geophysical Research: Atmospheres*, 119(9), pp.5128-5141. Available from: <https://doi.org/10.1002/2014JD021489>
- Brown, J.D. and Heuvelink, G.B., (2006). Assessing uncertainty propagation through physically-based models of soil water flow and solute transport. *Encyclopedia of hydrological sciences*. Available from: <https://doi.org/10.1002/0470848944.hsa081>
- Burnash, R.J.C., (1995). The NWS river forecast system-catchment modelling. In: Singh, V.J., (Ed.), *Computer Models of Watershed Hydrology, Water Resources Publication*, Highlands Ranch, Colorado, pp. 311–366. Available from: <https://geoscience.net/research/002/978/002978211.php> (Accessed: 10 January 2018).
- Canli, E., Loigge, B. and Glade, T., (2018). Spatially distributed rainfall information and its potential for regional landslide early warning systems. *Natural hazards*, 91(1), pp.103-127. Available from: <https://doi.org/10.1007/s11069-017-2953-9>
- Centre for Research on the Epidemiology of Disasters (CRED) (2016). *The Human Cost of Weather-Related Disasters 1995-2015*. Institute of Health and Society Université Catholique de Louvain (UCL), Belgium. Available from: [https://reliefweb.int/sites/reliefweb.int/files/resources/COP21\\_WeatherDisasters\\_Report\\_2015\\_FINAL.pdf](https://reliefweb.int/sites/reliefweb.int/files/resources/COP21_WeatherDisasters_Report_2015_FINAL.pdf) (Accessed on 20 July 2018)

- Chappell, A., Renzullo, L.J., Raupach, T.H. and Haylock, M., (2013). Evaluating Geostatistical methods of blending satellite and gauge data to estimate near real-time daily rainfall for Australia. *Journal of hydrology*, 493, pp.105-114. Available from:<https://doi.org/10.1016/j.jhydrol.2013.04.024>
- Chen, L. and Wang, L., (2018). *Recent advances in earth observation big data for hydrology*. *Big Earth Data*, 2(1), pp.86-107. Available from: <https://doi.org/10.1080/20964471.2018.1435072> (Accessed: 09 January 2018).
- Chen, L., Zheng, D., Liu, B., Yang, J. and Jin, Q., (2016). VFDB (2016). *Hierarchical and refined dataset for big data analysis—10 years on. Nucleic acids research*, 44(D1), pp. D694-D697. Available from: <https://doi.org/10.1093/nar/gkv1239>
- Chintalapudi, S., Sharif, H.O. and Furl, C., (2017). High-Resolution, Fully Distributed Hydrologic Event-Based Simulations Over a Large Watershed in Texas. *Arabian Journal for Science and Engineering*, 42(3), pp.1341-1357. Available from: <https://doi.org/10.1007/s13369-017-2446-x>
- China, G.D.P., 2013. “*Those who have knowledge don’t predict. Those who predict don’t have knowledge.*”—Lao Tzu, 600BC. *The Economist*. Available from: [https://www.brainyquote.com/quotes/lao\\_tzu\\_130286](https://www.brainyquote.com/quotes/lao_tzu_130286)
- Chow, Ven Te (2008). *Open-Channel Hydraulics (PDF)*. Caldwell, NJ: The Blackburn Press. ISBN 978-1932846188. Available from: <https://heidarpour.iut.ac.ir/sites/heidarpour.iut.ac.ir/files/u32/open-chow.pdf> (Accessed: 21 January 2019).



- Clark, K.E., Shanley, J.B., Scholl, M.A., Perdrial, N., Perdrial, J.N., Plante, A.F., McDowell W.H. (2017): Tropical river suspended sediment and solute dynamics in storms during an extreme drought. *Water Resources research*. Available from: <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016WR019737>
- Clark, P. and Connolly, E., (2012). Operational use of a grid-based model for flood forecasting. *Proceedings of the Institution of Civil Engineers*, 165(2), p.65. Available from: <https://www.icevirtuallibrary.com/doi/pdf/10.1680/wama.2012.165.2.65> (Accessed: 19 November 2018).
- Cunderlik, J., (2003). Hydrologic model selection for the CFCAS project: assessment of water resources risk and vulnerability to changing climatic conditions. *Department of Civil and Environmental Engineering*, The University of Western Ontario. Available from: <https://ir.lib.uwo.ca/wrrr/9/> (Accessed: 09 November 2018).
- Cloke, H.L. and Pappenberger, F., (2009). Ensemble flood forecasting: A review. *Journal of Hydrology*, 375(3-4), pp.613-626. Available from: <https://doi.org/10.1016/j.jhydrol.2009.06.005>
- Coccia, G. and Todini, E., (2011). Recent developments in predictive uncertainty assessment based on the model conditional processor approach. *Hydrology and Earth System Sciences*, 15(10), pp.3253-3274. Available from: <https://hess.copernicus.org/articles/15/3253/2011/hess-15-3253-2011.html> (Accessed: 29 November 2019).
- Cole, Steve; Robson, Alice; Moore, Bob. (2010). *The Grid-to-Grid Model for nationwide flood forecasting and its use of weather radar*. [Lecture] In Workshop on Polarimetric weather radar for quantitative rainfall estimation, University of Bristol, 28 July 2010. Available from: <https://core.ac.uk/download/pdf/56829.pdf> (Accessed: 19 November 2018).

- Coles, N. A., M. Sivapalan, J. E. Larsen, P. E. Linnet, and C. K. Fahrner (1997), Modelling runoff generation in small agricultural catchments: Can real-world runoff responses be captured? *Hydrol. Processes*, 11, 111– 136. Available from: [https://doi.org/10.1002/\(SICI\)1099-1085\(199702\)11:2<111:AID-HYP434>3.0.CO;2-M](https://doi.org/10.1002/(SICI)1099-1085(199702)11:2<111:AID-HYP434>3.0.CO;2-M) (Accessed: 29 January 2019).
- Coron, L., Thirel, G., Delaigue, O., Perrin, C. and Andréassian, V., (2017). The suite of lumped GR hydrological models in an R package. *Environmental modelling & software*, 94, pp.166-171. Available from: <https://doi.org/10.1016/j.envsoft.2017.05.002>
- Cristiano, E., Veldhuis, M.C.T. and Giesen, N.V.D., (2017). Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas—a review. *Hydrology and Earth System Sciences*, 21(7), pp.3859-3878. Available from: <https://doi.org/10.5194/hess-21-3859-2017>
- Daggupati, P., Yen, H., White, M.J., Srinivasan, R., Arnold, J.G., Keitzer, C.S. and Sowa, S.P., (2015). Impact of model development, calibration and validation decisions on hydrological simulations in West Lake Erie Basin. *Hydrological Processes*, 29(26), pp.5307-5320. Available at: <https://doi.org/10.1002/hyp.10536>
- Dai, Q., Han, D., Rico-Ramirez, M.A. and Islam, T., (2014). Modelling radar-rainfall estimation uncertainties using elliptical and Archimedean copulas with different marginal distributions. *Hydrological sciences journal*, 59(11), pp.1992-2008. Available from: <http://dx.doi.org/10.1080/02626667.2013.865841>
- Dawdy, D.R., Griffis, V.W. and Gupta, V.K., (2012). Regional flood frequency analysis: How we got here and where we are going. *Journal of Hydrologic Engineering*, 17(9), pp.953-959. Available from: [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0000584](https://doi.org/10.1061/(ASCE)HE.1943-5584.0000584)

- De Paiva, R.C.D., Buarque, D.C., Collischonn, W., Bonnet, M.P., Frappart, F., Calmant, S. and Bulhões Mendes, C.A., (2013). Large-scale hydrologic and hydrodynamic modelling of the Amazon River basin. *Water Resources Research*, 49(3), pp.1226-1243. Available from: <https://doi.org/10.1002/wrcr.20067>
- De Figueiredo, E.E. and Bathurst, J.C., (2007). *Runoff and sediment yield predictions in a semiarid region of Brazil using SHETRAN*. In Proceedings of the PUB Kick-off Meeting. Presented at the PUB Kick-off meeting, IAHS, Brasilia, Brazil. Available from: <https://iahs.info/uploads/dms/309030.pdf> (Accessed: 19 November 2018).
- Dembélé, M., Hrachowitz, M., Savenije, H.H., Mariéthoz, G. and Schaefli, B., (2020). Improving the Predictive Skill of a Distributed Hydrological Model by Calibration on Spatial Patterns with Multiple Satellite Data Sets. *Water Resources research*, 56(1), p. e2019WR026085. Available from: <https://doi.org/10.1029/2019WR026085>
- Deng, P.X.; Zhang, M.Y.; Bing, J.P.; Jia, J.W.; Zhang, D.D. Evaluation of the GSMaP\_Gauge products using rain gauge observations and SWAT model in the Upper Hanjiang River Basin. *Atmos. Res.* (2019), 219, 153–165. Available from: <https://doi.org/10.1016/j.atmosres.2018.12.032>
- Diaz-Ramirez JN, Johnson BE, McAnally WH, Martin JL, Alarcon VJ, et al. (2013) Estimation and Propagation of Parameter Uncertainty in Lumped Hydrological Models: A Case Study of HSPF Model Applied to Luxapallila Creek Watershed in Southeast USA. *J Hydrogeol Hydrol Eng* 2:1. doi:10.4172/2325-9647.1000105 (Accessed: 29 November 2018).
- Di Baldassarre, G., Schumann, G., Bates, P.D., Freer, J.E. and Beven, K.J., (2010). Floodplain mapping: a critical discussion of deterministic and probabilistic approaches. *Hydrological sciences journal–Journal des Sciences Hydrologiques*, 55(3), pp.364-376. Available from: <https://doi.org/10.1080/02626661003683389>

- Ding, J., Tarokh, V. and Yang, Y., (2018). *Model selection techniques: An overview*. *IEEE Signal Processing Magazine*, 35(6), pp.16-34. Available from: <https://signalprocessingsociety.org/publications-resources/ieee-signal-processing-magazine/model-selection-techniques-overview> (Accessed: 29 May 2018).
- Dinku, T., Block, P., Sharoff, J., Hailemariam, K., Osgood, D., del Corral, J., Cousin, R. and Thomson, M.C., (2014). Bridging critical gaps in climate services and applications in Africa. *Earth Perspectives*, 1(1), p.15. Available from: <https://doi.org/10.1186/2194-6434-1-15>
- Dukic, V., and Radic, Z. (2014). GIS-Based Estimation of Sediment Discharge and Areas of Soil Erosion and Deposition for the Torrential Lukovska River Catchment in Serbia. *Water Resources Management*, 28(13), 4567-4581. Available from: <https://link.springer.com/article/10.1007/s11269-014-0751-7>
- Dukic V., Radić, Z., (2016). Sensitivity analysis of a physically-based distributed model. *Water Resources Management*, 30(5), pp.1669-1684. Available from: <https://doi.org/10.1007/s11269-016-1243-8>
- Duan, Q., Sorooshian, S. and Gupta, V., (1992). Effective and efficient global optimisation for conceptual rainfall-runoff models. *Water Resources research*, 28(4), pp.1015-1031. Available from: <https://doi.org/10.1029/91WR02985>
- Duan, Q., Sorooshian, S. and Gupta, V.K., (1994). Optimal use of the SCE-UA global optimisation method for calibrating watershed models. *Journal of Hydrology*, 158(3-4), pp.265-284. Available from: [https://doi.org/10.1016/0022-1694\(94\)90057-4](https://doi.org/10.1016/0022-1694(94)90057-4)
- Elliott, A.H., Oehler, F., Schmidt, J. and Ekanayake, J.C., (2012). Sediment modelling with fine temporal and spatial resolution for a hilly catchment. *Hydrological Processes*, 26(24), pp.3645-3660. Available from: <https://doi.org/10.1002/hyp.8445>

- Ewen, J, O'Donnell, G., Burton, A. and O'Connell, P.E. (2007) Errors and uncertainty in physically-based rainfall-runoff modelling of catchment change effects. *Journal of Hydrology*, 330, 641-650. Available from: <https://doi.org/10.1016/j.jhydrol.2006.04.024>
- Ewen, J., O'Connell, E., Bathurst, J., Birkinshaw, S. J., Kilsby, C., Parkin, G., & O'Donnell, G. (2012). Physically-based modelling, uncertainty, and pragmatism. Comment on: Systeme Hydrologique European (SHE): review and perspectives after 30 years development in distributed physically-based hydrological modelling by Jens Christian Refsgaard, Borge Storm and Thomas Clausen. *Hydrology Research*, 43(6), 945-947.doi:10.2166/nh.2012.138. Available from: <https://doi.org/10.2166/nh.2012.001>
- Ewen, J., (1995). Contaminant transport component of the catchment modelling system SHETRAN (pp. 417-441). *Wiley, Chichester, UK*. Available on: [https://research.ncl.ac.uk/shetran/SHETRAN\\_ASCE\\_paper.pdf](https://research.ncl.ac.uk/shetran/SHETRAN_ASCE_paper.pdf). Accessed: 23 May 2019
- Grimes, D.I.F., Pardo-Iguzquiza, E. and Bonifacio, R., (1999). Optimal areal rainfall estimation using rain gauges and satellite data. *Journal of Hydrology*, 222(1-4), pp.93-108. Available from: [https://doi.org/10.1016/S0022-1694\(99\)00092-X](https://doi.org/10.1016/S0022-1694(99)00092-X)
- Famine Early Warning System Network, and the United States. (FEWS NET) (2018). *NET executive overview of food security*. [Washington, D.C.], [FEWS NET]. Available at: <http://fews.net/data> .(Accessed: 28 October 2018).
- FAO Soils Portal (2018): *Harmonized World Soil Database v 1.2 (HWSD) database* FAO, Rome, Italy and IIASA, Laxenburg, Austria. Available at: <http://www.fao.org/soils-portal/soil-survey/soil-maps-and-databases/harmonized-world-soil-database-v12/en/>.(Accessed: 18 December 2018).

- Fatichi, S., Vivoni, E.R., Ogden, F.L., Ivanov, V.Y., Mirus, B., Gochis, D., Downer, C.W., Camporese, M., Davison, J.H., Ebel, B. and Jones, N., (2016). An overview of current applications, challenges and future trends in distributed process-based models in hydrology. *Journal of Hydrology*, 537, pp.45-60. Available from: <https://doi.org/10.1016/j.jhydrol.2016.03.026>
- Furey, P.R. and Gupta, V.K., (2005). Effects of excess rainfall on the temporal variability of observed peak discharge power laws. *Advances in Water Resources*, 28(11), pp.1240-1253. Available from: <https://doi.org/10.1016/j.advwatres.2005.03.014>
- Gama, F., Ordoñez, L., E.M., Villanueva-García, C., Ortiz-Pérez, M.A., López, H.D., Torres, R.C. and Valadez, M.E.M., (2010). Floods in Tabasco Mexico: history and perspectives. Flood recovery, Innovation and Response II. *WIT Transactions on Ecology and the Environment*. Available from: <https://doi.org/10.2495/FRIAR100031>
- García-Pintado, J., Mason, D.C., Dance, S.L., Cloke, H.L., Neal, J.C., Freer, J. and Bates, P.D., 2015. Satellite-supported flood forecasting in river networks: A real case study. *Journal of Hydrology*, 523, pp.706-724. Available from: <https://doi.org/10.1016/j.jhydrol.2015.01.084>
- GoldSim Technology Group (2017), *GoldSim User's Guide (Version 12.1)*. Available from: <https://www.goldsim.com/Web/Customers/Education/Documentation/>, (Assessed: June 2018).
- Goovaerts, P., (1997). *Geostatistics for natural resources evaluation*. Oxford University Press, New York

- Grimaldi, S., Petroselli, A., Tauro, F. and Porfiri, M., (2012). Time of concentration: a paradox in modern hydrology. *Hydrological sciences journal*, 57(2), pp.217-228. Available from: <https://doi.org/10.1080/02626667.2011.644244>
- Grimes, D.I.F., Pardo-Iguzquiza, E. and Bonifacio, R., (1999). Optimal areal rainfall estimation using rain gauges and satellite data. *Journal of Hydrology*, 222(1-4), pp.93-108. Available from: [https://doi.org/10.1016/S0022-1694\(99\)00092-X](https://doi.org/10.1016/S0022-1694(99)00092-X)
- Gupta, V.K. and Waymire, E., (1998). *Spatial variability and scale invariance in hydrologic regionalization*, In: Scale Dependence and Scale Invariance in Hydrology, Cambridge University Press. Available from: <https://doi.org/10.1017/CBO9780511551864>
- Gupta, V.K., Ayalew, T.B., Mantilla, R. and Krajewski, W.F., (2015). Classical and generalised Horton laws for peak flows in rainfall-runoff events. *Chaos: An Interdisciplinary Journal of Non-linear Science*, 25(7), p.075408. Available from: <https://doi.org/10.1063/1.4922177>
- Gupta V.K., Castro S.L. & Over T.M. (1996) On scaling exponents of spatial peak flows from rainfall and river network geometry. *Journal of Hydrology* 1996, 187, (1–2), 81–104. Available from: [https://doi.org/10.1016/S0022-1694\(96\)03088-0](https://doi.org/10.1016/S0022-1694(96)03088-0)
- Gupta, V.K., Troutman, B.M. and Dawdy, D.R., 2007. Towards a non-linear geophysical theory of floods in river networks: an overview of 20 years of progress. *In Non-linear dynamics in geo-sciences* (pp. 121-151). Springer, New York, NY. Available from: [https://doi.org/10.1007/978-0-387-34918-3\\_8](https://doi.org/10.1007/978-0-387-34918-3_8)
- Gupta, V. K. (2004), Emergence of statistical scaling in floods on channel networks from complex runoff dynamics, *Chaos Solitons Fractals*, 19(2), 357–365. Available from: [https://doi.org/10.1016/S0960-0779\(03\)00048-1](https://doi.org/10.1016/S0960-0779(03)00048-1)

Gupta, V. K. and D. R. Dawdy (1995), Physical interpretations of regional variations in the scaling exponents of flood quantiles, *Hydrol. Processes*,9(3–4), 347–361. Available from: <https://doi.org/10.1002/hyp.3360090309>

Gupta, V. K., R. Mantilla, B. M. Troutman, D. Dawdy and W. F. Krajewski (2010), Generalising a non-linear geophysical flood theory to medium-sized river networks, *Geophys. Res. Lett.*,37, L11402. Available from: doi:10.1029/2009GL041540.

Gupta, V.K., (2017). Scaling Theory of Floods for Developing a Physical Basis of Statistical Flood Frequency Relations. In Oxford Research Encyclopaedia- *Flood Hazard Science*. Oxford University Press, Oxford, UK. Available from: DOI: 10.1093/acrefore/9780199389407.013.301

Guha-Sapir, D., Below R., Hoyois Ph, (2009). EM-DAT: *The CRED/OFDA International Disaster Database*. Université Catholique de Louvain, Brussels, Belgium. Available from: <https://www.emdat.be/> (Accessed: 7 December 2017)

Guse, B., Pfannerstill, M., Gafurov, A., Fohrer, N. and Gupta, H., (2016). Demasking the integrated information of discharge: Advancing sensitivity analysis to consider different hydrological components and their rates of change. *Water Resources Research*, 52(11), pp.8724-8743. Available from: <https://doi.org/10.1002/2016WR018894>

Harding, S., Trewin, D., Penny, A., Ziembicki, M., Dale, A., Ustan, T.S., Tan, D., Waltham, N., Laurance, W.F., Ceclia, R. and Yong, W.D., (2017). *State of the Tropics 2017 Report: Sustainable infrastructure in the tropics*. Available from: <https://www.jcu.edu.au/state-of-the-tropics/publications/2017> (Accessed: 7 December 2018)



- Helsel, D.R., Hirsch, R.M., (2002). Statistical methods in water resources, in Hydrologic Analysis and Interpretation. *US Geological Survey*, Reston, United States. Available from: <https://doi.org/10.3133/twri04A3>
- Heinze, G., Wallisch, C. and Dunkler, D., (2018). Variable selection—a review and recommendations for the practising statistician. *Biometrical Journal*, 60(3), pp.431-449. Available from: <https://doi.org/10.1002/bimj.201700067>
- Heistermann, M.; Kneis, (2011) Benchmarking quantitative precipitation estimation by conceptual rainfall-runoff modelling D. *Water Resour. Res.* 2011, 47, W06514. Available from: <https://doi.org/10.1029/2010WR009153>
- Henrik Sen, H.J., Roberts, M.J., van der Keur, P., Harjanne, A., Egilson, D. and Alfonso, L., (2018). Participatory early warning and monitoring systems: A Nordic framework for web-based Flood Risk Management. *International Journal of Disaster Risk Reduction*. Available from: <https://doi.org/10.1016/j.ijdr.2018.01.038>
- Hidayat, H., (2013). *Runoff, discharge and flood occurrence in a poorly gauged tropical basin: The Mahakam River*, PhD Thesis, Wageningen University Kalimantan. Available from: <https://library.wur.nl/WebQuery/wurpubs/443059>
- Hidayat, H., Teuling, A.J., Vermeulen, B., Taufik, M., Kastner, K., Geertsema, T.J., Bol, D.C., Hoekman, D.H., Haryani, G.S., Van Lanen, H.A. and Delinon, R.M., (2017). Hydrology of inland tropical lowlands: The Kapuas and Mahakam wetlands. *Hydrology and Earth System Sciences*, 21(5), p.2579-2594. Available from: <https://doi.org/10.5194/hess-21-2579>.

Horritt, M.S., Mason, D.C., Cobby, D.M., Davenport, I.J. and Bates, P.D., (2003).

Waterline mapping in flooded vegetation from airborne SAR imagery. *Remote Sensing of Environment*, 85(3), pp.271-281. Available from:

[https://doi.org/10.1016/S0034-4257\(03\)00006-3](https://doi.org/10.1016/S0034-4257(03)00006-3)

Horritt, M.S., Di Baldassarre, G., Bates, P.D. and Brath, A., (2007). Comparing the performance of a 2-D finite element and a 2-D finite volume model of floodplain inundation using airborne SAR imagery. *Hydrological Processes: An International Journal*, 21(20), pp.2745-2759. Available from:

<https://doi.org/10.1002/hyp.6486>

<https://doi.org/10.1002/hyp.6486>

Haddad, K., Rahman, A., (2011). Selection of the best fit flood frequency distribution and parameter estimation procedure: a case study for Tasmania in

Australia. *Stoch. Environ. Res. Risk Assess.* 25, 415–42. Available from:

<https://doi.org/10.1007/s00477-010-0412-1>

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D.,

Watanabe, S., Kim, H. and Kanae, S., (2013). Global flood risk under climate change. *Nature Climate Change*, 3(9), p.816. Available from:

<https://doi.org/10.1038/nclimate1911>

Hossain, F., Siddique-E-Akbor, A.H., Mazumder, L.C., ShahNewaz, S.M., Biancamaria, S., Lee, H. and Shum, C.K., (2013). Proof of concept of an altimeter-based river

forecasting system for trans-boundary flow inside Bangladesh. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(2),

pp.587-601. Available from: doi: 10.1109/JSTARS.2013.2283402.

Hossain, F., Katiyar, N., Hong, Y. et al. (2007). The emerging role of satellite rainfall data in improving the hydro-political situation of flood monitoring in the under-

developed regions of the world. *Nat Hazards* 43, 199–210. Available from:

<https://doi.org/10.1007/s11069-006-9094-x>

Hoyt, J.C. and Grover, N.C., (1916). *River Discharge: Prepared for the Use of Engineers and Students*. J. Wiley & Sons; Incorporated. Available from: <https://www.worldcat.org/title/river-discharge/oclc/3654546> (Accessed: 06 November 2019).

Huffman, G.J., R.F. Adler, D.T. Bolvin, E.J. Nelkin (2010). The TRMM Multi-satellite Precipitation Analysis (TMPA). Chapter 1 in *Satellite Rainfall Applications for Surface Hydrology*. Available from: <https://doi.org/10.1007/978-90-481-291>

International Centre for Water Hazard and Risk Management (ICHARM) Report, (2009). Global trends in water-related disasters: an insight for Policymakers. *International Centre for Water Hazard and Risk Management* (UNESCO), Tsukuba, Japan. Available from: <http://www.icharm.pwri.go.jp>. (Accessed on 13 April 2019)

Ibarra-Zavaleta, S.P., Landgrave, R., Romero-López, R., Poulin, A. and Arango-Miranda, R., (2017). Distributed hydrological modelling: Determination of theoretical hydraulic potential & streamflow simulation of extreme hydrometeorological events. *Water*, 9(8), p.602. Available from: <https://doi.org/10.3390/w9080602>

INAFED Gobierno del Estado de Tabasco. (2010) *Mexico Medio Físico*" [Environment]. Enciclopedia de los Municipios de México Tabasco (in Spanish). Available from: <http://inafed.gob.mx/work/enciclopedia/EMM27tabasco/index.html> (Accessed: 16 November 2019).

Instituto Nacional de Estadística, Geografía e Informática (INEGI) (Mexican Institute of Statistics and Geography), *Use of soil and vegetation Series 1 a 5*, (2011). Available from: <http://en.www.inegi.org.mx/temas/usosuelo/> (accessed on 15 November 2018)

Instituto Nacional de Estadística y Geografía (INEGI) (Mexican Institute of Statistics and Geography), (2012). “*Satellite Imagery Maps*” (“*Orthoimages*,”) Available from: <http://en.www.inegi.org.mx/temas/imagenes/espaciomapa/> (Accessed January 2019).

Instituto Nacional de Estadística y Geografía (INEGI) (Mexican Institute of Statistics and Geography), (2018). “*Topography*” (“*Hydrography*,”) Available from: <http://en.www.inegi.org.mx/temas/imagenes/espaciomapa/> (Accessed January 2019).

Intergovernmental Panel on Climate Change (IPCC), (2014): *Climate Change (2014): Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland 151 pp. Available from: <https://www.ipcc.ch/> (Accessed: 26 October 2019).

Jakeman, A.J., Letcher, R.A. and Norton, J.P., (2006). Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software*, 21(5), pp.602-614. Available from: <https://doi.org/10.1016/j.envsoft.2006.01.004>

Jain, S.K., Mani, P., Prakash, P., Singh, V.P., Tullós, D., Kumar, S., Agarwal, S.P. and Dimri, A.P., (2018). A Brief review of flood forecasting techniques and their applications. *International Journal of River Basin Management*, pp.1-16. Available from: <https://doi.org/10.1080/15715124.2017.1411920>

Jain, S.K., Jain, N. and Xu, C.Y., (2017). Hydrologic modelling of a Himalayan mountain basin by using the SWAT model. *Hydrology and Earth System Sciences Discussions*, pp.1-26. Available from: <https://doi.org/10.5194/hess-2017-100>, 2017.

Janes, V., Holman, I., Birkinshaw, S., O'donnell, G. and Kilsby, C., (2018). Improving bank erosion modelling at the catchment scale by incorporating temporal and spatial variability. *Earth Surface Processes and Landforms*, 43(1), pp.124-133. Available from: <https://doi.org/10.1002/esp.4149>

JCGM 200: (2012) International Vocabulary of Metrology, Basic and General Concepts and Associated Terms (VIM) Technical Report, *Joint Committee for Guides in Metrology (JCGM/WG1)* (2012). Available from: [https://www.bipm.org/utils/common/documents/jcgm/JCGM\\_200\\_2012.pdf](https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf) (Accessed: 26 November 2018).

Jenson, S.K. and Domingue, J.O., (1988). Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogrammetric engineering and remote sensing*, 54(11), pp.1593-1600. Available from: <https://pubs.er.usgs.gov/publication/70142175> (Accessed: 06 September 2018).

Jin, X., (2010). Parameter and modelling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. *Journal of Hydrology*, 383 (3–4), 147–155. Available from: <https://doi.org/10.1016/j.jhydrol.2009.12.028>

Jung, Y. and Merwade, V., (2015). Estimation of uncertainty propagation in flood inundation mapping using a 1-D hydraulic model. *Hydrological Processes*, 29(4), pp.624-640. Available from: <https://doi.org/10.1002/hyp.10185>

Jung, Y., Kim, D., Kim, D., Kim, M. and Lee, S., (2014). Simplified flood inundation mapping based on flood elevation-discharge rating curves using satellite images in gauged watersheds. *Water*, 6(5), pp.1280-1299. Liu, L., Hong, Y., Bednarczyk, C.N. Available from: <https://doi.org/10.3390/w6051280>

Johnson K, Ver Hoef JM, Krivoruchko K, Lucas N (2001) *Using ArcGIS geostatistical analyst*. GIS by ESRI, Redlands. Available from: [https://dusk.geo.orst.edu/gis/geostat\\_analyst.pdf](https://dusk.geo.orst.edu/gis/geostat_analyst.pdf) (Accessed: 16 November 2019).

Kahl, B. and Nachtnebel, H.P., (2008). *Online updating procedures for a real-time hydrological forecasting system*. In conference of the Danubian countries (Vol. 24). Available from: <https://iopscience.iop.org/article/10.1088/1755-1307/4/1/012001> (Accessed: 16 November 2019).

Kalyanapu, A.J., Shankar, S., Stephens, A., Judi, D.R., Burian, S.J., (2012). Assessment of PU computational enhancement to a 2D flood model. *Environ. Modell. Softw.* 26 (8), 1009e1016. Available on: <http://dx.doi.org/10.1016/j.envsoft.2011.02.014>

Kabir, Uddin., Deo Raj Gurung, Amarnath Giriraj, Basanta Shrestha, Application of Remote Sensing and GIS for Flood Hazard Management: A Case Study from Sindh Province, Pakistan, (2013) *American Journal of Geographic Information System*, Vol. 2 No. 1, pp. 1-5. Available from: doi: 10.5923/j.ajgis.20130201.01.

Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P. and Thielen, J., (2016). Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environmental Modelling & Software*, 75, pp.68-76. Available from: <https://doi.org/10.1016/j.envsoft.2015.09.009>

Khan, S.I., Adhikari, P., Hong, Y., Vergara, H., Adler, R.F., Policelli, F., Irwin, D., Korme, T. and Okello, L., (2011). Hydroclimatology of Lake Victoria region using hydrologic model and satellite remote sensing data. *Hydrology and Earth System Sciences*, 15(1), pp.107-117. Available from: <https://doi.org/10.5194/hess-15-107-2011>, 2011.

- Khan, S.I., Hong, Y., Vergara, H.J., Gourley, J.J., Adler, R.F. and Policelli, F., (2010). Multispectral and Microwave Satellite Remote Sensing for Flood monitoring in ungauged basins. *AGUFM*, 2010, pp.H23F-1280. Available from: [http://hydro.ou.edu/files/Crest\\_Workshops/refs/66.Sadiq\\_2011IEEE.pdf](http://hydro.ou.edu/files/Crest_Workshops/refs/66.Sadiq_2011IEEE.pdf) (Accessed: 20 November 2017).
- Khatibi, R. and Haywood, J., (2002). The role of flood forecasting and warning in sustainability of flood defence. In Proceedings of the Institution of Civil Engineers-Municipal Engineer (Vol. 151, No. 4, pp. 313-320). London: Published for the *Institution of Civil Engineers* by Thomas Telford Services, c1992. Available from: <https://doi.org/10.1680/muen.2002.151.4.313>
- Kherde, R.V., Sawant, P.H. and Patel, S.V.B., (2013). Integrating geographical information systems (GIS) with hydrological modelling-applicability and limitations. *International Journal of Engineering and Technology*, 5(4), pp.3374-3381. Available from: <http://creativecommons.org/licenses/by/4.0/> (Accessed: 06 November 2019).
- Kirchner, J.W., McDonnell, J.J., Savenije, H.H., Sivapalan, M. and Stump, C., (2019). Twenty-three Unsolved Problems in Hydrology (UPH)—a community perspective. *Hydrological sciences journal*, (just accepted). Available from: <https://doi.org/10.1080/02626667.2019.1620507>
- Komi, K., Neal, J., Trigg, M.A. and Diekkrüger, B., (2017). Modelling of flood hazard extent in data-sparse areas: a case study of the Oti River basin, West Africa. *Journal of Hydrology: Regional Studies*, 10, pp.122-132. Available from: <https://doi.org/10.1016/j.ejrh.2017.03.001>
- Komma, J., Reszler, C., Blöschl, G. and Haiden, T., (2007). Ensemble prediction of floods? Catchment non-linearity and forecast probabilities. *Natural Hazards and Earth System Science*, 7(4), pp.431-444. Available from: <https://doi.org/10.5194/nhess-7-431-2007>,

- Koutsoyiannis, D., Onof, C. and Wheater, H.S., (2003). Multivariate rainfall disaggregation at a fine timescale. *Water Resources Research*, 39(7). . Available from: <https://doi.org/10.1029/2002WR001600>
- Krasilnikov, P., del Carmen Gutiérrez-Castorena, M., Ahrens, R.J., Cruz-Gaistardo, C.O., Sedov, S. and Solleiro-Rebolledo, E., (2013). *The soils of Mexico*. Springer Science & Business Media. Available from: <https://book.cc/book/2203437/0ba496>
- Laerd Statistics (2018) *Friedman Test in SPSS*. [Online]. Available from: <https://statistics.laerd.com/spss-tutorials/friedman-test-using-SPSS-statistics.php> (Accessed: 7 December 2019).
- Lane, D.M., (2015). *Online statistics education: an interactive multimedia course of study (2015)*. Available: [http://onlinestatbook.com/2/calculators/normal\\_dist.html](http://onlinestatbook.com/2/calculators/normal_dist.html) (Accessed: 3 January 2018) (Accessed: 06 November 2019).
- Lee, H., Balin, D., Shrestha, R.R. and Rode, M., (2010). Streamflow prediction with uncertainty analysis, Weida catchment, Germany. *KSCE Journal of Civil Engineering*, 14(3), pp.413-420. Available from: <https://doi.org/10.1007/s12205-010-0413-0>
- Lewis, E., Birkinshaw, S., Kilsby, C. and Fowler, H.J., (2018). Development of a system for automated setup of a physically-based, spatially distributed hydrological model for catchments in Great Britain. *Environmental Modelling & Software*, 108, pp.102-110. Available from: <https://doi.org/10.1016/j.envsoft.2018.07.006>
- Lewis, E.A., (2016). *A robust multi-purpose hydrological model for Great Britain* Doctoral dissertation, Newcastle University. Available from: <http://theses.ncl.ac.uk/jspui/handle/10443/3290> (Accessed: 08 May 2017)



- Lewis, P.A. and Orav, E.J., (1989). *Simulation Methodology for Statisticians*. Operations Analysts, and Engineers, 1. Available from:  
<https://www.taylorfrancis.com/books/9780203710319/chapters/10.1201/9780203710319-4> (Accessed: 06 November 2019).
- Li, B., He, Y. and Ren, L., (2018). Multisource hydrologic modeling uncertainty analysis using the IBUNE framework in a humid catchment. *Stochastic environmental research and risk assessment*, 32(1), pp.37-50. Available from:  
<https://doi.org/10.1007/s00477-017-1424-x>
- Li, L., Xu, C.Y., Xia, J., Engeland K. and Reggiani, P., (2011). Uncertainty estimates by Bayesian method with likelihood of AR (1) plus Normal model and AR (1) plus Multi-Normal model in different time scales hydrological models. *Journal of Hydrology*, 406(1-2), pp.54-65. Available from:  
<https://doi.org/10.1016/j.jhydrol.2011.05.052>
- Li, M., Wang, Q.J., Bennett, J.C. and Robertson, D.E., (2016). Error reduction and representation in stages (ERRIS) in hydrological modelling for ensemble streamflow forecasting. *Hydrology and Earth System Sciences*, 20(9), p.3561.. Available from: <https://doi.org/10.5194/hess-20-3561-2016>
- Lim, J. and Lee, K.S., (2018). Flood mapping using multi-source remotely sensed data and logistic regression in the heterogeneous mountainous regions in North Korea. *Remote Sensing*, 10(7), p.1036. Available from:  
<https://doi.org/10.3390/rs10071036>
- Liu, Y., Weerts, A., Clark, M., Hendricks Franssen, H.J., Kumar, S., Moradkhani, H., Seo, D.J., Schwanenberg, D., Smith, P., Van Dijk, A.I.J.M. and Van Velzen, N., (2012). Advancing data assimilation in operational hydrologic forecasting: progress, challenges and emerging opportunities *Hydrol. Earth Syst. Sci.*, 16, 3863–3887. Available from: <https://doi.org/10.5194/hess-16-3863-2012>

- Liu, Y. and Gupta, H.V., (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water resources research*, 43(7). Available from: <https://doi.org/10.1029/2006WR005756>
- Loveridge, M., Rahman, A., Hill, P. and Babister, M., (2013). Investigation into probabilistic losses for design flood estimation: a case study for the Orara River catchment, NSW. *Australasian Journal of Water Resources*, 17(1), pp.13-24. Available from: DOI: 10.7158/13241583.2013.11465416
- Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., and Loumagne, C. (2014). When does higher spatial resolution rainfall information improve? streamflow simulation? An evaluation using 3620 flood events. *Hydrology and Earth System Sciences*, 18(2), 575–594, Available from: <https://doi.org/10.5194/hess-18-575-2014>
- López-Caloca, A.A., Tapia-Silva, F.O. and Rivera, G., (2017). *Sentinel-1 Satellite Data as a Tool for Monitoring Inundation Areas near Urban Areas in the Mexican Tropical Wet*. In *Water Challenges of an Urbanising World*. IntechOpen. Available from: <https://www.intechopen.com/books/water-challenges-of-an-urbanizing-world/sentinel-1-satellite-data-as-a-tool-for-monitoring-inundation-areas-near-urban-areas-in-the-mexican-> (Accessed: 06 September 2019).
- Lloyd, C.D. and Atkinson, P.M., 2004. Archaeology and geostatistics. *Journal of Archaeological Science*, 31(2), pp.151-165. Available from: <https://doi.org/10.1016/j.jas.2003.07.004>
- Lukey, B.T., Sheffield, J., Bathurst, J.C., Hiley, R.A. and Mathys, N., (2000). Test of the SHETRAN technology for modelling the impact of reforestation on badlands runoff and sediment yield at Draix, France. *Journal of Hydrology*, 235(1-2), pp.44-62. Available from: [https://doi.org/10.1016/S0022-1694\(00\)00260-2](https://doi.org/10.1016/S0022-1694(00)00260-2)

- Ma, J., Sun, W., Yang, G. and Zhang, D., (2018). Hydrological Analysis Using Satellite Remote Sensing Big Data and CREST Model. *IEEE Access*, 6, pp.9006-9016. Available from: DOI: 10.1109/ACCESS.2018.2810252
- Mamnoon, A., & Rahman, A. (2019). Uncertainty analysis in design rainfall estimation due to limited data length: a case study in Qatar. In A. M. Melesse, W. Abtew, & G. Se-nay (Eds.), *Extreme Hydrology and Climate Variability: Monitoring, Modelling, Adaptation and Mitigation* (pp. 37-45). Available from: <https://doi.org/10.1016/978-0-12-815998-9.00004-X>
- Mancini, M., Li, J., Corbari, C., Xin, J., Zhang, J., Zhang, X., Zhang, X. and Su, B., (2013). Satellite data and distributed hydrological model for water resources assessment in the upper Yangtze River basin. In Dragon 2 Final Results and Dragon 3 Kick-Off Symposium. *European Space Agency (ESA)*. Available from: <https://research.utwente.nl/en/publications/satellite-data-a...> (Accessed: 08 May 2019).
- Mandapaka, P.V., Krajewski, W.F., Ciach, G.J., Villarini, G. and Smith, J.A., (2009). Estimation of radar-rainfall error spatial correlation. *Advances in Water Resources*, 32(7), pp.1020-1030. Available from: <https://doi.org/10.1016/j.advwatres.2008.08.014>
- Mandapaka, P.V., Krajewski, W.F., Mantilla, R. and Gupta, V.K., (2009). Dissecting the effect of rainfall variability on the statistical structure of peak flows. *Advances in Water Resources*, 32(10), pp.1508-1525. Available from: <https://doi.org/10.1016/j.advwatres.2009.07.005>

Mantilla, R., (2007). *Physical basis of statistical scaling in peak flows and stream flow hydrographs for topologic and spatially embedded random self-similar channel networks*, Doctoral dissertation, University of Colorado at Boulder.

Available from:

[https://www.iuhr.uiowa.edu/rmantilla/files/2015/03/Thesis\\_Mantilla.pdf](https://www.iuhr.uiowa.edu/rmantilla/files/2015/03/Thesis_Mantilla.pdf)

(Accessed: 06 May 2019).

Mantilla, R., Gupta, V.K. and Troutman, B.M., (2011). Scaling of peak flows with constant flow velocity in random self-similar networks. *Non-linear Processes in Geophysics*, 18(4). Available from: <https://doi.org/10.5194/npg-18-489-2011>

Mercado, V.D., Perez, G.C., Solomatine, D. and Van Lanen, H.A., (2016). Spatio-temporal analysis of hydrological drought at catchment scale using a spatially distributed hydrological model. *Procedia engineering*, 154, pp.738-744. Available from: <https://doi.org/10.1016/j.proeng.2016.07.577>

Mason, D.C., Bates, P.D., Horritt, M.S. and Hunter, N.M. (2006) Improving models of river flood inundation using remote sensing. *Geospatial Today* (May 2006). pp. 40-44. Available from: <http://www.earsel.org/symposia/2006-symposium-Warsaw/pdf/1410.pdf> (Accessed: 16 June 2019).

Mason, D.C., Bates, P.D. and Dall'Amico, J.T., (2009). Calibration of uncertain flood inundation models using remotely sensed water levels. *Journal of Hydrology*, 368(1-4), pp.224-236. Available from: <https://doi.org/10.1016/j.jhydrol.2009.02.034>

Mason, D.C., GUY, J.S. and Bates, P.D., (2011). 11 Data Utilisation in Flood Inundation Modelling. *Flood risk science and management*, p.211. Available from: <https://doi.org/10.1002/9781444324846.ch11> (Accessed: 06 May 2019).

Matott, L.S., (2005). Ostrich: An optimization software tool, documentation and user's guide, Version 1.6. Department of Civil, Structural, and *Environmental Engineering*, State University of New York at Buffalo, Buffalo, NY Available from [www.groundwater.Buffalo.Edu](http://www.groundwater.Buffalo.Edu). Available from: <https://wiki.usask.ca/download/attachments/491552859/OstrichManual.pdf> (Accessed: 26 August 2019).

Matsumoto, M. and Nishimura, T. (1998) Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator. *ACM Transactions on Modelling and Computer Simulation*, 8, 3-30. Available from: <https://doi.org/10.1145/272991.272995>

Mazzoleni, M., Brandimarte, L. and Amaranto, A., (2019). Evaluating precipitation datasets for large-scale distributed hydrological modelling. *Journal of Hydrology*, 578, p.124076. Available from: <https://doi.org/10.1016/j.jhydrol.2019.124076>

McCabe, M.F., Rodell, M., Alsdorf, D.E., Gonzalez Miralles, D., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N., Franz, T.E. and Shi, J., (2017). The future of Earth observation in hydrology. *Hydrology and Earth System Sciences*, 21(7), pp.3879-3914. Available from: <https://doi.org/10.5194/hess-21-3879-2017>

MacLeod, D.A., Cloke, H.L., Pappenberger, F. and Weisheimer, A., (2016). Improved seasonal prediction of the hot summer of (2003) over Europe through better representation of uncertainty in the land surface. *Quarterly Journal of the Royal Meteorological Society*, 142(694), pp.79-90. Available from: <https://doi.org/10.1002/qj.2631>

McMillan, H., Jackson, B., Clark, M., Kavetski, D. and Woods, R., (2011). Rainfall uncertainty in hydrological modelling: An evaluation of multiplicative error models. *Journal of Hydrology*, 400(1-2), pp.83-94. Available from: <https://doi.org/10.1016/j.jhydrol.2011.01.026>

- Menabde, M. and Sivapalan, M., (2001). Linking space-time variability of river runoff and rainfall fields: a dynamic approach. *Advances in Water Resources*, 24(9-10), pp.1001-1014. Available from: [https://doi.org/10.1016/S0309-1708\(01\)00038-0](https://doi.org/10.1016/S0309-1708(01)00038-0).
- Merwade, V., Olivera, F., Arabi, M. and Edleman, S., (2008). Uncertainty in flood inundation mapping: current issues and future directions. *Journal of Hydrologic Engineering*, 13(7), pp.608-620. Available from: [https://doi.org/10.1061/\(ASCE\)1084-0699\(2008\)13:7\(608\)](https://doi.org/10.1061/(ASCE)1084-0699(2008)13:7(608))
- Met Office (2009). *The Hydrological Cycle and Drainage Basins*. [Photograph]. Available at: <http://www.georesource.co.uk/hydrosphere.html#> Assessed on 20 January 2017
- Mishra, A.K. and Coulibaly, P., (2009). Developments in hydrometric network design: A review. *Reviews of Geophysics*, 47(2). Available from: <https://doi.org/10.1029/2007RG000243>
- Montanari A., (2011) Uncertainty of Hydrological Predictions. In Peter Wilderer (ed.) *Treatise on Water Science*, vol. 2, pp. 459–478 Oxford: Academic Press. Available from: <https://doi.org/10.1016/B978-0-444-53199-5.00045-2>
- Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Bingner, R.L., Harmel, R.D. and Veith, T.L., (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3), pp.885-900. Available from: [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Mourato, S., Moreira, M. and Corte-Real, J., (2015). Water Resources impact assessment under climate change scenarios in Mediterranean watersheds. *Water Resources Management*, 29(7), pp.2377-2391. Available from: <https://link.springer.com/article/10.1007%2Fs11269-015-0947-5> (Accessed: 26 August 2018).

Munich Re, (2018). *Natural catastrophe review: Series of hurricanes makes 2017 year of highest insured losses ever*. Munich Re Press Release  
Available from: <https://www.munichre.com/us-non-life/en/company/media-relations/press-releases/2018/2018-01-04-natcat-2018.html> (Accessed: 24 February 2017)

Musálem-Castillejos, K., Laino Guanes, R., Bello-Mendoza, R., González-Espinosa, M., & Ramirez-Marcial, N. (2018). Water quality of the Grijalva river in the Chiapas-Tabasco border. *Ecosistemas y recursos agropecuarios*, 5(13) Available from: [oai: doaj.org/article:f2c86276b67a499fb1d0356f6ddaaaa2](https://oai.doaj.org/article:f2c86276b67a499fb1d0356f6ddaaaa2) (Accessed: 26 June 2018).

Nakakita, Yu, W., E., Kim, S. and Yamaguchi, K., (2016). Impact assessment of uncertainty propagation of ensemble NWP rainfall to flood forecasting with catchment scale. *Advances in Meteorology*, 2016. Available from: <https://doi.org/10.1155/2016/1384302>

Naseela, E.K., Dodamani, B.M. and Chandran, C., (2015). Estimation of runoff using NRCS-CN method and shetran model. *Int. Adv. Res. J. Sci. Eng. Technol*, 2, pp.23-28. Available from: <http://www.iaeme.com/IJCIET/issues.asp?JType=IJCIET&VType=7&IType=2> (Accessed: 16 January 2018).

Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models part I—A discussion of principles. *Journal of hydrology*, 10(3), pp.282-290. Available from: [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)

Néelz, S. and Pender, G., (2013). Benchmarking the latest generation of 2D hydraulic modelling packages. *Environment Agency*, Horizon House, Deanery Road, Bristol, BS1 9AH. Available from: <https://www.gov.uk/government/publications/benchmarking-the-latest-generation-of-2d-hydraulic-flood-modelling-packages> (Accessed: 26 June 2018).

- Nerini, D., Zulkafli, Z., Wang, L.P., Onof, C., Buytaert, W., Lavado-Casimiro, W. and Guyot, J.L., (2015). A comparative analysis of TRMM–rain gauge data merging techniques at the daily time scale for distributed rainfall-runoff modelling applications. *Journal of Hydrometeorology*, 16(5), pp.2153-2168. Available from: <https://doi.org/10.1175/JHM-D-14-0197.1>
- Niemczynowicz, J., (1988). The rainfall movement—A valuable complement to short-term rainfall data. *Journal of Hydrology*, 104(1-4), pp.311-326. Available from: [https://doi.org/10.1016/0022-1694\(88\)90172-2](https://doi.org/10.1016/0022-1694(88)90172-2)
- NIST, (2012). *Handbook 44, Specifications, Tolerances, and Other Technical Requirements for Weighing and Measuring Devices*, National Institute of Standards and Technology OR National Bureau of Standards; U.S. Department of Commerce: Gaithersburg, Available from: <https://www.nist.gov/> (Accessed: 26 October 2018).
- Norouzi Banis, Y., Bathurst, J.C. and Walling, D.E., 2004. Use of caesium-137 data to evaluate SHETRAN simulated long-term erosion patterns in arable lands. *Hydrological processes*, 18(10), pp.1795-1809. Available from: <https://doi.org/10.1002/hyp.1447>
- Notaro, V., Fontanazza, C.M., Freni, G., Puleo, V., (2013). Impact of rainfall data resolution in time and space on the urban flooding evaluation. *WaterSci. Technol.*68(9),1984–1993. Available from: <https://doi.org/10.6028/NIST.HB.44-2016>
- Ogden F.L. & Dawdy D.R. (2003). Peak discharge scaling in small Hortonian watershed. *J Hydrol Eng* 8, (2), 64–73. Available from: [https://doi.org/10.1061/\(ASCE\)1084-0699\(2003\)8:2\(64\)](https://doi.org/10.1061/(ASCE)1084-0699(2003)8:2(64)).



- Oliver, M.A. and Webster, R., 2014. A tutorial guide to geostatistics: Computing and modelling variograms and kriging. *Catena*, 113, pp.56-69. Available from: <https://doi.org/10.1016/j.catena.2013.09.006>
- Op de Hipt, F., Diekkrüger, B., Steup, G., Yira, Y., Hoffmann, T. and Rode, M., (2017). Applying SHETRAN in a tropical West African catchment (Dano, Burkina Faso), Calibration, validation, uncertainty assessment. *Water*, 9(2), p.101. Available from: <https://doi.org/10.3390/w9020101>
- Oubennaceur, K., Chokmani, K., Nastev, M., Tanguy, M. and Raymond, S., (2018). Uncertainty analysis of a two-dimensional hydraulic model. *Water*, 10(3), p.272. Available from: <https://doi.org/10.3390/w10030272>
- Pappenberger, F., Beven, K.J., Hunter, N.M., Bates, P.D., Gouweleeuw, B.T., Thielen, J. and De Roo, A.P.J., (2005). Cascading model uncertainty from medium-range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). *Hydrology and Earth System Sciences Discussions*, 9(4), pp.381-393. Available from: <https://doi.org/10.5194/hess-9-381-2005>
- Pappenberger, F., Dutra, E., Wetterhall, F., and Cloke, H. L. (2012). Deriving global flood hazard maps of fluvial floods through a physical model cascade, *Hydrol. Earth Syst. Sci.*, 16, 4143–4156, Available from: <https://doi.org/10.5194/hess-16-4143-2012>,
- Pappenberger, F., Bartholmes, J., Thielen, J., Cloke, H.L., Buizza, R. and de Roo, A., (2008). New dimensions in early flood warning across the globe using grand-ensemble weather predictions. *Geophysical Research Letters*, 35(10). Available from: <https://doi.org/10.1029/2008GL033837>

- Paiva, R.C., Collischonn, W. and Buarque, D.C., (2013). Validation of a full hydrodynamic model for large-scale hydrologic modelling in the Amazon. *Hydrological Processes*, 27(3), pp.333-346. Available from: <https://doi.org/10.1002/hyp.8425>
- Pechlivanidis, I.G., Jackson, B.M., McIntyre, N.R. and Wheeler, H.S., (2011). Catchment scale hydrological modelling: a review of model types, calibration approaches and uncertainty analysis methods in the context of recent developments in technology and applications. *Global NEST Journal*, 13(3), pp.193-214. Available from: <https://doi.org/10.30955/gnj.000778>
- Pedersen, L., Jensen, N.E., Christensen, L.E. and Madsen, H., (2010). Quantification of the spatial variability of rainfall based on a dense network of rain gauges. *Atmospheric research*, 95(4), pp.441-454. Available from: <https://doi.org/10.1016/j.atmosres.2009.11.007>
- Petropoulos, G.P., Ireland G. and Barrett, B., (2015). Surface soil moisture retrievals from remote sensing: Current status, products & future trends. *Physics and Chemistry of the Earth, Parts A/B/C*, 83, pp.36-56. Available from: <https://doi.org/10.1016/j.pce.2015.02.009>
- Povey, A.C. and Grainger, R.G., (2015). Known and unknown unknowns: uncertainty estimation in satellite remote sensing. *Atmospheric Measurement Techniques*, 8(11), pp.4699-4718. Available from: <https://doi.org/10.5194/amt-8-4699-2015>
- Priego-Hernández, G., Arias, H.R. and Trejo, F.R., (2018). Hydrodynamics and propagation of uncertainty in chemical fate and bioaccumulation models. *Environmental Toxicology and Chemistry* 21(4), 700-709. Available from: <https://pubmed.ncbi.nlm.nih.gov/11951941/> (Accessed: 26 July 2018).

Rabus, B., Eineder, M., Roth, A. and Bamler, R., (2003). The shuttle radar topography Mission a new class of digital elevation models acquired by spaceborne radar. *ISPRS journal of photogrammetry and remote sensing*, 57(4), pp.241-262. Available from: [https://doi.org/10.1016/S0924-2716\(02\)00124-7](https://doi.org/10.1016/S0924-2716(02)00124-7)

Ramos, J., Marrufo, L. and Gonzalez, F.J., (2009). Use of lidar data in floodplain risk management planning: the experience of Tabasco 2007 flood. *In Advances in Geoscience and Remote Sensing*. IntechOpen. Available from: <https://www.intechopen.com/books/advances-in-geoscience-and-remote-sensing> (Accessed: 26 July 2017).

Refsgaard, Jens Christian, Karsten Arnbjerg-Nielsen, Martin Drews, Kirsten Halsnæs, Erik Jeppesen, Henrik Madsen, Anil Markandya, Jørgen E. Olesen, John R. Porter and Jens H. Christensen. (2013) "The role of uncertainty in climate change adaptation strategies, A Danish water management example." *Mitigation and Adaptation Strategies for Global Change* 18, no. 3: 337-359. Available from: <https://doi.org/10.1007/s11027-012-9366-6>

Refsgaard, J.C., (1997). Parameterisation, calibration and validation of distributed hydrological models. *Journal of hydrology*, 198(1-4), pp.69-97. Available from: [https://doi.org/10.1016/S0022-1694\(96\)03329-X](https://doi.org/10.1016/S0022-1694(96)03329-X)

Revilla-Romero, B., Hirpa, F.A., Pozo, J.T.D., Salamon, P., Brakenridge, R., Pappenberger, F. and De Groeve, T., (2015). On the use of global flood forecasts and satellite-derived inundation maps for flood monitoring in data-sparse regions. *Remote Sensing*, 7(11), pp.15702-15728. Available from: <https://doi.org/10.3390/rs71115702>

- Ritter, A. and Muñoz-Carpena, R., (2013). Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments. *Journal of Hydrology*, 480, pp.33-45. Available from: <https://doi.org/10.1016/j.jhydrol.2012.12.004>Get rights and content
- Robinson J.S. and Sivapalan, M. (1997). An investigation into the physical causes of scaling and heterogeneity of regional flood frequency. *Water Resour Res*, 33, (5), 1045–1059. Available from: <https://doi.org/10.1029/97WR00044>
- Rodríguez-Rincón, J.P., Pedrozo-Acuña, A. and Breña-Naranjo, J.A., (2015). Propagation of hydrometeorological uncertainty in a model cascade framework to inundation prediction. *Hydrology and Earth System Sciences*, 19(7), p.2981. Available from: <https://doi.org/10.5194/hess-19-2981-2015>
- Roelofs, F. ed., (2018). Thermal-Hydraulics Aspects of Liquid Metal Cooled Nuclear Reactors. *Woodhead Publishing*. Available from: <https://doi.org/10.1016/B978-0-08-101980-1.00002-8>
- Rougier, J., Sparks, S., Aspinall, W., Cornell, S., Crossweller, S., Edwards, T., Freer, J., Hill, L. and Hincks, T., (2010). *SAPPUR: NERC Scoping study on uncertainty and risk in natural hazards*. BRISK, University of Bristol. Available from: <https://research-information.bris.ac.uk/en/publications/sappur-merc-scoping-study-on-uncertainty-and-risk-in-natural-haza> Accessed: 13 July 2019).
- Rubinstein, R. Y. and Kroese, D. P. (2007) Simulation and the Monte Carlo method (2nd ed.). New York: John Wiley & Sons. Available from: <https://doi.org/10.1002/9780470230381>

- Ryu D., (2014), *Integrating Ground and Remotely Sensed Observations into Operational Flood Forecasting System -Overview. Presentation under the theme Towards Improved Flood Forecasting Using Observations from Ground and Space*, Melbourne Australia. Available from: <https://www.mendeley.com/viewer/?fileId=51ef2625-c38e-6283-4f24-072b6ac9d4a1&documentId=7a114c84-66d1-3b57-93e9-> (Accessed: 06 July 2019).
- Sampson, C.C., Smith, A.M., Bates, P.D., Neal, J.C., Alfieri, L. and Freer, J.E., (2015). A high-resolution global flood hazard model. *Water Resources research*, 51 9, pp.7358-7381. Available from: <https://doi.org/10.1002/2015WR016954>
- Sandholt, I. andersen, J., Dybkjær, G., Lo, M., Rasmussen, K., Refsgaard, J.C. and Høgh-Jensen, K., (1999). Use of remote sensing data in distributed hydrological models: applications in the Senegal River basin. *Geografisk Tidsskrift-Danish Journal of Geography*, 99(1), pp.47-57. Available from: <https://doi.org/10.1016/j.jhydrol.2008.03.006>
- Schmalz, B., Tavares, F. and Fohrer, N., (2008). Modelling hydrological processes in mesoscale lowland river basins with SWAT—capabilities and challenges. *Hydrological sciences journal*, 53(5), pp.989-1000. Available from: <https://doi.org/10.1623/hysj.53.5.989>
- Schittkowski, K., (2002): EASY-FIT A software system for data fitting in dynamic systems. *Struct. Multidisc Optim.* 23, pp. 152-169. Springer-Verlag. Available from: <https://doi.org/10.1007/s00158-002-0174-6>
- Schumann, G. J.-P., J. C. Neal, N. Voisin, K. M. Andreadis, F. Pappenberger, N. Phanthuwongpakdee, A. C. Hall and P. D. Bates (2013), A first large-scale flood inundation forecasting model, *Water Resour. Res.*, 49, 6248–6257, doi:10.1002/wrcr.20521. Available from: <https://doi.org/10.1002/wrcr.20521>

- Schumann, G., Bates, P.D., Horritt, M.S., Matgen, P. and Pappenberger, F., (2009). Progress in integration of remote sensing–derived flood extent and stage data and hydraulic models. *Reviews of Geophysics*, 47(4). Available from: <https://doi.org/10.1029/2008RG000274>
- Schumann, G., Matgen, P., Hoffmann, L., Hostache, R., Pappenberger, F. and Pfister, L., (2007). Deriving distributed roughness values from satellite radar data for flood inundation modelling. *Journal of Hydrology*, 344(1-2), pp.96-111. Available from: <https://doi.org/10.1016/j.jhydrol.2007.06.024>
- Sharad, Kumar., Jain., Pankaj Mani, Sanjay K. Jain, Pavithra Prakash, Vijay P. Singh, Desiree Tullos, Sanjay Kumar, S. P. Agarwal & A. P. Dimri (2018). A Brief review of flood forecasting techniques and their applications, *International Journal of River Basin Management*, Available from: <https://doi.org/10.1080/15715124.2017.1411920>
- Shmueli, G., (2010). *To explain or to predict?*. *Statistical science*, 25(3), pp.289-310. Available from: <https://doi.org/10.1214/10-STS330>
- Shrestha, M., Acharya, S.C. and Shrestha, P.K., (2017). Bias correction of climate models for hydrological modelling—are simple methods still useful?. *Meteorological Applications*, 24(3), pp.531-539. Available from: <https://doi.org/10.1002/met.1655>
- Shrestha, P. (2017). *SWAT Calibration Helper v1.0*, HydroShare, Available from: <http://www.hydroshare.org/resource/8bb80a0aece94f81b52dcb719ada5a8f> (Accessed: 06 June 2018).

Shrestha, S.G., (2013). *Distributed Groundwater Modeling for Analysis of Groundwater Subsidence and its Implications* Institute of engineering (Doctoral dissertation), Pulchowk Campus Tribhuvan university, Lalitpur, Nepal. Available at [https://www.researchgate.net/profile/Pallav\\_Shrestha/publication/301204870\\_Distributed\\_groundwater\\_modeling\\_for\\_analysis\\_of\\_groundwater\\_subsidence\\_and\\_its\\_implications/links/570c98e308aee0660351ba52.pdf](https://www.researchgate.net/profile/Pallav_Shrestha/publication/301204870_Distributed_groundwater_modeling_for_analysis_of_groundwater_subsidence_and_its_implications/links/570c98e308aee0660351ba52.pdf) (Accessed 17 October 2018)

Shrestha, S., Bastola, S., Babel, M.S., Dulal, K.N., Magome, J., Hapuarachchi, H.A.P., Kazama, F., Ishidaira, H. and Takeuchi, K., (2007). The assessment of spatial and temporal transferability of a physically-based distributed hydrological model parameters in different physiographic regions of Nepal. *Journal of Hydrology*, 347(1-2), pp.153-172. Available from: <https://doi.org/10.1016/j.jhydrol.2007.09.016>

Sideris, I.V., Gabella, M., Erdin, R. and Germann, U., (2014). Real-time radar rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland. *Quarterly Journal of the Royal Meteorological Society*, 140(680), pp.1097-1111. Available from: <https://doi.org/10.1002/qj.2188>

Sivapalan, M., (2003). Prediction in ungauged basins: a grand challenge for theoretical hydrology. *Hydrological Processes*, 17(15), pp.3163-3170. Available from: <https://doi.org/10.1002/hyp.5155>

Smith, M. B., V. I. Koren, Z. Zhang, S. M. Reed, J.-J. Pan and F. Moreda, (2004): Runoff response to spatial variability in precipitation: An analysis of observed data. *Journal of Hydrology*,., 298, 267–286, doi: 10.1016/j.jhydrol.2004.03.039. Available from: <https://doi.org/10.1016/j.jhydrol.2004.03.039>

- Smith, M.B., Koren, V., Reed, S., Zhang, Z., Zhang, Y., Moreda, F., Cui, Z., Mizukami, N. Anderson, E.A. and Cosgrove, B.A., (2012). The distributed model intercomparison project–phase 2: Motivation and design of the Oklahoma experiments. *Journal of Hydrology*, 418, pp.3-16. Available from: <https://doi.org/10.1016/j.jhydrol.2011.08.055>
- Smith, M.B., Koren, V., Zhang, Z., Zhang, Y., Reed, S.M., Cui, Z., Moreda, F., Cosgrove, B.A., Mizukami, N. and Anderson, E.A., (2012). Results of the DMIP 2 Oklahoma experiments. *Journal of Hydrology*, 418, pp.17-48. Available from: <https://doi.org/10.1016/j.jhydrol.2011.08.056>
- Singh, V. P. (2018). Hydrologic modeling: Progress and future directions. *Geoscience Letters*, 5(1), 1–18. Available from: <https://doi.org/10.1186/s40562-018-0113-z>
- Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P. and Macomber, S.A., (2001). Classification and change detection using Landsat TM data: when and how to correct atmospheric effect. *Remote Sensing of Environment*, 75, pp. 230–244. Available from: [https://doi.org/10.1016/S0034-4257\(00\)00169-3](https://doi.org/10.1016/S0034-4257(00)00169-3)Get rights and content
- Spies, R.R., Franz, K.J., Hogue, T.S. and Bowman, A.L., (2015). Distributed hydrologic modelling using satellite-derived potential evapotranspiration. *Journal of Hydrometeorology*, 16(1), pp.129-146. Available from: <https://doi.org/10.1175/JHM-D-14-0047.1>
- Sreedevi, S. and Eldho, T.I., (2019). A two-stage sensitivity analysis for parameter identification and calibration of a physically-based distributed model in a river basin. *Hydrological Sciences Journal*, 64(6), pp.701-719. Available from: <https://doi.org/10.1080/02626667.2019.1602730>



- Statistics, L., (2018). *ANOVA with repeated measures using SPSS statistics*. IBM® SPSS® Statistics guides. Available at: <https://statistics.laerd.com/spss-tutorials/one-way-anova-repeated-measures-using-spss-statistics.php> [Accessed 31 Mar. 2019].
- Stephens, E., Schumann, G. and Bates, P., (2014). Problems with binary pattern measures for flood model evaluation. *Hydrological Processes*, 28(18), pp.4928-4937. Available from: <https://doi.org/10.1002/hyp.9979>
- Stephens, E.M., Bates, P.D., Freer, J. and Mason, D., (2012). Calibration of flood inundation models using uncertain satellite observed water levels. *J Hydrol.* doi, 10. Available from: <https://doi.org/10.1016/j.jhydrol.2009.02.034>
- Tanaka, T., Yoshioka, H., Siev, S., Fujii, H., Fujihara, Y., Hoshikawa, K., Ly, S. and Yoshimura, C., (2018). An Integrated hydrological-hydraulic model for simulating surface water flows of a shallow lake surrounded by large floodplains. *Water*, 10(9), p.1213. Available from: <https://doi.org/10.3390/w10091213>
- Tanoue, M., Hirabayashi, Y. and Ikeuchi, H., (2016). Global-scale river flood vulnerability in the last 50 years. *Scientific Reports*, 6, p.36021. Available from: <https://doi.org/10.1038/srep36021>
- Tegegne, G., Park, D.K. and Kim, Y.O., (2017). Comparison of hydrological models for the assessment of Water Resources in a data-scarce region, the Upper Blue Nile River Basin. *Journal of Hydrology: Regional Studies*, 14, pp.49-66. Available from: <https://doi.org/10.1016/j.ejrh.2017.10.002>
- Teng, J., Jakeman, A.J., Vaze, J., Croke, B.F., Dutta, D. and Kim, S., (2017). Flood inundation modelling: A review of methods, recent a study are advances and uncertainty analysis. *Environmental Modelling & Software*, 90, pp.201-216. Available from: <https://doi.org/10.1016/j.envsoft.2017.01.006>

The 2014 Pacific hurricane season' (2014). Available from:

[https://en.wikipedia.org/wiki/2014\\_Pacific\\_hurricane\\_season](https://en.wikipedia.org/wiki/2014_Pacific_hurricane_season) (Accessed: 06 June 2020).

Thielen, J., Bartholmes, J., Ramos, M.H. and Roo, A.D., (2009). The European flood alert system–part 1: concept and development. *Hydrology and Earth System Sciences*, 13(2), pp.125-140. Available from:

<https://doi.org/10.5194/hess-13-125-2009>

Thiemig, V., de Roo, A. and Gadain, H., (2011). Current status on flood forecasting and early warning in Africa. *Intl. J. River Basin Management*, 9(1), pp.63-78. Available from: <https://doi.org/10.1080/15715124.2011.555082>

Thiemig, V., Rojas, R., Zambrano-Bigiarini, M. and De Roo, A., (2013).

Hydrological evaluation of satellite-based rainfall estimates over the Volta and Baro-Akobo Basin. *Journal of Hydrology*, 499, pp.324-338. Available from: <https://doi.org/10.1016/j.jhydrol.2013.07.012>

Tolson B. and Shoemaker C. (2007). Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources research*, 43. Available from: <https://doi.org/10.1029/2005WR004723>.

Tripkovic, V., (2014). *Quantifying and up-scaling surface and subsurface runoff and nutrient flows under climate variability* (Doctoral dissertation, Newcastle University). Available from: <http://theses.ncl.ac.uk/jspui/handle/10443/2380> (Accessed: 09 May 2019).

Tscheikner-Gratl, F., Bellos, V., Schellart, A., Moreno-Rodenas, A., Muthusamy, M., Langeveld, J., Clemens, F., Benedetti, L., Rico-Ramirez, M.A., de Carvalho, R.F. and Breuer, L., (2018). Recent insights on uncertainties present in integrated catchment water quality modelling. *Water research*. Available from: <https://doi.org/10.1016/j.watres.2018.11.079>

- Tscheikner-Gratl, Moreno-Rodenas, A.M., F., Langeveld, J.G. and Clemens, F.H., (2019). Uncertainty analysis in a large-scale water quality integrated catchment modelling study. *Water research*, 158, pp.46-60. Available from: <https://doi.org/10.1016/j.watres.2019.04.016>
- VanShaar, J.R., Haddeland, I. and Lettenmaier, D.P., (2002). Effects of land-cover changes on the hydrological response of interior Columbia River basin forested catchments. *Hydrological Processes*, 16(13), pp.2499-2520. Available from: <https://doi.org/10.1002/hyp.1017>
- Van Wesemael, A., Gobeyn, S., Neal, J., Lievens, H., Van Eerdenbrugh, K., De Vleeschouwer, N., Schumann, G., Vernieuwe, H., Di Baldassarre, G., De Baets, B. and Bates, P., 2016, April. Calibration of a flood inundation model using a SAR image: Influence of acquisition time. In *EGU General Assembly Conference Abstracts* (Vol. 18). Available from: <https://ui.adsabs.harvard.edu/abs/2016EGUGA..18.8704V/abstract> (Accessed: 23 April 2019).
- Van Wesemael, A., Gobeyn, S., Neal, J., Lievens, H., Van Eerdenbrugh, K., De Vaze, J., Jordan, P., Beecham, R., Frost, A. and Summerell, G., (2011). *Guidelines for rainfall-runoff modelling: Towards best practice model application*. Bruce, Australia. Available from: [https://ewater.org.au/uploads/files/eWater-Guidelines-RRM-\(v1\\_0-Interim-Dec-2011\).pdf](https://ewater.org.au/uploads/files/eWater-Guidelines-RRM-(v1_0-Interim-Dec-2011).pdf) (Accessed: 23 November 2018).
- Verdin, A., Funk, C., Rajagopalan, B. and Kleiber, W., (2016). Kriging and local polynomial methods for blending satellite-derived and gauge precipitation estimates to support hydrologic early warning systems. *IEEE Transactions on Geoscience and Remote Sensing*, 54(5), pp.2552-2562. Available from: <https://doi.org/10.1109/TGRS.2015.2502956>

- Verhoeven, J. T. A. and Setter, T. L. (2010). Agricultural use of wetlands: opportunities and limitations, *Ann. Bot.*, 105, 155–163, Available from: <https://doi.org/10.1093/aob/mcp172>.
- Vila, D.A., De Goncalves, L.G.G., Toll, D.L. and Rozante, J.R., (2009). Statistical evaluation of combined daily gauge observations and rainfall satellite estimates over continental South America. *Journal of Hydrometeorology*, 10(2), pp.533-543. Available from: <https://doi.org/10.1175/2008JHM1048.1>
- Vogel, R.M., Yaindl, C. and Walter, M., (2011). Non-stationarity: flood magnification and recurrence reduction factors in the United States. *Journal of the American Water Resources Association*, 47 (3), 464–474. Available from: <https://doi.org/10.1111/j.1752-1688.2011.00541.x>.
- Volkman, T.H., Lyon, S.W., Gupta, H.V. and Troch, P.A., (2010). Multi criteria design of rain gauge networks for flash flood prediction in semiarid catchments with complex terrain. *Water Resources research*, 46(11). Available from: <https://doi.org/10.1029/2010WR009145>
- Wadsworth, Harrison M. (1990). *Handbook of Statistical Methods for Engineers and Scientists*. New York, New York: McGraw-Hill, Inc.
- Wang, S., Zhang, Z., Sun, G., Strauss, P., Guo, J. and Tang, Y., (2012). Distributed hydrological modelling in a large-scale watershed of Northern China: multi-site model calibration, validation and sensitivity analysis. *Hydrology & Earth System Sciences Discussions*, 9(5). Available from: <https://doi.org/10.5194/hess-16-4621-2012>

- Wang W, Lu H, Yang D, Sothea K, Jiao Y, Gao B, et al. (2016) Modelling Hydrologic Processes in the Mekong River Basin Using a Distributed Model Driven by Satellite Precipitation and Rain Gauge Observations. *PLoS ONE* 11(3): e0152229. <https://doi.org/10.1371/journal.pone.0152229>. Available from: <https://doi.org/10.1371/journal.pone.0152229>
- Warmink, J.J., Van der Klis, H., Booij, M.J. and Hulscher, S.J., (2011). Identification and quantification of uncertainties in a hydrodynamic river model using expert opinions. *Water Resources Management*, 25(2), pp.601-622. Available from: <https://doi.org/10.1007/s11269-010-9716-7>
- Weerts, A.H., Winsemius, H.C. and Verkade, J.S., (2011). Estimation of predictive hydrological uncertainty using quantile regression: examples from the National Flood Forecasting System (England and Wales). *Hydrology and Earth System Sciences*, 15(1), pp.255-265. Available from: <https://doi.org/10.5194/hess-15-255-2011>
- Wohl, E., Barros, A., Brunzell, N., Chappell, N.A., Coe, M., Giambelluca, T., Goldsmith, S., Harmon, R., Hendrickx, J.M., Juvik, J. and McDonnell, J., (2012). The hydrology of the humid tropics. *Nature Climate Change*, 2 9, p.655. Available from: <https://doi.org/10.1038/nclimate1556>
- World Health Organisation (WHO) (2012). Health resources availability mapping system (HeRAMS). Geneva: *World Health Organisation*. Available from: <https://www.who.int/initiatives/herams> (Accessed: 13 April 2017).
- Wicks, J.M.; Bathurst, J.C. (1996) SHESED: A physically based, distributed erosion and sediment yield component for the SHE hydrological modelling system. *Journal of Hydrology*, 1996, 175, 213–238. Available from: [https://doi.org/10.1016/S0022-1694\(96\)80012-6](https://doi.org/10.1016/S0022-1694(96)80012-6)

Wilby, R.L. and Keenan, R., (2012). Adapting to flood risk under climate change. *Progress in Physical Geography*, 36(3), pp.348-378. Available from: <https://doi.org/10.1177/0309133312438908>

Wilby, R.L., Clifford, N.J., De Luca, P., Harrigan, S., Hillier, J.K., Hodgkins, R., Johnson, M.F., Matthews, T.K., Murphy, C., Noone, S.J. and Parry, S., (2017). The 'dirty dozens of freshwater science": detecting then reconciling hydrological data biases and errors. *Wiley Interdisciplinary Reviews: Water*, 4(3), p.e1209. Available from: <https://doi.org/10.1002/wat2.1209>

Wilkinson, M.E. and Bathurst, J.C., (2018). A multi-scale nested experiment for understanding flood wave generation across four orders of magnitude of the catchment area. *Hydrology Research*, 49(3), pp.597-615. Available from: <https://doi.org/10.2166/nh.2017.070>

Willems, P., Olsson, J., Arnbjerg-Nielsen, K., Beecham, S., Assela, P., Gregersen, I.B., Madsen, H. and Van-Thanh-Van Nguyen, (2012). Impacts of climate change on rainfall extremes and urban drainage systems (p. 238). London, UK: IWA. Available from: <https://doi.org/10.2166/wst.2013.251>

Williams, P. and Berlamont, J., (2002). Accounting for the spatial rainfall Variability in urban modelling applications. *Water science and technology*, 45(2), pp.105-112. Available from: <https://doi.org/10.2166/wst.2002.0034>

Willner Sven N., Levermann Anders, Zhao Fang, Frieler Katja, (2018): Adaptation required to preserve future high-end river flood risk at present levels. *Science Advances*. Available from: [DOI:10.1126/sciadv.aao1914](https://doi.org/10.1126/sciadv.aao1914)

- Wittwer, C. (2013). 'WMO/UNESCO-IHP guideline for flood forecasting'. In *Regional Workshop on hydrological forecasting and impact of climate change on water resources*. Pržno/Budva, Montenegro. Available from: [https://sustainabledevelopment.un.org/content/documents/flood\\_guidelines.pdf](https://sustainabledevelopment.un.org/content/documents/flood_guidelines.pdf) (Accessed: 30 February 2018).
- Wigmosta, M.S., Vail, L.W. and Lettenmaier, D.P., (1994). A distributed Hydrology vegetation model for complex terrain. *Water Resources research*, 30(6), pp.1665-1679. Available from: <https://doi.org/10.1029/94WR00436>
- Wu, H., Adler, R.F., Tian, Y., Huffman, G.J., Li, H. and Wang, J., (2014). Real-time global flood estimation using satellite-based precipitation and a coupled land surface and routing model. *Water Resources Research*, 50(3), pp.2693-2717. Available from: <https://doi.org/10.1002/2013WR014710>
- Xevi, E., Christiaens, K., Espino, A., Sewnandan, W., Mallants, D., Sørensen, H. and Feyen, J., (1997). Calibration, validation and sensitivity analysis of the MIKE-SHE model using the Neuenkirchen catchment as case study. *Water Resources Management*, 11(3), pp.219-242. Available from: <https://doi.org/10.1023/A:1007977521604>
- Xu, C.Y., (2002). *Hydrologic models*. Textbooks of Uppsala University. Department of Earth Sciences Hydrology, Uppsala University, Vol 72
- Xu, T., Valocchi, A.J., Ye, M., Liang, F. and Lin, Y.F., (2017). Bayesian calibration of groundwater models with input data uncertainty. *Water Resources research*, 53(4), pp.3224-3245. Available from: <https://doi.org/10.1002/2016WR019512>
- Xu, X., Li, J. and Tolson, B.A., (2014). Progress in integrating remote sensing data and hydrologic modelling. *Progress in Physical Geography*, 38(4), pp.464-498. Available from: <https://doi.org/10.1177/0309133314536583>

- Xuan, Y., Cluckie, I.D. and Wang, Y., (2009). Uncertainty analysis of hydrological ensemble forecasts in a distributed model utilising short-range rainfall prediction. *Hydrology and Earth System Sciences*, 13(3), pp.293-303. Available from: <https://doi.org/10.5194/hess-13-293-2009>
- Yamazaki, D., Kanae, S., Kim, H. and Oki, T., (2011). A physically-based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research*, 47(4). Available from: <https://doi.org/10.1029/2010WR009726>
- Yan, K., (2015). *Low-cost space-borne data for inundation modelling: topography, flood extent and water level*; Doctoral Dissertation, UNESCO-IHE Institute for Water Education, Delft and Delft University of Technology. Available from: <https://cdm21063.contentdm.oclc.org/digital/collection/phd1/id/581> (Accessed: 10 April 2018).
- Yang, J., Jakeman, A., Fang, G. and Chen, X., (2018). Uncertainty analysis of the semi-distributed hydrologic model based on a Gaussian Process emulator. *Environmental Modelling & Software*, 101, pp.289-300. Available from: <https://doi.org/10.1016/j.envsoft.2017.11.037>
- Yu Xuan, Gopal Bhatt, Christopher Duffy, Denice Wardrop, Raymond Najjar, Andrew Ross, Matthew Rydzik (2015): A coupled surface-subsurface modelling framework to assess the impact of climate change on freshwater wetlands: a case study of 7 watersheds of the Susquehanna River Basin. *Climate Research*, 66:211-228. Available from: <https://doi.org/10.3354/cr01348>
- Yu, D., Yin, J. and Liu, M., (2016). Validating city-scale surface water flood modelling using crowd-sourced data. *Environmental Research Letters*, 11(12), p.124011. Available from: <https://doi.org/10.1088/1748-9326/11/12/124011>



- Zarzar, Christopher M., Hossein Hosseiny, Ridwan Siddique, Michael Gomez, Virginia Smith, Alfonso Mejia and Jamie Dyer, (2018). A hydraulic Multimodel Ensemble Framework for Visualising Flood Inundation Uncertainty. *Journal of the American Water Resources Association (JAWRA)* 1–13. Available from: <https://doi.org/10.1111/1752-1688.12656>
- Zhang, R., Santos, C.A., Moreira, M., Freire, P.K. and Corte-Real, J., (2013). Automatic calibration of the SHETRAN hydrological modelling system using MSCE. *Water resources management*, 27(11), pp.4053-4068. Available from: <https://doi.org/10.1007/s11269-013-0395-z>
- Zhang, A., Shi, H., Li, T. and Fu, X., (2018). Analysis of the Influence of Rainfall Spatial Uncertainty on Hydrological Simulations Using the Bootstrap Method. *Atmosphere*, 9(2), p.71. Available from: <https://doi.org/10.3390/atmos9020071>
- Zhang, R., Moreira, M. and Corte-Real, J., (2015). Multi-objective calibration of the physically based, spatially distributed SHETRAN hydrological model. *Journal of Hydroinformatics*, 18(3), pp.428-445. Available from: <https://doi.org/10.2166/hydro.2015.219>
- Zhuo, L., Han, D., & Dai, Q. (2016). Soil moisture deficit estimation using satellite multi-angle brightness temperature. *Journal of Hydrology*, 539, 392405. Available from: <https://doi.org/10.1016/j.jhydrol.2016.05.052>

## 8. APPENDICES

The following sections provide additional information on rainfall interpolation, error quantification of rainfall input correction for the SHETRAN hydrological model. These include average monthly rainfall plotted in the Grijalva basin since 1983, rainfall interpolation and error maps, mean monthly rainfall errors in Grijalva catchment and seasonal rainfall distribution in the Grijalva catchment area.

### Appendix A Rainfall input data

#### A.1 Average monthly rainfall for the Grijalva basin

Rainfall data were obtained from thirty-three rain gauges, which were unevenly distributed but almost covering the study area.

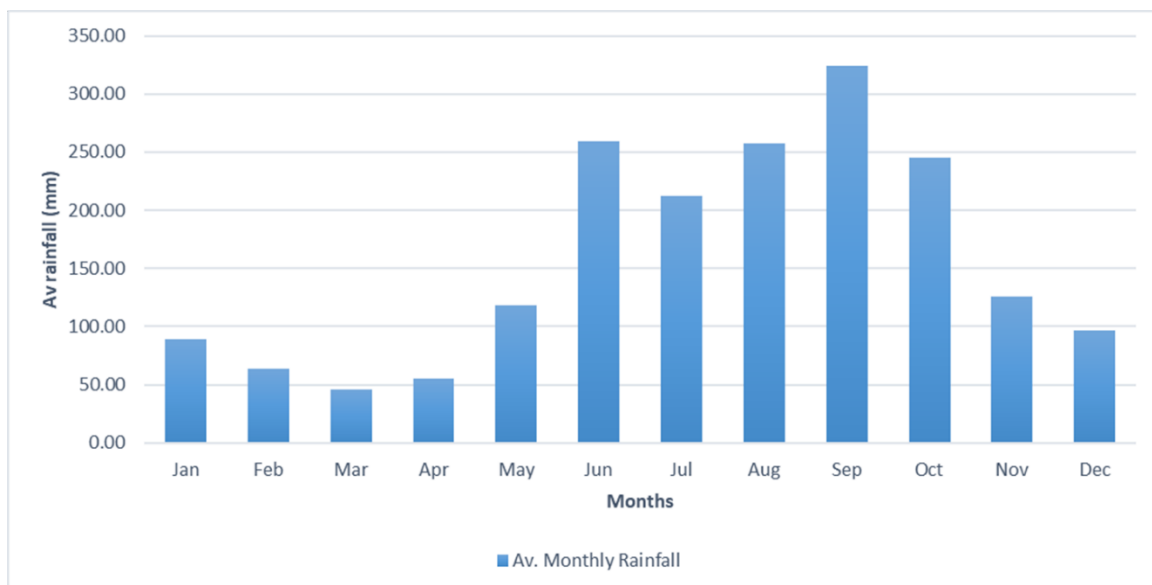
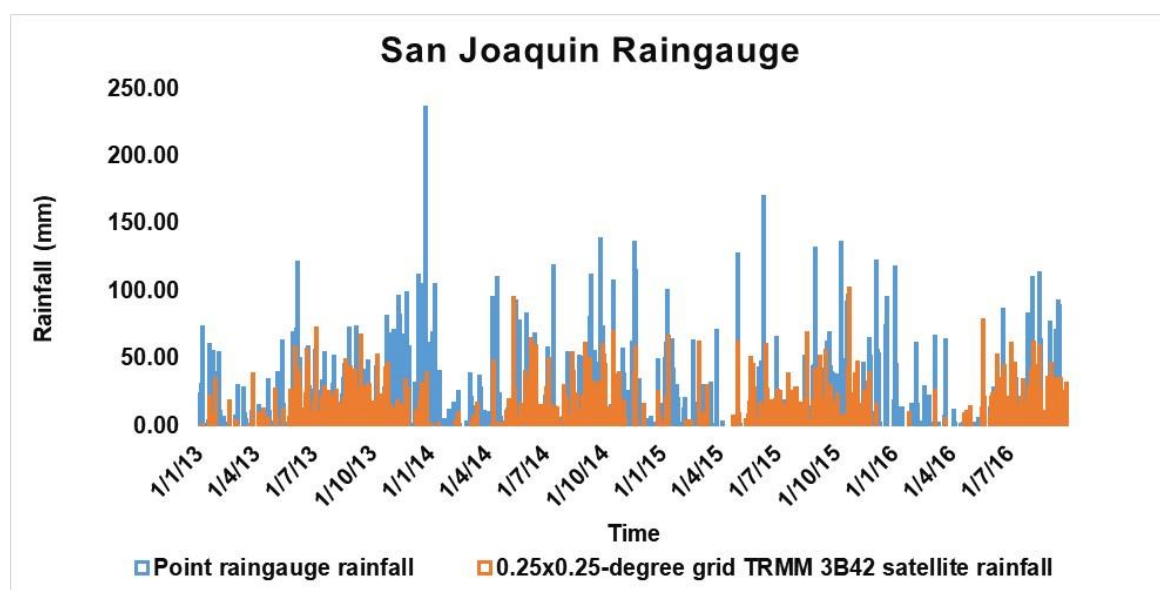
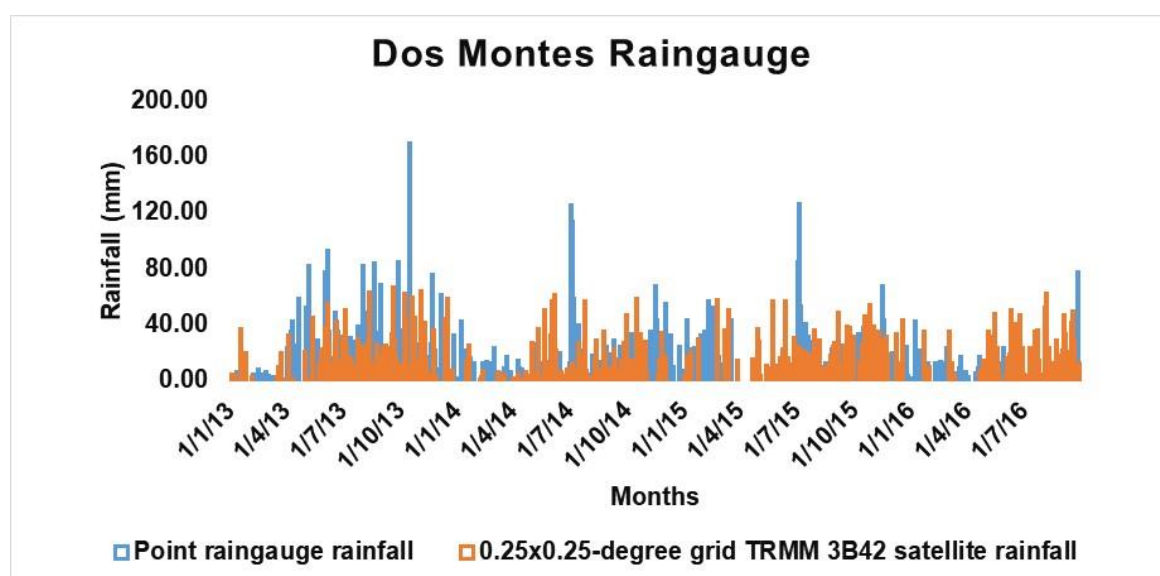
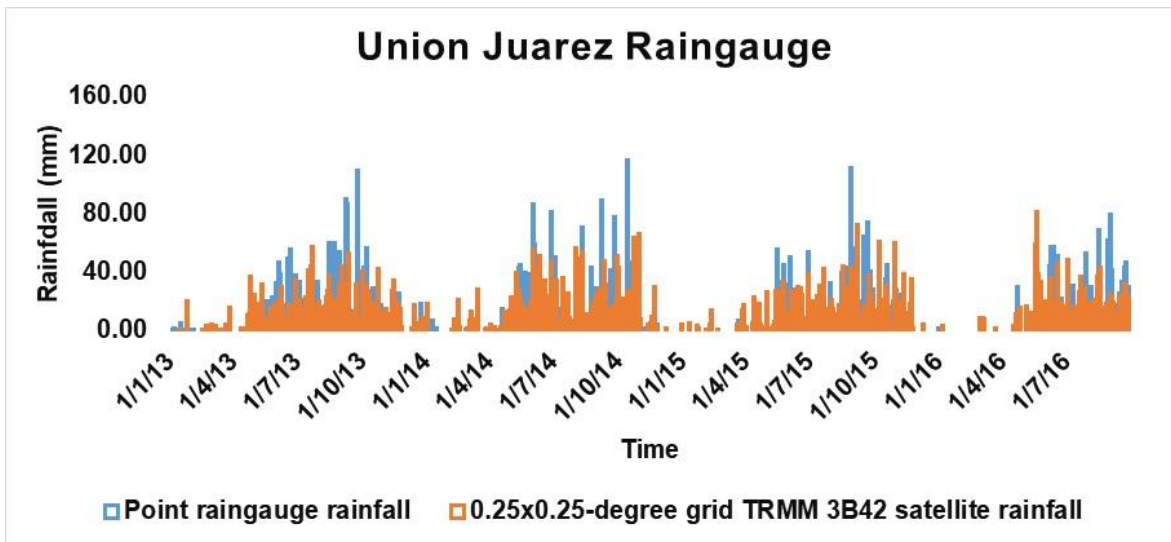
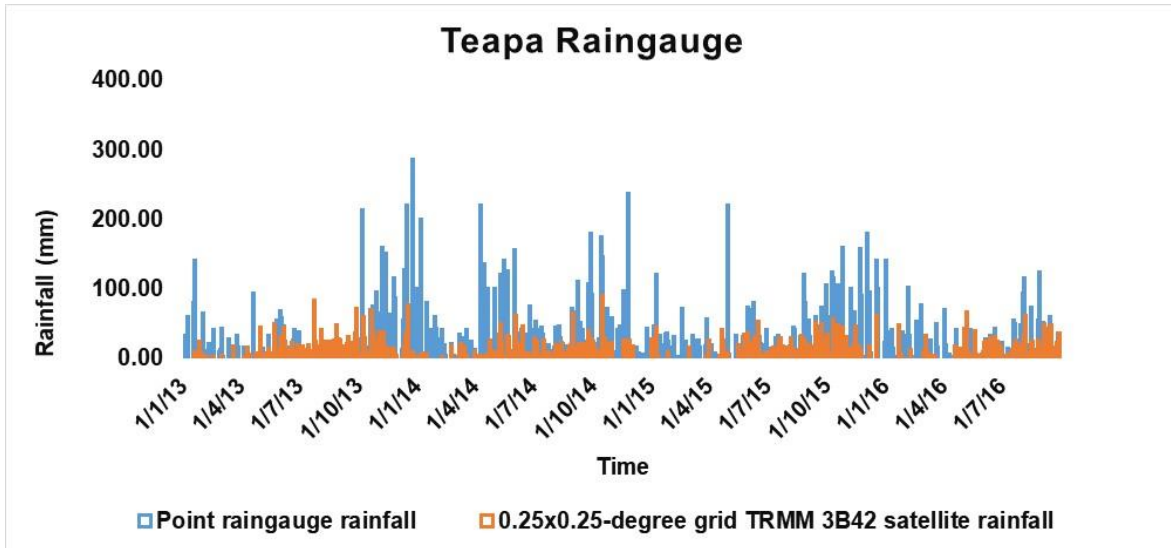


Figure A-1: Average monthly rainfall for the Grijalva basin, Mexico

## A.2 Rain gauge and TRMM 3B42 V7 rainfall comparisons

The figures below show selected rain gauge and pixel satellite (3B42V7, 0.25x0.25-degree) rainfall data for the calibration period 2013 to 2016 and offer the first impression of rainfall inter-annual and annual variability. The paired datasets follow a similar inter-annual trend and reasonably describe the annual rainfall pattern in the period under consideration. However, the 3B42V7 satellite rainfall data generally underestimate rainfall and is of low accuracy, particularly at higher temporal and spatial resolutions (Grimes et al., 1999; Dinku et al., 2014). Rain gauges provide direct point quantitative rainfall estimates, but they cannot describe the spatial variability of rainfall required as input data into gridded distributed hydrological models.





**Figure A-2: Direct comparison of selected raw rain gauge and satellite (pixel) rainfall over 2013 to 2016 calibration period to give the first impression of data inter-annual and annual variability**

### A.3 Monthly interpolated total rain gauge rainfall

A Geostatistical technique was used to obtain improved rainfall input data for the SHETRAN hydrological model. A Digital Elevation Model (DEM) was used to integrate topography as additional information to enhance the interpolation of point rain gauge rainfall to aerial catchment coverage (Lloyd and Atkinson, 2004). In this process mean monthly rainfall values from each rain gauge location upscaled to a 2 x 2 km aerial resolution dataset using the Block Kriging technique.

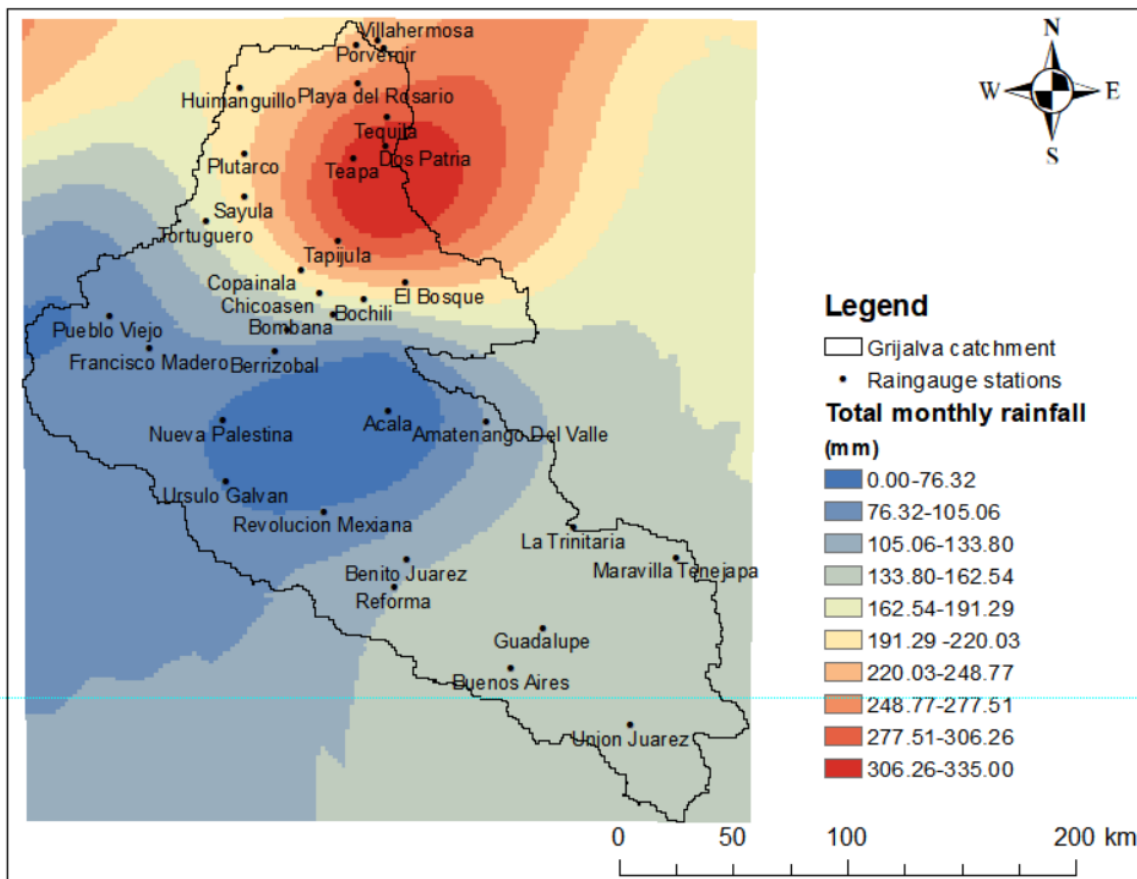


Figure A-3: Rain gauge distribution and mean monthly rain gauge rainfall distribution interpolated at a 2 x 2 grid-scale over 2013 to 2017 period

#### A.4 : Time-averaged TRMM TRMM\_3B42\_Daily v7 satellite rainfall data

Satellite rainfall data were available at 3-hour and also at daily (24 hours) and monthly temporal resolutions, however, the datasets were at the same spatial resolution of 0.25° x 0.25° (approximately 25 x 25 km) (Acker and Leptoukh, 2007; Huffman et al., 2010; Verdin et al., 2016). The latest Version 7 of the TRMM 3B42 of this satellite rainfall product (0.25° x 0.25°, approximately 25 x 25 km) provided the spatial description of rainfall information, particularly in areas where rain gauges were not available. (Dinku et al., 2014).

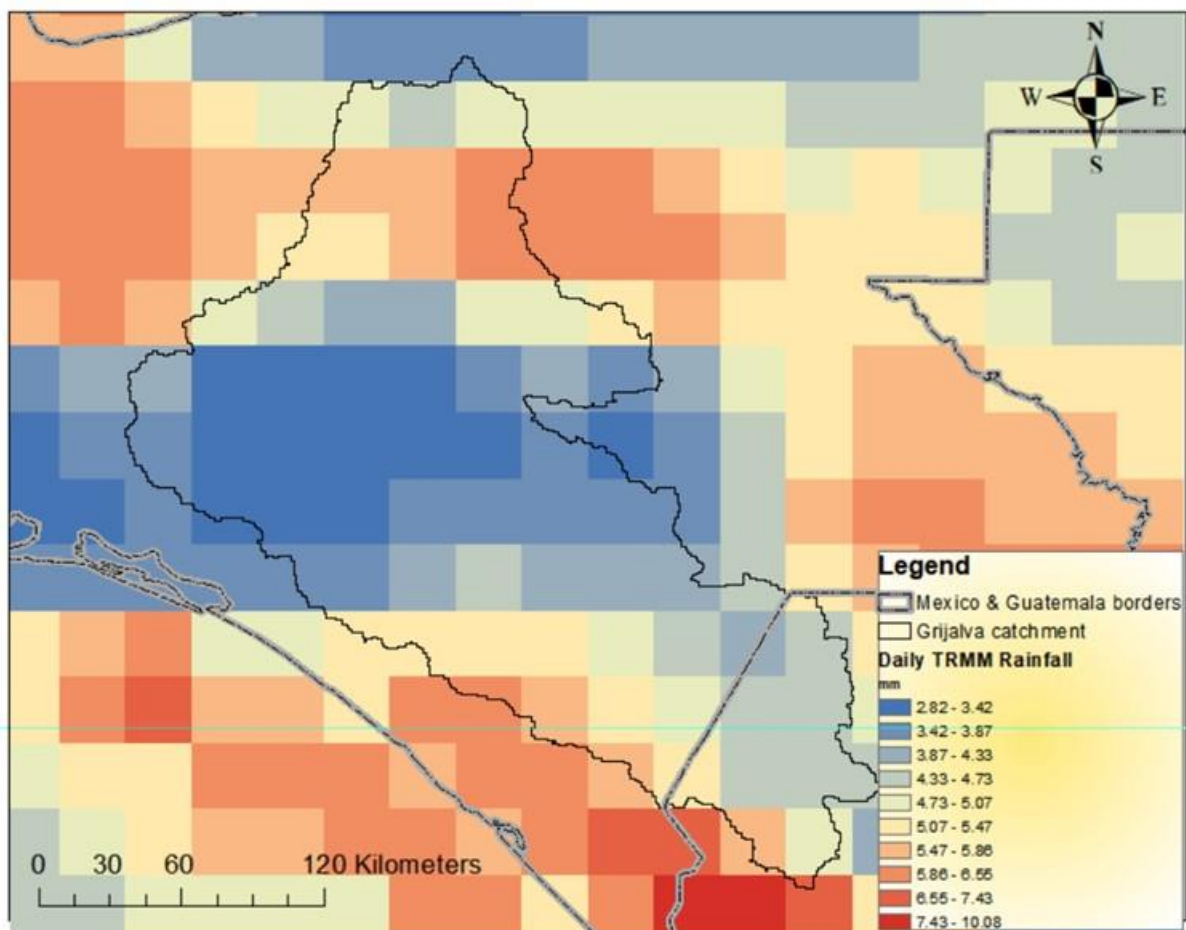


Figure A-4: Time Averaged Map of daily rainfall over 2013 to 2017 period at 0.25 degrees spatial scale (approx. 25 x 25 km) [TRMM TRMM\_3B42\_Daily v7] in Grijalva, Mexico.:

## A.5 Mean monthly rainfall and bias

Table A1: Mean rainfall for each gauging station based on four-year rainfall data.

	RAINGAUGE	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	mean rainfall	Merged	Bias
1	San Juan	0.25	1.25	2.25	3.25	4.25	5.25	2.90	3.87	8.45	3.47	0.39	0.28	2.99	2.53	0.46
2	Tequila	4.95	4.46	1.80	3.29	4.44	5.93	12.89	8.46	7.22	13.2	9.70	3.91	6.69	6.31	0.38
3	Dos Montes	6.29	0.82	2.77	2.21	4.36	4.49	4.44	4.97	6.68	10.1	10.9	8.70	5.57	5.44	0.13
4	Tapilula	2.48	0.92	1.12	2.20	4.98	7.66	3.09	6.48	10.5	7.61	3.70	3.34	4.51	4.37	0.14
5	Maravi Tene	0.05	1.05	2.05	3.05	4.05	5.05	1.09	5.25	5.49	3.62	0.45	0.12	2.61	2.03	0.58
6	Playa del R	5.47	1.82	1.44	0.98	3.36	7.93	4.17	6.00	8.53	14.9	8.29	5.93	5.74	5.5	0.24
7	Guadalupe	0.36	0.02	0.17	1.17	3.29	8.46	5.42	9.65	15.7	6.85	0.64	1.00	4.32	4.06	0.26
8	Porvenir	6.48	1.05	3.35	1.29	2.57	3.32	3.34	3.44	11.2	10.6	11.3	11.6	5.80	5.52	0.28
9	Dos Patria	12.40	4.05	5.66	4.00	6.38	12.33	7.36	14.01	18.1	16.9	15.5	17.5	11.20	10.86	0.34
10	Buenos Aires	0.19	0.02	0.17	0.82	4.52	11.65	6.82	12.11	17.5	8.47	2.35	0.24	5.41	5.12	0.29
11	Teapa	0.13	0.02	0.29	0.92	5.97	7.60	3.76	4.81	7.31	3.49	1.08	0.15	2.96	2.67	0.29
12	La Trinitaria	3.75	0.70	1.90	0.93	3.32	4.93	5.36	8.03	9.00	17.6	12.3	10.9	6.57	6.2	0.37
13	Tortuquero	3.71	0.75	2.44	2.58	1.77	8.95	5.09	9.05	8.30	16.5	16.2	16.5	7.66	7.3	0.36
14	Savula	0.20	0.40	0.40	1.78	6.66	16.57	15.11	9.30	19.7	8.58	2.40	0.00	6.76	6.42	0.34
15	Plutarco	0.20	0.03	0.50	0.42	3.39	9.02	6.98	5.78	12.6	7.31	2.92	0.14	4.11	3.79	0.32
16	Huimanquillo	0.80	1.80	2.80	3.80	4.80	9.92	6.09	5.11	17.1	0.77	0.15	1.00	4.43	3.6	0.83
17	Villahermosa	2.24	0.80	0.77	1.60	4.37	6.77	2.05	3.45	10.7	6.13	3.92	3.56	3.87	3.61	0.26
18	Francisco M	4.47	1.66	1.17	1.28	1.24	4.80	2.11	3.21	8.52	4.10	4.67	3.47	3.39	3.12	0.27
19	Las Flores	0.50	0.14	0.06	0.71	6.27	12.08	6.17	9.84	9.02	2.96	0.49	0.41	4.05	3.62	0.43
20	Berrizobal	2.42	3.73	1.84	0.45	3.70	7.22	2.49	6.14	8.98	2.95	1.25	0.18	3.44	3.12	0.32
21	Bombana	2.33	1.19	0.85	2.02	3.92	9.24	4.19	6.56	12.7	8.42	5.21	5.80	5.20	4.86	0.34
22	Pueblo Viejo	0.90	0.80	0.48	0.73	2.43	7.14	2.96	3.19	11.8	3.26	0.42	0.90	2.84	2.43	0.41
23	Acala	0.98	0.13	0.23	1.55	4.90	8.04	3.14	6.79	10.5	3.28	2.10	1.22	3.57	3.36	0.21
24	Bochili	0.90	1.00	2.00	3.00	4.00	5.00	4.76	6.07	10.4	1.97	0.18	0.25	3.29	3.06	0.23
25	Copainala	0.09	0.04	0.08	0.78	5.60	10.70	9.00	13.04	15.7	8.03	2.09	0.17	5.45	5.04	0.41
26	Benito Juárez	0.68	0.00	0.25	1.02	4.50	5.63	3.48	10.62	11.7	7.92	5.99	1.71	4.46	4.25	0.21
27	Reforma	0.19	0.02	0.17	0.82	4.47	11.65	6.82	12.11	17.5	8.47	2.35	0.24	5.41	5.06	0.35
28	Revolucion M	0.72	2.91	3.75	7.03	12.67	18.52	8.22	11.20	23.7	13.1	4.47	1.93	9.03	8.64	0.39
29	Union Juárez	0.07	0.08	0.48	1.29	5.38	12.47	7.44	10.69	13.8	8.13	2.14	0.25	5.19	4.89	0.3
30	San Francisco	0.13	0.02	0.29	0.92	5.93	7.60	3.76	4.81	7.31	3.49	1.08	0.15	2.96	2.69	0.27
31	Amatenango D	0.86	0.14	0.32	1.25	4.17	8.10	2.96	6.43	9.10	5.45	1.87	1.48	3.51	3.18	0.33
32	Chicoasen	2.26	1.06	0.80	1.31	4.43	9.28	5.95	7.80	13.5	5.84	4.11	4.35	5.06	4.71	0.35





## A.6 Seasonal rainfall distribution in Grijalva catchment

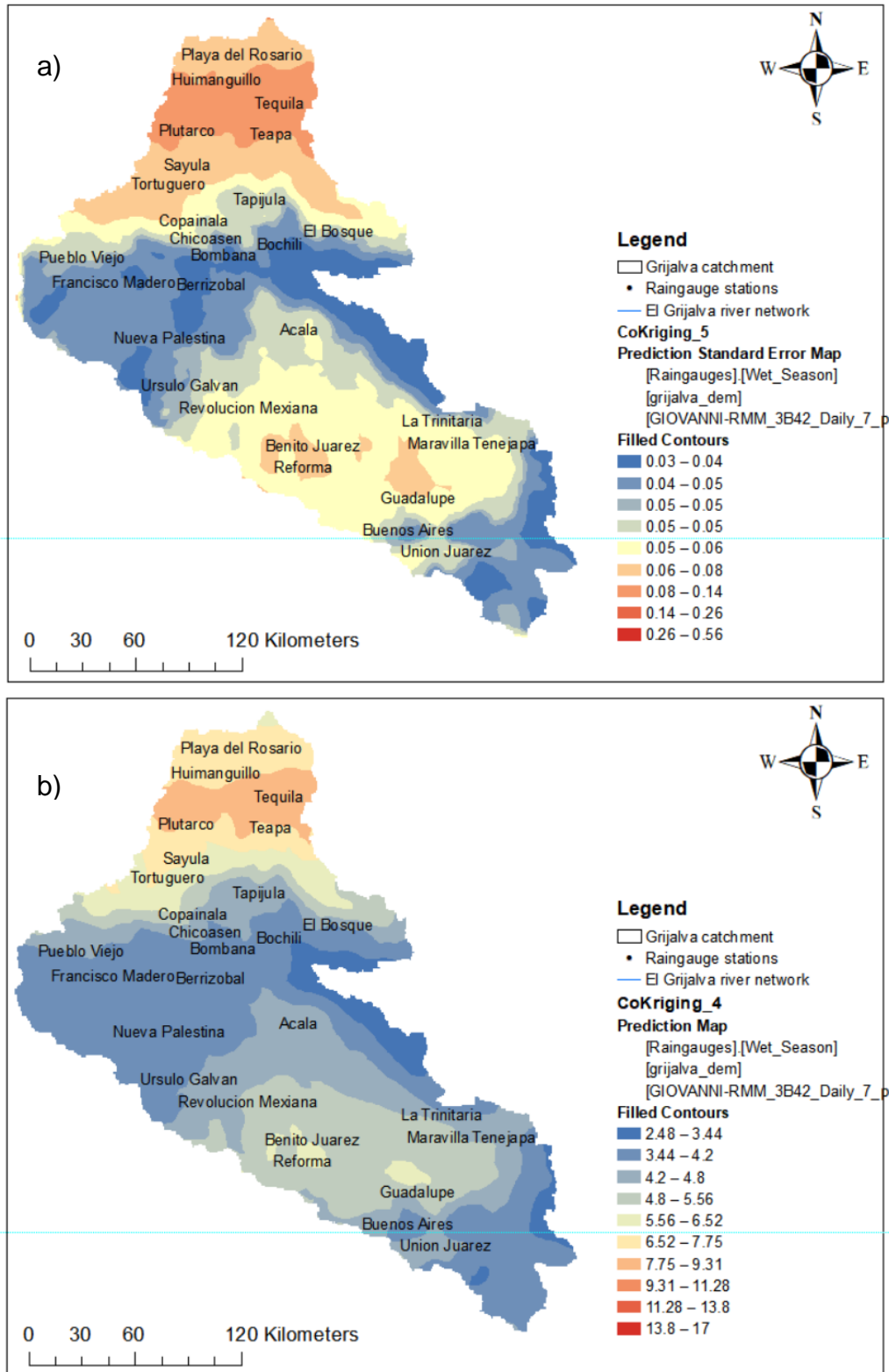


Figure A-5: Wet season predicted rainfall and errors for Grijalva catchment showing (a) Interpolated rainfall distribution (mm) (b) Mean rainfall error fields (mm)

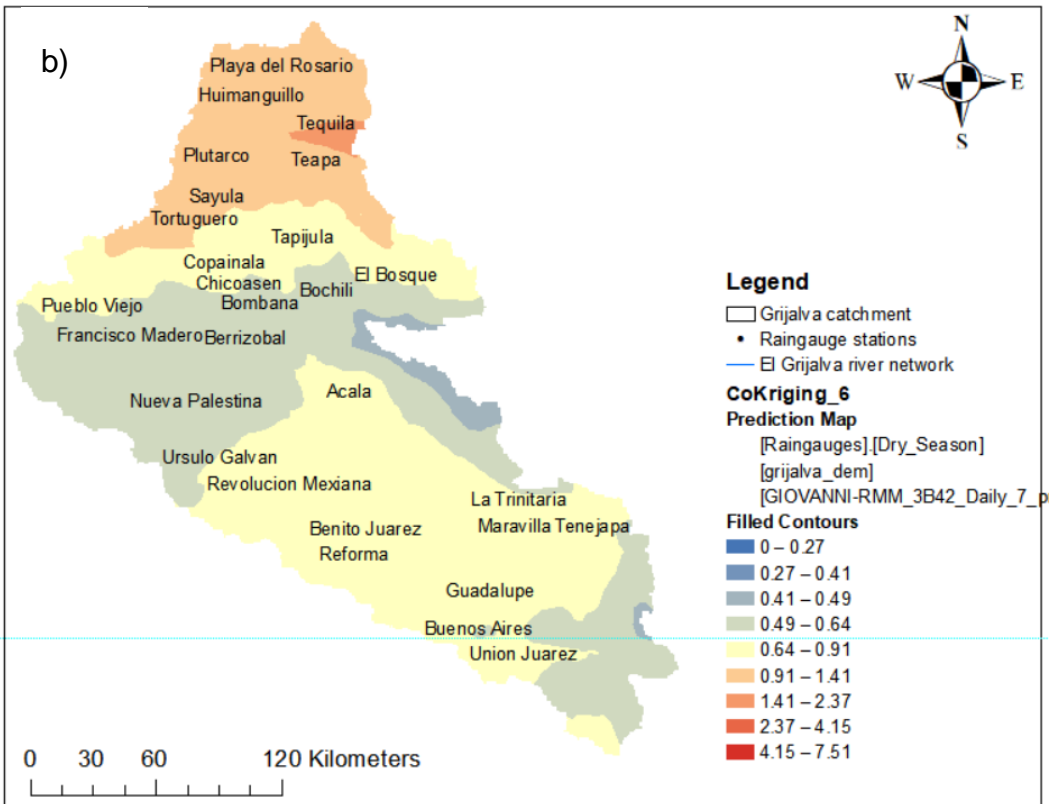
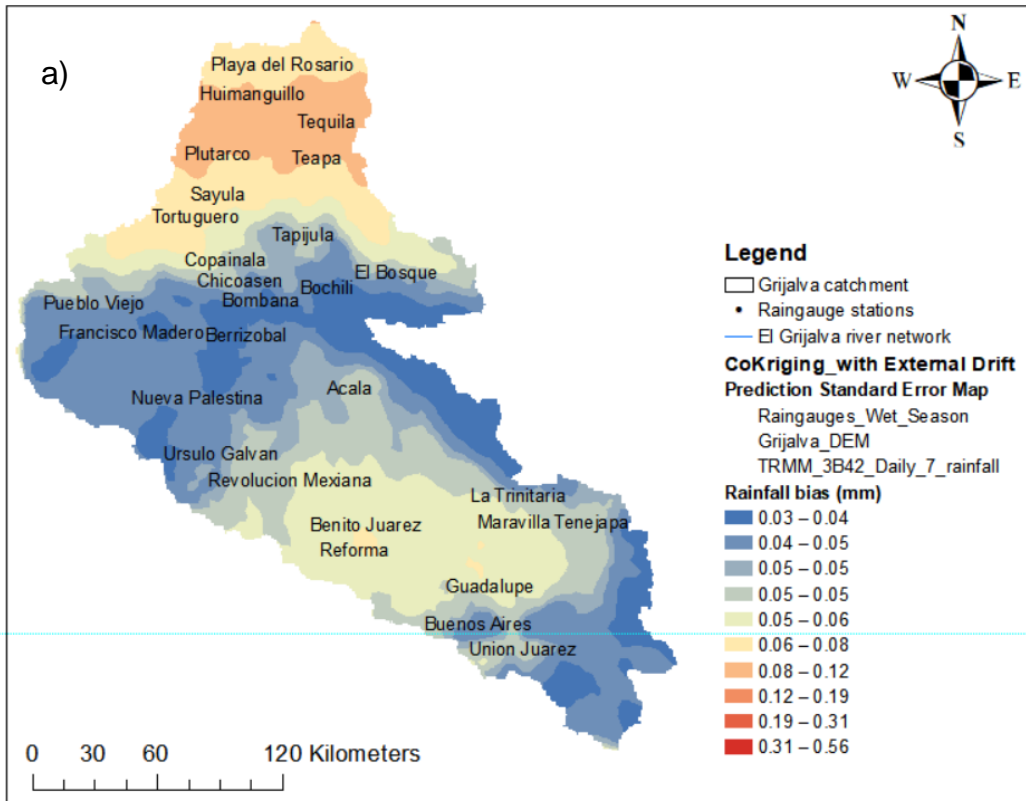


Figure A-6: Dry season predicted rainfall and errors for Grijalva catchment showing (a) Interpolated rainfall distribution (mm) (b) Mean rainfall error fields (mm)

## Appendix B : Diverted flows calculations

The maps below show an overview of the Grijalva catchment and the in the lower Grijalva River (in the red square), where diverted flows are diverted into Samaria River. The SHETRAN grid is also shown without the diverted flows and a map showing the actual situation of the diverted flows.



Figure B-1: An overview of the model domain showing the Grijalva River, tributaries and the main four burn

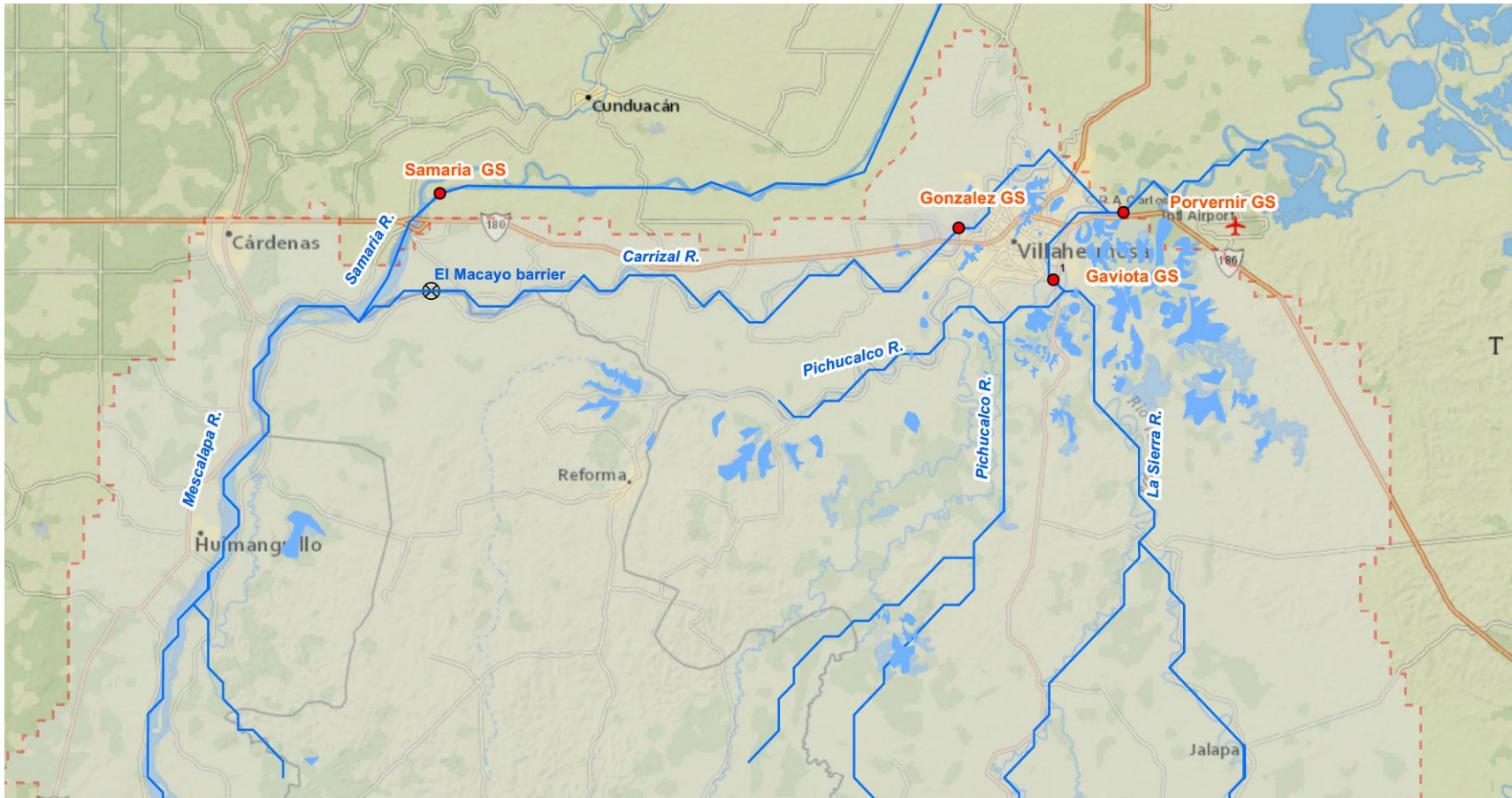


Figure B-2: In the lower Grijalva River, flows are diverted from the model domain into a distributary channel referred to as Samaria River

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	Year	Month	Q_dy_01	Q_dy_02	Q_dy_03	Q_dy_04	Q_dy_05	Q_dy_06	Q_dy_07	Q_dy_08	Q_dy_09	Q_dy_10	Q_dy_11	Q_dy_12	Q_dy_13	Q_dy_14	Q_dy_15	Q_dy_16	Q_dy_17	Q_dy_18	Q_dy_19	Q_dy_20	Q_dy_21
38	2012	9	634.208	632.552	623.623	606.757	611.728	585.514	579.416	567.156	562.312	562.088	575.679	494.279	418.728	398.171	389.017	397.396	393.604	435.403	518.273	549.098	613.435
39	2012	8	307.282	275.867	277.842	274.122	263.503	255.396	237.149	245.28	304.319	447.851	581.344	637.196	760.591	778.924	774.766	776.389	775.46	787.255	781.271	729.009	670.111
40	2012	7	354.866	341.228	332.614	329.646	357.646	358.271	349.292	345.733	343.439	326.78	353.814	366.349	365.115	363.094	360.255	336.958	327.004	334.733	318.809	305.109	263.939
41	2012	6	327.032	318.458	301.458	290.792	301.576	327.333	329.118	325.292	322.75	306.362	289.667	291.562	324.417	336.208	334.042	332.101	337.711	355.284	410.145	429.576	429.446
42	2012	5	326.104	327.284	315.022	324.902	324.25	321.693	318.026	306.388	316.091	322.962	322.929	320.923	319.714	324.944	330.105	393.857	430.375	444.542	435.625	415.939	364.275
43	2012	3	465.79	456.458	459.318	457.021	457.906	466.875	474.474	469.297	459.365	447.784	440.744	429.911	417.224	425.229	425.781	407.167	405.51	402.609	381.078	341.289	339.526
44	2012	1	377.17	364.562	441.508	563.134	583.057	569.979	550.339	529.516	473.628	394.995	401.777	392.801	391.221	491.875	589.811	620.55	618.938	616.044	604.945	600.214	593.203
45	2012	4	361.01	360.694	348.38	354.854	355.833	352.542	350.694	349.174	341.861	336.383	354.523	355.247	354.039	347.922	345.874	338.906	334.866	354.933	386.58	387.617	363.578
46	2013	12	941.521	926.637	885.188	874.59	859.167	841.594	829.879	814.917	789.263	785.725	645.23	645.358	680.863	704.958	783.862	941.577	979.988	1020.442	1063.212	1080.581	1081.094
47	2013	6	178	182.324	319.059	445.863	490.627	516.427	530.324	521.376	503.958	502.975	491.812	475.477	447.046	420.197	392.458	380.721	375.244	346.083	313.757	354.936	431.485
48	2013	5	103.007	104.597	103.193	109.595	139.131	157.203	159.93	174.265	163.932	160.016	164.394	175.153	243.688	192.06	270.72	224.221	164.661	144.14	140.583	133.854	144.613
49	2013	4	114.732	103.108	102.394	107.981	127.872	111.299	122.324	121.715	118.799	113.385	122.958	119.917	115.375	114.813	109.719	108.835	110.018	108.824	113.802	119.141	117.922
50	2013	3	163.62	152.135	59.043	65.987	57.43	51.942	42.581	42.34	40.015	39.693	99.115	164.776	164.129	161.501	163.76	167.886	164.666	164.906	164.03	157.824	158.057
51	2013	2	377.092	412.174	412.36	405.246	418.507	450.293	435.89	415.583	402.833	392.896	370.062	352.279	336.968	311.69	306.452	302.688	305.821	352.776	344.691	316.251	291.524
52	2013	1	488.627	477.936	457.994	455.274	465.104	476.958	490.874	537.49	544.31	535.241	499.495	476.271	439.232	382.231	365.754	360.856	392.829	485.544	505.375	473.833	466.817
53	2013	9	619.035	572.875	570.271	561.582	559.849	608.731	632.5	677	711.179	626.179	626.263	638.297	642.955	653.875	646.023	637.854	615.001	582.787	586.344	592.861	593.713
54	2013	8	330.004	351.974	359.479	334.997	346.113	332.341	333.261	323.648	275.389	258.101	265.256	286.819	309.405	338.57	378.445	446.635	467.963	493.875	510.368	490.198	477.316
55	2013	11	717.06	712	709.795	684.48	678.044	672.699	667.682	674.712	694.322	784.696	954.868	1001.949	1074.302	1107.357	1129.286	1120.642	1111.432	1079.026	1064.311	1067.616	1063.67
56	2013	10	510.229	591.968	614.349	678.067	671.283	647.459	575.678	592.029	622.539	639.462	633.539	629.459	628.099	597.02	595.053	587.43	573.082	556.393	555.198	555.688	570.966
57	2013	7	383.474	427.017	475.738	497.489	516.335	552.401	529.828	483.196	442.515	432.869	434.664	469.167	478.605	490.688	494.915	513.225	500.671	470.666	459.55	470.528	464.52
58	2014	5	464.583	513.096	508.967	535.638	551.887	526.85	513.75	499.108	488.749	492.685	470.396	478.042	466.854	472.723	492.866	574.432	602.5	597.101	580.736	566.802	596.375
59	2014	1	1009.643	998.855	993.043	992.563	1001.014	1004.258	1026.827	1067.992	1074.337	1072.078	1067.452	1060.732	1018.56	995.856	984.266	975.516	965.793	955.87	948.55	933.172	917.931
60	2014	2	623.354	610.765	592.804	584.545	577.936	569.583	566.339	564.563	557.971	550.457	556.605	532.907	522.543	520.781	520.031	515.681	500.534	504.754	495.457	485.598	471.696
61	2014	3	421.114	415.87	407.642	401.396	413.969	413.417	417.437	417.155	418.351	415.473	406.719	414.073	414.221	462.694	478.75	465.494	466.714	434.938	445.396	448.479	440.188
62	2014	12	995.518	972.99	962.551	952.978	942.199	926.611	908.722	893.5	881.167	877.5	873.167	864.5	856.389	846.5	821.444	800.722	740.956	543.408	508.622	503.558	494.944
63	2014	4	443.199	440.78	437.312	433.937	430.426	423.387	409.583	405.539	501.786	530.449	509.25	489.955	480.256	468.935	459.313	489.979	555.833	559.969	529.031	542.479	575.977
64	2014	9	619.604	676.361	635.262	605.218	589.375	574.307	454.681	408.442	477.31	534.378	552.281	554.068	566.848	577.179	666.443	677.09	675.871	662.791	647.549	631.646	631.396
65	2014	8	489.016	477.392	473.969	484.734	513.069	513.373	517.723	512.57	493.58	450.444	377.203	368.982	393.183	402.612	387.459	386.238	378.34	331.18	284.314	279.031	281.869
66	2014	7	644.375	659.59	644.25	648.474	672.767	679.114	691.485	697.963	699.63	685.466	672.567	661.768	674.67	619.3	650.813	644.125	643.615	632.167	626.274	619.592	609.695
67	2014	6	744.724	781.196	787.668	823.523	872.472	866.719	869.854	876.75	876	875.156	872.156	869.141	863.578	866.125	859.633	847.001	861.604	843.688	827.805	802.133	795.292
68	2014	10	907.447	904.729	892.778	873.611	835.829	806.347	785.647	746.722	715.478	706.463	693.316	688.556	688.407	682.729	712.394	783.627	881.453	883.756	859.082	800.857	781.259
69	2014	11	1009.018	1000.309	1010.442	984.278	944.832	908.889	883.222	867.722	855.444	831.889	745.711	650.232	620.563	614.688	616.325	617.72	617.113	649.684	800.248	955.566	991.801

Figure B-3: Raw observed discharge at Povenir gauging station

Estaciones	PORVENIR		SAMARIA		GONZALEZ		TAPIJULAPA		TEAPA		PUYACATENGO		Pueblo Nuevo		OXOLOTAN		Gaviotas	
	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge
16/05/2015	0.77	286.00	12.99	681.00	4.98	176.00	15.88	25.00	33.96	8.00	25.09	5.00	2.04	56.00	34.07	19.00	1.10	58.00
17/05/2015	0.75	283.00	12.99	681.00	4.98	176.00	15.17	17.00	34.09	11.00	25.09	5.00	1.96	54.00	34.17	21.00	1.06	54.00
18/05/2015	0.71	254.32	12.76	568.46	4.82	154.80	15.86	22.45	33.98	9.00	25.09	4.71	2.10	67.00	34.15	19.97	1.04	52.00
19/05/2015	0.66	249.00	12.97	673.43	4.94	172.30	15.85	22.11	33.95	7.87	25.09	4.65	2.02	61.00	34.12	20.84	0.94	48.00
20/05/2015	0.68	238.74	12.95	677.65	4.98	176.00	15.85	21.51	33.94	7.65	25.08	4.11	1.95	58.00	34.10	20.57	1.00	55.00
21/05/2015	0.71	253.84	12.95	675.54	5.00	182.00	15.84	21.83	33.93	7.24	25.08	4.05	1.94	57.00	34.11	0.00	1.11	0.00
22/05/2015	0.72	255.65	12.97	678.86	5.03	184.70	15.84	20.94	33.92	7.01	25.08	3.74	1.96	58.00	34.11	23.11	1.04	59.00
23/05/2015	0.72	255.00	12.96	676.00	5.05	185.00	15.84	21.00	33.91	6.00	25.08	4.00	1.92	54.00	34.10	22.00	1.06	61.00
24/05/2015	0.70	253.00	12.98	684.00	5.08	188.00	15.83	20.00	33.97	10.00	25.09	5.00	1.91	52.00	34.08	20.00	1.03	60.00
25/05/2015	0.72	255.00	12.72	599.15	4.90	161.60	15.88	22.59	34.02	8.63	25.08	3.79	1.92	53.00	34.31	36.24	1.02	59.00
26/05/2015	0.60	249.00	12.95	678.36	5.04	181.30	16.00	31.50	34.02	9.29	25.09	4.81	2.06	75.00	34.54	71.66	0.92	47.00
27/05/2015	0.76	276.38	12.95	674.76	5.06	190.90	15.90	24.37	33.91	7.03	25.08	3.93	2.36	82.00	34.28	37.28	1.12	62.00
28/05/2015	0.81	281.46	12.95	679.89	5.07	186.50	15.87	23.18	33.86	6.21	25.08	3.72	2.26	75.00	34.29	43.89	1.15	56.00
29/05/2015	0.79	277.00	12.95	679.00	5.06	188.10	15.84	21.78	33.81	5.66	25.08	3.59	2.18	61.00	34.17	27.81	1.12	55.00
30/05/2015	0.77	275.00	12.95	679.00	5.07	189.00	15.84	22.00	33.78	6.00	25.07	4.00	2.06	58.00	34.49	36.00	1.10	53.00
31/05/2015	0.72	270.00	13.00	699.00	5.08	188.00	16.35	34.00	34.10	14.00	25.09	5.00	2.12	59.00	34.31	31.00	1.04	47.00
01/06/2015	0.95	277.00	12.81	639.52	5.06	189.60	16.28	42.19	34.21	17.00	25.20	17.31	2.47	84.00	34.39	53.43	1.30	69.00
02/06/2015	0.94	276.00	12.98	680.09	5.08	191.40	16.15	40.47	33.97	8.15	25.12	5.88	2.69	95.00	34.26	35.00	1.32	73.00
03/06/2015	1.02	319.36	12.99	682.86	5.10	195.60	16.09	36.13	34.45	26.58	25.26	16.95	2.43	68.00	34.34	37.00	1.39	84.00
04/06/2015	1.04	321.50	13.00	685.29	5.13	202.10	16.51	55.67	34.25	20.12	25.25	18.44	2.62	59.97	34.33	43.06	1.40	85.00
05/06/2015	1.14	327.08	13.04	691.72	5.11	195.10	16.05	31.79	34.17	14.18	25.14	8.65	2.90	84.14	34.73	96.01	1.54	109.00
06/06/2015	1.19	332.00	13.06	700.00	5.16	197.00	16.73	49.00	34.27	16.00	25.18	10.00	3.09	89.00	34.66	94.00	1.67	112.00
07/06/2015	1.32	319.00	13.01	680.00	5.18	199.00	16.24	37.00	34.07	11.00	25.16	9.00	3.55	100.00	34.50	78.00	1.91	136.00
08/06/2015	1.26	314.00	12.76	570.00	5.01	185.20	16.01	31.74	34.00	8.65	25.14	10.12	3.12	71.18	34.38	43.67	1.85	129.00
09/06/2015	1.24	309.00	12.99	687.29	5.13	192.90	16.81	89.80	34.72	43.97	25.27	21.12	2.78	59.93	34.58	64.86	1.64	108.00
10/06/2015	1.50	444.12	13.01	698.15	5.20	213.70	17.36	156.72	34.93	60.05	25.47	59.62	4.24	118.41	34.76	100.24	2.02	146.00
11/06/2015	2.02	554.61	13.05	722.46	5.22	212.50	17.28	150.71	34.59	30.48	25.60	74.40	4.96	212.79	35.20	208.65	2.70	214.00
12/06/2015	2.35	587.00	13.03	708.31	5.28	225.40	17.61	211.88	34.53	27.64	25.32	25.63	5.60	272.17	35.02	166.79	3.14	238.00
13/06/2015	2.51	603.00	13.00	696.00	5.23	223.00	17.69	214.00	34.33	23.00	25.24	27.00	5.70	274.00	35.09	168.00	3.32	256.00
14/06/2015	2.45	597.00	12.99	692.00	5.22	222.00	17.10	199.00	34.20	20.00	25.19	25.00	5.16	260.00	34.80	161.00	3.21	245.00
15/06/2015	2.26	552.50	12.75	622.19	5.06	194.50	16.68	85.13	34.12	12.85	25.17	12.57	4.42	179.72	34.58	72.64	2.92	208.00
16/06/2015	2.03	531.65	12.96	676.37	5.19	208.50	16.38	50.86	34.12	12.00	25.16	11.52	3.84	137.09	34.64	91.42	2.68	193.68
17/06/2015	1.97	522.04	12.95	674.00	5.19	207.80	17.32	150.00	34.59	27.00	25.15	11.00	3.75	131.60	34.96	159.97	2.53	176.00
18/06/2015	1.99	523.64	12.98	686.00	5.16	204.30	17.06	60.89	34.22	17.00	25.18	13.00	4.15	143.74	34.76	109.00	2.60	185.00

Figure B-4: Raw observed discharge at selected gauging stations for diverted flow calculations and nested catchment modelling

Estaciones	PORVENIR		SAMARIA		GONZALEZ		TAPIJULAPA		TEAPA		PUYACATENGO		Pueblo Nuevo		OXOLOTAN		Gaviotas	
	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge
18/06/2015	1.99	523.64	12.98	686.00	5.16	204.30	17.06	60.89	34.22	17.00	25.18	13.00	4.15	143.74	34.76	109.00	2.60	185.00
19/06/2015	1.99	523.00	12.97	684.00	5.21	213.10	17.70	219.77	34.30	20.00	25.15	12.00	3.87	132.84	35.16	215.00	2.55	179.00
20/06/2015	1.96	520.00	12.94	672.00	5.20	213.00	16.88	198.00	34.14	16.00	25.14	12.00	4.21	140.00	34.69	203.00	2.56	179.00
21/06/2015	1.81	505.00	12.95	676.00	5.19	212.00	16.43	186.00	34.07	14.00	25.12	11.00	3.64	126.00	34.53	187.00	2.40	163.00
22/06/2015	1.67	491.00	12.70	551.00	5.02	189.49	16.41	51.59	34.03	11.00	25.12	7.04	3.16	73.53	34.44	56.86	2.12	135.00
23/06/2015	1.42	387.04	12.94	673.26	5.16	211.60	16.32	52.17	33.99	8.41	25.11	5.14	2.88	58.00	34.37	54.73	1.84	87.00
24/06/2015	1.36	350.33	12.93	671.37	5.16	210.80	16.26	44.20	33.97	7.00	25.11	5.01	2.69	50.00	34.32	41.22	1.77	75.00
25/06/2015	1.28	338.41	12.94	670.67	5.18	209.90	16.22	42.52	33.98	7.00	25.10	5.78	2.60	41.00	34.29	38.00	1.66	64.00
26/06/2015	1.23	335.00	12.92	672.86	5.15	209.00	16.19	40.35	33.96	6.00	25.10	5.51	2.50	39.00	34.30	39.25	1.62	60.00
27/06/2015	1.15	327.00	12.92	672.00	5.14	208.00	16.15	38.00	33.93	5.00	25.10	6.00	2.44	38.00	34.28	38.00	1.52	51.00
28/06/2015	1.04	300.00	12.90	667.00	5.12	207.00	16.10	37.00	33.93	5.00	25.10	6.00	2.35	37.00	34.25	36.00	1.44	43.00
29/06/2015	1.00	292.00	12.69	602.36	4.96	175.30	16.07	39.06	33.92	5.00	25.10	5.29	2.27	33.00	34.22	15.65	1.34	38.00
30/06/2015	0.86	278.00	12.95	675.56	5.16	206.40	16.06	38.98	34.21	15.00	25.11	5.16	2.21	31.00	34.24	15.77	1.22	36.00
01/07/2015	0.97	316.85	12.95	679.86	5.12	200.10	16.05	37.24	34.19	13.00	25.10	4.75	2.42	42.00	34.42	24.87	1.32	51.00
02/07/2015	1.02	318.14	12.95	679.00	5.10	202.30	16.04	36.10	34.00	9.13	25.09	4.95	2.54	54.00	34.26	18.75	1.38	61.00
03/07/2015	1.04	320.39	12.95	673.24	5.10	200.15	16.04	35.78	33.97	8.20	25.09	4.78	2.40	40.00	34.22	29.71	1.40	63.00
04/07/2015	0.99	315.00	12.98	685.00	5.10	200.00	16.03	35.00	34.11	11.00	25.14	6.00	2.26	36.00	34.21	29.00	1.36	62.00
05/07/2015	0.92	308.00	12.20	373.00	4.54	186.00	16.03	35.00	34.05	9.00	25.10	5.00	2.30	37.00	34.18	28.00	1.30	56.00
06/07/2015	0.73	256.93	12.16	386.00	4.51	124.50	16.01	24.66	33.98	7.00	25.08	4.45	2.16	75.00	34.18	26.54	1.06	69.00
07/07/2015	0.62	240.79	12.78	652.10	4.59	128.60	16.00	28.90	33.97	7.00	25.08	4.26	2.09	39.42	34.50	61.78	0.95	58.00
08/07/2015	0.67	263.29	12.78	652.00	4.55	130.80	16.13	30.28	33.96	7.00	25.08	4.03	2.14	39.57	34.32	34.70	1.02	69.00
09/07/2015	0.72	268.06	12.79	653.00	4.56	129.20	16.05	30.32	33.96	7.00	25.08	3.95	2.23	45.72	34.18	34.18	1.06	73.00
10/07/2015	0.68	238.43	12.79	653.00	4.56	130.40	16.01	28.12	33.95	7.56	25.08	4.07	2.06	39.94	34.22	31.27	1.02	71.00
11/07/2015	0.69	240.00	12.77	645.00	4.56	130.00	16.15	31.00	33.95	8.00	25.12	8.00	2.10	40.00	34.37	35.00	1.02	71.00
12/07/2015	0.66	237.00	12.20	417.00	4.49	128.00	16.00	27.00	33.96	8.00	25.09	6.00	2.28	44.00	34.32	33.00	1.04	72.00
13/07/2015	0.62	233.00	12.15	569.33	4.48	121.20	15.99	24.80	33.93	7.29	25.08	4.57	2.30	49.55	34.26	33.17	0.96	64.00
14/07/2015	0.62	233.00	12.80	615.16	4.58	134.30	16.14	16.13	34.00	10.00	25.10	6.47	2.19	47.80	34.43	56.27	0.96	64.00
15/07/2015	0.69	238.00	12.79	618.37	4.53	125.90	16.09	32.24	33.94	8.00	25.08	4.87	2.22	47.68	34.44	56.78	1.03	68.00
16/07/2015	0.76	250.75	12.81	608.42	4.53	126.20	16.03	30.41	33.92	6.98	25.08	4.56	2.40	53.97	34.62	90.16	1.19	82.95
17/07/2015	0.76	244.77	12.81	608.00	4.52	125.70	16.02	30.01	33.94	7.18	25.14	9.19	2.46	54.66	34.98	156.99	1.12	75.00
18/07/2015	0.84	253.00	12.31	407.00	4.43	122.00	16.67	46.00	33.91	6.00	25.09	7.00	3.26	77.00	34.59	118.00	1.26	89.00
19/07/2015	0.88	257.00	12.14	339.00	4.40	119.00	16.05	30.00	33.90	5.00	25.08	6.00	3.06	72.00	34.72	121.00	1.29	92.00
20/07/2015	0.72	241.00	12.10	538.88	4.39	112.30	17.11	13.63	33.93	9.00	25.07	3.90	2.82	70.29	34.86	144.05	1.10	127.58
21/07/2015	0.66	235.00	12.30	575.11	4.25	91.16	16.66	84.27	33.96	7.89	25.07	3.84	3.20	90.85	34.82	133.28	1.08	125.00

Observed Q data | CONAGUA Cuneo | correlations | 30002 GAVIOTAS | 30108 EL PORVENIR | 30003 BIGLIA | 30004 EL DORADO | 30005 SAMARIA | 30012 EL M...

Figure B-5: Raw observed discharge at selected gauging stations for diverted flow calculations and nested catchment modelling

Estaciones	PORVENIR		SAMARIA		GONZALEZ		TAPIJULAPA		TEAPA		PUYACATENGO		Pueblo Nuevo		OXOLOTAN		Gaviotas	
	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge
21/07/2015	0.66	235.00	12.30	575.11	4.25	91.16	16.66	84.27	33.96	7.89	25.07	3.84	3.20	90.85	34.82	133.28	1.08	125.00
22/07/2015	0.84	256.00	12.28	569.58	4.40	113.00	16.62	80.53	33.94	7.26	25.07	3.41	3.29	100.00	34.62	90.37	1.26	132.61
23/07/2015	0.84	238.51	12.29	571.26	4.41	115.00	16.39	51.79	33.92	7.00	25.07	3.15	2.88	80.00	34.45	67.01	1.20	72.79
24/07/2015	0.68	221.59	12.29	571.35	4.40	113.00	16.32	49.20	33.91	6.94	25.06	3.07	2.51	76.00	34.42	56.96	1.04	69.00
25/07/2015	0.58	212.00	12.29	571.00	4.41	115.00	16.34	51.00	34.23	15.00	25.06	3.00	2.32	71.00	34.39	56.00	0.90	55.00
26/07/2015	0.53	211.00	12.12	503.00	4.44	118.00	16.31	50.00	34.00	10.00	25.07	4.00	2.46	74.00	34.38	55.00	0.89	54.00
27/07/2015	0.56	214.00	12.06	524.99	4.38	113.00	16.28	46.20	33.97	7.48	25.07	3.52	2.41	68.00	34.30	39.15	0.91	56.00
28/07/2015	0.40	193.34	12.07	522.36	4.19	89.80	16.22	40.23	33.96	7.32	25.07	3.12	2.22	49.00	34.26	37.09	0.74	45.34
29/07/2015	0.36	183.28	12.09	523.68	4.30	103.60	16.12	36.10	33.94	7.24	25.07	3.55	2.09	43.00	34.22	32.96	0.68	45.58
30/07/2015	0.32	171.20	12.07	525.34	4.29	101.31	16.09	33.78	33.93	7.00	25.06	3.31	2.01	39.00	34.22	32.00	0.64	44.92
31/07/2015	0.29	169.00	12.08	527.16	4.28	102.90	16.07	32.47	33.93	7.00	25.08	4.06	1.93	37.00	34.20	29.78	0.60	39.72
01/08/2015	0.28	168.00	12.08	527.00	4.28	102.00	16.06	32.00	33.92	7.00	25.06	4.00	1.92	36.00	34.29	32.00	0.61	39.00
02/08/2015	0.28	168.00	12.09	531.00	4.28	102.00	16.05	31.00	33.91	7.00	25.06	4.00	1.93	37.00	34.22	30.00	0.61	39.00
03/08/2015	0.31	175.00	12.09	525.25	4.28	95.65	16.03	31.25	33.91	6.94	25.05	2.96	1.92	36.00	34.22	30.00	0.64	39.72
04/08/2015	0.24	165.42	12.09	522.24	4.12	74.80	16.02	30.78	33.90	6.79	25.05	2.78	1.86	31.00	34.38	50.00	0.56	43.27
05/08/2015	0.26	168.20	12.11	529.35	4.27	94.64	16.00	28.35	33.89	6.12	25.05	2.89	1.92	37.00	34.28	27.00	0.58	43.15
06/08/2015	0.26	167.14	12.10	525.80	4.26	94.48	15.97	27.60	33.87	5.91	25.05	2.82	1.91	36.00	34.22	21.00	0.59	42.20
07/08/2015	0.27	166.09	12.11	523.81	4.25	96.10	15.92	24.12	33.87	5.00	25.05	2.60	1.86	31.00	34.16	27.49	0.56	43.27
08/08/2015	0.25	164.00	12.13	532.00	4.26	96.00	15.90	24.00	33.86	5.00	25.05	3.00	1.79	29.00	34.16	27.00	0.55	43.00
09/08/2015	0.20	159.00	12.14	536.00	4.25	95.00	15.89	24.00	33.86	5.00	25.05	3.00	1.75	28.00	34.15	26.00	0.52	42.00
10/08/2015	0.22	162.98	12.15	538.00	4.24	92.20	15.86	19.99	33.85	5.72	25.04	2.40	1.73	27.00	34.14	29.18	0.53	43.00
11/08/2015	0.19	159.83	12.15	538.00	4.18	93.30	16.06	32.33	34.21	18.94	25.05	2.66	1.76	29.00	34.24	34.13	0.49	57.56
12/08/2015	0.22	162.65	12.19	554.00	4.24	95.80	16.02	29.94	34.00	14.00	25.05	2.53	2.06	36.00	34.27	37.89	0.55	42.87
13/08/2015	0.28	166.82	12.22	568.71	4.32	103.80	16.13	37.61	33.92	6.97	25.06	3.20	2.02	34.00	34.32	43.45	0.60	42.97
14/08/2015	0.29	168.11	12.19	556.38	4.26	94.74	15.98	24.80	34.26	17.23	25.13	7.98	2.14	42.00	34.24	38.40	0.62	44.00
15/08/2015	0.29	168.00	12.21	564.00	4.26	95.00	15.90	23.00	33.98	11.00	25.14	8.00	2.25	45.00	34.34	40.00	0.65	47.00
16/08/2015	0.36	161.00	12.21	564.00	4.27	97.00	16.02	26.00	34.04	12.00	25.15	9.00	2.06	40.00	34.30	39.00	0.68	50.00
17/08/2015	0.38	163.00	12.24	573.33	4.28	95.70	16.07	33.73	34.06	11.43	25.13	10.25	2.16	50.00	34.48	71.68	0.72	43.50
18/08/2015	0.36	161.00	12.23	566.86	4.13	77.24	15.99	28.50	34.14	14.08	25.06	3.35	2.38	76.00	34.46	91.13	0.72	45.34
19/08/2015	0.44	194.94	12.12	529.32	4.22	93.73	15.92	25.34	33.93	7.68	25.05	3.27	2.46	84.00	34.38	63.29	0.81	55.00
20/08/2015	0.40	192.00	12.08	500.91	4.18	87.67	15.88	20.29	33.90	7.04	25.05	3.05	2.30	75.00	34.28	49.23	0.73	54.90
21/08/2015	0.35	182.61	12.08	457.51	4.17	87.17	16.05	32.65	33.87	5.96	25.04	3.11	2.09	60.00	34.23	42.17	0.66	51.80
22/08/2015	0.27	175.00	12.09	461.00	4.17	87.00	16.10	34.00	33.90	7.00	25.04	3.00	1.96	57.00	34.38	46.00	0.56	42.00
23/08/2015	0.27	175.00	12.08	460.00	4.18	87.00	16.30	50.00	33.94	9.00	25.06	3.00	2.14	67.00	34.38	46.00	0.58	44.00

Figure B-6: Raw observed discharge at selected gauging stations for diverted flow calculations and nested catchment modelling



Estaciones	PORVENIR		SAMARIA		GONZALEZ		TAPIJULAPA		TEAPA		PUYACATENGO		Pueblo Nuevo		OXOLOTAN		Gaviotas	
	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge
23/08/2015	0.27	175.00	12.08	460.00	4.18	87.00	16.30	50.00	33.94	9.00	25.06	3.00	2.14	67.00	34.38	46.00	0.58	44.00
24/08/2015	0.30	178.00	12.06	456.00	4.16	91.63	16.01	28.50	33.87	5.92	25.05	3.44	2.26	73.00	34.23	38.86	0.64	51.18
25/08/2015	0.29	177.00	12.09	499.21	4.08	81.56	16.28	44.65	34.19	16.55	25.18	12.70	2.10	65.00	34.31	45.16	0.62	64.98
26/08/2015	0.34	173.80	12.09	499.00	4.21	96.02	17.32	173.47	34.16	16.23	25.10	6.32	2.44	81.00	34.96	141.00	0.66	71.70
27/08/2015	0.54	208.68	12.07	497.00	4.04	78.28	17.22	153.67	34.57	29.96	25.38	35.08	3.42	132.00	34.91	141.00	1.04	141.86
28/08/2015	1.00	291.98	12.16	532.02	4.08	81.50	17.02	128.89	34.78	43.56	25.33	31.42	4.22	152.00	35.03	192.01	1.67	229.00
29/08/2015	1.36	343.00	12.13	529.00	4.12	82.00	19.25	229.00	35.38	91.00	25.54	51.00	5.03	160.00	35.92	281.01	2.14	276.00
30/08/2015	2.00	407.00	12.34	613.00	4.20	84.00	17.79	192.00	34.56	70.00	25.28	44.00	5.83	180.00	35.16	262.00	3.01	363.00
31/08/2015	2.00	525.29	12.22	549.44	4.24	98.97	17.47	175.65	34.64	38.92	25.29	26.46	5.30	127.00	34.88	141.93	2.84	346.00
01/09/2015	1.89	509.64	12.24	561.41	4.25	101.50	17.23	154.60	34.50	27.64	25.27	19.41	4.83	80.00	34.81	134.00	2.67	329.00
02/09/2015	1.70	492.23	12.10	557.69	4.24	97.38	17.80	227.61	34.36	23.70	25.25	16.93	4.36	65.00	35.21	238.50	2.42	381.70
03/09/2015	1.61	467.29	12.02	542.76	4.13	86.26	17.74	215.53	34.31	20.00	25.25	15.59	4.60	89.00	35.21	237.70	2.41	378.00
04/09/2015	1.68	480.61	11.98	534.26	4.10	79.37	17.15	138.03	34.41	26.06	25.19	12.39	4.67	96.00	34.82	126.33	2.50	380.00
05/09/2015	1.49	462.00	12.02	550.00	4.08	77.00	17.25	140.00	34.37	25.00	25.22	13.00	4.05	80.00	34.77	124.00	2.19	349.00
06/09/2015	1.37	450.00	11.96	526.00	4.14	83.00	16.68	126.00	34.22	21.00	25.18	12.00	3.80	74.00	34.70	117.00	2.06	336.00
07/09/2015	1.31	438.00	11.96	513.15	4.08	81.61	16.57	76.52	34.16	15.59	25.15	9.03	3.60	71.00	34.70	108.05	1.94	356.97
08/09/2015	1.19	334.96	11.92	502.33	4.08	81.38	16.72	81.80	34.13	14.51	25.14	8.73	3.40	66.00	34.65	101.84	1.80	187.30
09/09/2015	1.10	316.58	11.90	491.36	4.06	80.43	16.52	67.93	34.09	13.38	25.12	7.35	3.22	61.00	34.59	91.20	1.68	192.42
10/09/2015	1.00	311.51	11.90	489.71	4.01	78.10	16.38	46.45	34.05	12.24	25.11	7.26	3.15	59.00	34.54	79.73	1.55	179.00
11/09/2015	0.88	285.35	11.90	491.23	3.98	72.77	16.49	66.80	34.03	10.12	25.11	6.56	2.84	51.00	34.52	82.21	1.43	120.15
12/09/2015	0.82	279.00	11.90	491.00	3.97	72.00	16.20	59.00	33.98	9.00	25.10	6.00	2.76	49.00	34.44	80.00	1.30	116.00
13/09/2015	0.74	271.00	11.90	491.00	4.09	84.00	16.18	58.00	34.01	10.00	25.10	6.00	2.64	43.00	34.50	86.00	1.23	109.00
14/09/2015	0.77	228.15	11.88	484.32	4.02	77.36	17.70	213.04	34.74	83.00	25.30	30.50	2.72	51.00	35.16	263.18	1.23	109.00
15/09/2015	1.27	339.34	11.99	516.96	4.08	89.39	19.66	612.50	35.03	112.00	25.34	34.02	5.46	325.00	36.56	1453.63	2.10	315.17
16/09/2015	2.12	560.31	11.94	509.65	4.10	91.81	18.00	285.30	34.61	25.74	25.22	14.57	5.86	396.00	35.35	293.13	3.15	713.60
17/09/2015	2.09	555.25	11.93	506.42	4.16	96.14	17.98	279.47	34.60	24.00	25.29	27.94	5.32	342.00	35.42	311.62	2.94	641.70
18/09/2015	2.03	541.19	11.98	530.46	4.13	89.47	17.28	149.79	34.42	24.39	25.20	15.48	5.30	340.00	34.96	162.52	2.91	548.85
19/09/2015	1.80	518.00	11.95	519.00	4.09	88.00	17.95	167.00	34.30	21.00	25.48	25.00	4.56	321.00	35.15	166.00	2.56	514.00
20/09/2015	1.91	529.00	11.97	527.00	4.10	89.00	18.15	172.00	34.47	25.00	25.34	21.00	5.18	336.00	35.32	170.00	2.68	526.00
21/09/2015	1.86	505.00	12.10	551.16	4.20	101.10	17.22	144.93	34.46	25.28	25.38	38.86	5.30	339.00	35.32	268.15	2.77	468.15
22/09/2015	2.00	536.58	12.05	546.96	4.19	100.30	19.83	613.39	35.52	116.03	25.54	67.36	5.72	357.00	36.38	1044.80	2.90	475.00
23/09/2015	2.48	576.83	12.11	554.12	4.18	103.40	20.28	946.73	35.98	206.86	25.76	106.06	6.83	564.00	36.82	1427.26	3.57	834.78
24/09/2015	2.72	654.50	12.15	560.55	4.24	108.50	19.10	502.50	35.46	112.90	25.56	76.25	6.76	557.00	35.94	587.44	3.76	849.00
25/09/2015	2.78	678.47	12.30	572.86	4.39	126.80	18.16	317.66	35.02	57.24	25.46	39.46	6.32	507.00	35.37	310.92	3.73	934.20

Figure B-7: Raw observed discharge at selected gauging stations for diverted flow calculations and nested catchment modelling

Estaciones	PORVENIR		SAMARIA		GONZALEZ		TAPIULAPA		TEAPA		PUYACATENGO		Pueblo Nuevo		OXOLOTAN		Gaviotas	
	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge
25/09/2015	2.78	678.47	12.30	572.86	4.39	126.80	18.16	317.66	35.02	57.24	25.46	39.46	6.32	507.00	35.37	310.92	3.73	934.20
26/09/2015	2.84	684.00	12.30	573.00	4.40	127.00	19.15	343.00	36.20	86.00	25.70	45.00	6.86	520.00	35.80	312.00	3.78	935.00
27/09/2015	2.94	694.00	12.98	707.00	4.42	129.00	17.35	300.00	35.28	63.00	25.48	40.00	6.46	480.00	35.24	298.00	3.90	947.00
28/09/2015	3.09	709.00	12.61	605.89	4.95	195.00	17.45	178.37	34.96	54.00	25.36	33.09	6.12	437.00	34.98	180.85	3.90	589.18
29/09/2015	2.87	184.86	12.14	554.36	4.34	118.80	17.32	152.43	34.71	37.00	25.30	28.28	5.87	363.00	35.06	212.43	3.74	566.92
30/09/2015	2.76	665.53	12.05	545.56	4.23	103.70	17.28	149.40	35.02	56.48	25.60	79.36	5.77	348.00	35.28	223.04	3.62	496.08
01/10/2015	2.98	731.39	12.00	538.42	4.22	100.50	19.84	575.40	35.96	179.56	25.72	92.63	6.72	681.00	36.26	950.88	3.92	737.70
02/10/2015	3.12	752.00	12.14	556.00	4.27	109.70	18.35	340.95	35.59	114.00	25.49	57.08	6.74	691.00	35.58	412.89	4.08	846.16
03/10/2015	3.14	750.00	11.90	460.00	4.21	104.00	17.96	302.00	35.17	103.00	25.43	56.00	6.38	655.00	35.27	397.00	4.06	840.00
04/10/2015	3.08	748.00	11.96	484.00	4.19	103.00	17.48	290.00	34.85	95.00	25.34	54.00	6.10	648.00	35.04	391.00	3.96	837.00
05/10/2015	2.75	715.00	11.90	496.02	4.11	90.78	17.10	138.95	34.67	28.00	25.28	20.81	5.85	623.00	34.90	154.08	3.87	685.16
06/10/2015	2.89	722.00	11.86	401.36	4.06	87.10	16.96	125.12	34.53	24.00	25.24	15.95	5.61	598.00	34.80	126.10	3.76	659.94
07/10/2015	2.79	672.66	11.83	395.00	4.03	83.97	16.80	105.51	34.45	24.29	25.22	14.70	5.37	564.00	34.72	111.46	3.66	623.53
08/10/2015	2.65	654.68	11.81	484.22	4.01	80.03	16.72	99.55	34.37	20.00	25.21	13.89	5.07	534.00	34.65	95.00	3.50	521.43
09/10/2015	2.54	599.68	11.80	482.00	3.98	78.50	16.60	93.56	34.32	17.00	25.20	13.07	4.78	495.00	34.62	88.99	3.46	497.68
10/10/2015	2.50	599.00	11.80	482.00	3.98	78.00	16.66	94.00	34.79	42.00	25.35	32.00	4.56	473.00	34.65	95.00	3.30	482.00
11/10/2015	2.55	601.00	12.08	594.00	4.10	81.00	17.72	120.00	35.22	53.00	25.86	45.00	5.07	486.00	34.86	100.00	3.43	485.00
12/10/2015	2.75	165.52	12.24	550.54	4.33	122.90	16.71	96.67	34.64	28.74	25.38	35.43	5.63	282.28	34.79	124.75	3.64	624.74
13/10/2015	2.64	652.87	12.11	538.32	4.38	129.40	16.55	79.01	34.39	22.17	25.28	21.20	5.16	235.90	34.56	87.71	3.50	576.43
14/10/2015	2.48	590.08	12.04	367.20	4.27	116.30	16.42	69.08	34.94	62.38	25.29	26.30	4.62	201.09	34.52	80.85	3.30	548.49
15/10/2015	2.35	552.50	12.00	359.23	4.24	109.40	17.08	147.19	34.80	42.00	25.56	71.83	4.56	191.17	34.75	118.05	3.16	475.14
16/10/2015	2.50	592.54	12.05	384.42	4.26	113.20	17.09	132.64	34.75	39.00	25.41	38.57	5.30	238.07	34.78	120.85	3.40	569.06
17/10/2015	2.57	599.00	12.04	384.00	4.28	113.00	17.32	147.00	34.75	39.00	25.44	41.00	5.55	258.00	34.95	138.00	3.51	580.00
18/10/2015	2.65	603.00	12.83	593.00	4.40	117.00	17.02	109.00	34.83	40.00	25.40	37.00	5.56	260.00	34.78	125.00	3.58	584.00
19/10/2015	2.88	690.61	12.52	440.22	4.69	161.90	19.21	493.62	35.87	91.00	25.74	99.01	6.15	289.24	36.05	669.49	3.80	583.64
20/10/2015	3.22	755.57	13.84	983.60	5.42	252.10	21.77	1321.81	36.42	209.54	25.89	114.60	6.78	385.57	37.64	2369.86	4.22	700.72
21/10/2015	3.89	875.22	15.11	1578.16	7.26	523.00	18.94	417.35	35.26	149.00	25.50	59.69	6.99	454.52	35.90	597.63	4.70	775.00
22/10/2015	4.16	909.64	13.98	1102.49	7.04	434.50	19.12	480.92	35.02	57.58	25.44	50.25	6.82	413.62	36.06	789.60	4.96	801.00
23/10/2015	4.02	841.30	13.14	723.71	5.86	285.20	18.72	355.72	34.79	49.16	25.36	33.44	6.71	396.57	35.71	440.96	4.86	847.50
24/10/2015	3.88	827.00	12.62	516.00	5.09	208.00	17.80	333.00	34.69	46.00	25.31	30.00	6.44	389.00	35.28	430.00	4.76	838.00
25/10/2015	3.82	821.00	12.37	503.00	4.83	173.00	17.45	298.00	34.55	31.00	25.28	28.00	6.21	421.00	35.10	412.00	4.70	832.00
26/10/2015	3.81	819.00	12.18	380.56	4.67	140.90	19.35	546.19	36.40	235.44	25.86	100.01	6.68	378.44	35.91	542.48	4.74	956.98
27/10/2015	3.90	822.09	12.09	365.29	4.62	133.90	18.12	258.96	35.27	69.37	25.44	47.15	6.62	385.49	35.40	316.44	4.86	1005.07
28/10/2015	3.91	818.41	12.01	362.26	4.58	125.25	17.47	197.00	34.93	55.39	25.35	32.39	6.29	340.99	35.08	200.71	4.85	801.29

Figure B-8: Raw observed discharge at selected gauging stations for diverted flow calculations and nested catchment modelling

Estaciones	PORVENIR		SAMARIA		GONZALEZ		TAPIJULAPA		TEAPA		PUYACATENGO		Pueblo Nuevo		OXOLOTAN		Gaviotas	
	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge
28/10/2015	3.91	818.41	12.01	362.26	4.58	125.25	17.47	197.00	34.93	55.39	25.35	32.39	6.29	340.99	35.08	200.71	4.85	801.29
29/10/2015	3.86	812.56	11.93	354.80	4.49	112.10	17.13	158.70	34.73	23.74	25.29	25.37	6.08	325.68	34.93	159.07	4.81	775.68
30/10/2015	3.78	801.21	11.90	356.76	4.40	109.50	16.85	119.12	34.60	20.00	25.27	21.27	5.86	299.04	34.82	148.00	4.71	750.70
31/10/2015	3.65	795.00	11.81	352.00	4.32	105.00	16.76	110.00	34.51	18.00	25.24	18.00	5.67	282.00	34.80	145.00	4.59	745.00
01/11/2015	3.48	778.00	11.76	332.00	4.27	103.00	16.73	108.00	34.43	17.00	25.22	17.00	5.53	278.00	34.71	142.00	4.43	729.00
02/11/2015	3.34	750.00	11.73	323.16	4.21	95.47	16.60	100.93	34.55	25.00	25.30	28.97	5.43	240.64	34.82	127.07	4.31	560.23
03/11/2015	3.26	772.84	11.79	315.59	4.24	103.20	16.67	107.10	34.43	21.00	25.24	16.09	5.43	241.29	34.66	100.77	4.22	475.14
04/11/2015	3.12	741.12	11.72	390.38	4.24	105.20	16.56	94.13	34.34	18.00	25.21	13.52	5.10	228.15	34.61	85.75	4.07	466.69
05/11/2015	2.96	714.17	11.71	302.12	4.24	104.00	16.52	82.82	34.29	16.00	25.19	12.60	4.79	204.09	34.56	82.35	3.90	459.74
06/11/2015	2.81	698.00	11.69	299.89	4.23	101.80	16.48	72.58	34.37	20.00	25.24	16.61	4.56	194.43	34.53	73.50	3.72	451.00
07/11/2015	2.76	693.00	11.70	300.00	4.24	102.00	16.40	71.00	34.28	18.00	25.20	16.00	4.54	194.00	34.76	79.00	3.64	449.00
08/11/2015	2.66	683.00	11.80	340.00	4.34	104.00	16.59	76.00	34.88	33.00	25.62	26.00	4.94	204.00	34.82	80.00	3.60	445.00
09/11/2015	2.85	681.64	11.89	342.15	4.42	125.30	16.68	108.12	34.69	20.00	25.50	57.60	5.42	243.30	34.60	60.00	3.84	469.00
10/11/2015	2.80	172.68	11.95	350.00	4.45	129.20	16.55	83.02	34.44	17.00	25.34	33.45	5.12	226.12	34.70	93.61	3.76	461.00
11/11/2015	2.64	654.68	11.77	328.87	4.40	123.10	16.52	80.45	34.31	15.00	25.28	21.59	4.69	201.27	34.54	78.63	3.56	536.16
12/11/2015	2.54	639.00	11.72	308.99	4.40	120.50	16.40	68.25	34.26	12.00	25.25	18.83	4.32	164.00	34.48	77.32	3.38	532.70
13/11/2015	2.39	537.26	11.67	288.22	4.40	123.20	16.68	109.11	34.78	64.00	25.40	38.86	4.13	145.00	34.49	79.00	3.24	514.14
14/11/2015	2.45	538.00	11.67	408.00	4.43	124.00	17.90	139.00	35.08	71.00	25.65	45.00	4.94	165.00	35.00	92.00	3.33	516.00
15/11/2015	2.89	549.00	12.78	852.00	5.10	140.00	17.58	131.00	34.75	63.00	25.42	39.00	5.96	190.00	35.13	95.00	3.91	530.00
16/11/2015	3.10	715.21	12.48	409.32	5.30	224.90	17.34	176.77	34.62	55.00	25.35	34.71	5.80	289.01	35.00	92.00	4.04	722.79
17/11/2015	3.00	712.19	12.01	364.26	4.74	151.60	17.13	155.75	34.50	51.00	25.30	25.38	5.66	264.48	34.90	89.00	3.96	647.40
18/11/2015	2.90	701.77	11.93	348.11	4.65	142.80	16.92	115.74	34.41	42.00	25.27	22.34	5.39	242.63	34.75	86.00	3.85	630.92
19/11/2015	2.79	692.50	11.86	337.99	4.60	138.20	16.74	110.08	34.47	44.00	25.25	19.56	4.97	224.16	34.89	144.48	3.70	633.04
20/11/2015	2.76	688.00	11.81	329.32	4.55	134.10	16.80	107.46	34.33	31.00	25.28	22.92	5.30	240.79	34.70	110.35	3.70	618.43
21/11/2015	2.86	690.00	11.80	325.00	4.53	131.00	17.29	119.00	34.27	29.00	25.25	21.00	5.28	239.00	35.05	145.00	3.80	619.00
22/11/2015	2.83	687.00	11.76	278.00	4.48	130.00	16.85	74.00	34.42	33.00	25.24	21.00	5.40	242.00	34.84	139.00	3.80	619.00
23/11/2015	2.87	698.81	12.07	373.26	4.64	143.20	19.87	695.65	37.00	400.00	26.20	210.93	6.00	269.38	36.11	709.33	3.78	550.00
24/11/2015	3.20	752.24	12.83	436.86	4.85	162.40	18.78	334.01	35.43	361.00	25.69	87.45	6.68	402.11	35.60	397.83	4.27	694.35
25/11/2015	3.52	783.03	13.50	852.90	6.05	314.30	17.88	254.60	34.93	315.00	25.41	40.60	6.33	351.85	35.22	233.02	4.42	694.34
26/11/2015	3.54	798.16	12.68	479.30	5.44	230.20	17.44	187.33	34.67	289.00	25.34	34.12	6.10	317.35	34.98	158.77	4.45	645.55
27/11/2015	3.40	773.83	12.25	448.08	5.10	198.80	16.95	117.06	34.54	276.00	25.29	22.82	5.86	297.60	34.82	137.63	4.34	669.97
28/11/2015	3.28	767.00	12.04	374.00	4.82	184.00	16.60	82.00	34.45	267.00	25.26	20.00	5.66	280.00	34.75	130.00	4.22	657.00
29/11/2015	3.18	764.00	12.00	358.00	4.70	181.00	16.47	79.00	34.39	265.00	25.25	19.00	5.47	275.00	34.78	131.00	4.12	654.00
30/11/2015	3.07	706.31	11.92	371.10	4.58	135.80	16.55	81.72	34.31	46.00	25.22	14.14	5.21	235.73	34.63	96.79	3.99	549.72

Figure B-9: Raw observed discharge at selected gauging stations for diverted flow calculations and nested catchment modelling

Estaciones	PORVENIR		SAMARIA		GONZALEZ		TAPIJULAPA		TEAPA		PUYACATENGO		Pueblo Nuevo		OXOLOTAN		Gaviotas	
	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge
30/11/2015	3.07	706.31	11.92	371.10	4.58	135.80	16.55	81.72	34.31	46.00	25.22	14.14	5.21	235.73	34.63	96.79	3.99	549.72
01/12/2015	2.96	698.81	11.82	333.52	4.50	131.30	16.49	80.01	34.25	41.00	25.21	13.88	4.92	219.70	34.60	90.50	3.86	545.77
02/12/2015	2.84	686.00	11.78	329.16	4.44	122.50	16.45	71.17	34.21	38.00	25.20	13.18	4.67	196.43	34.55	83.12	3.73	526.58
03/12/2015	2.75	679.69	11.74	324.96	4.40	119.50	16.44	73.41	35.30	83.00	25.44	39.70	4.46	185.36	34.54	81.04	3.60	512.86
04/12/2015	2.84	689.00	12.88	780.00	4.66	144.10	16.88	109.58	34.69	39.00	25.31	32.67	5.04	224.32	34.73	99.00	3.86	539.00
05/12/2015	3.19	724.00	12.56	652.00	5.58	225.00	17.16	116.00	34.82	42.00	25.36	37.00	5.28	236.00	34.86	118.00	4.07	544.00
06/12/2015	3.23	728.00	12.88	780.00	5.54	221.00	17.65	128.00	35.10	49.00	25.41	39.00	5.58	243.00	35.10	142.00	4.20	557.00
07/12/2015	3.34	740.00	12.44	400.60	5.36	219.30	17.76	239.82	35.62	118.00	25.64	78.89	5.84	276.93	35.27	159.00	4.32	646.30
08/12/2015	3.50	828.36	12.44	392.70	5.16	197.60	19.40	642.92	36.20	266.00	25.86	97.65	6.68	373.42	36.13	245.00	4.58	735.43
09/12/2015	3.68	844.47	12.94	468.08	5.61	255.50	18.07	298.56	35.30	171.00	25.48	55.09	6.51	367.82	35.36	170.00	4.72	748.51
10/12/2015	3.80	862.00	12.68	428.69	5.85	282.10	17.30	170.81	34.91	132.00	25.36	35.87	6.18	331.44	35.01	135.00	4.76	769.63
11/12/2015	3.74	850.00	12.38	363.38	5.42	231.20	16.97	116.05	34.71	112.00	25.32	33.36	5.96	298.08	34.84	152.00	4.72	740.31
12/12/2015	3.62	838.00	12.12	259.00	4.89	217.00	16.80	111.00	34.59	109.00	25.28	32.00	5.74	292.00	34.75	143.00	4.63	731.00
13/12/2015	3.54	830.00	12.00	211.00	4.77	205.00	16.72	109.00	34.48	106.00	25.25	31.00	5.56	287.00	34.67	135.00	4.56	724.00
14/12/2015	3.44	773.83	11.96	339.66	4.66	140.00	16.67	107.41	34.45	97.00	25.22	14.67	5.40	241.99	34.62	92.71	4.49	602.69
15/12/2015	3.25	753.00	11.86	333.12	4.59	135.30	16.50	77.78	34.40	83.00	25.21	14.01	5.18	232.92	34.57	86.24	4.38	699.31
16/12/2015	3.23	759.30	11.80	330.13	4.51	129.00	16.44	73.10	34.34	71.00	25.21	13.75	4.93	221.18	34.53	77.32	4.26	701.44
17/12/2015	3.10	719.92	11.63	301.69	4.38	118.20	16.40	71.16	34.27	57.00	25.20	13.31	4.73	201.32	34.50	72.33	4.10	641.16
18/12/2015	2.95	703.21	11.62	293.45	4.28	110.30	16.38	70.33	34.30	58.00	25.45	41.30	4.54	190.03	34.87	140.18	3.92	626.18
19/12/2015	3.34	742.00	11.70	325.00	4.42	124.00	22.05	212.00	36.65	254.00	26.02	56.00	6.92	249.00	37.74	212.00	4.58	692.00
20/12/2015	3.58	766.00	13.60	773.00	5.40	250.00	19.95	200.00	36.25	229.00	25.91	88.00	7.00	599.00	36.57	199.00	4.76	710.00
21/12/2015	4.20	919.24	14.10	1020.16	7.02	442.50	18.87	344.01	35.27	131.00	25.51	61.71	7.12	473.98	35.76	527.24	5.14	800.60
22/12/2015	4.28	959.31	12.98	483.15	5.94	284.30	17.88	224.00	34.93	122.00	25.39	36.32	6.80	417.05	35.32	281.42	5.18	805.00
23/12/2015	4.20	922.06	12.32	452.28	5.33	195.50	17.55	211.37	34.73	117.00	25.32	31.98	6.52	371.51	35.10	204.07	5.10	878.41
24/12/2015	4.06	889.74	11.94	379.22	4.95	153.80	17.38	180.70	34.62	114.00	25.29	30.68	6.27	343.42	34.96	170.60	4.99	867.00
25/12/2015	3.96	872.69	11.88	341.99	4.80	142.60	17.20	150.15	34.55	110.00	25.27	30.00	6.05	337.00	34.86	141.58	4.90	858.00
26/12/2015	3.84	860.00	11.76	294.00	4.64	135.00	17.00	145.00	34.46	107.00	25.25	27.00	5.86	332.00	34.78	133.00	4.82	850.00
27/12/2015	3.73	849.00	11.69	266.00	4.55	126.00	16.90	142.00	34.39	105.00	25.23	25.00	5.70	328.00	34.74	129.00	4.72	840.00
28/12/2015	3.60	825.88	11.63	297.45	4.46	108.70	16.75	125.40	34.36	93.00	25.22	14.30	5.56	266.65	34.69	105.03	4.60	726.84
29/12/2015	3.52	809.00	11.58	289.60	4.38	109.10	16.68	108.32	34.29	79.00	25.21	13.60	5.41	245.13	34.64	100.59	4.47	708.00
30/12/2015	3.29	730.68	11.56	283.60	4.32	108.30	16.65	101.72	34.26	71.00	25.20	13.25	5.20	228.97	34.61	95.84	4.30	684.00
31/12/2015	3.17	722.00	11.57	286.98	4.32	108.40	16.58	95.36	34.23	62.00	25.19	12.99	4.93	222.46	34.57	89.59	4.12	661.00
01/01/2016	3.18	723.00	11.63	298.36	4.32	108.00	16.54	90.18	34.24	62.00	25.18	12.00	4.73	202.00	34.55	92.71	3.92	631.00
02/01/2016	2.99	704.00	11.76	350.00	4.52	128.00	17.03	102.00	35.19	85.00	25.68	24.00	4.76	203.00	34.59	96.00	3.86	625.00

Figure B-10: Raw observed discharge at selected gauging stations for diverted flow calculations and nested catchment modelling

Estaciones	PORVENIR		SAMARIA		GONZALEZ		TAPIJULAPA		TEAPA		PUYACATENGO		Pueblo Nuevo		OXOLOTAN		Gaviotas	
	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge	Level	Discharge
02/01/2016	2.99	704.00	11.76	350.00	4.52	128.00	17.03	102.00	35.19	85.00	25.68	24.00	4.76	203.00	34.59	96.00	3.86	625.00
03/01/2016	3.17	722.00	12.45	520.00	4.70	142.00	16.66	90.00	34.81	70.00	25.41	21.00	5.68	284.00	34.54	90.00	4.19	358.00
04/01/2016	3.40	741.62	12.54	419.60	5.65	256.50	16.70	111.57	35.27	95.00	25.45	50.57	5.51	258.67	34.53	86.20	4.30	480.00
05/01/2016	3.40	749.52	12.72	433.39	5.48	240.10	17.15	165.80	35.22	90.00	25.53	64.65	5.72	264.49	34.81	133.20	4.36	509.00
06/01/2016	3.54	781.26	13.14	473.00	5.90	291.40	16.84	124.03	34.81	49.00	25.39	78.13	5.81	277.43	34.66	102.26	4.46	519.00
07/01/2016	3.52	775.20	12.40	412.00	5.42	227.10	16.67	107.41	34.61	34.00	25.32	31.33	5.59	262.68	34.54	88.41	4.43	516.00
08/01/2016	3.30	737.91	12.15	383.00	4.86	164.90	16.64	100.09	34.48	25.00	25.28	28.39	5.38	238.48	34.48	78.18	4.28	500.00
09/01/2016	3.15	722.00	11.98	315.00	4.68	160.00	16.60	99.00	34.41	23.00	25.25	28.00	5.00	228.00	34.45	75.00	4.09	481.00
10/01/2016	3.03	710.00	11.84	259.00	4.56	157.00	16.58	98.00	34.35	21.00	25.24	27.00	4.76	222.00	34.45	75.00	3.92	464.00
11/01/2016	2.98	702.00	11.78	345.90	4.45	110.10	16.66	101.69	34.59	33.00	25.34	33.85	4.82	196.05	34.54	74.39	3.82	454.00
12/01/2016	2.91	692.46	11.74	314.77	4.43	111.40	16.48	85.92	34.46	26.00	25.28	29.89	4.94	221.05	34.46	69.88	3.87	447.00
13/01/2016	2.82	670.11	11.70	301.40	4.38	105.30	16.40	78.37	34.44	25.00	25.28	27.55	4.65	197.51	34.42	64.62	3.73	433.00
14/01/2016	2.72	649.11	11.71	303.89	4.37	103.40	16.37	78.15	34.47	25.00	25.28	27.06	4.51	187.77	34.40	59.99	3.61	421.00
15/01/2016	2.64	631.67	11.67	296.90	4.37	102.10	16.36	77.51	34.41	22.00	25.25	20.76	4.41	178.48	34.38	58.50	3.52	419.00
16/01/2016	2.55	622.00	11.67	296.00	4.35	101.00	16.36	77.00	34.34	19.00	25.23	20.00	4.23	173.00	34.30	50.00	3.37	404.00
17/01/2016	2.44	611.00	11.76	332.00	4.40	102.00	16.35	77.00	34.29	17.00	25.21	18.00	4.07	169.00	34.35	55.00	3.25	392.00
18/01/2016	2.34	550.49	11.72	300.33	4.38	102.80	16.34	79.02	34.23	16.00	25.19	13.58	3.91	141.59	34.34	53.94	3.13	380.00
19/01/2016	2.28	545.00	11.70	298.20	4.36	98.41	16.34	78.14	34.18	15.00	25.18	13.00	3.77	134.53	34.32	48.27	2.97	368.00
20/01/2016	2.12	533.93	11.65	292.08	4.34	95.15	16.33	72.11	34.17	14.00	25.18	13.00	3.63	129.73	34.32	46.30	2.82	358.00
21/01/2016	1.99	500.85	11.65	295.66	4.32	98.10	16.33	68.45	34.15	13.00	25.18	13.00	3.52	117.34	34.31	44.45	2.70	254.76
22/01/2016	1.88	490.14	11.65	296.21	4.30	96.40	16.32	65.47	34.13	13.00	25.16	11.67	3.41	113.20	34.30	44.04	2.53	250.00
23/01/2016	1.72	474.00	11.64	292.00	4.29	95.00	16.40	67.00	34.41	20.00	25.16	12.00	3.31	103.00	34.44	46.00	2.40	237.00
24/01/2016	1.60	462.00	11.65	296.00	4.29	95.00	16.36	66.00	34.26	15.00	25.16	12.00	3.64	111.00	34.40	42.00	2.36	233.00
25/01/2016	1.62	464.00	11.65	294.35	4.27	99.10	16.35	73.70	34.17	13.00	25.15	11.27	3.48	114.38	34.34	51.97	2.24	221.00
26/01/2016	1.48	459.39	11.64	296.28	4.26	98.85	16.34	61.24	34.13	12.00	25.15	10.77	3.23	101.14	34.31	45.06	2.04	201.00
27/01/2016	1.35	446.88	11.64	295.22	4.26	97.40	16.33	57.95	34.10	11.00	25.15	10.46	3.04	95.19	34.28	44.56	1.87	184.00
28/01/2016	1.24	435.98	11.65	296.76	4.25	99.10	16.32	54.30	34.32	18.00	25.18	13.15	2.92	82.92	34.27	40.49	1.76	173.00
29/01/2016	1.25	437.00	11.70	303.65	4.28	102.70	16.31	52.88	34.61	31.00	25.28	30.63	3.07	91.08	34.30	45.23	1.74	171.00
30/01/2016	1.30	442.00	11.75	319.00	4.30	104.00	16.31	52.00	34.39	25.00	25.21	30.00	3.39	99.00	34.29	44.00	1.83	173.00
31/01/2016	1.30	442.00	11.71	307.00	4.30	103.00	16.30	51.00	34.28	22.00	25.19	29.00	3.26	96.00	34.27	42.00	1.85	175.00

Figure B-11: Raw observed discharge at selected gauging stations for diverted flow calculations and nested catchment modelling

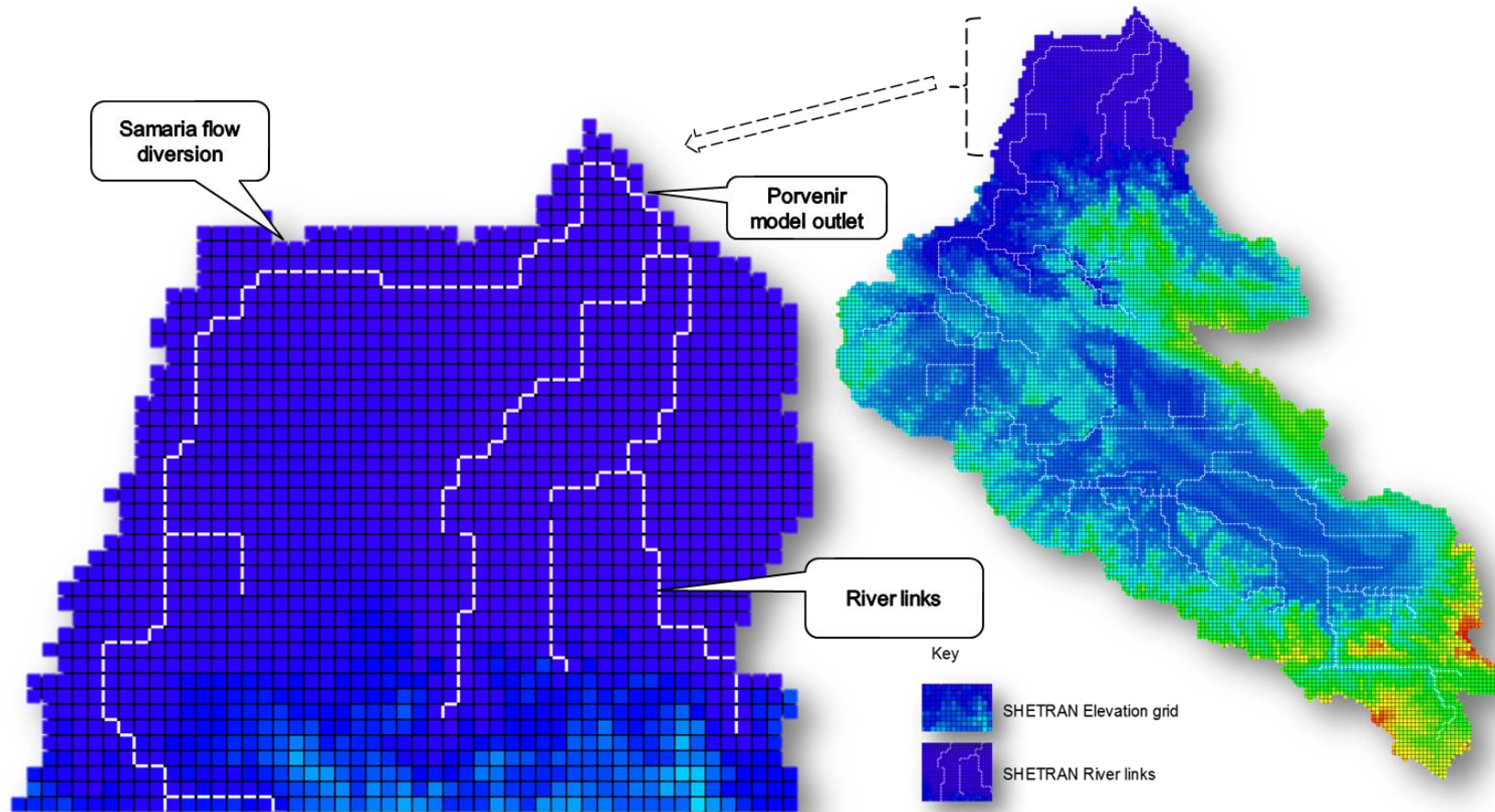


Figure B-12: The SHETRAN grid shows that diverted flows in Samaria River are not diverted but are wrongly part of simulated flows at Porvenir outlet



Figure B-13: On the ground, a fraction of Grijalva river flows are diverted from the model domain through Samaria River

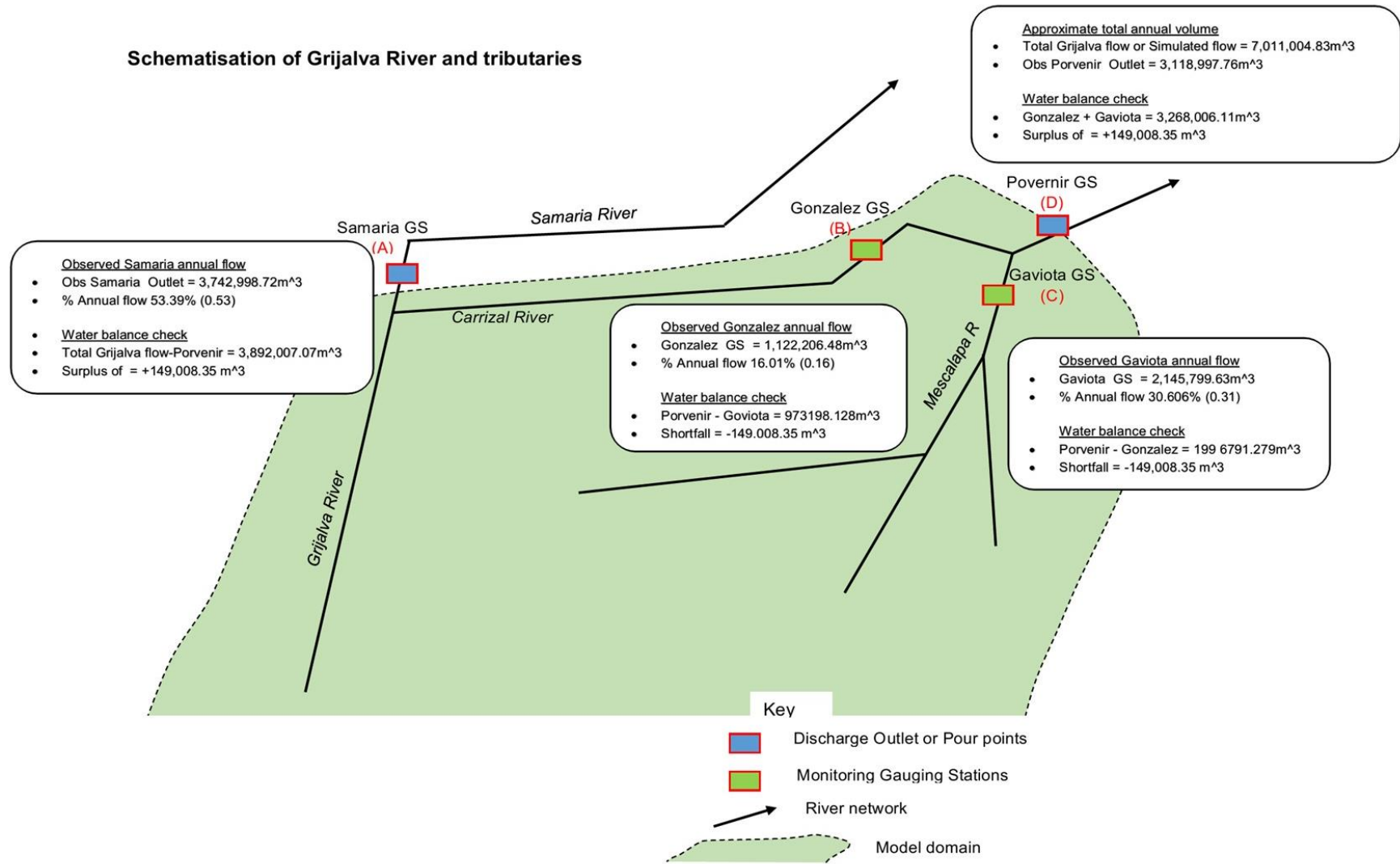


Figure B-14: Overview of Samaria River diverting flows off Grijalva catchment and El Macayo barrier that regulates remaining flows



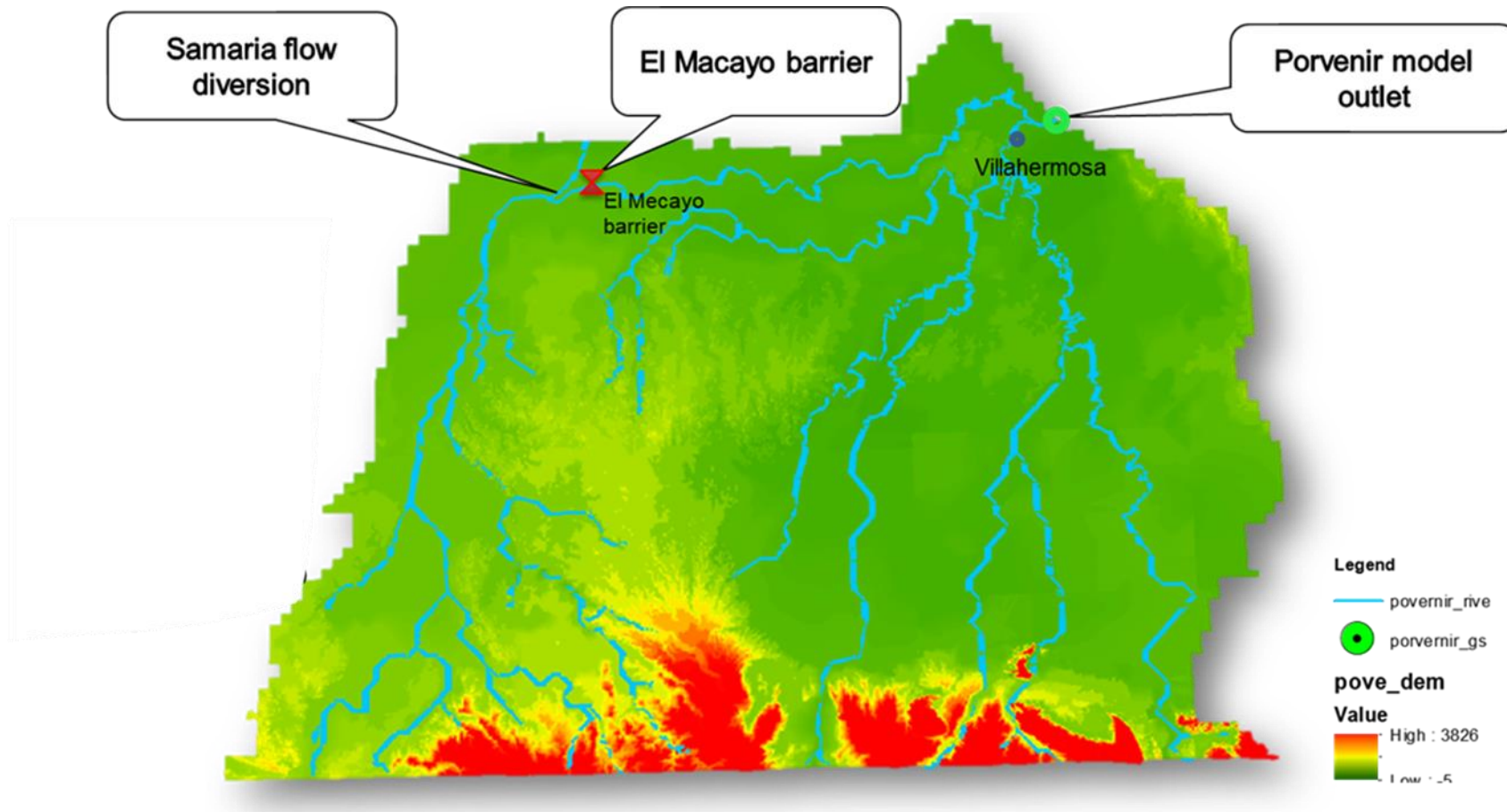


Figure B-15: In real-world a fraction of Grijalva flows area diverted from the model domain through Samaria and El Macayo barrier

Calculation of diverted flows from Grijalva Model Domain																		
Date	Observed Discharge (m <sup>3</sup> /s)				Simulated flow (m <sup>3</sup> /s)		Discharge volume estimates (m <sup>3</sup> )				Water Balance Analysis							
	Porvernir GS	Gaviota GS	Gonzalez GS	Samaria GS	Example of Simulated discharge	Corrected Discharge at Porvernir	$V_Q = \sum_{i=1}^{N-1} (t_{i+1} - t_i) \left( \frac{Q_{i+1} + Q_i}{2} \right)$				Gonzalez + Gaviota	Difference	Povernir - Gaviota =	Difference	Povernir - Gonzalez =	Difference	Samaria Net flow	Difference
					(Sim Q - % Diverted Q)	Gonzalez Area/Vol (B)	Samaria Area/Vol (A)	Porvernir Area/Vol (D)	Gaviota Area/Vol (C)	Porvernir B + C = D	Gonzalez flows D - C = B		Gaviota flows D - B = C		Samaria flow (A+B + C) -			
01/01/2015	366.00	154.95	165.64	594.00	301.27	160.85	3907.66	12749.76	8808.00	3740.89	7648.55	-1159.45	5067.11	1159.45	4900.34	1159.45	11590.31	-1159.45
02/01/2015	368.00	156.79	160.00	468.48	299.72	160.02	3746.13	12413.76	8400.00	3364.92	7111.05	-1288.95	5035.08	1288.95	4653.87	1288.95	11124.81	-1288.95
03/01/2015	332.00	123.62	152.18	566.00	298.84	159.55	3786.68	13656.00	8040.00	3077.41	6864.10	-1175.90	4962.59	1175.90	4253.32	1175.90	12480.10	-1175.90
04/01/2015	338.00	132.83	163.38	572.00	301.11	160.76	3948.21	12702.48	8208.00	4227.44	8175.65	-32.35	3980.56	32.35	4259.79	32.35	12670.13	-32.35
05/01/2015	346.00	219.45	165.64	486.54	305.19	162.94	6789.88	13403.66	10332.96	6306.35	13096.22	2763.26	4026.61	-2763.26	3543.08	-2763.26	16166.93	2763.26
06/01/2015	515.08	306.08	400.18	630.43	313.72	167.49	8168.07	14267.23	13376.52	8318.90	16486.98	3110.46	5057.62	-3110.46	5208.45	-3110.46	17377.69	3110.46
07/01/2015	599.63	387.17	280.49	558.50	321.49	171.64	6406.48	12678.62	14445.96	9645.86	16052.35	1606.39	4800.10	-1606.39	8039.48	-1606.39	14285.01	1606.39
08/01/2015	604.20	416.66	253.38	498.05	329.33	175.83	6081.22	13189.73	15266.40	11039.17	17120.39	1853.99	4227.23	-1853.99	9185.18	-1853.99	15043.72	1853.99
09/01/2015	668.00	503.28	253.38	601.10	331.25	176.86	6693.24	14499.55	16032.00	12476.71	19169.95	3137.95	3555.29	-3137.95	9338.76	-3137.95	17637.50	3137.95
10/01/2015	668.00	536.45	304.39	607.20	335.73	179.25	7113.28	14496.00	16176.00	12852.68	19965.96	3789.96	3232.32	-3789.96	9062.72	-3789.96	18285.96	3789.96
11/01/2015	680.00	534.61	288.39	600.80	347.07	185.30	6304.91	12050.16	16106.88	12565.18	18870.08	2763.20	3541.70	-2763.20	9801.97	-2763.20	14813.36	2763.20
12/01/2015	662.24	512.49	237.02	403.38	354.86	189.46	5554.83	9916.32	15390.84	11990.16	17544.99	2154.15	3400.68	-2154.15	9836.01	-2154.15	12070.47	2154.15
13/01/2015	620.33	486.69	225.88	422.98	357.00	190.60	5347.67	11731.68	14841.72	11503.61	16851.27	2009.55	3338.11	-2009.55	9494.05	-2009.55	13741.23	2009.55
14/01/2015	616.48	471.95	219.76	554.66	358.36	191.33	5274.19	13727.88	14801.76	11459.38	16733.56	1931.80	3426.38	-1931.80	9527.57	-1931.80	15659.68	1931.80
15/01/2015	617.00	483.00	219.76	589.33	359.51	191.94	5632.05	14163.12	15277.92	11945.93	17577.98	2300.06	3331.99	-2300.06	9645.87	-2300.06	16463.18	2300.06
16/01/2015	656.16	512.49	249.58	590.93	360.45	192.44	6204.70	14903.16	16069.92	12963.26	19167.96	3098.04	3106.66	-3098.04	9865.22	-3098.04	18001.20	3098.04
17/01/2015	683.00	567.78	267.48	651.00	361.50	193.00	6591.35	15624.00	16584.00	13980.60	20571.95	3987.95	2603.40	-3987.95	9992.65	-3987.95	19611.95	3987.95
18/01/2015	699.00	597.27	281.80	651.00	362.60	193.59	6285.88	13954.08	17476.20	14245.99	20531.87	3055.67	3230.21	-3055.67	11190.32	-3055.67	17009.75	3055.67
19/01/2015	757.35	589.90	242.02	511.84	364.04	194.36	5718.66	13268.40	17999.88	13958.48	19677.15	1677.27	4041.40	-1677.27	12281.22	-1677.27	14945.67	1677.27
20/01/2015	742.64	573.31	234.53	593.86	366.62	195.74	5510.23	13861.32	17734.08	13715.21	19225.43	1491.35	4018.87	-1491.35	12223.85	-1491.35	15352.67	1491.35
21/01/2015	735.20	569.62	224.65	561.25	370.82	197.98	5187.67	13035.00	17546.40	13471.93	18659.60	1113.20	4074.47	-1113.20	12358.73	-1113.20	14148.20	1113.20
22/01/2015	727.00	553.04	207.65	525.00	375.82	200.65	4854.88	12496.80	17364.00	13051.73	17906.61	542.61	4312.27	-542.61	12509.12	-542.61	13039.41	542.61
23/01/2015	720.00	534.61	196.92	516.40	379.75	202.75	4683.56	12340.80	17208.00	12786.34	17469.90	261.90	4421.66	-261.90	12524.44	-261.90	12602.70	261.90
24/01/2015	714.00	530.92	193.38	512.00	380.91	203.37	4598.72	12480.00	17172.00	12896.92	17495.64	323.64	4275.08	-323.64	12573.28	-323.64	12803.64	323.64
25/01/2015	717.00	543.82	189.85	528.00	378.72	202.20	4458.45	10865.52	17320.08	13272.89	17731.34	411.26	4047.19	-411.26	12861.63	-411.26	11276.78	411.26
26/01/2015	726.34	562.25	181.69	377.46	373.82	199.58	4374.43	10679.88	17140.08	13184.42	17558.85	418.77	3955.66	-418.77	12765.65	-418.77	11098.65	418.77
27/01/2015	702.00	536.45	182.85	512.53	367.24	196.07	4472.38	12173.04	16848.00	12808.45	17280.83	432.83	4039.55	-432.83	12375.62	-432.83	12605.87	432.83
28/01/2015	702.00	530.92	189.85	501.89	359.79	192.09	4486.32	11893.44	15777.48	12454.60	16940.92	1163.44	3322.88	-1163.44	11291.16	-1163.44	13056.88	1163.44
29/01/2015	612.79	506.96	184.01	489.23	352.25	188.07	4360.59	11814.60	14490.24	11857.46	16218.05	1727.81	2632.78	-1727.81	10129.65	-1727.81	13542.41	1727.81
30/01/2015	594.73	481.16	179.37	495.32	344.93	184.16	4249.64	11739.84	14120.76	11216.10	15465.74	1344.98	2904.66	-1344.98	9871.12	-1344.98	13084.82	1344.98
31/01/2015	582.00	453.52	174.76	483.00	337.81	180.36	4139.42	11448.00	13848.00	10641.08	14780.50	932.50	3206.92	-932.50	9708.58	-932.50	12380.50	932.50
01/02/2015	572.00	433.24	170.19	471.00	331.10	176.78	3908.47	9352.32	13596.00	10132.42	14040.89	444.89	3463.58	-444.89	9687.53	-444.89	9797.21	444.89
02/02/2015	561.00	411.13	155.52	308.36	324.92	173.48	3573.38	7312.32	13236.00	9402.59	12975.97	-260.03	3833.41	260.03	9662.62	260.03	7052.29	-260.03
03/02/2015	542.00	372.42	142.26	301.00	319.29	170.47	3533.30	7872.00	12839.28	8672.76	12206.06	-633.22	4166.52	633.22	9305.98	633.22	7238.78	-633.22
04/02/2015	527.94	350.31	152.18	355.00	314.21	167.76	4610.77	11112.00	13040.64	8672.76	13283.53	242.89	4367.88	-242.89	8429.87	-242.89	11354.89	242.89
06/02/2015	558.78	372.42	232.05	571.00	305.30	163.00	5703.94	13896.00	13773.36	9844.91	15548.85	1775.49	3928.45	-1775.49	8069.42	-1775.49	15671.49	1775.49
07/02/2015	589.00	447.99	243.28	587.00	301.23	160.83	5823.59	13992.00	14184.00	10795.90	16619.49	2435.49	3388.10	-2435.49	8360.41	-2435.49	16427.49	2435.49
08/02/2015	593.00	451.67	242.02	579.00	297.31	158.73	5196.55	13976.04	15104.40	10419.92	15616.48	512.08	4684.48	-512.08	9907.85	-512.08	14488.12	512.08
09/02/2015	665.70	416.66	191.02	585.67	293.53	156.71	5002.86	14927.16	15776.40	9601.63	14604.49	-1171.91	6174.77	1171.91	10773.54	1171.91	13755.25	-1171.91
10/02/2015	649.00	383.48	225.88	658.26	289.87	154.76	5391.69	15750.00	14469.60	8893.92	14285.61	-183.99	5575.68	183.99	9077.91	183.99	15566.01	-183.99
11/02/2015	556.80	357.68	223.43	654.24	286.35	152.88	5332.86	15788.28	13089.60	8252.56	13585.41	495.81	4837.04	-495.81	7756.74	-495.81	16284.09	495.81

Figure B-16: Calculation of diverted flows from Grijalva catchment

24/09/2015	654.50	849.00	99.02	560.55	749.49	400.15	2561.93	13600.92	15995.64	21398.40	23960.33	7964.69	-5402.76	-7964.69	13433.71	-7964.69	21565.61	7964.69
25/09/2015	678.47	934.20	114.47	572.86	761.80	406.73	3484.64	14145.00	16649.64	18280.56	21765.20	5115.56	-1630.92	-5115.56	13165.00	-5115.56	19260.56	5115.56
28/09/2015	709.00	589.18	175.91	605.89	723.84	386.46	3422.27	13923.00	17016.00	13873.20	17295.47	279.47	3142.80	-279.47	13593.73	-279.47	14202.47	279.47
29/09/2015	709.00	566.92	109.28	554.36	704.06	375.90	2487.38	13199.04	16494.36	12756.00	15243.38	-1250.98	3738.36	1250.98	14006.98	-1250.98	11948.06	-1250.98
30/09/2015	665.53	496.08	98.01	545.56	686.93	366.75	2340.00	13007.76	16763.04	14805.36	17145.36	382.32	1957.68	-382.32	14423.04	-382.32	13390.08	382.32
01/10/2015	731.39	737.70	96.99	538.42	673.34	359.50	2388.85	13133.04	17800.68	19006.32	21395.17	3594.49	-1205.64	-3594.49	15411.83	-3594.49	16727.53	3594.49
02/10/2015	752.00	846.16	102.08	556.00	659.25	351.97	2256.62	12624.24	17604.00	18375.84	20632.46	3028.46	-771.84	-3028.46	15347.38	-3028.46	15652.70	3028.46
05/10/2015	715.00	685.16	85.97	496.02	617.34	329.60	2004.17	10768.56	17244.00	16141.20	18145.37	901.37	1102.80	-901.37	15239.83	-901.37	11669.93	901.37
06/10/2015	722.00	659.94	81.04	401.36	607.17	324.17	1909.72	9556.32	16735.92	15401.64	17311.36	575.44	1334.28	-575.44	14826.20	-575.44	10131.76	575.44
07/10/2015	672.66	623.53	78.10	395.00	598.35	319.46	1851.09	10550.64	15928.08	13739.52	15590.61	-337.47	2188.56	337.47	14076.99	337.47	10213.17	-337.47
08/10/2015	654.68	521.43	76.15	484.22	591.64	315.87	1792.80	11594.64	15052.32	12229.32	14022.12	-1030.20	2823.00	1030.20	13259.52	1030.20	10564.44	-1030.20
09/10/2015	599.68	497.68	73.25	482.00	585.36	312.52	2177.84	12390.48	14408.16	13469.04	15646.88	1238.72	939.12	-1238.72	12230.32	-1238.72	13629.20	1238.72
12/10/2015	601.00	624.74	108.24	550.54	565.60	301.97	2660.05	13066.32	15046.44	14414.04	17074.09	2027.65	632.40	-2027.65	12386.39	-2027.65	15093.97	2027.65
13/10/2015	652.87	576.43	113.43	538.32	561.49	299.78	2586.09	10866.24	14915.40	13499.04	16085.13	1169.73	1416.36	-1169.73	12329.31	-1169.73	12035.97	1169.73
14/10/2015	590.08	548.49	102.08	367.20	558.85	298.37	2413.19	8717.16	14161.92	12283.56	14696.75	534.83	1878.36	-534.83	11748.73	-534.83	9251.99	534.83
15/10/2015	590.08	475.14	99.02	359.23	559.52	298.73	2400.94	8923.80	14191.44	12530.40	14931.34	739.90	1661.04	-739.90	11790.50	-739.90	9663.70	739.90
16/10/2015	592.54	569.06	101.06	384.42	573.63	306.26	2972.50	9895.68	15397.80	13832.40	16804.90	1407.10	1565.40	-1407.10	12425.30	-1407.10	11302.78	1407.10
19/10/2015	690.61	583.64	146.65	440.22	646.99	345.43	4544.43	14725.20	17354.16	15412.32	19956.75	2602.59	1941.84	-2602.59	12809.73	-2602.59	17327.79	2602.59
20/10/2015	755.57	700.72	232.05	786.88	685.29	365.87	6288.17	19792.90	19569.48	17708.64	23996.81	4427.33	1860.84	-4427.33	13281.31	-4427.33	24220.22	4427.33
21/10/2015	875.22	775.00	291.96	862.53	719.27	384.02	6593.80	19974.24	21418.32	18912.00	25505.80	4087.48	2506.32	-4087.48	14824.52	-4087.48	24061.72	4087.48
22/10/2015	909.64	801.00	257.52	801.99	730.86	390.20	6550.89	18308.42	21011.28	19782.00	26332.89	5321.61	1229.28	-5321.61	14460.39	-5321.61	23630.04	5321.61
23/10/2015	841.30	847.50	288.39	723.71	739.30	394.71	5194.08	13251.24	19923.60	21653.76	26847.84	6924.24	-1730.16	-6924.24	14729.52	-6924.24	20175.48	6924.24
26/10/2015	819.00	956.98	144.45	380.56	883.22	471.55	3401.33	8950.20	19693.08	23544.60	26945.93	7252.85	-3851.52	-7252.85	16291.75	-7252.85	16203.05	7252.85
27/10/2015	822.09	1005.07	138.99	365.29	948.40	506.35	3283.75	8730.60	19686.00	21676.32	24960.07	5274.07	-1990.32	-5274.07	16402.25	-5274.07	14004.67	5274.07
28/10/2015	818.41	801.29	134.66	362.26	1011.74	540.17	3115.98	8604.72	19571.64	18923.64	22039.62	2467.98	648.00	-2467.98	16455.66	-2467.98	11072.70	2467.98
29/10/2015	812.56	775.68	125.01	354.80	1060.55	566.23	2886.34	8538.72	19365.24	18316.56	21202.90	1837.66	1048.68	-1837.66	16478.90	-1837.66	10376.38	1837.66
30/10/2015	801.21	750.70	115.52	356.76	1079.37	576.28	2538.01	8159.04	18614.52	15731.16	18269.17	-345.35	2883.36	345.35	16076.51	345.35	7813.69	-345.35
02/11/2015	750.00	560.23	95.98	323.16	976.65	521.43	2340.04	7665.00	18274.08	12424.44	14764.48	-3509.60	5849.64	3509.60	15934.04	3509.60	4155.40	-3509.60
03/11/2015	772.84	475.14	99.02	315.59	962.16	513.70	2376.52	8471.64	18167.52	11301.96	13678.48	-4489.04	6865.56	4489.04	15791.00	4489.04	3982.60	-4489.04
04/11/2015	741.12	466.69	99.02	390.38	941.62	502.73	2376.52	8310.00	17463.48	11117.16	13493.68	-3969.80	6346.32	3969.80	15086.96	3969.80	4340.20	-3969.80
05/11/2015	714.17	459.74	99.02	302.12	908.79	485.20	2364.34	7224.12	16946.04	10928.88	13293.22	-3652.82	6017.16	3652.82	14581.70	3652.82	3571.30	-3652.82
06/11/2015	698.00	451.00	98.01	299.89	875.04	467.18	2587.45	7704.48	16555.68	11040.00	13627.45	-2928.23	5515.68	2928.23	13968.23	2928.23	4776.25	-2928.23
09/11/2015	681.64	469.00	117.61	342.15	859.74	459.02	2797.59	8052.24	16035.84	12061.92	14859.51	-1176.33	3973.92	1176.33	13238.25	1176.33	6875.91	-1176.33
11/11/2015	654.68	536.16	115.52	328.87	881.18	470.46	2772.44	7654.32	15524.16	12826.32	15598.76	74.60	2697.84	-74.60	12751.72	-74.60	7728.92	74.60
12/11/2015	639.00	532.70	115.52	308.99	878.86	469.22	2772.44	7166.52	14115.12	12562.08	15334.52	1219.40	1553.04	-1219.40	11342.68	-1219.40	8385.92	1219.40
13/11/2015	537.26	514.14	115.52	288.22	866.20	462.47	3994.08	8370.48	15029.64	14843.16	18837.24	3807.60	186.48	-3807.60	11035.56	-3807.60	12178.08	3807.60
16/11/2015	715.21	722.79	217.32	409.32	810.69	432.83	4434.01	9282.96	17128.80	16442.28	20876.29	3747.49	686.52	-3747.49	12694.79	-3747.49	13030.45	3747.49
17/11/2015	712.19	647.40	152.18	364.26	787.17	420.27	3533.30	8548.44	16967.52	15339.84	18873.14	1905.62	1627.68	-1905.62	13434.22	-1905.62	10454.06	1905.62
18/11/2015	701.77	630.92	142.26	348.11	764.30	408.06	3348.98	8233.20	16731.24	15167.52	18516.50	1785.26	1563.72	-1785.26	13382.26	-1785.26	10018.46	1785.26
19/11/2015	692.50	633.04	136.82	337.99	743.85	397.14	3218.91	8007.72	16566.00	15017.64	18236.55	1670.55	1548.36	-1670.55	13347.09	-1670.55	9678.27	1670.55
20/11/2015	688.00	618.43	131.42	329.32	726.65	387.96	3271.12	8430.96	16641.72	14021.16	17292.28	650.56	2620.56	-650.56	13370.60	-650.56	9081.52	650.56
23/11/2015	698.81	550.00	141.17	373.26	701.24	374.39	3668.13	9721.44	17412.60	14932.20	18600.33	1187.73	2480.40	-1187.73	13744.47	-1187.73	10909.17	1187.73

Figure B-17: Calculation of diverted flows from Grijalva catchment



Computing the volume of flow from Samaria and Gonzalez hydrographs				Water balance Checks			
Method : Computing the area under the curve of the hydrographs at Gonzalez and Samaria GS				Porvernir Flow balance check Gonzalez (B) + Gaviota ( C ) = Porvernir Outlet (D)			
$V_Q = \sum_{i=1}^{N-1} (t_{i+1} - t_i) \left( \frac{Q_{i+1} + Q_i}{2} \right)$				Gonzalez and Gaviota total 1,122,206.48 + 2,145,799.63 = 3,268,006.11			
				Obs at the outlet = 3,118,997.76			
(Boise State University lecture notes 2008)				Surplus = 149,008.35			
<b>Approx total Annual flow volumes</b>				<b>Gonzalez Flow balance check</b>			
		<b>Gaviota</b>	C	2,145,799.63 m <sup>3</sup> *	Porvernir (D) - Gaviota (C)	=	Gonzalez (B)
		<b>Gonzalez</b>	B	1,122,206.48 m <sup>3</sup> *	Porvernir and Gaviota Difference	3,118,997.76 - 2,145,799.63	= 973,198.13
		<b>Samaria</b>	A	3,742,998.72 m <sup>3</sup>	Obs at the outlet	=	1,122,206.48
		<b>Observed Porvernir Outlet</b>	D	3,118,997.76 m <sup>3</sup>	Shortfall	=	-149,008.35
Total Grijava and Carrizal annual flow volumes passing both gauges				Gaviota Flow balance check Porvernir (D) - Gonzalez (B) = Gaviota (C)			
		Gaviota		2,145,799.63	Porvernir and Gonzalez Difference	3,118,997.76 - 1,122,206.48	= 1,996,791.28
		Gonzalez		1,122,206.48	Obs at the outlet	=	2,145,799.63
		Samaria		3,742,998.72	Shortfall	=	-149,008.35
		Simulated Annual total Vol		<b>7,011,004.83 m<sup>3</sup></b>			
<b>Percentage annula flow volume calculations</b>				<b>Gonzalez and Samaria flow to check literature values</b>			
Annual % of discharge diverted through Gaviota				Total Gonzalez and Samaria daily flows 4,865,205.20			
				2,145,799.63	Gonzalez	B	1122206.481
				7,011,004.83	Samaria	A	3742998.72
		Annual % flow		<b>30.61 %</b>			
Annual % of discharge diverted through Gonzalez				Gonzalez 1122206.481			
				1,122,206.48			
				7,011,004.83			
		Annual % flow		<b>16.01 %</b>			
Annual % of discharge diverted through Samaria				Samaria 3742998.72			
				3,742,998.72			
				7,011,004.83			
		Annual % flow		<b>53.38748 %</b>			
<b>Approx total flow volumes for rain season</b>				<b>Wet Seasonal Total flows for Grijava and Carrizal passing both gauges</b>			
<b>Rain season</b>				Gaviota 1,027,462.68			
	June	150	Total seasonal rainfall	Gaviota			4865205.201
	Nov	276		Gonzalez			23.06596401
				Samaria			
				Porvernir Outlet			
					Samaria	A	3742998.72
							4865205.201
							76.93403599

Figure B-19: Calculation of diverted flows from Grijalva catchment

Wet Seasonal Total flows for Grijava and Carrizal passing both gauges			Gaviota	1,027,462.68
			Gonzalez	405,797.07
			Samaria	1,623,845.76
			wet seasonal total Vol	<b>3,057,105.51</b>
<b>Percentage seasonal flow volume calculations</b>				
Rain seasonal % of discharge diverted through Gaviota			1,027,462.68	
			<b>3,057,105.51</b>	
			% seasonal flow	<b>33.61 %</b>
Rain seasonal % of discharge diverted through Gonzalez			405,797.07	
			3,057,105.51	
			% seasonal flow	<b>13.27 %</b>
Rain seasonal % of discharge diverted through Samaria			1,623,845.76	
			3,057,105.51	
			% seasonal flow	<b>53.11710 %</b>
<b>Dry season</b>				
Total seasonal rainfall	Dec-01	277	Gaviota	364,019.64
	Dec-31	298	Gonzalez	754,317.31
	jan	9	Samaria	87,002.91
	may	149	Porvernir Outlet	629,406.50
				189,279.60
				1,929,873.36
				411,833.16
				1,287,001.44
Dry Seasonal total flows for Grijava and Carrizal			Gaviota	1,118,336.95
			Gonzalez	716,409.41
			Samaria	2,119,152.96
			wet seasonal total Vol	<b>3,953,899.33</b>
<b>Percentage seasonal flow volume calculations</b>				
Rain seasonal % of discharge diverted through Gaviota			1,118,336.95	
			<b>3,044,240.55</b>	
			% seasonal flow	<b>36.74 %</b>
Rain seasonal % of discharge diverted through Gonzalez			716,409.41	

4865205.201  
76.93403599

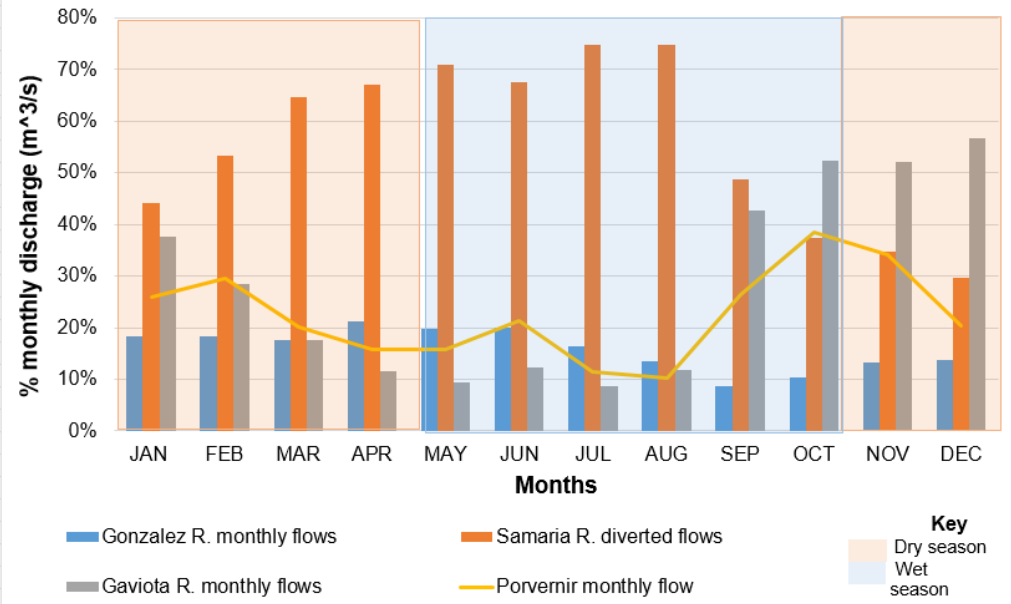


Figure B-20: Calculation of diverted flows from Grijalva catchment

## Appendix C : Simulated and observed discharge

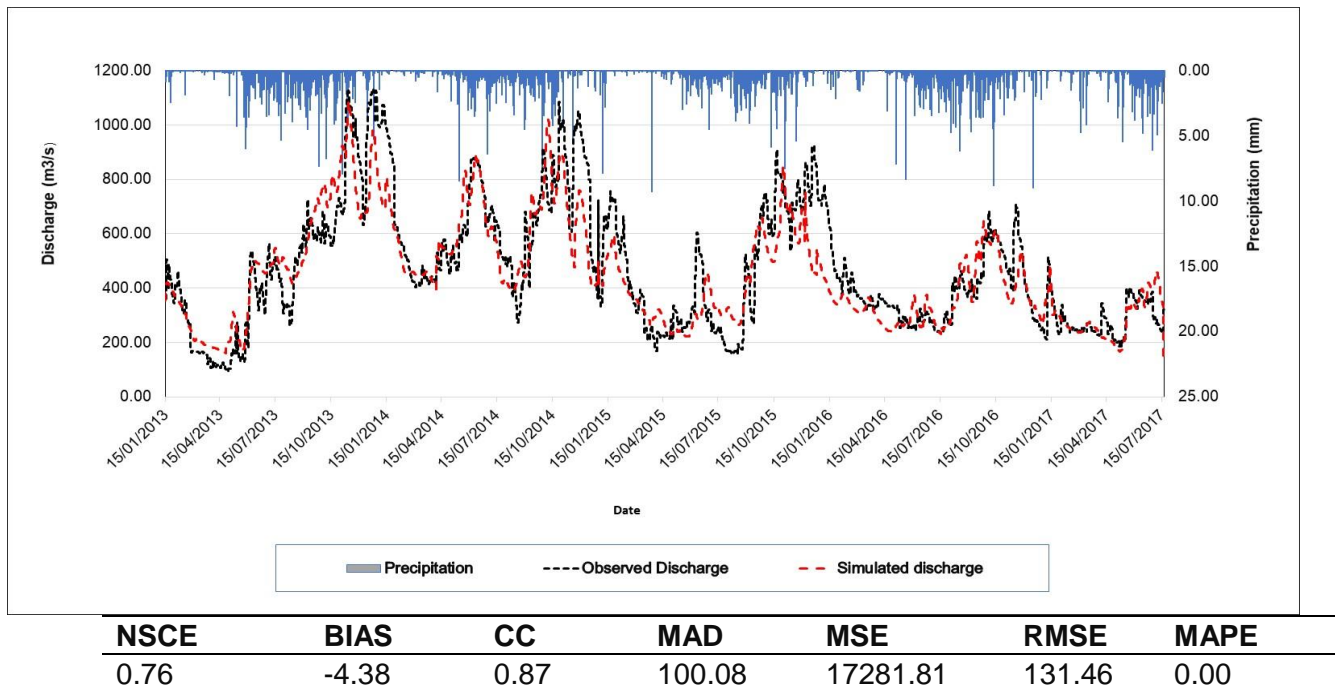


Figure C-1: Comparison of simulated and observed discharge at Porvenir gauging station 2013-2017

## Appendix D : Nested catchment modelling

Table D1: Physiographic characteristics of nested catchments.

	<b>Gaviota</b>	<b>Pueblo Nuevo</b>	<b>Puente</b>	<b>Teapa</b>	<b>Puyacatengo</b>	<b>Tapijulapa</b>	<b>Oxolotan</b>
<b>Topography</b>	Southern 9m and in the northern Sierra Mountains 1 673m	Southern 9m and in the northern Sierra Mountains 1 673m	Southern 9m and in the north Sierra Mountains 1 673m	Northern Sierra Mountains 18 to 1 777m	Northern Sierra Mountains 1 777m	Northern Sierra Mts 66-1179m	Northern Sierra Mts 66-1179m
<b>Soils</b>	Deep alluvial clay	Deep alluvial clay	Deep alluvial clay	Thin mountain	Thin mountain	Thin mountain	Thin mountain
<b>Slope</b>	Gentle to nearly flat. U-shaped slopes in the south. Very steep V-shaped in north	Gentle to nearly flat U-shaped slopes in the south. Very steep in north	Gentle to nearly flat U-shaped slopes in the south. Very steep V-shaped in north	Very steep V-shaped slopes	Very steep V-shaped slopes	Very steep V-shaped slopes	Very steep V-shaped slopes
<b>Climate</b>	Warm humid in the north Semi-warm Sub-mist in the south	Warm humid in the north Semi-warm Sub-mist in the south	Warm humid in the north Semi-warm Sub-mist in the south	Semi-warm Sub	Warm to sub-humid	Semi-warm Sub and Moist Temperate	Semi-warm Sub and Moist Temperate
<b>Temperatures</b>	24 in lowland area to 14 °C in upland areas	24 in lowland area to 14 in upland areas	24 in lowland area to 14 in upland areas	20 to 14 °C	18 to 14 °C	18 to 14 °C	18 to 14 °C
<b>Av. annual rainfall</b>	4 000 mm in 4 500mm south	3000 in the north 4 500mm south	2 500mm north and 4 500mm south	3 500mm lowland south 4 500 north of	4 500mm	4 500mm south	4 500mm south
<b>Vegetation</b>	Northern arable & Grasslands. Southern mountain forest	Northern arable & Grasslands. Southern mountain forest	Northern arable Grasslands. Southern mountain forest	Mountain forest	Mountain forest	Mountain forest	Mountain forest
<b>Runoff Coef.</b>	20 to 30%	20 to 30%	20 to 30%	>30%	20 to 30%	>30%	>30%



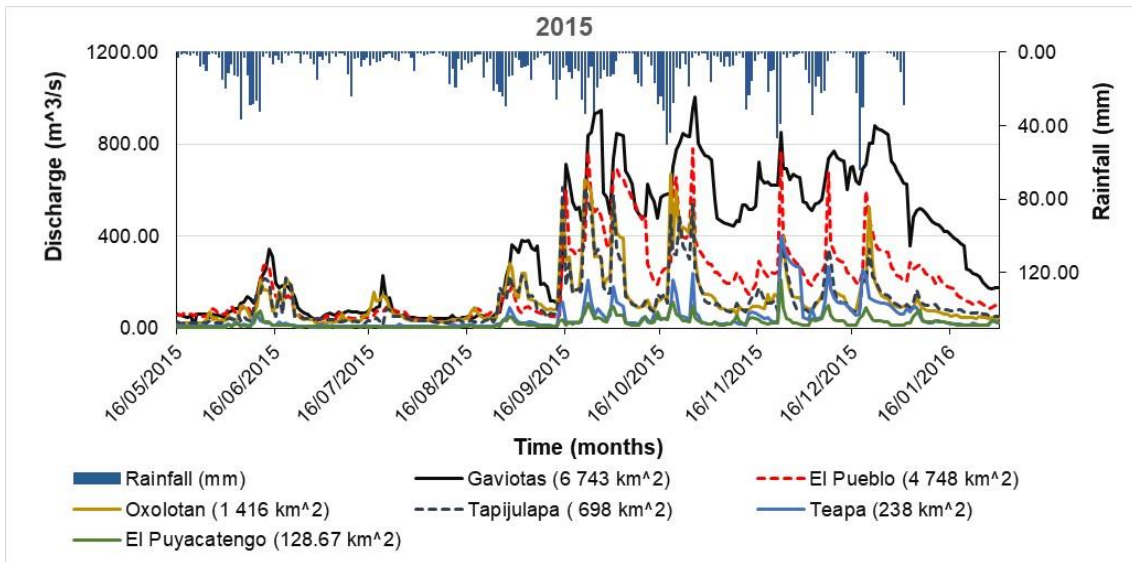


Figure D-1: 2015 Nested catchment hydrographs

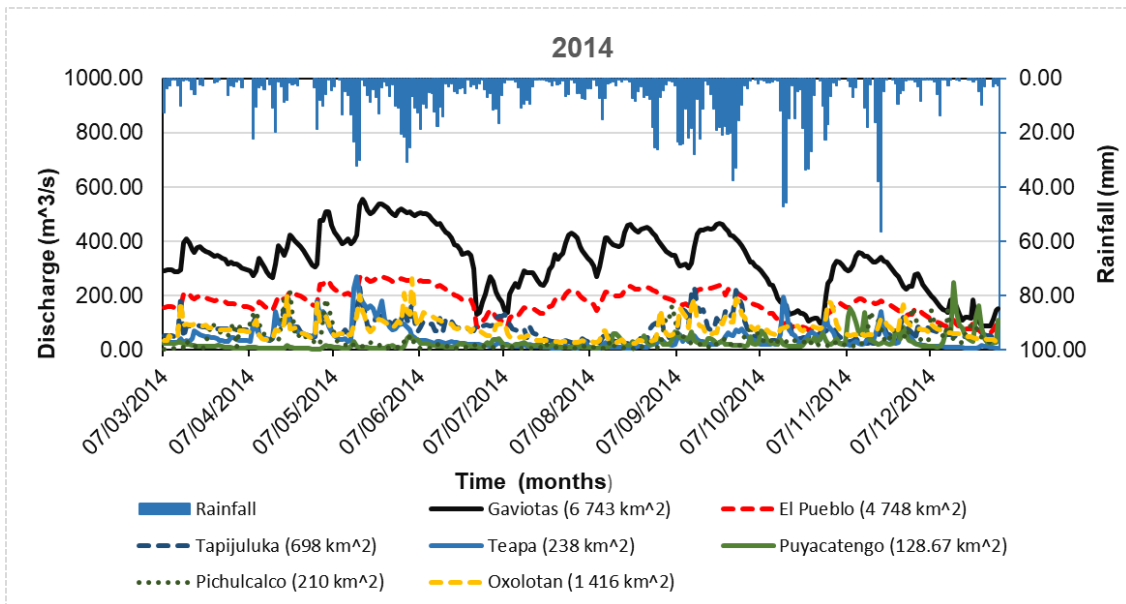
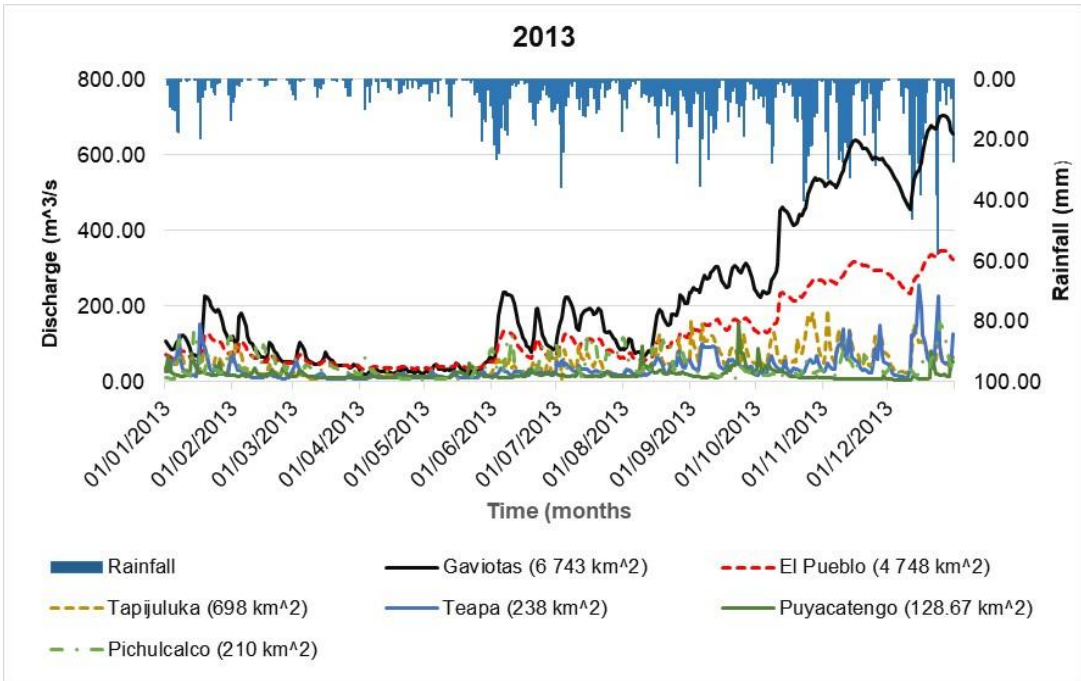
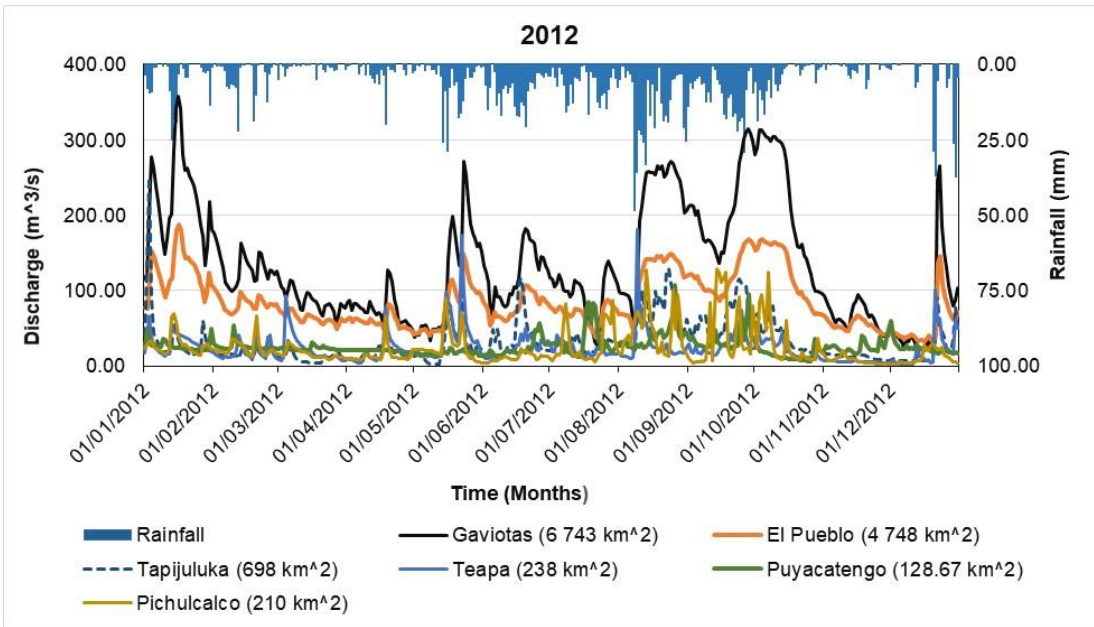


Figure D-2: 2014 Nested catchment hydrographs. The well-above-average rainfall-runoff activity shown was a result of the 2014 Pacific hurricane season in the study area. The 2014 season experienced fourteen named storms, seven hurricanes, and five major hurricanes in one year. ‘The 2014 Pacific hurricane season’ (2014)



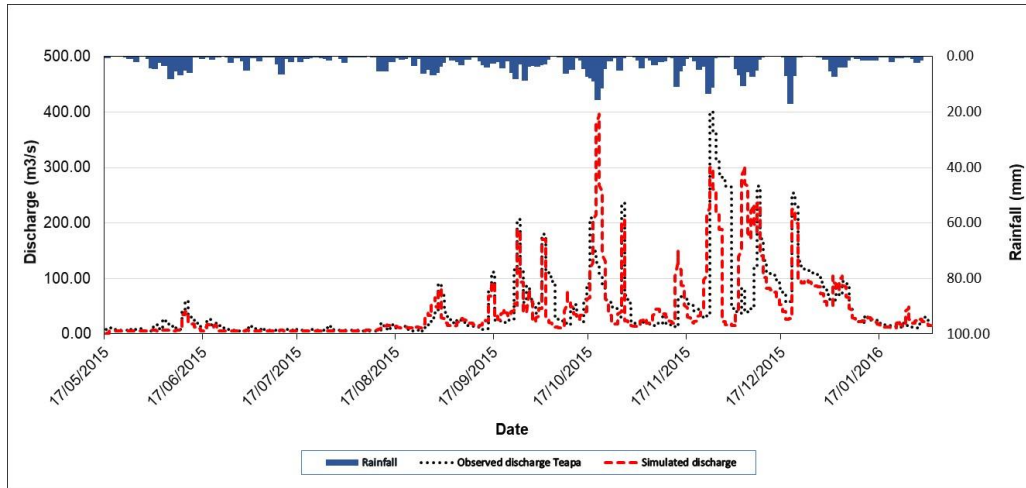
**Figure D-3: 2013 nested catchment hydrographs**



**Figure D-4: 2012 nested catchment hydrographs**

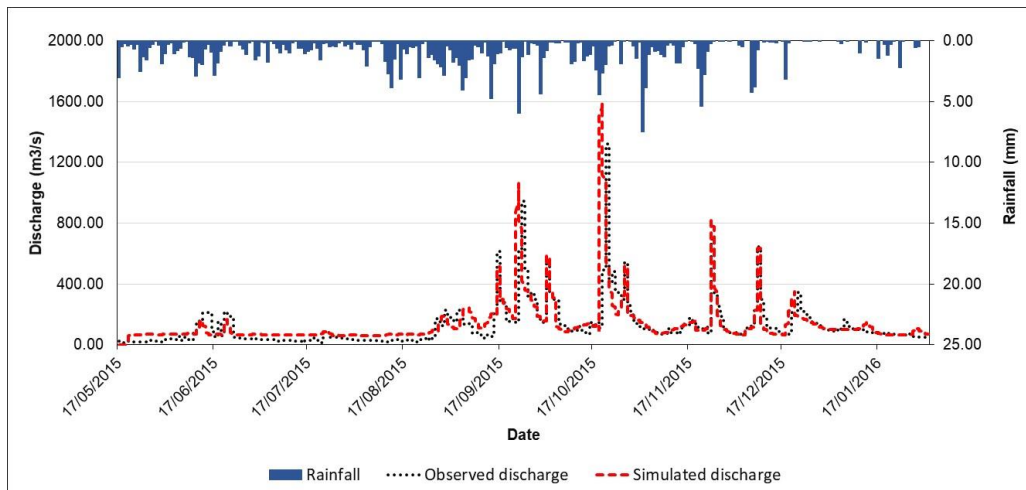
## Appendix E : Comparison of simulated and observed discharge in selected nested catchments

All nested catchment models were automatically calibrated using the OSTRICH (Optimisation Software Toolkit for Research Involving Computational Heuristics) (v17.12.19) tool, (Matott, 2005). A model-independent program that was integrated with the SHETRAN hydrological model for automatic calibration.



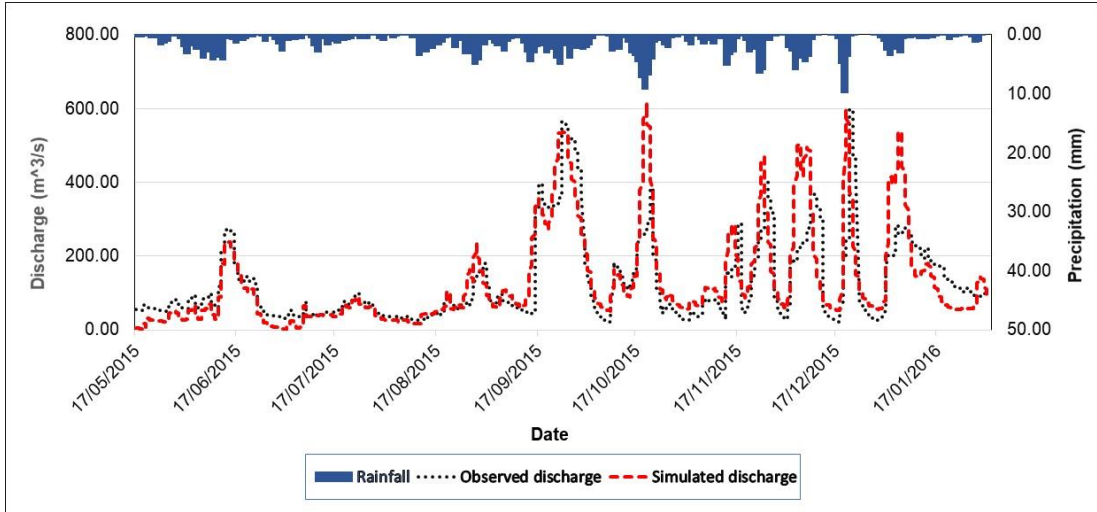
NSCE	BIAS	CC	MAD	RMSE	MAPE
0.44	-8.60	0.71	0.00	48.03	46.49

Figure E-1: Comparison of observed and simulated Teapa discharge



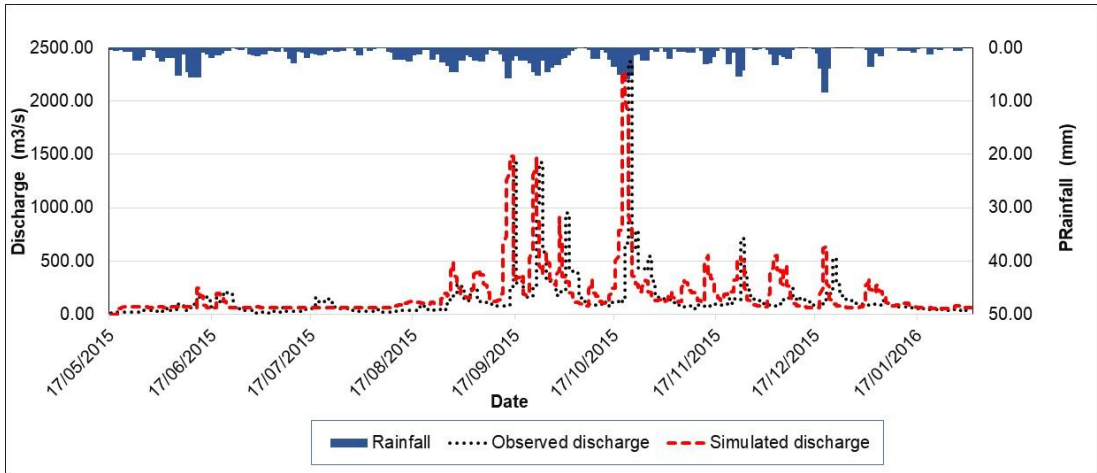
NSCE	BIAS	CC	MAD	RMSE	MAPE
0.49	9.18	0.76	0.02	104.93	61.52

Figure E-2: Comparison of observed and simulated Tapijulapa discharge



NSCE	BIAS	CC	MAD	RMSE	MAPE
0.60	4.68	0.83	0.01	73.34	43.56

Figure E-3: Comparison of observed and simulated Pueblo discharge



NSE	BIAS	CC	MAD	RMSE	MAPE
0.55	29.64	0.51	0.02	244.42	98.30

Figure E-4: Comparison of observed and simulated Oxolotan discharge