CRANFIELD UNIVERSITY

ALI ALDERETE PERALTA

SPATIO-TEMPORAL MODELLING OF DIFFUSION OF ELECTRIC VEHICLES AND SOLAR PHOTOVOLTAIC PANELS: AN INTEGRATED AGENT-BASED AND ARTIFICIAL NEURAL NETWORKS MODEL

SCHOOL OF WATER, ENERGY AND ENVIRONMENT
PhD Environment And Agrifood

PhD
Academic Year: 2015 - 2020

Supervisor: Dr Nazmiye Ozkan
Associate Supervisor: Professor Athanasios Kolios
Associate Supervisor: Professor Philip Longhurst
July 2020

CRANFIELD UNIVERSITY

SCHOOL OF WATER, ENERGY AND ENVIRONMENT
PhD Environment And Agrifood

PhD

Academic Year 2015 - 2020

ALI ALDERETE PERALTA

SPATIO-TEMPORAL MODELLING OF DIFFUSION OF ELECTRIC VEHICLES AND SOLAR PHOTOVOLTAIC PANELS: AN INTEGRATED AGENT-BASED AND ARTIFICIAL NEURAL NETWORKS MODEL

Supervisor: Dr Nazmiye Ozkan
Associate Supervisor: Professor Athanasios Kolios
Associate Supervisor: Professor Philip Longhurst
July 2020

This thesis is submitted in partial fulfilment of the requirements for the degree of PhD
(NB. This section can be removed if the award of the degree is based solely on examination of the thesis)

ABSTRACT

This work investigates the spatio-temporal patterns of EV and PV adoption, attending to the need of informing the management of the distribution network with spatially explicit estimations of the EV and PV adoption rates. The research assesses the strengths and weaknesses of current modelling approaches and finds three recurrent techniques: agent-based model (ABM), the spatial regression (SR) and the Poisson model. The research addresses the limitations of the current modelling approaches to characterise the different factors that drive the decision-making for the adoption of these technologies: spatial and temporal dependence and social dynamics. The framework addresses some of the limitations of ABMs that use a rule or equation-based decision-making, by adopting an aggregated agents' definition and using artificial neural networks (ANN) as decision-making criteria. The integrated model provides a more realistic characterisation of the decision-making and its evolution over time, moreover, the results can inform network operators and policymakers explicitly about the location and pace of EV and PV adoption.

The research develops a spatio-temporally explicit ABM that accounts for: (i) spatial and (ii) temporal dependence, (iii) peer-effect, (iv) spillover effect and (iv) preferences towards other technologies. The development of the model follows a sequential approach to managing the complexity of constructing this new approach. First, two autoregressive models are developed to analyse the adoption patterns of EV and PV patterns, using the postcode and monthly data resolution. The temporal validation uses the Mean Absolute Percentage Error to measure the model's capability to replicate the time-series of the adoption rates. The spatial validation compares the actual and estimated spatial pattern of adoption by calculating the Moran's I index. Besides, the results are benchmarked against the Bass model, a commonly used tool for this purpose by ABM experts. The results show that in most of the cases the ABM and ANN integrated models perform better than the Bass model especially for those months with high fluctuations in the adoption rates. These models can estimate upmost three months with an accuracy higher than 80%, however, the models

present a significant accumulation of errors that limits the results for a longer forecast. To reduce the error accumulation and produce a longer forecast, the autoregressive PV model is extended by including socioeconomic variables. The resulting model improves the performance by 5% by the incorporation of variables including income, electricity consumption and average household size.

Lastly, the framework combines the EV and PV autoregressive models with a view to characterising the exchange of knowledge between EVs and PVs. This reflects the influences of owning one of those technologies on the preference for the second technology. The exchange of knowledge improves the performance of the model significantly with results above 80% of accuracy for eight months into the future. Given the high spatial resolution of the model, the results may help to design policies that recognise the socioeconomic differences within a geographical area. The research shows how the results can inform the management of the distribution network, by considering the worst-case scenario where the PV generation surplus is injected to the grid, and where the entire fleet of EVs are charged at home during the night. Also, the results of the hot spot analysis can inform the network operators about the emergence of clusters of EVs and PVs in the future.

The research finds that a spatio-temporally explicit ABM can characterise the EV and PV adoption process at the aggregated level, which also accounts for social effects. Such a model can also integrate heterogeneity amongst the population, whilst being resilient to changes in the size of the study area. The research also produces data-driven insights into the spatio-temporal patterns of EV and PV adoption, and how the adaptive capabilities of the ANN address some of the limitations of the ruled-based ABM. Lastly, the research finds that knowledge exchange takes place between the EV and PV adoption process. These findings are relevant to other low carbon technologies and for the modelling of other sociotechnical systems.

Keywords: Knowledge exchange; Decision-making; Energy systems; Complex system modelling; Policymaking.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my supervisor Dr Nazmiye Ozkan for her outstanding guidance and encouragement during this journey, and to my co-supervisors Professor Athanasios Kolios and Professor Philip Longhurst for their critical advice. I also acknowledge the National Council of Science and Technology in México (CONACYT) for the trust and financial support.

Also, my foremost thanks to all the people who were with me at the beginning, to those who stayed by my side, to those who have left and to those who will continue. Thanks to the people that have made my life brighter by sharing their own life with me. And thanks to those who are family by choice.

Edit: I want to explicitly acknowledge the help and support of those who pushed me upwards and onwards this last year: Fany, Erika, Citlalli, Daniela, Zoi, Meike, Roberta, Natalia, Ana, Becky, Omar, Gustavo, Gerard, Washington, Pablo, Roberto, Eduardo, and Daniel.

Finally, all my love and gratitude to those who were the farthest, yet, were the closest, Josué, Lilia and Isay.

De qué sirve ser más fuerte sin saber ser mejor.

...

There's no use in being stronger unless you know how to be better.

TABLE OF CONTENTS

ABSTRACT i
ACKNOWLEDGEMENTS iii
LIST OF FIGURES vii
LIST OF TABLES xi
1 Introduction 12
1.1 Literature review 15
1.1.1 Review method 16
1.1.2 Literature review questions 17
1.1.3 Localization and identification of relevant studies 18
1.1.4 The screening and selection of relevant literature 19
1.1.5 Analysis and synthesis 20
1.2 EV and PV adoption modelling approaches 21
1.2.1 Agent-based modelling 22
1.2.2 Spatial regression 25
1.2.3 Poisson model 29
1.2.4 Other modelling approaches 31
1.3 Review of strengths and weaknesses of the modelling approaches 31
1.4 Review Limitations 35
1.5 Discussion 35
1.5.1 Research gaps 39
1.6 Thesis aim and objectives 40
2 Explicit spatio-temporally characterisation of an agent-based and artificial neural networks integrated model 42
2.1 Introduction 42
2.1.1 Spatio-temporal characterisation of the technology diffusion 43
2.1.2 Artificial neural networks as a decision-making process 46
2.1.3 The spatial nature of the social effects 49
2.2 Methods and materials 50
2.2.1 Data 50
2.2.2 Spatio-temporal resolution 51
2.2.3 Temporal dependency 55
2.2.4 Spatial dependency 57
2.2.5 Model design 59
2.3 Model development 61
2.3.1 Decision-making process 66
2.3.2 Model validation 70
2.4 Results 71
2.4.1 Temporal validation 71
2.4.2 Spatial validation 75
2.4.3 Predictive accuracy 78

2.4.4 Agents explicit spatio-temporal characterisation 81
2.4.5 Social effects and decision-making process.................................... 82
2.4.6 Model validation 83
2.4.7 Predictive capability.. 84
2.5 Reflective summary 85
3 Dynamic characterisation of population heterogeneity in a spatio-temporally ABM and ANN integrated model 87
3.1 Introduction 87
3.1.1 Characterisation of heterogeneous agents...................................... 88
3.1.2 Modifiable areal unit problem 90
3.2 Methods and materials.. 91
3.2.1 Decision making and social effects 95
3.3 Results.. 98
3.3.1 Temporal validation.. 99
3.3.2 Spatial validation 107
3.3.3 Predictive accuracy 111
3.4 Reflective summary 116
4 A spatio-temporally explicit ABM and ANN integrated model: the EV adoption case 117
4.1 Introduction 117
4.2 Methods and materials... 118
4.2.1 Data.. 118
4.2.2 Spatio-temporal resolution 122
4.2.3 Spatio-temporal dependency 125
4.2.4 Decision-making process 127
4.2.5 Model validation 128
4.3 Results... 129
4.3.1 Temporal validation... 129
4.3.2 Spatial validation 132
4.3.3 Predictive accuracy 137
4.4 Reflective summary 138
5 Integrating knowledge exchange between EV and PV adoption into a spatio-temporally explicit ABM 140
5.1.1 EV and PV diffusion in Birmingham... 141
5.2 Methods 147
5.2.1 Decision-making and social effects 147
5.2.2 Model validation 149
5.3 Results... 149
5.3.1 Temporal validation... 149
5.3.2 Spatial validation 152
5.3.3 Predictive accuracy 155
5.4 Potential of applicability: network management 158

5.5 Reflective summary 164
6 Discussion and conclusions 166
6.1 Summary of findings 166
6.1.1 Model development 166
6.1.2 Temporal dependence 168
6.1.3 Spatial dependence 170
6.1.4 Results 173
6.2 Discussion 178
6.2.1 Contribution to knowledge 178
6.2.2 Review of the research question 188
6.2.3 Implications of the research 189
6.3 Research limitations 196
6.4 Future work 198
6.5 Concluding remarks 200
REFERENCES 202
APPENDICES 221

LIST OF FIGURES

Figure 1. A conceptual framework for the literature review. 16

Figure 2. The methodology followed in carrying out a systematic literature review. 17

Figure 3. Funnel diagram of selected articles that are included in the review. . 20

Figure 4. Annual cumulative number of publications by type of technology modelled. 22

Figure 5. Annual cumulative number of publications by type modelling approach. 22

Figure 6. Boxplot of the total PV installations in Birmingham. 52

Figure 7. Normalised monthly cumulative adoption rates of solar PV at the PC level. 54

Figure 8. Normalised monthly distribution of solar PV adoption at LSOA level. 54

Figure 9. Spatial distribution of the PV installations in the Birmingham postcodes (Dec 2015). 55

Figure 10. ACF of the PV installation data at the PC level. 56

Figure 11. PACF of the PV installation data at the PC level. 57

Figure 12. Hot spot analysis of the PV installations at PC level (Dec 2015)..... 59

Figure 13. Conceptual model of a spatio-temporal explicit ABM. 60

Figure 14. Methodological framework and information flow between the layers. 62

Figure 15. A four-step process for the implementation of the model. 64

Figure 16. Implementation of the conceptual model using Anylogic software. . 66

Figure 17. Artificial neural network design for a spatio-temporal explicit ABM. 69

Figure 18. Cumulative PV adoption rates estimated by ANN and Bass model vs. actual data. 73

Figure 19. The error of estimation for the ANN and Bass models. 73

Figure 20. The temporal pattern of the marginal changes of the estimation errors overtime for the Autoregressive model. 75

Figure 21. Spatial distribution of the ANN estimation’s error by the end of the training - Jul 2015. 76

Figure 22. Spatial distribution of the Bass estimation’s error by Jul 2015. 77

Figure 23. Hot spot analysis of the ANN estimation's error by the end of the training - Jul 2015. 77

Figure 24. Hot spot analysis of the Bass estimation's error by Jul 2015. 78

Figure 25. Spatial distribution of the ANN estimation's error – (a) First, (b) Third and (c) Fifth forecasted period. 79

Figure 26. Spatial distribution of the Bass estimation's error – (a) First, (b) Third and (c) Fifth forecasted period. 80

Figure 27. Methodology to characterise the evolution of the population heterogeneity, and assessment of MAUP effect. 92

Figure 28. PCs included in the different study scales for the multivariable models. 93

Figure 29. Implementation of the model in Anylogic software – (a) MV-Birmingham, (b) MV-Extended, and (c) MV-LADs. 94

Figure 30. Example of an artificial neural network fed by the autoregressive elements and the socioeconomic variables. 99

Figure 31. Cumulative adoption rates of PVs estimated by the MV-Birmingham and the Bass model vs. actual data. 103

Figure 32. Estimation error of the autoregressive and MV-Birmingham, and the Bass model. 104

Figure 33. The temporal pattern of the marginal changes of the estimation errors overtime for the MV-Birmingham model. 105

Figure 34. Spatial distribution of the impact of FiT on the adoption rates by November 2011. 106

Figure 35. Estimation errors for the multivariable and Bass models. 107

Figure 36. Spatial distribution of estimation errors - (a) Autoregressive model, (b) MV-Birmingham model. 109

Figure 37. Spatial distribution of the estimation errors - (a) MV-Birmingham; (b) MV-Extended; (c) MV-LADs. 110

Figure 38. Hot-spot analysis of the MV-Birmingham (a) and MV-LADs (b) model's MAPE by the end of the training - Jul 2015. 111

Figure 39. Estimation error for the forecasted periods - autoregressive model vs. multivariable. 112

Figure 40. Estimation error for the forecasted periods for the multivariable models. 113

Figure 41. Spatial distribution of the error of estimation. First forecast - (a) autoregressive model; (b) MV-Extended; Third forecast – (c) autoregressive

model; (d) MV-Extended; Fifth forecast - (e) autoregressive model (3rd forecast); (f) MV-Extended.. 115

Figure 42. Boxplot of the total EV registration in 96 PCs in the West Midlands. 120

Figure 43. Spatial distribution of the EV registrations by September 2018..... 121

Figure 44. Hot spot analysis of the EV registration at PC level by Sept 2018. 122

Figure 45. Monthly cumulative adoption rates EV registration for the 96 PCs.124

Figure 46. Spatial representation of the EV autoregressive model in AnyLogic software. 125

Figure 47. ACF of the EV registration data at the PC level. 126

Figure 48. PACF of the EV registration data at the PC level. 127

Figure 49. Decision-making process defined by the ANN, EV adoption process. 128

Figure 50. Cumulative adoption rates of EVs estimated by ANN and Bass model vs. actual data. 131

Figure 51. Estimation error for the ANN and Bass model for the EV data...... 131

Figure 52. Spatial distribution of the EV estimations of the (a) ANN and (b) Bass model by the end of the training period - December 2016....................... 133

Figure 53. Hot spot analysis of the EVs estimation for the (a) ANN and (b) Bass model by the end of the training period - December 2016....................... 134

Figure 54. Spatial distribution of the errors of estimation for the (a) ANN and (b) Bass model by the end of the training period – December 2016. 135

Figure 55. Hot spot analysis of the error of estimation for the (a) ANN and (b) Bass models error of estimation by December 2016 136

Figure 56. Estimation error for the forecast period ANN vs Bass model. 137

Figure 57. The cumulative adoption rate of PV installations and EV registrations in 96 PCs in the West Midlands, UK.. 142

Figure 58. Spatial distribution of (a) the EV registrations and (b) PV installations for 99 PCs in the West Midlands (Sept 2018)... 144

Figure 59. Hot spot analysis of the distribution of (a) the EV registrations and (b) PV installations for 9 local authorities in the West Midlands by Sept 2018. 145

Figure 60. Multi-output ANN design for the EV and PV adoption process...... 148

Figure 61. Estimations for the PV installations and EV registrations, for the 96 PCs.. 151

Figure 62. The error of estimation for the EV and PV series - MAPE. 151

Figure 63. Spatial distribution of the estimation error for the (a) PV and (b) EV, for the 96 PCs in the West Midlands. 153

Figure 64. Hot spot analysis of the estimation error for (a)PV and (b) EV, for the 96 PCs in the West Midlands. 154

Figure 65. Total PV generation injected to the network after self-consumption. 159

Figure 66. Spatial distribution of the total PV generation injected to the grid – (a) April 2017; (b) November 2017. 160

Figure 67. Total domestic demand during EV charging. 161

Figure 68. Spatial distribution of the total domestic energy demand during EV charging – (a) April 2017; (b) November 2017. 162

Figure 69. Spatial distribution of the combined impact of EVs and PVs. 164

Figure 70. Historical PV data at MSOA level and 6-month basis for Birmingham city. 226

Figure 71. Historical PV data at LSOA level and 6-month basis for Birmingham city. 227

Figure 72. Historical PV data at PC level and 6-month basis for Birmingham city. 228

Figure 73. Historical PV data at MSOA level and monthly basis for Birmingham city. 229

Figure 74. Bass model's estimations for the PV uptake in the Birmingham's PCs. 243

Figure 75. Bass model's estimations for the EV uptake in the Birmingham's PCs. 243

LIST OF TABLES

Table 1 Lists of keywords used for each research question 18

Table 2. List of modelling elements of ABMs 23

Table 3. The criteria to assess the strengths and weaknesses of different models 32

Table 4.Key characteristics of reviewed studies 34

Table 6. The Moran's I index value, z-score and p-value of the PV installations data 58

Table 5. Comparison of three implementations of ABM and ANN hybrid models. 61

Table 7. List of independent socioeconomic variables and their resolution. 97

Table 8. Results summary from the stepwise process of variable selection ... 101

Table 9. The Moran's I index value, z-score and p-value of the EV registration data 122

Table 10. Summary of OLS Results - Model Variables 146

Table 11. Forecasted MAPE for the autoregressive and combined models ... 157

Table 12. Summary of models' performance behaviour of the estimation errors. 176

Table 13. Reviwed articles by research question or snowball criteria 221

Table 14. Variables used in the modelling approaches 232

1 Introduction

The UK government has set ambitious energy policy and revised its greenhouse gases (GHG) emissions target to mitigate global warming. Via the *Net Zero* policy, the UK aims to reduce its emissions to net-zero, contributing to keeping global temperatures rise below 2C [1,2]. This is an amendment to the Climate Change Act 2008, which bound the UK to reduce its GHG emissions by at least 80% from the 1990 levels by 2050. This also required setting the Committee on Climate Change and five-years periodical carbon budgets. Then, because the transport and energy supply sectors account for 32% and 27% of the GHG emissions in the UK by 2018 [2,3], it is imperative to decarbonise such sectors. Therefore, electrification of transport via low carbon electricity produced from renewable resources will play a key role to deliver on this goal [1,2].

More specifically, it has been highlighted that electric vehicles (EVs) [4–6] and domestic solar photovoltaic panels (PVs) [6–8] will be crucial towards achieving the government's ambitious emissions targets whilst also reducing dependence on fossil fuels [4,5]. The future of a low carbon energy system has been described in a number of scenarios that assess technological feasibility or overall trends, and develop pathways of requirements to achieve a specific goal (i.e. reduction of GHG) [9–11]. The main limitation of these scenarios is that their underlying modelling approaches overlook the complex network of interactions between the different actors that the energy system comprises [5,9]. Moreover, because these approaches focus on national-level analysis, they are limited to inform about the local evolution of these technologies' regularities and how the adoption of EVs and PVs may impact on the distribution network. As mentioned by a growing number of authors [8,12–21], the adoption patterns of EVs and PVs present spatial regularities, and that a high geographical concentration of these technologies may cause issues on the low voltage lines. By creating reverse flows, solar PVs can diminish predictability of load, voltage and demand flows [22–24]. Uncontrolled charging of EVs may intensify the stress in the distribution networks and cause faults and power cuts [25–27].

Consequently, the adoption of EVs and PVs will shape the evolution and characteristics of the energy system [28] and creates a challenge for the management of the distribution networks. That is why the development of the tools and methods to predict where these technologies will appear and at what pace they will evolve is important [8,29]. Yet, the development of these technologies is highly uncertain as the adoption of EVs and PVs is driven by subjective factors such as perceived affordability [30], social influences from other individuals and one's neighbourhoods [8,20], or personal believes [31]. The adoption process is also affected by objective factors like income [12,19,25], energy cost [16,25], fuel cost [32,33], available infrastructure [25,32], or available policies [32].

Moreover, these two technologies present clear differences. Solar PVs require a one-time effort when installing, and provides an intermittent generation profile [34]. On the other hand, EVs are an alternative technology for mobility services. Depending on the charging state of the battery, they may require the individuals to change their driving behaviour [35] which may potentially result in new charging patterns. Regardless of these differences, emergent literature points out that there are empirical regularities between EV and PV ownership [14,26,36], highlighting these technologies are adopted by the same kind of consumers (those with a high environmental concern) [34]. However, the population growth theory presents limitations when accounting for (i) the spatial regularities; (ii) explicit temporal dynamics; and (iii) the variety of factors that drive the adoption process [37].

Despite these limitations, this theory has been widely applied to the diffusion of innovation. More specifically, three implementations of this theory have been used to characterise the adoption of EVs and PVs [38]:

- The Logistic model is used to forecast the adoption of innovation, using the population growth theory. The model considers three elements, (i) the number of adopter at a given time, (ii) the potential adopters, and (iii) the coefficient of adoption [39]. This adoption model has been extended into the following two models [40].

- The Bass model characterises the market diffusion of EVs and PV assuming that these technologies are adopted by imitation or media-effect influence [40,41]. Al-Alawi and Bradley [40] mention that the Bass model is the most extensively used approach for the adoption of innovation.
- Lastly, the Gompertz model estimates the count of EVs or PV, by fitting the time series from the EVs and PVs sales to a saturation curve (Gompertz distribution) [40,42].

The modelling of EV and PV adoption has experienced changes in the last decade, adopting new approaches that consider the overlooked complexity of actors interacting in the energy system or trying to explain the factors driving the adoption process using aggregate data. Also, as demonstrated by this research, there is potential in integrating elements of the human cognition into the modelling approaches. The review of the literature presented in Section 1.2 identifies three recurrent contemporary approaches. Namely, the agent-based modelling (ABM), part of the complexity science discipline, characterises the individual decision-making of whether or not to adopt such technologies, as the result of the social interaction of the agents [9,10,43–45]. The spatial regression[1] (SR), part of the econometric techniques, seeks to understand the effect of the different factors that drive the adoption process, whilst considering the spatial regularities of the adoption process [8,12,16,19,46,47]. And the Poisson model that is used extensively in epidemiology characterises the adoption of EVs and PVs as the product of stochastic social interaction among individuals [48–51].

This research adopts the ABM modelling approach because of its capabilities to characterise the complexity of the energy system evolution as a social-technical process [52,53]. Indeed, the applications of the ABM to analyse the adoption of EVs and PVs have produced useful insights on how the influence of the social interaction caused by the interaction between individuals drive the evolution of the energy system [32,43,53]. Section 1.3 presents the strengths and weaknesses of the ABM and the other two modelling approaches in detail. For instance, most of the existing applications of ABMs have an explorative nature

[1] The spatial regression will be referred as SR from now on.

that focuses on the emergent behaviour of the agents. However, these ABMs do not consider the spatial regularities of the adoption process as represented by the SR models, which in contrast disregards the temporal nature of the adoption process. Thus, these approaches have limitations to inform about spatially and temporally explicit estimations of adoption rates of EVs and PVs. In recognition of these limitations and the need for spatio-temporally explicit estimations of the adoption rates of PVs and EVs, to inform network planning and management, the following initial research question formulated:

To what extent is it possible to characterise the adoption process of EVs and/or PVs, whilst integrating the spatio-temporal regularities and the different factors that drive the adoption process?

The following section carries out a literature review, with the view of informing about the current modelling approaches for EV and PV adoption, whilst identifying the gap in knowledge. Then, this refines and reformulates the research question as a hypothesis and define specific objectives to test it.

1.1 Literature review

This section reviews the literature by breaking down the initial research question into smaller and more specific questions. Firstly, as shown in Figure 1, the review focuses on the combination of EVs **OR** PVs studies (blue and yellow circle), because focusing only in the intersection of EVs **AND** PVs studies (green intersection) may exclude the insights from the single-technology studies. Hence, this review undertakes an analytical review of the contemporary EV and PV market diffusion approaches that consider spatio-temporal regularities of the adoption process. As illustrated in Figure 1, the review aims to identify their modelling strengths and weaknesses, along with the different factors that drive the adoption process, and how (if so) the relationship between their ownership has been modelled (Section 1.3). Consequently, the review also informs about the modelling elements that future modelling approaches should consider. Lastly, the review identifies knowledge gaps and proposes suitable approaches to address the limitations of the current modelling approaches and test the thesis hypothesis (Section 1.5).

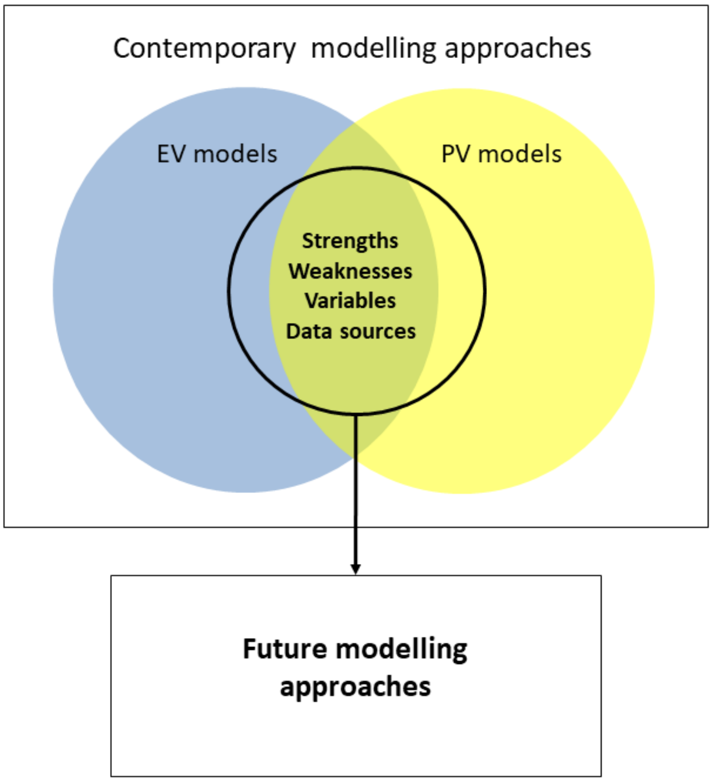


Figure 1. A conceptual framework for the literature review.

1.1.1 Review method

A systematic approach is followed because this review seeks to identify the gap in knowledge, whilst defining specific objectives. The systematic reviews are argued to be transparent to the reader and replicable [52–54]. Narrative reviews are ideal for broader research questions, and usually do not require clear inclusion nor exclusion criteria. Thus, findings produced by narrative reviews are potentially biased, as the review evaluation is not uniform [52–54]. Therefore, this review builds upon Denyer and Tranfield systematic review methodology [55] and expands it further by considering other studies on the design of query, construction of database and presentations of the results [56–62]. The systematic process is summarised in Figure 2 and fully discussed in the following sections.

As the spatial patterns of EV and PV diffusion are newly developing fields, the systematic review yields only 24 relevant studies (see section 1.1.4). That is why, following Adams et al. [63], the *Snowballing* technique is incorporated into the search method. Snowballing is a forward and backwards process that retrieves relevant articles cited by the preliminary search and increases the reliability of the review [64,65]. This technique allows also to keep the list of references up to date (04 Dec 2018) [66].

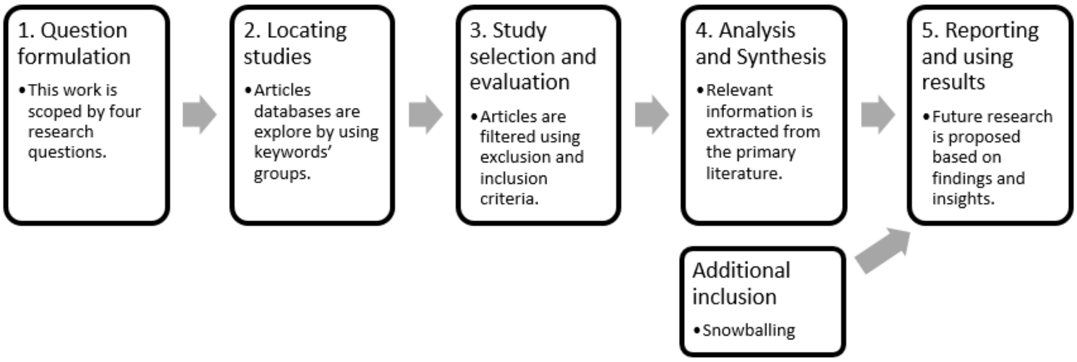


Figure 2. The methodology followed in carrying out a systematic literature review.

Source: Adapted from Denyer and Tranfield [55] and integrating Sayers [64].

1.1.2 Literature review questions

As presented in Figure 2., firstly, a base of articles is created using research questions (RQ)1 and RQ2, scoping the review. The next three research questions build on this research base which includes the articles added via the snowballing method. RQ3 focuses on assessing these models' strengths and weaknesses, whilst RQ4 identifies the variables and data sources used. Finally, RQ5 focuses on how to build on the strength of these alternative modelling approaches to guide future research. Research questions are defined as follows:

- RQ1: How do different approaches used for modelling of EV adoption compare and contrast in terms of spatial and temporal dependence and social effects?

- RQ2: How do different approaches used for modelling of PV adoption compare and contrast in terms of spatial and temporal dependence and social effects?
- RQ3: What are the strengths and weaknesses of different approaches to account for spatial and temporal dependence and social effects?
- RQ4: What variables and data sources have been used to characterise the adoption process of EVs and PVs?
- RQ5: What an alternative modelling framework can be developed to address the limitation of current modelling approaches?

1.1.3 Localization and identification of relevant studies

A list of keywords is created for RQ1 and RQ2 to ensure any relevant study is not left out. The literature review assumes that the studies under the *adoption, diffusion and penetration* keywords consider the temporal variable already, without excluding those that do not. These lists are used to design the search queries. Groups of keywords used are listed in Table 1. Four online databases, Web of Science, ScienceDirect, Scopus and Google Scholar are explored to identify relevant studies. Search queries are designed through the combination of the keywords and logical functions.

Table 1 Lists of keywords used for each research question

RQ1:	EV, electric vehicle, electric car, adoption, diffusion, penetration, space, spatial.
RQ2:	PV, solar panel, solar photovoltaic, adoption, diffusion, penetration, space, spatial.

1.1.4 The screening and selection of relevant literature

The review process is characterised in a funnel diagram in Figure 3. During the first stage, a total of 10,385 document results are retrieved from the online databases The second stage applies the exclusion criteria, as described below and reducing the numbers by two thirds. The next stage involves screening the article's title and abstract, leaving only 1% of the original sample.. At this point, the rest of the articles are fully reviewed, resulting in a total of 24 studies. Finally, by using the *snowballing* criteria, another 18 articles are added to the number of studies included in the review. Given that the analysis of the spatio-temporal patterns of EVs/PVs adoption is a recent and emerging field, it was found that most of the reviewed studies were published from 2010. The list of studies included in the review is provided in Appendix 1.

Exclusion criteria

- Duplicated articles across different online databases are omitted.
- The material is not related to spatial distribution.
- It is related to Physics or Chemistry sciences, e.g. fuel performance, surface analysis.
- It is related to engineering analysis or design, e.g. battery performance or composition, molecular dynamics.
- It is related to broader environmental issues, e.g. air pollution, climate change, ecosystem services.
- It is not a specific article, but a journal or compendium of materials, i.e. journals abstract, indexes of conferences.

Inclusion criteria

The snowballing technique is included as additional screening, which is carried out in a forward and backward direction as listed below. The process starts with the studies resulting from the exclusion criteria and stops when the review becomes cyclical. In other words, it stops when the potential new articles have been already screened (included or excluded) and there are no more new studies.

- **Forwards**: papers which cite an already selected article are analysed. These articles were screened considering the Exclusion criteria.
- **Backwards**: studies cited by an already selected article are screened also. Initial exclusion criteria are applied to these studies.

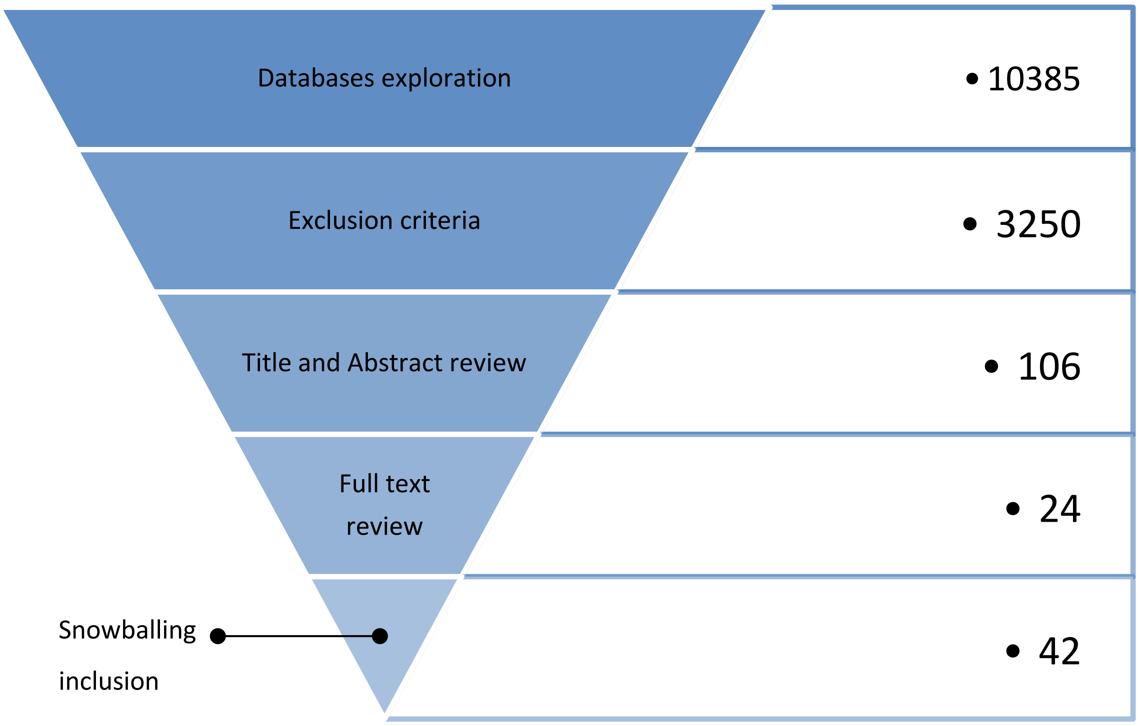


Figure 3. Funnel diagram of selected articles that are included in the review.

1.1.5 Analysis and synthesis

From the selected studies, there are four relevant elements to assess the model's strengths and weakness: First, the model's capabilities to capture the spatial dependency of the adoption process. Second, the model's effectiveness to integrate the different factors driving the adoption process. Third, whether the model can integrate the relationship between EV and PV ownership. And finally, the analysis also looks at the model’s goodness of fit to score and identify the best suitable approach.

1.2 EV and PV adoption modelling approaches

As shown in Figure 4, the number of studies has been increasing in the last five years. Most of these studies are published in the following journals: Energy, Technological Forecasting and Social Change, Renewable and Sustainable Energy Reviews, Energy Policy, Environmental Innovation and Societal Transitions. In general, there is a similar number of studies for EV and PV technologies. At the beginning of the 2010s EV, studies are slightly greater, but in recent years, the academic interest in PV articles has increased, resulting in more PV than EV studies. The use of snowballing criteria meant the inclusion of other technologies like environmental and energy innovations and heating systems in the analysis, though there were a few numbers of articles focusing on them. Then, Figure 5 shows the distribution of the modelling approaches used. There are three main recurrent approaches: The agent-based modelling, the spatial regression, and the Poisson model. Besides these models, there are other models such as the Discrete choice, Logistic regression or Network theory. Yet, because there are a few applications of these approaches, they are presented broadly in Section 1.2.4.

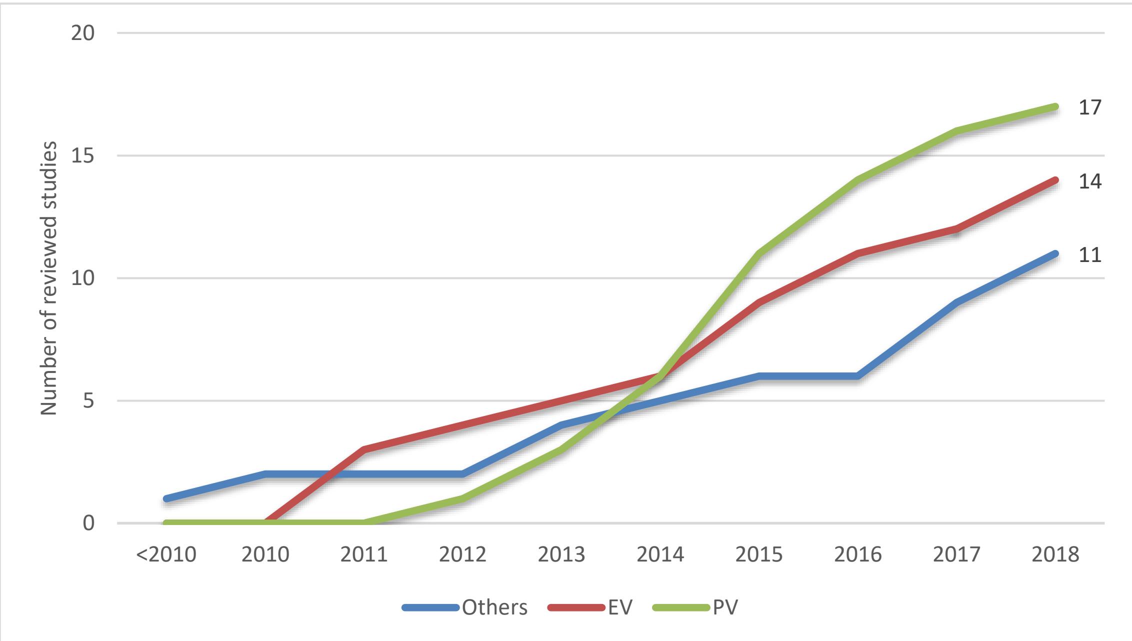


Figure 4. Annual cumulative number of publications by type of technology modelled.

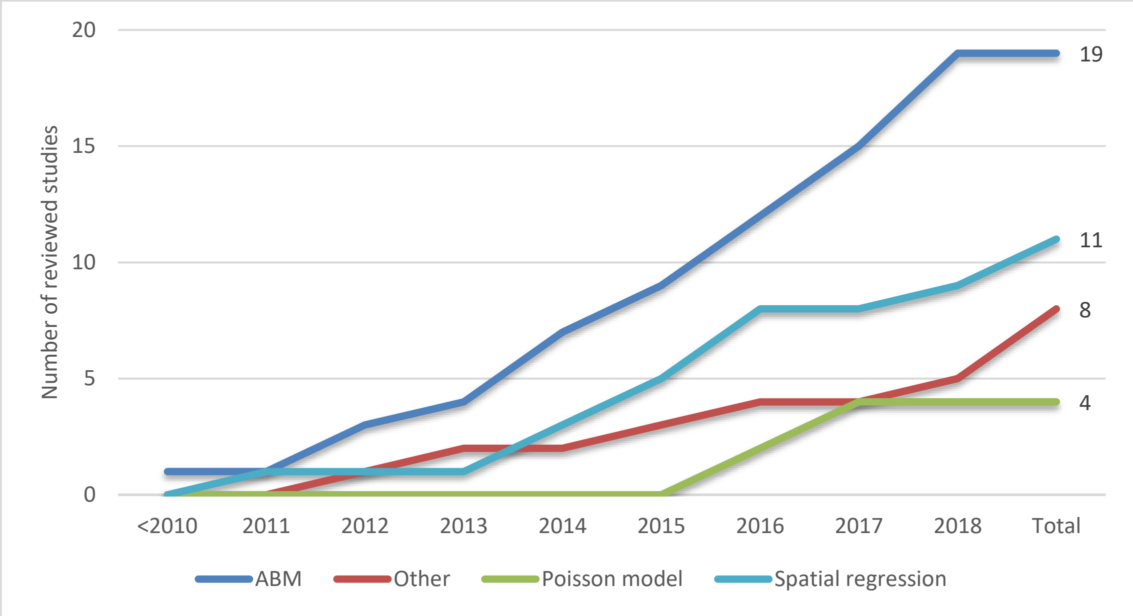


Figure 5. Annual cumulative number of publications by type modelling approach.

1.2.1 Agent-based modelling

The ABM has become one of the most used techniques to analyse social-technical systems, as this allows the modelling of entire systems by characterising the individual actors that comprise them [67,68]. The ABM simulates the interaction of these actors through physical and social networks, then, the evolution of the system is the results of the cumulative decisions taken by the agents [32,43,68]. The applications of the ABM comprise air traffic management, biomedical and epidemic research, crime analysis, stock market analysis, organizational decision-making, ecology, energy modelling and diffusion of innovation and adoption dynamics [69,70].

Because the methodologies to design and implement an ABM are as extensive as the number of applications for it, these studies are reviewed using an ABM development protocol. The Grimm et al. [71] protocol is a widely used standard protocol. This methodology is a comprehensive framework that allows mapping the: **The entities** that represent the agents' definition. **The object** is the technology or behaviour introduced into the system. **Characterisation** refers to the techniques to assign individual features to the agents, this also includes. **Their location** of the agents in the virtual environment. **Decision-making** that defines the agents’ behaviour and the rules to modify such behaviour. **The social effect** refers to how the agents interact. Table 2 summarises the different elements used to implement the reviewed ABMs, further Section 2.1 presents in detail these modelling elements, focusing on the empirical implementation.

Table 2. List of modelling elements of ABMs

Modelling elements	**Description**
Entities	Individuals or households
Object	EVs, PVs and heating systems, Energy saving practices
Characterisation	Empirically if the data comes from a survey Semi-empirically in the case of simulating or calculating any value based on other current data such as the census And assigning features randomly
Location	Same as above
Decision-making	Financial benefits (utility) and/or on the social benefits
Social effects	Direct contact with other agents (peer-effect) or from the influence of collective behaviour (social norms).

The ABMs commonly work under the assumption that the agents can evaluate the benefits of their decisions or that these follow rule-based criteria, which is defined by the modeller. However, this characterisation of decision-making is limited. First, because the assessment of those drivers assume that the agents possess the perfect market information, which is rare, and secondly because some of the drivers such as personal believes and preferences have subjective values [43,72–74]. The decision-making of the agents is usually not implemented temporally or spatially explicit simultaneously; from now on this is referred to as spatio-temporally explicit characterisation. However, the literature offers examples of ABMs that analyse those elements independently.

The ABMs representation of the temporal variable gives those models a predictive or explorative nature. The predictive models explicitly characterise the temporal variability, as the models aim to replicate the actual behaviour of the agents over time. Then, during the simulations, the models search for the combination of parameters values that best fits historical data of adoption rates. Each of those combinations is to produce a fitness parameter or accuracy level, which in turn defines which is the best mode. Krebs [75] uses PV data at a monthly basis, Robinson and Rai [76] (94%) and use time-series in quarterly basis, whilst Adepetu and Keshav [68] (95%) analyse the adoption of PV in a 6-month basis. Instead, those ABMs with an explorative nature don't present an explicit time horizon but are run until new behaviours emerge. For instance, McCoy and Lyons [26] run the simulation 100 times and report the average of the results, or Noori and Tatari [49] who run their model for 10,000. In both cases, the authors rather explore the results of the model under each of the combinations of each possible value of the parameters. Thus, because the simulation lack of an explicit time horizon, the models have limited applicability for forecasting future rates of adoption.

In general, the spatio-temporal regularities have not been yet integrated into the decision-making process. Yet, some of the reviewed ABMs studies aggregate the data to analyse the spatial validity of the model, showing that the models have limitations to adapt to changes in the data trends [75,77,78]. Moreover, because the ABMs work at the individual level, the ABM requires lots of data to empirically characterise individuals. An empirical characterisation would require having data for each of the households to keep the geographical layout of the agents. In this sense, the ABM is limited to keep the spatial layout because of the lack of data to represent the whole population [51,79]. Consequently, the ABM studies rather use a sample of the population or escalate empirical data up to reaching the actual population numbers.

Additional to the ABM, two other models were identified by the literature review, which shares characteristics with the ABM. First, Bale et al. [79] analyse the adoption of more energy-efficient technologies using the Network Theory evaluates the total utility of adopting based on personal benefits, the average value of adopters in the individual social network, and the average of adopters in the whole population. Yet, the adoption criteria are based on an equation that weights the former elements [79]. Then, Dimatulac and Maoh [80] use the Discrete choice model to characterises the adoption process of EVs at the individual level. The model substitutes the utility or social threshold, the adoption criteria is a probabilistic function that considers the available options of vehicles. The model also partially implements the neighbourhood effect, by implementing the count of EVs in a specific zone for each of the individual in that zone. Then, the decision-making of the agents within a zone will be informed by the number of EVs in the adjacent areas. However, this effect disregards the diminishing effect of the distance between zones, neither considers the influence of the agents in the same zone.

The Network theory and Discreet choice model may be seen a simplified version of the agent-based model, as they have similar components such as the adoption criteria in the form of an equation, and the utility and social threshold as fitted parameters for each case study or probability function. Similarly, both approaches present the same limitations as the ABM not to account specifically for the temporal dimension and disregard the spatial dependence of the adoption process.

1.2.2 Spatial regression

The SR has been used to understand the effect of socioeconomic variables on the adoption of EVs and PVs, whilst considering the spatial regularities of those variable and the EV/PV distribution. Alike classical regression, this method analyses the effect of the dependent variables on the variable of interest. At an aggregated level, the spatial regression model considers that this effect on dependent variables varies as a function of distance. In general, the authors conclude that there is a significant relationship between the adoption rate in a local area and its surrounding neighbourhoods. The spatial regression model can be defined as follows:

$$= \rho W y + X\beta + \varepsilon \tag{1-1}$$

Where,

y is the dependent variable

ρ is the autoregressive coefficient[2]

W is the spatial weights matrix

X is the vector of independent variables

β is the vector of regression coefficients

ε is the stochastic error term

The dependent variable y is the cumulative number of objects in a specific location. Authors use two cumulative measurements, the total/absolute count and the relative count of a specific technology (adopters) in a specific location. Most of the authors apply the total cumulative count of objects, except for [12,18,29], who acknowledge the disparity in locations' size and population density. Langheim [12] normalises the number of PV installations by considering the number of single-family buildings, so the model considers the market size in each location. Schaffer and Brun [18] acknowledge the variation in locations' size, so

[2] This reflect how much the data is spatially correlated, because this is a multidimensional index (accounting for distances in deferent directions), its values are not limited to [-1,1].]

they calculate the total PV systems divided by the location's area. Most of the authors include population density as an explanatory variable. The last exception is Morton et al. [29], who model the number of EVs per thousand of cars.

A similar application approach to the SR is the Geographically Weighted Regression. The model locally weights the variation of the relationships between the dependent and independent variable, rather than the overall spatial regularities as the SR [12]. Yet, the model is limited to account for the temporal non-stationarity[3] [81], which is a characteristic for the PV or EV adoption process (as further explained in Section 2.2.3 and 4.2.3, respectively).

Then, the independent variable *W* is a matrix of spatial lags, reflecting the distances between locations. This matrix has been characterised in different ways; which may affect the magnitude of the correlation, albeit not significantly in some instances [8]. Badi and Bartlomiej [82], who carry out a sensitivity analysis on the spatial correlation by changing the way the matrix weight is defined. They find that the magnitude of the spatial autoregressive coefficient changes up to a 20%, but in all the cases the direction of the relationship does not change (i.e. from -.047 to -0.58). Authors commonly calculate matrix weights considering the actual distance between two locations (Euclidean distance). On the other hand, an alternative to the Euclidean distance is the adjacency approach, which reflects whether two locations are physically contiguous or not. An alternative definition is to consider the closest locations as contiguous locations [18,19].

The characterisation of the dependent variable and the spatial weights matrix vary across the studies, the former is discussed in detail in Section 3.1.1. The latter is closely related to the analysis' definition of geographical area, as the number of locations and the distance between them depends on the spatial resolution. Frequently, these authors use different geographical areas at the regional level, which could be standard or local definitions. For instance, the Nomenclature of Territorial Units for Statistics NUTS3 or local level, which is the standard geographical definition for the members of the European Union. Such a

[3] A stationary timeseries is assumed that its statistical properties such as mean, variance, etc. are all constant over time.

definition allows comparing statistics between countries, yet this definition might not correspond to a local definition or political division. On the other hand, local definitions such as census tracks or administrative definitions can be used depending on the objective and scope of the study, which might depend on the local government.

For example, highlight the lack of socio-economic data at smaller spatial units than the regional level. On the other hand, bigger units of analysis may exhibit loss of spatial heterogeneity as data are aggregated. Yet, across Europe, the use of NUTS3 classification is not uncommon (e.g. [18]). On the other hand, a higher resolution such as ZIP codes, block-groups or census tracts, are also used as spatial units. Authors use census data to match the total PV installations and aggregated socio-economic data [12,16,19,46,47]. Overall, authors agree on the spatial regularities in low carbon technology adoption patterns, even though the studies differ from each other about the magnitude of each variable's influence, due to the context and variables included.

A limitation of the spatial analysis is the fact that it overlooks the temporal dynamic of the diffusion process [8,18,29]. Authors such as Graziano and Gillingham [16] alternatively introduce a temporal lag, considering that the spatial effect decreases with the distance and time. Similarly, Richter [20] develops an econometric model at a high spatial resolution, using the PC layout as well. Then, using the postcode resolution the author characterises the social influence as the effect within each area through the lagged number of PVs. Yet, there is no influence from one area to another. Other authors like Morton, Wilson and Anable [47] replicate their analysis for different points in time. Although the temporal dynamics of the adoption process are not considered, the authors present a series of independent snap-shot results. Also, because the statistics values of the models vary from one spatial resolution to another, the change in the study's scale may introduce bias to the study, as any statistics is dependent on study scale and spatial resolution [15,72,83,84], which is known as the Modifiable Areal Unit Problem (MAUP). Therefore, Balta-Ozkan, Yildirim and Connor [8] propose working with a finer resolution (more local), as these spatial differences are important for the application of local policies. Besides Noonan, Hsieh and Matisoff

[46] point out that the aggregation process may imply a loss in spatial heterogeneity.

1.2.2.1 Social effects

The social effects are also considered by the SR and the Poisson model (discussed in the following section) as the neighbourhood effect and the peer-effect. The neighbourhood effect comprises the spatial effect of the surrounding location, the effect of the total adoption rate (alike the social norms of the ABMs), or the exposure of surrounding locations. On the other hand, the peer-effect works at the individual level and is present only in some Poisson models and the ABMs, which characterises the mechanisms of these influence. Contrary, because of the spatially explicit nature of the SR and the Poisson model, these modelling elements inform about the presence of spatial effects and how these decreases with the distance, even though the distance is characterised in different ways. One can argue that this distance can be compared with the techniques used to associate agents (e.g. similar income, similar socio-economic characteristics, random connections, etc.) in the ABMs.

1.2.3 Poisson model

This approach assumes that there is no explanation for any decision taken during the adoption of innovation. Instead, the individuals' interaction is not fixed to a particular network, and the decision-making criteria are not deterministic, a property shared with the individual level models that these elements are still unknown to the modeller. There are three modelling considerations. First, there are a limited number of contacts made at each period. Second, it is not essential to know the specific individual’s network to understand the general diffusion process. Finally, the last consideration is the degree of randomness, which depends on the scale of the study, as defined by the modeller. Depending on the focus of the study the modeller defines whether the model reflects individuals’ or small-groups’ behaviour. The microscale (individual) can potentially assess the probability of adoption. On the other hand, macro-scale (group) studies may provide a general adoption pattern.

The Poisson model assumes that the drivers of the decision-making process are unknown. Then, the adoption rate is defined through spatial exposure to the innovation and a demographic variable [51,85]; the latter is characterised as the employed population density or housing density. De Groote, Pepermans and Verboven [51] argue to explore the temporal dynamics of the adoption rate as its drivers change over time. They acknowledge that as more detailed data become available, the models can produce more robust results. Similar to the ABM, the Poisson model also takes into account the peer-effect. While the ABM considers the structure of the social network, the Poisson model reflects this via exposure to the innovation. Yet rather than to specify a social network that defines the interaction between individuals, these studies define a probability of contact (e.g. how many contacts occurs among the population). The authors characterise the threshold of adoption as dependent on the number of adopters in the adopters' social network [48–50].

The Poisson model includes a temporal variable, yet, this is a reference to the simulation period rather than the actual time horizon. This synthetic temporality allows investigating which factors drive the timing of adoption [51] as well as studying the adoption rate over time. Adjemian, Lin and Williams [19] define a virtual population which adopts a specific vehicle from a set of options and maximises the utility from it. Instead of a distance variable, they identify cumulative variables with the five closest locations. Bansal, Kockelman and Wang [85] models the spatial diffusion of EVs by applying a Poisson-lognormal model, to quantify the spatial effect of lagged independent variables. They use census data for Texas (US) to model the spatial diffusion of EVs. For the spatial effects, the authors define contiguous neighbours if two (or more) blocks share a border, similar to [18]. De Groote, Pepermans and Verboven [51] model the PV spatial diffusion with a Poisson distribution. They calculate the specific factor of adoption (elasticities) for each neighbourhood and then estimate the spatial effect of those variables. They use the covariance of the dependent variables to construct these coefficients.

1.2.4 Other modelling approaches

There are a few other modelling approaches employed among the reviewed articles, for instance, the Logistic model or hybrid models. The Logistic model overlooks the spatial dimension of the diffusion process, yet, the authors overcome this limitation by creating “buffers” per areas. These “buffers” accounts the number of PVs in different radius, to capture the spatial regularities. [86]. The Logistic model has also been used to calculate individual regressions for different areas [17]. Yet, the spatial dependence is not considered, as each of the regressions is independent. As well, the Bass model [41] has been used to characterised the EV adoption as the result of interactions between individuals. The interactions are defined considering the local population size and the share of adopters at each neighbourhood. Additionally, an influencing factor of media coverage has been included as the imitation parameter of the Bass model [87].

On the other hand, Zhao et al. and Zhou et al. hybrid models [88,89] define the adoption process locally using artificial intelligence and the Bass model principle. The models use cellular automata to identify and classify the PV saturation in each area using historical data. Then, the individual estimations are generated using a fitted to time-series S-curve. These approaches also consider the estate of adjacent areas during the classification of the areas; however, the effect of the neighbouring areas is not weighted against the distance between them. Additionally, as noted before the Bass model has limitations to considers social nor spatial effects.

1.3 Review of strengths and weaknesses of the modelling approaches

The review questions aim to identify four modelling elements: (i) spatial regularities, (ii) temporal regularities, (iii) social dynamics and (iv) the relationship between PVs and EVs ownership. However, from the reviewed methods, it’s noted that these are not completely independent or exclusive from one method to another. Firstly, as noted by the SR studies, the modelling of the adoption process requires to include the time variable, as there is a delay between the decision to adoption PV and the actual installation. Therefore, future modelling approaches should include spatial and temporal regularities of the adoption process. Furthermore, alternative methods are required to reflect the spatio-temporal regularities of the different factors that drive the adoption of such technologies. Then, from the reviewed articles, it is noted that despite the wide number of variables included in the analyses, the social effects are a constant factor that authors recognise to be critical. The entire group of reviewed studies implemented the spillover effect, either in the form of the neighbourhood effect [8,16,29] or social norms [25,79,87]. On the other hand, only 30% of the studies characterised the peer-effect [14,25–27,32,33,50,51,68,75,77,78,90–93]

On the other hand, the empirical regularities between the EV and PV ownership have not been empirically tested in-depth, meaning that the influence of adopting one technology on the decision-making towards the other is still unexplored. Thus, future approaches may empirically and explicitly test the regularities between the ownership of these technologies. In the next section, the reviewed methods are assessed in terms of these modelling aspects. As it is argued that some there is knowledge to be transferred between some of these approaches, the following section assesses the strengths and weaknesses of the reviewed models. Therefore, the review informs about future modelling approaches.

The studies analysing the modelling of PVs and EVs are assessed in terms of the following four aspects: (i) spatial regularities, (ii) temporal regularities, (iii) social dynamics and (iv) the relationship between the ownership of PVs and EVs. These features are assessed following the criteria shown in Table 3, where scores are defined from [1-5] depending on the capabilities to explicitly characterise each of the elements.

Table 3. The criteria to assess the strengths and weaknesses of different models

Score	**Description**
1	The model disregards this modelling element.
2	The authors have implemented a proxy for this element.

Score	Description
3	The model has limitations to characterise this element explicitly.
4	The model can characterise this element explicitly but uses simulated or semi-empirical data to implement it.
5	The model can characterise this element explicitly and empirically.

Table 4 summarises of the assessment of the reviewed models, which are sorted considering the scores for the three elements. Additionally, the models' fitness is reported (if provided), typically, this evaluates the differences between model estimates and the actual data; however, this is not reported in all the models.

On the one hand, the SR and the Poisson models are suitable to represent the population heterogeneity and capture actual spatial layouts. Nevertheless, these tend to oversee the time variable and to have marginal fitness. On the other hand, the ABMs approaches have limitations to represent the heterogeneity of the populations (mainly because of data availability, as mentioned in [51,79]), this comes also with limitations to represent the actual spatial distribution of the agents. Yet, these studies can handle the time variable and they exhibit greater fitness than the aggregated approaches. An exception to this is the ABMs without a defined time horizon, which are more of explorative studies than actual forecasting methods (therefore the lack of a fitness indicator). Furthermore, these models do not consider the evolution of the agents' characteristics nor spatial dependency. Furthermore, it is noted that when integrating a measure of spatial regularities to the ABMs, models tend to perform significantly worse [76].

Table 4.Key characteristics of reviewed studies

Modelling approach	Ref	Obj.	Spatial regularities	Temporal correlation	EV and PV relationship	Fitness
Spatial regression	[14]	PV	5	1	5	55%
ABM	[26]	EV	3	2	5	-
ABM	[75]	Engy Inn.	4	5	1	98%
ABM	[76]	PV	4	5	1	94%
ABM	[91]	PV	4	5	1	96%
ABM	[78]	Engy Inn.	4	5	1	-
Poisson model	[92]	PV	5	+	1	40%
ABM	[77]	Engy Inn.	3	5	1	-
ABM	[90]	PV	3	5	1	94%
ABM	[25]	EV	3	5	1	-
ABM	[49]	EV	4	4	1	-
ABM	[93]	PV	4	3	1	-
ABM	[27]	EV	4	3	1	-
Mathematical	[94]	PV	4	3	1	-
Poisson model	[51]	PV	5	2	1	5%
Bass model	[41]	EV	5	2	1	-
ABM	[50]	Heating	3	3	1	-
Geographical regression	[12]	PV	5	1	1	54%
Spatial regression	[8]	PV	5	1	1	75%
Spatial regression	[15]	PV	5	1	1	61%
Spatial regression	[95]	PV	5	1	1	-
Spatial regression	[18]	PV	5	1	1	71%
Spatial regression	[29]	EV	5	1	1	72%
Spatial regression	[46]	Engy Inn	5	1	1	60%
Spatial regression	[19]	Automobile	5	1	1	-
Poisson model	[85]	EV	5	1	1	-
Poisson model	[96]	EV	5	1	1	-
Hybrid model (ANN)	[88]	PV	5	1	1	94%
Hybrid model (ANN)	[89]	PV	5	1	1	90%
ABM	[68]	PV	1	4	1	95
ABM	[33]	EV	3	2	1	-
ABM	[32]	EV	3	2	1	-
Network theory	[79]	Engy Inn	3	2	1	-
ABM	[48]	EV	3	2	1	-
Discrete choice	[80]	EV	3	2	1	-

Modelling approach	Ref	Obj.	Spatial regularities	Temporal correlation	EV and PV relationship	Fitness
Mathematical	[87]	Engy Inn.	1	3	2	-
Logistic regression	[86]	PV	3	1	1	-
Statistical analysis	[17]	PV	3	1	1	-

*Env. Inn: Environmental Innovation; Engy Inn: Energy Innovation.

1.4 Review Limitations

It is acknowledged that the review methodology uses a strict list of keywords and that the screening of studies by a single reviewer might have had an effect on the results [97]. Yet, the snowballing criteria improve the reliability of the review and fallow a systematic update of the emerging literature into the analysis [64,65]. Finally, it needs to be considered that this review may have introduced two types of bias into the results. First, the review may be subject to a publication bias, which implies that it is more likely for those studies with results in a preferred direction are published than those without relevance. Thus, the available (positive) papers may lead to over-overstatement of the results [98]. Secondly, the interpretation bias states that researchers with different backgrounds and/or objectives/interests may assess and discuss the same result in different ways; thus, they may draw opposite conclusions and insights [99].

1.5 Discussion

This section answers the review question and produces three main insights to inform the thesis research question. First, the modelling approaches for EV and PV diffusion are summarised. Then, potential avenues for future research are presented, focusing on how the ABM can be refined to account for the spatio-temporal dynamics of the adoption process. Finally, this section discusses the evidence to implement the influence of one technology over the other.

Modelling approaches for analysing the EV and PV adoption patterns

Three common modelling approaches that characterise the spatial patterns of EV adoption are presented in Section 1.2: spatial regression, Poisson model and ABMs. However, there is a clear difference in the share of these, being the ABM that accounts for 58%, whilst the SR or the Poisson model account for less than 20% each. This suggests that the spatial nature of the EV adoption still understudied. Thus, the need for further understanding of the spatial dependence of the EV adoption (and drivers).

Like the EV studies, the PV spatial diffusion patterns are also studied by these three common approaches. However, the number of SR and Poison model is more extensive for PV than for the EV, consequently, there is a potential for the PV studies to inform the adoption of EVs. The spatial analysis of PVs has highlighted the degree of spatial dependence and neighbour effect, the socioeconomic characteristics that drive de adoption, and the impact of incentives and rebates. Yet, the spatial analysis of the EV adoption is still emerging.

In this regard, this systematic review only retrieved one spatial study on EVs, Morton et al. [29], and to our acknowledgement, no other study has been published as of 04 Dec 2018. Although this study contributes to understanding the drivers of EV adoption in the UK and the spatial patterns of adoption, the model's data has limitations. Low data availability and lack of consistency in the time frame (vehicle registration for 2016, income data for 2015, other socioeconomic data for 2011), introduces a bias into the results. These limitations result in a moderate estimation accuracy (72%), which suggests that some of the spatial dependence is still not explained by the model. Other insights that may be integrated into Morton et al. SR study are from Adjemian et al. [19], who analyse the car ownership and its spatial dependency, highlight that the consumers' decision is driven by their peers' car ownership decisions [19]. This can be integrated into Morton's modelling, which only considers the neighbourhood effect.

The temporal dimension is recognised to be essential on the modelling of EVs and PVs adoption. On the one hand, approaches such as the SR try to integrate this in creative ways [16,20,95], such as including the time-lagged total number of these technologies, see 1.2.2 and 1.2.3. On the other hand, the ABMs have the potential to integrate this dimension, as the agents make decisions in discrete time steps, but this potential has not been exploited yet.

Explicit spatio-temporal characterisation of the decision-making

The spatio-temporal nature of low carbon energy transitions (e.g. [13,100]) necessitates developing methods to embed these regularities in the modelling framework. Furthermore, as the data availability is recognised as a constraining factor for the empirical characterisation of ABMs, it could be argued that the aggregation of agents in small geographical areas is a suitable alternative approach to the common individual ABMs. Indeed, there are a few numbers of ABMs that characterise agents as geographical areas or group of individuals whose behaviour is characterised as a singular decision-making unit, including Bierkandt et al. [101] and Kunz [102]. The former considers the production and consumption sites as agents, disregarding the individual characteristics and instead considering total supply and demand capacities. The latter model disregards the specific behaviour of individuals and assumes that there is a common behaviour within a group where the groups interact with each other. The authors' [102] assumption is that the group members behave similarly.

Thus, it is argued that it is possible to characterise the decision-making at an aggregated level, to integrate the strengths of the ABMs and SR approaches. This could allow capturing the location of agents accurately and their behaviour, while the requirement of data may not be as exhaustive as that for the individual level. This approach will enable the inclusion of the spatial, temporal and social dynamics that drive the adoption process. However, one can argue that integrating insights from the SR to address some of the ABM limitations may also imply new challenges, for instance, it is necessary to investigate whether a spatially explicit ABM is also subject to the MAUP, like the SR.

Moreover, given the limitation of the ABM to calculate both the financial and social utility, as it is argued that there is potential in the artificial neural networks (ANN) to improve the decision-making process. Rather than the rational choice or the stochastic preferences, there is a need for new approaches that capture the nonlinearity of correlation between adoption drivers and the adoption process [20,21]. Because of the ANN training algorithms can extract the complex behaviour of data sets, the ANNs are universal estimators. Therefore, they can approximate non-linear behaviours such as the PV and EV datasets. Yet, the full potential of the neural model to increase the reliability of the model results has not been applied to the analysis of spatio-temporal analysis of PV and EVs diffusion. It is expected that data-driven knowledge generated by neural networks help to estimate the spatio-temporal rates of adoption of these technologies and increase the robustness of the model outputs by using spatially explicit data sets. As well, it is necessary to recognise the wide variety of drivers of the adoption process and their effect on the model performance. Moreover, as the literature suggests empirical regularities between PV and EV adoption [29,36] raise whether there is a way to transfer knowledge from one process to another. More explicitly, whether the PV adoption patterns can inform the spatial diffusion patterns of the EVs have been overlooked in the literature.

Knowledge exchange between PV and EV adoption patterns

Emerging literature notes empirical regularities between EV and PV adoption decisions. McCoy and Lyons [26] model the EV adopters' environmental utility based on the previous adoption of energy-saving technologies which might be extended to include PVs as they can offset household demands to draw electricity from the grid. More specifically, Davidson et al. [14] have introduced the number of EVs registered in a location as an explanatory variable to analyse residential PV installations. This together with Cohen and Kollman's insights [36], that suggests the presence of a relationship between EV and PV adoption, however, the nature and scope of this relationship have not been empirically quantified.

Moreover, there is evidence of regularities between the ownership of these technologies. Despite these technologies have been analysed simultaneously, they have been modelled independently for the same time horizon. As a result, there are no available insights on whether agents can transfer learning from the decision on one technology to another. This emerges as an area for future research to model knowledge transmission between domains, or how the spatial patterns of EV adoption can be informed by another technology's pattern. This may help to extend the understanding of the decision-making process of EV adopters based on previous low carbon technology decisions taken.

1.5.1 Research gaps

The spatio-temporal analysis of EV and PV (or other low-carbon technologies) is an emerging area of study. The contemporary modelling approaches allow to explore the spatial, temporal and social dynamics elements of the adoption process, yet, there is still the need for new approaches that characterise these elements in an integrated framework. More explicitly, the ABM, which in previous sections has been shown to have the potential to address such need, by integrating elements from the SR and the ANN. Most of ABMs geographic location is simulated or semi-empirical, meaning that are limited to capture the spatial regularities of the adoption process or energy system. Because the spatial regularities in the adoption of low carbon technologies adoption are noted in different contexts [46,47], it is imperative to explicitly capture the spatial dimension of the EV and PV adoption process. This raises the prospects to adopt the spatially explicit models' capabilities to capture the spatial dependence of the adoption process as well as considering the spatial variation in socio-economic variables (agents' heterogeneity). One implication of such spatial ABM models might be the use of aggregate data which might be less demanding than those at the individual level. However, these spatially explicit models (i.e. spatial regression) tend to overlook the temporal dynamics of the adoption process.

Therefore, the research question is revisited considering the literature review, formulating the following hypothesis:

It is possible to explicitly characterise the spatio-temporal dynamics of the decision-making towards EVs and PVs, whilst including the social dynamics and the relationship between these two technologies.

1.6 Thesis aim and objectives

Therefore, this research aims to develop, test, and validate a modelling framework to characterise the spatio-temporal dynamics of the decision-making towards EVs and PVs, whilst including the social dynamics and the knowledge exchange between these two technologies. In particular, the research develops an integrated ABM and ANN model that addresses some of the limitations of the common rule-based ABMs. Besides, to integrate the (i) spatial and (ii) temporal dependence, (iii) peer-effect, (iv) spillover effect and (iv) preferences towards other technologies, such model would build upon insights from the SR, resulting in a novel spatio-temporally explicit ABM, which can analyse the EV and PV adoption process. Thus, the aim is unfolded in the following four objectives:

1. Investigate how the spatio-temporal and social dynamics of the adoption process can be captured explicitly by an ABM, drawing insights from the SR and integrating the ANN approach as the decision-making process.
2. Analyse the effect of reflecting the population heterogeneity, by integrating the agents' socioeconomic variables into the agents' characterisation.
3. Investigate the spatio-temporal patterns of EV adoption with the integrated model and assess the model’s flexibility and transferability between different technologies.
4. Finally, investigate whether it is possible to exchange knowledge between two adoption processes, by integrating PV and EV data into the same decision-making.

Following chapters (2-5) develop each of these objectives, building upon the insights of the literature review and on those which precede. Chapter 2 develops a novel PV ABM and ANN integrated model, which characterises the spatio-temporal and social dynamics that drive the adoption process. Chapter 3 extends this model to account for the agents' heterogeneity and investigate whether the model is subject to the MAUP. Chapter 4 investigates the spatio-temporal patterns of EV adoption and assesses whether the model developed for PVs can be used for other low carbon technologies. Chapter 5 tests the thesis hypothesis by integrating the PV and EV models, to evaluate whether the PV adoption process can inform the decision-making towards EV. Each of the modelling chapters (2-5) presents a data analysis that underlines the model's design and validates the results temporally and spatially. Chapter 6 is devoted to the discussion of the overall results and findings, acting as the intersection of the individual insights produced in previous chapters. Finally, the main conclusions are presented at the end of Chapter 6.

2 Explicit spatio-temporally characterisation of an agent-based and artificial neural networks integrated model

2.1 Introduction

In Chapter 1 the main elements and applications of the ABM and SR are presented. This chapter presents in detail the design and development of the explicit spatio-temporal model. Given the need for new approaches to characterise the adoption process of EVs and PVs, this chapter seeks to outline how the temporal and spatial regularities of PV adoption can be explicitly captured into the ABM. Moreover, the chapter presents in detail the limitations of the ABM decision-making, and how the ANNs will allow agents to learn from their past decisions (Section 2.1.2). Therefore, this chapter develops a novel approach that characterises the diffusion of PVs. The model's design is based on agent-based modelling and integrates insights from the spatial regression model.

Namely, this chapter's objectives are as follows:

1. To integrate the artificial neural networks model as the agents' decision-making approach and its capabilities to handle time-series.
2. To make use of spatio-temporal explicit data sets to incorporate the spatial and temporal dependence of the adoption process into the agents' decision-making.
3. To implement the effect of social dynamics into the decision-making process, by characterising the peer-effect and the neighbourhood effect.

The model is expected to estimate rates of PV adoption at a specific time and place, whilst addressing the limitations of the ABMs. The model is empirically tested and validated for the city of Birmingham in England, using PV installations data at high spatial and temporal resolution. Birmingham is a metropolitan city in the West Midlands, England. It was selected as a study location because it is representative of the typical population and is one of the most populous British cities besides London.

The chapter is organised as follows: the remaining of Section 2.1 presents a review of the relevant literature which underlines the theoretical basis for the model, focusing on the SR and ANN's insights that can address the limitations of ABMs. Then, section 2.2 explores the spatio-temporal characteristics of the PV data, which in turn will define the model design. Section 2.2.5 builds upon these findings and present the spatio-temporally explicit ABM, as well as the technique for validation. In Sections 2.4 the results are presented and discussed, focusing on the model validation, the applicability of the results, and limitations. Finally, Sections 2.5 summarises the findings produced and sets a pathway for the following chapters.

2.1.1 Spatio-temporal characterisation of the technology diffusion

Building upon the literature review in Chapter 1, this section presents in detail the different ABM characterisations of the adoption process and underlines the modelling elements for the spatio-temporally explicit ABM. Such elements are entities, objects, agents characterisation, decision-making, and social effects.

Objects and Entities

The modelled objects in the reviewed ABMs comprise EVs [25,27,33], PVs [68,94], and others such as environmental innovations [93], energy innovations [43,75,87], and heating systems [50]. Yet, the entities (agents) characterised only include individuals or households, as the reviewed studies consider consumer to consumer interactions, disregarding the merchandising activity and the seller-consumer interaction.

Characterisation, Location and Data sources

Agents' characterisation comprises the assignment of individual characteristics and their location in a virtual representation of the physical world. Modellers commonly use georeferenced data (and geographical information systems) to create a world-like layout and locate the agents. The reviewed ABMs commonly characterises agents as individuals or households, whilst the number of agents varies from a small proportion of the population to entire populations. For instance, Adepetu and Keshav [68] create a semi-empirical population which is assigned with real socioeconomic characteristics of 100 households, yet it is not clear whether the sample is statistically significant for the total population in Ontario. Instead, Rai and Robinson [90], Robinson and Rai [76], and Robinson et al. [91] model the total households from Austin (Texas, US), using data from the census. Then, they differentiate adopters from non-adopters, applying the findings from a previous survey carried out for PV adopters.

An alternative approach is the one which combines current survey data with other methods or software to create a virtual population. Survey data has been used to characterise significant samples of the population, which is then simulated as many times as needed to capture the actual population size [25,77,78]. Similarly, Cui et al. [27] generate a virtual population of households based on actual aggregated data. By applying the *copula-based household synthesiser*[4], they generate an individual virtual household population with similar characteristics (with intra-group variance), resulting in the simulation of 190,965 agents. Eppstein et al. [33] create a full virtual population of households. First, they generate a virtual distribution of *income*, considering five hypothetical cities. Drawing from previous experimentations that show no significant differences between simulation with 1,000 or 10,000 agents, the authors use 1,000 in the interest of computational efficiency. These agents are created using the *turning bands method*.

The location of the agents is fully related to the agents' characterisation, yet empirical or semi-empirical characterisation may imply a simulated location. Hence, the actual location may differ from the simulation and may involve a loss in data accuracy [25]. Although 'the agents' location needs to be accurate, it is a challenging task to obtain personal characteristics of the entire population [27].

Decision-making criteria

Commonly, the ABMs use rational choice principles, assuming that agents have access to perfect market information and can evaluate the benefit of their

[4] This method simulates households using local distributions (i.e. the statistics of each census block), which allows to keep empirical correlations.

decisions. However, this is rather limiting as some of the drivers such as peer-effect or personal beliefs have subjective value, and individuals rarely possess perfect market information [43,72–74]. Then, the authors use a utility or social threshold to characterise their decision-making process [26,33]. The utility threshold considers exogenous and objective elements to the agents, such as electricity prices or government subsidies, and this is usually associated with the financial benefits of adopting a certain technology. On the other hand, the social threshold reflects subjective elements such as the agents' personal beliefs, values, and reflects how adopting a certain technology satisfies the interpersonal preferences of the individual.

In most studies, the agent's decision-making process is defined by a single criterion [25–27,32,68]. A single criterion approach is most commonly used, however, other studies have also suggested employing a multi-step process, for instance, Eppstein et al. [33] characterise the agents with an equation to calculate the relative cost of buying a new vehicle (among a set of options). They consider the purchase cost, financing, fuel and electricity cost. Then, the agent compares this cost with net annual income. If the sum of these costs exceeds 20% of the annual income, this is considered unaffordable. However, if the vehicle is affordable, the agent then considers the social benefits. The agent then calculates the perceived social benefit of each vehicle; if these benefits exceed the social threshold, the agent adopts that specific option. Others such as Robinson et al. [76,91] model the decision-making process using both social and utility thresholds. Financial benefits are calculated based on the financial payback, considering the energy produced by the solar panel, the price per kWh generated, and the government subsidies. The social threshold is related to the attitudinal variables which change according to the social network effect.

Social effects

Regardless of the type of threshold used to characterise the decision-making process, most of the ABM studies implement the interaction among agents. These interactions can be described in two approaches, the peer-effect and the social norms. The former refers to the direct influence of the social network in the decisions (perception). The latter comprises the impact of the collective behaviour on the agent's decision-making. The agent's social network may be characterised using different approaches, for instance, defining a spatial neighbourhood or choosing between those agents with similar characteristics. Eppstein et al. [33] define the agent's social network as the *k* nearest neighbourhood (starting with *k*=2) and generate random social networks. Similarly, Adepetu, Keshav and Arya [25] and Schwarz and Ernst [93] define spatial proximity and connect the agents with similar socio-economic characteristics. Robinson and Rai [76] combine the distance-based and agents similarity criteria with a random connection so that agents are connected with other agents anywhere in the area. Exceptionally, Cui et al. [27] do not provide detail on the agents' social-network structure.

Alternatively, the social norm has been modelled by considering the adoption rate of the total population as a proxy for the social effect [26,32,68]. Similarly, Eppstein et al. [33] use the media coverage in an area as an indication of the environmental awareness which increases over time, whilst Robinson and Rai [76] implement the perception of the technology as a measurement of social awareness of the social norm

2.1.2 Artificial neural networks as a decision-making process

The ABM faces some limitations (Section 1.2.1) because the decision-making process assumes that the agents decide by assessing the financial benefits from a set of alternatives, and social utility based on personal and social norms. This is because the social utility is subjective, and consumers rarely possess the perfect information of whole alternatives to make comparisons economically. To address these uncertainties, some authors such as [40] have run a sensitivity analysis on market conditions. These authors modelled the adoption of EVs under different assumptions of the macro-economic adoption drivers (i.e. gasoline price, energy cost, vehicle taxes, subsidies, vehicle cost). Additionally, some scholars have pointed to the need for alternative decision-making methods that consider elements from the human cognition [49].

Noori and Tatari [49] propose considering alternative and more realistic behavioural models by including other characteristics of human decision-making, however, the literature for ABMs using an alternative adoption criterion is scarce. For instance, Kang and Choi [103] propose a theoretical ABM and ANN integrated model as an alternative to the decision-making process. The artificial neural network model emulates how the human brain generates information, by associating experience (past data) and the associated decision (output). The generated experience-based knowledge reduces uncertainty, as the model can adapt to new conditions that are not available during the training. Kang and Choi's [103] model optimises the combination of each individual's decision which is chosen from a set of possible decisions. This approach uses the ANN to optimise the strategies that the agents execute, evaluating the global fitness of the model to solve a specific problem. However, the decision-making is not done by the ANN but is chosen from a list of possible actions. Then, through the training process, the ANN learn which combination of actions produces the best model fitness and communicates that to the agents. In this sense, the ANN interact with the agents in a hierarchical way, which is in principle a contradiction to the ABM philosophy. The ABM assumes that the agents are autonomous [68] and that the system cannot be controlled in its whole [43].

Given the limitations of the ABMs decision-making, and that the agents rely on experience and perception more than from complicated calculation, as informed by the bounded rationality [76,102,104]. This is expected to provide a more realistic characterisation of the adoption process, whilst providing an explicit time horizon for the decision-making. Thus, it is argued that the ANN concept of experienced-based knowledge [105–107] applies to the decision-making of the ABMs. From the pedagogy point of view, this has been described by Kolb's learning framework, which defines learning as the combination of experience and reflection. The former refers to the degree of involvement/engagement, the relevance for the individual, and the type of activity (task or interaction). The latter comprises the mechanism for reflection available to the individual, behaviours and attitudes [108,109]. These experience and reflection elements can be found in both the ABM and the adoption process. First, the experience includes the presentation of information to the ANN, meaning the EV/PV time-series. This also comprises the actual first-hand experience of the agents with the EVs or PVs technologies, for instance, trials [110] or observation [20]. The experience considers the relative relevance for the individuals, for instance, environmental concern [36] or green party tendencies. Secondly, the reflection elements comprise how the ANN assesses the fitness of its estimations while reflecting attitudinal preference towards those EVs/PVs (financial and social utility), and the way the individual assesses such benefits. Then, the learning of the ABM and the individuals come from the interaction of such elements.

Other computational models such as the Fuzzy Logic or the Genetic Algorithms also implement elements of the human cognition. The former generates knowledge from a set of inference rules, by assuming that decision-making is uncertain and imprecise [111–113], while the latter is a biology-inspired optimisation model that can extract knowledge from complex datasets [111,112,114]. However, these techniques present limitations to handle the time variable. Moreover, the capabilities of the ANNs to inform the decision-making process in ABMs is still unexplored in detail yet even though, they present a number of capabilities noted by Haykin [115]:

- Universal estimation capability - ANN can describe nonlinear functions, usually corresponding to complex behaviour.
- Mapping of associations between input and output - ANNs do not require specifying the association rules between inputs and desired outputs, rather they are created through training.
- Adaptivity - ANN can learn from disruptive changes in the data trends, by modifying its characteristics.
- Contextual Information - each ANN element is potentially affected by each other input; thus, correlations are considered.

In a particular study, the ANN has been used at the aggregated level to identify the “adoption state” of a location [88,89]. In these studies, the neural network model is trained to memorise the socioeconomic characteristics of the areas with PVs, and then identify when a location reaches these characteristics. The areas in this model are uniformly defined grid cells, disregarding the real-world layout and impacts of areas on each other. The forecasting process calculates the total amount of PV in each region using an S-curve, which shows the historical growth of the PV in the entire population. The assignment of local values from regional characteristics do not allow the model to consider local characteristics. Consequently, the potential of ANN to improve the decision-making process of PV adopters by capturing characteristics of areas explicitly is still overlooked which is expected to be addressed using empirical data.

2.1.3 The spatial nature of the social effects

The ABM allows the representation of the social dynamics that drive the decision-making. Namely, these social dynamics include the peer-effect and social norms, capturing the influence that one individual has on another or the influence of the overall society's trend, respectively. Although some authors have defined the agent's social network using the distance between agents as criteria [33,93], whether the distance between agents reinforces or decreases the social effects has not yet been tested.

Contrary to the SR allows to characterise explicitly the spatial layout of the diffusion process, yet, these approaches tend to overlook the temporal dimension of the adoption process. In spatial regression, the so-called spillover effect captures the information flow between or within specific locations [8,20,46]. This effect is implemented in different means, for instance, Balta-Ozkan, Yildirim and Connor [8] characterise this as the influence of one area over the adjacent locations. Conversely, Richter [20] defines this as the transmission of information between the individuals in the same area. Like the ABM, areas are influenced by those in their "social network", yet, this social network is based on geographical elements. The spatial regression models identify this relationship in two ways: adjacency and distance. In the former, this influence has the same magnitude for all areas that share a border, whilst disregarding the influence of areas that are not adjacent [13]. The distance-based approach disregards the adjacency and assumes that the social influence decreases with the distance such that the influence of one area over another will have a greater impact the closer they are. The research argues that the spillover effect as a representation of the spatial dependence can be homologous to the social influence of the ABMs. Then, the following sections present the step-by-step development of an ABM model that implements the spatio-temporal and social dynamics of the adoption process.

2.2 Methods and materials

This section outlines the construction of a postcode district level agent-based model that can reproduce the spatio-temporal patterns of the PV adoption, by using ANN as the agents' decision-making criteria, to reduce the uncertainty in the decision-making process.

2.2.1 Data

The model uses historical PV installation data and ignores any activity done by the sellers/car retailers [8,14,20,68]. The dataset comprises individual PV registrations by their registration date as compiled by the Office of Gas and Electricity Markets (Ofgem) as of 30 December 2015. The *Feed-in Tariff Installation Report* accounts for the geo-referenced[5] database of all renewable energy installations such as anaerobic digestion, combined heat and power, the hydro, wind, and solar photovoltaic. The analysis uses all domestic PV registrations as explained below, the choice of the database's spatial and temporal resolution is critical to accurately represent the actual spatio-temporal patterns of PV adoption as discussed further in section 3.1.1. Therefore, the ABM and ANN model characterises the agents as the 49 Postcodes (PCs) in the City of Birmingham in England. The simulation comprises data from Jan 2011 to Dec 2015, resulting in a total of 60 observations (months) for each PC; over this period the number of PVs installed increased from $N = 214$ to 8236[6]. Figure 6 shows the range of values among the PCs, which have an average of 151 PVs. Despite most of the areas have a total number of PVs between 100 and 200, there are areas with extreme values (more than 400 and 0 PVs), resulting in a large variance in the data across the areas. It is expected that that given the

[5] Each registration is geo-referenced to a postcode, which can be later be associated with other spatial resolution, such as Local Authorities.
[6] Note that because of the low level of PV adoption before Jan 2011, this data was excluded from the analysis (Total PV < 3% before 2011; Average PV per PC < 4)

individuality of the ANNs, the agents can adapt to the local variations, instead of fitting their behaviour to the overall data trend. Moreover, Chapter 3 studies the effect of changing the number of agents in the simulation, consequently introducing variation in the socioeconomic characteristics of the population.

2.2.2 Spatio-temporal resolution

The aim is to generate a model able to analyse the patterns of the PV adoption at a high spatio-temporal resolution, yet the identification of what defines 'high' resolution requires critical analysis of the data. Yet, given the different range of values across the areas, data sets are normalised in a [0-1] range[7], following equation (2-1).

$$\widehat{PV}_{i,t} = \frac{PV_{i,t} - MinPV_i}{MaxPV_i - MinPV_i} \qquad \textbf{(2-1)}$$

Where

$\widehat{PV}_{i,t}$is the *t-th* PV estimation in the *i-th* area

Min_i is the minimum number of PVs in the *i-th* area over the study period

Max_i is the maximum number of PVs in the *i-th* area over the study period

[7] The total PV for Birmingham is also presented as a reference, normalisation follows the same definition as the individual areal definition.

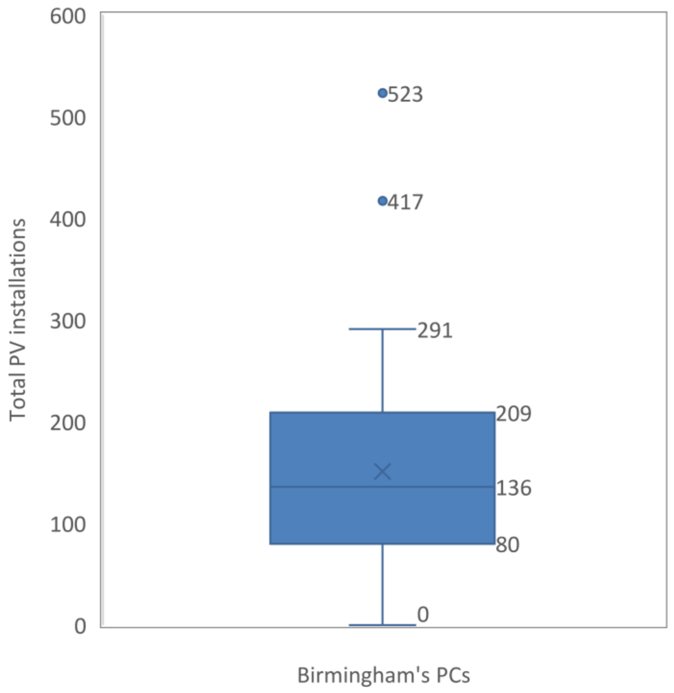


Figure 6. Boxplot of the total PV installations in Birmingham.

This transformation helps to visualise that there is a trade-off between the spatio-temporal resolution and the data variability and availability, as high spatial resolutions present limited data variability, whilst low temporal resolutions may overlook specific changes in the agents' behaviour. Figure 7 and Figure 8 show the adoption rates of PVs for the postcodes districts and Local Superoutput Area in Birmingham, being the former bigger areas than the latter. Figure 7 displays time-series with high variability, exhibiting similar behaviour to the total of Birmingham, on the other hand, Figure 8 shows a stepped function with low variability. As seen, the data may present different behaviours depending on the spatial resolutions (PV adoption is proportional to spatial areas). On the other hand, the temporal resolution helps to increase the data variability, as the data accumulates and changes over time are more evident (i.e. ranging from monthly to six-months), however, the temporal resolution limits the number of observations. On one extreme of the resolutions, the annual adoption rates at the national level have been used by those approaches that focus on the technological choices to achieve certain levels of GHGs, and as mentioned before, those have limitations to inform about local regularities in the EV and PV adoption. On the other side, characterising the individual adoption on daily basis would result in a significant amount of data. Moreover, the results of using such resolution would not be able to inform policymaking or DNOs unless these were aggregated [77,78]. Alternative combinations of spatial and temporal resolutions are available in Figure 70, Figure 71, Figure 72, and Figure 73 (Appendix 2).

This study uses a combination of the PC and monthly resolutions, as this is the combination with the most variability and the most possible number of observations; this resolution is consistent with Richter [20]. Figure 7 and Figure 8 also present the cumulative adoption rates of PVs adoption for Birmingham, (solid black line), to visualise how the aggregated figures disregard the specific local behaviours. However, this is not clear if there is a seasonal behaviour or any repetitive patterns, thus, the following section presents an in-depth analysis of spatial and temporal dependences. The spatial distribution of PVs at PC resolution is shown in Figure 9. As seen, there are a few areas with less than 50 PVs, most of which correspond to the city centre areas where the number of residential buildings is low. The following sections present a statistical analysis to explore the spatial and temporal dependence of the datasets.

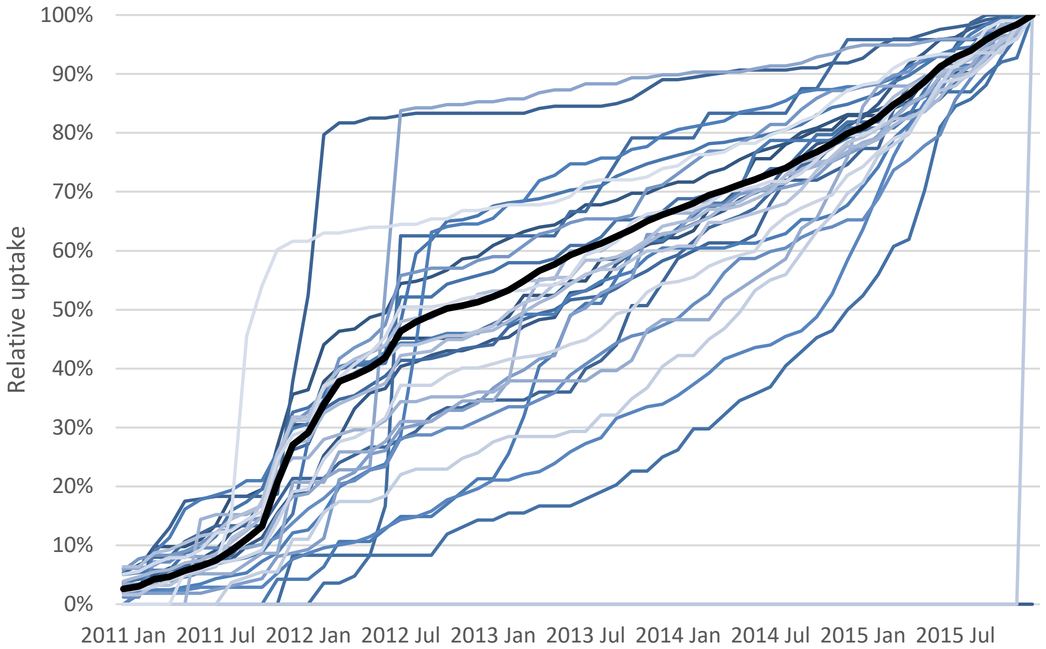


Figure 7. Normalised monthly cumulative adoption rates of solar PV at the PC level.

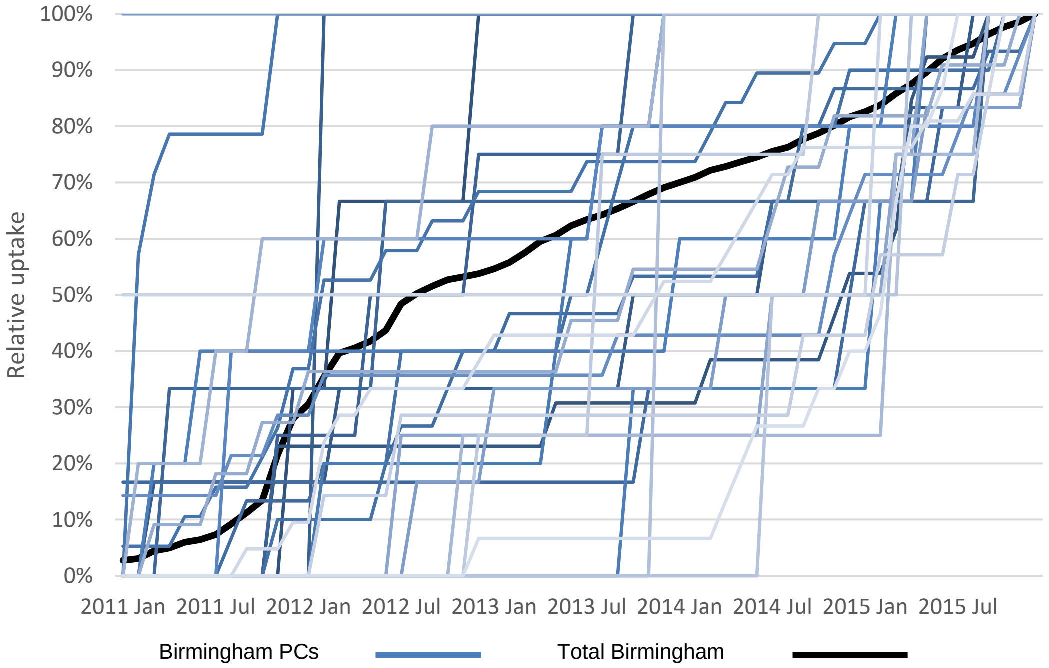


Figure 8. Normalised monthly distribution of solar PV adoption at LSOA level.

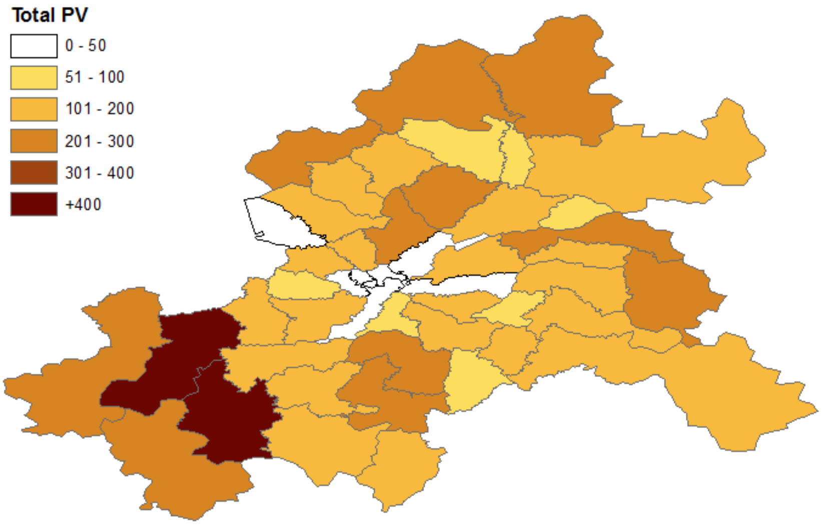


Figure 9. Spatial distribution of the PV installations in the Birmingham postcodes (Dec 2015)[8].

2.2.3 Temporal dependency

Because literature does not provide clear criteria or method to identify the temporal dependency of the adoption process, other than trial and error, the Autocorrelation function (ACF) and the Partial autocorrelation function (PACF) are calculated. Both statistics are used in the time-series analysis to define the number and order of lags in the data. For instance, the Box-Jenkins methodology uses the ACF and PACF to define the ARIMA[9] model parameters. The ACF measures the correlation between PV_t and PV_{t-1}, the second lag measures PV_t and PV_{t-2}, and so on. Figure 10 shows the ACF functions for the cumulative rates of PV adoption, which plots the degree of the autocorrelation of each lag and the significance interval. Each boxplot represents the number of PCs with a significant lag at the *i-th* lag. For instance, at the mean and median of the correlation or *order one* (t-1) is very close to the significance level. The size of the boxes decreases over time, suggesting that the PV datasets have an autoregressive nature.

---

[8] All maps presented in this thesis are self-elaborated using the data stated in each of the chapters, using the ArcMap v10.4.1 software.
[9] Auto Regressive Integrated Moving Average

On the other hand, the PACF looks at the effect that has not been yet explained by the autoregressive element. Then, the PACF looks at the autocorrelation of the residuals after calculating the AFC. Figure 11, the datasets have a significant correlation with the first lag, this means that the autoregressive model is of order 1, thus, the model's inputs are elements at *t*-1. This is used in Section 2.3 to define the structure of the inputs, defining the output of the decision-making process in terms of the preceding month. There are other significant lags, yet these are present in less than 10% of the PCs and their mean and median are close to cero, suggesting that most of the PCs have a near to zero influence of decisions made more than a month ago. Also, the absence of significant lags in a quarterly manner nor the twelfth lag suggests that the adoption of PVs does not present a seasonal behaviour. Both functions are calculated using Python, considering monthly resolution and up to 12 lags.

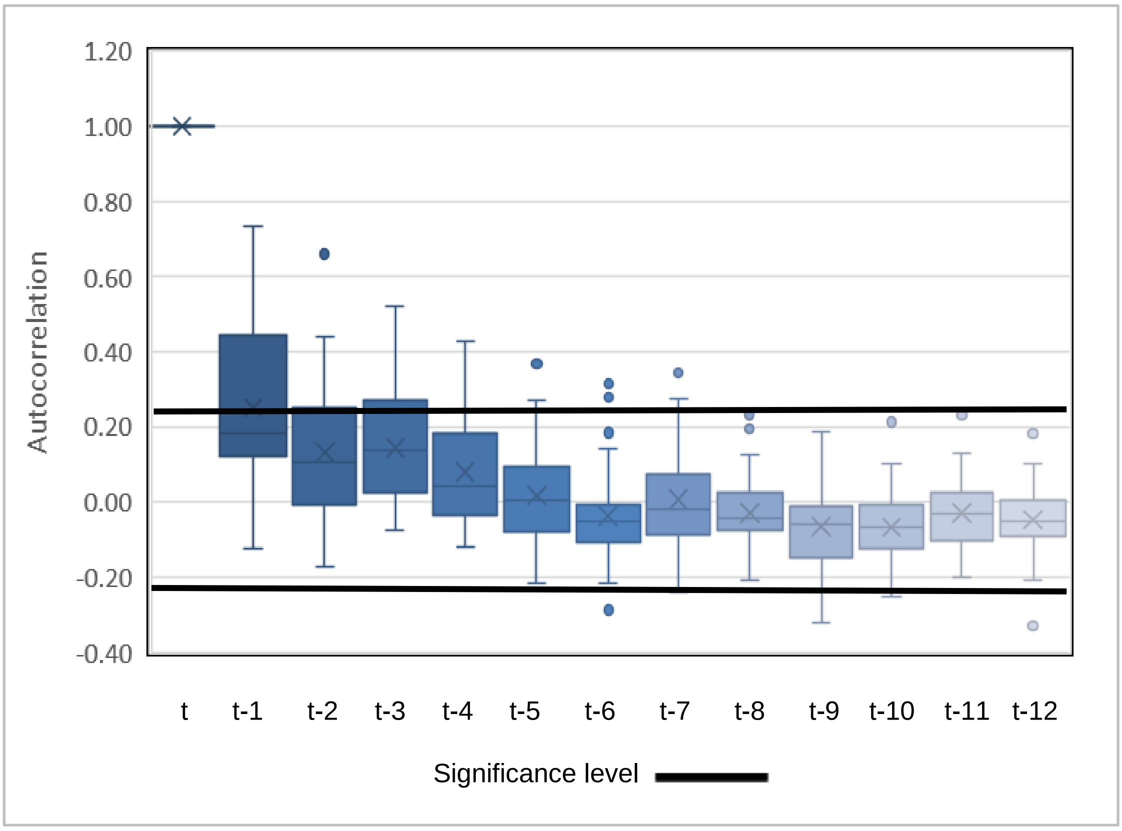


Figure 10. ACF of the PV installation data at the PC level.

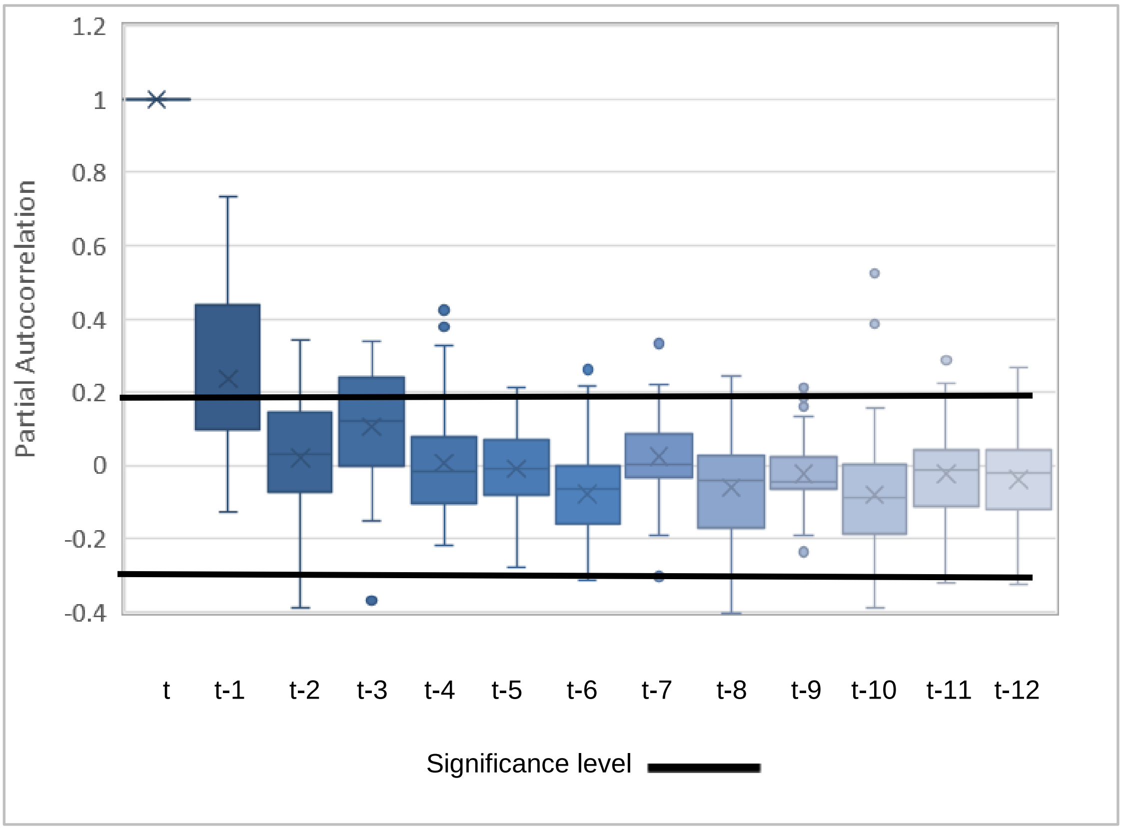


Figure 11. PACF of the PV installation data at the PC level.

2.2.4 Spatial dependency

The spatial association between the observations is analysed via the Moran's I statistic [8,15,19,80], which is a global indicator of spatial association. This is a multi-directional indicator that accounts for the distances between locations instead of temporal lags. A Moran's I value > 0 indicates that the data tend to cluster, while a value < 0 indicates a tendency to dispersion. The Moran's I statistic is calculated using the ArcMap software (version 10.4.1) for the monthly data at PC resolution, considering the distance from the population centroids[10]. To reflect how the correlation between areas decays along with the distance, the weights are calculated as the inverse of the distance: $w_{ij} = {}^{1}/_{d_{ij}}$, where d_{ij} is the distance between any two PCs. Table 5 summarises the Moran's I statistics for

[10] These centroids are a weighted reference point for the centre of the population.

the total number of PV installations at each PC by Dec 2015, which suggests spatial clustering patterns. Then Figure 12 displays the clusters and their types[11], as seen, the PV datasets exhibit Low-Low clusters (cold spots) located in the City centre having a low number of PVs, while the (High-High cluster) hot spots are in the south-west of the City accounting for the max number of PVs and its surroundings (see Figure 9).

Table 5. The Moran's I index value, z-score and p-value of the PV installations data

Statistics	Moran's Index	z-cores	p-value
Value	0.3640	5.0341	0.0000*
Given the z-core of 5.0341, there is less than 1% likelihood that this clustered pattern could be the results of random chance.			

The statistical significance is marked with asterisks.

**p<0.05*

[11] Clusters of low PV values (Low-Low) or cold spots, and or high (High-High) data values or hot spots [164–167].

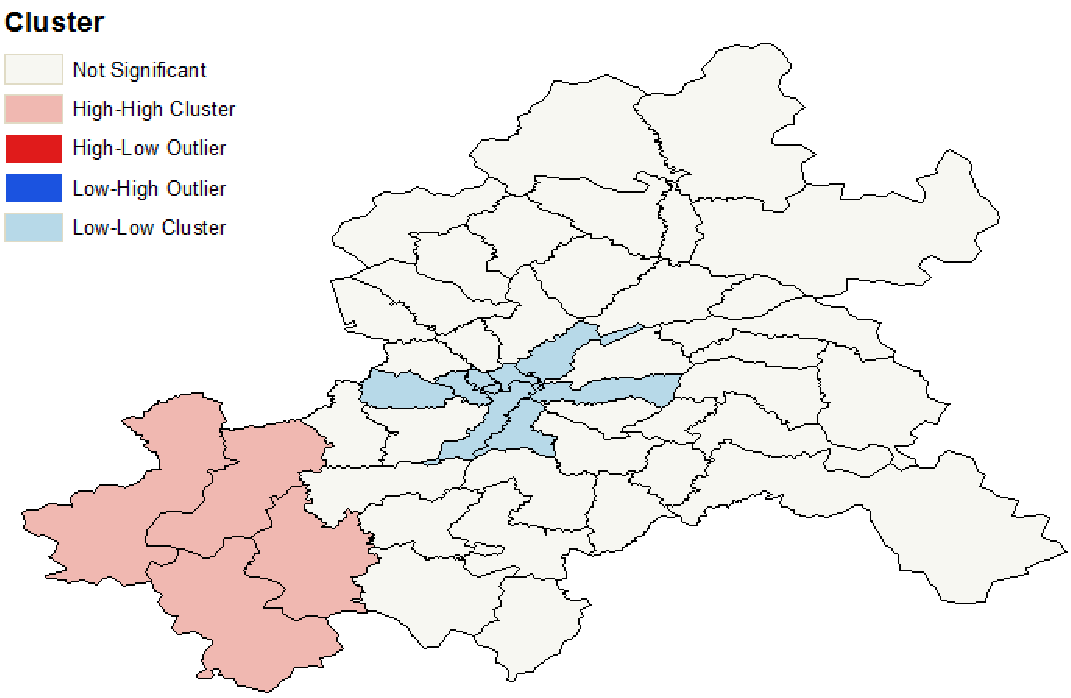


Figure 12. Hot spot analysis of the PV installations at PC level (Dec 2015).

2.2.5 Model design

Figure 13 exemplifies the key elements of the spatio-temporally explicit ABM and ANN approach. As seen, the model characterises agents as geographical areas ($Area_i$) that represent the cumulativive decision making of the individuals living in that area [20,102]. The model accounts for the spillover effect and peer-effect, the former reflects the influence of one area over others, and the latter is the pairwise influence of one area over another. Additionally, the model reflects how the spillover effect decreases with the distance, by considering the distance between locations.

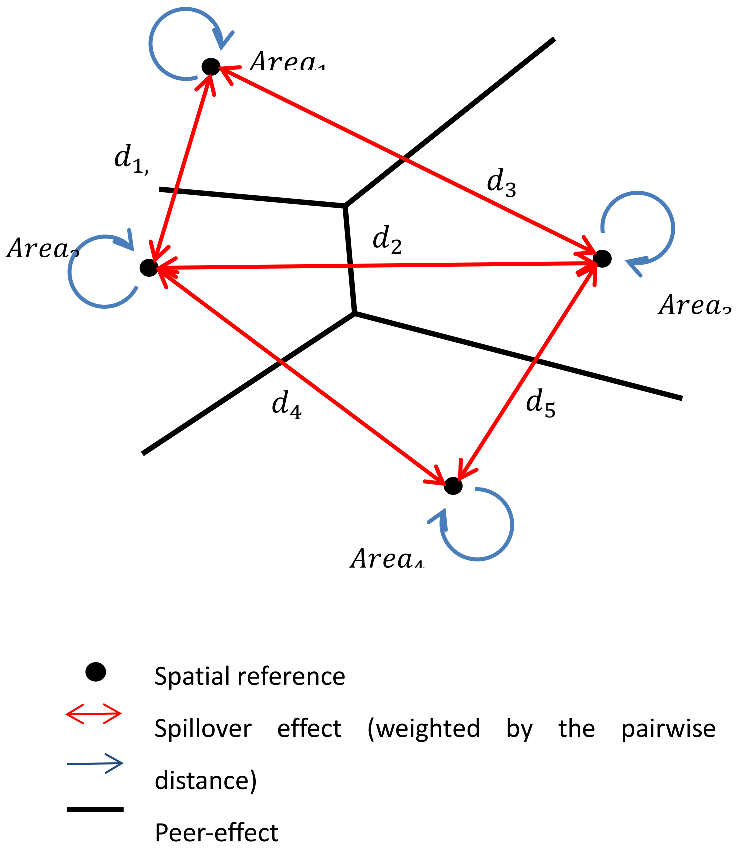
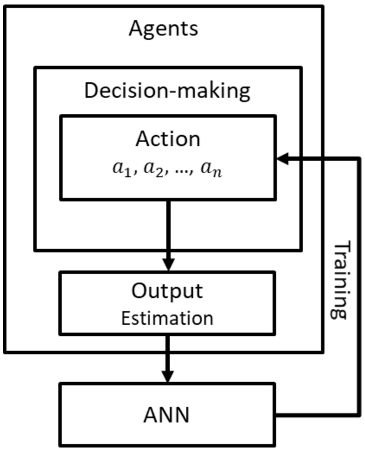
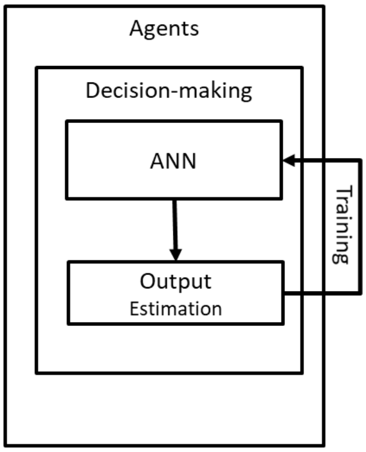
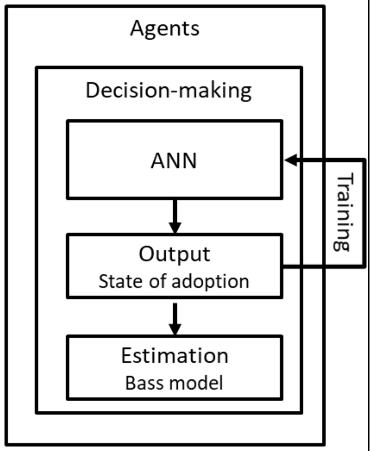


Figure 13. Conceptual model of a spatio-temporal explicit ABM.

The literature offers studies that have integrated the ABM and ANN models to create intelligent agents, however, these have not fully utilised the decision-making capabilities of the ANNs. Table 6 shows the abstraction of two of those models and the model proposed in this research. Kang and Choi [103] theoretical model use the ANN to improve the overall model performance, optimising the available options for the agents. In a sense, this could be seen as governing or directing the behaviour of the agents instead of keeping their autonomy. Zhao et al. [88] use the ANN to memorise the adoption state of a geographical area. Then, if the output of the ANN is that the area is in one of the adoption states, the model uses the Bass model to estimate the total number of PV. Despite these initial attempts to integrate the ANN with the ABM, the decision-making is still not fully characterised by the ANN. As seen in Table 6, the proposed model uses the ANN for both the decision making and the estimation of the adoption rates.

Table 6. Comparison of three implementations of ABM and ANN hybrid models.

ABM + ANN [103]	Spatio-temporally explicit ABM + ANN	Hybrid model (ANN) [88]
		

2.3 Model development

The integrated ABM and ANN model is structured in two layers: the spatial layout and the decision-making process. Information is transmitted from one layer to another as shown in Figure 14, where the second layer is embedded in the first one. The spatial layout layer is an ABM module that fulfils the following three main functions: (i) spatio-temporal characterisation of agents; (ii) definition of social effects; and (iii) management of information flows between the agents.

The second layer contains a population of ANN that replaces the common rule-based approach by processing the inputs obtained from the spatial layout layer. The decision-making process is driven by the temporal resolution, reflecting the temporal dependence of the number of PVs on those in a preceding period.

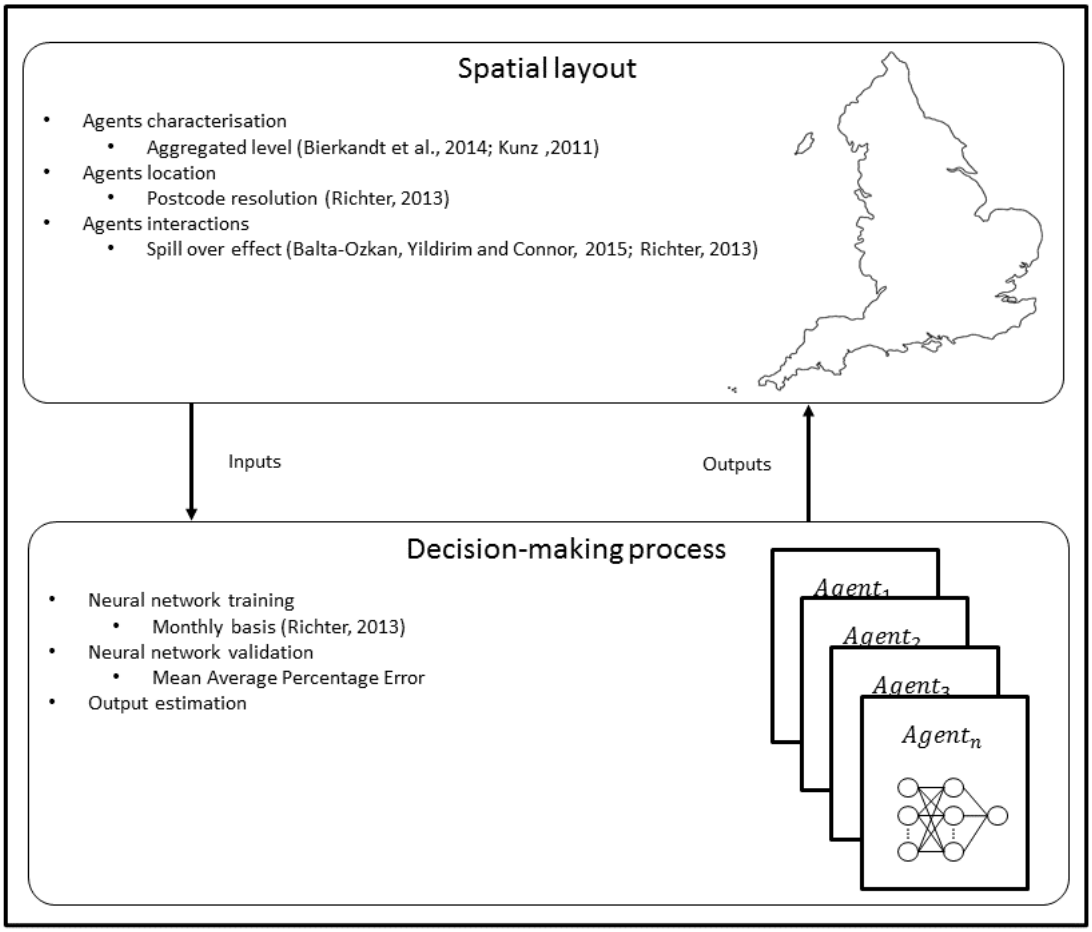


Figure 14. Methodological framework and information flow between the layers.

The model implementation follows the four steps shown in Figure 15. These start with the agents' characterisation, creating an agent for each of the PCs in the Birmingham area. The model uses the boundary files (.shp) to define the shape of each area and to assign a georeference according to the population centroid. Each of these agents makes a query to the PV installation database, loading their individual time-series. Using the adjacency principle, the model associates neighbouring PCs and calculate the distance between their centroids. This is then used to calculate the spill-over effect. Afterwards, each of the agents is assigned with a neural network, which is initialised using random values.Then, the ANN is presented with pairs of inputs and outputs to the neural network, through the learning algorithm it generates knowledge (see Algorithm 2 in Appendix 4).

Because the decision-making considers the actions of other agents, the algorithm must communicate each individual decision to the population. This information flows occurs considering the social-network of each agent and the distance. To reduce computation time during the training, instead of calculating the number of PVs in the surrounding areas, this number is calculated beforehand. Thus, the time-series for the training comprises the total number of PVs in the adjacent areas which are weighted by the distance between the areas. During the forecast, the communication process takes place after every agent has estimated the adoption rates. The model also considers the random element of the synaptic weights, by running the whole simulation 100 times, then, the results report the average behaviour of the model's output.

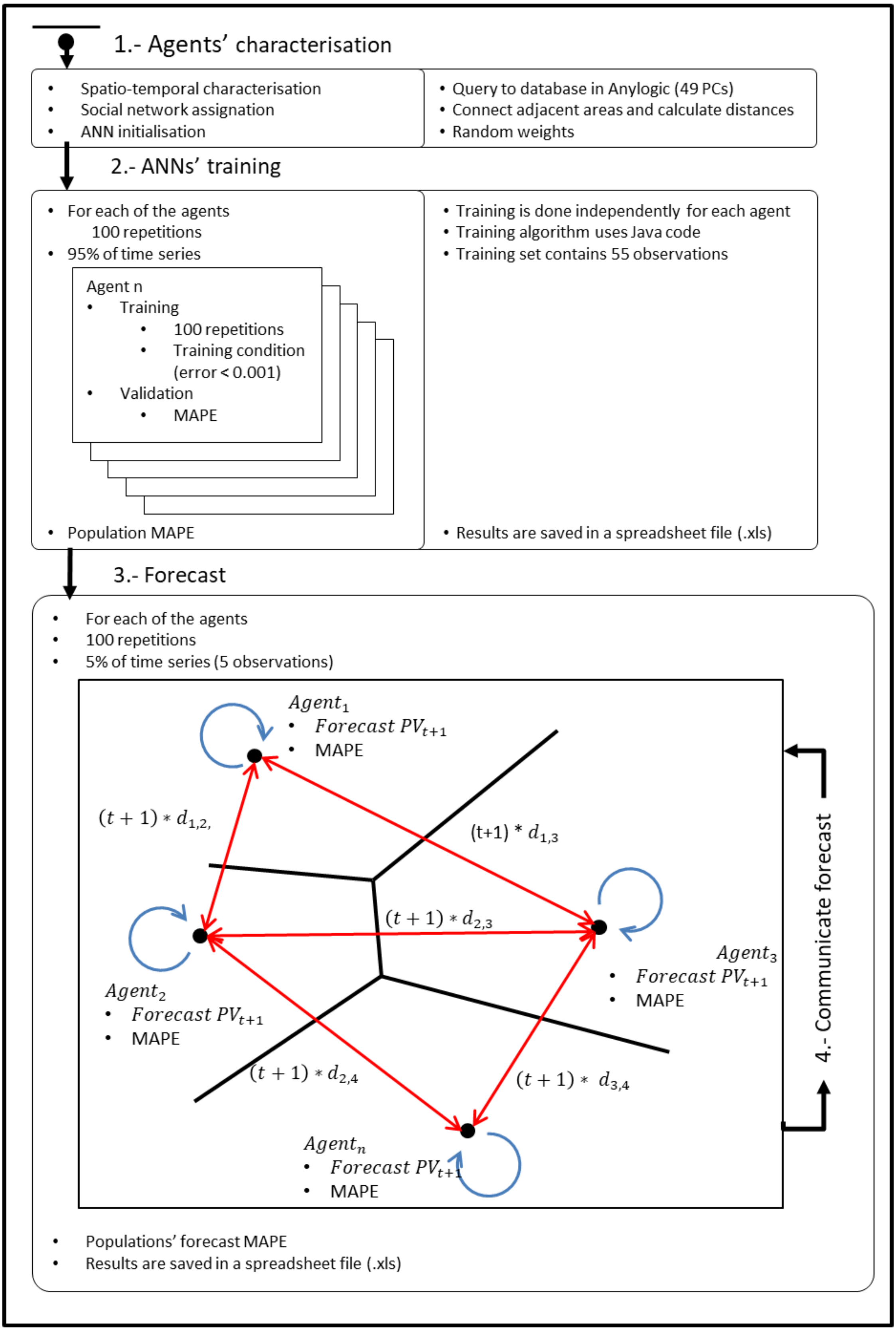


Figure 15. A four-step process for the implementation of the model.

The model was implemented using the *AnyLogic* v7.3.2 software, Java programming language for the ANN module, Python and ArcMap for the statistical tests. Algorithm 1 shows the pseudocode for the step-by-step main procedure for the simulations. Figure 16 shows the spatial layout of the agents in the Birmingham area and their social-networks.

Algorithm 1 The PV adoption process

Agents characterisation

```
for each agent PC in Birmingham do
    PC.location ← actual population centroid
    PC.PV_t ← PV installation dataset
    function AGENT_NEIGHBOURS()
        for each PC in agent_aux.Neighbours do
            agent_aux.calculatePVNeighbourhood()
            agent_aux.calculateDistance()
        end for
    end function
    function AGENT_ANN()
        for each PC in Birmingham do
            agent_aux.ANN(weight) ← random_between(0,1)
        end for
    end function
end for
```

Training

```
function TRAIN()
    for each agent PC in Birmingham do
        PC.train()
        PC.estimationError ← Mean absolute percentage error
    end for
end function
```

Forecasting

```
for each PC in Birmingham do
    PC.PV_t+1 ← PC.forecastPV()
    PC. calculatePVNeighbourhood()
end for
```

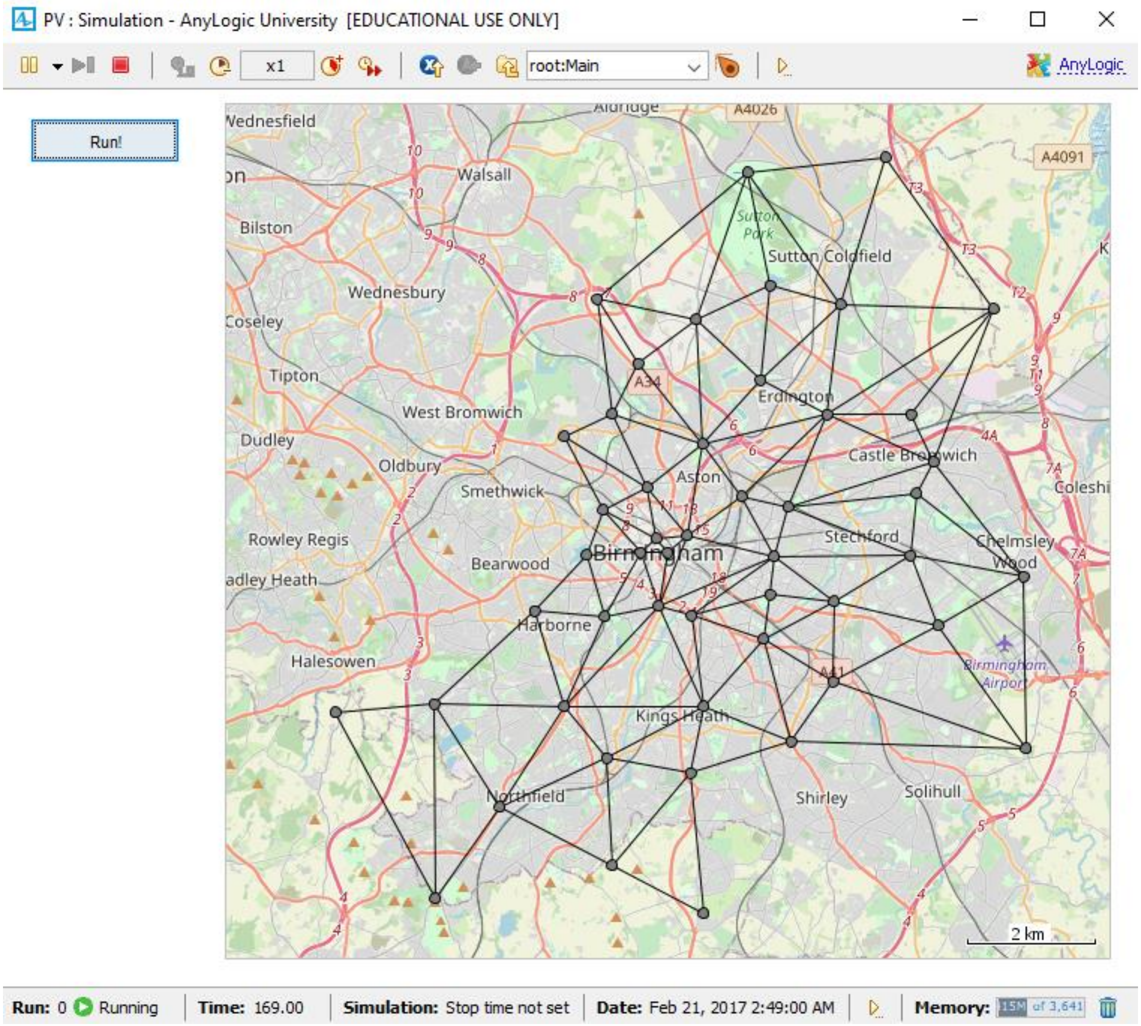


Figure 16. Implementation of the conceptual model using Anylogic software.

2.3.1 Decision-making process

The agent's decision-making is characterised by an artificial neural network, which emulates some elements of the human cognition. In the brain, this strength is modified by repetitive stimulation or by lack of activity in the specific neurons. In the computational models, the synaptic weights are fitted through a process called training. Both social effects, the influences between (i.e. spillover) and within (i.e. peer-effect) areas, are characterised independently. The spillover effect follows Balta-Ozkan, Yildirim and Connor [8], whereas the influence within areas is defined by Richter [20]. Both spillover and peer-effect are defined by the total number of PVs at a specific location and the other agents in the social network. Although the effect of risk or uncertainty on the decision-making is not explicitly considered, it has been shown the perception of risk is driven by the individuals' and peers' experiences [74]. The assumption of the lack of perfect market foresight can be extended to the assessment of risk, as individuals may not be able to calculate the probability of extreme events (i.e. drought or flooding) nor the impact on their expected benefit [74]. Then, the aversion to technology adoption due to risk and uncertainty is biased and driven by the individual experiences and those in their social networks [26,74,95]. Therefore, the model accounts for a degree of risk and uncertainty by integrating the peer-effect into the decision-making.

Social networks are defined using the adjacency criteria, whilst considering the distance between connected areas, thus, the social networks and the spatial weights are fixed at the beginning of the simulation, Then, a spatial coefficient is calculated based on the inverse distance between locations [8]. Figure 13 shows the ABM module, where each agent has a fixed position and explicit boundaries. The red lines represent the spillover effect, which is the social effect that spatially adjacent geographical areas have on each other [8], which is affected by the distance between the agents. The blue lines represent the peer effect, which is the social effect that occurs within geographical locations [20] and is not sensitive to the distance.

Neural network training

ANNs are like linear (or non-linear) least squares regression, in the sense that both attempt to minimize the sum of squared errors. In this case, the fitting takes place through a training process, which consists of presenting pairs of inputs and outputs, calculating the error of estimation, and subsequently adjusting the ANNs' weights. This process happens iteratively by splitting the datasets into three subsets: training, validation, and test. A common training approach is the Backpropagation method (BP), which is a two-phase process. During the forward phase, inputs from the training subset are passed through the neural network to determine the output. Once an output has been estimated, an error of estimation is calculated against the expected value. Then, this error is propagated through the network in the backward phase, by adjusting the weights and minimizing the error of estimation. Once the whole training subset has been presented, the model validates the ANN against the validation subset, if the ANN does not meet the stopping criteria, the process repeats; presenting the whole training subset again. Commonly, the stopping criteria are related to the level of error, in other words, if the ANN can estimate the actual data at a certain accuracy level then the training is over. The training algorithm used for this analysis is shown in Appendix 4; for a detailed mathematical description of the backpropagation see [105,116,117].

This research designs the ANNs to have a rather simple linear and sequential structure, with three layers of neurons, the input, the hidden and the output layers. A linear structure with a single hidden layer has been proven to be sufficient to approach to any nonlinear function [118] and used to forecast the PV generation [119,120]. As seen in Figure 17, the input layer (yellow) includes a node for each of the social effects for *t*-1; the hidden layer (green) comprises one neuron for each of the inputs, and the output (red) layer contains a single neuron to produce a single output. The temporal dependency is captured by the time lag between input and output, while the spatial dependency is captured by the spillover effect (via the neighbouring PV node). Then, the layers are connected to the next layer using synaptic weights, which are randomly initialized with values between [0- 1]**.** Additionally, the model includes bias nodes in the input and hidden layers. These are nodes with a constant value of (1) and can be interpreted as the β_0 in a Linear Regression; which is a constant term reflecting the intercept of the function.

To produce these outputs, the neurons use an activation function, which is similar to the biological activation threshold [105,121,122]. Because the model accounts for the decision of whether to adopt or not (a binary decision), the model implements the sigmoid function (logistic function) for the neurons in the hidden and output layers[12]. This function, presented in equation (2-2), is ideal to account

[12] No calculations are made at the input layer, this can be seen as if the input layer uses the linear function ($f(x) = x$)

for probabilities since the output values are between 0 and 1 as shown in equation (2-2).

$$f(x) = \frac{1}{1 + e^{-x}} \tag{2-2}$$

Where

x is the number of PVs in a specific PC

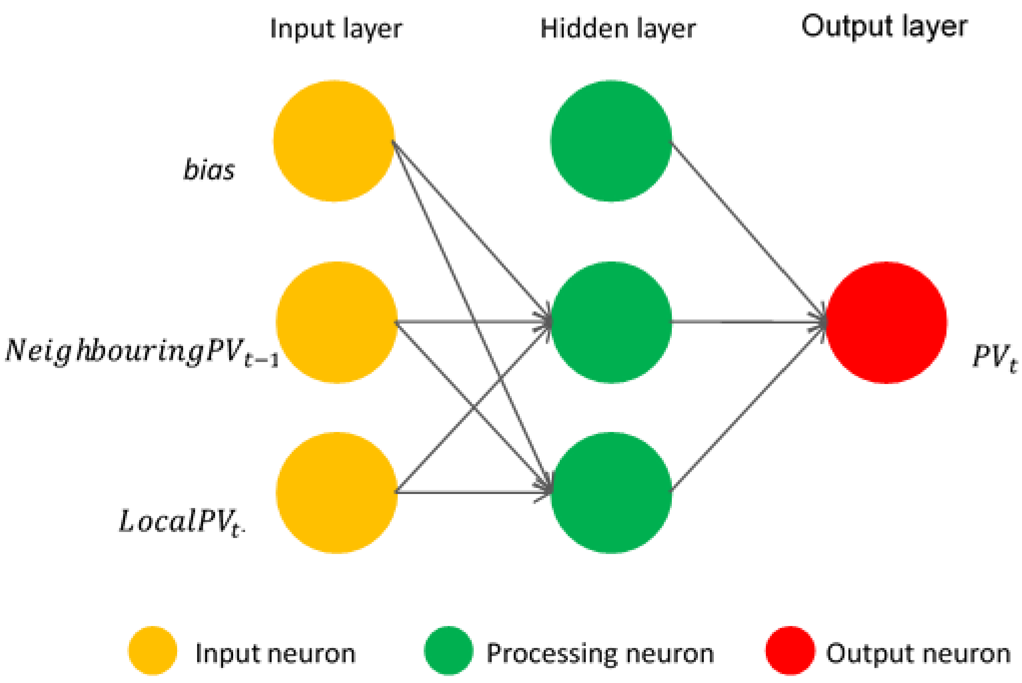


Figure 17. Artificial neural network design for a spatio-temporal explicit ABM.

The training process assesses the ANN's performance using the validation subset and evaluating the average error. If this error meets the stopping condition, the training stops, otherwise, the training continues. This validation is an internal process of the training process and is independent of the overall model validation. Instead of the common 70%, 15%, 15% split of the data for the training, validation and forecast, because of the limited number of observations (N=60), the sample is divided into 90%, 5%, 5% sets following [105,121,122]. After optimising the ANN synaptic weights, the training and the validation subsets are merged, enlarging the training set (95%-5%, 55-5 observations).

2.3.2 Model validation

The model is assessed based on its capabilities to reproduce the spatio-temporal patterns of PV adoption. Thus, its performance is measured in space and time. A common measure to assess the ANN performance is the Mean Absolute Percentage Error (MAPE) [105], which measure the overall performance over the specific time horizon (time-series). However, this calculation is for a single neural network. As each area has its neural network, this definition is applied for each of the agents whilst accounting for the population size. Individual and population MAPE's definition is shown in equations (2-3) and (2-4).

$$MAPE_j = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{PV_t - \widehat{PV}_t}{PV_t}\right| \quad \textbf{(2-3)}$$

Where

n is the time series size

PV_t is the current number of PVs in the month *t*

$\widehat{PV}_t$is the estimation of the number of PVs in the month *t*

i is the specific month (among the time-series)

j is the specific postcode district

$$Population\ MAPE = \frac{1}{m}\sum_{k=1}^{m} MAPE_k \qquad \textbf{(2-4)}$$

Where

m is the population size

k is the specific area (among the 49 PCs)

Additionally, as a benchmark for comparison the analysis estimate s-curves using the same datasets, these are presented where relevant. The Bass model is commonly used to model the adoption of innovation [40,42], yet this does not consider the spatial dependence and have limited capabilities to account for the temporal dynamics of the adoption process [37]. Appendix 6 summarises the methodology followed; the errors of estimation are calculated in the same way as those for the ANNs.

2.4 Results

2.4.1 Temporal validation

Figure 18 compares actual PV numbers against the estimations produced by the model and those from the Bass model. As seen, both models' estimations exhibit different behaviours over time, for instance, the Bass model is closer to the actual data at the beginning and at the end of the time series. Yet the rate of adoption is nearly linear with large deviations in the mid-periods. On the other hand, the ANN estimation follows the actual PV installations more closely most of the time, yet its forecasts are largely overestimated in comparison with the Bass model. In general, the ANN has better capabilities to replicate the temporal dynamics, with an accuracy around 89% of underestimation against the 82% of overestimation of the Bass model by the last months of the training phase. The model's performance is ~5% lower than those from the reviewed ABMs [25,76,90,91], yet, those models cannot inform about the local estimation rates. The errors of estimations are shown in Figure 19, where the ANNs are more likely to produce extreme values because at the beginning of the training neural network they have been fed with a small proportion of the time-series data [118,121–123]. Yet, the MAPE decreases over time and stabilises at the end of the sample. However, the errors of estimation for the forecasted periods start accumulating and then diverge.

The MAPE for the first forecasted values are like the model's training, but because the estimation is affected not only by the own agent's error but by those in their social networks, it quickly reaches 70% by the fifth step (further details are in section 4.3). Similarly, the Bass model also presents large errors when the ANNs do, however, its errors are at least 20% larger. Then, at the end of the time-series, the errors decrease and diverge. These results of the autoregressive model are then compared with those of the models in Chapters 3, 4 and 5 in Section 5.3.3 and considering those from the literature in Section 6.1.4. Despite both models having a similar behaviour at the end of the training period, the performance differs more than 20%. The ABM integrated model being an iterative method aims to reduce the estimation error at each time step, adapting to the data behaviour over time. However, the learning of the model depends on the amount of data that has been presented to the ANN, having the best performance at the end of the training. On the other hand, the Bass model minimises the overall estimation error, meaning that the estimations with high accuracy will offset the estimations with low accuracy. Then, on average the model minimises the estimation error but has limited capabilities to adapt to local changes in the data behaviour.

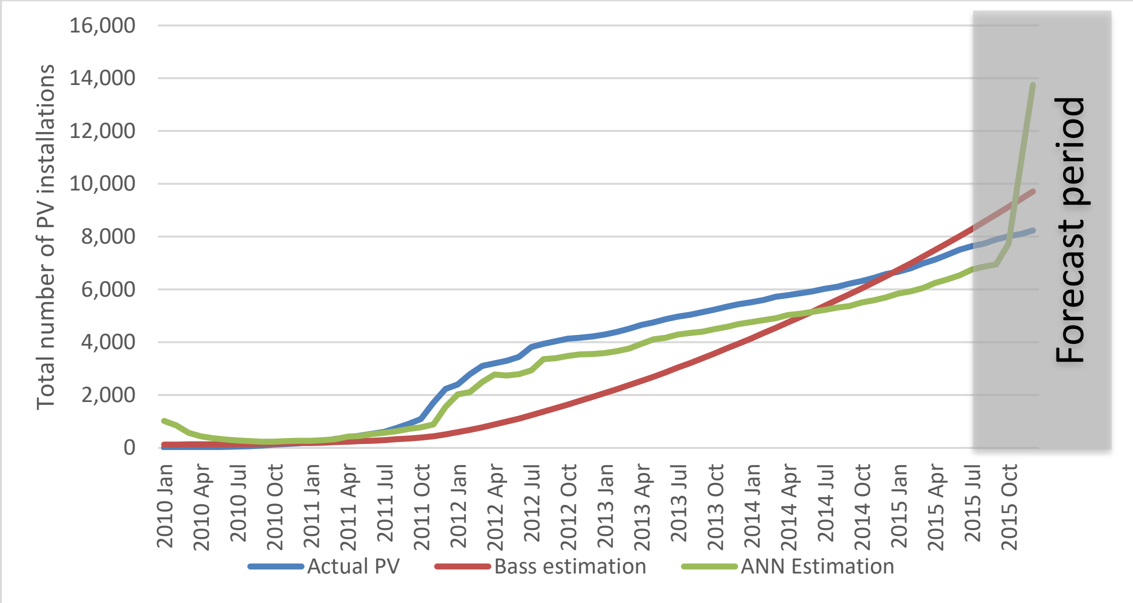


Figure 18. Cumulative PV adoption rates estimated by ANN and Bass model vs. actual data.

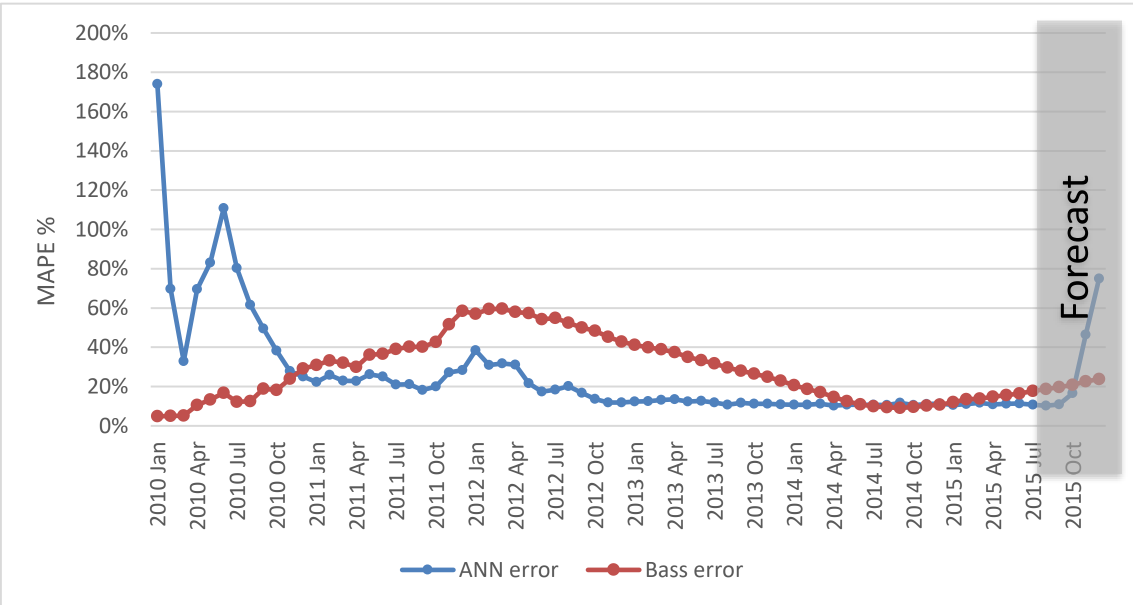


Figure 19. The error of estimation for the ANN and Bass models.

The errors of estimations present some peaks, which suggest shifts in the agents' behaviour in late 2011s, yet, the ANNs can adapt to these changes and adjust to the new data trends. To identify these disturbances, the analysis calculates the absolute marginal change over time and the average of those changes, using equations (2-5) and (2-6):

$$Absolute\ marginal\ change_i\ =\ |MAPE_i - MAPE_{i-1}| \tag{2-5}$$

$$Average\ marginal\ change = \frac{1}{n}\sum_{i=1}^{n} Absolute\ marginal\ change_i \tag{2-6}$$

Figure 20 shows the concentration of the marginal changes that were higher than the average change. As seen, there are three points where ~20% of the PCs present peaks in their estimations. The concentration of these changes suggests that some of the agents' behaviour is not captured by the autoregressive model and that there might be other factors that are driving this shift which can be captured by a multivariable model. It is been noted that one of the drivers for the PV adoption is the Feed-in Tarrif's (FiT) financial incentive [20,124–126], thus, it may be argued that these disturbances may be caused by the variation in the FiT levels. For instance, the residuals present one disturbance at the end of November 2011, matching with the revision and announcement of reducing such incentive. However, these could be also caused by another external event or by the variation in the population socioeconomic characteristics. Therefore, Chapter 3 analyses such disturbances after extending the model to a multivariable model, so disregarding the effect of other socioeconomic variables. Nevertheless, the model can adapt to those changes and stabilise the errors of estimation by the end of the training process. Chapter 2 aims to increase the accuracy of the model by introducing socioeconomic data and investigates its impact on the errors over time.

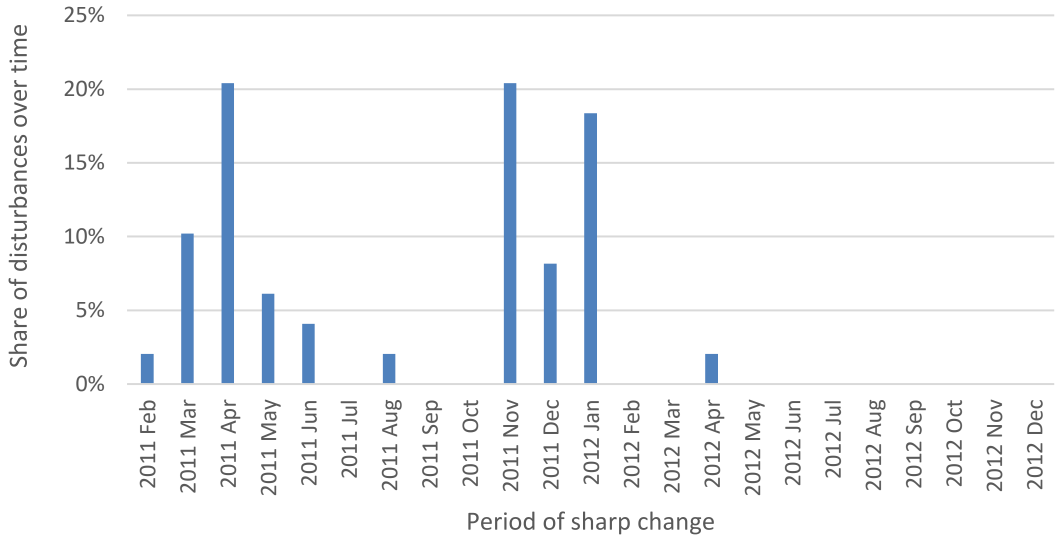


Figure 20. The temporal pattern of the marginal changes of the estimation errors overtime for the Autoregressive model.

2.4.2 Spatial validation

Figure 21 and Figure 22 show the distribution of the errors of estimations for both ANN-ABM and Bass models, respectively. As seen, the Bass model yields a more uniform distribution of the errors, with most of them between 11% and 30%. On the other hand, the ANN estimations present a more random spatial distribution of the estimation errors, with most of them between 0% and 10%. Despite the model has around 90% of accuracy and the errors tend to converge at the end of the training, the distribution of the estimation errors presents spatial regularities. Figure 21 and Figure 22 displays the largest estimation errors in the PCs in the city centre. Since these PCs have a low density of residential buildings, they also have a low or null number of PVs.

Then, because the MAPE associated with small numbers produces large errors even with relatively small under or overestimations, results are analysed by considering the spatial regularities of the errors instead [127,128]. For a clearer picture of the spatial patters of the estimation errors [16,29]. Figure 23 and Figure 24 present the maps for the hot spot analysis for both estimations. In both cases, most of the areas present non-significant regularities except for the central PCs, meaning that the MAPE is mostly randomly distributed and that the model can capture most of the spatial dependence [128]. Then, the clusters exhibited by the estimation errors correspond to the areas with low residential buildings, as mentioned before. Therefore, further analysis could focus on identifying local spatial association terms that account for these regularities [127]. In the following chapter, the model is extended by introducing socioeconomic variables to the agents' characterisation and investigate the effect on the capabilities of the model to capture the spatial dependence.

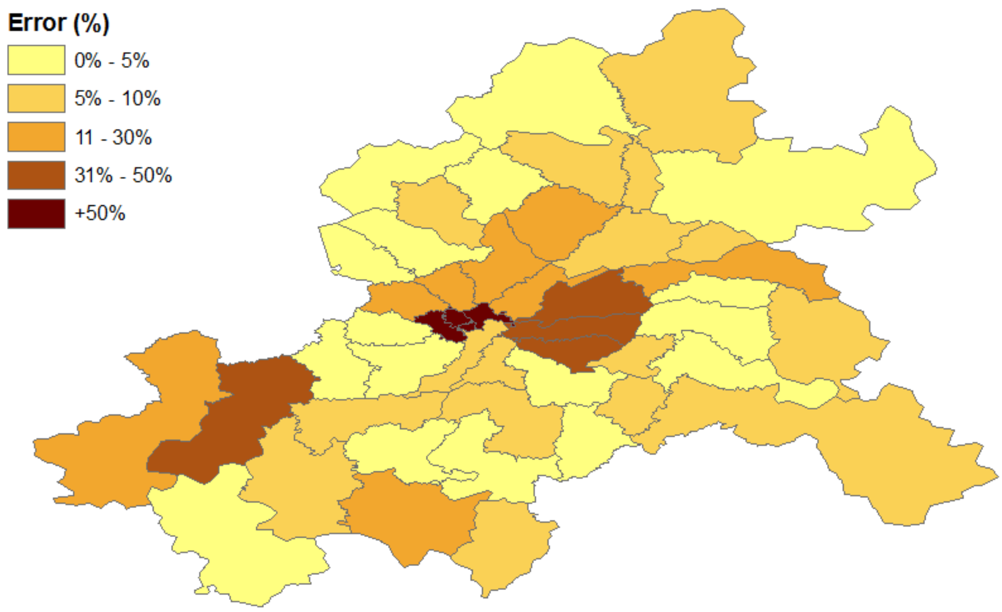


Figure 21. Spatial distribution of the ANN estimation's error by the end of the training - Jul 2015.

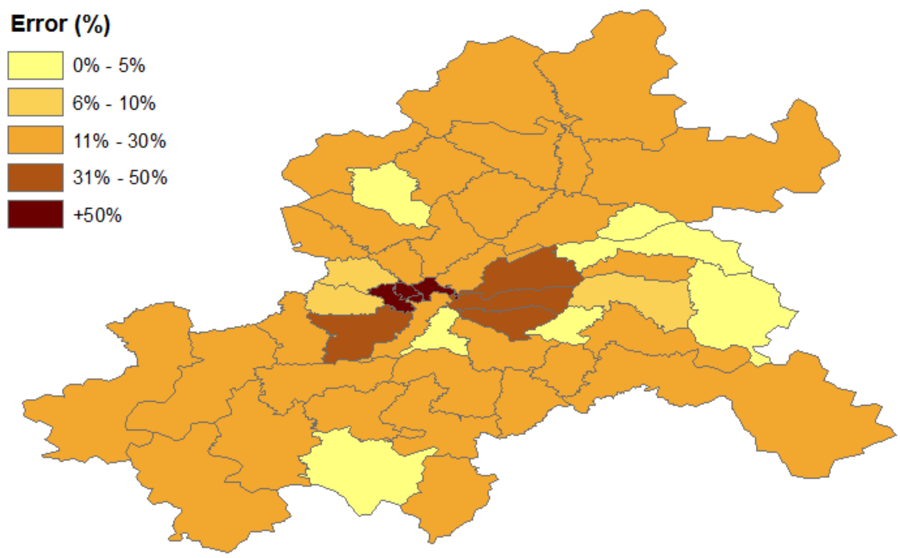


Figure 22. Spatial distribution of the Bass estimation's error by Jul 2015.

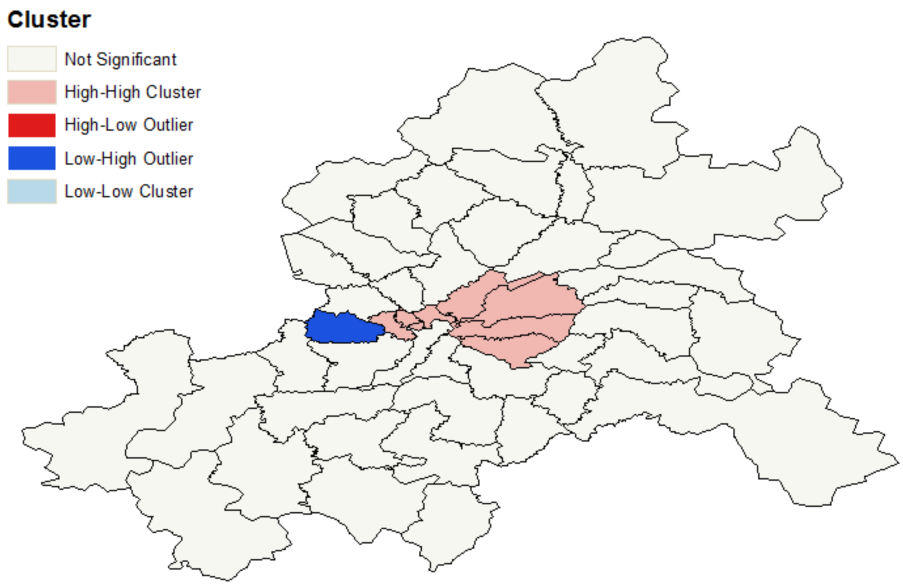


Figure 23. Hot spot analysis of the ANN estimation's error by the end of the training - Jul 2015.

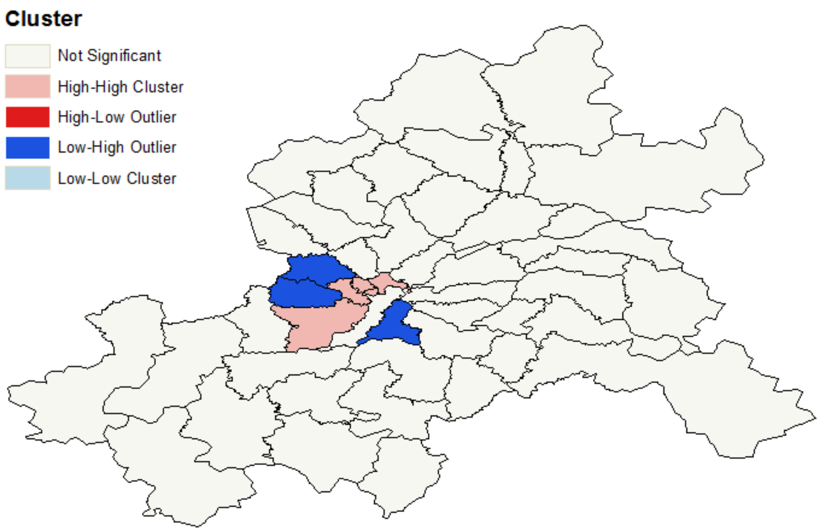


Figure 24. Hot spot analysis of the Bass estimation's error by Jul 2015.

2.4.3 Predictive accuracy

This section presents the forecasts, the ANN's estimations by PCs are shown in Figure 25a, Figure 25b, and Figure 25c. For a better visual comparison, only the MAPE for the 1st, 3rd and 5th forecasts are displayed, corresponding for August, October and December 2015; whilst the Bass model's forecasts are seen in Figure 26a, Figure 26b, and Figure 26c. During the first month of the forecast, most of the areas have an error below 10%. However, by the fifth forecast, the error almost doubles, resulting in almost 80% of the agents having more than 10% errors in estimations. Thus, while the model can capture the spatio-temporal nature of the PV adoption process, it is only able to produce short term forecasts. As seen in Figure 18, because of the error accumulation the model has a better estimation during the first three periods than the Bass model. In particular, the Bass model requires an explicit time horizon for the whole potential adopters of PVs to adopt. In other words, because the Bass model uses a time horizon further than 2015, the forecast months are also included in this period and the estimation is also minimised for them. Instead, the forecast of the ANN is done dynamically step by step after the training, considering the previous forecast to forecast the *t+1-th* period, which causes the accumulation of errors.

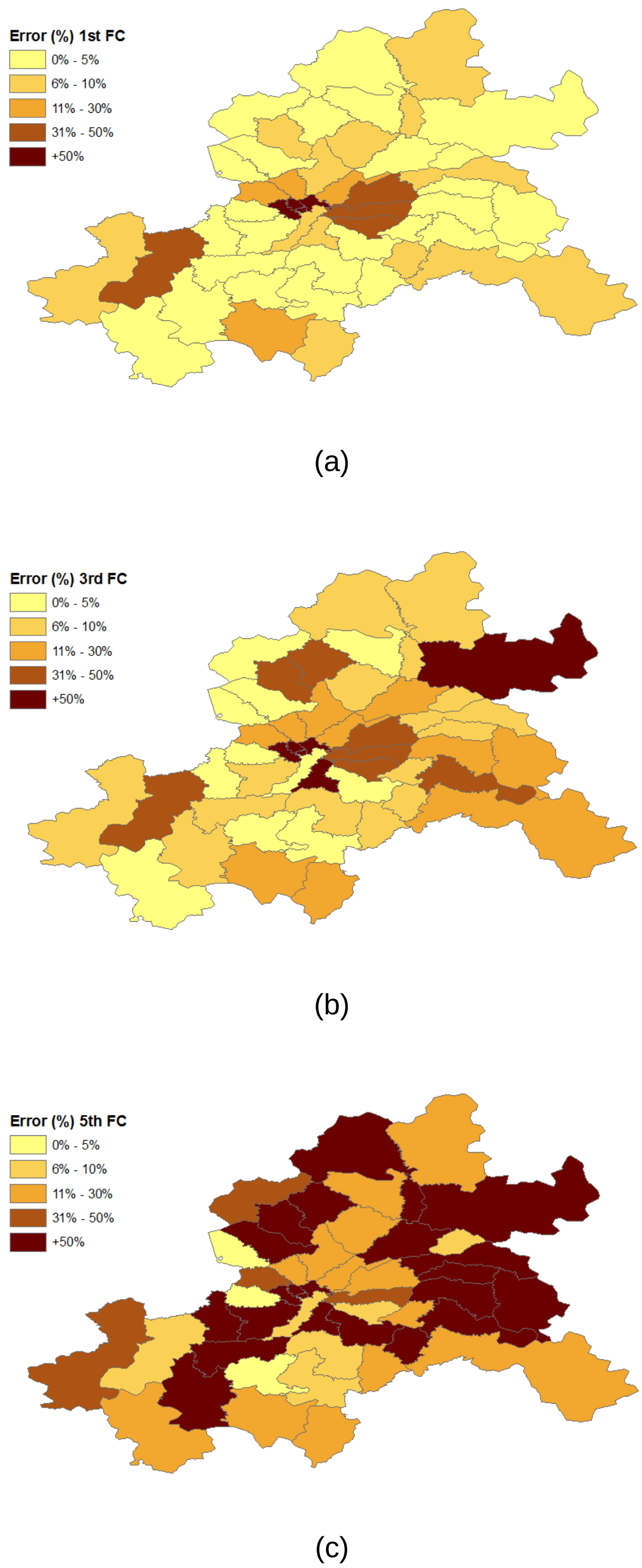
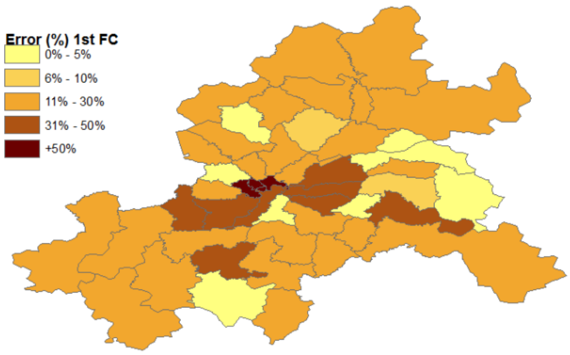
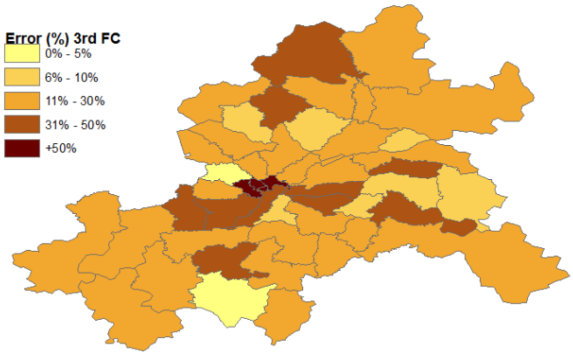


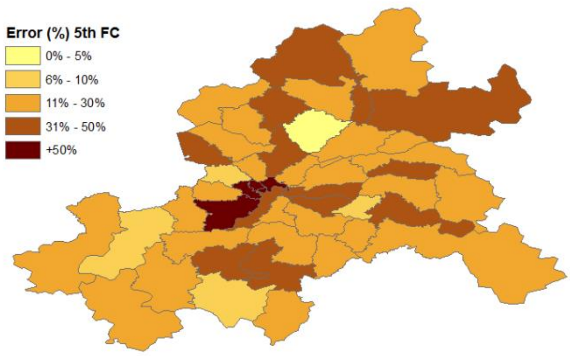
Figure 25. Spatial distribution of the ANN estimation's error – (a) First, (b) Third and (c) Fifth forecasted period.



(a)



(b)



(c)

Figure 26. Spatial distribution of the Bass estimation's error – (a) First, (b) Third and (c) Fifth forecasted period.

2.4.4 Agents explicit spatio-temporal characterisation

The modelling literature of the PV adoption has overlooked the spatial dependence of the drivers whilst focusing on the social dynamics and the possible emergent behaviours. Despite ABMs' strength in providing insights on emergent system behaviour, two main limitations are still outstanding: use of *rational choice*-based decision-making [72–74], and synthetic characterisation of their temporal dynamics, rather than utilising actual time horizons.

The former is limiting the approach as some of the drivers such as social utility have subjective value, and individuals rarely possess perfect market information [43,72–74]. Instead, the ANNs generate knowledge by considering the previously-made decisions. The second limitation is on the potential of the ABMs to inform about possible pathways for the PV development (emergent behaviour), to state the actual time horizon or a specific time for the adoption to take place. Instead, this spatio-temporal explicit ABM can provide forecasts for each location at specific times.

Then, the model is validated temporally and spatially based on the model performance, which is measured by the population's average Mean Absolute Percentage Error. The following sections discuss (i) the specific results for the spatio-temporal characterisation of the agents, (ii) the modelling of social effects, and decision-making, (iii) the model's validation, and (iv) the potential applicability of the model. The last two subjects are discussed in the interest of informing the next chapter and the thesis objectives

Although the spatial accuracy of the model inputs is important [27], the gathering of this information for entire populations is a challenge. Therefore, those ABMs that simulate data or use semi-empirical data, due to the high data-intensive needs, may present a loss in the spatial accuracy of results [25]. Alternatively, the model defines agents as geographical areas, recognising that they have similar socioeconomic characteristics, common interests and present similar behaviours [101,102]. The use of spatially explicit data at postcode district level allows the modelling of real-world layout, without excluding any locations. This spatio-temporal characterisation also captures the change of social effects over time. However, this characterisation is not without imperfections. Yet, because of the aggregated nature of the data, there is potentially a loss of heterogeneity in the individuals' own decision choices in an area.

The spatial resolution is driven by the data availability and temporal variability within it, highlighting a trade-off between these two elements. While higher spatial resolution improves spatial accuracy, this is at the expense of lower temporal variability in the adoption rates. Conversely, higher temporal resolution results in less data variability whilst the number of observations increase. Additionally, the model developed keeps the real layout and does not exclude any area to ensure the accuracy of spatial layout, contrary to Richter [20] who excludes the areas with close to zero installations from the analysis. However, the results and insights presented are specific for the Birmingham city. The area of study is bounded, and the model assumes no external effect from the surrounding areas of Birmingham city. Another issue is whether a spatially explicit model like the one developed in this research would be subject to the *Modifiable areal unit problem*, which suggests that the change in scale and data resolution may introduce bias to the study [15,72,83,84]. Therefore, the model scalability could be assessed in further analysis extending the area of studies, including more PCs.

Compared to other ABMs, this approach enables the simulation of system behaviour for discrete time series. Hence, the model allows the investigation of the evolution of the adoption patterns spatially and over time and recognises any variation within them.

2.4.5 Social effects and decision-making process

Our model assumes that individuals with similar decision-making, attitudes and interests tend to spatially cluster [16], creating social networks [77] and that knowledge is generated through experience-based learning [102,129,130]. On the other hand, in conventional applications of ABMs, the agents are assigned with an adoption rule or adoption probability at the beginning of the simulation. Yet, as the actual behaviour in a specific location is still unknown to the modeller (early stages of adoption), it is argued that assigning an adoption rule to areas in early stages of adoption may increase the uncertainty. Alternatively, an aggregated characterisation of the agents and the use of ANNs can allow modelling the entire population. Such an approach could disregard the specific individuals' preferences, by generating knowledge over time. The model takes into consideration the historical data, even for those areas with a low number of PVs. This can be seen in the error produced by the model, as the performance of ANN increases as it is fed with more data and starts adapting to the changes in data behaviours, i.e. as individuals have experienced more situations [118,121–123].

Additionally, the aggregated characterisation of the agents allows the ABM to provide a more realistic representation of the decision-making process. The characterisation of agents as geographical areas allows integrating the ANN as their decision-making, as the high data demand from the ANN is reduced [131]. Besides, adjusting the strength of the spillover effect as a function of the distance between the areas reflects the spatial dependence of the PV adoption. Thus, the model accounts for both the influence between and within locations [20]. Then, the model aggregates the number of PVs registered by month, then the social influence is defined by the number of PVS in the adjacent areas on monthly basis. Richter's [20] econometric model estimated a lead time between adoption decision and installation to be between two and three months. This difference could be because the econometric model looks at the parameters that best fit the overall data, whilst the data analysis carried out in Section 2.2.3 looks at the most common behaviour among the ACF and PACF. Moreover, the PACF shows no significant lags at t-2, t-3, t-6 or t-12, pointing out the lack of any seasonal effect.

2.4.6 Model validation

The model is tested and validated for Birmingham city area. The errors of estimations over time are used to assess the temporal accuracy, whilst the Moran's I index and clustering maps were used to evaluate the spatial accuracy of the model estimations (see Table 5, Figure 23 and Figure 25). The results suggest that the model can capture the spatio-temporal dependence of the PV adoption process. The model training error stabilises and converges to ~10% MAPE, yet, there are peaks at the same time as the data present changes in behaviour over time. This implies that there is a behaviour that is not captured by the model, suggesting the need for a multivariable model. As highlighted by Samarasinghe [105], the autoregressive models may increase their accuracy by introducing exogenous variables such as income, electricity consumption, etc.

2.4.7 Predictive capability

The model has a ~90% of accuracy for the training period, yet, the errors in the forecasting phase rapidly increase. This restricts the model's capability to yield long term forecasts, making the model's forecast only relevant at most three months forward, as the 3rd forecasted period's accuracy is still above 80%, and better than the Bass model. The reviewed models in Section 1.3 have a performance around 95% of accuracy [25,76,90,91], resulting in a ~5% lower for the ABM and ANN model, nevertheless, the results of the ABM and ANN are spatio-temporally explicit. Thus, it could be argued that trying to predict the specific behaviour of smaller areas (other than national or county) is more challenging. Moreover, because the estimation of new values considers not only agent's past decisions but also those in its social network, a better understanding of the error accumulation is needed, as this accumulates not only over time but also across space. In the case of the reviewed models, it is not clear whether the results consider this error accumulation or what techniques are used to handle or minimise the effect.

As data availability increases over time, future work may consider longer data sets to improve the predictive capability and include other socio-economic data to characterise the agents. Because this is a data-driven approach, there could be some issues with the data, for instance (i) disjointed and inconsistent data sources, (ii) different temporal resolution and spatial reference, (iii) poor quality or incomplete data. Thus, future analysis should consider how to inform the model of possible bias due to the quality of the datasets or instrumental errors. Despite these limitations, the model still has the potential to inform the industry. For instance, by the time the FiT report was published on 31 Dec 2015, the results could have advised to the DNOs about the extra load that new solar panels would produce and potentially bring into the network a quarter later. Yet, because of its short-term forecast, this does not have much of potential to inform strategic planning, yet, it still has the potential to inform demand and supply balancing, for instance, where to locate calls for demand-side response strategies and flexibility tenders[13].

2.5 Reflective summary

This chapter designs, developments and empirically validates a novel spatio-temporal explicit agent-based model that integrates artificial neural networks into the agent's decision-making. The model advances ABMs by characterising the spatio-temporal dependence of the PV adoption process, whilst improving the agents' decision-making procedure. This approach does not only build upon ABM literature but also draws from disciplines such as SR. The use of spatially explicit data sets allows reflecting individual behaviours of each location and follow the real-world layout for the city of Birmingham. The model utilises the ANN's capabilities to approximate historical PV data in the generation of knowledge and adapts to changes in data trends. Therefore, the model can reduce uncertainties in the agents' decision-making.

The results suggest that the model can account for the spatial, temporal and social dynamics that drive the adoption process. Furthermore, the ABM and ANN model can produce a better estimation than a Bass model, this could be because this model does not consider the social effect nor the spatial dependence of the adoption process. In principle, the spatio-temporally explicit forecasts could inform network planning and investment decisions of the energy industry. Yet, this potential currently is limited as the model is only able to produce short-term forecasts due to limited availability of data for 60 time periods.

Future work could include the development and validation of a multivariable model, to improve the model accuracy and produce longer-term forecasts. Other future work could also investigate if the model is flexible enough to handle different space-time resolutions and social network structures, and handle data from other diffusion process or systems that are affected by the social effects.

[13] Here you can refer to UKPN and WPD flexibility tenders - https://www.ukpowernetworks.co.uk/internet/en/have-your-say/listening-to-our-connections-customers/flexibility-services.HTML

Despite these limitations, the idea to integrate ANN and ABMs for the first time and develop spatio-temporally explicit, empirical results offers a new method to capture the complexity of the energy system decarbonisation process.

3 Dynamic characterisation of population heterogeneity in a spatio-temporally ABM and ANN integrated model

3.1 Introduction

The previous chapter outline how to integrate explicitly the spatio-temporal and social dynamics into an ABM for the adoption of PVs, developing a model that uses ANNs as the decision-making process. Although the model reaches 95% of estimation accuracy over the training period, the model is only able to forecasts three months ahead. The chapter demonstrates that the model disregards the importance of multiple socio-economic factors that might influence agents decision-making [51,67,104,132]. Additionally, the model disregards the agent's heterogeneity (socioeconomic characteristics) and assumes no inputs from outside the analysis area. The ABMs capture agents heterogeneity by characterising their different socioeconomic characteristics, also the utility or social threshold allows outlining the differences in preferences towards PVs [67,104]. Alternatively, the SR captures the population's heterogeneity with the independent variables, which may present spatio-temporal regularities [72,133,134]. Yet, both SR and ABM have limitations to characterise the dynamic nature of those characteristics.

Buchmann, Grossman and Schwarz [67] suggest three ways to implement the heterogeneity of the agents: (i) different characteristics or preferences. (ii) different adoption rules, and (iii) multiple agents. Therefore, this chapter extends the characterisation of the agents' decision-making, by integrating socio-economic variables into the model. This is due to a lack of data to characterise multiple agents (i.e. PV sellers or PV producers), and the fact that the ANN is supposed to create individual adoption rules (knowledge). Moreover, because of the spatio-temporally explicit nature of the model, the model also implements the local differences and evolution of the population heterogeneity. This approach aims to address the limitations of the ABMs that fix the values of the agents' preferences at the beginning of the simulation and disregard its evolution. Additionally, the analysis assesses the flexibility of the model to handle changes to the geographical boundary of the model, by systematically changing the number of agents in the simulations. It is expected that adding socioeconomics variables to the model can yield a better performance and potentially longer forecast [105,115].

Namely, this chapter's objectives are as follows:

1. To extend the model's decision-making characterisation by introducing socioeconomic variables to the autoregressive model and reflect their spatio-temporal dynamic.
2. Assess the effect of the *Modifiable areal unit problem* by systematically broadening the study area.

The chapter is organised as follows: the remainder of this section present relevant literature used to implement the changes to the model. Then, section 3.2 presents the implementation of such modifications and the potential variables to be included. The results of the extended model are presented and discussed in section 3.3. Finally, section 3.4 is devoted to reflecting on the findings of the chapter.

3.1.1 Characterisation of heterogeneous agents

The adoption decision-making is driven by factors such as energy prices, government policies, peer-effects, age, gender, income, vehicle price, incentives, etc. [8,15,18,68,86]. The agents' decision-making process is usually characterised by a utility or social function, and a threshold that represents the criteria of whether to adopt or not. The utility function considers the financial benefits of adopting PV, such as electricity savings or financial incentives. On the other hand, the social function is subjective, reflecting the agents' personal believes, values, and social benefits [25,33]. Both approaches, utility and social functions capture the individual's preferences, providing a robust representation of the population's heterogeneity.

The authors implement those with using fixed values, statistical distributions, or statistical regression. Most of those approaches consider previous analysis and data surveys, which are used to create population segmentation or estimate the preferences to adopt certain technology. For instance, Bale et al. [79,132]

assigned fixed weights based on population segmentation, using priority surveys. Krebs and Ernst, and Ernst and Briegel [75,77,78] create weights using the continuous uniform distribution, using empirical data (survey) and type of lifestyles segmentations (taken from previous surveys). Robinson et al. [76,90,91], on the other hand, estimate the weights using an ad hoc spatial regression model based on survey data. Regarding the adoption criteria, authors alike Bale et al. [79,132] define the adoption threshold as the perceived financial utility, fixing this value at four levels: low, mid, high, and not able to adopt (0.25, 0.45, 0.75 and 1, respectively). Krebs and Ernst, and Ernst and Briegel [75,77,78] implement the expected financial benefits (i.e. payback) by considering the household characteristics. Robinson et al. [76,90,91], calculate the perceived financial affordability against the investment's payback. These criteria are calculated at each step of the simulation, and if the condition is met the agents adopt PV.

On the other hand, SR does focus on identifying the importance of socio-economic variables on the adoption process. Similar to the ABMs, the correlation coefficients calculated for each independent variable are similar to the preference weighting assigned to each of the agents' characteristics. The definition of these independent variables and the characteristics of agents utilised in ABMs vary across different studies, and there are no clear criteria for variable selection. However, Bale et al. [79] and De Groote, Pepermans and Verboven [51] recognise that data availability limits the number and type of agents characteristics and the selection of independent variables; for the ABMs and SRs respectively. Among the reviewed studies, some variables are implemented in similar ways, for instance, solar radiation (resource availability), as in the case of [8,18]. Other variables such as *income* are characterised in different ways across different studies: the gross domestic product [8], gross regional product per capita [18], mean income [12], the median income [12,16,19,46], mortgage vs. income ratio [14], population share by income bands [85]. Interestingly, the SR is likely to include endogenous variables, such as age, distribution of the population by gender or income. Instead, the ABMs tend to include exogenous variables, such as house characteristics or energy cost, and subjective variables, such as social and environmental utility [25,33]. Political tendency and environmental awareness are included only in the SR. These observations highlight the potential gains from merging SR insights into the ABM to make the best use of available data, as SR can inform about the variables that are statistically significant for the PV adoption. Data selection, collection and processing procedure is fully described in Section 3.2.1, whilst a detailed summary of all variables included in the analysis is presented in Appendix 5.

3.1.2 Modifiable areal unit problem

Because the ABM and ANN model has a spatially explicit nature, it is potentially affected by changes in scale. Despite the ABMs present a wide range of agents characterisation and study scale, given the non-spatial nature of the ABM, these studies do not provide any evidence of being resilient to changes in scale. Moreover, because the model looks at the total number of PV systems, the local regularities of the adoption process are disregarded. Therefore, the study does not inform whether the survey is representative at the local level nor whether the model can fit the agents' parameters at the local level.

On the other hand, the SR is directly affected by the study scale and resolution, as statistics and correlation depend on the number and size of areas that compound the study in its whole [72,133,134]. There is a wide range of areal definition (scale) and size of units of analysis (resolution) available in the literature, ranging from local areas (census tracks or political boundaries) to standardised statistical units (NUTS3[14]), and scales such as counties or countries. [8,12,15,16,18,20,46,47].

Similarly, the ABM studies characterise the adoption process in different study scales (number of agents). For instance, Adepetu and Keshav [38] create a semi-empirical population which is assigned with real values for 100 households. Then, the authors arbitrarily populate their simulation with 26,160 agents, scaling down the number of PV contracts. Rai and Robinson [40], Robinson and Rai [41], and Robinson et al. [42] model the total households from Austin (Texas, US). Cui et

---

[14] Nomenclature of Territorial Units for Statistics NUTS3.

al. [46] generate a virtual population of households based on actual aggregated data. By applying the copula-based household synthesiser, they generate an individual virtual household population with similar characteristics (with intra-group variance); resulting in the simulation of 190,965 agents. Eppstein et al. [47] create a full virtual population of households. First, they generate a virtual distribution of income, considering five hypothetical cities, and then they create 1,000 and 10,000 agents (for initial model runs). The authors choose 1,000 as the final number of agents due to computational efficiency and the lack of significant difference between the simulation using both numbers. These agents are created using the *turning bands* method, which simulates random observations using the annual salary covariance function.

One can argue that the ABM may not be fully subject to the MAUP, however, this raises the question of whether the emergence of behaviour is dependant to the number of agents in the simulation, more specifically, the effect on the spatio-temporal patterns of adoption. Despite the wide variety of ABM implementations, the effect of the number of agents on the model's performance is not discussed entirely.

3.2 Methods and materials

The model seeks to represent the dynamic nature of the heterogeneity, whilst assessing whether the change in scale has an impact on the model outputs. Therefore, this chapter uses time-series of socioeconomic variables to capture the evolution of the agents' preferences.

As seen in Figure 27 the analysis is carried out in two phases. **Phase I** builds upon the autoregressive model, using the same spatio-temporal resolution and ANN design. Then, this phase redefines the decision-making process considering a multivariable characterisation, including socioeconomic variables and updating the values at each step of the simulation. **Phase II** carries the multivariable characterisation of Phase I and systematically increases the size of the study area. This increment first considers the adjacent PCs to the autoregressive model, and then the PCs in the adjacent LADs to Birmingham, resulting in three multivariable models. The results of Phase I are compared with those from the autoregressive model, whilst those of Phase II consider the results of Phase I.

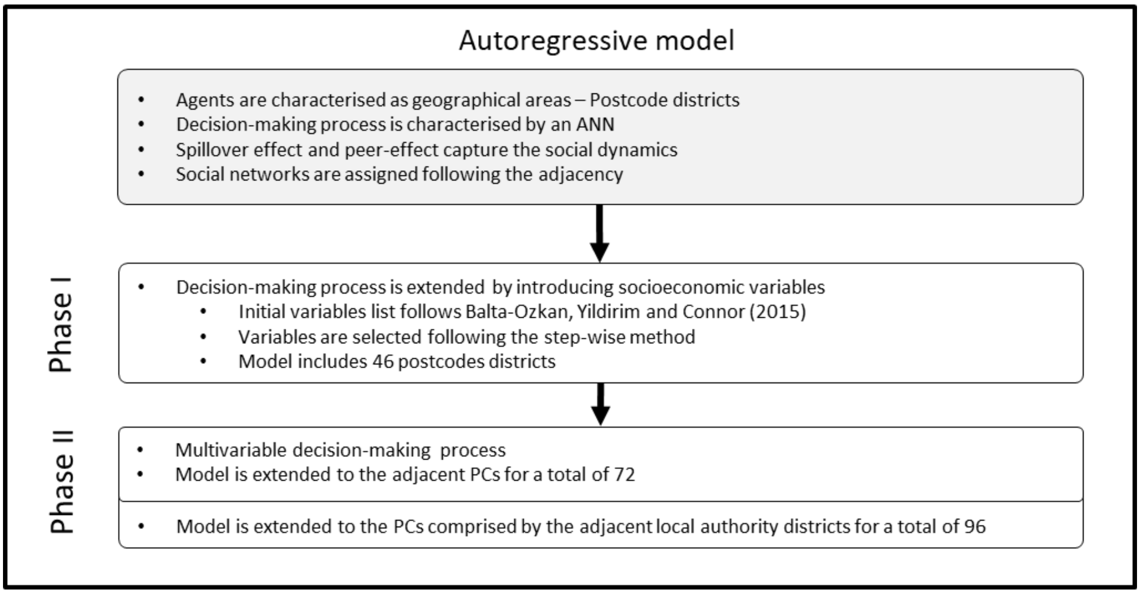


Figure 27. Methodology to characterise the evolution of the population heterogeneity, and assessment of MAUP effect.

Figure 28 shows the progression of the three areal definitions for this chapter.

(i) **MV-Birmingham** – The first model (black perimeter) comprises the base area of the city of Birmingham, made up of 49 PCs

(ii) **MV-Extended** – The second model (blue perimeter) keeps the multivariable characterisation, and cover the adjacent PCs to the **MV-**Birmingham, for a total of 72 PCs.

(iii) **MV-LADs** – The third model (red perimeter) extends the previous model to include 96 PCs. These comprise the adjacent LADs: Lichfield, Tamworth, North Warwickshire, Solihull, Bromsgrove, Dudley, Sandwell, Walsall.

As seen, the southern PCs does not increase further than the blue perimeter, because the North Warwickshire, Solihull, Bromsgrove PCs have been included already in the **MV-Extended**. Data corresponding to PV installations covers from January 2011 to December 2015. When referring to three models as a group, they will be called *multivariable models.* Figure 29 shows the implementation of the spatial distribution of the agents and the configuration of their social networks in the Anylogic v7.3.2 software.

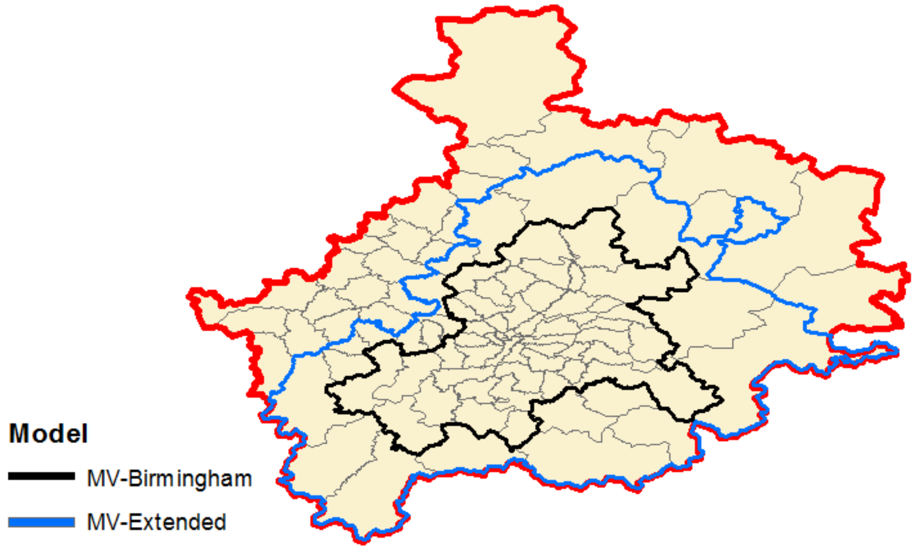
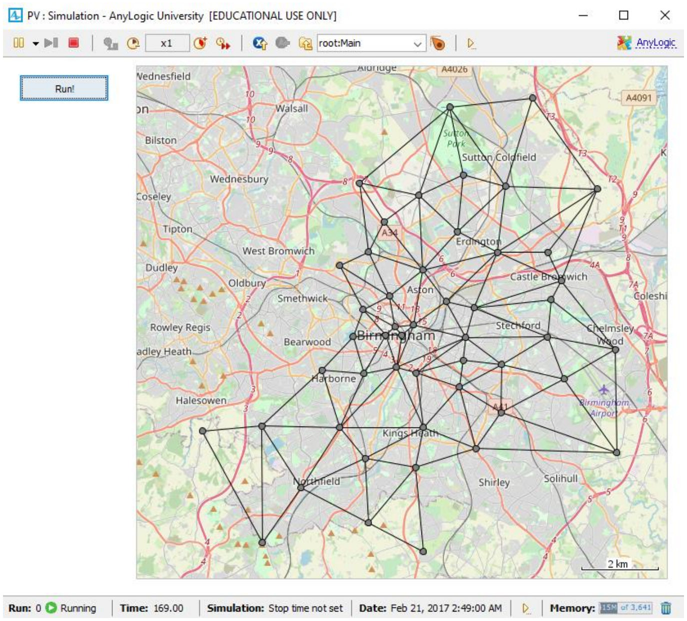
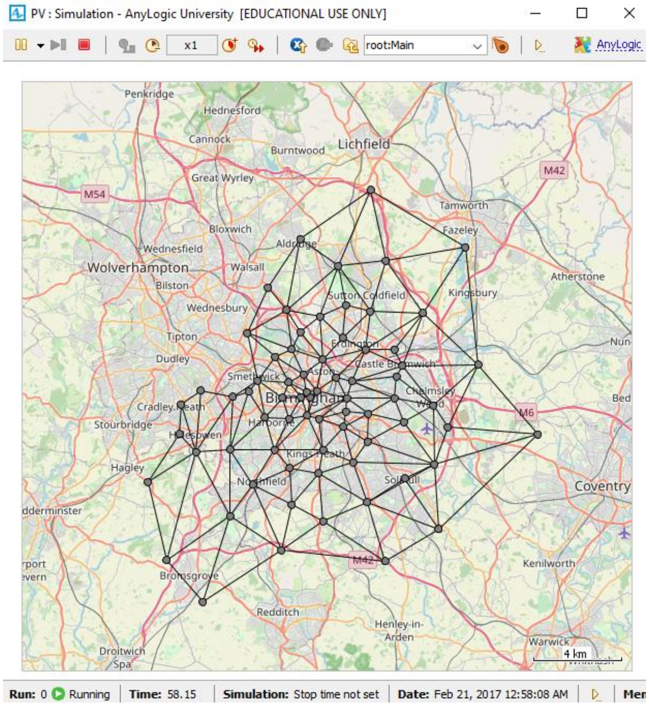


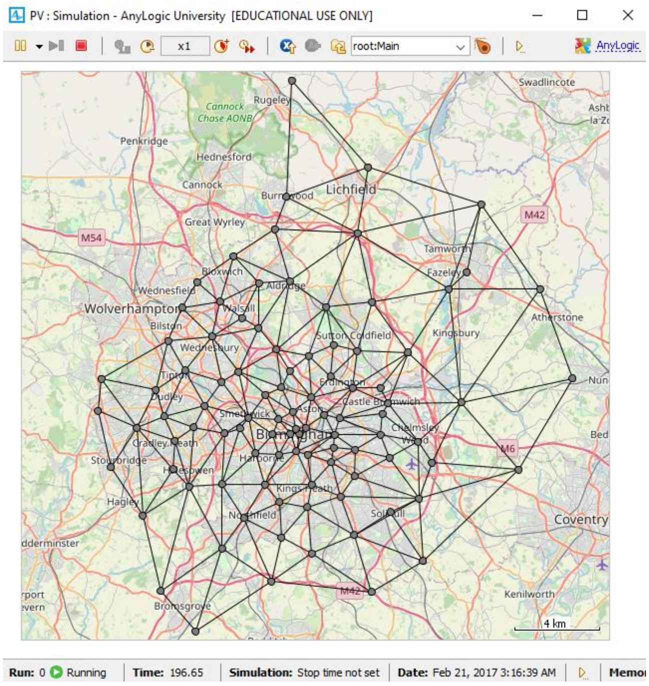
Figure 28. PCs included in the different study scales for the multivariable models.



(a)



(b)



(c)

Figure 29. Implementation of the model in Anylogic software – (a) MV-Birmingham, (b) MV-Extended, and (c) MV-LADs.

3.2.1 Decision making and social effects

To abstract the multivariable definition of the adoption process, the ANN's design is modified as seen in equation (2-1).

$$PV_t = f(LocalPV_{t-1}, NeighbouringPV_{t-1}, Socioeconomic_{t-1}) \qquad \textbf{(3-1)}$$

Where

PV_t is the total number of PVs in a specific time

$LocalPV_{t-1}$ is the autoregressive element

$NeighbouringPV_{t-1}$ is the number of PVs in the adjacent areas

$Socioeconomic_{t-1}$; is a set of independent variables

The temporal dependence is denoted by lagging these inputs in *t-1*, and the set of variables is discussed in the following section.

3.2.1.1 Variable selection

The literature reviewed in section 1.2 offers a wide variety of variables included in the analysis, furthermore, some of these variables are implemented in different ways depending on the data availability. The full list of potential variables included in the literature review is summarised in Appendix 5, these are grouped into broad categories to identify similar variables with different implementations. However, because this is the first attempt to integrate socioeconomic variables into the ABM and ANN model, the analysis builds upon Balta-Ozkan, Yildirim and Connor's spatial-econometric model [8] for the diffusion of PVs in the UK. Then, the analysis follows the econometric stepwise method to select the final list of variables [21], starting from their initial list of socioeconomic variables. This method starts with a single variable and assesses the model fitness (MAPE in this case), then the variable with the best fitness is fixed for the next iteration. In a reiterative process, a new variable is introduced and the fitness assessed, discarding if it does not improve the fitness; then stopping when the fitness can't be improved any further. The evolution of the local socioeconomic variables may also be affected by households changing address, thus, the model may also capture a degree of the changes in the structure of the social networks.

3.2.1.2 Data collection and processing

Table 7 shows the summary table of the potential variables included Balta-Ozkan, Yildirim and Connor's model [8], together with the available spatio-temporal resolution and data source. As seen in Table 7, most of the variables are available at PC level and Census basis. An exception to this is the PV installations, weekly income, electricity consumption, and CO_2 emissions. To meet the spatial resolution requirements, the variables at Lower Layer Super Output Area (LSOA) or Medium Layer Super Output Area (MSOA) level were aggregated to PC level using Office of National Statistics reference lookup tables[15]. The only exception is the solar irradiation data because the changes in solar irradiation between one area to another, in this study scope, is negligible[16].

On the other hand, because the temporal resolution of some of these variables' is not in the required resolution, the data is interpolated to produce monthly observations. This process is carried out following the UK Office of National Statistics' (ONS) methodologies for temporal disaggregation [135]. Temporal disaggregation is a process that generates a time series at a higher frequency from data with a lower temporal resolution. Monthly gross domestic product (GDP) observations have been estimated from the annual time-series, by applying Fernandez 's technique and using the Index of Services[17] [135]. Then, following ONS's methodology, monthly observations are estimated with the Fernandez algorithm. As this estimation is at the national level, the Index of

[15]Office of National Statistics Source: https://ons.maps.arcgis.com/home/item.html?id=ef72efd6adf64b11a2228f7b3e95deea

[16]The distance between Birmingham and Edinburgh is ~394km, whilst the horizontal solar irradiance changes only by less than 10% [125]. Yet, the longitude of a cross section of Birmingham is around 30km.

[17] The Index of Services measures the quantity of output from all UK services industries, and accounts for more than three-quarters of the output approach to the measurement of Gross Domestic Product.

Services is replaced by an index of house pricing, which has a high spatio-temporal resolution (monthly/LAD).

PV data is available in the Feed-in Tariff Installation Report published by the Office of Gas and Electricity Markets (Ofgem) on a quarterly basis, which contains the registration date (dd/mm/yyyy) of domestic PVs. Socioeconomic data is published by the Office of National Statistics, containing data such as income and homeownership which are available for 2001 and 2011. Appendix 7 summarises the statistics of the potential explicative variables, differentiating between the smallest and largets. The boxplots fo the variables show that as more PCs are included, the statistics fluctuate. It is expected that the ANN adapts to the wide variety of data behaviour, improving the confidence in the model.

Table 7. List of independent socioeconomic variables and their resolution.

Variable	**Spatial resolution**	**Temporal resolution**	**Data points**	**Data source**
PV installations	LSOA	Daily	2011-15	OFGEM
Weekly income	MSOA	Weekly	2013, 2015	ONS
Population density	PC	Census	2001,2011	ONS
% Owned household	PC	Census	2001,2011	ONS
% Detached household	PC	Census	2001,2011	ONS
Electricity consumption	LSOA	Annual	2001, 2011-15	ONS
Education level	PC	Census	2010,2015	ONS
Average household size	PC	Census	2001,2011	ONS
CO_2 emissions proxy	LSOA	Annual	2001,2011	ONS

3.2.1.3 Neural network

This chapter builds on the ANN structure presented in Section 2.2, increasing the number of input neurons, as seen in Figure 30. The number of neurons for the $SocioeconomicVar_t$ will depend on the number of socioeconomic variables, these are included or excluded as discussed in Section 3.2.1.1. The initialisation of the simulation follows the process described in Section 2.2.5, except that the query to the database brings both the time-series for the number of PVs and socioeconomic variables. The training follows a similar process as the autoregressive model, except that the number of neurons has increased. Thus, the algorithm repeats the adjustment of the weights for the PV inputs for the weights of the socioeconomic variables, yet, these weights are independent and unique. The training period uses 95% of the time-series (55 months) and 5% for the forecast (5 months); the testing and validation of the model keeps the same process as presented in Section 2.3.2.

3.3 Results

The results from the characterisation of heterogeneous agents are shown first. It is expected that the model's performance will improve, whilst allowing a longer forecast. Samarasinghe [105] notes that a multivariable model allows the ANN to have multiple predictors for the same phenomena, therefore, the accuracy of the model may improve. Additionally, previous results (Bass model and autoregressive model) are shown where relevant. Second, the results for the *extended models* are presented and compared with the original model, to evaluate the effect of increasing the number of agents in the simulation.

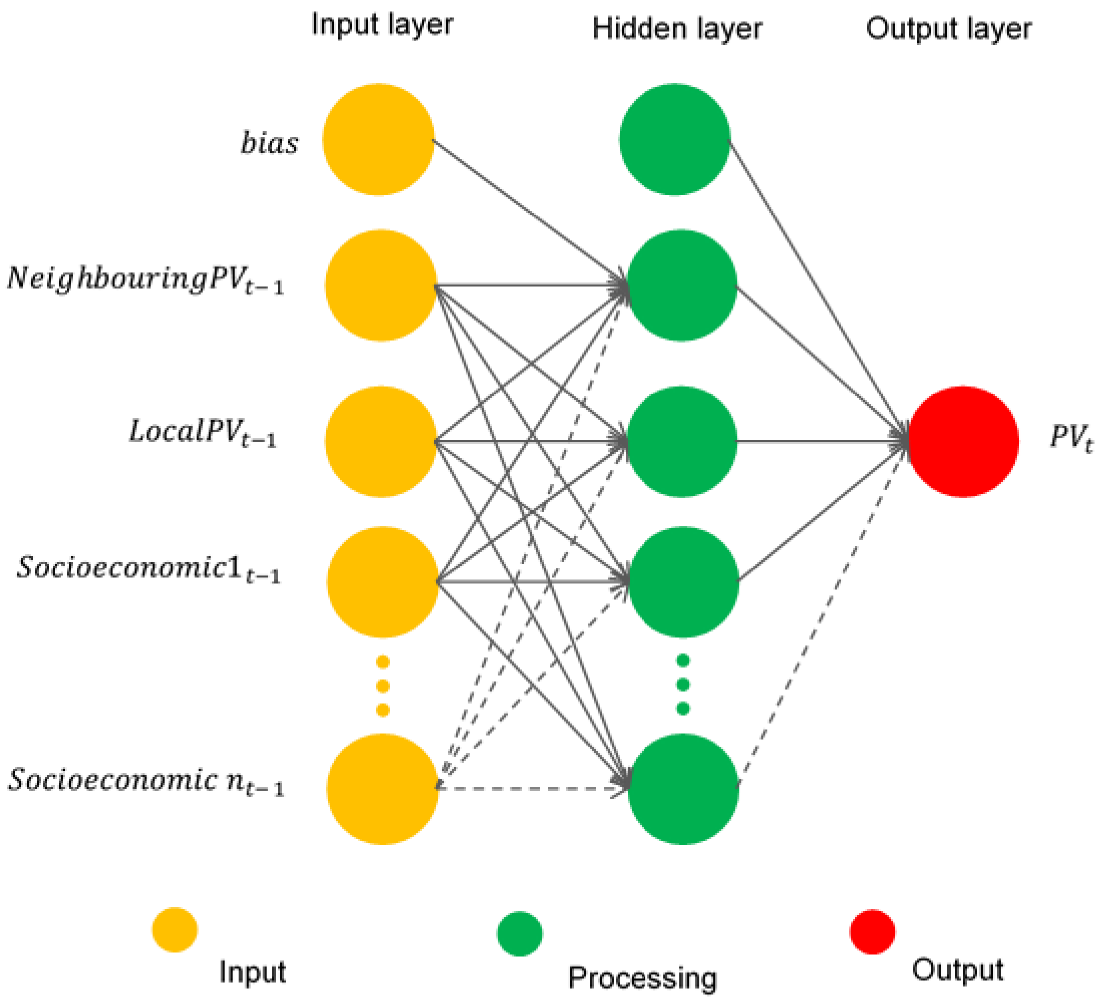


Figure 30. Example of an artificial neural network fed by the autoregressive elements and the socioeconomic variables.

3.3.1 Temporal validation

Table 8 shows the stepwise process of variable selection, which in this case stops at the fourth step. During the first step, the **income** variable is selected as this improves the model accuracy the highest, by 5.2%. During the second step, two variables present the same improvement, **electricity consumption** and the **average household size**, both of which are carried to the next step. Step three selects the model with both variables, whilst step 4 shows no improvement when introducing any other fourth variable; thus the process ends.

The results highlight that income, electricity consumption and average household size are the variables that yield the best fitness. This combination of variables increases estimations' accuracy from 90% to 95%, compared with the autoregressive model. Drawing from the SR, one can argue that the extra inputs work as the explanatory variables, in the way that the SR uses the independent variables to explain the variance of the dependent variable. Then, the output of the SR is the value of the intercept (β_0) plus the effect of each other variable. In the case of the ANN, the extra inputs and their synaptic weights contribute to the estimation of the output, contrary to the SR, the process to understand these values is not transparent.

Nevertheless, the results can be explained through a contextual analysis of whether the selected variables are in line with those that have been proven to drive the adoption process in SR studies. For instance, **income** has been used to define the agents' utility or social threshold [25], in line with [8,94,136] finding that income is a key decision variable for households to adopt the PV technology. The **electricity consumption** variable reflects the findings of households with high energy usage are more likely to be concerned about being self-sufficient [8], this is extended to the energy cost [16]. **Average household size** has a statistically significant negative impact on PV diffusion, as bigger families may have less cash flow for a PV installation [8,14,19]. The increase in fitness from 90% to 95% suggests that increasing the agents' heterogeneity improves model performance. This could be because the decision-making does not consider only the experience with the PV technology or the social influences, but is also driven by the socioeconomic characteristics of the individuals [30]. For instance, the affordability and expected benefits depend on the individuals' socioeconomics, but the perception of these is influenced by the social interactions [8,20] and past experiences [31]. Furthermore, updating the socioeconomic data at each step provides a more realistic decision-making process than that characterisation that does not update the agents' variables (see Section 3.1.1).

Table 8. Results summary from the stepwise process of variable selection.

MAPE	GDHI proxy	Population density	% Owned household share	% Detached household share	Electricity usage	Education level	Avg household size	Irradiation	CO2 proxy	Criteria
11.0%	Autoregressive model									
Step 1										
5.80%	X									Fixed
9.40%		X								Carry
8.60%			X							Carry
9.40%				X						Carry
9.00%					X					Carry
9.20%						X				Carry
8.80%							X			Carry
10.40%								X		Out
9.00%									X	Carry
Step 2										
6.2%	X	X								Carry
5.8%	X		X							Carry
6.2%	X			X						Carry
5.6%	X				X					Fixed
6.4%	X					X				Carry
5.6%	X						X			Fixed
6.6%	X								X	Out
Step 3										
5.8%	X	X			X					Carry
5.8%	X		X		X					Carry
6.6%	X			X	X					Out
5.8%	X				X	X				Carry
5.4%	X				X		X			Fixed
6.0%	X	X					X			Carry
6.4%	X		X				X			Carry
5.8%	X			X			X			Carry
6.8%	X					X	X			Out
Step 4										
6.0%	X	X			X		X			Out
6.2%	X		X		X		X			Out
6.2%	X				X	X	X			Out

Figure 31 shows the estimations made by the Bass, autoregressive and multivariable models, where the last two presenting similar behaviour. Because of the change in the number of PCs between the autoregressive, **MV-Extended** and **MV-LADs**, the results are not fully comparable. The results of the multivariable and autoregressive (See Section 2.4.1) are similar, even though the multivariable model improves the accuracy and reduces the disturbances arguably related to the FiT changes. Figure 32 displays the error of estimation for the multivariable model, and the Bass model and autoregressive model as references. The ANN estimations keep an estimation error of ~5% over most of the training, especially at the end of the training where both histograms stabilise. Both the Bass and the ABM are likely to produce extreme values at the beginning of the training, as the neural networks have not been fed with much information yet [118,121–123]. Also, during the first half of the training, there are disturbances in the MAPE, which matches with the period[18] of maximum Feed-in-Tarif rate and the announcement of its upcoming reduction [124]. On the other hand, the Bass model presents a different behaviour, this is because this model looks at the time-series as a whole and adjusts its parameter to minimise the overall error. In contrast, the ANN adjusts its parameters (synaptic weights) at each time step. Therefore, the Bass model has limitations to reflect changes in data trends, whilst the ANN uses its adaptive capabilities to learn from them.

[18] Between October 2011 and January 2012.

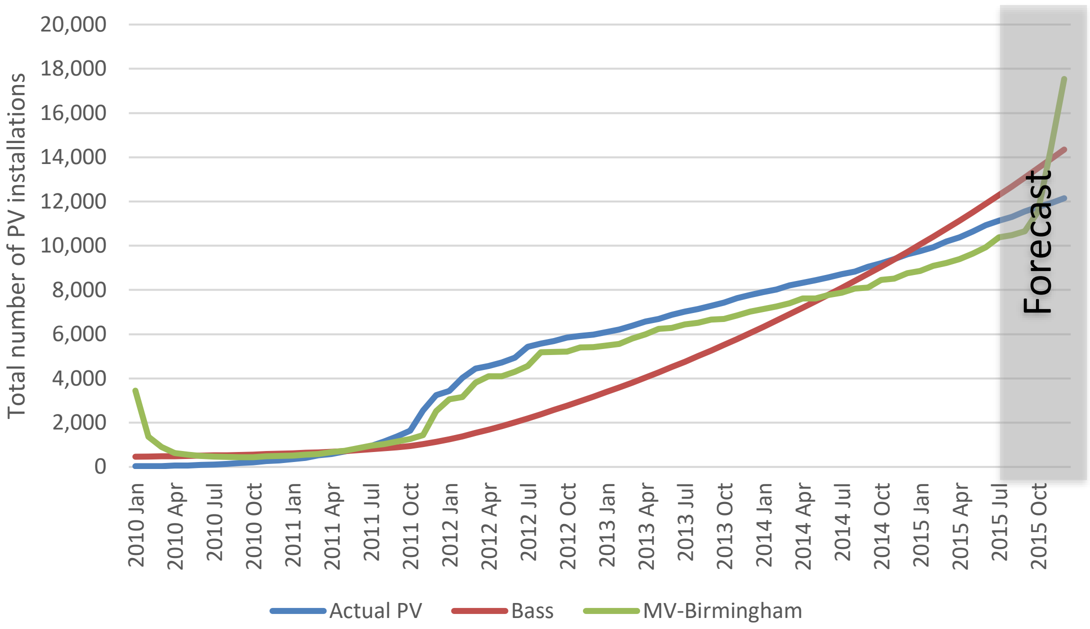


Figure 31. Cumulative adoption rates of PVs estimated by the MV-Birmingham and the Bass model vs. actual data.

Further analysis of the residuals confirms a stronger temporal pattern, clarifying the origin and nature of the extending disturbances exhibited by the MAPE histogram. Following the same approach for the autoregressive model, equations (3-2) and (3-3) are used to calculate the marginal changes month-to-month:

$$Absolute\ marginal\ change_i = |MAPE_i - MAPE_{i-1}| \quad \textbf{(3-2)}$$

$$Average\ marginal\ change = \frac{1}{n}\sum_{i=1}^{n} Absolute\ marginal\ change_i \quad \textbf{(3-3)}$$

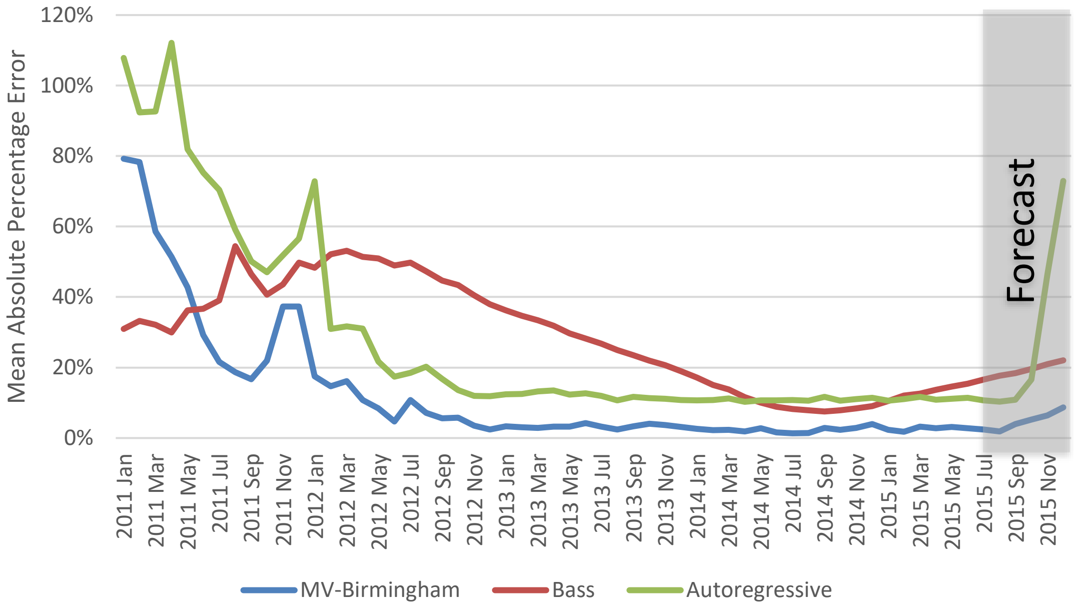


Figure 32. Estimation error of the autoregressive and MV-Birmingham, and the Bass model.

Figure 33 shows the temporal patterns of the marginal changes that were higher than the average marginal change. The autoregressive residuals show three major disturbances or sharp changes in the agents' behaviour. In contrast, the multivariable model presents only one major disturbance at the end of November 2011, matching with the changes in FiT rates, highlighting improvements of the latter model to capture more of the predictable data behaviour. Because these disturbances are not present in all the PCs, it can be argued that the effect of the FiTs rates may present spatial regularities only for some of the PCs. This highlights the importance of considering local socioeconomics when designing new policies. Figure 34 shows the marginal changes of each PC from October to November 2011, reflecting the local responses to the FiT. As shown, the most responsive areas are those at the edges of the city, while PCs near the centre are less responsive. Because the multivariable model has already accounted for the effects of **income**, **household size** and **total electricity consumption**, one can argue that this response can be associated with some of the excluded variables. For instance, the share of the owned house [8,14,18] has been found to have a negative effect on the adoption of PVs, reflecting that renting households are less likely (indeed incapable) of adopting.

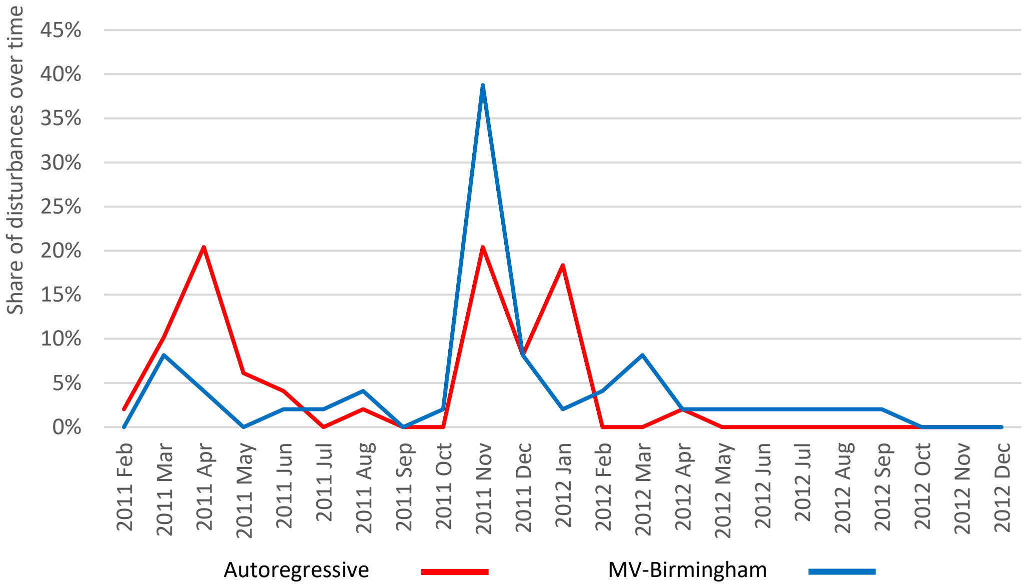


Figure 33. The temporal pattern of the marginal changes of the estimation errors overtime for the MV-Birmingham model.

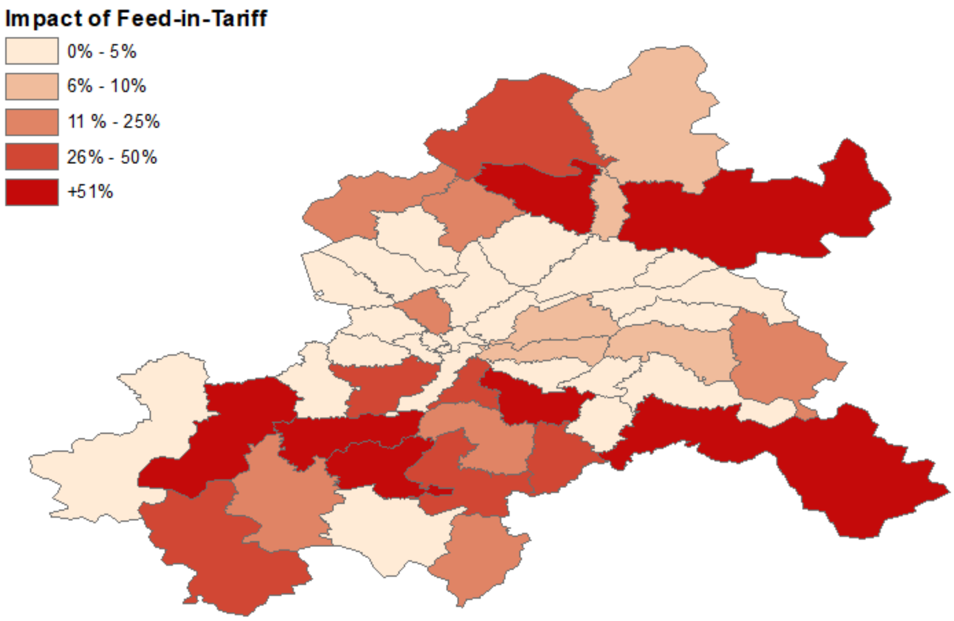


Figure 34. Spatial distribution of the impact of FiT on the adoption rates by November 2011.

The same rationale is followed for the **MV-Extended** and **MV-LADs**, which keep the same multivariable characterisation of MV-Birmingham (income, electricity consumption and average household). However, because this model increases the number of PCs, the total number of PVs are not fully comparable. Instead, Figure 35 presents the comparison between the estimation errors, as these are relative and comparable. The fact that there is no more than 1% difference between the three models' performance confirms that the ANNs' adaptive capabilities make the model resilient to the changes in geographical scale. The models stabilise at similar levels, having 95%, 94% 94.5% of accuracy at the end of the training period (55 observations). Moreover, because the time series are spatially explicit, results suggest that characterising the evolution of the population's heterogeneity improves the modelling of decision-making. On the other hand, the estimation errors of the Bass model are larger than those of ANN especially when the trend in the data changes. after this point, the estimation errors decrease to similar levels to the ANN. Despite both the multivariable and the Bass model overestimating the adoption rates, the ANN reduces the error accumulation and produces smaller errors than the Bass and autoregressive model.

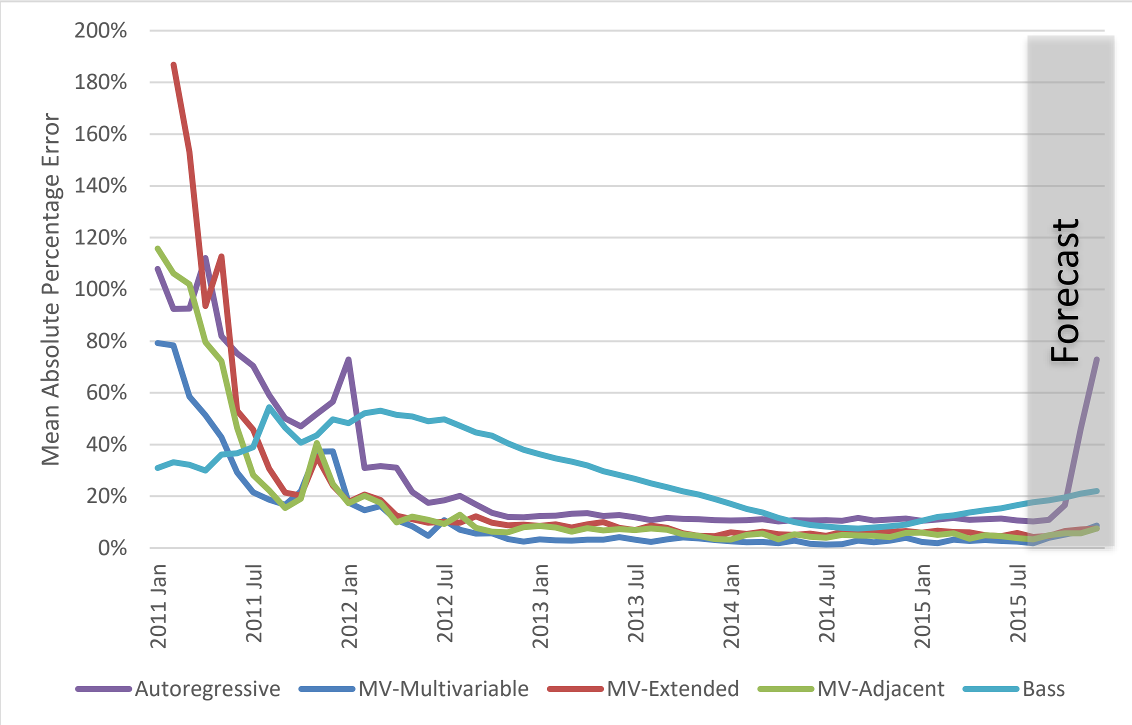
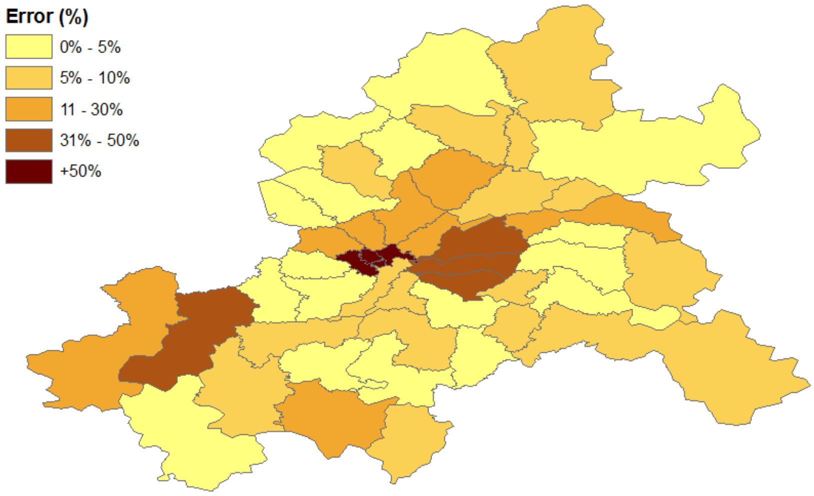


Figure 35. Estimation errors for the multivariable and Bass models.

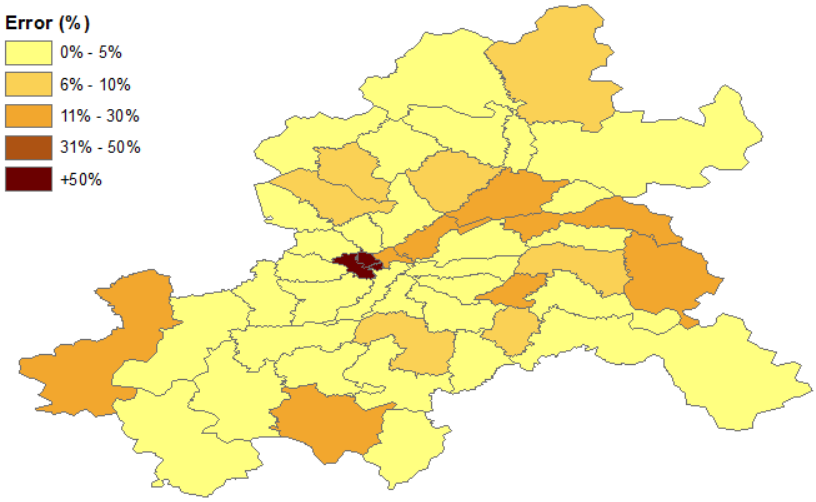
3.3.2 Spatial validation

This section examines the spatial regularities of the results previously presented, as seen in Figure 36a and Figure 36b, most of the areas in the **MV-Birmingham model** present an improvement in their estimations in comparison with the autoregressive model, having a MAPE below 10%. Yet, the areas with the largest errors (+25%) are the ones with less than five PV installations which are in line with Richter [20]. This behaviour is similar to the autoregressive model, as the MAPE calculation is sensitive to any minor change in small numbers, such as in the case of the central PCs, where the relative under or overestimation in one unit will produce a larger error than in areas with a higher number of PVs. Moreover, the errors of estimation stay at low levels even when the area of analysis (the number of agents) increases. Figure 37 shows the spatial distribution of the MAPE for the ***multivariable models***, suggesting that the models are resilient to the change in scale.

For the **MV-Birmingham** and **MV-LADs**, smallest and largest models, most of the areas do not present spatial regularities, as seen in Figure 38, there is only one cluster of high errors. The multivariable characterisation of the agents helps to reduce the number of areas that cluster. While Figure 23 shows that the PV autoregressive model presents errors of estimation in 9 out of the 49 PCs, the **MV-Birmingham model** only presents 3 PCs with a High-High clustering nature, these errors correspond to central PCs, those with a low number of PVs. Moreover, the **MV-LADs** with 96 PCs exhibits the same 3 PCs plus the Birmingham Airport PCs. This suggests that the socioeconomic variables can supplement the model and increase the amount of spatial dependence captured by the model [128]. However, the prevalence of clustering in the central PCs suggests that there may be specific terms of spatial association that may be disregarded by the model [127]. Later when Chapter 3 investigates the spatio-temporal patterns of EV adoption, these insights are brought to the discussion of the spatial validation of the EV model (Section 3.3.2).



(a)



(b)

Figure 36. Spatial distribution of estimation errors - (a) Autoregressive model, (b) MV-Birmingham model.

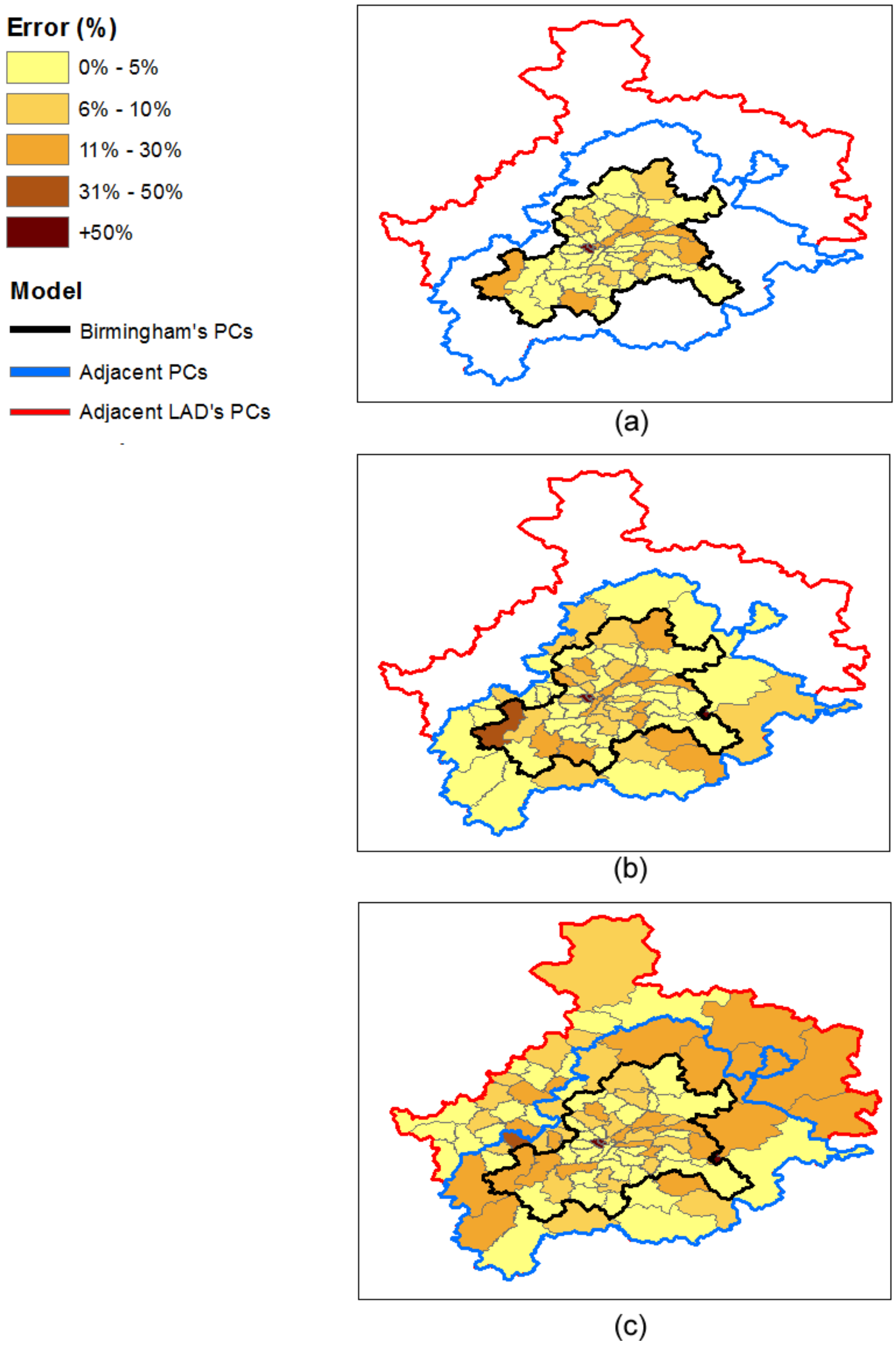
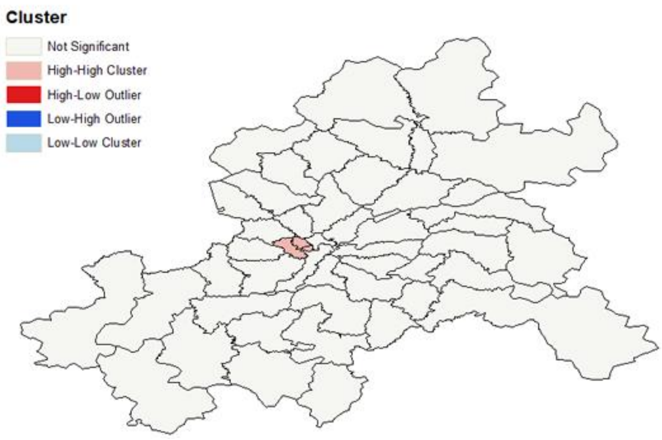


Figure 37. Spatial distribution of the estimation errors - (a) MV-Birmingham; (b) MV-Extended; (c) MV-LADs.



(a)

(b)

Figure 38. Hot-spot analysis of the MV-Birmingham (a) and MV-LADs (b) model's MAPE by the end of the training - Jul 2015.

3.3.3 Predictive accuracy

After the model is been validated temporally and spatially, the training phase proceeds to the forecasting phase. The model forecasts the last 5% of each spatially explicit time-series which are excluded during the training to assess the model's capability to estimate future diffusion of PVs. The predictive accuracy follows the same calculation as described in Section 2.4.3.

Then, to appreciate the error accumulation and the models' predictive accuracy, the estimation errors for the forecasted periods are presented in Figure 39 and Figure 40. Figure 39 shows the forecasting MAPE histogram of the autoregressive and **MV-Birmingham** models, which exhibit an opposite behaviour to the training phase. The errors of estimation diverge and accumulate, yet, the errors of estimation of the multivariable model increases at a significantly lower rate than that for the autoregressive one. This is because the multivariable model reduces the error's magnitude, having a 10% error at the end of the 5^{th} period against 75% error accumulation of the autoregressive model.

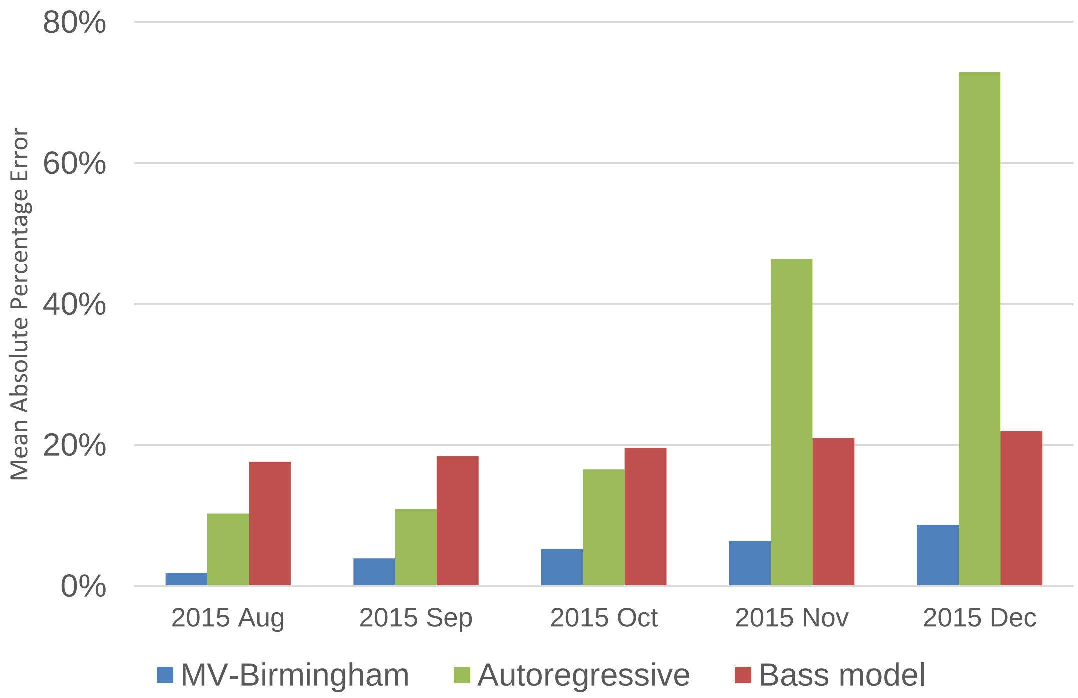


Figure 39. Estimation error for the forecasted periods - autoregressive model vs. multivariable.

Following the same rationale, the predictive accuracy of the extended models is shown in Figure 40. In general, all three multivariate models behave in similar ways, whit only a neglectable difference of 1% by the end of the forecast period. The three models have similar levels of error during the training (see Figure 35). However, the forecasts of the **MV-Birmingham** model accumulates error at a rate of ~2% faster than the other two models, going from 2% to 9%. On the other hand, **MV-Extended** and **MV-LADs** go from 4% to 8%, at rates of ~1-1.5%. There are two possible reasons for this, first, because the MAPE is affected by the number of agents, thus, the average of the estimation errors may be stabilised. One can argue that by extending the area of study, areas that may be considered outliers are also included and this may affect the MAPE (i.e. the airport and adjacent PCs). However, results show that increasing the number of agents reduce the error accumulation instead.

Secondly, because a principle of the ABM is that the behaviour emerges from the interaction between agents, the fact that **MV-Extended** and **MV-LADs** increase the number of agents increases the number of interactions [33,75]. Therefore, the behaviour of the agents appears quicker than the number of agents and interactions increase. Then, one can expect that extending the area of study to a national level could improve the performance of the model. Moreover, this could address the limitation of the model to treat areas as an independent unit of decision-making that disregards what happens in the adjacent areas.

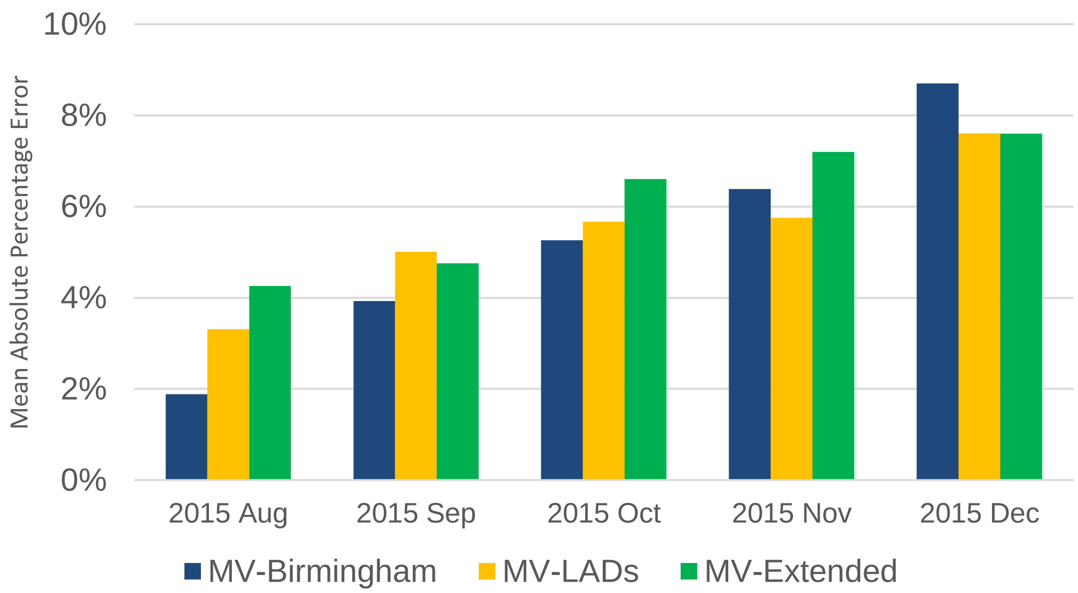


Figure 40. Estimation error for the forecasted periods for the multivariable models.

The individual results at the PC level for the 1st, 3rd and 5th forecasts (autoregressive model and the MV-Extended multivariable model) are shown in Figure 41, respectively. As seen, during the first forecast period most of the areas have an error below 10%. However, by the fifth forecast, the error significantly increases where almost 30% of the areas having more than 10% of errors. Even though the multivariable model duplicates its forecasting error by the fifth estimation, the errors' accumulation over time is reduced in comparison with the autoregressive model; as shown in Figure 41, highlighting the robustness of the model results to changes in scale. The same behaviour is exhibited by the **MV-Extended** and **MV-LADs** models. The temporal results suggest that characterising the evolution of agents' heterogeneity improves the model's performance as they generate a lower MAPE (see Figure 35). This suggests that the preferences of the agents are dynamic and so is their decision-making. It could be argued that the PV autoregressive model disregards the evolution of the preferences and focuses only on the experiences and social dynamics, thus the large error accumulation. Moreover, those ABMs that fix the agents' preferences (adoption thresholds or adoption rules) may be subject to similar limitations.

Even though the data processing to interpolate the temporal resolution may introduce a degree of instrumental error to the model, the multivariable model is more efficient than the autoregressive model, during both training and forecasting. This model implements the temporal dynamic of the agents' heterogeneity by updating the agents' characteristics at each simulation step (month), instead of to the static implementation of agents heterogeneity and fixed specific functional groups [67,104]. Also, by considering the spatially explicit data sets the models do not require a prior classification of customers or their decision-making criteria.

Besides, the results suggest that the spatio-temporal explicit ABM may not be subject to the MAUP, and the methodology can be extended to larger areas. In other words, contrary to the spatial regression, this model is not affected by the change in scale, as the model does not depend on the overall statistics of the population, but generates spatially explicit knowledge.

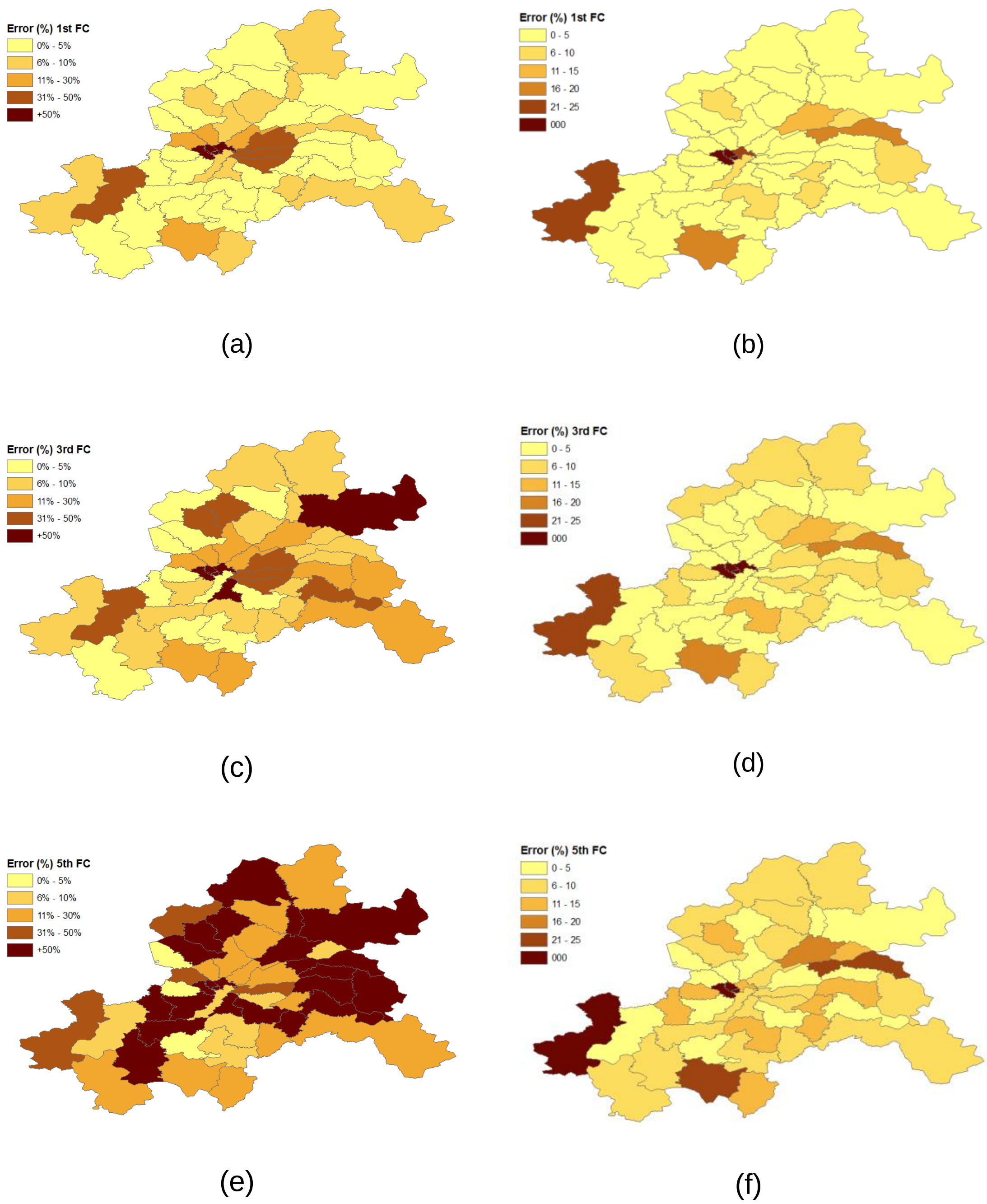


Figure 41. Spatial distribution of the error of estimation. First forecast - (a) autoregressive model; (b) MV-Extended; Third forecast – (c) autoregressive model; (d) MV-Extended; Fifth forecast - (e) autoregressive model (3rd forecast); (f) MV-Extended.

3.4 Reflective summary

This chapter extends the agents characterisation by capturing the evolution of the population heterogeneity. Expecting to improve the model predictive accuracy, the autoregressive model was extended by introducing socioeconomic variables. In line with the literature, results suggest that income is a key decision variable for households to adopt the PV technology. Another important variable is the electricity consumption indicating that households with high energy usage are more likely to adopt PVs, similarly to those concerned about being self-sufficient. Then, the average household size variable captures the negative impact of PV ownership, as bigger families may have less cash flow for a PV installation. Although this multivariable characterisation does not increase the models' accuracy significantly (only 5%), it allows reducing drastically the error from the forecast, from 10% to 2% for the first forecast and from 73% to 9% for the fifth forecast.

Because the heterogeneity of the agents' characteristics and preferences is dynamic, the model captures the evolution of those over time. The autoregressive model has captured the evolution of the agents' preferences using the ANN. Then, the multivariable model captures this by updating the time-series of the socioeconomic variables at each step, providing a more realistic characterisation of the adoption process and agents decision-making process. Additionally, the model addresses the limitation of the autoregressive model which ignores inputs from outside the study area and explores the possibility of the model being susceptible to the MAUP. Thus, the model performance was assessed using three different scales, results show that due to the adaptive capabilities of the ANN, the model is flexible to handle different study area sizes.

Therefore, attending to the aim of this thesis, the following chapters focus on identifying further modifications required to analyse other technologies adoption process (chapter 4), and on finding ways in which the model can inform one's technology adoption process with other's decision-making (chapter 5).

4 A spatio-temporally explicit ABM and ANN integrated model: the EV adoption case

4.1 Introduction

This chapter analyses the spatio-temporal patterns of EV adoption, by implementing the methodology used for the PV autoregressive model. The chapter first explores the historical EV data, highlighting the differences with the PV data and discusses the expected differences in the results. Then, the model estimates the adoption rates of EV adoption and discusses whether further modifications to the (PV) autoregressive model are required. Previously, Chapter 2 designs, develops and implements a model that characterise the spatio-temporal regularities and social dynamics into a novel ABM for the adoption of PVs. Subsequently, Chapter 3 extends on the model so as that it recognises the heterogeneity of the population and analyses the effects of the changes in the geographical scale. Besides, the model is validated spatially and temporally, and the results show short-term applicability of the results (1-5 months). Because the research aims to study the regularities between the EV and PV adoption, this chapter also works as a first step toward a combined model to characterise the influence of owning PV on the decision-making towards adopting an EV, or the other way around.

The literature review, in Section 1.1, presents both SR and ABMs studies that analyses the market diffusion of EVs. From the SR perspective, Morton et al. [29] provide the only EV study, to our acknowledgement as of 04 Dec 2018. Besides this study contributes to understanding the drivers of EV adoption in the UK, the study faces a common limitation that is the low data availability, as the model presents inconsistency in the timestamps. The vehicle registration for 2016 is modelled with income data for 2015 and other socioeconomic data for 2011. Besides the results are subject to bias, the model yields a modest accuracy of 60%, suggesting that some of the spatial dependence still need to be explained. On the other hand, the ABM studies comprise a wider number of applications [26,27,32,33,49], yet, these studies are purely explorative. The authors present the EV market diffusion under different scenarios, such as vehicle price [26,27,33], rebates or subsidies [33], fuel cost [27,32,33,49], discount rate [49]. The agents' decision-making includes both the financial [26,27,32,33,49]and social utilities [26,32,33], using multivariable characterisation. Additionally, the Bass model has been used, though, with no information on its fitness, Linder and Wirges [41] construct s-curves for the EV uptake using different levels of imitation rate and the initial number of adopters.

This chapter seeks first to analyse the spatio-temporal patterns of EV adoption, using the autoregressive ABM and ANN integrating model. The chapter analyses the EV adoption using the same modelling framework employed to analyse the PV data. The following section starts the analysis looking at the EV data, providing an overlook of the spatial and temporal dependence, and comparing with those of the PV. Moreover, this chapter discusses whether it is suitable to combine both autoregressive PV and EV models into a combined model.

4.2 Methods and materials

This section first analyses the EV spatial and temporal regularities, and follows the design process developed in Chapter 2[19] made for the PV adoption process. The model uses the PC and monthly resolutions and includes the 96 PCs of the **MV-LADs** model. The ANN design and training follows the same approach that has been followed in the previous chapters.

4.2.1 Data

The model uses data from the *Stock Vehicle Database* held by the *Department of Transport,* accounting for vehicle registration as of July 2018. These datasets include the total number of registered vehicles by propulsion type[20] and the month of first registration at Lower layer super out area (LSOA) level. This data is

[19] Figure 13, in Section 2.2, displays the conceptual model for the autoregressive PV model.
[20] i.e. Diesel, Electric, Hybrid-electric, Petrol, etc.

aggregated to PC level using ONS reference lookup tables [21]. Although the spatio-temporal resolution is PC and monthly basis, the EV model increases both the size of the time-series and the number of PC in the analysis. Therefore, the ABM and ANN model characterises the EV adoption process for 96 PCs, and including data from January 2011 to September 2018, resulting in 93 observations (months). Over this period the number of EVs registered increased from $N = 4{,}960$ to 15,066[22], where the two PCs adjacent to the airport accounting for more than a third of these new registrations[23]. Figure 42 shows the range of values among the PCs, excluding these two PCs. The average number of EVs by PC is 99 with most of the values being between 61 and 134, and outliners of more than double the average of EVs.

[21] Available at the Open Geography portal from the Office for National Statistics, source: http://geoportal.statistics.gov.uk/datasets/local-authority-district-to-combined-authority-december-2015-lookup-in-england

[22] Following the same criteria that with the PVs, those dates with low number of EVs are excluded, limiting the dataset to post Jan 2011.

[23] The B92 and CV7 accounts for 5,246 and 633 EVs (for a total of 5879)

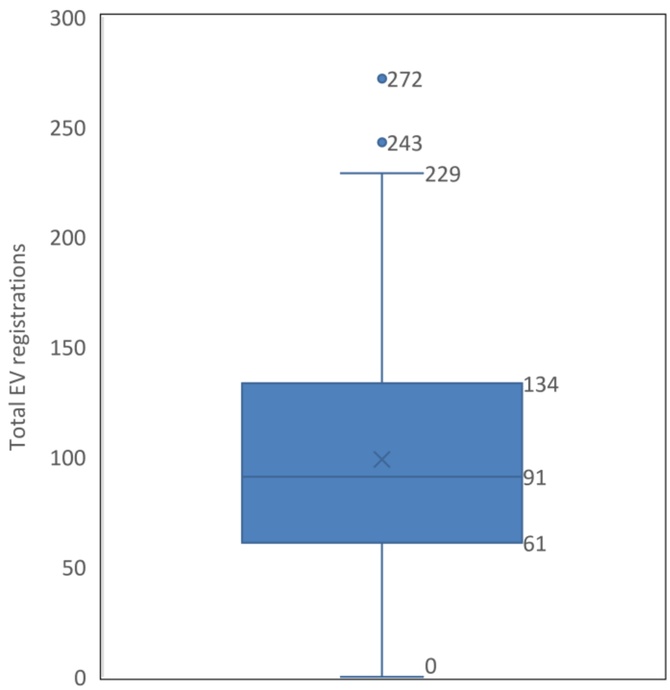


Figure 42. Boxplot of the total EV registration in 96 PCs in the West Midlands.

Figure 43 shows the spatial distribution of the EV registrations. Like the PV spatial distribution, the areas in the centre of Birmingham exhibits a low number of EVs, arguably because of the low number of residential buildings, whilst the areas in the South East have a high concentration of EVs. These areas are some of the PCs adjacent to the Birmingham Airport where EVs in these areas might be used as a taxi service by the passengers. Both of those PCs have almost doubled the quantity of registered EVs from January 2011 to September 2018. These results are shown in Table 9, which suggest that the EVs exhibit spatial regularities.

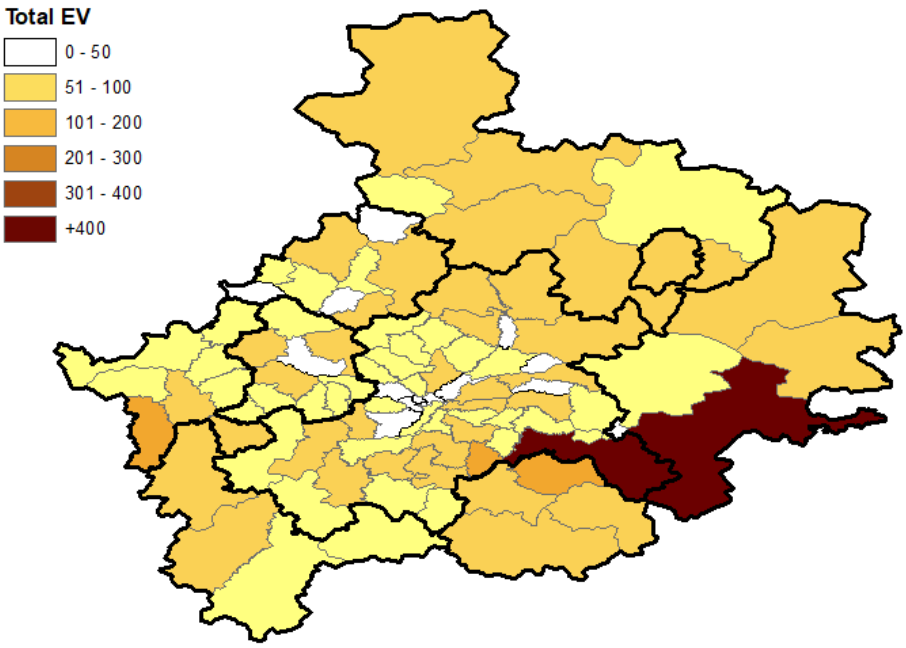


Figure 43. Spatial distribution of the EV registrations by September 2018.

Figure 44 displays clusters of EVs where low values are concentrated in the central PCs, whilst the high values are located in the South and South-West. Although, the EVs distribution present spatial regularities, the degree of clustering is lower than for the PV case; being close to 0 for EVs and ~0.4 for PVs. The difference between the Moran's index of the EVs and PVs can be explained with the difference in the adoption stage of both technologies. van der Kam et al. [28] use Rogers categories[24] to understand the stages of EVs and PVs adoption, pointing out that the PVs are in the early adopters phase, whilst the EVs are in the innovators phase.

[24] Rogers [37] classification includes innovators, early adopters, early majority, late majority, and laggards. Each of these categories has differenct characteristics and are distributed normally, being innovators those accounting for 2.5% of population and ealy adopters 13.5% of the population.

Table 9. The Moran's I index value, z-score and p-value of the EV registration data

Statistics	Moran's Index	z-cores	p-value
Value	0.0193	3.9628	0.00007*
Given the z-core of 3.9628, there is less than 1% likelihood that this clustered pattern could be the result of random chance.			

The statistical significance is marked with asterisks.

**p<0.05*

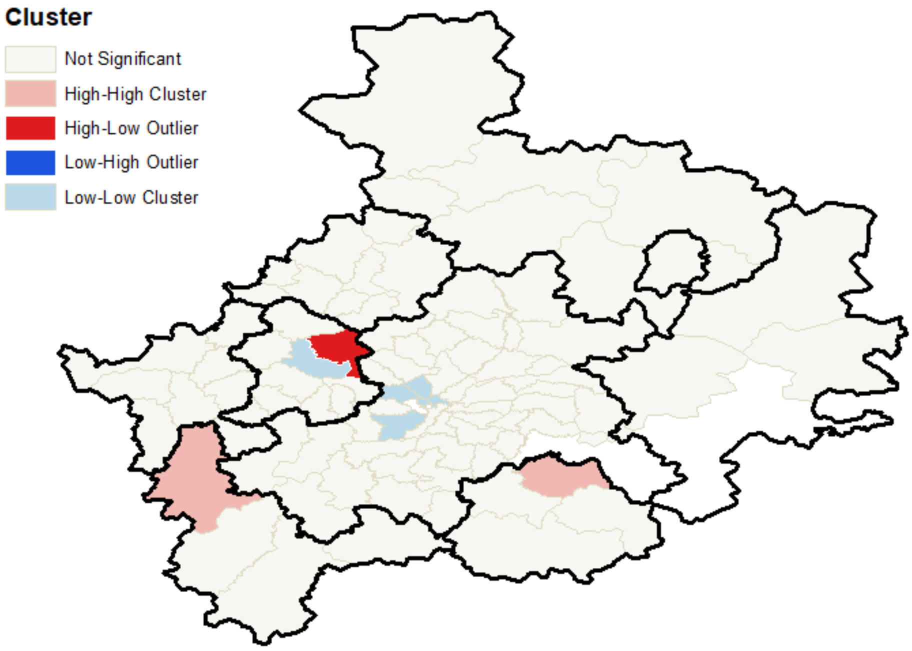


Figure 44. Hot spot analysis of the EV registration at PC level by Sept 2018.

4.2.2 Spatio-temporal resolution

Given the wide range of values in the EV data, the EV time-series are normalised into values between [0-1] following the insights from Chapter 2. Then, the transformed data is comparable across areas. This transformation was done considering the overall population's range of values using the following formula:

$$\widehat{EV}_{i,t} = \frac{EV_{i,t} - MinEV_i}{MaxEV_i - MinEV_i} \quad \textbf{(4-1)}$$

Where

$\widehat{EV}_{i,t}$ is the *t-th* EV estimation in the *i-th* area

Min_i is the minimum number of EVs in the *i-th* area over the study period

Max_i is the maximum number of EVs in the *i-th* area over the study period

Figure 45 shows the share of the household of EV adopters, displaying a wide number of behaviours among the PCs. Probably the most important feature is the negative trends of some of the PCs, this could be explained for the possibility of registering any vehicle *off the road*[25], or the fact that cars can be traded between users of different locations. This is contrary to the PV data, which presents a constant positive increase month by month, whilst the EVs registration may fluctuate over time. It is expected that the ANNs capability to estimate any type of function helps to adapt to these differences [105,116].

[25] This is done when an user wants to stop taxing and insuring it. S**ource**: https://www.gov.uk/make-a-sorn

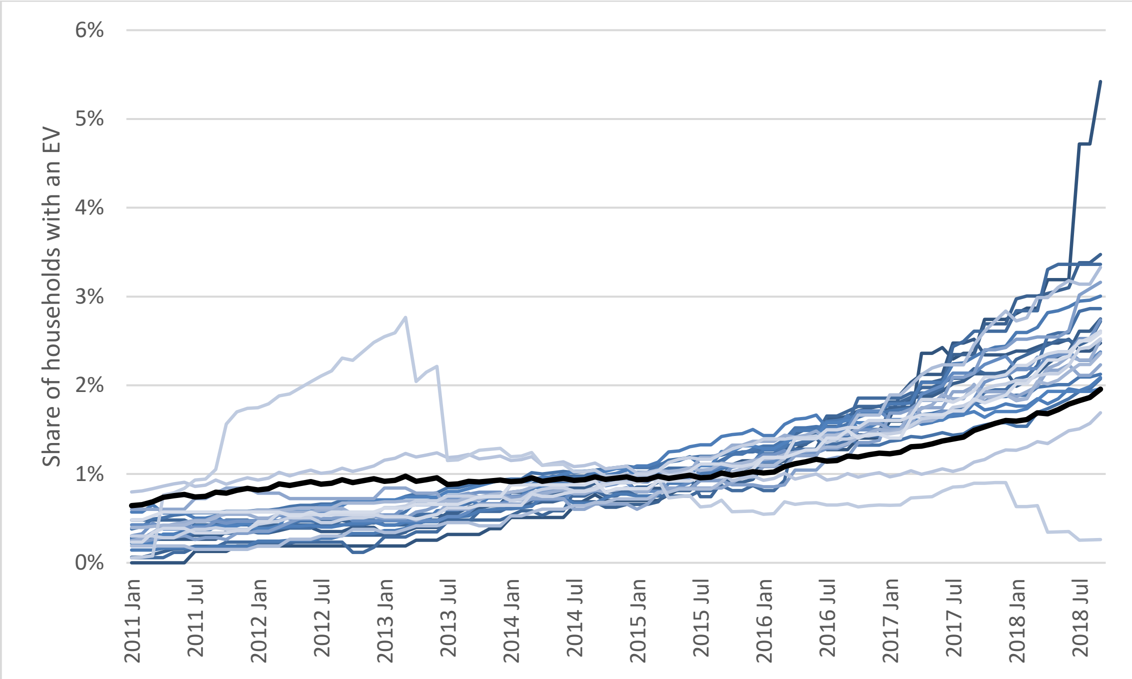


Figure 45. Monthly cumulative adoption rates EV registration for the 96 PCs.

The analysis uses the same study size as the **MV-LADs** described in Section 3.1.2, Figure 46 shows the 96 PCs characterisation in Anylogic v7.3.2 software.

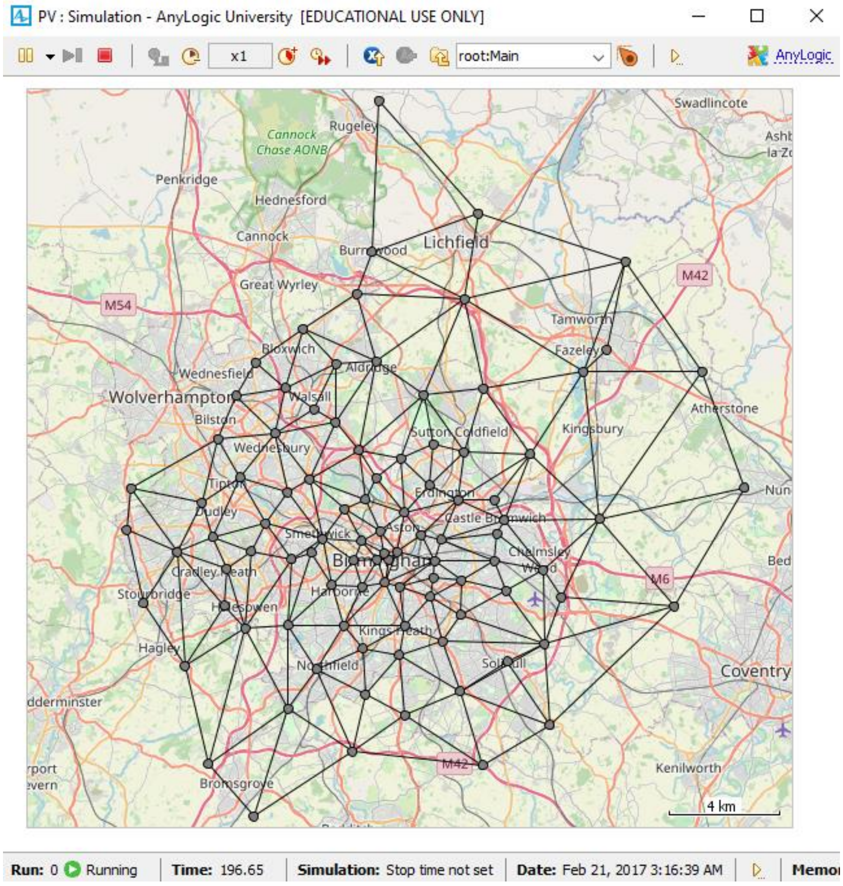


Figure 46. Spatial representation of the EV autoregressive model in AnyLogic software.

4.2.3 Spatio-temporal dependency

The Moran's I statistic is shown in Table 9 and the clustering map in Figure 44 demonstrates that EV registrations present spatial regularities. These features are similar to the PV case, which also exhibits clusters of low and high values. However, the locations of these are different, see Figure 12. On the other hand, ACF and PACF, from the Box-Jenkins methodology, are used to analyse the temporal lag of the EV adoption. Both functions are calculated using Python, considering monthly resolution and up to 12 lags. Figure 47 shows the ACF functions, where each of the boxplots shows the mean and median of the temporal behaviour of the population at each period of time. Similar to the PV case, the significant lags present a slow decrease over time. This suggests that the model has an autoregressive nature. However, the EV's ACF presents negative values, this means that some observations have a negative influence on the total number of EVs, maybe reflective the effect of *off the road* registration. Figure 48 presents the PACFs, which presents significant lags in the 1st, 4th, 7th and 10th lags. However, this semi-quarterly regularity is not uniform among the population, as the share of PCs decreases up to less than 50% for the positive lag and less than 20% for the negative values. On the other hand, the entire population present a significant lag of first-order (*t*-1), therefore the model's autoregressive element will be kept the same as for the PV model. Later, Section 4.3 discusses the implications of disregarding the other temporal lags, as this may introduce bias into the estimations. Thus, either the temporal validation may show disturbances in the estimation errors with a quarterly periodicity, or the ANNs to adapt and learn from the data sets and minimise the errors.

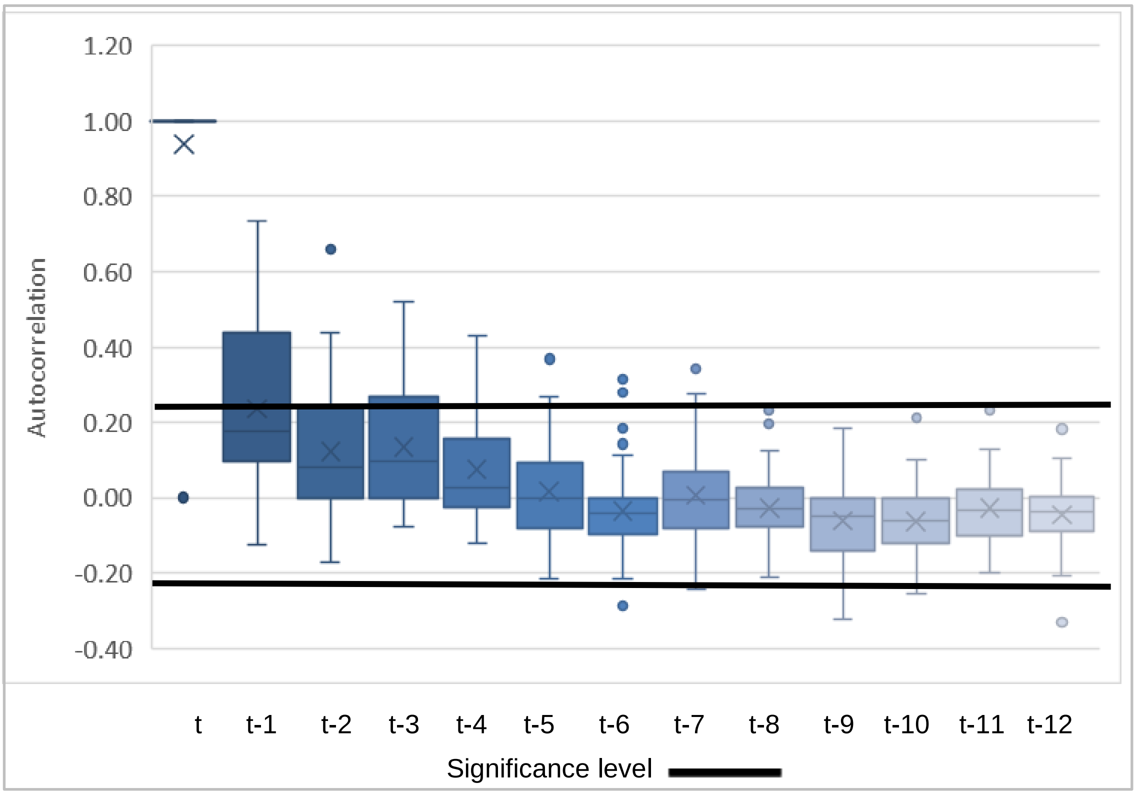


Figure 47. ACF of the EV registration data at the PC level.

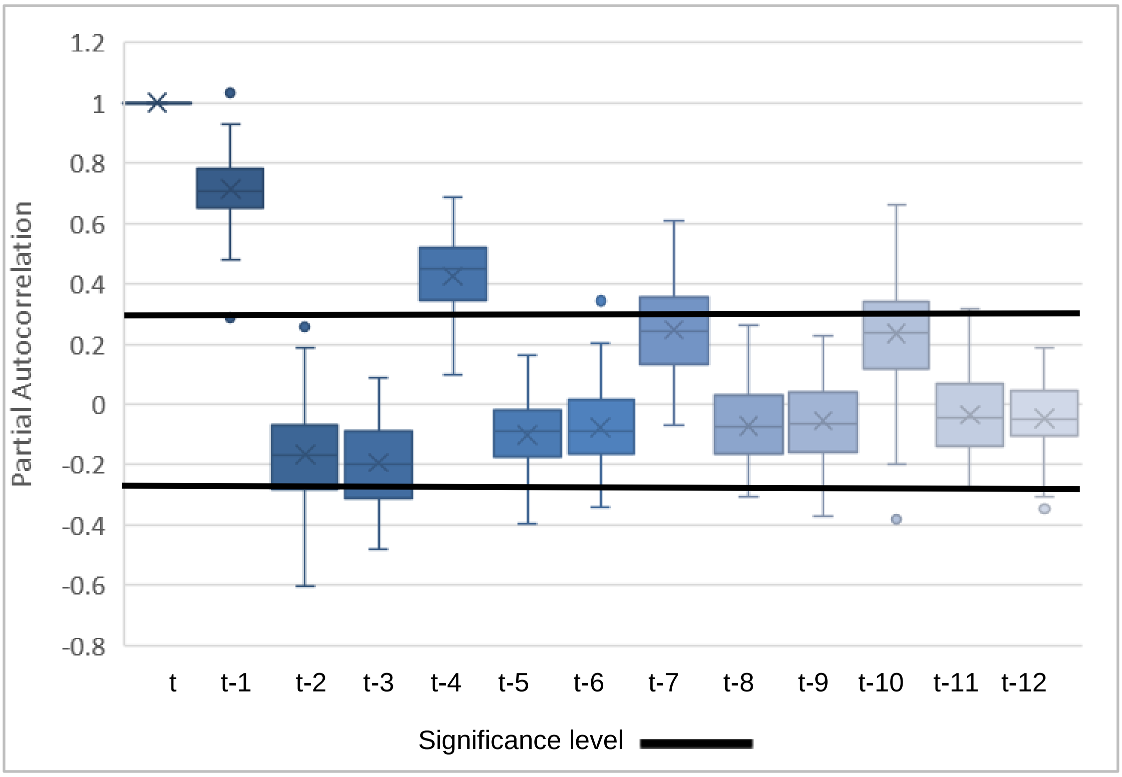


Figure 48. PACF of the EV registration data at the PC level.

4.2.4 Decision-making process

The ANN that characterises the decision-making process implements back-propagation training. The conceptual design is identical to the PV autoregressive model (see Section 2.2.5), as shown next in Figure 49.

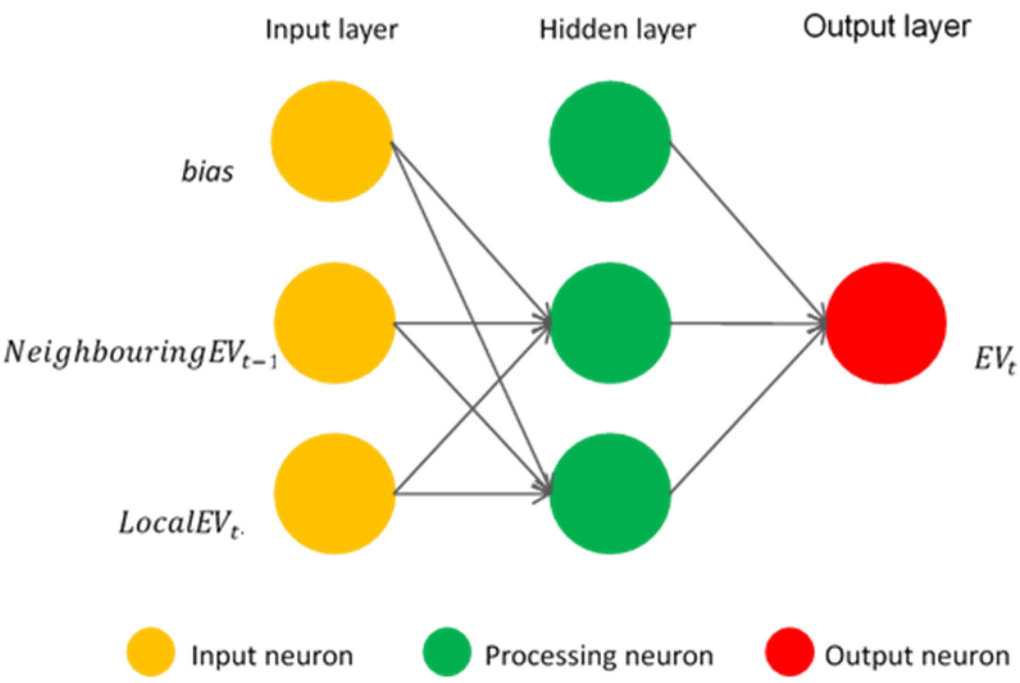


Figure 49. Decision-making process defined by the ANN, EV adoption process.

4.2.5 Model validation

The model's fitness is assessed using the MAPE, accounting for the entire population using the equations (4-2) and (4-3):

$$MAPE_j = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{EV_t - \widehat{EV}_t}{EV_t}\right| \quad \textbf{(4-2)}$$

Where

n is the time series size

EV_t is the current number of EVs in the month *t*

$\widehat{EV}_t$ is the estimation of the number of EVs in the month *t* *i* is the specific month

j is the specific area

$$Population\ MAPE = \frac{1}{m}\sum_{k=1}^{m} MAPE_k \qquad \textbf{(4-3)}$$

Where

m is the population size

k is the specific area

4.3 Results

4.3.1 Temporal validation

Figure 50 shows the estimations for the ANN, Bass model and the actual data for the training period. As seen, while the ANN overestimates the total number of EVs, the Bass model underestimates. The ANNs can reflect the changes in the data trends, those around April 2013 and January 2016, while the Bass model can just assume a positive trend (see Appendix 6 for the entire Bass estimation). One thing to consider is the limitation of the Bass model to describe negative tendencies, as its structure is an exponential function. Despite the model seems to fit the actual EV data, if the EV growth would continue with the significant increase seen from 2017, the model then would significantly underestimate the number of EVs. Moreover, if the growth of EVs would change its trend, a similar case to the PV would be seen in Figure 18, where the model cannot adapt to the new data behaviour. On the other hand, the ANN estimations follow the trends in actual data over time and the estimation errors stabilise, yet, the results present a negative tendency at the end of the forecasted period. Maybe because the model can recognise the negative effect of the *off-road* registration, the error accumulation results in the underestimation of EVs. Additional to the error accumulation, the EV growth rates increase significantly during the forecast period, having a ~10%, ~20%, ~30% annual growth for 2016, 2017 and 2018, whilst the annual rate before that was around 2%. This could be the results of the *Low-emission vehicles grant*, which has had a positive impact on the number of EVs since its establishment in 2011. Besides, the rental car industry has recognised the potential role of this industry to improve the sustainability of the UK transport sector [137,138]. For instance, *Enterprise Rent-A-Car*[26] has almost duplicated its fleets of EVs in the UK and Ireland during 2017 and 2018, which shows how the industry has identified a business opportunity for the rental car [139,140]. Additionally, given that the decision-making is also driven by experience and perception, common users of rental cars may have a higher preference to buy an EV in the future [140].

Figure 51 shows the level of errors for the ANN and Bass model, although the ANN estimations produce greater errors at the beginning of the training. However, the error levels rapidly decrease and converge levels lower than the Bass model, without presenting any peak or disturbance in the estimation errors. The data analysis performed in section 4.2.3 noted for the 4^{th}, 7^{th} and 10^{th} lags were also significant, yet, these were not included in the model. These were excluded because that behaviour is not uniform across the agents, accounting for less than 90%, 50% and 50% respectively. Nevertheless, even when this behaviour in the data was overlooked, the capabilities of the ANN to approach to any nonlinear functions yield more than 95% of accuracy.

[26] Enterprise is a mobility solutions company. **Source**:https://www.enterprise.co.uk/en/home.html

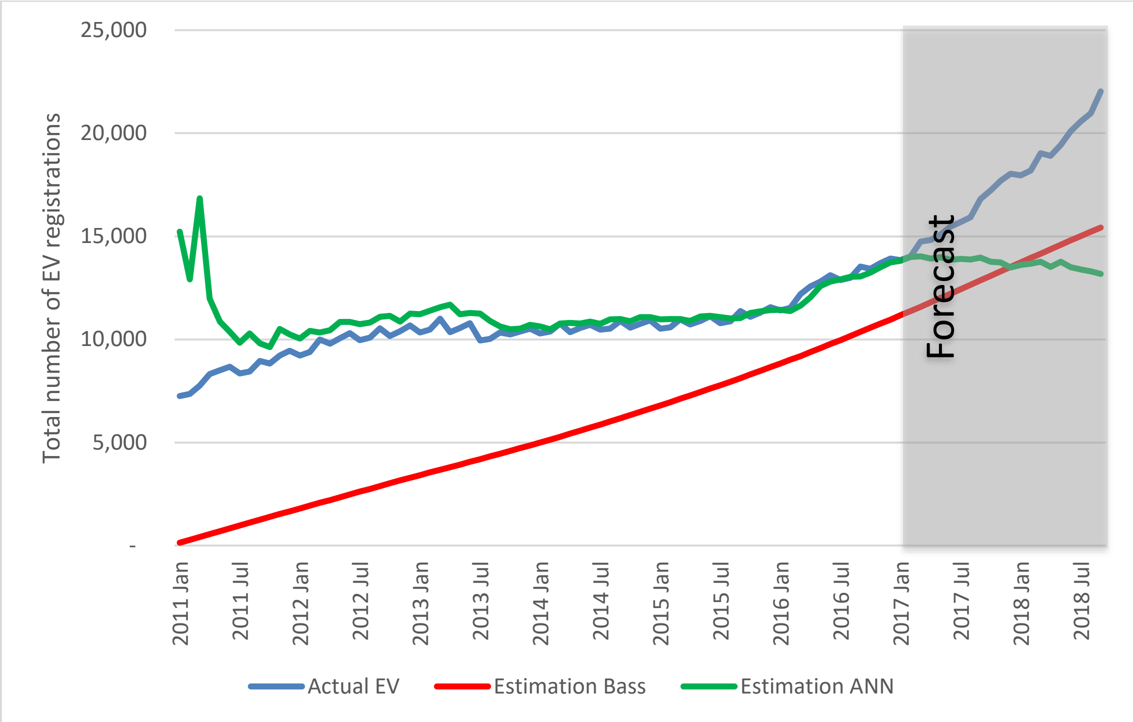


Figure 50. Cumulative adoption rates of EVs estimated by ANN and Bass model vs. actual data.

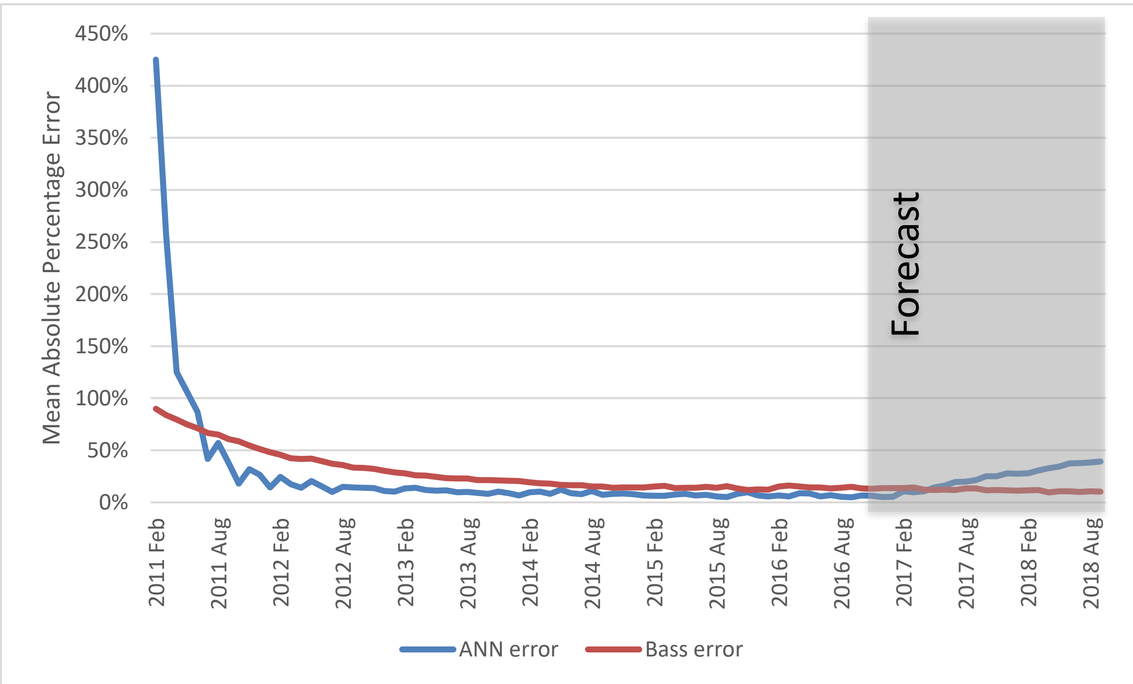
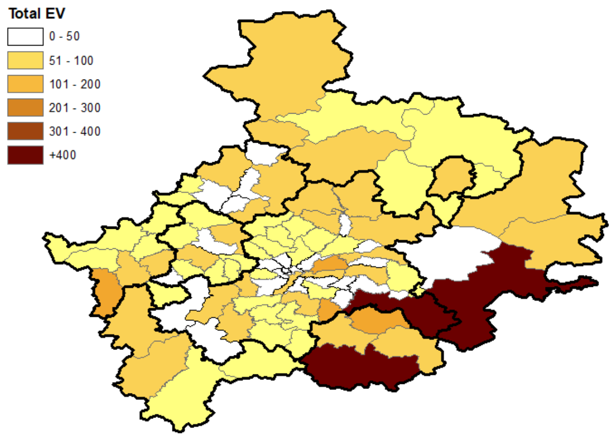


Figure 51. Estimation error for the ANN and Bass model for the EV data.

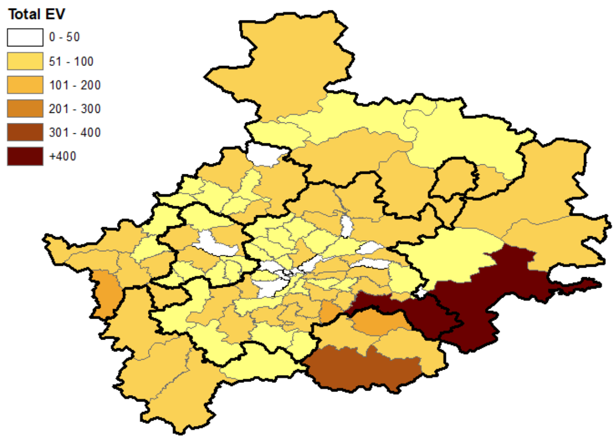
4.3.2 Spatial validation

Figure 52a and Figure 52b present the spatial distribution of the ANN and Bass estimations at the end of the training. Both models exhibit similar behaviour with high forecasts of EVs in South and South-West, except that the ANN present more extreme values; both low and high. Although, these clusters are not the same as the ones in the actual data (see Figure 44 for comparison), the errors of estimation only present spatial regularities around Birmingham's airport PC, suggesting that there is a spatial effect originating from the airport. Figure 53a and Figure 53b show that both models present these spatial regularities. Considering that the degree of clustering of the EV data is close to zero (see Section 4.2.2) and that the Bass model disregards the spatial dependence, this confirm that the distribution of the EVs has yet not developed spatial regularities. Moreover, this suggests that the clusters shown near the airport are caused by other factors other than spatial regularities. Therefore, the estimations of the Bass model and the ABM produce similar clustering patterns.

The spatial distribution of the errors is analysed to further understand how well the model captures the spatial dependence of the EV diffusion. As seen in Figure 54a and Figure 54b the error from the ANN estimations vs. those forms the Bass model are significantly different. As seen, the Bass model's errors are lower than those from the ANN. Yet, most of the ANN estimations present errors below 6% or higher than 30%, whilst the Bass estimations have a more uniform error distribution with errors between 6% and 10%; see Figure 54b. Then, Figure 55a and Figure 55b are used to verify that the errors of estimation fo the ANN model tend to present lower values in the central PCs, and higher errors in the Northwest. This could mean that as the EV lacks strong spatial regularities, the model may overestimating this effect, by creating random patterns of error estimation. Instead, the Bass model presents lower clustering and only two outliers in the central PCs, this follows the fact that the Bass model does not consider the spatial dependence.

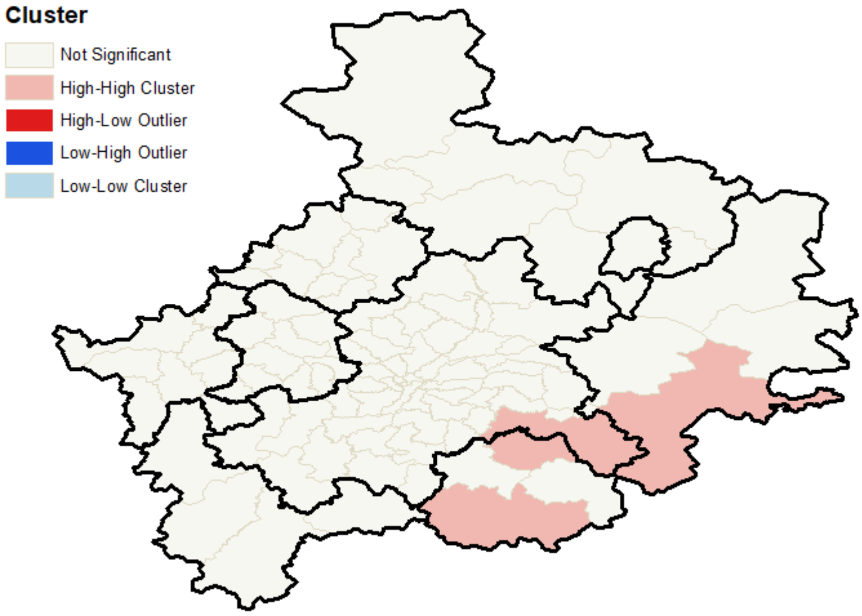


(a)

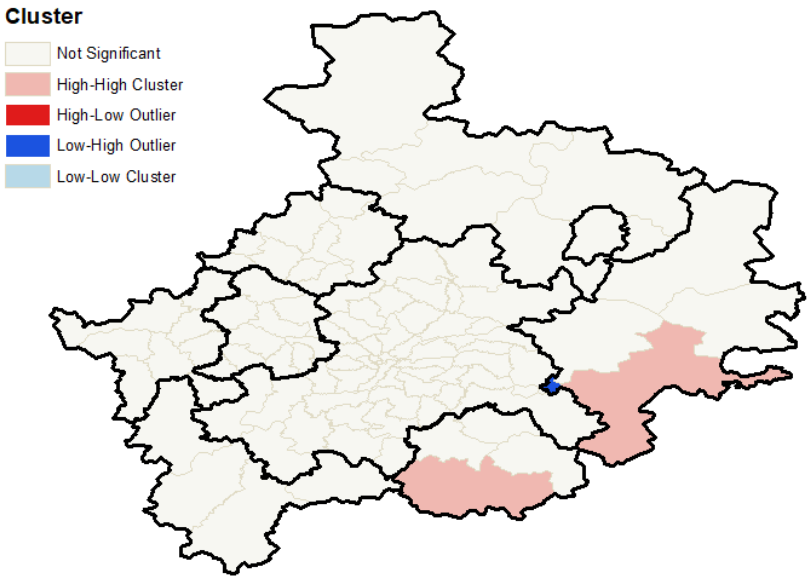


(b)

Figure 52. Spatial distribution of the EV estimations of the (a) ANN and (b) Bass model by the end of the training period - December 2016.

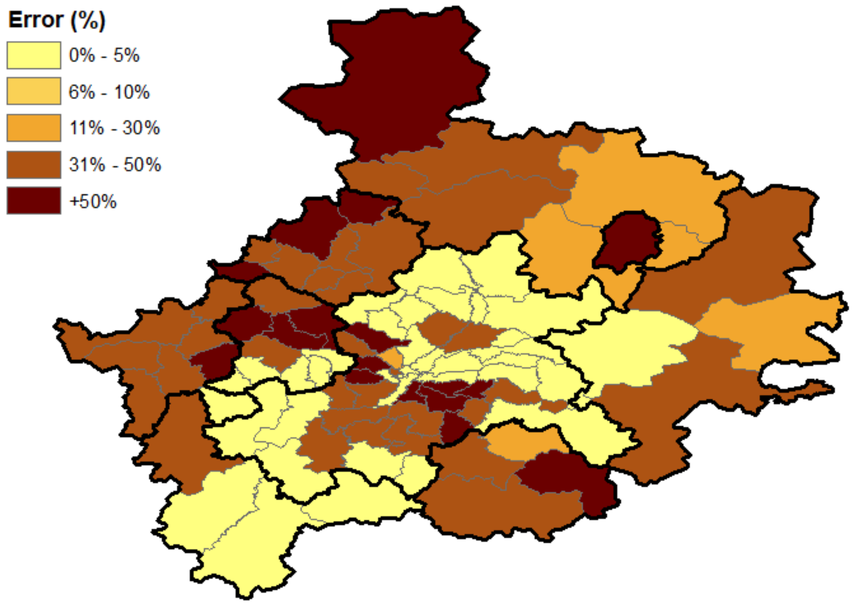


(a)

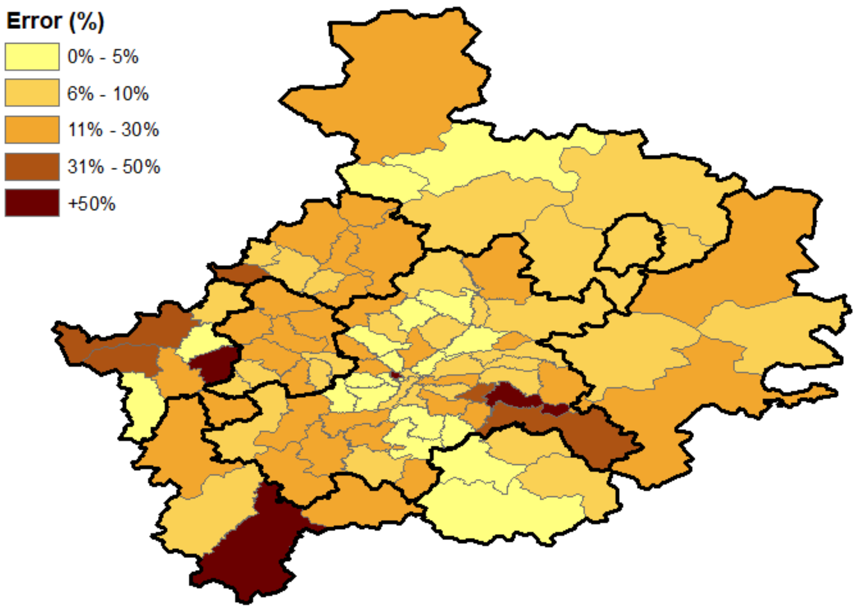


(b)

Figure 53. Hot spot analysis of the EVs estimation for the (a) ANN and (b) Bass model by the end of the training period - December 2016.

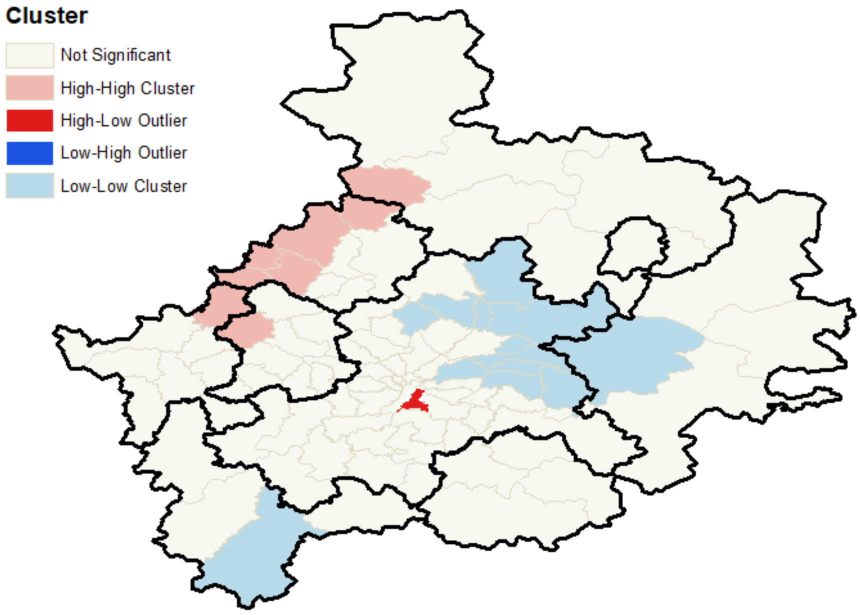


(a)

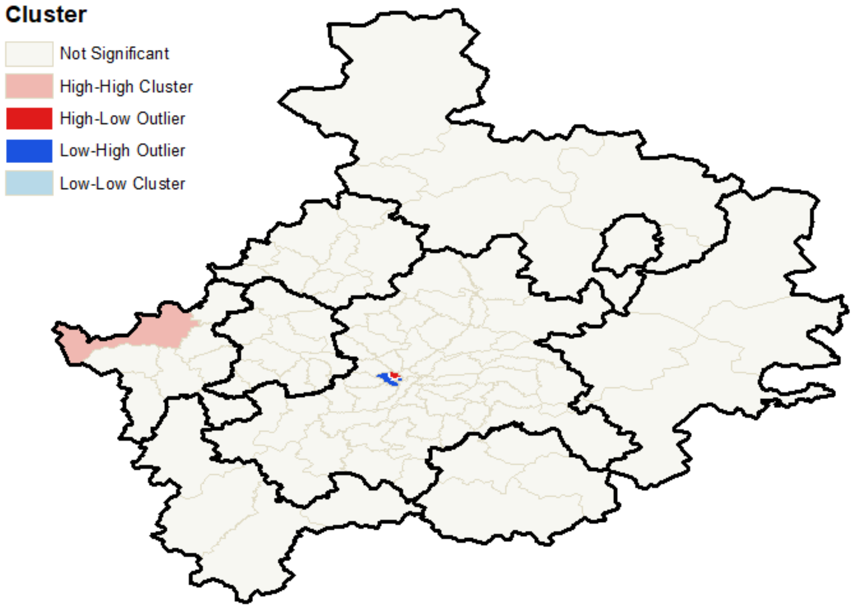


(b)

Figure 54. Spatial distribution of the errors of estimation for the (a) ANN and (b) Bass model by the end of the training period – December 2016.



(a)



(b)

Figure 55. Hot spot analysis of the error of estimation for the (a) ANN and (b) Bass models error of estimation by December 2016

4.3.3 Predictive accuracy

Figure 56 shows the MAPE of the 18 forecasted months, whilst Figure 51 displays that during the training period the error of estimations converges to levels around the ~15%. However, the errors start accumulating during the forecasting period. Both the EV and PV estimations presents this accumulation of errors, however, the magnitude of those are different. In the PV case the errors of estimation are lower than 20% by the 3rd forecast, yet, it jumps to +40% for the 4th and 5th estimations. In the EV case, the MAPE stays below 20% up to the 8th forecast, and lower than 40% by the end of the forecast period (September 2018). Although the ANN estimations can replicate the behaviour of the EV data over time, the accumulation of the underestimation during the forecast period results in higher estimation errors than those of the Bass model. Nevertheless, the ANN model can produce lower errors of estimation than the Bass model for the first three periods, after that the Bass model is better at estimating the EV adoption rates. Because the exponential equation used in the Bass model, their estimations cannot react to the negative trends of the data, however, it can fit the last months where the behaviour is entirely positive [41].

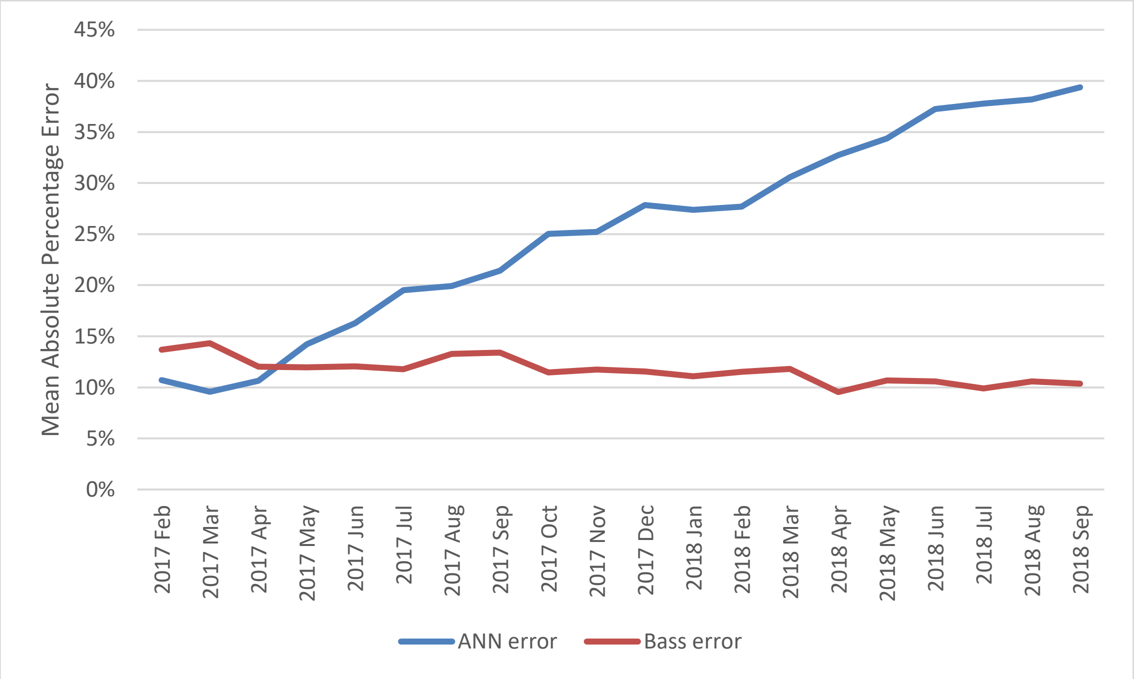


Figure 56. Estimation error for the forecast period ANN vs Bass model.

4.4 Reflective summary

This chapter analyses the EV spatio-temporal patterns of adoption, using an ABM and ANN integrated model. The results suggest that the model can characterise these patterns, moreover, the model shows benefits over the Bass model. Although the model has some limitations to capture the spatial dependence of the adoption process, the adaptive capabilities of the ANN can reflect the changes in the adoption trends; which is a limitation for the Bass model. Moreover, because these spatio-temporal regularities are only shown at high data resolution, the results reiterate the importance of generating local estimations at the high spatio-temporal resolution, as shown in Section 4.2.1, rather than regional or country estimations. Morton et al. [29] also generates a map of the EV clustering for the UK at LAD level, showing Birmingham LAD with no significant spatial clustering, Sandwell and Walsall with a Low-Low clustering and Lichfield with a High-High pattern; these three LADs are the areas in north Birmingham. Conversely, this analysis shows that only one of those areas has a High-Low outlier. Therefore, using results at LAD level may disregard local specific behaviours, thus, the model has limitations to inform planning decisions.

The model has the potential to generate spatio-temporally explicit estimations of EVs, as the model can produce estimations of up to 8 months ahead with 80% of accuracy. However, the potential to inform policymaking or network planning is limited, as the model is only able to produce up to three months ahead with less than 20% of error. Despite the differences of the EV and PV data, both models are limited to short-term applicability, which may suggest that an autoregressive ABM and ANN model may not be able to inform the medium- or long-term management of the distribution network. Nevertheless, the results suggest that the model is flexible and transferable to other technologies; even if those have a different spatial or temporal structure, For instance, the temporal dependency tests for the EV time-series show that there are other significant lags. Yet, it is shown that the model can estimate the EV uptake by using the same one-lag autoregressive model as the PV one. Also, the EV and PV present different spatial regularities, and they may be affected by different spatially explicit economic activities. The model can still capture different degrees of spatial dependence without making it explicit to the model or fixing that effect into the model decision-making. One avenue for future research could be to implement the lags that were disregarded, by increasing the number of input neurons. Although the inclusion of longer lags, i.e. the 10^{th} lag, means losing ten observations of the time-series, this trade-off is negligible as the available data increases.

Second, the spatial validation shows that the estimation errors present a higher number of areas with spatial regularities (see Figure 23), suggesting that the model has limitations to capture the spatial dependence of the EVs. Because, these spatial regularities can be explained by specific conditions of those areas, rather than a clustering coming from the model, the results suggest that the framework used for the autoregressive PV model is also able to characterise the EV case without any further modification. However, given the limitations for the EV model to capture the spatial regularities of the adoption patterns, further work is required. Chapter 3 has already shown the effect of including socioeconomic variables into the agents' characterisation, demonstrating that the model performance can be improved by characterising the temporal dynamic of the population heterogeneity. Therefore, the following chapter rather focuses on investigating the effect of considering the effect of a second adoption process on the decision-making. Then, the model combines the insights produced by both autoregressive models, to investigate whether the exchange of information can improve the model's performance.

5 Integrating knowledge exchange between EV and PV adoption into a spatio-temporally explicit ABM

Previous chapters have contributed to test the hypothesis of this thesis. Chapter 2 shows that it is possible to integrate the ABM and ANN while integrating insights from the SR. Chapter 3 demonstrates that a multivariable characterisation of the agents improves the model performance and helps to reduce the error accumulation of the forecast. Finally, Chapter 4 analyses the adoption patterns of EVs and assesses the transferability of the model. Results have shown that by generating experience-based knowledge, the ANNs can adapt to the different spatio-temporal patterns of EVs and PVs adoption (autoregressive or multivariable model). However, it is still pending to test whether this knowledge can be exchanged between domains. Authors such as McCoy and Lyons [26] assume that owners of PVs have a higher preference for EVs than other agents, defining the agents' social utility based on whether the individuals have adopted any energy-saving technologies previously. Because the households can offset its energy consumption from the grid with the PV produced energy, the PV can be seen as energy-saving technology. Cohen and Kollmann [36] suggest that households who already own a PV have the intention to purchase an EV during the following five years. Davidson et al.'s diffusion model note the significant and positive impact of the number of EVs registered in a location to explain the number of PV installations [14].

Building upon these insights into the model's design, this chapter investigates how to capture the higher preference to the adoption of another associated technology, in this case, the associated technologies are those adopted by consumers with high environmental concern [34]. Therefore, this chapter aims to integrate both autoregressive models to capture the knowledge exchange between EVs and PVs and assess to what extent one technology can inform the decision-making for another technology. The analysis assumes that the knowledge generated for one technology, during the training, can inform the decision-making for another technology. The analysis builds upon the insights of both autoregressive models and combines the EV and PV time-series. The model captures the information flow shared between EVs and PVs, but also the unique information for each decision-making. The former is characterised by the shared connections in the artificial neural network, whilst the latter is captured by the unique synaptic weights that calculate the outcome of the neural network.

5.1.1 EV and PV diffusion in Birmingham

Figure 57 shows the number of EVs registered and PVs installed annually from 2011 to 2018 for the 96 PCs in Birmingham, and each time-series comprises 60 monthly observations. The diffusion of both technologies exhibits different patterns. The EV registrations almost increased at a steady rate from 2011 until 2016, whereas PVs had a sudden increase between 2012 and 2015. Before 2012, there were almost no PV installations. Then, the number of PVs had a significant increase and almost double the number between 2014 and 2016, whilst the number of EVs registered kept stable until 2016. At this point, both technologies changed trends almost like exchanging trends. Although both the EV and PV growth rates show a positive trend until July 2018, the number of PVs increased steadily, whilst the EVs had a more rapid increase.

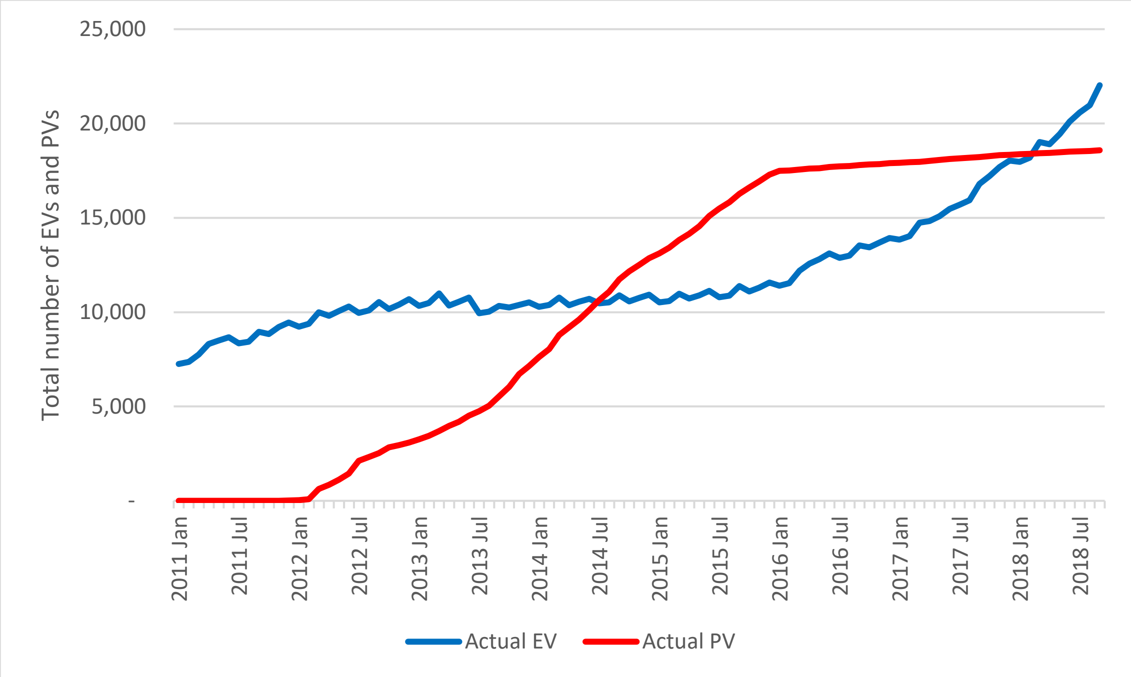
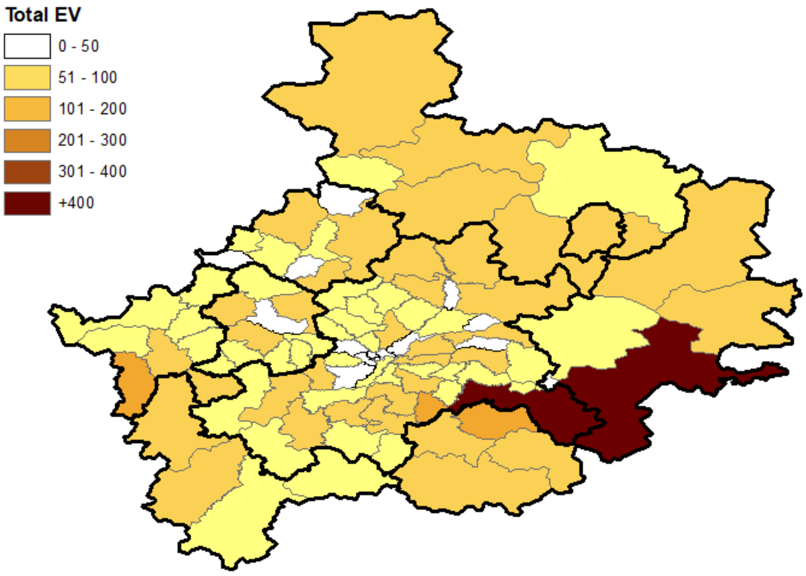


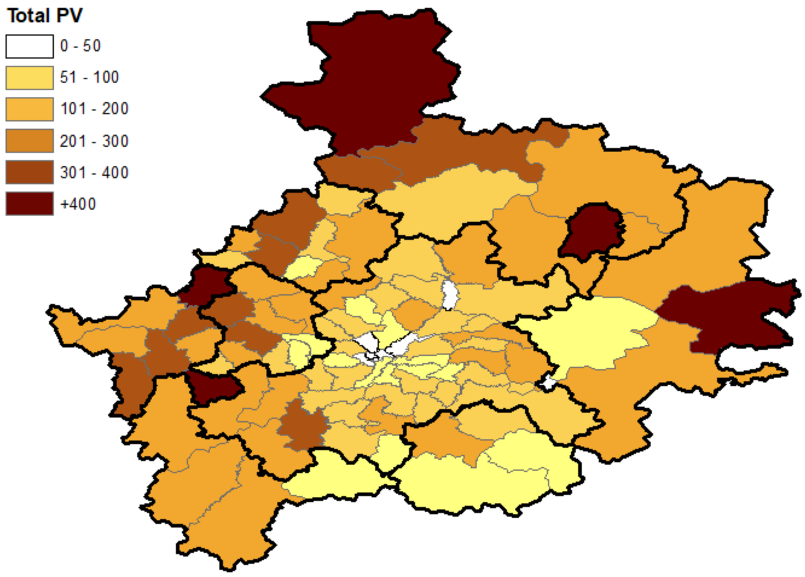
Figure 57. The cumulative adoption rate of PV installations and EV registrations in 96 PCs in the West Midlands, UK.

On the one hand, the change in the PV trend can be explained with the drop in the FiT rates [7,8,124]. This scheme has been greatly successful especially for the adoption of PVs, which comprised mostly domestic installations. Then, the government announced the review of the FiT rates, aiming to match real-world cost reductions. The announcement of the reduction in the FiT scheme caused an urge from households to install PVs before the deadline on the 1st of April 2012 [124]. Finally, further reductions to these rates in 2016 resulted in a dramatic decrease in the number of new PV installations [141]. On the other hand, the increase in the EV numbers can be associated with the *Low-emission vehicles grant*, which has successfully increased the number of EVs since its establishment in 2011. Likewise the FiT scheme, this grant was revised and reduced at the end of 2012, at which point one can see a deceleration in the EV adoption [142,143]. Despite further cuts to this grant in recent years, the number of EVs has regained a significant increase in the adoption rate. This late change in the trend can be attributed to changes in the car rental market, which has recognised the potential role of this industry to improve the sustainability of the UK transport sector [137,138]. Enterprises have identified a business opportunity for the rental car sector and created new business models [139,140], increasing the EV fleet [137,138]. Additionally, it has been noted that common users of rental cars may have a higher preference to buy an EV in the future [140].

Figure 58a and Figure 58b shows the spatial distribution of EVs and PVs, which in general differ from each other. Then, Figure 59a and Figure 59b show statistically significant clusters and the nature of such clusters, which differ between the two technologies.

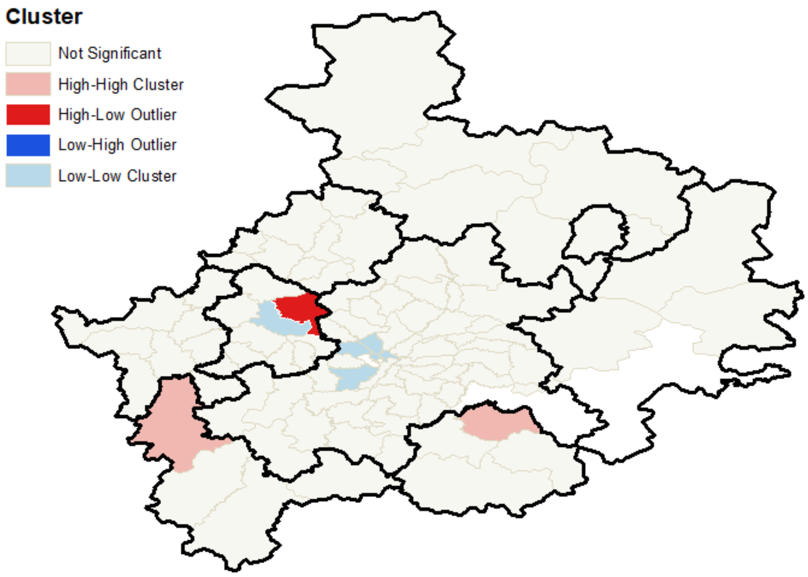


(a)

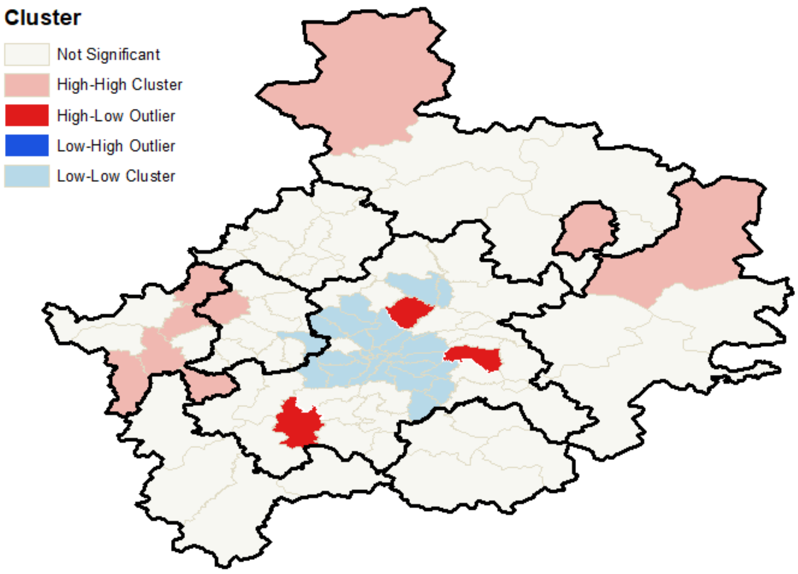


(b)

Figure 58. Spatial distribution of (a) the EV registrations and (b) PV installations for 99 PCs in the West Midlands (Sept 2018).



(a)



(b)

Figure 59. Hot spot analysis of the distribution of (a) the EV registrations and (b) PV installations for 9 local authorities in the West Midlands by Sept 2018.

Because of the lack of information and evidence on whether the PV spatio-temporal patterns of adoption may inform the EV decision-making, and *vice versa*, a diagnostic analysis is performed. Again, drawing from the spatial-econometrics discipline, an ordinary least-squares (OLS) regression is used to inform the model and provide more information about the regularities between these technologies. The OLS has been used to estimate the relationship between the EVs or PVs and a set of independent variables [8,28,29,46,47]. The OLS was constructed using the following a linear equation:

$$EV_i = \beta_0 + \beta_1 PV_i + \varepsilon_i \quad \textbf{(5-1)}$$

Table 10 summarises the results of the OLS estimation. Although the results show that the PV can marginally explain only 16% of the EV adoption rate, the PV variable is found to have a positive and significant impact on EV adoption and highlights the need to include other variables into the model.

Table 10. Summary of OLS Results - Model Variables

Variable	Coefficient	t-Statistic	Probability	Robust Prob
Intercept	65.794	7.423	0.00000*	0.000000*
PV	0.165	4.331	0.00004*	0.000001*
Multiple R-Squared	0.169383		Adjusted R-Squared	0.160354

The statistical significance is marked with asterisks.

**p<0.05*

Because any degree of correlation is a two-way measurement of association between two observed phenomena, by considering the possibility that the spatio-temporal patterns of PV can inform the EV diffusion, also the *vice versa* scenario is considered. Then, it is assumed that the decision of adopting EVs and PV are not mutually exclusive, contrary to the case of vehicle selection where choosing a specific vehicle type will exclude the other types [19]. To compute this multiple decision, the model uses the ANNs capabilities to estimate multiple outputs, while the connections between neurons can capture the knowledge exchange.

5.2 Methods

5.2.1 Decision-making and social effects

The autoregressive EV model developed in Chapter 4 is here extended to recognise the influence of PV adoption, whilst considering the spatio-temporal dynamics and social effects; the same principle is applied to characterise the PV adoption. This chapter combines the insights from both autoregressive models, to generate a combined model that characterises both decision-making processes simultaneously. Chapter 4 has shown that despite the differences in the spatio-temporal patterns of EVs and PVs, the framework can analyse both processes using the same model. Figure 60 exemplifies the multi-output ANN, which processes the same inputs for both outputs. However, the ANN estimates the EV and PV estimations through unique weights from the hidden to the output layer. Equations (5-2) and (5-3) describe the factors that define the estimation of each technology.

The model is initialised following the process described in Section 2.2.5, in this case, the query done to the database retrieves both time-series, number of EVs and PVs. Then, the ANN learns through the iterative process of presenting input-output pairs and adjusting the synaptic weights. This adjustment occurs when the ANN calculates the associated estimation error and propagates this through the network in the backward phase (see Algorithm 2 in Appendix 4). Then, as shown in Figure 60, output nodes have individual synaptic weights connected to the node in the hidden layer, and these are adjusted based only on the node they are connected to. However, the weights connecting the hidden layer and the input layers are modified by both outputs. On the one hand, the individual estimation of EVs and PVs are provided by the unique combination of connections from the hidden layer to the output layer, providing the unique learning to estimate either EVs or PVs. On the other hand, the shared connections from the input to the hidden layer transfer knowledge between both technologies.

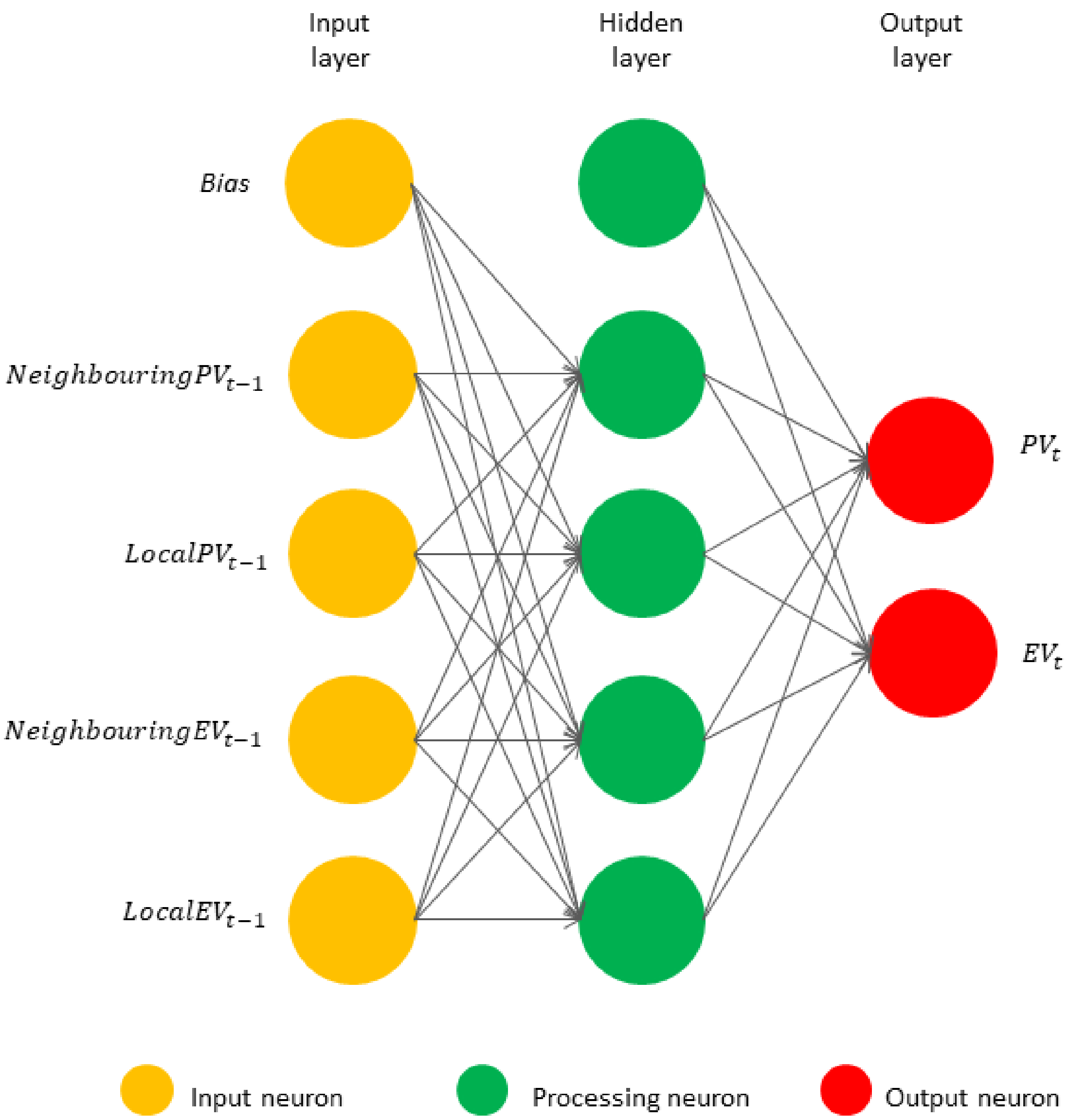


Figure 60. Multi-output ANN design for the EV and PV adoption process.

$$EV_t = f(LocalEV_{t-1}, NeighbouringEV_{t-1}, LocalPV_{t-1}, NeighbouringPV_{t-1}) \quad \textbf{(5-2)}$$

$$PV_t = f(LocalPV_{t-1}, NeighbouringPV_{t-1}, LocalEV_{t-1}, NeighbouringEV_{t-1}) \quad \textbf{(5-3)}$$

Where

PV_t *and* EV_t *are the total number of PVs and EVs in a specific time*

$LocalPV_{t-1}$ *and* $LocalEV_{t-1}$ *are the autoregressive elements*

$NeighbouringPV_{t-1}$ *and* $NeighbouringEV_{t-1}$ *are the number of PVs and EVs in the adjacent areas*

5.2.2 Model validation

Because of the multi-output design of the ANN, it is required to evaluate the errors of estimation for EVs and PVs individually. Therefore, this analysis presents the capabilities of the model to reproduce the spatio-temporal patterns of both technologies, using the validation method in Sections 2.3.2 for PVs and in Section 4.2.5 for EVs. Furthermore, as a benchmark for comparison, these results are compared with the ones produced by the autoregressive models presented in sections 2.4 and 4.3, respectively.

5.3 Results

5.3.1 Temporal validation

Figure 61 shows the actual data and the estimations for both PV and EVs. Although both estimations are very similar to the actual data during the training period, they exhibit different behaviours during the forecast period. Like the autoregressive PV model, the combined model overestimates the number of PVs. This overestimation is reduced by the PV multivariable models, thus, these variables[27] could be included in the model to reduce the error accumulation. On the other hand, the number of EVs presents the same underestimations as in the EV autoregressive model's estimations. However, any knowledge being transferred from the PV helps to eliminate the negative trend forecasted for the EV autoregressive model (see Figure 50). Indeed, the combined model forecasts a positive trend, yet, this still does not reflect the dramatic increase in the number of EVs. Then, future work could include socioeconomic variables in the EV decision-making.

The PV estimations also present an overestimation at the beginning of the training, yet contrary to the EV estimations, these keep large until the rates of adoption start to increase around January 2012. There is a large difference between the EVs, and PVs estimation errors are because the MAPE is sensitive to the estimation of low numbers. Then, as the PVs present values close to zero, under or overestimating one-unit results in a large error, whilst the EVs are in the thousands and the MAPE would not be affected as much as the PV estimation.

Then, by July 2013 the ANN can react and adapt to the changes in the adoption behaviour. However, when the behaviour changes again around 2016, the model starts overestimating the rates of adoption. This could be due to two reasons, firstly, because it takes a few timesteps for the ANN to adapt to the new network. Secondly, for this specific case, the forecast period starts before showing whether the results are under or overestimated. The overestimation of the results is similar to the PV autoregressive, which accumulates the errors of estimations and quickly overestimates the number of PVs. Chapter 3 shows that a multivariable model helps to reduce the error of estimation during the training period and slows down the error accumulation during the forecast period. Therefore, further analysis may focus on including socioeconomic variables into the combined model, moreover, to investigate whether both technologies can be explained using the same socioeconomic variables.

---

[27] As presented in Section 3.3, these variables include electricity consumption, the average household size, and income.

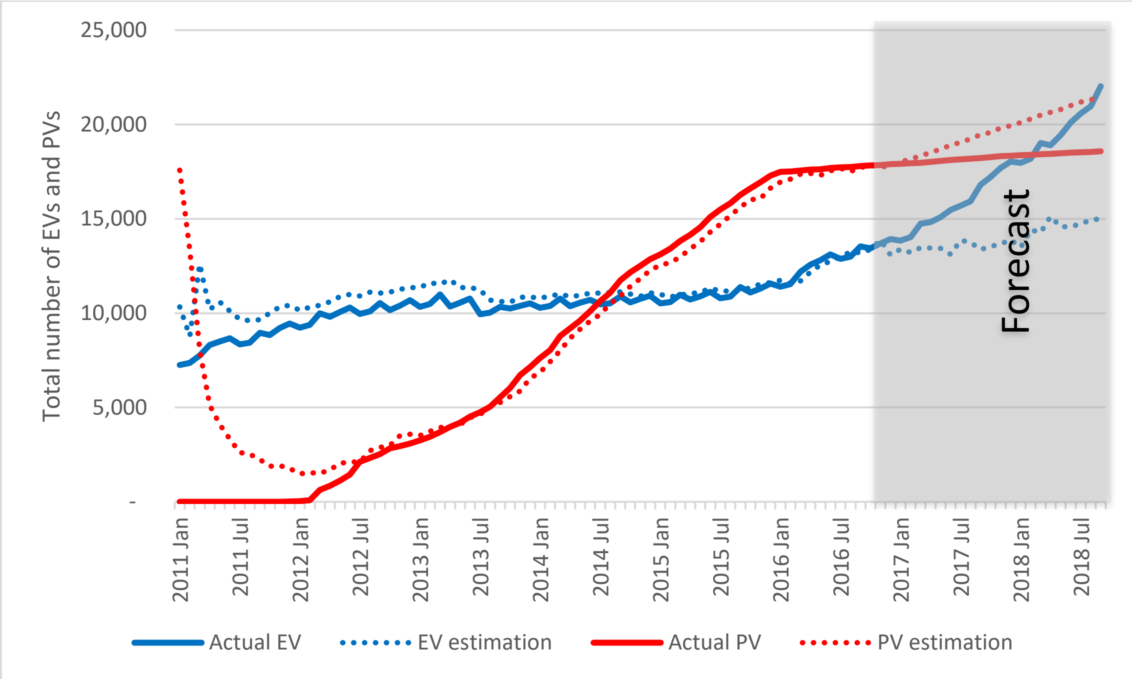


Figure 61. Estimations for the PV installations and EV registrations, for the 96 PCs.

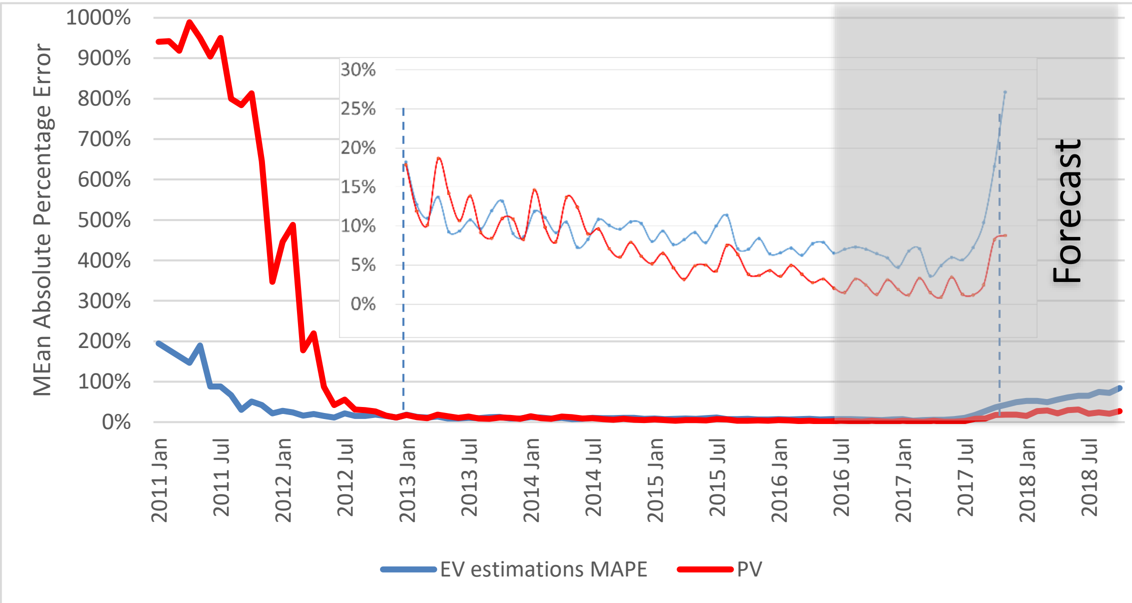


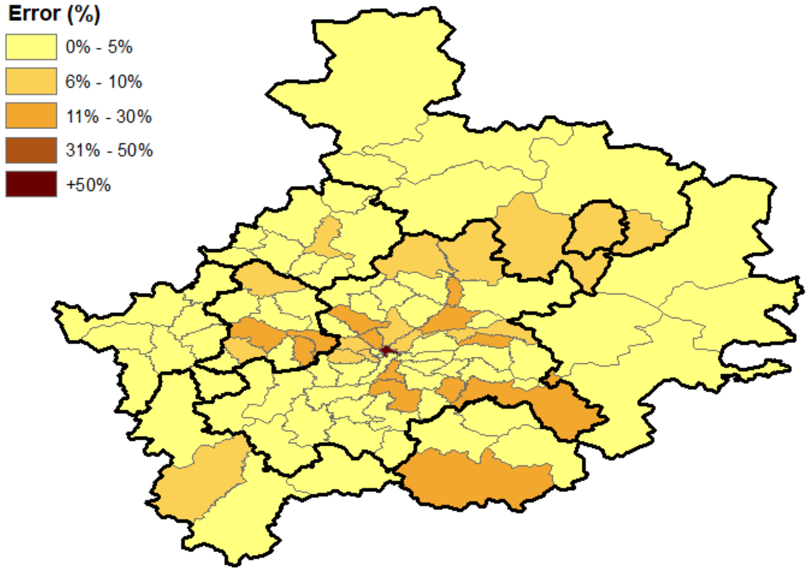
Figure 62. The error of estimation for the EV and PV series - MAPE.

*** the second graph is a zoom in to the estimation errors**

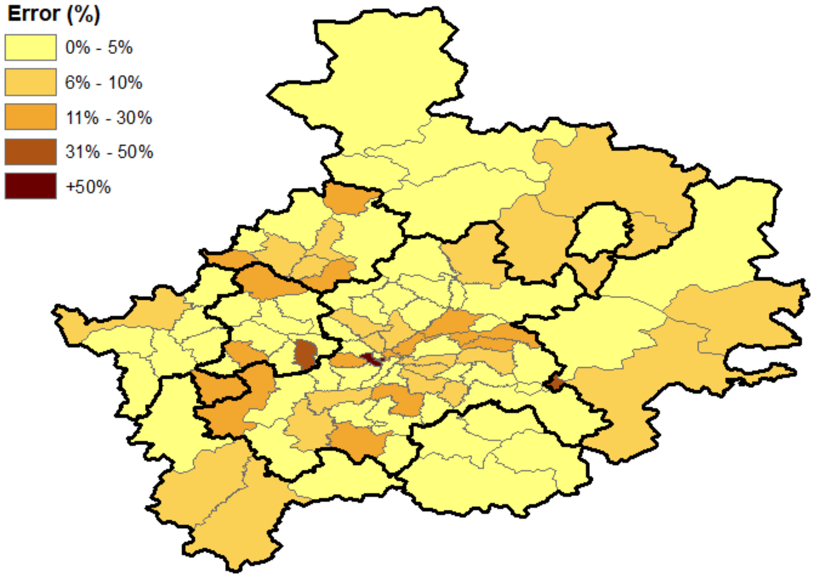
Both EV and PV estimations are likely to produce extreme values at the beginning of the training, as the neural networks have not been fed with sufficient information [118,121–123]. Figure 63 shows the errors of the PV and EV estimations, which exhibit similar behaviour. Both estimations produce a high level of errors initially. Then, as the ANNs are fed with more data, the errors decrease and stabilise without any disturbance in the mid periods. Although there are fluctuations in the errors of estimation, these do not exceed a difference of 2% on average. This is contrary to the PV autoregressive model that presents marginal changes greater than 10%. The absence of disturbances in the PV estimation errors suggests that the knowledge associated with the EV decision-making also informs the estimation of PVs. This could be because the positive trend of the adoption rates of EVs up to March 2012 is followed by a steady number of PVs matches with the changes in the FiTs rates. Despite the results presented in Table 10 are from a linear regression between, which disregards the temporal and spatial dependence of the adoption process, the results show a 16% of correlation between the EVs and PVs adoption rate.

5.3.2 Spatial validation

Figure 63a and Figure 63b show the distribution of the estimation error at the end of the training, where most of the PCs have an estimation error lower than 10%; 81 PCs for PV, and 77 PCs for EV. There are few areas with more than 50% of error (1 and 2 PCs for each technology), which are located in the city centre. Then, Figure 64a and Figure 64b present the spatial regularities of the estimation errors, which present a random distribution, except for the clusters of high values in the city centre PCs. These spatial regularities show that the combined model can explain more of the spatial dependence than the autoregressive models and the PV multivariable model. The combined model can reduce the number of clustered and outlier PCs, see Figure 23, and Figure 53a and Figure 64a for comparison. The areas exhibiting spatial regularities are those with low numbers of EVs and PVs, and even a small under or overestimation may result in a high MAPE.

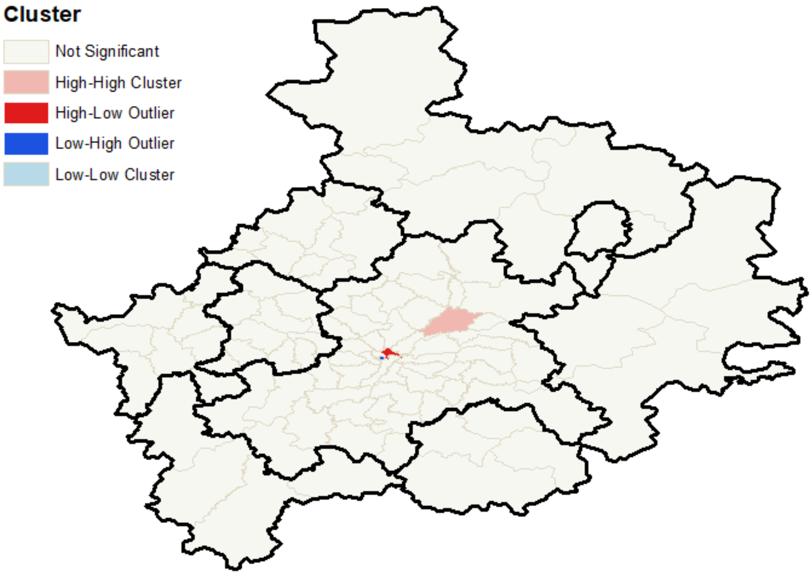


(a)

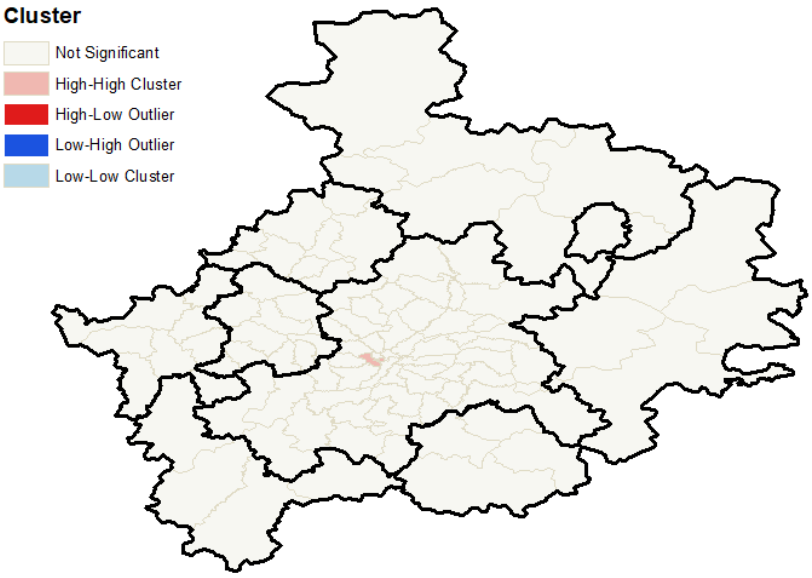


(b)

Figure 63. Spatial distribution of the estimation error for the (a) PV and (b) EV, for the 96 PCs in the West Midlands.



(a)



(b)

Figure 64. Hot spot analysis of the estimation error for (a)PV and (b) EV, for the 96 PCs in the West Midlands.

5.3.3 Predictive accuracy

Table 11 shows the estimation errors for the combined model and both autoregressive versions, presenting the results in intervals of five months to facilitate the comparison with these. The combined and autoregressive EV model estimations are similar until August 2017 (8^{th} forecast). Then the errors for the combined model accumulate, reaching more than twice the error of the autoregressive model for a total of 85% vs. 39%, by September 2018. Chapter 3 showed that the magnitude of the error accumulation can be reduced by introducing other explanatory variables. Already, the spatio-temporal patterns of PVs have informed the EV adoption decision-making, eliminating the negative trend from the autoregressive model. However, the model fitness requires improvement, which could be done by introducing socioeconomic variables to the model; as shown to work for the PV multivariable model in chapter 3.

On the other hand, the influence of the PV over the EV adoption has a larger impact on the spatial dependence, this can be seen in the reduction of spatial regularities exhibited by the estimation errors. The EV autoregressive model presents clusters of high values in the city centre and the southwest, these are the areas with the lowest and the highest values of EVs (see Figure 43). However, the combined model can capture some of the missing spatial dependence, and reduce the clustering of high MAPE values (see Figure 64b). This suggests that the influence of the spatio-temporal patterns of PVs on the EV adoption decision-making has a stronger effect spatially than temporally. The ACF and PACF showed in Chapter 2 and Chapter 4 suggest that most of the time-series present a 1-lag temporal dependency, however, few areas present other significant lags. Therefore, a future analysis could focus on generating a model that increases the agents' heterogeneity by introducing those lags into the decision-making.

For the modelling of PV diffusion, the PV autoregressive model is capable of producing short-term forecasts, as the estimation errors start to accumulate after the third month. Table 11 illustrates the (only) five forecasts produced by the PV autoregressive and MV-LADs models, these may not be fully comparable with the PV estimations of the combined model, but they exemplify the quick accumulation fo the estimation error. As seen, the first and fifth PV autoregressive estimations have larger errors than the first and 18th estimations of the combined model. The combined PV model produces an estimation error of 4% and keeps error levels lower than 30% by the end of the forecasting period. These levels of error are lower than the autoregressive and multivariable models. Likewise, the spatial distribution of the errors improves, meaning that the combined model captures more of the spatial dependence of the adoption process. On the one hand, the PV autoregressive model presents up to ten PCs which are part of a high-value cluster or outliers in the city centre (see Figure 23a). Then, the PV multivariable model can reduce the number of these PCs in the city centre, yet, the Birmingham airport PC is still shown to be an outlier, as this also has a minimum number of PVs (see Figure 38b). On the other hand, the combined model can further reduce the number of PCs with high-value errors and the outlier in the airport PC (see Figure 63a). As the number of EVs and PVs are equally low in the central PCs, this might be due to the decision-making of each technology is being informed by the other. Then the model can reinforce the ANN learning based on both behaviours, in other words, the model can exchange knowledge between the same area about the low or high number of EVs and PVs.

The fact that the combined model can reduce the spatial regularities of both estimation errors suggests that the knowledge exchange between PVs and EVs may inform the spatial dependence of their decision-making. On the other hand, the temporal patterns of the errors change differently in comparison with the autoregressive models. The EV estimations are marginally improved, whilst significantly improved for the PVs. This may suggest that the PV adoption is more sensitive to the influence of EV. This suggests that the *t-1* lag combined with the exchange of knowledge may not capture the individual decision-making towards the EVs. Therefore, it is needed to investigate the effect of the other significant lags on the EV decision-making, furthermore, whether the combined model can characterise two decision-making with a different time span.

Table 11. Forecasted MAPE for the autoregressive and combined models

Model	First forecast*	Eighth forecast	Fifteenth forecast	Eighteenth forecast
EV – autoregressive	5%	20%	27%	39%
EV – combined model	4%	18%	52%	85%
PV – combined model	1%	8%	28%	27%
*PV – autoregressive***	*10%, 11%, 17%, 46%, 73%*			
MV-LADs	*4%, 5%, 7%, 7%, 8%*			

*First month: April 2017; Eighth month: October 2017; Fifteenth month: March 2018; Last month: September 2018.

** The autoregressive model estimates only five forecasts monthly, from September 2015 to December 2015.

Nevertheless, these results are in line with the emergent evidence of regularities between EVs and PVs. However, as pointed by Lanzini and Thøgersen [144], it is not clear what the temporal dynamics of the behavioural spillover are. This could be because of the differences in the interaction and engagement that the individuals have with these technologies. One can argue that, from the installation of the PV and the purchase of the EV, the observational element is present in the agents' experience (neighbourhood effect), however, it is not clear when an agent starts talking about one or the other with its peers (peer-effect). Based on the data analysis performed in sections 2.2.1 and 4.2.1, this work studies the immediate effect month-by-month, suggesting that the adoption decision-making is influenced by the decisions made in the previous month. This differs from other authors, for instance, Cohen and Kollmann [36] suggest that these householders owning a PV have the intention to purchase an EV during the following five years. Richter [20] assumes that the nature of this delay is three-months (t-3), associating it with the time between the decision of adopting and the completion of installation. However, it is not clear whether the decision is influenced by last-month decisions and delayed by external factors, or if the decision is influenced by a longer temporal lag. Given that the existing literature focuses on the time span of behavioural spillover but not in the spatio-temporal dynamics, future work can focus on understanding the spatio-temporal patterns of the behavioural spillover.

5.4 Potential of applicability: network management

This section presents a rather simple application of the spatio-temporally explicit estimations of the model, focusing on how this may enhance management of the distribution network. The following analysis explores two scenarios that inform about the total capacity needed in the distribution lines to either accommodate the PV generation or to supply the EV demand [8].

PV generation

This scenario assumes that the PV generation is first used for domestic purpose and the surplus is injected into the network during the peak generation at noon. Thus, the analysis considers a 4kW PV system [8,124], which accounts for the 96% of the FiT registration [8], and domestic consumption of 0.8kW [145], resulting in a total of 3.2kW surplus for each PV installation. Figure 65 shows the total monthly load for the first eight months of the forecast, which have an accuracy greater than 80%. This can inform about the overall trend of PV adoption, however, this fails to inform whether any specific areas should be prioritised. Then, Figure 66 shows the evolution of the distribution of the capacity required to accommodate the extra PV load, providing a sense of the areas that may be prioritised. By April 2017 there are some spots at the edge of the study area with more than 1MW of solar energy surplus. Yet, seven months later, the areas with < 1MW increases and their spatial distribution changes, being present closer to the City centre. Although this can inform about the PV penetration up to eight months and complement the overall estimations, the model disregards the characteristics of the distribution network. This may misinform the DNOs decision-making, as the areas with the highest loads injected to the grid may already have enough capacity or the other way around. Therefore, a future analysis may consider the distribution characteristics to produce more accurate insights.

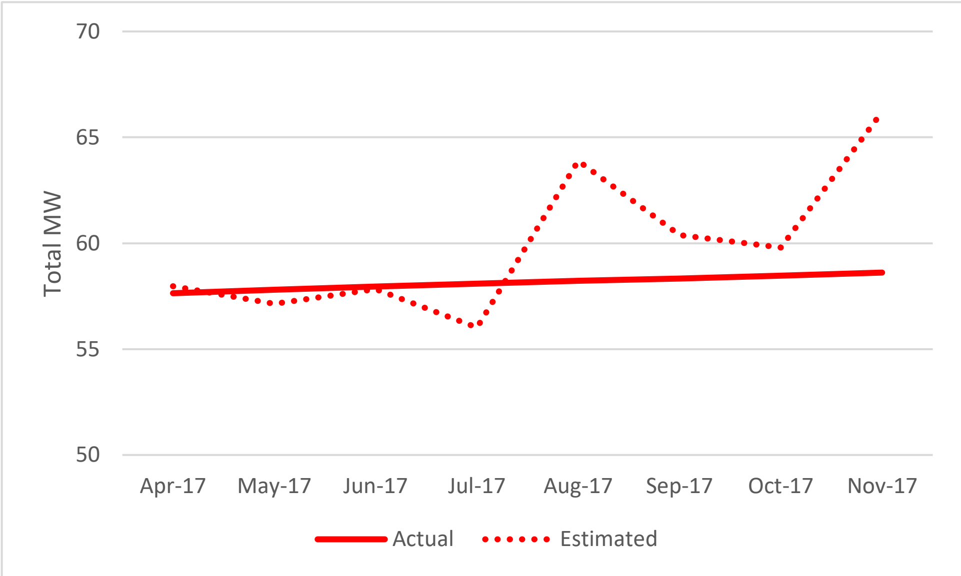
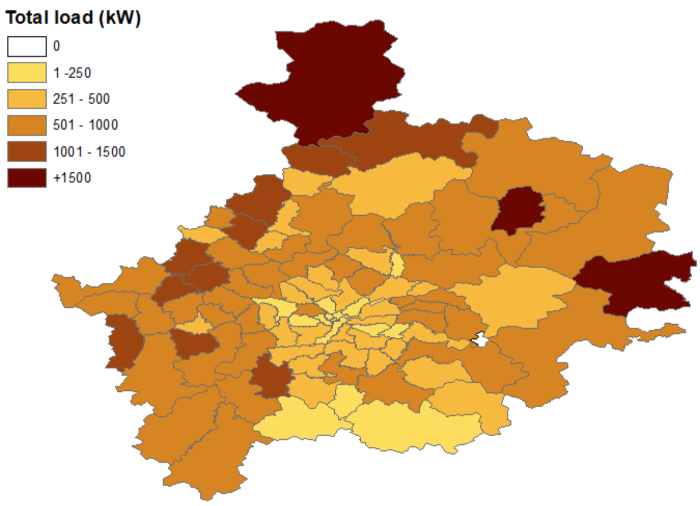
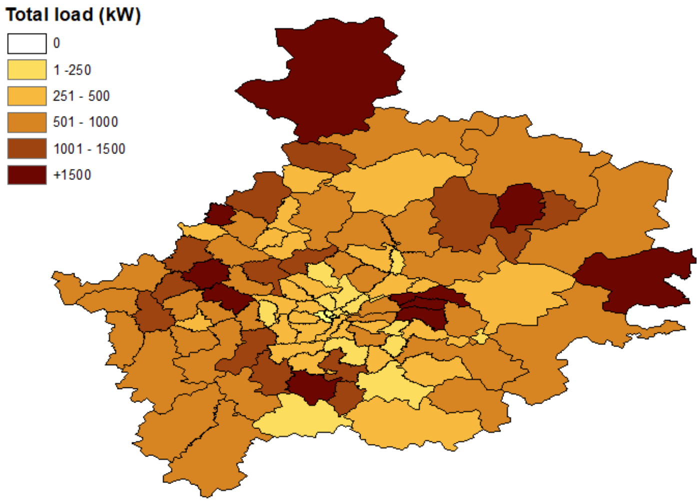


Figure 65. Total PV generation injected to the network after self-consumption.



(a)



(b)

Figure 66. Spatial distribution of the total PV generation injected to the grid – (a) April 2017; (b) November 2017.

EV energy demand

The EV scenario considers that all the EVs are charged at home, or where they are registered at night time. The charging is done using a Type 1 or Type 2 connection with a load of 7.4kW [146,147], and the domestic usage is assumed as 1kW. Figure 67 and Figure 68 show the evolution of the overall energy demand due to domestic EV charging and its spatial resolution. Similar to the PV case, the number of the PCs with high energy demand changes, those start in the west of the study area by April 2017, and then in the PCs closer to the City centre by November 2017. This highlights the areas with a higher chance to have energy faults, assuming uniform levels of headroom capacity at low voltage networks. This scenario presents the same limitations as the PV case because the model does not consider the variation in the distribution network, the results cannot inform accurately which of the areas are to be prioritised. Even when the analysis of the worst-case scenarios can be the bottom line for the planning of network capacity requirements, the model has the potential to produce other insights by dropping the assumptions and integrating empirical behavioural patters in an intraday simulation. For instance, the model could be improved by integrating the location of EV charging stations, travel patterns, or the possibility of using EVs as energy storage. Moreover, the interactions between EVs and PVs have potential benefits for balancing the energy demand and supply [44,148].

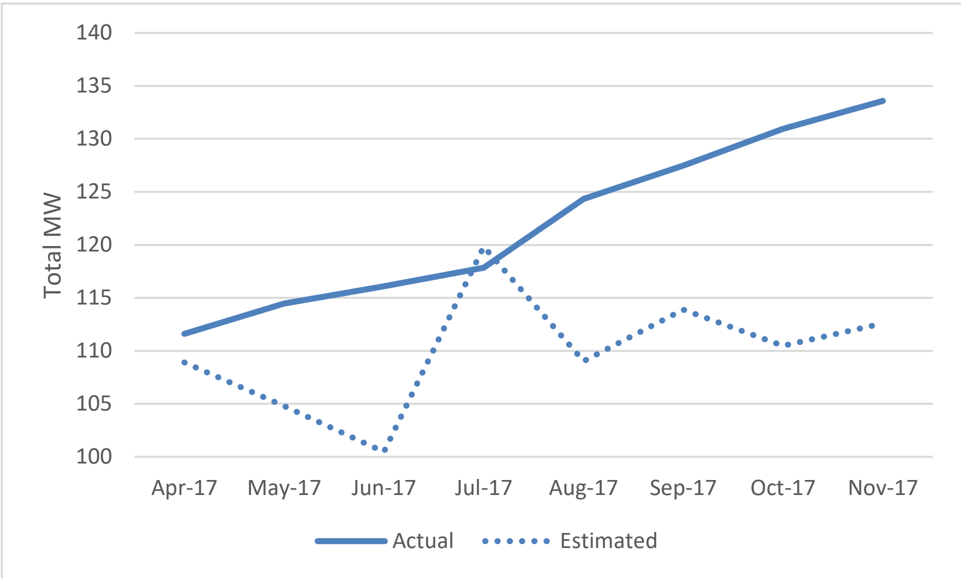
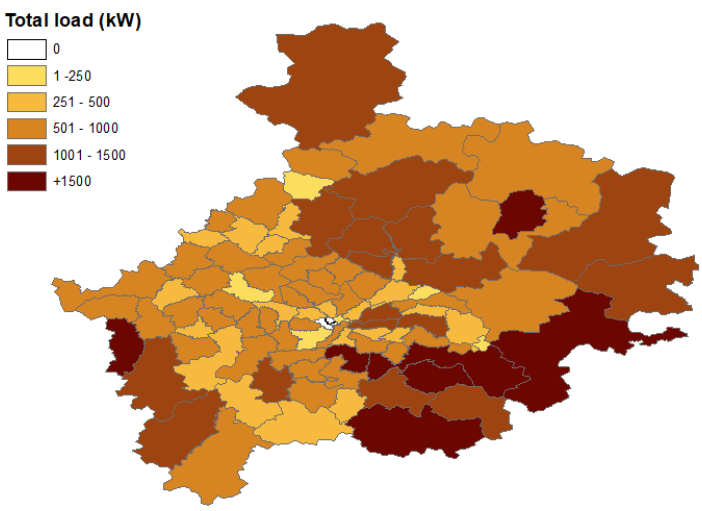
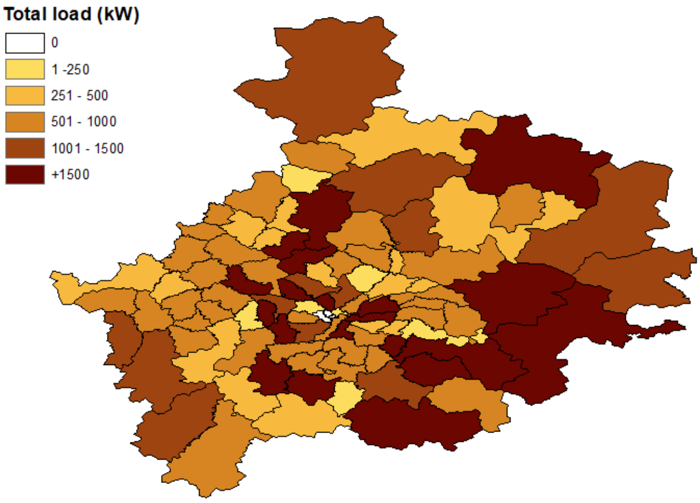


Figure 67. Total domestic demand during EV charging.



(a)

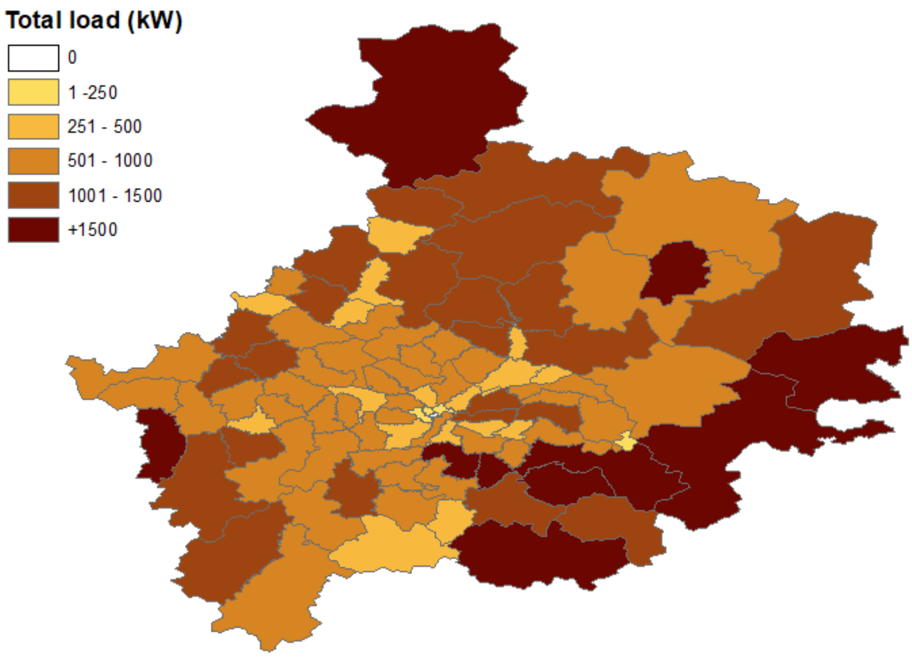


(b)

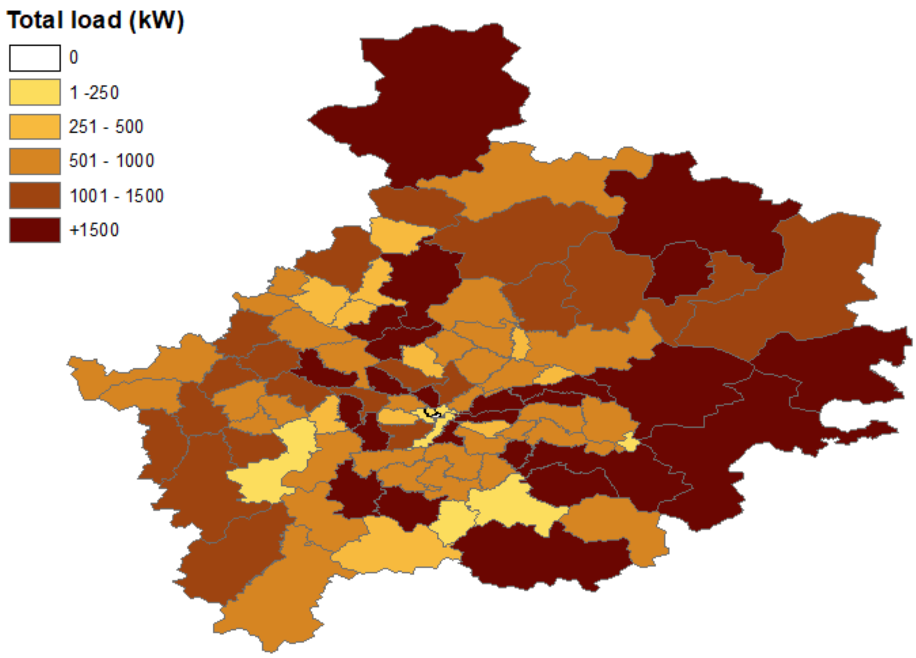
Figure 68. Spatial distribution of the total domestic energy demand during EV charging – (a) April 2017; (b) November 2017.

Combined EV and PV impact

Logically, the results should be able to inform the impacts of both EVs and PVs on the grid, rather than independently. To test a potential worst-case scenario where the timing of EV charging does take place in a different period than peak PV generation and excluding battery storage. Figure 65 and Figure 67 show that the total PV load is lower than the EV demand. However, because of the lack of spatial dimension, these do not inform whether an area has a higher PV generation than EV demand. Then, Figure 69 shows the combined impact of both PVs and EVs for the first and eighth forecast, accounting for the maximum value among them. As seen, the areas with high values are scattered around the study area, suggesting that even when the model can replicate the spatio-temporal patterns of adoption, there is a synergy between the impact of EVs and PVs effects on the grid, as the patterns of the combined effect are different of those of each of EVs of PVs individually. Despite the broad assumptions made to produce these insights, this is a first application of the framework that shows its potential as a decision-making tool for both industry and policy as discussed in Section 3.3.1.



(a)



(b)

Figure 69. Spatial distribution of the combined impact of EVs and PVs.

5.5 Reflective summary

This chapter presents a spatio-temporally explicit ABM that characterises the adoption of EV and PV simultaneously, investigating whether the decision to adopt one technology can influence the decision to adopt the other. The model assumes that both decisions may happen at the same time and that they are not mutually excluded/exclusive. Therefore, at each iteration of the training, the knowledge generated is shared and influences both technologies, then, the multi-output ANN estimates the number of EVs and PVs individually. In other words, the model simulates how the agents learn over time from multiple decisions. It is argued that because the learning processes are done simultaneously, using the same inputs, the model can capture the influence of one decision on another. Then, the results show that the estimations of EVs and PVs improve compared to the autoregressive models, suggesting the possibility to exchange knowledge from one technology to another.

The model improves the forecast of the number of PVs of the autoregressive model, providing a framework that is capable of forecasting more than one year ahead, the model only improves marginally the EV estimations, accounting for up to eight months forward. Although, there is a potential for the model to inform policymakers and DNOs' decision-making, by estimating spatially explicit adoption rates of EVs, potential applications are discussed in detail in Section 6.2.3. Additionally, because this is the first attempt to explicitly combine the spatio-temporal regularities of EV and PV correlation, further fine-tuning of the model structure is required, i.e. to capture the specific adoption drivers of each technology; this is further discussed in Section 6.3.

The following chapter contextualises this work within the existing literature, highlighting the contributions of the results and research design, as well as the potential applicability and limitations of the models. The chapter focuses on the potential of the model to inform decision-making for the design of policy incentives, network planning and system reinforcement. Additionally, the potential of ANN to inform the decision-making of other technologies and future work are also discussed.

6 Discussion and conclusions

This chapter discusses the findings of the thesis in context with previous research and concludes about the potential benefits for DNOs and policymakers. First, the Chapter summarises the findings produced on each of the chapters. Secondly, Section 6.2 discusses the contribution to the knowledge and the extent to which this study addresses the limitations of the existing modelling approaches. The contribution to knowledge is organised by reviewing the objectives and research question. The discussion focuses on the implication of the research and the potential of spatially explicit prediction of PVs and EVs adoption. For instance, how the results can inform about the total load PV generation injected into low voltage lines at the local level [22–24], and the risk of faults and power cuts in the distribution networks caused by the uncontrolled charging of EVs [25–27]. Thirdly, Section 6.3 discusses the limitations of the research and avenues for further work and concludes. Lastly, Section 6.4 is devoted to drawing conclusions for this research.

6.1 Summary of findings

6.1.1 Model development

This thesis reports on research that addresses the limitations of the ABM to estimate spatio-temporally explicit rates of EVs and PVs adoption. Chapter 2 refers to the theory of bounded rationality and introduces ANNs as a decision-making process with attributes that enable an aggregated characterisation of the agents. The ANN is used here to replace the current common rule or equation-based decision-making, which assume that agents possess perfect market information to assess the benefits and implications of their choices. Moreover, the ANN allows analysts and researchers to understand the temporal dependence of the adoption process and examine an explicit time horizon within the model. Secondly, the model integrates insights from the spatial regression (SR) to characterise agents as geographical areas and generate spatially explicit estimations of EVs and PVs. Furthermore, the aggregated characterisation allows the model to capture and understand better spatial dependence.

Table 6 (in 2.2.5) shows the comparison between previous integrations of ABMs and ANN, highlighting how the developed model integrates directly the ANN as the agent’s decision-making, as well as to estimate the number of PVs/EVs. Therefore, the direct integration of ANNs as decision-making extends Kang and Choi’s [103] concept of “intelligent agents”, from those which actions are directed by an ANN to those who are independent and generate their own knowledge.

After integrating the SR and ANN into a spatio-temporally explicit ABM, this work investigates the effect of the heterogeneity amongst the population on the model by integrating socioeconomic variables into the decision-making. The multivariable models developed in Chapter 3 shows the evolution of the agents' preferences by dynamically updating the agents' heterogeneity. This addresses the limitation of ABMs that function with adoption rules and adoption threshold constant along with the simulations. Thus, the model captures the evolution of the agents’ presences. In earlier work, Krebs [75] has implemented heterogeneity amongst the population by using weightings for social and utility preferences, thus creating different rules of adoption. Nevertheless, these weights remain constant during the entire simulation. Ernst and Briegel [77] also recognise that when the model differentiates agents' behaviour from one another, even these agents follow the same decision rules over time. Hence, there is no learning from the agents over time. Robinson et al. [76,90,91] partly address this limitation, by implementing the agents' attitude towards PVs using a dynamic parameter. This parameter changes at each step of the simulation accordingly to the interaction with other agents. However, the influence of each agent is uniform, whereas one can argue that the influence of each agent has the potential to be different. Instead, the model developed and explained in this thesis captures the changes in the agents' characteristics by updating the socioeconomic variables at each step, whilst the adoption rules are adjusted at each step by the ANN.

Chapter 4 reports on the development of the autoregressive model that is used to analyse the spatio-temporal patterns of EV adoption. The data analysis produces similar insights to those from the PV analysis, presenting temporal regularities of the first order. However, the EV time-series presented shows evidence of significant lags of fourth and seventh order. Despite those extra lags being excluded from the decision-making, the model can replicate the patterns of adoption. The results also highlight the evidence that clusters of EVs around the Birmingham Airport cause spatial regularities that are not directly related to the density of residential buildings, as in the case of the PVs. The fact that the model has attributes that enable estimation of the number of EVs up to eight months ahead, with an accuracy higher than 80%, suggests that this modelling approach is flexible enough to characterise the adoption of other low-carbon technologies.

The last section of Chapter 4 presents the investigation into whether it is possible to exchange knowledge between two adoption processes. Then, Chapter 5 builds on the findings reported in previous chapters and combines both autoregressive models. The model uses the attributes of ANN to make estimations of multiple outputs from the same inputs, thus capturing knowledge exchange between the PV and EV adoption. The results suggest that when characterising the EVs and PVs simultaneously within the adoption process, one technology's decision-making informs the others and improves the performance of the model.

6.1.2 Temporal dependence

The model draws on the ideas of bounded rationality theory [76,102] to adopt the ANN as the agents' decision-making. This theory informs that the individuals generate knowledge over time to inform their decision-making. This is extended within the model using Kolb's learning framework and the ANN's principle of emulating the human brain [108,109. The former defines learning as the combination of experience and reflection [108,109], whilst the latter bases its learning algorithm on how the brain generates knowledge by pairing experience (inputs) and the decisions (outputs). Kang and Choi [103] develop a theoretical ABM and ANN model that optimises the set of choices available for the agents, however, these options are fixed rules defined by the authors. Despite the ANN learns and adapts through the simulations (experience), the behaviour of the agents is fixed once the optimal set of options has been found. Therefore, Kang and Choi's ANN [103] could be seen as a central agent that governs the system and dictates the behaviour of the other agents. It could be argued that this contradicts the ABM's bottom-up nature, as the ABM assumes that the behaviour of a system emerges as the results of cumulative actions [26] of autonomous agents [68], and that the system cannot be controlled in its whole [43]. Instead, the model here developed allows each agent to generate individual behaviour based on spatio-temporal explicit experience (datasets).

The use of the ANN allows the model to account for the temporal dependence, which is captured by introducing lags in the input data. The model uses monthly resolution and only one temporal lag in the inputs, suggesting that the time span of the decision-making process is one month. In contrast, in the reviewed studies, there is a recurrent quarterly characterisation of the adoption process [68,88,89,91], suggesting that the adoption has a temporal dependency of three months. Instead, this thesis shows evidence for a shorter time-span to model the decision-making. First, the data analysis for the PVs and EVs time-series reveals that using quarterly basis can result in inconsistencies with the actual behaviour. Because of the data aggregation, specific local behaviour and the effect of interventions are softened and may be disregarded (see Sections 2.2.3 and 4.2.2). Secondly, the data analysis reveals a dependency of one lag for the PV and at least one lag for the EV, suggesting that the influence of modelled elements is almost immediate. This can also be seen in the response of the agents to the FiT revisions (see Figure 7), as the shift in the data behaviour happens almost immediately after the official announcement of changes in the FiT rates. This is consistent with Richter [20], as this study is one of the exceptions for a monthly definition of the adoption process.

In summary, the thesis has built on the theoretical work of Kang and Choi's [103] to integrate the ABM and ANN. Furthermore, the thesis has transferred those insights from the econometrics and artificial intelligence disciplines into the energy planning and policymaking fields, making use of others decision-making and learning theories. Therefore, the model has addressed the need for more realistic approaches, pointed by Noori and Tatari [49], producing the following insights:

- Integrating the ABM and ANN is possible assumng that the agents' decision-making considers past experiences rather than complicated calculations.
- The use of ANN allows to include an explicit time horizon into the simulations, as well as to capture the temporal dependence of the adoption process.

6.1.3 Spatial dependence

The model builds on the theoretical work of Kunz [102] in the field of organisational learning, and the empirical work of Bierkandt et al. [101] in the fields of Supply Chain and Environmental Economics. Kunz [102] states that it is possible to characterise groups of individuals as a single unit of decision-making. These groups share information with other groups producing learning that resembles the spillover effect from the SR, which reflect the information flow between areas [8,47]. The groups also share information within the same group's individuals, similar to the peer-effect to capture the information flow within the same area [20,76]. Bierkandt et al. [101] model perturbations to the supply/demand chain due to environmental disasters, and the reaction of production and consumptions sites to these. These sites are comprised of agents that represent regional industries (production) and aggregated regional consumption (i.e. households, government, other production sites, etc.). The authors assume that the agents in the same site behave and make decisions in similar ways, then, the model characterises the single decision-making of the whole site. This disregards the individual characteristics of the agents and rather consider only cumulative supply and demand capacities. Therefore, the developed model integrates the peer-effect as the social influence within an area by assuming that the agents in the same PC share similar socioeconomic characteristics and decision-making. The model uses the PC resolution which accounts for 10,297 households on average, which is greater than 600 and 3,000 households for LSOAs and MSOAs. This could also involve a loss of heterogeneity of the agents' characteristics, however, the PC resolution has been used before to analyse the adoption process of PVs [20]. Moreover, as seen in

Sections 2.2.2 and 4.2.2, the use of PC and monthly resolution provides the best combination of data variability and availability.

Richter [20] notes that households with high environmental concern tend to live in environmentally friendly neighbourhoods, resulting in similar preferences towards PVs adoption. Moreover, Graziano and Gillingham [16] mention that individuals actively select their peers by moving into neighbourhoods with the same views and interests. Considering that PVs and EVs are adopted by those households with high environmental concerns [34], Richter's [20] insights about neighbours' similarities may also be true for EVs. Although the similarities in the agents' characteristics, values, attitudes and preferences allow defining singular units of decision-making, one can argue that this is inconsistent with the definition of the ABM. In principle, the ABM characterises the autonomous behaviour and interaction of individuals, which leads to the emergence of new patterns of behaviours. In this sense, the ABM and ANN model keeps these elements, except that it loses heterogeneity when aggregating the agents. The common ABMs face a similar issue when simulating agents from existing data (i.e. surveys, interviews, etc.), except that the use of aggregate data alleviates the intensive data demand of the ABM to characterise each individual or household.

At the individual level, Eppstein et al. [33] highlight that the agents tend to associate with those with similar attitudes and behaviours, resulting in social-networks of individuals with similar preferences towards EVs. However, Ernst and Briegel [77] note that individuals may associate with peers even if these are spatially far away. Therefore, because this work uses the adjacency criteria for the spillover effect, the model has not yet implemented the social influence from peers in distant PCs. Furthermore, future work could check the influence of different spatial weight or metric not necessary based on Euclidean distance.

The thesis aggregates households using PC resolution, keeping the real-world layout and making the model spatially explicit. Then, the spillover effect informs the decision-making of how PVs/EVs in one area may influence the decision-making in the adjacent areas. This influence is then weighted against the distance between the population centroid of each area, so the model includes the spatial dependence of the adoption process. The PC resolution follows Richter [20] work, however, this thesis improves the spatial characterisation of the adoption process by not omitting any of the PCs. Richter [20] excludes those areas without any PV installations from the analysis, omitting three Birmingham PCs with (6% of the total PCs). These areas are in the city centre and they have a low number of residential buildings. In contrast, the AMB and ANN's results consider these areas and show that these areas tend to create spatial regularities of estimation errors (see Sections 2.4.2, 3.3.2, 4.3.2 and 5.3.2) because the MAPE is sensitive to low numbers. Nevertheless, the estimation errors for the rest of the areas does not exhibit spatial regularities, demonstrating that the model can replicate the spatial patterns of adoption.

Besides accounting for the actual spatial layout, the model also contributes to improving the decision-making by including both the spillover effect and the peer-effect as the social influences for the adoption of PVs and EVs. While Kunz [102] characterisation of information flow between groups is similar to the spillover effect between areas, Richter [20] implements the peer-effect as the flow of information within the area.

In this sense, the thesis has developed an empirical characterisation of the Kunz [102] model and has improved the Bierkandt et al. [101] model through the introduction of learning into the model. Furthermore, the thesis has transferred those insights from their fields of work, Organisational Learning and Environmental Economics, into the Energy Planning and Policy fields by considering a group of agents as a single unit of decision-making. Consequently, the model has addressed the need for spatio-temporally explicit estimations of EVs/PVs, whilst producing the following findings:

- An aggregated characterisation of the ABM is possible assuming that the group of individual share similarities.

- The aggregated characterisation allows the use of the spillover effect, which can be a proxy of the spatial dependence if it is weighted against the distance between areas.
- The aggregated characterisation can implement the peer-effect as the influence of past adoption rates within the same area.

6.1.4 Results

The model has shown to be able to replicate the spatio-temporal patterns of EV and PV adoption. The spatial validation of the models is based on the model's capabilities to replicate the current spatial distribution of the data; these results are presented in sections 2.4.2, 3.3.2, 4.3.2 and 5.3.2. Then, Table 12 summarises the accuracy of the four models and compares against models from the literature. The ABM and ANN are highlighted in grey for visual identification. It is also indicated whether the errors of estimation converge or diverge during the training or fitting period of the models. Because SR models look at how much of the spatial dependence can be explained by the independent variables, the accuracy of the models may not be fully comparable between SR and the ABMs, however, this may work as a benchmark for the models' results. Table 12 also shows the significant variables and whether they have a negative or positive effect on the SR.

Table 12 shows that the ABM and ANN model has higher accuracy than the SR or epidemic models. However, the results show that the estimation errors present spatial regularities still: see Figure 23, Figure 38, Figure 53 and Figure 59. This means that the model may not be capturing some of the spatial dependences. These clustering patterns exhibited by the errors are mostly the areas with a low number of residential buildings, including the city centre and the PCs adjacent to the airport. One may suggest integrating a parameter to limit the top or ceiling of a maximum number of EVs and PVs in an area, much like the Bass model considers the total of the potential number of adopters. However, the case of the PCs adjacent to the airport is different, as this specific economic activity influences the number of EVs registered in the adjacent PCs. This suggests that the spatial distribution of the EVs and PVs may be subject to disturbances.

This is similar to the temporal validation which reveals some disturbances in the estimation errors, as the agents react to the government's interventions to facilitate PV uptake [8,20]. It is important to look at the clustering of the error to understand the nature of any effects that have not been included yet. Then, investigating the patterns of the estimation errors can inform about the nature of these factors. For instance, a cluster of errors could mean that variables such as the FiTs have not yet been investigated, as the influence on the PCs may vary across the entire area of study. In contrast, outliers in the estimation errors could mean the presence of different socio-economic activities such as an airport.

The need for analysing the adoption process at high spatio-temporal resolution is also highlighted by the temporal validation, as these disturbances may disappear when data is aggregated. Table 12 also shows that, in general, the ABM and ANN models perform better than those in the literature and give an explicit time horizon for the adoption process. The temporal validation is based on the model's capabilities to replicate the current data behaviour over time; these results are presented in sections 2.4.1, 3.3.1, 4.3.1 and 5.3.1. The models developed here are shown to perform as well as the ABMs in the literature [75,76,90,91] and can adapt to the changes in the data trends, which are associated with the fluctuation of the government incentives (i.e. FiT).

However, similar to the spatial validation, the quarterly resolution of these studies may ignore some of the temporal behaviour exhibited in higher resolutions [76,88–90]. Therefore, when these models estimate the parameters that drive the decision-making, they may be ignoring some of the temporal behaviour. This is also related to the stage of the technology adoption, as the available data may not reflect any change in the data trends, as in the case of [88,89]. These two studies analyse the PV data for a period where the data only exhibits a rapid increase in the number of installations. However, looking at the other reviewed studies, one may argue that the data will exhibit changes in the adoption rate, like in the case of the reviewed ABMs [75,76,90,91]. Therefore, it is not clear whether Zhao et al. and Zhou et al. models can adapt to these fluctuations [88,89]. This highlights that approaches for the adoption of PV and EV, and other low carbon technologies need adaptive features.

This adaptive behaviour has been pointed out by the complexity science [9,10,43–45], which informs that a socio-technical system such as the energy system is adaptive, in the sense that the evolution of the system emerges from the collective behaviour of the agents. However, the rules that dictate that behaviour of the agents from the relevant ABMs are fixed at the beginning of the simulations, which constraints the adaptability of the agents and does not allow any type of learning [77].

Even though the model emulates the experience-based knowledge generation [105,121,122], the errors of estimation still present disturbances. The results of the autoregressive and multivariable PV models show that the estimation errors exhibit peak matching with those of the FiT changes, highlighting the need to include this as an explicative variable to fully account for the changes in the agents' behaviour. Figure 33 shows the share of the PCs that are sensitive to the changes in the FiT and matching the disturbances in the errors with the announcements of reducing the FiT. Therefore, there is still a need to inform the model about these changes in behaviour.

On the other hand, the combined model developed in Chapter 5 reduces these disturbances, suggesting that at some degree the knowledge associated with the EV decision-making accounts for the change in the agent preference for PVs. This could be associated with the evidence of empirical regularities between the ownership of these technologies [36].

Table 12. Summary of models' performance behaviour of the estimation errors.

Model	Ref	Model's fitness	Estimation errors and significant variables
EV autoregressive	§ 4.3.1	99%	• Errors converge • Estimation errors do not present disturbances over time
ABM - Engy Inn.	[75]	98%	• Error diverges
PV combined model	§ 5.3.1	96%	• Errors converge • Estimation errors do not present disturbances over time
ABM – PV	[91]	96%	• Errors diverge
PV multivariable	§ 3.3.1	95%	• Errors converge • The model can reduce the disturbances in the estimation errors, yet, they present a peak in Nov 2011 • Income, electricity consumption and average household size are the variables that yield the best fitness
EV combined model	§ 5.3.1	95%	• Error converges • Estimation errors do not present disturbances over time.
ABM - PV	[68]	95%	• Error converges
Hybrid model (ANN) – PV	[88]	94%	• Errors slightly diverge • Current data does not present changes in the data trend
ABM - PV	[76]	94%	• Errors converge
ABM - PV	[90]	94%	• Errors converge
PV autoregressive	§ 2.4.1	90%	• Errors converge • Estimation errors present disturbances at the point where the FiT is announced to be reduced
Mathematical – PV	[89]	82.5%	• Errors converge

Model	Ref	Model's fitness	Estimation errors and significant variables
			• Current data does not present changes in the data trend
Spatial regression - PV	[8]	75%	• Positive effect: income, education level, electricity sales, irradiation, the share of detached houses • Negative effect: share of owned houses, population density and average number of households
Spatial regression - EV	[29]	72%	• Positive effect: the proportion of cohabiting couples under 35, household size • Negative effect: income, education level, financial incentives.
Spatial regression - PV	[18]	71%	• Positive effect: solar radiation, house density, the share of owned houses, GRP per capita
Spatial regression - PV	[15]	61%	• Positive effect: ROI, income, education • Negative effect: unemployment, age <20, the share of new buildings per 10,000 existing units
Spatial regression - Engy Inn.	[46]	60%	• Positive effect: house value, square footage, share of graduates, share of vacant housing unit, population density, wood heating, • Negative effect: taxes
Spatial regression - PV	[14]	55%	• Positive effect: share of houses with a mortgage or equity loan, number of rooms, owner of an EV, share of graduates, household size, foreclosure risk score • Negative effect: share of houses with a mortgage between 20% and 40%, share of PV third-party-owned
Geographical regression - PV	[12]	54%	• Positive effect: income, age
Bass model - EV	§ 4.3.1	18%	• Errors diverge
Bass model - PV	§ 2.4.1	14%	• Errors diverge

6.2 Discussion

6.2.1 Contribution to knowledge

This work analyses the spatio-temporal patterns of EV and PV adoption, by developing an integrated agent-based and artificial neural networks model. To integrate the ANNs as decision-making one must assume that the agents make decisions based on experiences and perception rather than rational or complicated calculations, as inform by the bounded rationality theory [76,102,104]. This is the first time ANN neural networks are integrated directly as the agents' decision-making to address a real-world problem, whilst allowing to predict future adoptions rates. The framework integrated the ANN as decision-making, first, because of the need to produce temporally explicit estimations of the adoption rates of EVs and PVs. Secondly, because the relationship between the adoption process and the factors that drive it are not linear, the model makes use of the ANN's universal estimation capabilities. Moreover, given that the behaviour of the EVs and PVs data is non-linear and presents changes in the data trends, the model uses the ANN's adaptive capabilities to learn from the changes in their tendencies over time.

Then, the model also integrates the analysis at aggregated level from the spatial regression, to address the need of capturing the spatial dependence into the decision-making. Thus, the agents are defined as geographical areas and their characteristics are the aggregated (i.e. total electricity consumption at a time period, number of EVs and PVs adopted at a particular time period) or average value of these (i.e. size of the household, income)..

The research also contributes to advancing the understanding of the spatial and temporal dependence of the adoption of EVs and PVs, the effect of socioeconomic variables, and the regularities between the ownership of EVs and PVs. The contribution to knowledge is revised following each of the objectives, and discussing to what extent the thesis answers the research question. Then, section 6.2.3 discusses the implication of the research and produces new insights that may inform DNOs and Policy Makers' decision-making.

Objective 1. *Investigate how the spatio-temporal and social dynamics of the adoption process can be captured explicitly by an ABM, drawing insights from the SR and integrating the ANN approach as the decision-making process.*

This research develops the first ABM and ANN that uses the neural model to characterise the agents' decision-making. This approach was selected as the agents' decision-making because of the inherent adaptiveness and capacity of the ANNs to analyse time-series. This allows the model to account for the temporal dependence of the adoption process, which is captured by introducing lags in the model's input. The results shown in the Chapters 2,3,4 and 5 demonstrate the importance of learning and adaptivity of the decision-making. However, one can argue that other techniques emulate human reasoning or that are inspired in biology. For instance, Fuzzy Logic generates knowledge from a set of inference rules, by assuming that decision-making is uncertain and imprecise [111–113]; or the Genetic Algorithms is a biology-inspired optimisation model that can extract knowledge from complex datasets [111,112,114]. Therefore, the use of ANN directly as the agents' decision-making opens the door to other techniques that might increase the efficiency and confidence in the insights produced by ABM.

The reviewed ABMs select and estimate certain parameters to find the combination of those that best describe the reality (i.e. adoption threshold, energy prices, social-network size). Then, these models can inform about the differences in the values of such parameters by performing sensitivity analysis, making a transparent tool for decision-making [26,68,76,93]. On the other hand, it is a challenge to explain the effect or influence of each of the ANN's input variables. The knowledge created by the ANN is stored in numerical form for the synoptic weights, which cannot be translated or understood therefore, the acceptance of the ANN is limited. [149–151]. Therefore, there is a trade-off between the transparency and explanation of the ruled-based ABMs and the performance and complexity of the ANNs. Nevertheless, complementary techniques can be added to the framework developed in this research that helps to understand the structure of the ANNs. For instance, data visualization techniques have been used to give the user a sense of reference to where to start the interpretation of the synaptic weights. [152]. Other authors such as Olden and Jackson [151] have performed a sensitivity analysis on the inputs, so to understand the effect of these from outside the model instead of trying to understand the values of the synaptic weights.

To implement the temporal variable into the decision-making, it is also necessary to unveil the degree of the temporal dependence. The statistical tests ran in Sections 2.2.3 and 4.2.3 show that the most common temporal dependence is of order one (*t-1*). Nevertheless, the result presented in Sections 2.4.1 and 4.3.1 show that introducing one lag in the model's input is enough to account for the temporal dependence of the adoption process. Therefore, the integrated model addresses the limitation of the ABMs to provide an explicit time horizon to the simulations

The autoregressive models can estimate monthly PV and EV adoption rate upmost three months, with an accuracy higher than 80%. The Bass model is used as a benchmark to validate the models, a common approach used to model the adoption of innovation [153]. The results show that the autoregressive models perform better than the Bass model especially for the months in the middle of the time series, those with most of the fluctuations in the number of EVs and PVs. In the case of the PVs, the results show disturbances in the estimation errors, which may be caused by the changes in the FiT rates. Because these rates do not vary across space, this variable was not included. However, the model can adapt to the changes in data behaviour. For the EV case, the estimation errors do not exhibit these disturbances, suggesting that the adoption of EVs are not as sensitive to incentives as the PV. Despite both models performing better than the Bass model and having an accuracy of 90% and 99% for the PV and EV cases, the PV model presents a quick accumulation of errors having an accuracy lower than 25% by the fifth forecast. Then, it is expected that a multivariable characterisation of the adoption process could reduce error accumulation.

To account for the spatial and social dynamics that drive the adoption process, the model adopts an aggregated characterisation of the agents Drawing from the bounded rationality theory, the model assumes that a group of individuals with similar socioeconomic characteristics make similar decisions, thus, groups can be characterised as singular decision-making units. Then, this aggregated characterisation of the agents, allows the model to integrate the spatial dependence as a function of the spillover effect and peer-effect. First, the spillover effect which is weighted against the distance between areas captures the spatial dependence and the influence of adjacent areas on the decision-making. Secondly, the peer-effect is captured by the effect that the individuals within an area have over those in the same neighbourhood and reflect how individuals tend to associate with those of similar attitudes and values. Therefore, the model addresses the limitation of the ABMs to produce spatio-temporally explicit estimations of EVs and PVs rates of adoption.

Therefore, the first objective of the research is considerably achieved, as Chapter 2 informs about the theories that underline the integration of ANN and ABM, as well as the procedures to identify and implement the spatio-temporal and social dynamics of the adoption process. However, the aggregated characterisation introduces two main issues, the loss of heterogeneity and the possibility of the model being subject to the MAUP. Both issues are related to the size and number of agents in the study, in turn, these are dependent on the number data availability and variability. In Sections 2.2.2 and 4.2.2, the datasets for EVs and PVs are analysed descriptively, suggesting the PC resolution. This implies that a single ANN will account for an average of 10,297 households, without accounting for (standard) deviation from the average values.

An alternative point of view could be to assume that a fraction of the adoption rates at each location or for the whole population is due to a random effect [26], thus, the model could represent a higher degree of heterogeneity. Additionally, the adoption process could be improved too, for instance, the simulations could have an extra step, before the agents' characterisation, which informs about the different number of significant lags. Then, the model could define different ANN structures for each of the agents, increasing the uniqueness (heterogeneity) of the agents and accounting for the whole of significant temporal lags. Other alternatives to face the loss of heterogeneity and the transparency of the ANN’s elements are discussed below, along with Objective 2.

Objective 2. *Extend the characterisation of the decision-making process, by integrating the agents' socioeconomic variables into the model, to capture the effect of the population heterogeneity.*

Once one has assumed that the agents’ decision-making is driven by experience and perceptions, the model should also account for the influence of the socioeconomic characteristics of the agents. Although, instead of complicated evaluation about affordability [30], energy/fuel economics [32,33] or payback period [75,77,78], etc., the ANN is fed with a list of variables directly. Moreover, given the need to reflect the evolution of the agents’ preferences, the values of such socioeconomic variables are updated at each time step during the simulation. Given the error accumulation observed during the forecast of the PV autoregressive model, it is expected that the multivariable characterisation of the decision-making process improves the model’s accuracy. Additionally, the model outputs may be sensitive to the changes in the study area, Chapter 3 also systematically increases the number of agents in the simulations (extending the area of the study).

Chapter 3 extends the model by including socioeconomic variables, which improves the model performance from 90% to 95%. The model multivariable model slows down the accumulation of the estimation errors, forecasting up to five months with an accuracy higher than 90%. This shows that a relatively small improvement in the model’s performance leads to a lower accumulation of errors. This may suggest that the more variables are included in the agents’ decision-making is marginally improved in the short-term, whilst resulting in more confident decision-making into the future.

The variables are selected from a list that has been found to be relevant for the adoption of PVs, these include factors that are internal (i.e. income, education level) and external to the agents (i.e. population density). The common ABMs would try either to find a value for the coefficients related to each variable that best describe the adoption process, or to analyse the effect of such coefficients on the adoption patterns in the long term. This would reflect the preferences of the agents and inform about the effect of each of the variables.

However, as mentioned before, an ABM and ANN model would have difficulties to validate or to explain the meaning of the selected variables without the assist of other studies. For instance, this research builds on the insights of the SR to make sense of the **electricity consumption** variable being selected, reflecting that those households with a high energy usage tend to be concerned about being self-sufficient [8] and reduce their energy bills [16]. This suggests that besides improving the performance of the model, the multivariable model can also identify some of the contexts of the adoption process. However, the ABM and ANN may not be self-sufficient to provide a sense of meaning to the selected variables without the assist of more contextual studies. For instance, the SR can translate the selected variables and the value of their coefficient to straight clear insights. Moreover, because the initial list of variables corresponds to a single relevant study ([8]), one can argue that a different list of variables may result in a different configuration and insights produced by the model. Also, one can argue that the results and variables selected may be exclusive to the area of Birmingham and surrounding Local Authorities. Therefore, a future analysis may include other cities of the UK to investigate whether the same list of variables produces similar model's results.

The results also suggest that, even when the socioeconomic data varies among the models of different sizes, original and extended ones (see Chapter 3), increasing the number of agents does not change the performance of the models more than 2%. Consequently, despite the aggregated characterisation of the agents and the implementation of the spillover effect and peer-effect, the model is arguably resilient to the MAUP. This may be because of the adaptability of ANNs and that they are characterised as independent decision-making units. However, similar considerations to the multivariable characterisation are needed for further validation, as these results may be particular for this specific case study.

Hence, Chapter 3 substantially accomplishes the second objective of the research, as this extends the autoregressive model into the multivariable model, increasing the degree of heterogeneity of the agents' characterisation; also attending to some of the limitations of the autoregressive model. Nevertheless, the multivariable model also introduces two main limitations, the initial list of where variables are selected and the need for contextual justification of the results. The first could be alleviated by choosing a different list or initial variables, or by merging multiple lists of variables. The second implies that the variables to be introduced in the model are delimited to those that have been already studied and shown to influence PV adoption. Because of the data-driven nature of ANN, if the model is to include the amount of apple consumption in a PC, it would be hard to explain the nature of the influence of such variables on the agents' preferences.

Objective 3. *Analyse the spatio-temporal patterns of EV adoption drawing from the insights of the PV model, and assess whether the model is flexible enough to characterise other technologies' adoption process.*

Because the research aims to model the spatio-temporal patterns of EVs and PVs of adoption, the research builds on the insights of the PV autoregressive model and analyses the adoption patterns of EVs. The analysis carries out a statistical test to identify the temporal dependence of the EV time-series, finding temporal regularities of first, fourth, and seventh order. The analysis also finds clustering of EVs in the adjacent areas to the Birmingham airport. Despite that the EV and PV spatio-temporal regularities are different, the model can estimate EV adoption rates up to eight months ahead, with an accuracy higher than 80%.

Chapter 4 follows the same framework developed for the autoregressive PV model. The results are also used to investigate if the framework is transferable or replicable to EVs. In general, the methodology behind both autoregressive models presents similar results but also similar limitations. This could be because they are both low-carbon technologies, and it is noted that they are likely to be adopted by the same type of individuals, those with high environmental concern. Then, Chapter 4 analyses the data from the adoption rates of EVs, showing three main differences compared to characteristics of PV data: (i) the number of extra significant lags is significantly greater, (ii) the EV data presents positive and negative adoption rates, (iii) the high concentration of EVs is arguably driven by the rental business around Birmingham airport.

Firstly, it could be argued that by ignoring the extra significant lags from the EV data the framework is simplifying the EV adoption process in a greater degree than for the PV case. Secondly, the results from the EV model complement those from the PV model, in the sense that the former informs that the model can characterise positive and negative data trends. This means that the agents' decision-making can reflect not only the decision to adopt but also the decision to undo that decision; in this case, some of the EVs being taken *off-road*. Similar to the negative trends, the third difference suggests that the model can identify positive and negative spatial dependence, as the number of EVs in the airport and surroundings PCs is inversely proportional.

Therefore, the third objective is greatly achieved, as regardless of the differences between EV and PV data, Chapter 4 characterises the adoption process with 95% accuracy. However, future research should focus on extending the model so as this includes the omitted temporal lags, which also fall in place with the recommendations done for the PV models. First, the model could be informed about the number and order of temporal lags, having them stored together with the EV/PV time-series. Then, these values would be used to define the structure of each ANN and the inputs for them, this would create a more unique and realistic decision-making for the agents. And secondly, to implement the *off-road* effect, the model can assume that a fraction of the adoption rates is a random effect. However, contrary to the PV case, this effect may be negative or positive, as some of the cars may also change the PC of registration. In summary, the framework developed to this point addresses the limitation of the common ABMs, which require fixed rules of behaviours for the agents. Instead of these rules for the adoption of EVs and PVs, the experience-based approaches allow the model to generate the adoption rules based on the data presented to the model. T The order which the framework developed this research was due to data availability, being the PV data publicly available whilst the EV data required a data request to the DfT. If data availability would have been the opposite, probably the autoregressive model would have a different structure. Nevertheless, it would be expected that a complex ANN's design with dynamic structures and inputs driven by the EV data, would handle also simple adoption processes such as the PV's one.

Objective 4. *Investigate whether the spatio-temporal patterns of PV can inform the EV diffusion process, by integrating PV data into the decision-making towards EV adoption.*

Given the need to predict the adoption rates of EVs and PVs to assist the management of the distribution network, and the potential of one technology informing another technology adoption process, Chapter 5 investigates the effect of exchanging knowledge from one adoption process to another. Chapter 3 shows the positive impact of a multivariable characterisation of the agents on the model's performance. However, to distinguish between the effect of knowledge exchange and the multivariable characterisation, Chapter 5 only looks at the integration of both autoregressive models. Therefore, because the adoption of both technologies can be characterised using an autoregressive model with the same elements: (i) spatio-temporal resolution, (ii) temporal dependence, (iii) spatial dependence, and (iv) social effects, the model developed in Chapter 5 assumes that both decision-making can be characterised simultaneously.

Thus, Chapter 5 extends the EV autoregressive model by integrating the input nodes for the PV data and adds one extra output node for the estimation of PVs rates of adoption. Figure 60 shows the design of the ANN that implements the abstraction of the adoption process described in equations (5-2) and (5-3). The knowledge exchange happens during the second phase of the learning algorithm (see Section 2.3.1) when the algorithm adjusts the synaptic weights. Then, both learnings propagate to the entire network because the output nodes are connected to the same neurons in the middle layer, and yet the model also captures the knowledge that is generated exclusively for each technology. Then, the model that exchanges of knowledge yields a model performance with a higher than 80% accuracy for the first eight forecasted months (see Section 5.3.3). This is a higher accuracy than the individual autoregressive models, reducing the disturbances in the PV estimation errors and both error accumulations. Thus, drawing from the behavioural spillover theory, this suggests that when the agents consider previous decisions made in a near past (one month) their decisions in the long term are mode accurate.

The concept of knowledge exchange is introduced to the ANN as the behavioural spillover effect [144], which reflects the tendency of some agents to have a higher preference towards EVs if they already possess a PV and the other way around. The model implements this by having both EV and PV datasets as inputs for the ANN and follows the same temporal dependence of the than for the rest of the inputs. However, might be argued that it has been pointed that it is not clear what is the time spam for the spillover to occur, thus the model again faces the issue of fully represent the temporal behaviour of the adoption process. In this case, future work would require a similar analysis to the temporal correlation of the EV and PV, so to complement the results of the OLS carried out in section 5.1.1.

Moreover, building from the discussion of objective 2 and 3, the model could be extended by introducing socioeconomic variables for both technologies. These variables may be similar as in the case of van der Kam et al. [28] or be from completely different sets. Because the ANNs are flexible in their structure, they can represent even a complex configuration of the inputs.

In sum, these results and those from previous chapters provide evidence to validate the initial hypothesis, as explained in the following section. Indeed, the modelling of the adoption process of low carbon technologies has evolved into more complex approaches, from the population growth theory to the recognition of spatio-temporal and social dynamics. Then, this research takes forward the ABM by integrating the ANN directly as the agents' decision-making, yet, recognising that this is just the first step towards a more realistic representation of the adoption process. finally, as demonstrated in this research, there is understudied potential of integrating elements of the human cognition into the modelling approaches.

6.2.2 Review of the research question

Chapter 1 introduces the following research question:

To what extent is it possible to characterise the adoption process of EVs and/or PVs, whilst integrating the spatio-temporal regularities and the different factors that drive the adoption process?

To inform this question, the systematic literature review included in Section 1.1 highlights which characteristics of agents' may also present spatio-temporal regularities. It also shows that the adoption process is driven by social dynamics and that, further, there is an influence from adopting other technologies. Consequently, the research question is developed into the following hypothesis:

It is possible to explicitly characterise the spatio-temporal dynamics of the decision-making towards EVs and PVs, whilst including the social dynamics and the relationship between these technologies.

The thesis breaks down the hypothesis into fours objectives (see Section 1.6), each of one is developed in Chapters (2-5). The thesis builds on the relevant ABM and SR studies, developing a novel model to characterise the decision-making process of the adoption of low carbon technologies at high spatio-temporal resolution. Therefore, the results and insights are synthesised in the following statements that attempt to answer the research question:

1. Indeed, it is possible to explicitly characterise the spatio-temporal dynamics of the decision-making towards EVs and PVs (Objective 1 and 3). The approach uses spatially explicit datasets to characterise agents as geographical areas and feed this to an ANN to have an explicit time horizon. The spatial characterisation requires keeping the actual spatial layout and high resolution to be applicable/useful for informing policymakers and network operators. This research uses the postcode and monthly resolution, which is the combination that best trade-off data availability and variability.
2. It is also possible to include the social dynamics that influence the adoption of EVs and PVs (Objective 2), the model includes the peer-effect and the spillover effect. The former can be characterised as an autoregressive element that reflects the influence of the number of EVs/PVs on the individuals within the same area. The latter is modelled as the influence of the total of EVs/PVs in the adjacent areas and weighted against the distance to the population centre.
3. Lastly, it is possible to capture the regularities between EVs and PVs ownership, reflecting the influence of the agents' preferences towards other technologies (Objective 4). This thesis reports on the use of the features of the ANN to characterise multi-output decision-making and characterise knowledge exchange between both decision-making processes. The results present potential to inform DNOs and policymakers about the location and pace of EV and PV adoption at high spatio-temporal resolution.

6.2.3 Implications of the research

The model has been proved to estimate spatio-temporally explicit rates of EVs/PVs adoption, providing a more realistic characterisation of the decision-making and using empirical data. Consequently, the model can be used as a tool to inform policymaking and network management. The model could also be extended by mapping the energy assets or substations into the geographical areas while calculating the use and production of electricity based on the estimated adoption. Moreover, given the spatially explicit layout of the model, it could also include the distance between the users and the substations allowing it to account for the distribution losses by [154], or even help to model peer-to-peer energy trading [155,156].

Network management

It has been pointed that with penetration of renewable energy sources the probability of reverse flows significantly increases [157]. Those reverse flows happen because of the difference in the loads' power quality and consistency, which results in stress for the low voltage network (distribution network). Moreover, solar PVs is the renewable source technology with the highest variability in the energy output [24], which makes balancing demand and supplies an issue for the DNOs [158]. Consequently, the management of the network requires to account for the direction of the flows in relation to the components of the network as well as the local technical capacity of the lines [22]. Furthermore, the DNOs require tools that help to foresee the location of high concentration of PVs across the network.

Similarly, the DNOs face issues resulting from a high concentration of EVs, in particular, if the EVs charge from the grid. This implies extra stress for the low voltage network especially during the peak demand period [159], which can happen even if the overall penetration of EVs is very low [26]. It has been demonstrated that the diversification of the charging times can reduce the stress on the local energy infrastructure [27], however, this requires strategies to incentivise users to charge EVs during off-peak time. Therefore, the DNOs must have a tool that informs about the location and pace of the EVs adoption. Moreover, such tools must consider the spatial heterogeneity of the local energy infrastructure, to consider the differences in the lines and substations capacities.

Thus, this research is paragon for network management and DNOs decision-making. The model has shown to be able to forecast between three to eight forecasts with an accuracy greater than 80%. After this point, the errors start accumulating before diverging. Although the other ABMs also present some error accumulation [76,91], the spatially explicit nature of the ABM and ANN model makes the errors to accumulate across space as well. From that perspective, the ABM estimations reported by studies such as Krebs [75], Robinson et al. [91], and Adepetu and Keshav [68] underestimate the error accumulation over space, as they disregard the spatial dependence. Then, even with lower predictive accuracy, the ABM and ANN integrated model produces more realistic estimations.

Therefore, to investigate how to reduce this error accumulation, the reporting of the model's estimations and errors require further understanding and alternative presentations. Because the results point that the agents' heterogeneity and the knowledge exchange help to reduce the error accumulation. These two factors may be interpreted as follow. First, the results of the multivariable model suggest that the more elements integrated into the decision-making, the more accurate is the forecasting [105,121,122]. Secondly, the results of the combined model suggest that as the agents are exposed to more experiences, the more elements of reflection for the agents to include in the decision-making. However, to avoid unnecessary variables, the models can follow the stepwise method used in Chapter 3. The thesis initial list of variables follows Balta-Ozkan, Yildirim and Connor's model spatial analysis of the PV diffusion in the UK [8], however, other variables are significant for the adoption of PVs. Moreover, the effect of some of those variables may differ from study to study. For instance, the share of owned houses has been pointed to have a significant effect, however, the direction of these effects differs. Whilst Balta-Ozkan, Yildirim and Connor [8] find a positive effect and Schaffer and Brun [18] model accounts for a negative effect, the multivariable model developed does not include this variable. Moreover, the effect of these variables may also differ from one technology to another, as the income variable is noted to have a positive impact in the PV uptake [8], while a negative effect for the EV adoption [29].

Secondly, extending the findings of studies that note empirical regularities between EV and PV ownership [14,26,36], the results of the ABM and ANN model suggest that decision of whether to adopt EV or PV is not independent of the other technology decision-making. Instead, the experience of each decision-making intertwines and exchanges knowledge from one to another. When the model integrates the socioeconomic variables or the other technical data series, the model can reduce the disturbances in the errors of estimation. Moreover, the dynamic characterisation of those inputs may reflect how agents also modify their preferences over time, while also reflecting the influence of new technologies in the markets. These new products or services could include new PV products, new EV models, or innovative features in both.

Therefore, this work advances the existing knowledge by proving spatio-temporally explicit estimations of EVs and PVs, moreover, these insights can address the limitations of other current approaches. For instance, Krebs [75] analysis of the adoption of energy innovation can estimate the annual rates of adoption with a 98% accuracy. The model covers the entire German territory, simulating a representative total of 40 million households, implementing a semi-empirical characterisation of defined types of lifestyles. The decision-making is based on the financial and social utility, assuming that different lifestyles have different preferences for each adoption factor/variable. The authors find that individuals close to the manufacturer's headquarters rapidly account for the early adopters and increase rapidly. This could be because of the exposure of the PVs in the area or because employees living in the surroundings are default adopters.

Krebs [75] recognises that this approach based on the bounded rationality is limited, as this assumes agents to have perfect market information. A second limitation is that the factors that reflect the agents' preferences are set to be constant across the types of lifestyles. Nevertheless, the authors point out that the applicability of predictive ABMs increases when real-world geography is included in the model. Accordingly, the ABM and ANN can address the limitations of Krebs [75] framework. First, the use of spatially explicit data sets allows substituting the fixed types of lifestyles for individual and unique agents' characterisations, producing spatio-temporally explicit estimations and creating individual knowledge for each of the areas. Secondly, because the model considers the spatial dependence, the model could inform about the spatial influence of PVs manufacturing companies on the local adoption.

Furthermore, the model could consider the influences of other agents in the simulation. As discussed in Section 1.2.1, the reviewed articles only consider individuals or households as units of decision-making, disregarding the effect of other agents such as government, energy companies, PVs/EVs manufacturers, rental cars companies, etc. (see Section 5.3.2). Thus, future research could focus on characterising other agents and the spatio-temporal patterns of its influence on the adoption process.

Robinson and Rai [76] develop a model at the individual level and quarterly basis for the adoption of PVs in the US. The authors use empirical data to characterise the entire population, estimating the rates of adoption at 94% accuracy. The results show that the best estimations are produced by the model that randomly fits the parameters of the decision-making. The authors point out that this is due to the one atypical over-estimation of the PVs, as the results of a drop in the energy cost, then, the authors remove this observation to improve the model's performance. The authors point out a second limitation is that the adoption threshold is static over time.

The ABM and ANN integrated model addresses these limitations, as the model does not exclude any observations and uses the ANN's adaptive capabilities against interventions/disturbances in the data. Then, the second limitation is addressed by employing the experienced-based approach, as the rules (knowledge) are generated and adjusted over time.

Although these studies provide useful insights on the adoption of EVs and PVs that could help DNOs to improve the management of the local infrastructure, these studies disregard the influence of other low carbon technologies on the adoption of EVs/PVs. Moreover, these disregard the impact of the combined effect of EVs and PVs on the distribution network, which could be positive or negative depending on the network management strategies. On the other hand, frameworks like those of Bhatti et al. [159] and Chaouachi et al. [44] study the potential benefits from the synergy between the extra loads injected to the network by the PVs and the extra energy demand from charging EVs.

Chaouachi et al. [44] develop a distributed system architecture that allows finding the maximum number of EVs and PVs that a local grid can allocate assuming a coordinated charging. The model accounts for the total capacity of the lines, number of costumers, number of substations, and the profiles of PVs production and EVs charging. Despite Chaouachi et al.'s model finds an optimal charging strategy considering the network capacity, it assumes a maximum rate of adoption of EVs and PVs in a typical European city. Moreover, the model does not consider the spatial regularities of the PVs and EVs, instead, it looks at the overall energy supply and demand. Consequently, the model has limitations to inform which substations are more or less likely to face reverse flows or powercuts. Therefore, this framework could build upon the results of the ABM and ANN model to characterise the spatio-temporal patters of EVs and PVs adoption, whilst considering the heterogeneous capacities of the distribution network. Such a framework could inform the reinforcement of certain assets based on on the level of risk of each substation.

Policymaking

Given that the model recognises the disturbances in time and space (such as the FiT revisions, see Sections 2.3, 3.3, 4.3 and 5.3), the model can be used to run sensitivity analysis under different scenarios and help policymaking. Moreover, the multivariable models and combined model show the flexibility of the model to implement a different number of inputs. Therefore, the datasets for the current variables could be modified for the forecast period [40] to reflects hypothetical scenarios such as an increase in the population income. Also, because of the high spatial resolution of the model, these assumptions can be made locally, allowing the model to contribute to the design of policies that recognise intraregional socioeconomic inequalities.

Therefore, this thesis is critical for policymakers and local government. The model improves the existing work the explorative ABMs listed in Table 4. For instance, Adepetu, Keshav and Arya [25] simulate the uptake of EVs under three scenarios: (i) base case, (ii) no government incentives, and (iii) extra government incentives. The authors find that increasing the available incentives is more important than the actual affordability of the EVs as well as the number of available charging stations. However, because the model uses a semi-empirical characterisation, the agents are located uniformly in a regular grid. Thus, the model is limited to explicitly inform if those policies result in spatial regularities of EVs, neither when are those rebates should be announced, be available or and revised.

However, given that the results have shown that not all the areas response in the same way to the change in the governmental incentives (see Figure 20), the framework can be used to analyse the effect of such policies at the local level. On the other hand, the spatio-temporally explicit ABM could be extended to recognise the influence of the charging stations around the city and the average driving distance [25]. Hence, the method could help to investigate the spatio-temporal influence of these on the patterns of adoption, or even find the optimal number and location of the recharge station needed to maximise the rates of adoption.

This last feature is similar to McCoy and Lyons [26] framework that investigates different scenarios of adoption. These scenarios are based on the initial distribution of early adopter and the number of connections in the agents' social-network. The authors find that the location of the early adopters has a significant impact on the adoption rates in the long term. Also, the insertion of random adopters increases the likelihood of adoption of those agents with a low probability to adopt. Nevertheless, the agents' characterisation and location are semi-empirical, resulting in an initial configuration that lacks spatio-temporal accuracy. The ABM and ANN integrated model could address this limitation providing an explicit location for those random adopters.

Moreover, given the temporally explicit nature of the model, future research could investigate the effect of introducing a few EVs in a specific PC at a given time (i.e. as a marketing strategy). Section 3.3.1 shows the effect of increasing the number of agents in the study, suggesting that as the number of agents increases the error of estimation is stabilised during the forecasting error, see Figure 35. Because the number of connections increases proportionally to the number of agents, the model could investigate the effect of connecting each PC with PCs others than those that are adjacent to it [26].

Ernst and Briegel [77] propose an experiment on the social influence of the model to investigate the effect of increasing environmental awareness on the adoption of Energy Innovation (i.e. PVs). The authors increase the frequency of peer-to-peer communication, doubling the peer-effect in the simulations. Then, they find that this experiment increases the adoption rates in the long term (one-two years). However, the model completely disregards the spatial dimension of the adoption process, making these results not applicable to local policy design. The ABM and ANN model could address the limitations of Ernst and Briegel's [77] framework to explicitly locate those areas that are more sensitive to increase the frequency of peer-to-peer communication. Moreover, the model could help to investigate the effect of a higher influence from the peer-effect by doubling the influence of the autoregressive element of decision-making, only in certain areas if necessary.

In general, the findings from the literature can be improved by the model developed in this research because of the explicit and high resolution of the spatio-temporal ABM. Because those ABMs have limited capabilities to inform of the time and location EVs/PVs are adopted, the policy recommendations produced by those models disregard the local differences in the population's heterogeneity. Therefore, the model can help policymakers to understand the impact of different policies at the national level and locally design policies, by enabling multilevel decision-making.

6.3 Research limitations

Addressing the need for spatio-temporal estimations of the EVs and PVs adoption rates has come without imperfections. To integrates the AMB and ANN the framework faces the following main limitations: (i) loss of heterogeneity, (ii) transparency of the results, (iii) and complexity and simplification of the decision-making process.

First, when the model adopts the aggregated definition of agents, the model trades off specific behaviour of the agents to alleviate the high data demand of empirical characterisation of each individual. The framework defines the model resolution by finding the best combination of data variability and availability, thus, the characterisation of the decision-making is dependent on the data availability. Despite the model being validated for different areal definitions, the results presented in Chapters 2-5 are specific to the City of Birmingham and the surrounding local authorities. Therefore, it is important to note that the results are data-driven, especially for temporal dependence. One can argue that different study areas may present different temporal patterns of adoption with longer temporal lags.

Secondly, because the limitations of the ANN to explain the meaning of its elements, the use of ANN to characterise the decision-making limits the ABM to inform about the significance and meaning of the variables by itself. Moreover, as seen in Chapter 3, the framework relies on previous contextual studies to select variables from, as well as to make sense of those which are selected. Also, because the model relies on data variability, variables with low variability across space were omitted. For instance, the literature has found other variables such as the FiT rates or the solar irradiation, however, these have not been included yet because these may not vary locally. Given the FiT is homogeneous across the whole country, one may focus on the temporal effect that this implies, so to plan future interventions of the government (i.e. one-time incentives). Similarly, because the model considers households as the main actors, the model disregards the effect of other actors on the decision-making process. Moreover, given that the understanding of the effect of these actors stills understudied, the inclusion of these effects requires further analysis.

Third, although the ANN structure is shown to be capable of characterising the decision-making and replicating the spatio-temporal patterns of adoption, the design of the neural network is rather simple. Moreover, the implementation of the temporal dependence assumes the all the agents have similar decision-making. Therefore, the design of the ANN could move forward not only in the dynamic design of it but also move to designs that may improve the performance of the model. For instance, due to the significant increase in the use of ANN in the last decade, there is potential to improve the performance of the model by implementing a more complex design. For instance, to increase the model performance, the framework could substitute the simple backpropagation with Long Short-Term Memory [160] or Deep Neural Networks [161]; that have been used to model energy consumption.

This research has addressed some of the agent-based model limitations, by integrating the ABM and ANN. However, given the limitations aforementioned and the potential of the model to enable multi-level decision-making, the following avenues for future research as proposed.

6.4 Future work

Increasing the degree of heterogeneity given the loss of heterogeneity in the model due to the aggregated characterisation, future research may focus on ways to implement elements that compensate for this. Two main alternatives are proposed: (i) unique design for each ANN and (ii) alternative list of variables. The former builds upon the significantly increased use of ANN in the last decade which has resulted in more complex designs. This thesis assumes that the EV owners have a higher preference for PVs (and the other way around), thus a possible scenario could be that the adoption of PVs blocks automatically the adoption of EVs (or the opposite). This could reflect the affluence and affordability of these technologies, where agents can afford only one of the technologies at a time or only one of them permanently. Then, both technologies could be seen as competitive, mutually exclusive from each other. This scenario could be characterised by competitive neural networks [162,163]. In this approach, only one of the multiple output nodes is activated and the rest of the synoptic weights are inhibited. The second point to increase the heterogeneity is to use a different list of initial variables to the analysis, and analyse whether the same variables can predict both the EVs and PVs adoption rates. This also includes the fact that future work may investigate how to account for variables that do not vary across space.

Understanding the simplification and complexity of the model that allows a wider characterisation of the adoption process, by including elements of the human cognition. However, paradoxically, the model takes the ABM a step further from being a transparent approach. Moreover, the aggregated characterisation means that the decision-making is simplified. As seen in the data analysis of the different combinations of resolutions, the aggregation of values softens specific behaviours. Therefore, a future analysis may investigate the trade-off between the simplification of the decision-making due to data aggregation to the decision-making and the opportunity to characterise a more realistic adoption-process. Also, this may study how the model softens or overlooks some effects such as random choices or *off-road* registrations.

Refining the knowledge exchange framework: this research provides empirical evidence that characterising two decision-making processes simultaneously can improve the performance of an ABM. However, because this is the first application that considers this exchange of knowledge, future research may analyse this from the perspective of social or neural science. Although there is literature that reports empirical regularities between EVs and PVs ownership, these are data-driven studies that look at the correlation between these variables. Therefore, there is a need to underline the social mechanics of the knowledge exchange process, and how the knowledge generated by one experience spills-over another decision-making process.

Increasing the type of agents and social influence: the model can be extended by introducing a second type of agent that drives the decision-making by macro-economic variables (i.e. government agent), and which output is the level of incentives. This could also include other agents such as energy companies, EV/PV sellers, green NGOs, etc., and outputs like energy prices, marketing strategies, EV trials and communication events. Temporal changes in the intensity of certain influence could also be included, such as higher social influence to reflect a higher communication frequency or areas with higher transit of users.

Temporary influence and controlled interventions: the model could introduce short term effect/interventions such as focused advertisements (displays and outdoor advertising) as dummy time-series for the seller agents. The model could also investigate the long-term effect of seeding strategies, by introducing EVs/PVs in areas with lower usage of these technologies. Also, further research could temporally change the intensity of certain influences, for instance, higher social influence to reflect a higher communication frequency or areas with higher transit of users.

6.5 Concluding remarks

The research addresses the limitations of the agent-based modelling to inform explicitly the location and pace of EVs and PVs adoption, whilst providing a more realistic characterisation of the decision-making and its evolution over time. A model that integrates the agent-based modelling and artificial neural networks are proposed to inform network operators and policymakers with insights of spatio-temporal patterns of EV and PV adoption. This results in a framework that characterises the decision-making process whilst considering the spatial, temporal, social dynamics, and preferences towards other technologies that drive the adoption process.

This work draws from the spatial regression and the artificial neural networks model to account for the spatial and temporal dependence. An aggregated characterisation of the agents, like the spatial regression, allows implementing the social effects. Besides, the preference's heterogeneity amongst the population is captured by the artificial neural networks, reflecting the evolution of the individuals' preferences.

The model is compared with other approaches based on spatial and temporal accuracy, showing accuracy levels higher than 90% for the training period. In general, the spatio-temporally explicit models developed here perform as well as the existing ones, especially the model that characterises the preferences towards other technologies (knowledge exchange). This work characterises multiple decision-making for the first time, moreover, the model does not require different rules for each technology to be defined by the researcher. Instead, the model generates individual knowledge for each of the areas, and in turn, each area generates unique knowledge for each of the technologies. This exchange of knowledge between technologies results in higher accuracy than the autoregressive models for EVs and PVs that the model builds upon. The model that accounts for knowledge exchange can estimate mid-term rates of EVs and PVs with accuracy levels higher than 80%. The thesis reports how these results are relevant to the energy industry, especially for informing policymaking and network management of the location and pace of EVs and PVs adoption.

Finally, the research is paragon for the modelling of the spatio-temporal adoption patterns of low carbon technologies, which can be exploited as confidence in artificial intelligence models increases and empirical datasets become more and more available. Moreover, this research leads towards more complex approaches of decision-making that recognise the multiple dynamics driving the adoption process of EVs and PVs. Furthermore, this works highlights the new insights they can generate to address the impact of clustering of these technologies on the management of distribution networks. For instance, the energy industry can benefit to predict the location of reverse flows caused solar PVs and extra demand caused by uncontrolled charging of EVs.

REFERENCES

[1] Gummer RHJ, Brown B, Bell K, Chater N, Forster P, Heaton R, et al. Net Zero The UK's contribution to stopping global warming Committee on Climate Change. 2019.

[2] Department of Energy and Climate Change. The Carbon Plan: Delivering our low carbon future. 2011.

[3] Department of Business E& IS. 2018 UK GREENHOUSE GAS EMISSIONS. 2019.

[4] Van der Zwaan B, Keppo I, Johnsson F. How to decarbonize the transport sector? Energy Policy 2013;61:562–73. doi:10.1016/j.enpol.2013.05.118.

[5] Lah O. Decarbonizing the transportation sector: policy options, synergies, and institutions to deliver on a low-carbon stabilization pathway. Wiley Interdiscip Rev Energy Environ 2017;6:e257. doi:10.1002/wene.257.

[6] Foxon TJ. Transition pathways for a UK low carbon electricity future. Energy Policy 2013;52:10–24. doi:10.1016/j.enpol.2012.04.001.

[7] Sweetnam T, Spataru C, Cliffen B, Zikos S, Barrett M. PV System Performance and the Potential Impact of the Green Deal Policy on Market Growth in London, UK. Energy Procedia 2013;42:347–56. doi:10.1016/j.egypro.2013.11.035.

[8] Balta-Ozkan N, Yildirim J, Connor PM. Regional distribution of photovoltaic deployment in the UK and its determinants: A spatial econometric approach. Energy Econ 2015;51:417–29. doi:10.1016/j.eneco.2015.08.003.

[9] Hughes N, Strachan N, Gross R. The structure of uncertainty in future low carbon pathways. Energy Policy 2013;52:45–54. doi:10.1016/j.enpol.2012.04.028.

[10] Ault G, Frame D, Hughes N, Strachan N. Electricity Network Scenarios for

Great Britain in 2050. 2008.

[11] Hughes N, Strachan N. Methodological review of UK and international low carbon scenarios. Energy Policy 2010;38:6056–65. doi:10.1016/j.enpol.2010.05.061.

[12] Langheim R. Third Party-Owned PV Systems: Understanding Market Diffusion with Geospatial Tools. Energy Mark Predict Pap from 2014 AAAI Fall Symp Third 2014.

[13] Balta-Ozkan N, Watson T, Mocca E. Spatially uneven development and low carbon transitions: Insights from urban and regional planning. Energy Policy 2015;85:500–10. doi:10.1016/j.enpol.2015.05.013.

[14] Davidson C, Drury E, Lopez A, Elmore R, Margolis R, W BR, et al. Modeling photovoltaic diffusion: an analysis of geospatial datasets. Environ Res Lett 2014;9:074009. doi:10.1088/1748-9326/9/7/074009.

[15] Dharshing S. Household dynamics of technology adoption: A spatial econometric analysis of residential solar photovoltaic (PV) systems in Germany. Energy Res Soc Sci 2017;23:113–24. doi:10.1016/j.erss.2016.10.012.

[16] Graziano M, Gillingham K. Spatial patterns of solar photovoltaic system adoption: The influence of neighbors and the built environment. J Econ Geogr 2015;15:815–39. doi:10.1093/jeg/lbu036.

[17] Kwan CL. Influence of local environmental, social, economic and political variables on the spatial distribution of residential solar PV arrays across the United States. Energy Policy 2012;47:332–44. doi:10.1016/j.enpol.2012.04.074.

[18] Schaffer AJ, Brun S. Beyond the sun—Socioeconomic drivers of the adoption of small-scale photovoltaic installations in Germany. Energy Res Soc Sci 2015;10:220–7. doi:10.1016/j.erss.2015.06.010.

[19] Adjemian MK, Cynthia Lin CY, Williams J. Estimating spatial

interdependence in automobile type choice with survey data. Transp Res Part A Policy Pract 2010;44:661–75. doi:10.1016/j.tra.2010.06.001.

[20] Richter LL. Social Effects in the Diffusion of Solar Photovoltaic Technology in the UK. Cambidge Work Pap Econ 2013:40.

[21] Copiello S, Grillenzoni C. Solar Photovoltaic Energy and Its Spatial Dependence. Energy Procedia, vol. 141, Elsevier; 2017, p. 86–90. doi:10.1016/j.egypro.2017.11.017.

[22] Jupe SCE. Active Management of Distributed Generation based on Component Thermal Properties. Durham University, 2010.

[23] Masters CL. Voltage rise: the big issue when connecting embedded generation to long 11 kV overhead lines. Power Eng J 2002;16:5–12. doi:10.1049/pe:20020101.

[24] Barnacle M, Robertson E, Galloway S, Barton J, Ault G. Modelling generation and infrastructure requirements for transition pathways. Energy Policy 2013;52:60–75. doi:10.1016/j.enpol.2012.04.031.

[25] Adepetu A, Keshav S, Arya V. An agent-based electric vehicle ecosystem model: San Francisco case study. Transp Policy 2016;46:109–22. doi:10.1016/j.tranpol.2015.11.012.

[26] McCoy D, Lyons S. Consumer preferences and the influence of networks in electric vehicle diffusion: An agent-based microsimulation in Ireland. Energy Res Soc Sci 2014;3:89–101. doi:10.1016/j.erss.2014.07.008.

[27] Cui X, Kim HK, Liu C, Kao S-C, Bhaduri BL. Simulating the household plug-in hybrid electric vehicle distribution and its electric distribution network impacts. Transp Res Part D Transp Environ 2012;17:548–54. doi:10.1016/j.trd.2012.05.011.

[28] van der Kam MJ, Meelen AAH, van Sark WGJHM, Alkemade F. Diffusion of solar photovoltaic systems and electric vehicles among Dutch consumers: Implications for the energy transition. Energy Res Soc Sci

2018;46:68–85. doi:10.1016/j.erss.2018.06.003.

[29] Morton C, Anable J, Yeboah G, Cottrill C. The spatial pattern of demand in the early market for electric vehicles: Evidence from the United Kingdom. J Transp Geogr 2018;72:119–30. doi:10.1016/j.jtrangeo.2018.08.020.

[30] Namdeo A, Tiwary A, Dziurla R. Spatial planning of public charging points using multi-dimensional analysis of early adopters of electric vehicles for a city region. Technol Forecast Soc Change 2014;89:188–200. doi:10.1016/j.techfore.2013.08.032.

[31] Shafiei E, Thorkelsson H, Ásgeirsson EI, Davidsdottir B, Raberto M, Stefansson H. An agent-based modeling approach to predict the evolution of market share of electric vehicles: A case study from Iceland. Technol Forecast Soc Change 2012;79:1638–53. doi:10.1016/j.techfore.2012.05.011.

[32] Sweda T, Klabjan D. An agent-based decision support system for electric vehicle charging infrastructure deployment. 2011 IEEE Veh. Power Propuls. Conf. VPPC 2011, IEEE; 2011, p. 1–5. doi:10.1109/VPPC.2011.6043201.

[33] Eppstein MJ, Grover DK, Marshall JS, Rizzo DM. An agent-based model to study market penetration of plug-in hybrid electric vehicles. Energy Policy 2011;39:3789–802. doi:10.1016/j.enpol.2011.04.007.

[34] Axsen J, TyreeHageman J, Lentz A. Lifestyle practices and pro-environmental technology. Ecol Econ 2012;82:64–74. doi:10.1016/j.ecolecon.2012.07.013.

[35] Sundaram SM. Proposal for establishing an Electric Highway to show the way for future of personal mobility. 2015 IEEE Int. Transp. Electrif. Conf., IEEE; 2015, p. 1–5. doi:10.1109/ITEC-India.2015.7386881.

[36] Cohen J, Kollmann A. HOUSEHOLD APPLIANCE PURCHASES AS A JOINT-DECISION PROCESS: EVIDENCE FROM UPPER-AUSTRIAN

SOLAR ADOPTION. In: Hackenfort M, Carabias-Hütter V, Hartmann C, Janser M, Schwarz N, Stücheli-Herlach P, editors. 5th Eur. Conf. Behav. Energy Effic., Zurich: 2018, p. 91–2.

[37] Rogers EM. Diffusion of innovations. New York : Free Press, 2003.; 2003.

[38] Di Crescenzo A, Paraggio P. Logistic Growth Described by Birth-Death and Diffusion Processes. Mathematics 2019;7:489. doi:10.3390/math7060489.

[39] Frank LD. An analysis of the effect of the economic situation on modeling and forecasting the diffusion of wireless communications in Finland. Technol Forecast Soc Change 2004;71:391–403. doi:10.1016/S0040-1625(02)00392-X.

[40] Al-Alawi BM, Bradley TH. Review of hybrid, plug-in hybrid, and electric vehicle market modeling Studies. Renew Sustain Energy Rev 2013;21:190–203. doi:10.1016/j.rser.2012.12.048.

[41] Linder S, Wirges J. Spatial diffusion of electric vehicles in the German metropolitan region of Stuttgart. ERSA Conf Pap 2011:1–27.

[42] Bodo P. MADness in the method: On the volatility and irregularity of technology diffusion. Technol Forecast Soc Change 2016;111:2–11. doi:10.1016/j.techfore.2016.05.011.

[43] Bale CSE, Varga L, Foxon TJ. Energy and complexity: New ways forward. Appl Energy 2015;138:150–9. doi:10.1016/j.apenergy.2014.10.057.

[44] Chaouachi A, Bompard E, Fulli G, Masera M, De Gennaro M, Paffumi E. Assessment framework for EV and PV synergies in emerging distribution systems. Renew Sustain Energy Rev 2016;55:719–28. doi:10.1016/j.rser.2015.09.093.

[45] Kahraman C. Energy Management—Collective and Computational Intelligence with Theory and Applications. vol. 149. 2018. doi:10.1007/978-3-319-75690-5.

[46] Noonan DS, Hsieh L-HC, Matisoff D. Spatial Effects in Energy-Efficient Residential HVAC Technology Adoption. Environ Behav 2013;45:476–503. doi:10.1177/0013916511421664.

[47] Morton C, Wilson C, Anable J. The diffusion of domestic energy efficiency policies: A spatial perspective. Energy Policy 2018;114:77–88. doi:10.1016/j.enpol.2017.11.057.

[48] Eppstein MJ, Rizzo DM, Lee BHY, Krupa JS, Manukyan N. Using national survey respondents as consumers in an agent-based model of plug-in hybrid vehicle adoption. IEEE Access 2015;3:457–68. doi:10.1109/ACCESS.2015.2427252.

[49] Noori M, Tatari O. Development of an agent-based model for regional market penetration projections of electric vehicles in the United States. Energy 2016;96:215–30. doi:10.1016/j.energy.2015.12.018.

[50] Sopha BM, Klöckner CA, Hertwich EG. Adoption and diffusion of heating systems in Norway: Coupling agent-based modeling with empirical research. Environ Innov Soc Transitions 2013;8:42–61. doi:10.1016/j.eist.2013.06.001.

[51] De Groote O, Pepermans G, Verboven F. Heterogeneity in the adoption of photovoltaic systems in Flanders. Energy Econ 2016;59:45–57. doi:10.1016/j.eneco.2016.07.008.

[52] Drucker AM, Fleming P, Chan A-W. Research Techniques Made Simple: Assessing Risk of Bias in Systematic Reviews. J Invest Dermatol 2016;136:e109–14. doi:10.1016/J.JID.2016.08.021.

[53] Rother ET. Revisão sistemática X revisão narrativa. Acta Paul Enferm 2007;20:v–vi. doi:10.1590/S0103-21002007000200001.

[54] Viswanathan M, Patnode CD, Berkman ND, Bass EB, Chang S, Hartling L, et al. Assessing the Risk of Bias in Systematic Reviews of Health Care Interventions. Agency for Healthcare Research and Quality (US); 2008.

[55] Denyer D, Tranfield D. Producing a systematic review. Sage Handb. Organ. Res. methods, London: Sage Publications Ltd; 2009, p. 671–689.

[56] Rasool G, Ehsan F, Shahbaz M. A systematic literature review on electricity management systems. Renew Sustain Energy Rev 2015;49:975–89. doi:10.1016/j.rser.2015.04.054.

[57] Torreglosa JP, García-Triviño P, Fernández-Ramirez LM, Jurado F. Control strategies for DC networks: A systematic literature review. Renew Sustain Energy Rev 2016;58:319–30. doi:10.1016/j.rser.2015.12.314.

[58] Liu HC, Liu L, Liu N. Risk evaluation approaches in failure mode and effects analysis: A literature review. Expert Syst Appl 2013;40:828–38. doi:10.1016/j.eswa.2012.08.010.

[59] Martinez-Luengo M, Kolios A, Wang L. Structural health monitoring of offshore wind turbines: A review through the Statistical Pattern Recognition Paradigm. Renew Sustain Energy Rev 2016;64:91–105. doi:10.1016/j.rser.2016.05.085.

[60] Moretti M, Djomo SN, Azadi H, May K, De Vos K, Van Passel S, et al. A systematic review of environmental and economic impacts of smart grids. Renew Sustain Energy Rev 2017;68:888–98. doi:10.1016/j.rser.2016.03.039.

[61] Pittaway L, Robertson M, Munir K, Denyer D, Neely A. Networking and innovation: a systematic review of the evidence. Int J Manag Rev 2004;5–6:137–68. doi:10.1111/j.1460-8545.2004.00101.x.

[62] Ioannou A, Angus A, Brennan F. Risk-based methods for sustainable energy system planning: A review. Renew Sustain Energy Rev 2017;74:602–15. doi:10.1016/J.RSER.2017.02.082.

[63] Adams R, Jeanrenaud S, Bessant J, Denyer D, Overy P. Sustainability-oriented Innovation: A Systematic Review. Int J Manag Rev 2016;18:180–205. doi:10.1111/ijmr.12068.

[64] Sayers A. Tips and tricks in performing a systematic review. Br J Gen Pract 2007;57:759.

[65] Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. Br Med J 2005;331:1064–5. doi:10.1136/bmj.38636.593461.68.

[66] Felizardo KR, Mendes E, Kalinowski M, Souza ÉF, Vijaykumar NL. Using Forward Snowballing to update Systematic Reviews in Software Engineering. Proc. 10th ACM/IEEE Int. Symp. Empir. Softw. Eng. Meas. - ESEM '16, New York, New York, USA: ACM Press; 2016, p. 1–6. doi:10.1145/2961111.2962630.

[67] Buchmann CM, Grossmann K, Schwarz N. How agent heterogeneity, model structure and input data determine the performance of an empirical ABM - A real-world case study on residential mobility. Environ Model Softw 2016;75:77–93. doi:10.1016/j.envsoft.2015.10.005.

[68] Adepetu A, Keshav S. Understanding solar PV and battery adoption in Ontario. Proc. Seventh Int. Conf. Futur. Energy Syst. - e-Energy '16, New York, New York, USA: ACM Press; 2016, p. 1–12. doi:10.1145/2934328.2934333.

[69] Bonabeau E. Agent-based modeling: Methods and techniques for simulating human systems. Proc Natl Acad Sci U S A 2002;99:7280–7. doi:10.1073/pnas.082080899.

[70] Macal CM, North MJ. Agent-based modeling and simulation. Proc. - Winter Simul. Conf., IEEE; 2009, p. 86–98. doi:10.1109/WSC.2009.5429318.

[71] Grimm V, Berger U, Bastiansen F, Eliassen S, Ginot V, Giske J, et al. A standard protocol for describing individual-based and agent-based models. Ecol Modell 2006;198:115–26. doi:10.1016/j.ecolmodel.2006.04.023.

[72] Amblard F, Quattrociocchi W. Simulating social complexity: Social Networks and Spatial Distribution. In: Edmonds B, Meyer R, editors.

Simulating Complex. second, 2018, p. 763–800.

[73] Farmer JD, Foley D. The economy needs agent-based modelling. Nature 2009;460:685–6. doi:10.1038/460685a.

[74] van Duinen R, Filatova T, Jager W, van der Veen A. Going beyond perfect rationality: drought risk, economic choices and the influence of social networks. Ann Reg Sci 2016;57:335–69. doi:10.1007/s00168-015-0699-4.

[75] Krebs F. An empirically grounded model of green electricity adoption in Germany: Calibration, validation and insights into patterns of diffusion. JASSS 2017;20:10. doi:10.18564/jasss.3429.

[76] Robinson SA, Rai V. Determinants of spatio-temporal patterns of energy technology adoption: An agent-based modeling approach. Appl Energy 2015;151:273–84. doi:10.1016/j.apenergy.2015.04.071.

[77] Ernst A, Briegel R. A dynamic and spatially explicit psychological model of the diffusion of green electricity across Germany. J Environ Psychol 2017;52:183–93. doi:10.1016/j.jenvp.2016.12.003.

[78] Krebs F, Ernst A. A spatially explicit agent-based model of the diffusion of green electricity: Model setup and retrodictive validation. Adv. Intell. Syst. Comput., vol. 528, Springer, Cham; 2017, p. 217–30. doi:10.1007/978-3-319-47253-9_19.

[79] Bale C, Gale W, McCullen N, Rucklidge A, Foxon T. Harnessing social networks for promoting adoption of energy technologies in the domestic sector. Energy Policy 2013.

[80] Dimatulac T, Maoh H. The spatial distribution of hybrid electric vehicles in a sprawled mid-size Canadian city: Evidence from Windsor, Canada. J Transp Geogr 2017;60:59–67. doi:10.1016/j.jtrangeo.2017.02.010.

[81] Huang B, Wu B, Barry M. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. Int J Geogr Inf Sci 2010;24:383–401. doi:10.1080/13658810802672469.

[82] Baltagi B, Rokicki B. The spatial Polish wage curve with gender effects: Evidence from the Polish Labor Survey. Reg Sci Urban Econ 2014;49:36–47. doi:10.1016/j.regsciurbeco.2014.08.001.

[83] Coll M-H, Vandersmissen M-H, Thériault M. Modeling spatio-temporal diffusion of carsharing membership in Québec City. J Transp Geogr 2014;38:22–37. doi:10.1016/j.jtrangeo.2014.04.017.

[84] Kopczewska K. Cluster-based measures of regional concentration. Critical overview. Spat Stat 2018;27:31–57. doi:10.1016/j.spasta.2018.07.008.

[85] Bansal P, Kockelman KM, Wang Y. Hybrid Electric Vehicle Ownership and Fuel Economy Across Texas. Transp Res Rec J Transp Res Board 2015;2495:53–64. doi:10.3141/2495-06.

[86] Haifeng Z, Yevgeniy V, Letchford J, Lakkaraju K. Predicting Rooftop Solar Adoption Using Agent-Based Modeling. Energy Mark. Predict. Pap. from 2014 AAAI Fall Symp. Predict., 2014.

[87] Abbar S, Boumaiza A, Mohandes N, Sanfilippo A. Modeling innovation diffusion for renewable energy technologies in city neighborhoods. 2018 9th Int. Renew. Energy Congr. IREC 2018, IEEE; 2018, p. 1–6. doi:10.1109/IREC.2018.8362523.

[88] Zhao T, Zhou Z, Zhang Y, Ling P, Tian Y. Spatio-Temporal Analysis and Forecasting of Distributed PV Systems Diffusion: A Case Study of Shanghai Using a Data-Driven Approach. IEEE Access 2017;5:5135–48. doi:10.1109/ACCESS.2017.2694009.

[89] Zhou Z, Zhao T, Zhang Y, Su Y. A data-driven approach to forecasting the distribution of distributed photovoltaic systems. Proc. IEEE Int. Conf. Ind. Technol., vol. 2018- Febru, IEEE; 2018, p. 867–72. doi:10.1109/ICIT.2018.8352292.

[90] Rai V, Robinson SA. Agent-based modeling of energy technology adoption: Empirical integration of social, behavioral, economic, and environmental

factors. Environ Model Softw 2015;70:163–77. doi:10.1016/j.envsoft.2015.04.014.

[91] Robinson SA, Stringer M, Rai V, Tondon A. GIS-Integrated Agent-Based Model of Residential Solar PV Diffusion. USAEE 2013, 2013.

[92] Rode J, Weber A. Does localized imitation drive technology adoption? A case study on rooftop photovoltaic systems in Germany. J Environ Econ Manage 2016;78:38–48. doi:10.1016/j.jeem.2016.02.001.

[93] Schwarz N, Ernst A. Agent-based modeling of the diffusion of environmental innovations - An empirical approach. Technol Forecast Soc Change 2009;76:497–511. doi:10.1016/j.techfore.2008.03.024.

[94] Müller S, Rode J. The adoption of photovoltaic systems in Wiesbaden, Germany. Econ Innov New Technol 2013;22:519–35. doi:10.1080/10438599.2013.804333.

[95] Macal CM, Graziano DJ, Ozik J. Modeling solar PV adoption: A social-behavioral agent-based framework. 2014 AAAI Fall Symp 2014;FS:24–31.

[96] Chen TD, Wang Y, Kockelman KM. Where are the electric vehicles? A spatial model for vehicle-choice count data. J Transp Geogr 2015;43:181–8. doi:10.1016/j.jtrangeo.2015.02.005.

[97] Whitaker E. UCSF Guides: Systematic Review: More about pros and cons of doing a systematic review 2017.

[98] Song F, Parekh, S, Hooper L, Loke Y, Ryder J, Sutton A, et al. Dissemination and publication of research findings: an updated review of related biases. Health Technol Assess (Rockv) 2010;14. doi:10.3310/hta14080.

[99] MacCoun RJ. BIASES IN THE INTERPRETATION AND USE OF RESEARCH RESULTS. Annu Rev Psychol 1998;49:259–87. doi:10.1146/annurev.psych.49.1.259.

[100] Bridge G, Bouzarovski S, Bradshaw M, Eyre N. Geographies of energy transition: Space, place and the low-carbon economy. Energy Policy 2013;53:331–40. doi:10.1016/j.enpol.2012.10.066.

[101] Bierkandt R, Wenz L, Willner SN, Levermann A. Acclimate—a model for economic damage propagation. Part 1: basic formulation of damage transfer within a global supply network and damage conserving dynamics. Environ Syst Decis 2014;34:507–24. doi:10.1007/s10669-014-9523-4.

[102] Kunz J. Group-Level Exploration and Exploitation: A Computer Simulation-Based Analysis. 2010:29:3 2011.

[103] Kang SG, Choi SH. Multi-Agent Based Beam Search for Real-Time Production Scheduling and Control. The Multi-Agent Based Beam Search Method, London: Springer London; 2013. doi:10.1007/978-1-4471-4576-9.

[104] Pahl-Wostl C. Agent based simulation in integrated assessment and resources management. Int. Environ. Model. Softw. Soc. (iEMSs 2002), vol. 229, 2002, p. 239–44.

[105] Samarasinghe S. Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition. Auerbach; 2016.

[106] Kök BV, Yilmaz M, Çakiroglu M, Kuloglu N, Sengür A. Neural network modeling of SBS modified bitumen produced with different methods. Fuel 2013;106:265–70. doi:10.1016/j.fuel.2012.12.073.

[107] Dematos G, Boyd MS, Kermanshahi B, Kohzadi N, Kaastra I. Feedforward versus recurrent neural networks for forecasting monthly japanese yen exchange rates. Financ Eng Japanese Mark 1996;3:59–75. doi:10.1007/BF00868008.

[108] Fowler J. Experiential learning and its facilitation. Nurse Educ Today 2008;28:427–33. doi:10.1016/j.nedt.2007.07.007.

[109] Bergsteiner H, Avery GC, Neumann R. Kolb's experiential learning model: Critique from a modelling perspective. Stud Contin Educ 2010;32:29–46.

doi:10.1080/01580370903534355.

[110] Burgess M, King N, Harris M, Lewis E. Electric vehicle drivers' reported interactions with the public: Driving stereotype change? Transp Res Part F Traffic Psychol Behav 2013;17:33–44. doi:10.1016/j.trf.2012.09.003.

[111] Machin M, Sanguesa JA, Garrido P, Martinez FJ. On the use of artificial intelligence techniques in intelligent transportation systems. 2018 IEEE Wirel. Commun. Netw. Conf. Work. WCNCW 2018, Institute of Electrical and Electronics Engineers Inc.; 2018, p. 332–7. doi:10.1109/WCNCW.2018.8369029.

[112] Raza MQ, Khosravi A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renew Sustain Energy Rev 2015;50:1352–72. doi:10.1016/j.rser.2015.04.065.

[113] Kanungo DP, Arora MK, Sarkar S, Gupta RP. A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas. Eng Geol 2006;85:347–66. doi:10.1016/j.enggeo.2006.03.004.

[114] Aghaei J, Nezhad AE, Rabiee A, Rahimi E. Contribution of Plug-in Hybrid Electric Vehicles in power system uncertainty management. Renew Sustain Energy Rev 2016;59:450–8. doi:10.1016/j.rser.2015.12.207.

[115] Haykin S. Neural networks : a comprehensive foundation. 2nd ed. 1999.

[116] Kaastra I, Boyd M. Designing a neural network for forecasting financial and economic time series. Neurocomputing 1996;10:215–36. doi:10.1016/0925-2312(95)00039-9.

[117] Werbos PJ. Backpropagation Through Time: What It Does and How to Do It. Proc IEEE 1990;78:1550–60. doi:10.1109/5.58337.

[118] Silva IN da, Hernane Spatti D, Andrade Flauzino R, Bartocci Liboni LH, dos Reis Alves SF. Multilayer Perceptron Networks. Artif. Neural Networks, Cham: Springer International Publishing; 2017, p. 55–115.

doi:10.1007/978-3-319-43162-8.

[119] Saraiva F de O, Bernardes WMS, Asada EN. A framework for classification of non-linear loads in smart grids using Artificial Neural Networks and Multi-Agent Systems. Neurocomputing 2015;170:328–38. doi:10.1016/j.neucom.2015.02.090.

[120] Liu L, Zhao Y, Chang D, Xie J, Ma Z, Sun Q, et al. Prediction of short-term PV power output and uncertainty analysis. Appl Energy 2018;228:700–11. doi:10.1016/j.apenergy.2018.06.112.

[121] Lai LYWK. Basic Learning Principles of Artificial Neural Networks. Foreign-Exchange-Rate Forecast. With Artif. Neural Networks, Springer, Boston, MA; 2007, p. 27–37. doi:10.1007/978-0-387-71720-3_2.

[122] Hill T, Marquez L, O'Connor M, Remus W. Artificial neural network models for forecasting and decision making. Int J Forecast 1994;10:5–15. doi:10.1016/0169-2070(94)90045-0.

[123] Zou J, Han Y, So SS. Overview of artificial neural networks. Artif. Neural Networks, vol. 458, Humana Press; 2008, p. 15–23. doi:10.1007/978-1-60327-101-1_2.

[124] Cherrington R, Goodship V, Longfield A, Kirwan K. The feed-in tariff in the UK: A case study focus on domestic photovoltaic systems. Renew Energy 2013;50:421–6. doi:10.1016/j.renene.2012.06.055.

[125] Burnett D, Barbour E, Harrison GP. The UK solar energy resource and the impact of climate change. Renew Energy 2014;71:333–43. doi:10.1016/j.renene.2014.05.034.

[126] Islam T, Meade N. The impact of attribute preferences on adoption timing: The case of photo-voltaic (PV) solar cells for household electricity generation. Energy Policy 2013;55:521–30. doi:10.1016/j.enpol.2012.12.041.

[127] Lin G, Zhang T. Loglinear residual tests of Moran's I autocorrelation and

their applications to Kentucky breast cancer data. Geogr Anal 2007;39:293–310. doi:10.1111/j.1538-4632.2007.00705.x.

[128] Chen Y. Spatial autocorrelation approaches to testing residuals from least squares regression. PLoS One 2016;11:e0146865. doi:10.1371/journal.pone.0146865.

[129] Simon HA. Bounded Rationality. Util. Probab., London: Palgrave Macmillan UK; 1990, p. 15–8. doi:10.1007/978-1-349-20568-4_5.

[130] Arthur WB. Inductive reasoning and bounded rationality (The El Farol problem). Am Econ Rev 1994;84:406–11. doi:10.2307/2117868.

[131] Kyriakarakos G, Piromalis DD, Dounis AI, Arvanitis KG, Papadakis G. Intelligent demand side energy management system for autonomous polygeneration microgrids. Appl Energy 2013;103:39–51. doi:10.1016/j.apenergy.2012.10.011.

[132] Bale CSE, Mccullen NJ, Foxon TJ, Rucklidge AM, Gale WF. Modeling diffusion of energy innovations on a heterogeneous social network and approaches to integration of real-world data. Complexity 2014;19:83–94. doi:10.1002/cplx.21523.

[133] Lee S-I, Lee M, Chun Y, Griffith DA. Uncertainty in the effects of the modifiable areal unit problem under different levels of spatial autocorrelation: a simulation study. Int J Geogr Inf Sci 2018:1–20. doi:10.1080/13658816.2018.1542699.

[134] Arbia G, Petrarca F. Effects of MAUP on spatial econometric models. Lett Spat Resour Sci 2011;4:173–85. doi:10.1007/s12076-011-0065-9.

[135] Chamberlin G. Temporal Disaggregation. 2010.

[136] DECC D of E and CC. Electricity System: Assessment of Future Challenges. 2012.

[137] Market Research Future. Electric Car Rental Market 2019 Industry Analysis

By Size, Share, Demand, Trends, Business Growth, Opportunity, Statistics, Regional Analysis With Global Forecast To 2025 - MarketWatch 2019. https://www.abnewswire.com/pressreleases/electric-car-rental-market-2019-industry-analysis-by-size-share-demand-trends-business-growth-opportunity-statistics-regional-analysis-with-global-forecast-to-2025_441075.html (accessed April 9, 2020).

[138] Roberts G. Enterprise doubles hybrid and electric vehicles on fleet | Fleet Industry News 2018. https://www.fleetnews.co.uk/news/fleet-industry-news/2018/10/18/enterprise-doubles-hybrid-and-electric-vehicles-on-fleet (accessed April 9, 2020).

[139] Roberts G. Nexus launches electric vehicle rental platform | Fleet Industry News 2020. https://www.fleetnews.co.uk/news/fleet-industry-news/2020/02/10/nexus-launches-electric-vehicle-rental-platform (accessed April 9, 2020).

[140] Langbroek JHM, Cebecauer M, Malmsten J, Franklin JP, Susilo YO, Georén P. Electric vehicle rental and electric vehicle adoption. Res Transp Econ 2019;73:72–82. doi:10.1016/j.retrec.2019.02.002.

[141] McKenna E, Pless J, Darby SJ. Solar photovoltaic self-consumption in the UK residential sector: New estimates from a smart grid demonstration project. Energy Policy 2018;118:482–91. doi:10.1016/j.enpol.2018.04.006.

[142] Brand C, Anable J, Tran M. Accelerating the transformation to a low carbon passenger transport system: The role of car purchase taxes, feebates, road taxes and scrappage incentives in the UK. Transp Res Part A Policy Pract 2013;49:132–48. doi:10.1016/j.tra.2013.01.010.

[143] Dept of Transport. Low-emission vehicles eligible for a plug-in grant. GovUk Website 2017:2. https://www.gov.uk/plug-in-car-van-grants (accessed April 8, 2020).

[144] Lanzini P, Thøgersen J. Behavioural spillover in the environmental domain:

An intervention study. J Environ Psychol 2014;40:381–90. doi:10.1016/j.jenvp.2014.09.006.

[145] Yohanis YG, Mondol JD, Wright A, Norton B. Real-life energy use in the UK: How occupancy and dwelling characteristics affect domestic electricity use. Energy Build 2008;40:1053–9. doi:10.1016/j.enbuild.2007.09.001.

[146] Mwasilu F, Justo JJ, Kim E-K, Do TD, Jung J-W. Electric vehicles and smart grid interaction: A review on vehicle to grid and renewable energy sources integration. Renew Sustain Energy Rev 2014;34:501–16. doi:10.1016/j.rser.2014.03.031.

[147] Green RC, Wang L, Alam M. The impact of plug-in hybrid electric vehicles on distribution networks: A review and outlook. Renew Sustain Energy Rev 2011;15:544–53. doi:10.1016/j.rser.2010.08.015.

[148] Hoarau Q, Perez Y. Interactions between electric mobility and photovoltaic generation: A review. Renew Sustain Energy Rev 2018;94:510–22. doi:10.1016/j.rser.2018.06.039.

[149] Rai A. Explainable AI: from black box to glass box. J Acad Mark Sci 2020;48:137–41. doi:10.1007/s11747-019-00710-5.

[150] Benitez JM, Castro JL, Requena I. Are artificial neural networks black boxes? IEEE Trans Neural Networks 1997;8:1156–64. doi:10.1109/72.623216.

[151] Olden JD, Jackson DA. Illuminating the “black box”: A randomization approach for understanding variable contributions in artificial neural networks. Ecol Modell 2002;154:135–50. doi:10.1016/S0304-3800(02)00064-9.

[152] Fan-Yin Tzeng, Kwan-Liu Ma. Opening the Black Box - Data Driven Visualization of Neural Networks, Institute of Electrical and Electronics Engineers (IEEE); 2006, p. 383–90. doi:10.1109/visual.2005.1532820.

[153] Moglia M, Cook S, McGregor J. A review of Agent-Based Modelling of

technology diffusion with special reference to residential energy efficiency. Sustain Cities Soc 2017;31:173–82. doi:10.1016/j.scs.2017.03.006.

[154] Nykamp S, Molderink A, Hurink JL, Smit GJM. Statistics for PV, wind and biomass generators and their impact on distribution grid planning. Energy 2012;45:924–32. doi:10.1016/j.energy.2012.06.067.

[155] McCullen N, Rucklidge A, Bale C, Foxon T, Gale W. Multiparameter Models of Innovation Diffusion on Complex Networks. SIAM J Appl Dyn Syst 2013.

[156] Lopez-Rodriguez I, Hernandez-Tejera M. Infrastructure based on supernodes and software agents for the implementation of energy markets in demand-response programs. Appl Energy 2015;158:1–11. doi:10.1016/j.apenergy.2015.08.039.

[157] Mokhtari G, Nourbakhsh G, Zare F, Ghosh A. Overvoltage prevention in LV smart grid using customer resources coordination. Energy Build 2013;61:387–95. doi:10.1016/j.enbuild.2013.02.015.

[158] National Grid. UK Future Energy Escenarios 2017.

[159] Bhatti AR, Salam Z, Aziz MJBA, Yee KP, Ashique RH. Electric vehicles charging using photovoltaic: Status and technological review. Renew Sustain Energy Rev 2016;54:34–47. doi:10.1016/j.rser.2015.09.091.

[160] Wang W, Hong T, Xu X, Chen J, Liu Z, Xu N. Forecasting district-scale energy dynamics through integrating building network and long short-term memory learning algorithm. Appl Energy 2019;248:217–30. doi:10.1016/j.apenergy.2019.04.085.

[161] Jana RK, Ghosh I, Sanyal MK. A granular deep learning approach for predicting energy consumption. Appl Soft Comput J 2020;89:106091. doi:10.1016/j.asoc.2020.106091.

[162] Konar A. Competitive Learning Using Neural Nets. Comput. Intell., Berlin, Heidelberg: Springer Berlin Heidelberg; 2005, p. 267–93. doi:10.1007/3-540-27335-2_10.

[163] Abas AR. Adaptive competitive learning neural networks. Egypt Informatics J 2013;14:183–94. doi:10.1016/j.eij.2013.08.001.

[164] Peeters A, Zude M, Käthner J, Ünlü M, Kanber R, Hetzroni A, et al. Getis–Ord's hot- and cold-spot statistics as a basis for multivariate spatial clustering of orchard tree data. Comput Electron Agric 2015;111:140–50. doi:10.1016/J.COMPAG.2014.12.011.

[165] Schilling J-G, Kingsley RH, Devine JD. Galapagos Hot Spot-Spreading Center System: 1. Spatial petrological and geochemical variations (83°W-101°W). J Geophys Res Solid Earth 1982;87:5593–610. doi:10.1029/JB087iB07p05593.

[166] Tibshirani R, Wang P. Spatial smoothing and hot spot detection for CGH data using the fused lasso. Biostatistics 2008;9:18–29. doi:10.1093/biostatistics/kxm013.

[167] Nelson TA, Boots B. Detecting spatial hot spots in landscape ecology. Ecography (Cop) 2008;31:556–66. doi:10.1111/j.0906-7590.2008.05548.x.

APPENDICES

Appendix 1 Full list of articles by research question and Snowballing

Table 13. Reviwed articles by research question or snowball criteria

Source	Author	Article	Journal
RQ1	[25]	An agent-based electric vehicle ecosystem model: San Francisco case study	Transport Policy
	[27]	Simulating the household plug-in hybrid electric vehicle distribution and its electric distribution network impacts.	Transportation Research Part D: Transport and Environment
	[14]	Modeling photovoltaic diffusion: an analysis of geospatial datasets.	Environmental Research Letters
	[16]	Spatial patterns of solar photovoltaic system adoption: The influence of neighbors and the built environment.	Journal of Economic Geography
	[86]	Predicting Rooftop Solar Adoption Using Agent-Based Modeling.	Energy Market Prediction: Papers from the 2014 AAA Fall Symposium
	[78]	A spatially explicit agent-based model of the diffusion of green electricity: Model setup and retrodictive validation. Adv. Intell. Syst. Comput.	Advances in Intelligent Systems and Computing
	[12]	Third Party-Owned PV Systems: Understanding Market Diffusion with Geospatial Tools	Energy Market Prediction: Papers from the 2014 AAAI Fall Symposium Third

	[94]	The adoption of photovoltaic systems in Wiesbaden, Germany.	Economics of Innovation and New Technology
	[8]	Regional distribution of photovoltaic deployment in the UK and its determinants: A spatial econometric approach. Energy Econ	Energy Economics
	[20]	Social Effects in the Diffusion of Solar Photovoltaic Technology in the UK.	Economics of Innovation and New Technology
	[92]	Does localized imitation drive technology adoption? A case study on rooftop photovoltaic systems in Germany.	Journal of Economics of Innovation and New Technology
	[18]	Beyond the sun—Socioeconomic drivers of the adoption of small-scale photovoltaic installations in Germany.	Energy Research and Social Science
RQ2	[68]	Understanding solar PV and battery adoption in Ontario.	Proceedings of the Seventh International Conference on Future Energy Systems - e-Energy '16
	[19]	Estimating spatial interdependence in automobile type choice with survey data. Transp Res Part A Policy Pract	
	[132]	Modeling diffusion of energy innovations on a heterogeneous social network and approaches to integration of real-world data. Complexity	Complexity

	[85]	Hybrid Electric Vehicle Ownership and Fuel Economy Across Texas. Transp Res Rec J Transp Res Board	Transportation Research Part A: Policy and Practice
	[96]	Where are the electric vehicles? A spatial model for vehicle-choice count data.	Journal of Transport Geography
	[15]	Household dynamics of technology adoption: A spatial econometric analysis of residential solar photovoltaic (PV) systems in Germany.	Energy Research and Social Science
	[41]	Spatial diffusion of electric vehicles in the German metropolitan region of Stuttgart.	ERSA conference papers
	[26]	Consumer preferences and the influence of networks in electric vehicle diffusion: An agent-based microsimulation in Ireland. Energy Res Soc Sci	Energy Research and Social Science
	[46]	Spatial Effects in Energy-Efficient Residential HVAC Technology Adoption.	Environment and Behavior
	[49]	Development of an agent-based model for regional market penetration projections of electric vehicles in the United States.	Energy
	[90]	Agent-based modeling of energy technology adoption: Empirical integration of social, behavioral, economic, and environmental factors.	Environmental Modelling & Software
	[32]	An agent-based decision support system for electric vehicle charging infrastructure deployment. Power Propuls. Conf., IEEE	2011 IEEE Vehicle Power and Propulsion Conference
Snowballing	[87]	Modeling innovation diffusion for renewable energy technologies in city neighborhoods	2018 9th International Renewable Energy Congress, IREC 2018

	[43]	Energy and complexity: New ways forward.	Applied Energy
	[51]	Heterogeneity in the adoption of photovoltaic systems in Flanders.	Energy Economics
	[80]	The spatial distribution of hybrid electric vehicles in a sprawled mid-size Canadian city: Evidence from Windsor, Canada.	Journal of Transport Geography
	[48]	USING NATIONAL SURVEY RESPONDENTS AS CONSUMERS IN AN AGENT-BASED MODEL OF PLUG-IN HYBRID VEHICLE ADOPTION.	IEEE Access
	[33]	An agent-based model to study market penetration of plug-in hybrid electric vehicles.	Energy Policy
	[77]	A dynamic and spatially explicit psychological model of the diffusion of green electricity across Germany.	Journal of Environmental Psychology
	[75]	An empirically grounded model of green electricity adoption in Germany: Calibration, validation and insights into patterns of diffusion.	JASSS
	[17]	Influence of local environmental, social, economic and political variables on the spatial distribution of residential solar PV arrays across the United States.	Energy Policy
	[47]	The diffusion of domestic energy efficiency policies: A spatial perspective.	Energy Policy
	[29]	The spatial pattern of demand in the early market for electric vehicles: Evidence from the United Kingdom.	Journal of Transport Geography
	[76]	Determinants of spatio-temporal patterns of energy technology adoption: An agent-based modeling approach.	Applied Energy

	[91]	GIS-Integrated Agent-Based Model of Residential Solar PV Diffusion	USAEE Online Proceedings 2013
	[93]	Agent-based modeling of the diffusion of environmental innovations - An empirical approach.	Technological Forecasting and Social Change
	[50]	Adoption and diffusion of heating systems in Norway: Coupling agent-based modeling with empirical research.	Environmental Innovation and Societal Transitions
	[88]	Spatio-Temporal Analysis and Forecasting of Distributed PV Systems Diffusion: A Case Study of Shanghai Using a Data-Driven Approach.	IEEE Access
	[89]	A data-driven approach to forecasting the distribution of distributed photovoltaic systems.	Proceedings of the IEEE International Conference on Industrial Technology

Appendix 2 Data analysis

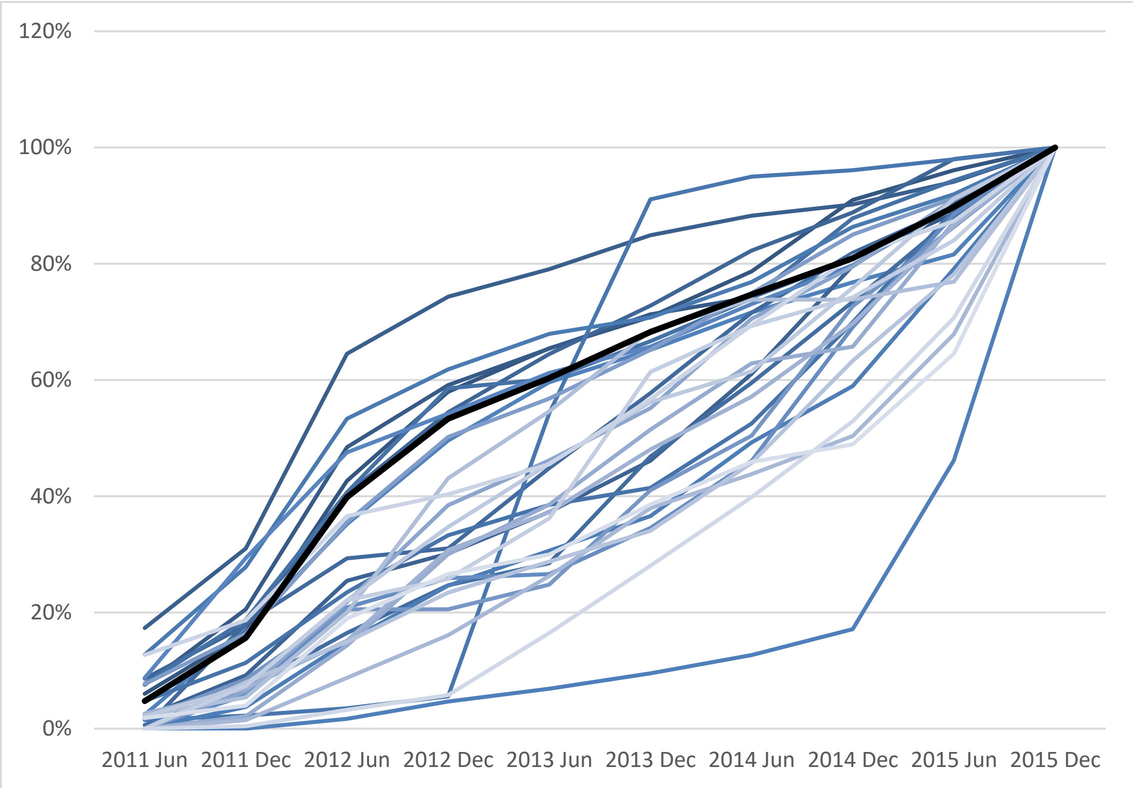


Figure 70. Historical PV data at MSOA level and 6-month basis for Birmingham city.

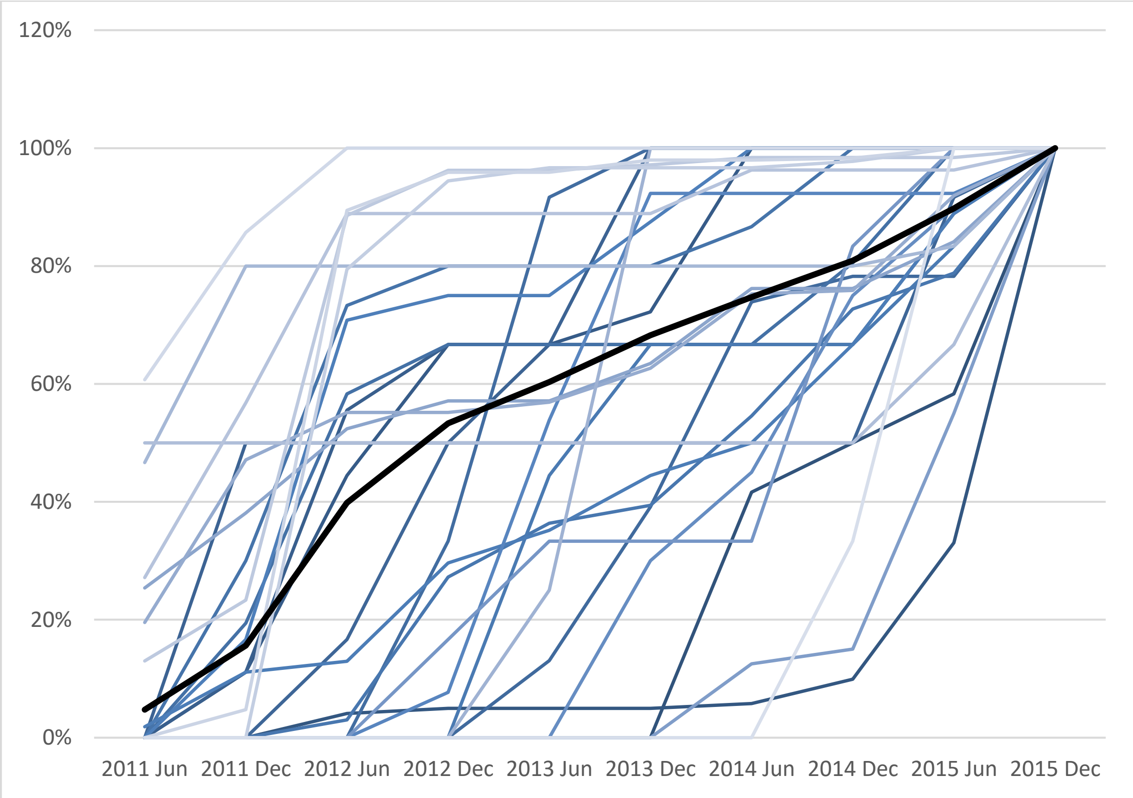


Figure 71. Historical PV data at LSOA level and 6-month basis for Birmingham city.

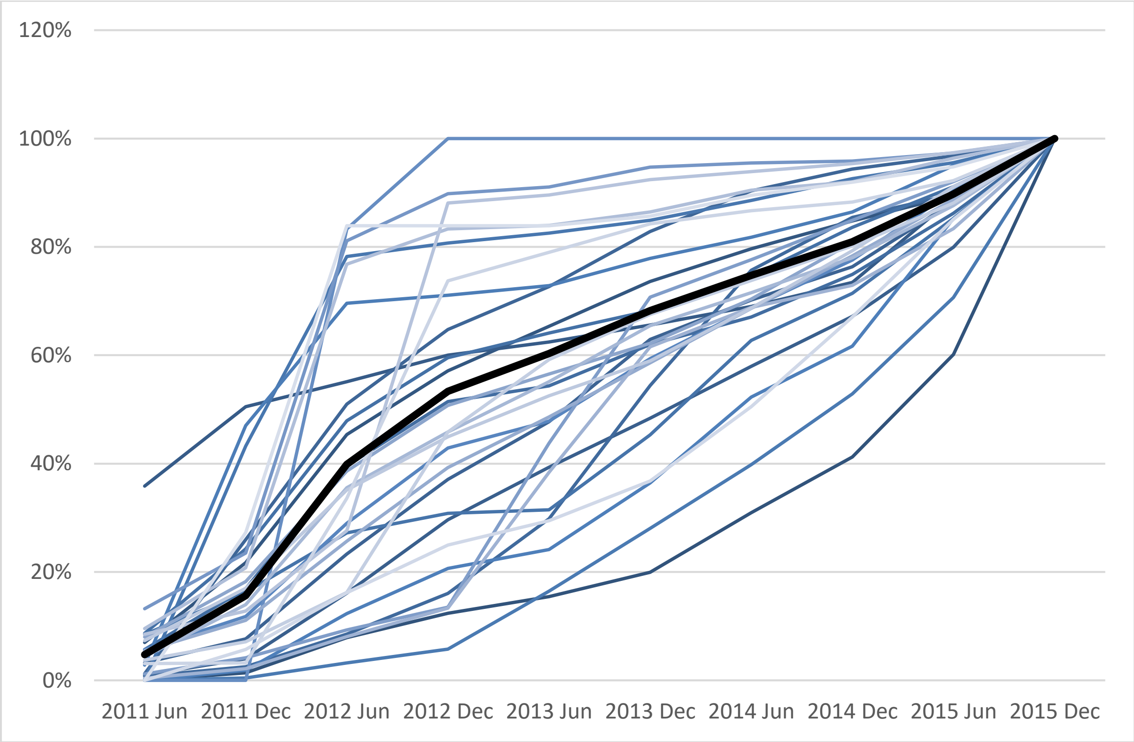


Figure 72. Historical PV data at PC level and 6-month basis for Birmingham city.

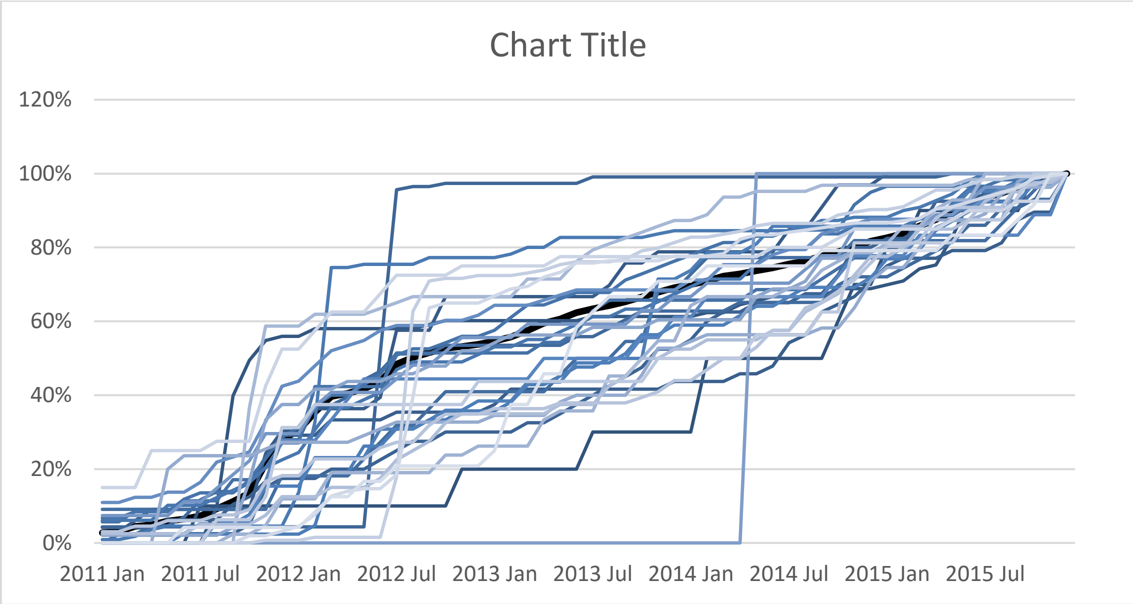


Figure 73. Historical PV data at MSOA level and monthly basis for Birmingham city.

Appendix 3 Simulation algorithm

Algorithm 1 The PV adoption process

Initialisation

```
for each agent PC in Birmingham do
    PC.location ← actual population centroid
    PC.PV_t ← PV installation dataset
    function AGENT_NEIGHBOURS()
        for each PC in agent_aux.Neighbours do
            agent_aux.calculateDistance()
            agent_aux.calculatePVNeighbourhood()
        end for
    end function
    function AGENT_ANN()
        for each PC in Birmingham do
            agent_aux.ANN(weight) ← random_between(0,1)
        end for
    end function
end for
```

Training

```
function TRAIN()
    for each agent PC in Birmingham do
        PC.train()
        PC.estimationError ← Mean absolute percentage error
    end for
end function
```

Forecasting

```
for each PC in Birmingham do
    PC.PV_{t+1} ← PC.forecastPV()
    PC. calculatePVNeighbourhood()
end for
```

Appendix 4 Training Algorithm

Algorithm 2 Backpropagation learning algorithm

Initialisation

```
for each agent PC in Birmingham do
    PC.data ← historical PV installation data by month
    for each month m in PC.data do
        FORWARDS PASS
        Starting from the input layer, use each activation function to compute
        the outputs
        Use the synaptic weights to pass the outputs from each layer to the
        following one
        Calculate the network output and the error of estimation
        Check for stop condition
        BACKWARDS PASS
        Beginning from the last layer, compute the derivates of the output
        layer's function with respect to the estimation error
        Compute the derivates of each other hidden layer with respect to the
        previous layer neurons function
        Calculate the adjustment coefficient for each synaptic weight
        considering the previous layer neurons function
    end for
end for
```

Appendix 5 Variables used in the modelling approaches by broaden categories.

Table 14. Variables used in the modelling approaches

Ref.	Variable	Description
Adepetu and Keshav, 2016	Social Utility	Social threshold
	Energy cost	electricity consumption
	Income	Assigned budget for purchasing
Adepetu, Keshav and Arya, 2016	Environmental awareness	Greenness (fitted parameter)
	Social Utility	Social threshold
	Age	Age
	Economics	Discount rate
	Energy cost	Cost of electricity
	Energy cost	Cost of gas
	Income	Income

Ref.	**Variable**	**Description**
	Policies	Incentives
	Available Infrastructure	Charging point
	Household characteristics	House location
	Vehicle characteristics	Vehicle type
Adjemian, Cynthia Lin and Williams, 2010	Age	Mean population age
	Education	Population with a college degree (%)
	Ethnicity	Asian population %
	Ethnicity	Latino population %
	Gender	Female population %
	Household characteristics	Family size
	Household economics	Median income
	Marital status	Married population (%)

Ref.	Variable	Description
	Population density	Population density
Balta-Ozkan, Yildirim and Connor, 2015	Environmental awareness	Share of green party voters per region
	Age	Mean population age
	Education	Education level proxy (QL2)
	Electricity consumption	Electricity sales
	Household characteristics	Households per zone (NUTS3)
	Household economics	Gross domestic household income
	Household physical characteristics	House size
	Household physical characteristics	House type
	Solar energy	Solar irradiation (kWh/m2)
Bansal, Kockelman and Wang, 2015b	Age	Median age

Ref.	Variable	Description
	Age	Population under 16 %
	Education	Population with a bachelor degree %
	Ethnicity	African-American population %
	Gender	Male population %
	Income	Families in poverty %
	Income	Population with high income %
	Population density	Population density
	Household characteristics	Mean house size
Chen, Wang and Kockelman, 2015	Household characteristics	House density
	Household characteristics	Household workers
	Income	Mean income
	Income	Population with income >$35K %

Ref.	**Variable**	**Description**
	Population density	Population density
	Workers density	Jobs per acre
	Workers density	Resident workers density
Davidson et al., 2014	Environmental awareness	Total HEV registered
	Education	Population with a college degree (%)
	Education	Population with a postgraduate (master and/or phd) degree %)
	Ethnicity	White population %
	Homeownership	Foreclosure risk score
	Household characteristics	Family size
	Household economics	Mortgage vs Income ratio (<40) %
	Household physical characteristics	Heating (wood)

Ref.	**Variable**	**Description**
	Household physical characteristics	House value
	Household physical characteristics	Number of rooms
	Homeownership	Owned household with a mortgage %
De Groote, Pepermans and Verboven, 2016	Environmental awareness	Surveyed Proxy (roof insulation)
	Political tendency	Left party votes %
	Age	Age
	Education	Population with a College degree %
	Ethnicity	Foreigner population %
	Homeownership	Homeownership %
	Household characteristics	Family size
	Income	Mean income

Ref.	**Variable**	**Description**
	Income	Variation (std deviation)
	Policies	Subsidies
	Household characteristics	House size
	Household characteristics	House type
	Household characteristics	House value
Eppstein et al., 2011a	Social Utility	Social threshold (EV market share %)
	Income	Income
	Household characteristics	House location
	Vehicle characteristics	Car age
Graziano and Gillingham, 2015	Political tendency	Democrat voters (%)
	Political tendency	Minor parties (%)
	Age	Median age of older population (5%)

Ref.	Variable	Description
	Age	Median population age
	Electricity consumption	Electricity cost
	Ethnicity	Asian population %
	Ethnicity	Black population %
	Ethnicity	White population %
	Homeownership	Rented house %
	Household characteristics	Households per concentred radius
	Household economics	Median income
	Population density	Neighbours per concentred radius
Langheim, 2014b	Ethnicity	Native born %
	Household characteristics	Median age
	Household economics	Mean income

Ref.	**Variable**	**Description**
	Household economics	Median income
McCoy and Lyons, 2014	Environmental awareness	Environmental utility
	Income	Income utility (fitted parameter)
Noonan, Hsieh and Matisoff, 2013	Homeownership	30-years fixed mortgage %
	Homeownership	Vacant house ratio in the area%
	Household economics	Median income
	Population density	Population density
	Household physical characteristics	House age
	Household physical characteristics	House size
	Household physical characteristics	Median house value

Ref.	Variable	Description
	Household physical characteristics	Rehabilitated buildings %
Schaffer and Brun, 2015	Homeownership	Owned households per sqkm
	Household characteristics	Buildings per sqkm
	Household economics	Gross regional product per capita
	Solar energy	Installed capacity (kWp/sqm)
Sweda and Klabjan, 2011	Environmental awareness	Greenness
	Income	Income
	Household characteristics	Household location
	Vehicle characteristics	Car age
	Vehicle characteristics	Car fuel type
	Vehicle characteristics	Stated preferred vehicle

Appendix 6 Bass model estimation

The Bass model is used to estimate the S-curve of each area, using Ordinary Least squares to estimates the parameter of the model. The model assumes 2050 as the horizon for total uptake of the PVs, and the number of residential building owned by the householder (per area) as the number of potential adopters. Then, following [40,41] the Bass model is defined as:

$$S(t) = (p + q * \frac{Y(t)}{m})(m - (Y(t))$$

Where:

S(t) is the number of new PV installations

p is the coefficient of innovation

q is the coefficient of imitation

Y(t) is the number of adopters

m is potential adopter

Figure 74 shows the different S-curves for the PCs in Birmingham city, the line in black represent the estimation for the Total number of PVs in Birmingham.

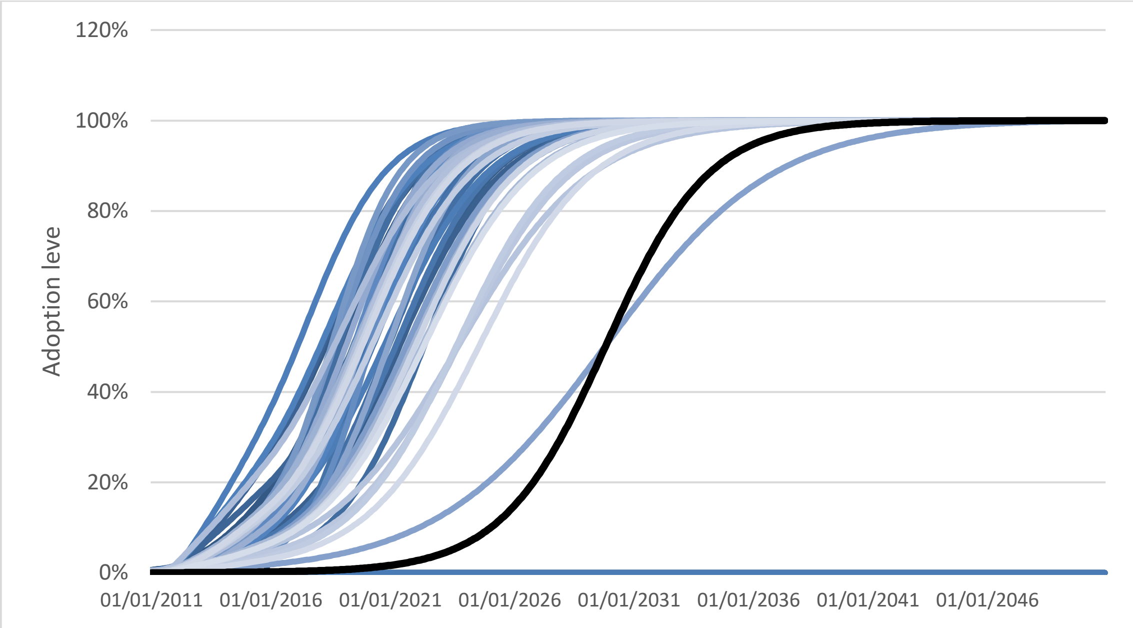


Figure 74. Bass model's estimations for the PV uptake in the Birmingham's PCs.

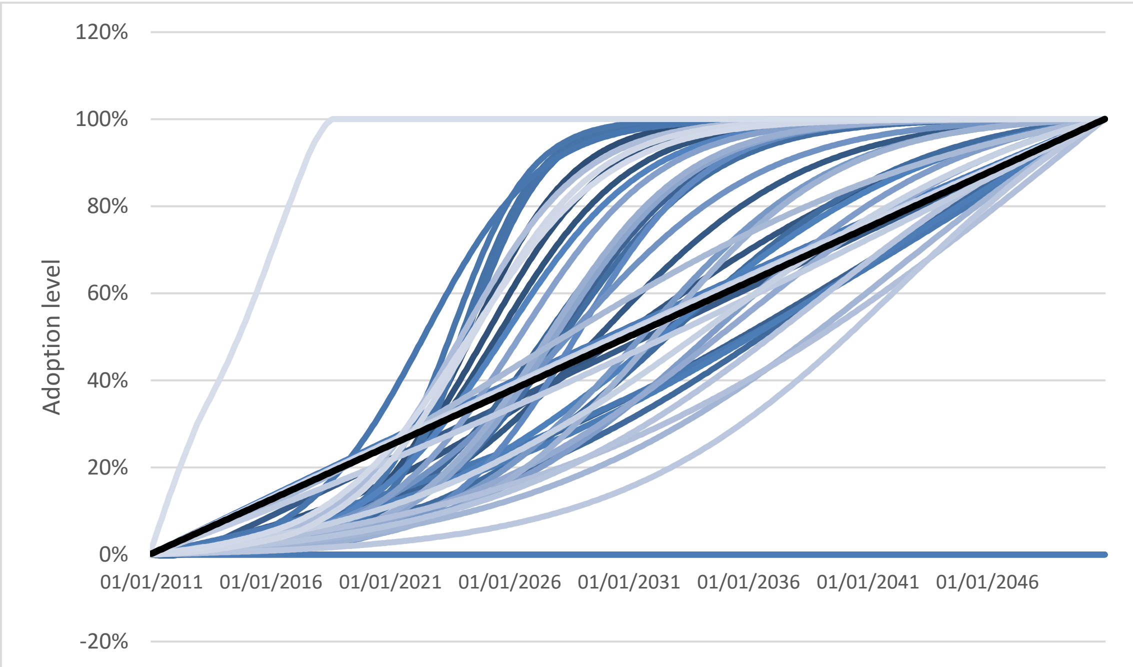
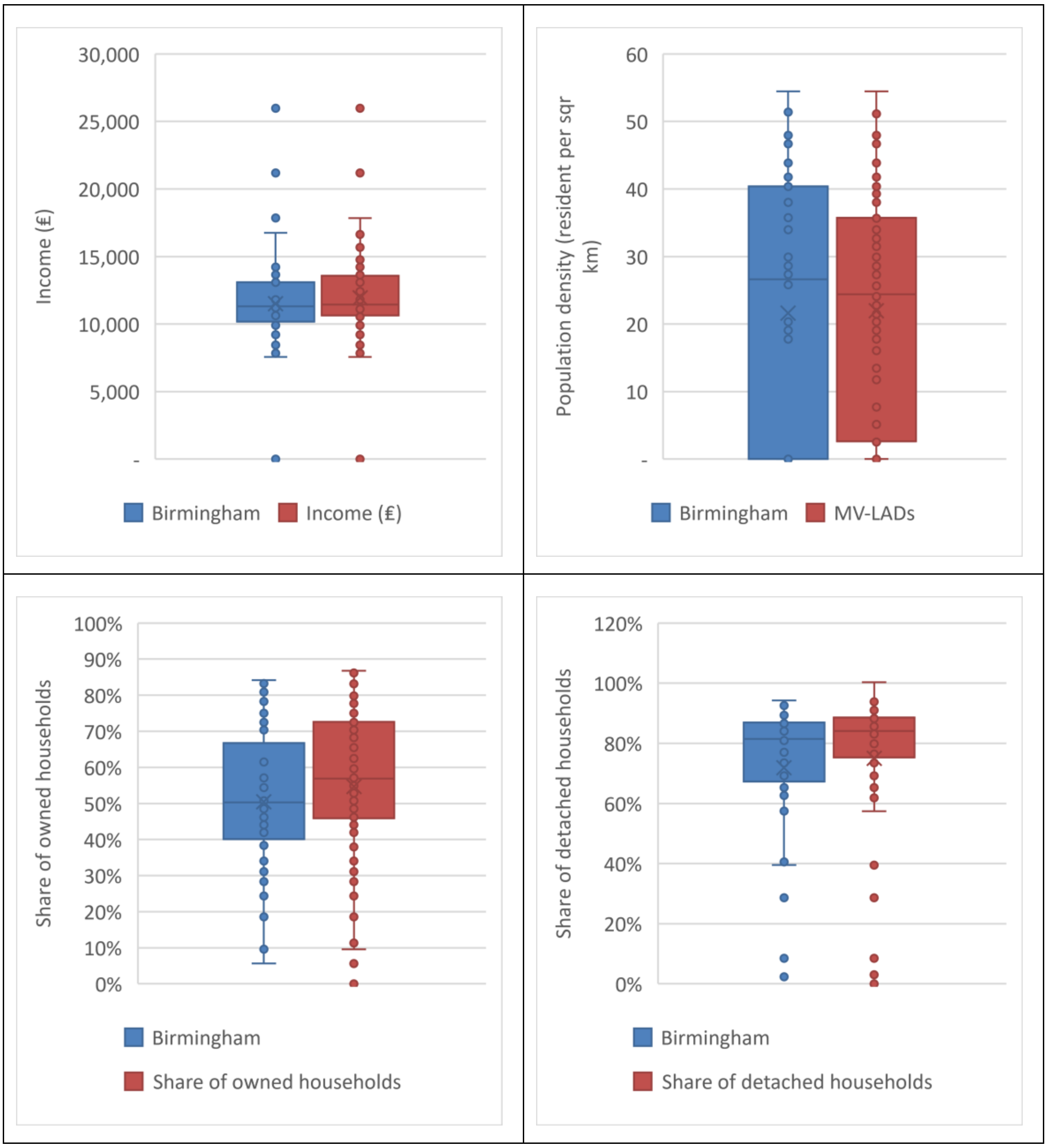
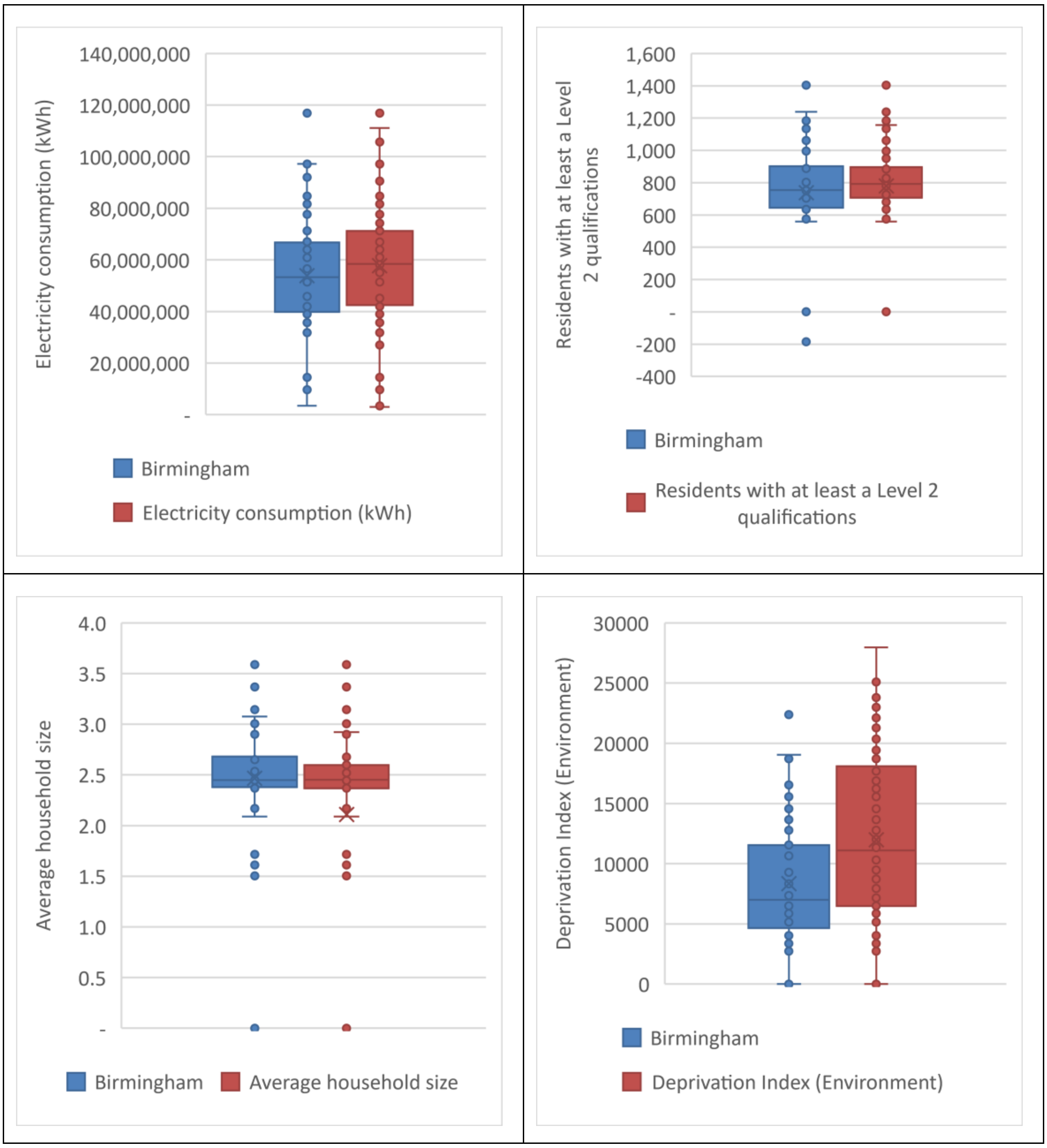


Figure 75. Bass model's estimations for the EV uptake in the Birmingham's PCs.

Appendix 7 Summary of the potential socioeconomic variables




Electricity consumption (kWh)
140,000,000
120,000,000
100,000,000
80,000,000
60,000,000
40,000,000
20,000,000
-
Birmingham
Electricity consumption (kWh)
Residents with at least a Level 2 qualifications
1,600
1,400
1,200
1,000
800
600
400
200
-
-200
-400
Birmingham
Residents with at least a Level 2 qualifications
Average household size
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
-
Birmingham
Average household size
Deprivation Index (Environment)
30000
25000
20000
15000
10000
5000
0
Birmingham
Deprivation Index (Environment)