CRANFIELD UNIVERSITY

Khoi Quang Ngo (Lucas)

Road Traffic Emission Dispersion Modelling: An Application to Hanoi and Ho Chi Minh City Using ADMS

School of Water, Energy and Environment

MSc by Research in Environment and Agrifood
Academic Year: 2018 - 2020

Supervisor: Dr Iq Mead
Associate Supervisor: Prof Neil Harris
August 2020

CRANFIELD UNIVERSITY

School of Water, Energy and Environment

MSc by Research in Environment and Agrifood

Academic Year 2018 - 2020

Khoi Quang Ngo (Lucas)

Road Traffic Emission Dispersion Modelling: An Application to Hanoi and Ho Chi Minh City Using ADMS

Supervisor: Dr Iq Mead
Associate Supervisor: Prof Neil Harris
August 2020

This thesis is submitted in partial fulfilment of the requirements for the degree of MSc by Research in Environment and Agrifood

ABSTRACT

Urban air quality in Vietnam has become a pressing matter that require immediate attention to ensure a sustainable development. However due to the overreliance on in-situ observations, which only measure the end result, there is limited understanding of the connection between pollution sources and concentrations. This in turn hinders the effectiveness of environmental law enforcement and management. Since road traffic is widely regarded as the main polluter, attempts have been made to adopt atmospheric dispersion models to traffic emission in Vietnam. Most however, suggest that due to input data scarcity, model applications are limited. This work therefore employed ADMS, an advanced dispersion model that is highly adaptable to produce a full mapping of road traffic derived emission for Hanoi and HCMC, i.e. Vietnam’s 2 most populated cities. Also, a modelling framework, which exploits existing, quality traffic data to generate suitable model inputs, was developed. With this framework, a detailed GIS-based road network dataset that contains road parameters, vehicle count and travel-condition-depending emission factor was produced. Carbon Monoxide was modelled as a pilot pollutant species. Resulted concentrations show an overall moderate positive correlation with observations (r = 0.4). Inadequate information on background pollution however prevents in-depth model validation to be conducted. In overall, this work demonstrates the compatibility of ADMS with the circumstance of Vietnam. Combined with an improved data processing framework, applications of dispersion model in developing countries can be greatly expanded.

Keywords:

ADMS; Dispersion Model; Traffic Emission; Urban Air Pollution; Vietnam

ACKNOWLEDGEMENTS

I would like to express profound thanks to my supervisors, Dr Iq Mead and Professor Neil Harris for their patience and guidance. Their insightful suggestions and unwavering supports have been a tremendous help to the thesis as well as to me as a young researcher.

Most sincere appreciations are extended to Professor Hoang Anh Le (VNU University of Science, VNU), Professor To Thi Hien (VNUHCM – University of Science) and Professor Ho Quoc Bang (VNUHCM) for their help in sourcing/ sharing data and their valuable comments.

Lastly, the highest gratitude is to my family and friends. Their love, care and understanding have provided me the energy and positivity that fuelled this work.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

LIST OF ACRONYMS AND ABBREVIATIONS

ADMS	Advanced Dispersion Modelling System
CEM	Center for Environmental Monitoring
CFD	Computational Fluid Dynamics
EF	Emission Factor
EFT	Emission Factors Toolkit
FB	Fractional Bias
GFS	Global Forecast System
GIS	Geographic Information System
gROADS	Global Roads Open Access Data Set
HCMC	Ho Chi Minh City
ISD	Integrated Surface Database
JICA	Japan International Cooperation Agency
MC	Motorcycle
US EPA	United States Environmental Protection Agency
MONRE	Minister of Natural Resources and Environment
NCEI	National Centers for Environmental Information
NCEP	National Centers for Environmental Prediction
NMSE	Normalised Mean Square Error
NOAA	National Oceanic and Atmospheric Administration
OCHA	Office for the Coordination of Human Affairs
OSM	Open Street Map
PM	Particulate matter
ROAP	Regional Office for Asia and the Pacific
SEDAC	Socioeconomic Data and Applications Center
UN	United Nations
VEA	Vietnam Environment Administration

1 Introduction

1.1 Motivation

For the last two centuries, thanks to global industrialisation (including in the agricultural sector which is mainly supported by the Haber-Bosch process), the human species has been able to carry its exponential growth and the global population has increased by a factor of more than 7(Roser, 2019). As a result there is an escalating pressure on the Earth's environment and natural resources. Among other concerns namely climate change and conservation, urbanisation and urban pollution have been continually discussed by both the media and major reports. According to the United Nations (UN), there were globally only 15 cities with over 1 million inhabitants in 1900, whereas by 2018, this number has grown to 548 (United Nations, 2019). 55% of the world's population currently lives in urban areas, a proportion that is predicted to rise to 68% by 2050 (United Nations, 2019). Over 90% of this increase is in the developing world. Sustainable development has been repeatedly emphasised (UN-Habitat, 2006) but is not necessarily implemented.

Over the last decade, Vietnam has entered a crucial period of urbanisation which corresponds to its economic development (World Bank, 2011). It exhibits the highest urbanisation rate among countries of the Southeast Asia region, i.e. 2.8% physical expansion and 3.4 % population growth (Chu & Thi, 2017). However, similar to other emerging economies namely China, India and Brazil, Vietnam has difficulties matching wealth growth with infrastructure development (e.g. roads, housings, public amenities and transportations) (Pojani & Stead, 2017). As a result, motorised private transport is preferred because it is unsafe, unpleasant or impossible to travel otherwise (Pucher, 2017); (Table 1). In detail, Vietnam's motorcycle (MC) fleet grew 400% between 1996 and 2006; the current annual growth rate is 16% for MC and 20% for personal cars (Huynh & Gomez-Ibanez, 2017); (Trang, et al., 2015). Rapid motorisation, with its consumption of fossil fuel and well-associated emission of atmospheric pollutants, is thus presented as a major subsequent issue of urbanisation.

Table 1. Key social-economic characteristics of selected countries (Pojani & Stead, 2017).

Country	Total population, 2013 (million)	Urban population, 2013 (%)	Urban population growth 1990-2013 (%)	Passenger cars, 2011 (per 1000 people)	Passenger cars increase 2000-2011 (per 1000 people)	Air pollution, PM_{10} levels (μg/m^3)
Brazil	200	85	11	179	45	36
China	1400	53	27	54	47	82
India	1300	32	6	11	5	27
Indonesia	250	52	22	39	25	100
Mexico	122	79	7	195	93	115
Nigeria	174	51	16	31	21	46
Vietnam	90	32	12	14	N/a	65

As for adverse health impacts of air pollution, it is suggested that human exposure to road traffic derived emission causes approximately 0.8 million premature deaths per annum globally (Amorim, et al., 2014). In Vietnam, over 60,000 deaths in 2016 were linked to air pollution (World Health Organization, n.d.); respiratory illness also causes an annual loss of 20 million USD for Hanoi and 50 million for Ho Chi Minh City (HCMC).

The Minister of Natural Resources and Environment (MONRE) has confirmed high levels of air pollution in urban areas of Vietnam, making it of higher priority compared to water and soil contamination (MONRE, 2013). Particulate matter (PM) has been identified as the most concerning threat. In 2013, Hanoi experienced an average value of total suspended particle (TSP) of over 900 μg/m^3, i.e. 4.5 times the legal threshold. Similarly, events of haze and smog have become increasingly common. Decision makers have put an emphasis on vehicle emissions since it is suggested that vehicles are responsible for 70% of all urban air pollution in Vietnam (MONRE, 2016). One key evidence of this claim is that pollutant concentrations have shown very similar diurnal patterns compared to traffic pattern (Figure 1). With the number of vehicles still rapidly increasing, finding a solution is a governmental priority.

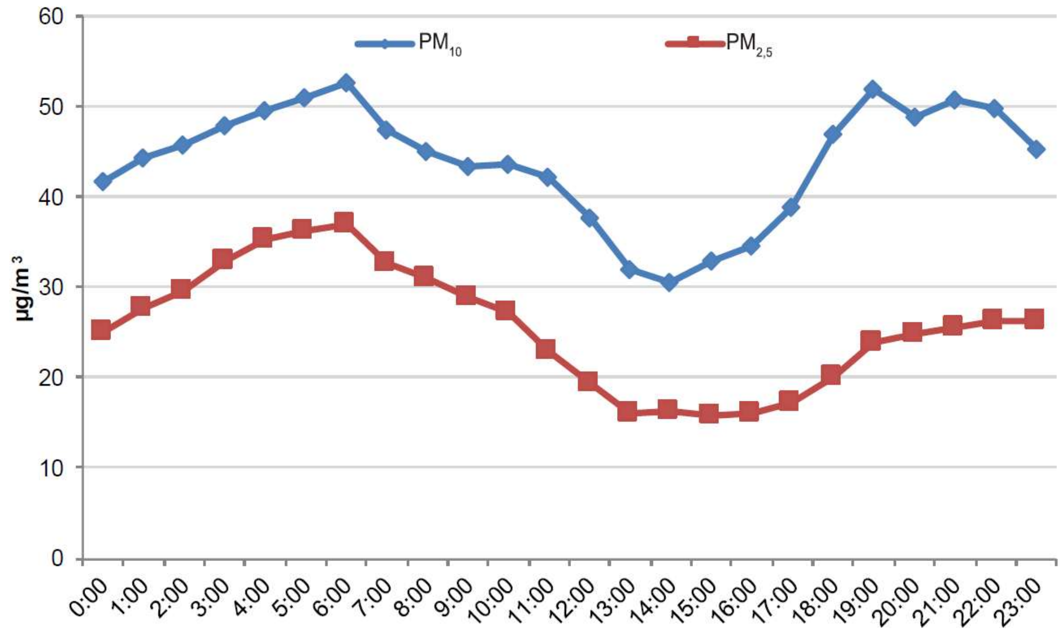


Figure 1. Representative PM_{10} and $PM_{2.5}$ hourly record in Hanoi (MONRE, 2016).

Even though ‘exposure to air pollution is now an almost inescapable part of urban life throughout the world’, residents of cities in developing countries are at higher risk due to the lack of pollution control measures (World Health Organization, 1992). Vietnam relies on a limited number of air quality monitoring stations, these stations used to be operated by different governmental units and had difficulties in communication. Prior to 2011, a monitoring network therefore could not be established effectively, this essentially isolated stations to their limited location, further restricting the use of observations (Hanoi People's Committee , 2012).

Although the monitoring network has been lately improved, recent incidents still proved that Vietnam’s air quality management system needs other tools apart from stationary observations, namely dispersion modelling. This is because dispersion modelling can provide information on the origin, source contribution and dispersion of pollutants. For example, during the late 2019 haze event in HCMC, measurements alone failed to find the pollution source. As such, the Indonesia forest fire was initially announced to the public, however, with insufficient evident, other causes were later suggested (Samuel, 2019). Dispersion modelling could have been employed in this case to simulate each potential cause and to find the most plausible explanation by comparing against observations. Similarly, a MC ban was proposed in Hanoi as an attempt to reduce congestion and pollution by 2030. This would require a substantial investment in alternative means of transportation, namely 8 rail routes, 200 bus routes and thousands of other public vehicles (Phong, 2019). However, the projected environmental impact could yet be explicitly valued. Model can therefore be useful to test the scenarios where MC is replaced by cars and buses.

Dispersion modelling has been widely uses by developed countries to simulate the physical and chemical processes that involve the dispersion of pollutants. By estimating the concentration of a substance in the atmosphere, given specific input, it provide a means of investigating the fate of an emission (Tiwary & Colls, 2010). There are various modelling techniques, each with a different principle and thus is designed for different purposes. Similarly, a great number of software has been developed to employ these techniques whilst offer additional functionalities.

The key issue with applying dispersion modelling in developing countries is that these countries do not have enough quality data to fulfil either the input requirement or model validation (Biggart, et al., 2019). As such, even though there have been some modelling attempts in Vietnam, the scarcity of data has prevented software from achieving their optimal performance. Being marketed as an adaptive model that welcomes non-specialist users, Atmospheric Dispersion Modelling System (ADMS) is well suited to the road traffic network of Hanoi and HCMC.

1.2 Aim of the Research

The aim of this Master research project has been to use ADMS to produce a full mapping of road traffic derived emission for Hanoi and HCMC. **The emphasis is placed on Vietnam's data limitation. Thus, the idea is to develop a framework that utilises online, public accessible data in combination with official, quality measurements to create an ADMS-suitable input dataset**. This framework will firstly demonstrate that ADMS, i.e. an advanced dispersion model, is usable in Vietnam and developing countries.

Furthermore, the created dataset aims to put all of ADMS's features to use. As such, it will showcase the superiority in functionality compared to previous modelling attempts. On the other hand, it has not been the intention of this research to produce an in-depth, accuracy optimised investigation of the cities' pollution, but to present a potential tool that can bring great assistance to Vietnam's air quality management system.

1.3 Objectives

- Assess the requirements of Vietnam, as an example of a developing country, laid upon a modelling framework
- Address the issues that previous attempts had met with dispersion modelling in developing country and propose improvements
- Develop an optimised ADMS modelling framework
- Using CO as a pilot pollutant species, validate the model output
- Identify future improvements and applications

2 Literature Review

This chapter provides definitions and concepts leading to the employment of ADMS to simulate road-traffic derived air pollutants in Hanoi and HCMC. Case study of previous applications of ADMS and other notable modelling tools are also included to evaluate their strengths as well as potential issues when applied to Vietnam, as an example of a developing country. Lastly, ADMS's structure, features and input requirement are briefed.

2.1 Air Pollution and Emission Sources

According to (US Environmental Protection Agency (EPA), 2008), "An air pollutant is any substance in the air that could in high enough concentrations harm human, animals, vegetation or other natural materials". The emission of air pollutants can be from both natural and anthropogenic sources.

Air pollution sources can be divided into 4 types based on their geometry: point, line, area and volume. This classification allows sources to be located and defined in any three-dimensional space using coordinates.

2.2 Dispersion Modelling as an Urban Air Quality Management Tool

As regard to urban air pollution in developing countries, the emphasis is placed on providing health protection to the large, ever growing population, which then benefits the sustainability of development (Metcalfe & Derwent, 2005). Traditionally, epidemiologic researches rely on real world observations as to quantify exposure and health impact (e.g. death, hospital statistics) (Russell, et al., 2014). Measurements are however a passive method of air pollution detection. By monitoring the end-result, measurements alone "provide little information on the origin of the pollutants in question, on the dispersal process in the atmosphere and on the impact of new sources or the benefits of controls" (Williams, 2014). As such, whilst a pollutant can be used as an indicator of a process, e.g. NO_x is mostly produced by fossil fuel combustion, identifying the largest emitter from a group of sources that all involve said process would face various challenges. Resource constrains in Vietnam, namely the lack of comprehensive emission inventories and unreliable monitoring data also amplify the difficulties in finding the responsible source. Additionally, observations study historical events. Trends and patterns can therefore be addressed; predictions are however very limited.

Being 'an attempt to replicate a simplified representation of a part of the real world – the system of interest – and its behaviour' (Ortuzar & Willumsen, cited in Palmer, 2007, p.15), models take account of the entire process rather than just the result. As such, dispersion modelling mimics the way air pollutants behave in the atmosphere through the use of mathematical equations; attempt to describe the fate of emissions by calculating concentrations at desired locations and times (Holmes & Morawska, 2006). This offers a proactive method of study that by adjusting variables (e.g. meteorology, emission sources/ factors, chemical reactions), can fill the limitations of using observations alone. Subsequently, dispersion modelling improves the estimate of the contributions of current sources and allows future predictions. Scenarios are particularly important when the rapid developing rate and its following changes are taken into considered.

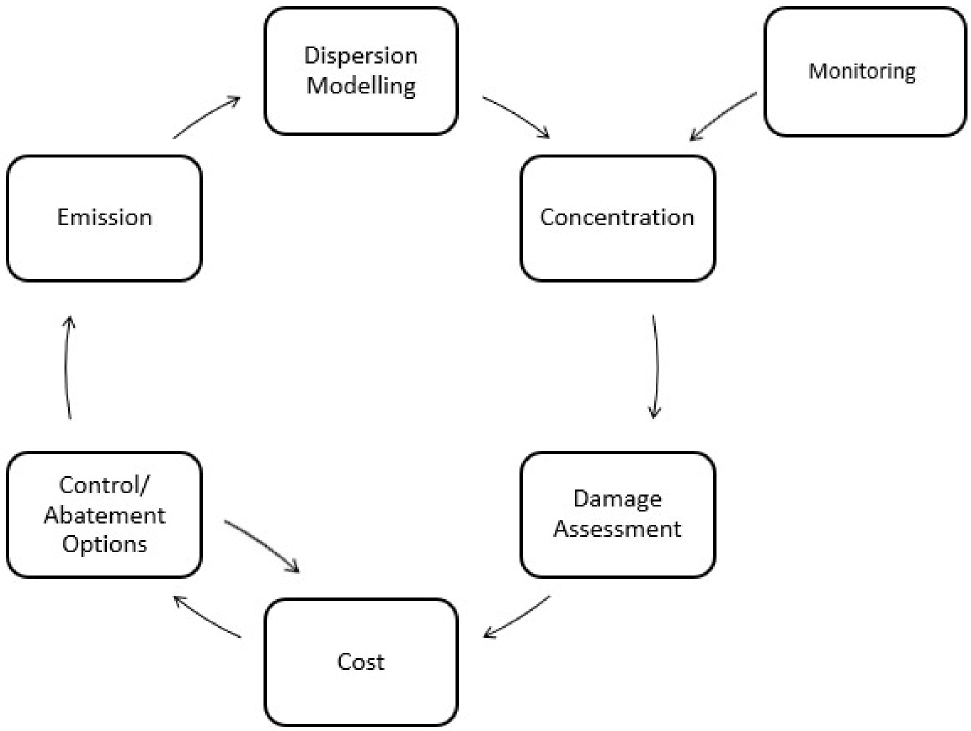


Figure 2. Elements of the Air Quality Management Strategy (Steinar, et.al., 1997).

Like any other management systems, an urban air quality management is dictated by its analysis of the cost – benefit relationships. Figure 2, which is a simplified version of the diagram shown in (Sternar, et al., 1997), suggested that modelling is the core of said analysis. Without dispersion modelling, there would be a missing link between emission and concentration. In turn, changes in emission – either reduction (e.g. vehicle ban) or increasement (e.g. new factories) – cannot be effectively evaluated and valued. Decision makers therefore lack the information to offer definite solutions. For example, Vietnam's National Environmental Report (2013), which relied on monitoring and emission inventories, could only ranked air quality as fail or pass the official standards. Their proposed investment strategy therefore remained unassertive, relying on updating the policies and performing physical inspection to individual sources. To a developing country, unspecified solutions are both impractical and wasteful.

As according to (Hung, 2010), Vietnam had made little progress in managing urban air pollution because whilst practices of air quality monitoring were fragmented, no other tools are available either. As such, even though environmental laws had been regularly updated based on international regulations, policy enforcers had limited resources and evidences to uphold legal standards.

Employing dispersion modelling requires a clear understanding of the real world. Both the model and the modeller's input must firstly represent the actual system. The modeller therefore have to break down the related variables, study their roles and allocate them with justified values. Due to the nature of this project, i.e. applying a known method to a foreign environment, data issues ought to be anticipated (e.g. unavailability, incompatibility). Assumptions will thus be made, accompanied by best attempts of logical reasoning.

As to kept the scenarios making function, built models should however remain adaptive, allowing for future adjustments, improvements and manipulations. An initial objective should also be repeated: this work is designed to demonstrate a potential tool that help the air quality management system of a developing country; it does not aim to be an in-depth assessment of the current air pollution issue.

2.3 Techniques of Dispersion Modelling

Theoretically, a perfect model that could accurately calculate concentrations at any chosen time and location, would replace measurements from all purposes (Tiwary & Colls, 2010). Since such ideal is not yet available, air quality models still have limitations, that arise from errors in the input (e.g. data assumptions made by modeller that introduced uncertainties and biases) and model parameter and structure (Russell, et al., 2014). Figure 2 indicates this by having the monitoring module separated from the cycle, to be used as a validation tool.

Choosing a model is thus essential. This section demonstrates the researcher's understanding of the four most referred air dispersion modelling techniques: Box model, Eulerian model and Lagrangian model, Gaussian model, and Computational Fluid Dynamics (CFD) model.

Box model treats the modelling site as a defined volume of the atmosphere – a box. Being the simplest type, pollutants within this box are assumed to be homogeneously distributed. Concentrations are thus uniform throughout (Tiwary & Colls, 2010). Input requirement is also minimum, including emission, meteorology and movement in and out of the box. This simplicity whilst restricts the usage of box model, also allows more detailed chemical and physical equations to be applied on the aerosol dynamics (Holmes & Morawska, 2006). For example, box model has been incorporated to address the effects of the re-circulation zone in street canyons (Hung, 2010).

Eulerian model and Lagrangian model are of stochastic type, which is based on semi-empirical mathematical relationships between the concentrations and any influencing factors (Tiwary & Colls, 2010). Whilst Eulerian has a fixed reference system, Lagrangian's travels accordingly to the dispersion of pollutants. The principle is to let a pollution parcel move downwind with inputted average speed and direction. Along its course, concentrations are diluted by ambient air and volume is increased. Turbulences are applied in the form of random forcing onto the motion of the airmass. The resulted wind vectors are recorded as a trajectory. This process is then repeated multiple times to obtain a sizable sample of potential trajectories. The sample is used to statistically compute the probability that said parcel would reach a chosen location at a particular time, i.e. the destination. Finally, concentrations at the destination are calculated using the original emission mass and that probability (Anfossi & Physick, 2004). These models specialise at estimating the dispersion of a specific emission source over large distances and/ or at the effect of complex meteorology conditions. A notable example of Lagrangian model is UK's NAME, that was developed originally in prepare of nuclear accidents. NAME has been adopted into other applications such as regional scale prediction of nitrate and sulphate aerosol (Redington &

Derwent, 2002), and dispersion of volcanic ash (Leadbetter & Hort, 2011). Similarly, Austria's GRAL model has been applied to simulate emission derived from a specific highway (>30km) in complex terrain at difficult low wind conditions (Oettl, et al., 2012).

Gaussian model in principle is a subset of Eulerian. Instead of re-calculating the probability for each destination, Gaussian dispersion theory assumes the fluctuations in wind vector to be normally distributed (Tiwary & Colls, 2010). This enables concentrations to be conveniently calculated at any location. In other words, it offers an 'efficient compromise between reasonable accuracy and manageable computational time' (Fallah-Shorshani, et al., 2017). There are great number of working Gaussian models, each aims to improve an aspect of the theory's limitations such as the lack of consideration given to physical (e.g. deposition) and chemical (e.g. NO_x and SO_x) processes (Vardoulakis, et al., 2003). The three most important drawbacks with Gaussian models are: Inaccuracies have been consistently observed in low wind conditions or at close approximates (<100m) (Holmes & Morawska, 2006). Long distance (>100km) modelling is also not recommended since Gaussian equations assume winds would be the same across the domain. Obstacles (e.g. buildings, infrastructures) are defined poorly via surface roughness parameters, applications in small scale urban areas require extra assisting modules.

CFD models are 3D based simulations of fluid flow, heat transfer and related processes (e.g. chemical reactions) (Vardoulakis, et al., 2003). Its purpose is to study complex flows and turbulences caused by obstacles using flexible fine-scale grids. Conventional models, being less tractable, thus fall short in small-scale dispersion compared to CFD (Williams, 2014). As a trade-off for high resolution, this technique is resource demanding. Examples of CFD's best use can be the modelling of wind and microclimate in complex urban areas (Antoniou, et al., 2019) (Kaseb, et al., 2020), or the impact of certain urban micro areas (e.g. green space, traffic hot-spot) onto air quality (Moradpour & Hosseini, 2020) (Sanchez, et al., 2017).

2.4 The Modelling of Road-traffic Derived Emission and the Issue when applying to Developing Countries

This section revisits studies that employed those above mentioned techniques on the modelling of urban road-traffic pollution, preferably in developing countries. Combined with recent findings in Vietnam, the objective is to evaluate the applicability of these techniques to Hanoi and HCMC.

An example of the box modelling approach was in Santiago, Chile (Jorquera, 2002). A Box model was used because of 'a lack of reliable emission inventory databases' and of the information on related processes, most namely chemical reactions. Apart from other sources, the city's vehicles were all grouped together and emission data was extrapolated from fuel consumption. Estimates were produced on a monthly basis. From those results, further evaluations were made regarding seasonal and yearly variations. Comparing modelled concentrations against a number of monitoring stations across the city showed slight overpredictions for annually averaged CO and NO_x (2.5% to 21%). Result indicated the dominant contribution of mobile sources to air pollution. This could preliminarily guide decision makers, e.g. traffic emission needed to reduce by 75% for concentrations in winter to be as low as in summer; old and heavy vehicles were responsible for 65% and >50% of CO and NO_x respectively.

Due to its principle, i.e. modelling the dispersion of a particular emitting source rather than multiple explicitly identified sources, Eulerian/ Lagrangian models have limited usage when considering an entire road network. Conversely, they are more suitable with environmental impact evaluation of known emission hotspots or future infrastructure developments. (Rafiei & Sturm, 2018) employed this technique to investigate the impact of ventilation from a tunnel in Tehran, Iran. Input requirement included: the portals' location, the desired emission inventories (this project used the worst case scenario – congested traffic), meteorological data from nearby stations and some simple topographical parameters. The resulted CO contours could be used to provide further understanding of the health impact at the close vicinity of portals (<200m).

There are many applications of the Gaussian modelling technique to road-traffic emission. Studies have implemented different models (e.g. ADMS-Urban, AERMOD, OPSM) and varied levels of complexity/ detail. The core inputs are similar to which described above with the Lagrangian case study. Depended on purposes, large-scale dispersion mapping (e.g. (Biggart, et al., 2019)) or in-depth evaluations using hybrid models/ integrated modelling chains (e.g. (Misra, et al., 2013)), can all be achieved with corresponding additional input.

Thanks to recent computational advancements, CFD models have become more practical (Tominaga & Stathopoulos, 2011). Applications studying pollutant transport in the vicinity of roads are more frequently conducted. Other than human exposure to pollutants, these works can assess the impact of wind field upon monitoring sites, essentially re-evaluate the quality of observations (Costabile, et al., 2006); (Santiago & Martin, 2015). Models are therefore very specific and input hungry, e.g. using blueprint measurements to map a building (Harmanto, et al., 2015); sensor monitored traffic and automatic congestion indicator (Dixon, et al., 2006). There are thus cases that rejected CFD modelling due to data inadequacy (Rafiei & Sturm, 2018).

TAPOM, a Eulerian based model, has been successfully employed to simulate the dispersion of pollutants in HCMC (Bang, 2010). Grid source is used in TAPOM and the emission of each cell is applied to an equation, which contains consideration for meteorology, turbulence and chemical reactivity (Brulfert, 2004). Evidently, to employ TAPOM, a method of aggregating individual emission sources into cells are needed, thus the use of EMISENS, a tool that computes new sets of gridded data from both top-down and bottom-up emission inventories (Bang, 2010). This method, designed to tackle the lack of emission data in developing countries, has proven to be very practical. However, by only using amassed cells, TAPOM has certain drawbacks, namely reduced details (e.g. road canyons), inconvenience when adding new data/ variables (e.g. emission temporal variation, changes in vehicle fleet), deviation from standardised formats (e.g. GIS-defined roads). In short, even though TAPOM coupled with EMISENS is an able framework, it lacks the flexibility for large scale applications such as quick predictions or fine adjustments because every input changes would involve recalculate the entire process.

Another notable application of dispersion modelling in Vietnam is with OSPM. This street canyon model couples a Gaussian model for the flume derived from vehicles and a box model for the effects of turbulence (Holmes & Morawska, 2006). Introduced by OSPM, the recirculation zone within road canyon has been employed by many air quality models namely AEOLIUS and ADMS (Chatzimichailidis, et al., 2019). Since it is extensively referenced, time efficient and gives attention to many contributing variables (e.g. traffic produced turbulence), it has been widely employed to study pollution level at street scale (Kakosimos, et al., 2010) (Vardoulakis, et al., 2007) (Berkowicz, et al., 2008) (Lazić, et al., 2016). (Hung, 2010) calculated the pollutant concentrations on 5 roads in Hanoi to evaluate how suitable OSPM is with limited input data of varying quality. As such, whilst each road was individually configured, detailed traffic count was sourced from previously conducted studies.

Also by (Hung, 2010), OML was applied as to estimate and map the pollution at urban scale in Hanoi. Similar to (Bang, 2010), roads were grouped, multiplied with surveyed emission factor then divided into grid cells. Point and area sources were then inputted alongside traffic. OML is Gaussian based and uses boundary layer scaling instead of Pasquill stability classification. The main limitations with this model are currently the lack of consideration for deposition and the simplified method used to describe building effects.

(Hung, 2010)'s conclusion was that even though models could be implemented, the efficiency was mostly influenced by the lack of 'a GIS data set with detail parameters of the road network, vehicle volume for each street and studies of emission factors for each vehicle category'. Next, validating modelled result against observations faced difficulties because the contribution of other emitting sources (i.e. industrial, domestic) was poorly understood. Lastly, the monitoring network was not effectively configurated, e.g. stations were operated differently or built too high instead of at street level and there was no regional background measurement.

HoanKiemAir project has recently been initiated as to produce a visualised 3D simulation of the effects of traffic upon air quality in Hoan Kiem district, Hanoi (Figure 3). Limited documentation was found on the project, however it may prompt the applications of more sophisticated models in Vietnam, most namely CFD.

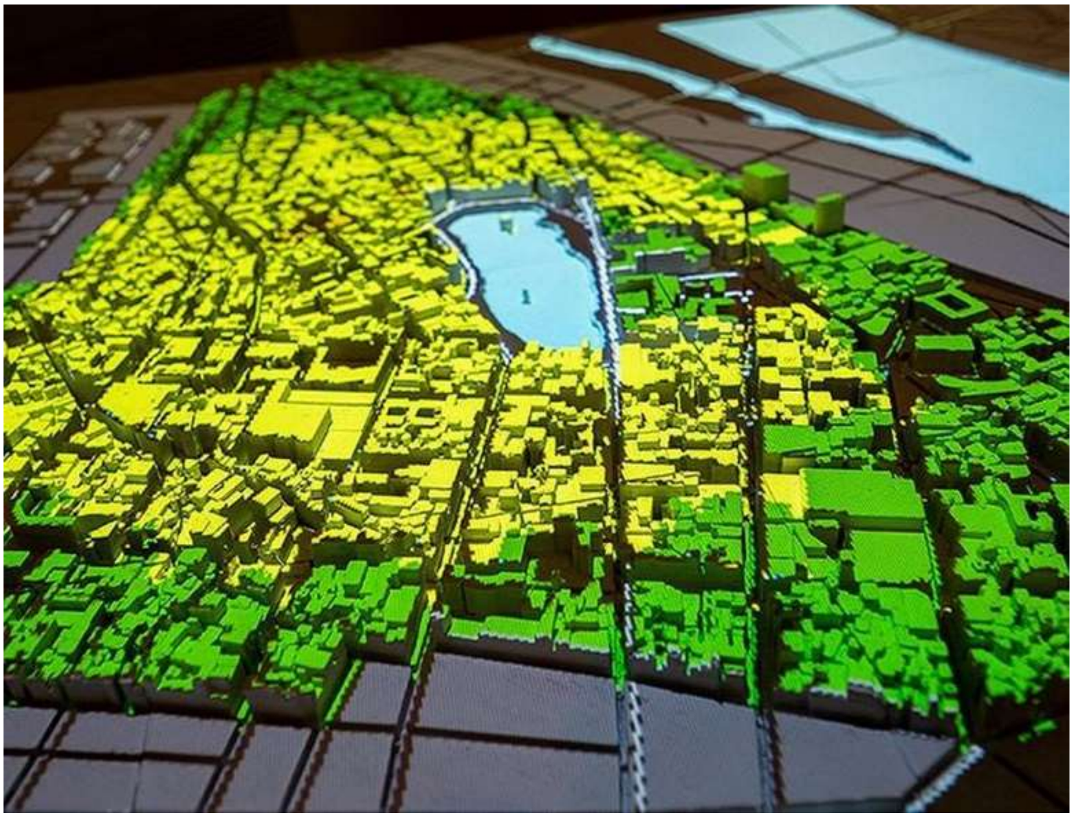


Figure 3. A demonstration of the HoanKiemAir simulation (Khanh, 2019). Air Quality Index displayed on buildings, less green is more polluted.

Out of the four mentioned dispersion modelling techniques, Gaussian based models seem to be the most fitting for the road network in Hanoi and HCMC. Not only have there been successfully employments, outputs are produced at detail levels adequate for current regulatory purposes (e.g. urban scale environmental impact assessments, preliminary human exposure evaluations). Also, the input requirements are manageable. Gaussian models are preferred over Eulerian/ Lagrangian based because they are more dynamic. As such, with the current rapid changes in traffic, the modelling of emission is best suited with a fast, versatile process. There is also a wide variety of Gaussian models, each offers its own additional modules to improve functionality.

2.5 The Development of ADMS

ADMS-Urban is the complete version of Atmospheric Dispersion Modelling System (ADMS or UK-ADMS) and is a Gaussian based model that has an emphasis on versatility. As such, realising that local authorities do not have a uniform, adequately equipped tool to be used for their air quality assessments, ADMS is developed by the Cambridge Environmental Research Consultants (CERC) to 'bring the latest modelling techniques to the non-specialist user' (Carruthers, et al., 1997). The software has been commercially released since 2002.

Described as an advanced dispersion model, ADMS incorporates the Monin-Obukhov length similarity theory to describe the structure of the boundary layer (Douglas, et al., 2016). It allows point, line, area, volume and grid sources to be explicitly inputted and modelled simultaneously within a referenced domain. Other features such as street canyon, road tunnel, chemical reactions, up-to-date EFs, GIS compatibility and domain nesting are also available (CERC, 2017a). Similarly, resulted pollutant concentrations can be produced at various frequencies (e.g. 10 minutes – 1 year) and resolutions (i.e. up to 2001 points each horizontal direction, 501 vertical levels). These aspects of ADMS provide great adaptability, i.e. simulations can be configurated to match specific modelling purposes, e.g. study the simplest sources (a singular stack or line) to complex scenarios (an entire urban area); or increase/ decrease the temporal output to suit the need of model evaluation.

ADMS has been extensively validated against similar models developed in the UK and worldwide. For example, (Stocker, et al., 2013) studied 4 different models, including ADMS-Roads, US-EPA's regulatory model AERMOD, California's CALINE and US-EPA research model RLINE. Performances were assessed through 2 experiments, one mimicked the real-world scenarios whilst the other was more controlled. ADMS exhibited the best correlation/ highest r value at 0.78 and 0.88 for each experiment, respectively. With the real-world scenarios, ADMS also achieved the best Normalised Mean Square Error (NMSE) (i.e. 0.2 with ideal = 0) and prediction within the factor of two of the observations (Fac2) (i.e. 0.85 with ideal =1).

As for applications of ADMS, London's air quality in 2012 was simulated by (Hood, et al., 2018). Input emission data were gathered from the London Atmospheric Emissions Inventory and projected to the modelling year (e.g. from 2010 to 2012). In light of the Volkswagen emissions scandal, some adjustments were made to recalculate the vehicular EFs for all road traffic sources in London. This increased the total annual NO_x emissions by 55%. The result suggested a consistently good agreement between modelled concentration and observation across all 6 pollutants, including NO_x, NO_2, O_3, CO_2, PM_{10}, and $PM_{2.5}$. In specific, the r – value ranged between 0.504 and 0.777, Fac2 was between 0.664 and 0.882, and the Fractional bias (Fb) was below ± 0.12. Similarly, apart from NO_x with NMSE of 0.728, which was likely caused by the uncertainties of the EFs adjustment, the remaining was below 0.4. On the other hand, concentrations modelled using adjusted EFs showed substantial improvements over which of raw EFs, especially for near-road sites: Fb from -0.3 to 0; NMSE from 0.88 to 0.62. This comparison demonstrated an important potential application of ADMS, which is to predict/ quantify the impact of improved EFs, via better fuels (e.g. cleaner coal for generators), or technical advancement (e.g. phase-out old vehicles, alternative transportations). In overall, (Hood, et al., 2018) is a good example of a dispersion modelling work using ADMS, featuring a typical method of sourcing input data, i.e. temporal projecting, and the elevated level of attention that can be included to better depict the real-world conditions.

The work of (Biggart, et al., 2019) shared a very similar concept with this project. As such, it aimed to use ADMS to produce pollutant concentrations across urban and suburban Beijing, using freely available geographical data to explicitly identify road traffic sources. Realising the inefficiency of manual vehicle counting, the study took advantage of the Multi-resolution Emission Inventory for China (MEIC). MEIC, which is a gridded emission inventory, was optimised and apportioned cell-by-cell to individual roads. The optimisation process contained two stages: 1. Temporal adjustments (i.e. re-scaling MEIC from 2013 to 2016), and 2. Spatial adjustments (i.e. re-distributing part of the urban emission to the surrounding suburban areas). This method in essence was proposed as a combination of improved spatial detail (e.g. to calculate near road concentrations) and reduced input data intensity. For $PM_{2.5}$, the result showed good agreements between modelled and observed concentrations with NMSE, Fb and r at 0.37, −0.04 and 0.76, respectively. NO_x and NO_2 exhibited lower r values (0.53 and 0.41), mostly due to MEIC not fully account for the night-time influx of heavy diesel vehicles. In turn, thanks to ADMS's included chemistry scheme, refining NO_x emission was shown to improve model performance with O_3. On the other hand, inadequate background observation in Beijing, as with many developing countries, contributed to the overall uncertainties. The study also pointed out several directions of further development namely refining explicit road emissions network for higher detail (e.g. vehicle type, road type, congestion, and different diurnal profiles), fine scale evaluation for health exposure-related studies, or develop model modifications tailored for specific metrologies (e.g. very stable, low wind conditions commonly found during East Asia winter monsoon).

2.6 Model Features

There are four variables to be considered for the modelling of road traffic emission: source location, source intensity, dispersion and background concentration. Even though this project aims to demonstrate the full capacity of ADMS in order to propose its applicability to Vietnam, some of the optional features had to left unused due to the time and data limitation. Those were namely chemical reaction schemes, advanced road canyon, trajectory model and deposition. Chemistry scheme in particular, should be included in following works since it simulates 1. the simplified $NO_x – O_3$ relationship, which can change the ratios of NO/NO_2, NO_2/NO_x, and NO_2/O_3; and 2. the conversion of gaseous SO_2

to sulphate. In turn, if chemistry were applied, modelled result would likely exhibit 1. sharp increases in pollutant species concentration near major roads; 2. a slower response of NO_2 concentrations to NO_x emission changes; and 3. increase in PMs concentration.

Figure 4 illustrates how the remaining ADMS's features are used to define simulations. Table 2 lists the primary input file formats. The property of these features and their corresponding input requirement are discussed below. This work uses ADMS-Urban version 4.1.

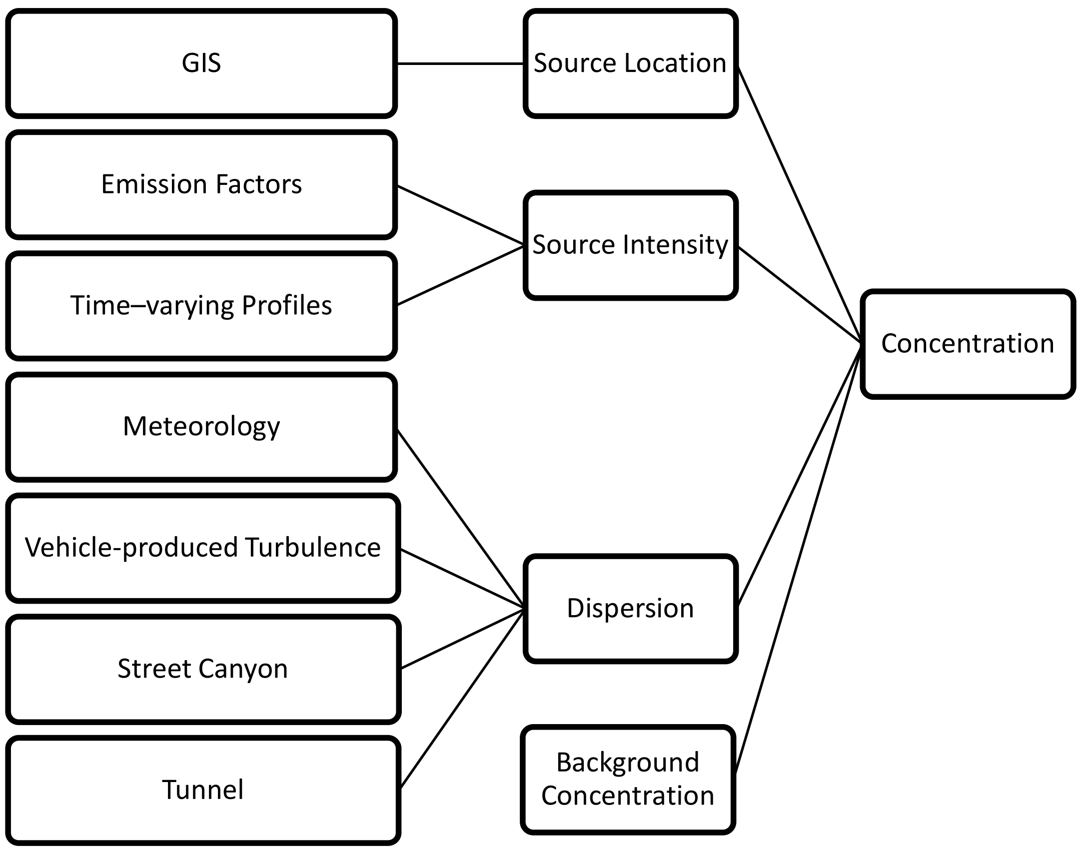


Figure 4. Demonstration of ADMS's modules used to calculate the emission of road sources.

Table 2. ADMS's input files.

Extension	Description
.vgt	Vertex information for all sources
.spt	Sources' properties (e.g. type, efflux parameters)
.eit	Emission information for all sources
.tft	Traffic flow for road sources
.fac	Emission diurnal variation factor
.bgd	Background pollution data

Source Location:

To large-scale capable dispersion models, the integration of Geographic Information System (GIS) technology is a very advanced functionality. As such, other than ADMS, only few others are GIS-equipped, namely AERMOD, DUSTRAN and OSPM (Gulliver & Briggs, 2011). ADMS is therefore able to produce results of high resolution that are directly useable in geospatial data manipulating software such as ArcGIS, QGIS and MapInfo. Additionally, many fields of information can be attached to cartographic data, meaning GIS is not only beneficial to source location but also to other physical depictions of the real world (e.g. terrain, buildings). As for road sources in ADMS, their coordinates must be listed in a .vgt file. Each road can only have a maximum of 50 vertices and there can only be 3000 roads per model. Consequently, if domain contains too many roads, modeler have to either reduce the input resolution, or shrink the area.

Source Intensity:

Previous modelling attempts show that in developing countries, line sources' emission data, which are extrapolated from other datasets (e.g. regional emission inventory, fuel consumption), have great uncertainties (Bang, 2010) (Jorquera, 2002). Optimisations can be made to substantially improve the data quality and the suitability for street-scale modelling (Biggart, et al., 2019). Still, certain intriguing details (e.g. diurnal variation, contribution of vehicle classes) remain missing from proxy-based emission inventories. Vehicle counts, even though are more difficult/ laborious to obtain, should therefore be used for emission input. Each road's traffic data including number of vehicles per hour and average speed, must be listed in the .tft file for ADMS to process.

Although it is known that vehicular emission varies according to fuel type, vehicle condition and travelling condition (i.e. speed, on-off motion, road quality), such variations are poorly understood for the vehicles in Vietnam. This is due to 1. The lack of vehicle quality control and 2. The chaotic traffic pattern (e.g. MC filtering between queues) (Trang, et al., 2015). Many studies in Vietnam must therefore compensate by adopting EF elsewhere, e.g. EPA, Bangkok, Euro standard (Trang, et al., 2015), . ADMS's integration of various EF datasets from the UK and China is a very practical feature. As such, once an emission dataset is addressed in the .spt file, the EF of each road would be scaled with the average speed on said road. Conversely, through a .eit file, it also remain possible for modellers to define fixed emission for each road.

The ability to define time-varying EFs is very important to dispersion modelling in Vietnam. This is because, as demonstrated by (Chen, et al., 2009) and (Cai & Xie, 2011), the function allows impact of traffic control policies to be evaluated. ADMS requires diurnal profile (i.e. a set of 24 hourly factors) to be inputted into .fac files.

A road source's total EF at a particular time is given by:

$$t_{EF} = \left(\sum_{i}^{n_v} N_i E_U \right) F_t$$

where:

nv	= number of vehicle categories
Ni	= number of vehicles per hour for that vehicle category
EU	= emission rate scaled with speed
Ft	= hourly factor at the time

Equation 1. Calculation of a road source's emission factor at a particular time

Based on the above, all vehicle types on a road would share an identical diurnal variation pattern. This is not always the case with megacities in developing countries (Chen, et al., 2009).

Dispersion:

ADMS's meteorological pre-processor module is called to calculate the boundary layer meteorological parameters required for running the dispersion simulation (CERC, 2017b). This module is designed to allow both hourly sequential and non-chronological, statistical data can be processed. Likewise, the input requirement can be adaptable to both specialised and more standard datasets. The minimum input consists of 4 variables: wind speed, wind direction, cloud cover and time (hour/day/year). Based on data availability, users can improve the estimate with other variables. For example, the surface sensible heat flux, if supplied, would be preferred over cloud cover; when data are scarce, temperature should be added to the minimum requirement. All variables must be written in a .met file as input.

To ADMS, roads are configured as line sources that have additional consideration given to traffic produced turbulence and street canyons (CERC, 2017c). Essentially, the busier the traffic flow, the more turbulence is induced thus the faster emission is dispersed. This extra turbulence is calculated using the number of vehicle per second, speed and vehicle cross sectional area. Based on UK's vehicle classification, there can be up to 8 vehicle classes varying between MC to specialised HGV; ADMS assigns each with an area value, ranging from 2m^2 to 16m^2 (CERC, 2017a). Compared to OSPM, which only uses 2 values for either small or large vehicles (Vardoulakis, et al., 2007) and ignores turbulence produced by MC (Hung, 2010), ADMS suggests better detail.

ADMS is sometimes referred as a quasi-Gaussian model, mostly because of the integrated box model based street canyon modules (Holmes & Morawska, 2006); (Hood, et al., 2018). This function is therefore a signature addition over other Gaussian based models. ADMS offers two options including basic and advanced street canyon. The basic module is based on OSPM and requires less input data, whereas the advanced one has a wider range of functionality (CERC, 2017d). The key difference between these two options is that the basic street canyon function assumes canyon walls are symmetrical and nonporous. It also can only model the effects upon points inside the canyon, meaning it is designed for health exposure related evaluations rather than urban-scale pollution mapping. As a result, the inclusion of this function would have limited impact on the overall output here, but would be beneficial as a steppingstone to apply advanced street canyon in the future. With that said, testing both options on a known contour would be interesting as it may obtain further findings compared to (Hung, 2010), which employed OSPM.

To mimic the accumulated emissions within tunnels and how they are released through tunnel portals by following the traffic and affect the surrounding concentrations, the road tunnel module is available in ADMS as an optional feature (CERC, 2017b). This module requires a .csv file that contains the location, depth and width of portals, as well as the name of the connected roads (i.e. outflows). Each outflow is then modelled as 3 volume sources, with emission obtained from the tunnel.

Background Concentration:

Due to limited time and available data, chemical reactions are not implemented in this work. Information on the underlying pollution therefore has reduced importance and is optional. Still, some background concentrations should be added, either as a consistent value, or as an hourly sequential series (written in a .bgd file). This would help with the model validation (Biggart, et al., 2019).

2.7 Hanoi and Ho Chi Minh City Background

Sharing similar morphologies consisting of a flat terrain, multiple rivers and suitable climates, Hanoi and HCMC are heavily populated and are the core nodes of development in Vietnam: Hanoi hosts most of the governmental institutions; HCMC is the headquarter of manufacturing and logistic activities (World Bank, 2011).

From a geographical perspective, the main difference between these two cities is their distances from the sea. Being inland, Hanoi is less involved in trade related activities compared to HCMC, which is coastal. The dissimilarity in function and physical profile results in different urban planning approaches. Whilst Hanoi aims for a metropolitan area comprised of a core and serval satellite cities, HCMC targets the development of an efficient, focalised urban living space. An important landmark supporting this was the 2008 expansion, which increased the area of Hanoi by 3.6 times. Simultaneously, this added another 2.8 million to Hanoi's population and made it on a par with HCMC.

The population are thus distributed differently for these cities. Hanoi is 37% larger, but it holds approximately 13% less population. Similarly, the net migration rate of Hanoi is -0.3‰ whilst HCMC is 5.3‰ (Vietnam General Statistical Office, 2016).

In turn, the road network is affected. Hanoi has much higher total length of trunk and motorway to fulfil the demand of long distance travelling, whilst HCMC is crowded with smaller roads.

There are also differences in the weather regime. Hanoi, like the majority of the Northern region, has 4 distinct seasons: summer, winter and two transitions. Whereas HCMC only has wet and dry seasons. Hanoi has prevailing winds from the South East during summer and North East during winter. West South-West winds dominate HCMC from June to October, North North-East winds from November to February and South South-East for the remaining months.

Overall, whilst Hanoi and HCMC face similar issues that result in their shared concern over road traffic emissions, their differing socio-economic functions and natural conditions demand individual attention. As such, for in-depth, accuracy-optimised evaluation of the cities' atmospheric pollution, studies are needed for each individual city. However, with the scope of this project and that the availability of emission data were initially uncertain, data were anticipated to be clustered and shared between cities. Some generalisations would therefore have to be made and certain variations between the two cities would be neglected.

2.8 Summary

It is evident that Vietnam's urban air quality management system can benefit from dispersion modelling as a proactive response to its issues. However, like many other developing countries, it has difficulties obtaining high quality data to support the applications. Out of the four most referred modelling techniques, Gaussian models appear to be the most suitable as they have a good combination of detail/ accuracy and resource requirement. ADMS, being a Gaussian model with improved adaptability, is thus chosen for the dispersion modelling of road traffic emission in Hanoi and HCMC.

3 Methodology

This chapter consists of 2 sections: input data and concentration data and experiments. As such, the first section summarises the process needed for data sourcing and processing in order to satisfy ADMS's input requirement. The second briefly describes the available concentration datasets, which set the scope for model validation.

3.1 Input Data

3.1.1 Modelling Boundary

As discussed above regarding Gaussian models, a domain with medium to large spatial scale is desired. The objective is to define boundaries which cover the cities' road traffic activities (i.e. internal and partial of the external travelling needs).

Thus, a square with edge of 120 kilometres was allocated to each city (from this point onward, they are referred as Hanoi's and HCMC's square). As a trade-off for exceeding the recommended size for domains in a Gaussian model, these boundaries also include some nearest provinces to the cities.

For Hanoi, The upper and left margins have minimum distances (i.e. approx. 7km) away from the city's border. The right and lower ones respectively go through the central of Hai Duong and Nam Dinh. Including Hanoi itself, there are 15 first level administrative regions that contributes in the expanse of this square, some of which have insignificant presence, e.g. Ninh Binh, Thai Nguyen, Lang Son. Similarly, Figure 5 illustrates the square assigned for HCMC, which overlays the share of 8 provinces as well as a part of the ocean.

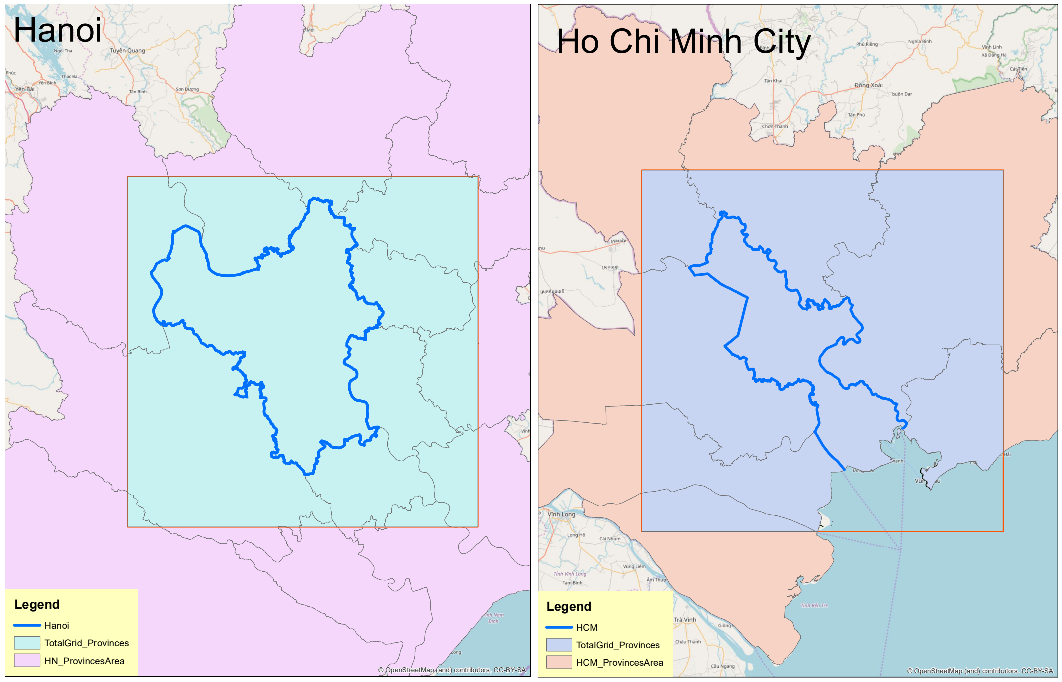


Figure 5. Modelling boundary for Hanoi and HCMC.

ArcMap was used to estimate how much area of a province is included within a square. By assuming the population density of a province stays constant across its area, the population contributed by each province can thus be calculated as:

$$p = \frac{\text{P x a}}{\text{A}}$$

Where:

p = Population contributed to the modelling boundary by a province

P = Total population of said province

a = Spatial area of the province overlaid by the modelling boundary

A = Total area of the province

Equation 2. Calculation of population from each province

With that, the population contained within each square was obtained: 14.2 million for Hanoi's and 13.6 million for HCMC's. This 6.3% difference in population counts was used as a proxy to indicate that these two squares would have similar number of vehicles and thus total road-traffic emission. In essence, this was an attempt to normalise the two cities, despite their differences as discussed in Section 2.7. Ultimately, it would allow the use of traffic data from either city on another, a necessary preparation considering how difficult these data were to retrieve (Section 3.1.3.1).

3.1.2 Geographical Data and Modelling Domain

3.1.2.1 Road Data Collection

Initially Tried Online Accessible Sources:

Aligned with the preference of data from public accessible sources, several sets of road data in the format of shapefiles were tried.

Data was firstly sourced from GIS.vn, a website which claims to be the official channel for Vietnamese government to share geographical data for the purpose of mutual development. Whilst the shapefiles were detailed, the server was a beta version with a very unreliable connection.

Next, the Global Roads Open Access Data Set (gROADS), v1 (1980-2010) developed by the Socioeconomic Data and Applications Center (SEDAC) – NASA was tested. However, the coordinates were inaccurate whilst the level of detail was low.

The Center of measuring and mapping data, MONRE also offered geographical data for major Vietnamese cities. Although this seemed to be the most official channel, it required a fee and a complicated procedure, including a request document from a governmental unit.

OSM Data – its Versatility and Compatibility with ADMS's Input Processing:

The dataset of choice was sourced from OpenStreetMap (OSM) via Geofabrik . Its advantages include being detailed, free, readily available, regularly updated and most importantly, widely used as a source of geographical data for academic and researching purposes. However, as warned by Geofabrik, in order to achieve such efficiency with the data collection, OSM must relies on the flexibility in the mapping process of features. As such, rules "are often not well defined and there is no mandatory quality control" (Ramm, 2017). This statement also refers to the collection technique, which is fundamentally similar to which of Wikipedia and Wikidata, i.e., importing data from other sources then having volunteers to perform ground surveys and improvements. Subsequently, whilst promoting ground truth and local knowledge, the database is exposed to individual level inaccuracies and uncertainties that ought to be acknowledged thoroughly. As a result, further care was taken when extracting more elaborated details from the database, e.g. comparing HCMC's data with multiple other cities' when looking for the number of lanes per carriageway (Section 3.1.2.5).

OSM attaches details onto map by having each geographical entity, also called 'feature', to contain specific tags. These tags exist in the form of key-value pairs (e.g. entity = building, type = school). This allows great flexibility in data usage as tags can be explicitly extracted for tailored purposes. As such, whilst most users can find frequently used tags from the simplified .shp version, the raw .osm files can be exploited for further details. The latter point had proven to be beneficial to this work.

Extraction of Main Roads:

For the modelling of road traffic emission, features that are unrelated or have limited importance to the travelling condition of the motorised vehicle fleet, are ignored. These includes cycleway, footway, path, track, residential and others (i.e. properties are unknown or unclassified). Service roads are also removed from the model due to their inseparable association with industrial activities, which should requires distinctive attentions compared to those of general uses. The remaining: motorway, trunk, tertiary, secondary and primary roads are therefore extracted as ‘main roads’.

However, since ‘main roads’ is a broad term, there are other methods of categorisation. As such, the task of classification is subjective to the dataset and the definition of choice, which is poorly articulated in Vietnam (JICA, 2010). For example, (Bang, 2018) used a class called provincial road, which is similar to national road in purpose, i.e. to connect administrative centres to important infrastructures, but is managed by individual provinces rather than the central government. Thus, they tend to be smaller and have reduced traffic flow compared to national highways, i.e. trunk. Provincial roads are equivalent to mostly tertiary and secondary in OSM. Similarly, a blurred distinction between secondary and tertiary roads is found in most of the conducted works. Subsequently, in order to merge data from multiple sources, this work converts all roads to OSM’s classification which includes Motorway, Trunk, Primary, Secondary and Tertiary.

3.1.2.2 Administrative Boundaries

The shapefiles, in both polygon and line format, for administrative boundaries were sourced from the Regional Office for Asia and the Pacific (ROAP), United Nations Office for the Coordination of Human Affairs (OCHA). The Office focuses is on assisting international responses to emergency situations, most likely large-scale disasters or humanitarian crises. Their archive of maps and other geographical illustrations are respectable both in term of diversity and quality.

There were three administrative levels included within this dataset. This was in line with the current hierarchy set by the National Assembly of Vietnam (Figure 6). The last data update was in 2015, thus these shapefiles covered most major boundary adjustments, namely the Hanoi expansion in 2008. However, due to the rapid increase of administrative units, which is unusual for most countries, there were certainly some outdated insignificant details (e.g. at the 3rd – communal level).

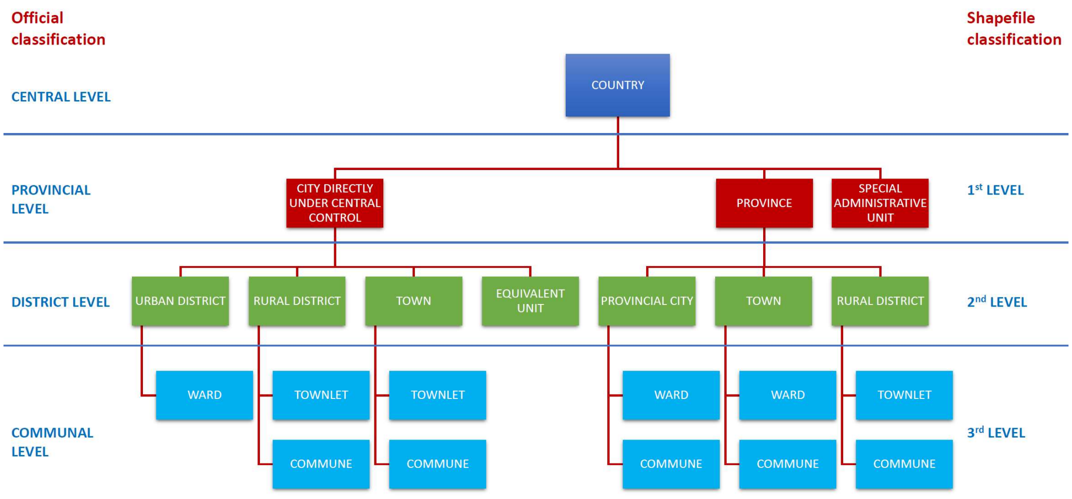


Figure 6. Vietnam's official classification of administrative levels and OSM's equivalent.

3.1.2.3 Road Shapefile Processing

This section covers the process of converting freely available geographical data into formats that satisfy ADMS's input requirement. There were 3 main steps: 1. Coordinate projection, 2. Detail reduction & defining domains and 3. Automated file format conversion.

From the original geographic coordinate system, roads were projected onto the VN2000/ UTM zone 48N using ArcMap. This projected coordinate system was chosen for its fitted bounds, which neatly cover both the Northern and Southern regions of Vietnam, thus minimising distortions during the projection process.

As discussed, ADMS has a limit of 3000 road features per model run. Because Hanoi's square has 7537 road features whilst HCMC's has 9363, a process that reduces these counts is required. The concept is to exploit ADMS's allowance of 50 vertices per road. Using various ArcMap's functions, connecting roads with the same name and type (e.g. 3 short tertiary roads), were merged into long lines. A Python script then iterated these lines and cut them if the length exceeds 50 vertices, producing new roads. In essence, this method simply finds segments of the same road and redefines them as one; it keeps all geographical details intact and does not introduce any distortion. Consequently, it would not have any effect on the resolution of the ADMS input. Since short roads are targeted, the process is particularly suitable for urban areas with densely packed road network. This is demonstrated by that whilst only 6.5% was deducted from Vietnam's entire road network, 15,5% and 23% reduction were the result for the Hanoi's and HCMC's square respectively.

A limitation of the above technique is that it relies on the attributions included within the shapefiles (i.e. road name), its efficiency would therefore vary if data other than OSM was used. Detail reduction can be improved, however it would requires more geometrical heavy processing. For example, a script can be built using ArcPy (ArcGIS's Python package created specifically to dynamically exploit built-in functions for the analysis of spatial data) to merge connecting roads, regardless of name, or to search parallel lines and fuse them into duel carriageways. This can potentially reduce over half the number of roads. However, considering the complexity of the road network at this scale, such a tool was overwhelmed for the current project, thus should be reserved to be developed in the future.

3.1.2.4 Domains

Since further detail reduction was too complicated and there were still over 6000 roads for Hanoi and 7000 for HCMC, the initial modelling boundaries were divided each into 5 domains. All domains contains less than 3000 roads. The smallest, yet most important domains situate in the middle of each square and are referred as the 'cores'. They were drawn as fitted overlay of the most densely populated districts of either city, i.e. 11 of Hanoi's districts and 13 of HCMC's (Figure 7). The cores would have great compatibility with data in Vietnam since, due to limited resource, most studies tend to target the main districts only. The remaining four domains shared edges with their respective core and modelling square so that minimum overlapping is introduced. This shows the consideration given to double counting when merging adjacent results.

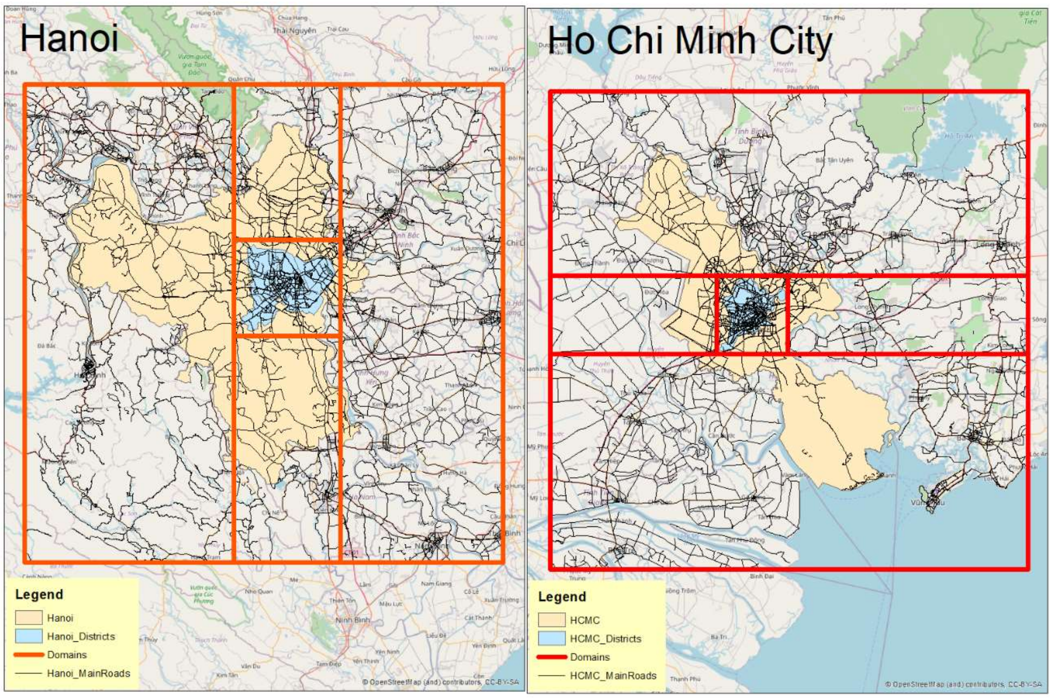


Figure 7. Hanoi and HCMC's main roads and domains.

The main difference between these two sets of domains are their orientation, in which Hanoi's square was divided into 3 main vertical pieces whilst HCMC's were horizontal. This configuration reflects the shift in economic function across the areas. In details, whilst the middle sections are the most developed, packed with industrials and residential establishments, Hanoi's East and HCMC's South have elevated growth due to the location of ports. This leaves the Hanoi's West and HCMC's North with mainly agricultural activities.

The population-based definition of boundaries (Section 3.1.1) and this orientation of domains were intended as to compare the overall pollution level between Hanoi and HCMC. However, with the current focus and data availability, these points could not be fully exploited. They therefore have limited impact in this work and are reserved for future uses as perhaps, with improved inputs, further evaluations can be made (e.g. model the emission dispersed from different economic function zones such as ports, agricultural or residential activities).

Lastly, an executable script was written using Python to systematically convert shapefiles into the required input format, i.e. iterating through the road network, listing the coordinate of all vertices of each road, and writing them into a .vgt file.

3.1.2.5 Road Parameters

Due to the scarcity of data, ADMS's basic street canyon function was employed for this study. As such, unlike the advanced canyon module, only road width, road height and average building height were required by the .spt file.

Road Width and Canyon Height:

The issue with sourcing data for road width and canyon height is that whilst regulations exist, Vietnamese policy enforcement is weak thus allows variations as exampled in Figure 8. Survey data is therefore preferred. However, other than the raw .osm files, none that contain road width parameter was found. Using QGIS, OSM data for Bangkok, Berlin, HCMC, London and Tokyo from OSM were studied to find any pattern in the number of lane per carriageway. The result was that excepts for Bangkok roads, other cities were quite similar: Most curves had a single spike at 2 lanes per carriageway and were slightly right skewed. London and Tokyo also suggested great symmetrical tendencies, resembling a bell shaped distribution.

Photo-1 Condition of NH21 Photo-2 Condition of NH12B

Figure 8. Two different national highway (JICA, 2010).

For this project's domains, the fluctuation in road condition was therefore assumed to be normally distributed. A Gaussian generator was incorporated in the Python script to assign each polyline, which often represents a single carriageway, with a random lane number. This number was produced using a mean value and standard deviation which, based on the road's type, were taken from the official construction guidance as shown in Table 3. Each lane number was then used to calculate the carriageway width and through Table 4, create a range, which allowed similar approach to be done to canyon height.

Table 3. Lane width and number of lanes (Ministry of Construction, 2007).

Road type		Designed speed (km/h)							Minimum lane no.	Desired lane no.
		100	80	70	60	50	40	30		
		Lane width								
Motorway (Trunk)		3.75			3.5				4	6-10
Urban main road	Primary		3.75		3.5				6	8-10
	Secondary			3.5					4	6-8
Collector (Tertiary)					3.5		3.25		2	4-6

Table 4. Construction height limit (Ho Chi Minh City People's Committee, 2009).

Road width - L (m)	Maximum height to the base of 1st floor (m)	Maximum construction height (m)					
		3 Stories	4 Stories	5 Stories	6 Stories	7 Stories	8 Stories
L ≥ 25	7.0	-	-	21.6	25.0	28.4	31.8
20 ≤ L < 25	7.0	-	-	21.6	25.0	28.4	31.8
12 ≤ L < 20	5.8	-	17.0	20.4	23.8	27.2	-
7 ≤ L < 12	5.8	-	17.0	20.4	23.8	-	-
3.5 ≤ L < 7	5.8	13.6	17.0	-	-	-	-
L < 3.5	5.8	11.6	-	-	-	-	-

Ultimately, the generation of these parameters all relied on OSM's definition of road type, of which quality is not guaranteed. It was critically the most rational method considering the current state of data. However, future works should source some records as to gain further assurance.

Some field surveyed road width data from HCMC were later acquired, however their quantity was limited, i.e., only contained 3 road types: primary, secondary and trunk; each type had less than 40 measurements in 2 different units (Bang, 2018). Compared to OSM, these new data showed slightly higher mean value of lane per road. This was expected since surveys tend to be held within the city core where roads are enlarged. Also, the distribution of secondary roads had 2 spikes perhaps because this class is bundled by both secondary and tertiary.

Bridges:

Bridges were easily identified and integrated into the input by using the shapefile's attribution called 'layer', which uses values from -5 to 5 to represent the levels of road elevation compared to the ground (i.e., layer = 0). A 5 meter vertical difference between each level was applied.

Tunnels:

Tunnels were not as straightforward since other than vertical depth of the bore, each required at least one connected road to be identified as the tunnel portal or outflow. Due to data limitation, tunnels were assumed to not have any other exhaust vents, other than the outflows.

To automatically identify outflows, couple script modules were added into the Python executable. They in turn performed: 1. Find tunnels (i.e. line features with layer value < 0). 2. Merge connecting tunnel lines (These separation points may serve as indicator for vents however they were ignored due to the above reason). 3. Find connecting roads. Identify inflow and outflow using OSM's attribution for traffic direction. 4. Discard special/ error cases (e.g. tunnels that have no attached line, or illogical traffic direction). 5. Write ADMS required input file. The process is also illustrated by Figure 9.

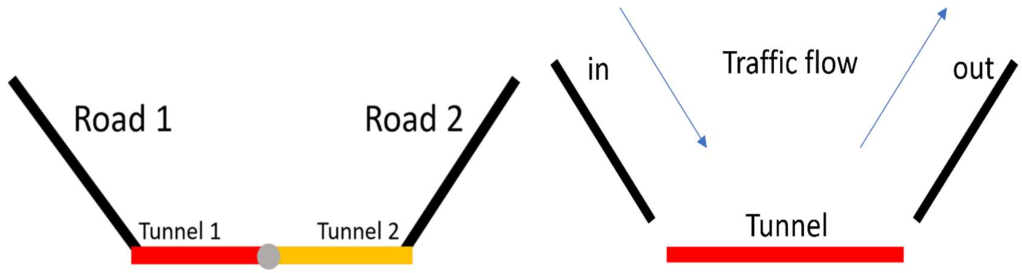


Figure 9. Identifying tunnel's outflows.

3.1.3 Emission Data

3.1.3.1 The Issues with Obtaining Emission Data

Emission data were critically the most challenging segment of the input requirement. This was because even though the data are available, they tend to only be accessible through physical copies. For example, whilst the transportation master plan by (JICA, 2010) was found online, the contributing vehicle data taken from (JICA, 2007) was digitally absent. As such, great amount of time was spent finding researchers or governmental units that can potentially agree to data sharing. This subsequently became one of the major objectives of the trip to Vietnam in May 2019.

There are other methods of estimating road emission factors that bypass the need of physical vehicle count. For example: extrapolating from fuel consumption or total travelling demand; automatic counting using camera, GPS or traffic stimulating models. However, due to poorly regulated vehicle qualities and chaotic traffic patterns, current studies in Vietnam face great difficulties applying those methods. This leaves emission data solely in the format of physical vehicle counts, which are usually collected by filming, playing back and manual counting. Evidently, they are very valuable and laborious. Inefficient data collection also causes cluttered manners of data sorting and storing. For instance, there were couple units, who confirmed to have multiple relevant urban transporting datasets (e.g. vehicles were counted to serve some particular campaigns); yet the staffs were unable to retrace those from piles of physical documents. Combining with the recognised mindset of holding back information, it was important to demonstrate the quality of this work and how suitable it is to Vietnam and developing countries, to encourage data support.

The data requesting scheme emphasised on the semi-automated processing of collated data. As such, this project would merge emission data from both city. The developed framework had the capacity to effectively manipulate large, differently formatted datasets. Also, with obtained records all averaged together, researchers might feel more secured when offering their data.

3.1.3.2 Processing of Vehicle Data

Ideally, each city would have a large vehicle dataset, dedicated models and means for validation. However whilst HCMC lacked in situ concentration data, Hanoi's vehicle counts were outdated. Critically, the acquired data were insufficient to respectively support each city's standalone evaluations. The vehicle data merge was therefore initiated to firstly filling the missing data gap; and secondary, to connect Hanoi and HCMC during model validation.

The emission input used in this work was the averaged result of four datasets including: 1. (Bang, 2018), 2. (Phuc & Oanh, 2018), 3. (JICA, 2010), 4. (JICA, 2007). Being the largest set, (Bang, 2018) contained close to 150 locations in HCMC. Number 2, 3 and 4 in turn contained 2, 35 and 10 roads. They were all conducted in Hanoi with 3 and 4 used in the construction of the urban transportation masterplans (JICA, 2010). Other datasets namely (Hung, 2010), (Trang, et al., 2015) and (Oanh, et al., 2008) had insightful information yet were eventually removed from the input due to sub-desired levels of detail. In overall, thanks to its large number of sampling points, traffic count from HCMC has a higher contribution toward this work's ADMS emission input. This may have an impact when evaluating the model output against Hanoi's observations.

These four datasets were chosen because they included hourly measurements for each road. Because each dataset categorised vehicle differently, some of which contained vehicle classes that are incompatible with ADMS's integrated format, traffic data were converted into a more conventional classification. This work therefore divided vehicle into 5 types: Motorcycle, Car Light (<12 seats), Car Heavy (>12 seats), Truck Light (<3.5t), and Truck Heavy (>3.5t).

Datasets were also made in different years: the oldest being (JICA, 2007) with data that dated back to 2004 whereas the latest was (Bang, 2018), counted in 2017. To close the time gap, a value of 3.8% was used as yearly increasement of total urban travelling demand (JICA, 2010). To break down this value into vehicle growth rates, the fleet's composition was assessed. Because Car Light (20% annual growth) is growing faster than MC (16%), the dominance of MC had gradually declined over the years (Figure 10). Still, changes were critically insignificant (<10%). Vehicle classes were therefore deemed to have consistent annual growths (Table 5).

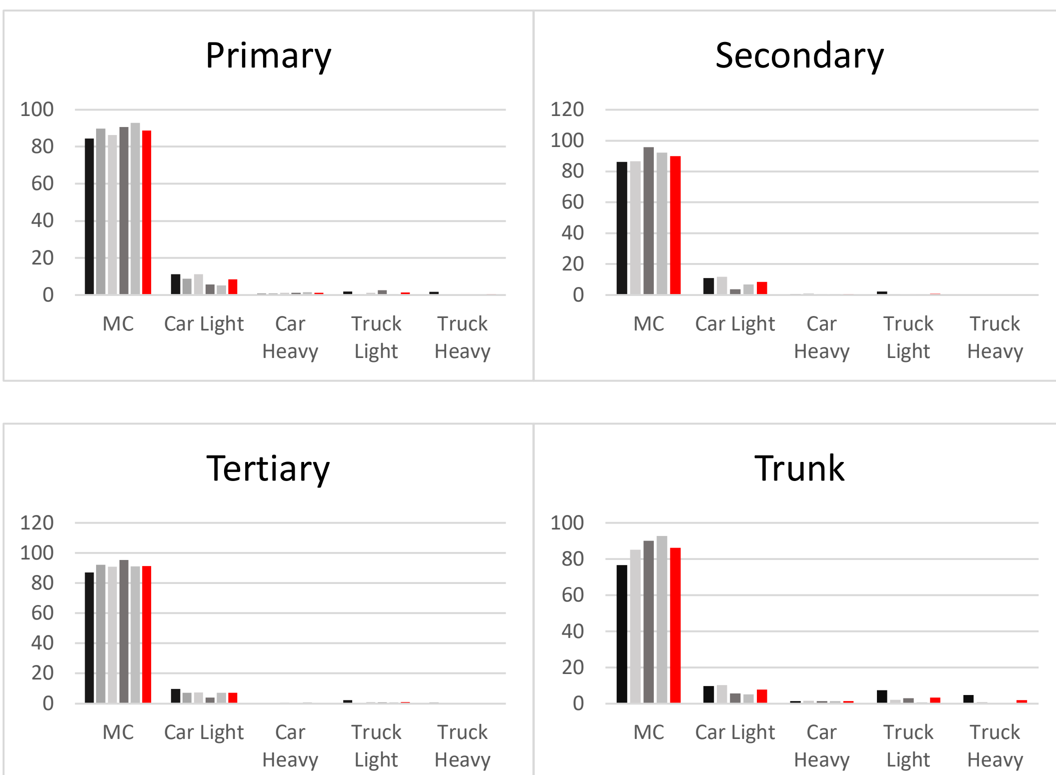


Figure 10. Vehicle composition in different road types and years.

Table 5. Calculated annual growth rate of each vehicle class.

	MC	Car Light	Car Heavy	Truck Light	Truck Heavy	Total
Averaged Vehicle Composition	89%	7.9%	0.9%	1.6%	0.7%	100.1%
Averaged Growth per Year	3.38%	0.3%	0.03%	0.06%	0.03%	3.8%

Once reformatted, the four datasets were used to produce Table 6, which shows this work's finalised ADMS vehicle input. Motorways shared identical data with Trunk as there was insufficient measurements. Each vehicle class had its own diurnal variation pattern. By counting vehicles repeatedly on only 2 sampling sites, (Phuc & Oanh, 2018) suggested that traffic rush hours on weekends had reduced magnitude and shifted slightly later. Due to data limitation (i.e. most roads were only observed for 1 day), traffic patterns were assumed to be indifferent between weekday and weekend. Consequently, model may overpredict emission at 7 and 18 o'clock on Saturday and Sunday.

A Python script used the average value and standard deviation (Table 6) to randomly assign each road with 5 hourly vehicle counts, which corresponded to 5 vehicle classes. According to Equation 1, these values were then multiplied with the hourly factor at a particular time to produce the number of vehicle at said time. For example, with primary roads having ≈ 5895 ± 2922 MC/hour (Table 6), the script therefore assigns 2 roads with 6000 and 5500 MC per hour respectively. The hourly factor at 1 o'clock is 0.09. These roads therefore have 540 and 495 MC at 1 o'clock. In overall, this process whilst retained a uniform diurnal pattern throughout the road network, still introduced a sense of randomisation which mimics the real-world situation. The drawback was that there may arise roads with extreme values, e.g. 1600 Heavy Truck/ hour.

Table 6. Finalised vehicle data.

Road Type	Primary					Secondary					Tertiary					Trunk				
Vehicle	MC	Car light	Car Heavy	Truck Light	Truck Heavy	MC	Car light	Car Heavy	Truck Light	Truck Heavy	MC	Car light	Car Heavy	Truck Light	Truck Heavy	MC	Car light	Car Heavy	Truck Light	Truck Heavy
Time																				
0	834	265	26	88	326	851	271	4	85	33	617	150	4	48	29	631	161	15	221	610
1	459	159	3	70	244	496	160	2	64	20	346	91	3	45	22	390	110	14	204	483
2	332	103	2	59	214	388	109	2	62	15	261	61	2	38	16	444	86	13	197	433
3	438	122	7	61	187	482	111	7	55	11	333	71	5	37	16	657	115	34	190	394
4	779	181	14	63	229	735	166	8	50	10	621	106	6	51	21	1048	156	38	188	483
5	1892	291	41	91	208	1937	284	32	82	15	1312	155	14	45	26	2147	260	85	236	450
6	5353	593	95	33	29	4817	497	62	31	6	2945	245	25	32	13	5663	536	166	248	135
7	11189	941	99	42	27	10530	680	55	33	7	5808	333	25	102	12	11425	830	155	233	97
8	11050	999	80	123	33	10368	820	48	130	7	6000	420	25	69	16	9118	846	123	440	120
9	8671	1022	78	204	37	8390	865	56	205	6	4685	417	24	102	16	7023	808	120	535	142
10	7847	1006	73	243	39	7501	897	47	218	6	4458	417	22	98	16	6571	789	114	607	162
11	7466	948	72	238	39	7771	825	47	228	10	4659	386	23	101	22	6220	706	109	589	170
12	6813	887	79	221	40	6788	830	48	212	10	3814	390	21	90	19	5540	684	112	556	181
13	5869	812	68	209	43	5835	706	42	196	12	3464	332	19	79	18	4909	671	104	529	162
14	6920	961	71	224	43	6738	868	42	208	10	3877	401	20	83	19	5595	684	104	578	177
15	7725	966	72	200	39	7126	859	44	198	12	4236	401	23	84	19	6053	744	106	562	178
16	8884	1009	76	70	32	7984	845	53	68	10	5265	415	25	60	17	6754	806	122	462	161
17	11120	1020	79	62	29	10162	844	48	65	8	6258	409	25	46	15	9766	826	144	347	148
18	9524	1001	75	36	21	8423	836	47	47	12	5325	406	25	42	18	8170	822	106	301	131
19	6874	869	56	37	29	6759	801	44	58	2	4148	375	19	46	14	6137	707	106	291	156
20	6879	835	50	87	35	6468	763	36	94	1	4042	413	16	62	13	5460	626	82	312	177
21	6033	776	30	85	29	4736	688	17	85	3	3373	394	11	52	21	4230	542	57	278	172
22	4142	654	23	86	70	3145	592	14	107	32	1956	295	10	53	19	3121	485	38	295	240
23	1865	420	9	80	96	1342	383	3	120	40	888	185	5	55	20	1312	269	14	253	274
Average	5895	714	56	114	84	5685	638	36	110	11	3465	302	17	61	16	4933	553	87	360	243
StDev	2922	509	41	131	380	2262	634	29	120	23	2530	495	18	91	42	2681	464	54	328	529

3.1.3.3 Emission Factors and Vehicle Speed

This work used the ADMS's built in Emission Factors Toolkit (EFT) v7.0 issues by UK Defra. The reason is that Vietnam has not been able to produce a similar conversion tool that calculates emission based on vehicle/ road condition and speed.

3.1.4 Meteorological Data

3.1.4.1 Data Sources

Meteorological data were downloaded from two sources: 1. Met Office Hadley Centre observations datasets (HadISD) for cloud cover and temperature and 2. The Global Forecast System (GFS) for wind speed and direction.

HadISD is based on the National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Information (NCEI) Integrated Surface Dataset (ISD), which consists of hourly synoptic climatological observations from over 35,000 stations worldwide. The advantages of using HadISD are namely global coverage, uniform format, long duration and daily updating frequency. However the dataset is susceptive to instrumental variations. For example, 2 HadISD's stations used in this project record wind at different heights: 11.9m and 10.1m. HadISD has its own quality control code, which encourages users' evaluation based on individual purposes. Within the scope of this project, it was most suitable to use HadISD as an archive of readily available meteorological data.

GFS is a weather forecast model produced by the NOAA's National Centers for Environmental Prediction (NCEP). Being a simulated, gridded wind dataset instead of observation, GFS offers no chronologically missing record, uniform quality and improved spatial definition. As such, it allows each domain to have a respective wind condition. The version of GFS used in this project is surface wind (i.e., at 10m of elevation) and has a resolution of 0.5 degree (approx. 50km) (Fernandez-Lopez & Schliep, 2019).

3.1.4.2 Data Processing

HadISD:

A HadISD's station were chosen for each city: Noi Bai International Airport for Hanoi and Tan Son Nhat International Airport for HCMC. Future studies may benefit from that several other Vietnamese cities also have their corresponding stations.

Downloaded from the Metoffice, HadISD files were fed into RStudio to be extracted to more conventional formats. Since there were missing data rows, it was important to fill these chronologically with blank values. Else, once merged with GFS winds, there would be time offsets and eventually errors. ADMS will automatically skip meteorological lines that have inadequate parameters.

GFS:

Data were obtained using RStudio equipped with rWind, which is a R package designed to download and manage GFS wind data (Fernandez-Lopez & Schliep, 2019). rWind only downloads GFS at 0.5 degree resolution, even though other levels of resolution were found from the NCEP archive (i.e. 1 and 0.25 degree).

Approx. 7 years of wind data (i.e. from 2013 to the end of September 2019) were downloaded for each of the 10 domains.

Each domain's winds were calculated as the average of its contained GFS grid points. This was done using the wind's U and V component. Once calculated, the averaged U and V were fed into the followings:

$$\text{Speed} = \sqrt{\mathrm{u}^2 + \mathrm{v}^2}$$

$$\text{Direction} = \alpha = \frac{180}{\pi} \times \text{atan2}\frac{u}{v}$$

Equation 3. Calculation of average wind from u and v components

When applied, the second equation should be checked carefully since different math packages have their arctan2 functions perform differently. For instance:

$$\text{Excel:} \qquad \alpha = \text{Degrees}(\text{atan2}(V,U))$$

$$\text{Python:} \qquad \alpha = \text{Degrees}(\text{atan2}(U,V))$$

Equation 4. Difference between Excel and Python in calculating Atan

Note that this direction remained as the wind vector azimuth (i.e. winds blowing toward with an angle of α) instead of the meteorological wind direction (i.e. winds blowing from an angle of β), which is the desired ADMS's input. The finalised direction was then calculated as:

$$\beta = \alpha + 180$$

Equation 5. Calculation of meteorological wind direction from wind vector azimuth

The entire process consisting of grouping, averaging U, V and calculating wind speed and direction was computed by a dedicated Python script. Higher levels of wind detail (e.g. GFS at 0.25 degree resolution) would logically enhance the definition of each domain's wind and help better segregate them. However, the options are currently unavailable.

3.2 Concentration Data, Modelling Experiments and Layers

3.2.1 In Situ Concentration Data

Being the primary method of model validation, in situ concentration data and the potential difficulties in acquiring them had been considered across the duration of this project.

Cranfield University deployed some CO sensors in both cities: 2 stations in Hanoi and 3 in HCMC. These were operated for 6 months: from October 2018 to April 2019.

With the understanding that Vietnam had performed multiple air pollution monitoring projects, it was suggested to be further obtained concentration data by contacting related personnel and proposing data sharing requests. As a result, another 2 datasets from Hanoi were sourced.

The first set was from the Center for Environment Monitoring (CEM), Vietnam Environment Administration (VEA), located at Nguyen Van Cu Street, Hanoi (from here onward referred as NVC data). It contained hourly measurements of 9 commonly studied pollutant species and covered 5 years, from 2010 to 2015. Meteorological records were also included. Produced by one of the first government owned continuous monitoring stations, numerous studies have employed this dataset, including multiple national environmental evaluations (MONRE, 2016). However, similar to many older Vietnamese projects, related documentations and reports tend to be strictly kept for internal circulation. Therefore, whilst the dataset is valuable, extra care was given since the overall quality could not be critically evaluated.

The second set was from a network of 10 static continuous stations. Full name and assigned acronym of stations were presented in Table 7. According to Nam, et.al. (2018), which was among few accessible reports regarding these stations, TY and MK employ Environment S.A's equipment and standardised methodology for high accuracy observations. The remaining 8 are equipped with small, relatively low cost CAirPol's Cairsens sensors. They use TY and MK as reference. The exact procedure was not documented however considering that TY locates on the 6th floor whilst others may only be built on sample masts with less than 10m of elevation, the dataset critically does not have the highest quality. The obtained data were from May 2017 to May 2019.

Table 7. Name, measured pollutants, and data duration of monitoring stations.

ID	Measured pollutant						Duration						
	PM_{10}	$PM_{2.5}$	CO	NO_2	SO_2	O_3	2013	2014	2015	2016	2017	2018	2019
TY													
MK													
HD													
HK													
KL													
MD													
PVD													
TC													
TM1													
TM2													
NVC													
132													
151													
138													
145													
146													

In overall, both datasets showed the lack of associated documentation. Still, with the country's limited resources allocated for environment management, they remained the best that could be offered. As such, out of 12 available monitoring stations in Hanoi, listed by MONRE(2016), this work obtained data from 11.

3.2.2 Modelling Experiments

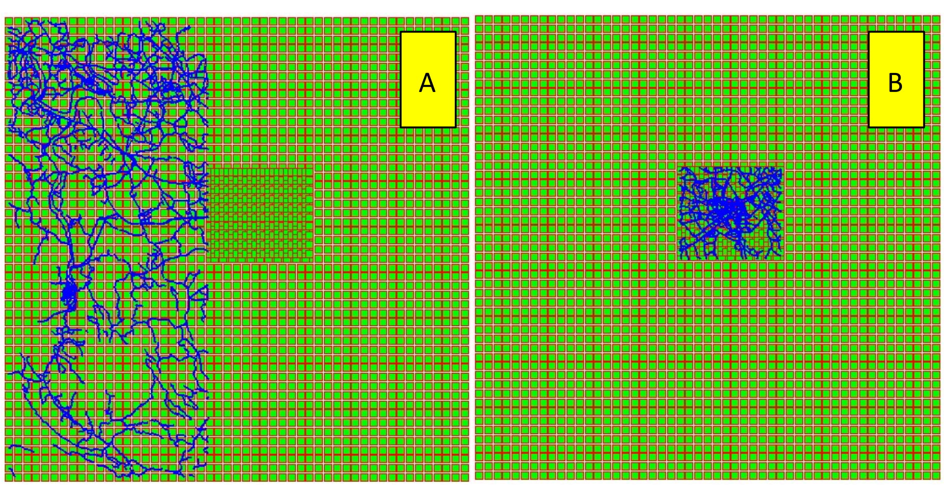
Once all data were collected, it became clear that the available datasets are too centralised in the cities' core that except for the main districts, those surrounding domains would not have any data to be used for model validation. The idea of having two experiments therefore was initialised.

3.2.2.1 Experiment 1

The first one focuses on model validation by trying to engage as much of the obtained data as possible. As such, it simulates the dispersion of pollutants in the core districts only. Models in Experiment 1 also used meteorological data of longer duration (7 years, from 2013 to 2019) to improve statistical significances and to match the span of available in-situ observations,

3.2.2.2 Experiment 2

On the other hand, the second experiment simulates all 5 domains separately as nested domains (Figure 11). The results of suburban domains were layered on top of the core's, i.e. total concentration at core = core + West + East + North + South. This would allow conducting a coarse assessment of the contribution of long-distance travelling to concentration at the cores. Similarly, Experiment 2 would indicate how effective long distance modelling is, considering the current data scarcity. To save running time however, models in Experiment 2 only used 2 years' worth of meteorological data (from 2017 to 2019).



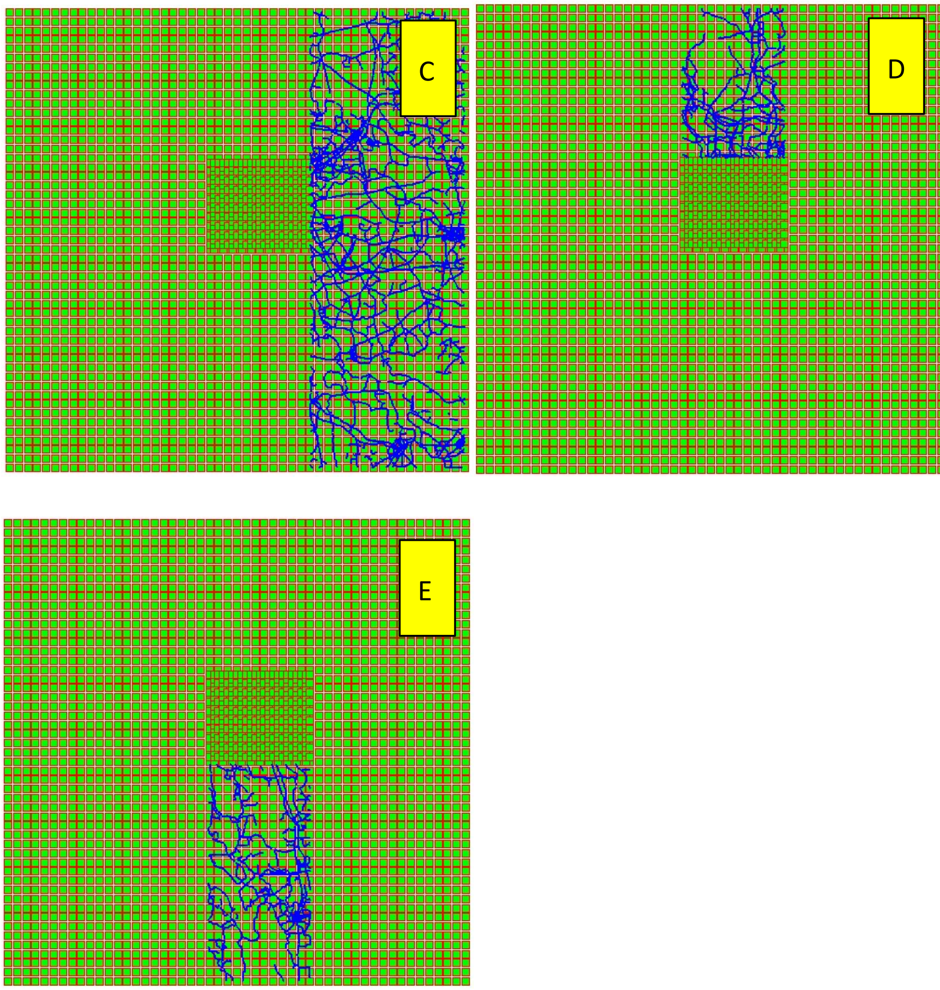


Figure 11. Nested domains used as layers in Experiment 2: A – West; B – Core; C – East; D – North; E – South.

To help compare experiments' result, both Experiment 1 and 2 were inputted with identical road data and were set to output at the same resolution and grid points. However, wind data for Experiment 1 were GFS averaged for the district domain only, whereas Experiment 2 used winds averaged for the entire square of 120x120km. The impact of this difference should be minimal and is briefly discussed in Section 4.3.2.1.

3.2.3 Modelling Layers

With vehicle classes exhibiting various diurnal variation patterns, e.g. high influx of Heavy Truck in night-time, similar to which found in Beijing (Biggart, et al., 2019), it was deemed interesting to investigate the emission contribution of each class explicitly. As such, each vehicle class was inputted into a separated model, containing an individual hourly factor profile (.fac file). The total simulated concentration would be the sum of all 5 layering models, i.e. MC + Car Light + Car Heavy + Truck Light + Truck Heavy. For Experiment 2 in particular (Section 3.2.2.2), this would mean the total concentration being the sum of 25 layers (i.e. 5 nested domains, each contains 5 vehicle layers).

3.3 Background Pollution

As discussed above, background concentration is essential to the model as it help validating the accuracy of prediction. ADMS would prefer background information inputted along with the data during the modelling. However, since this project would have multiple models laying on each other to calculate the sum modelled concentration (Section 3.2.2 and 3.2.3), background data would be doubled counted if add directly to the model. As a result, background pollution was added at the end, which was part of the reason why chemical reactions could not be included in this work.

Background data were sourced from a recent monitoring campaign in Ha Giang, a province Northern of Hanoi (So Tai Nguyen va Moi Truong, 2019). They claimed that measurements were made outside of the traffic busy hours and the average values for each of their 34 stations were published. The background concentrations used in work are: 35.15μg/m^3 for $PM_{2.5}$, 49.1μg/m^3 for PM_{10} and 1940.41μg/m^3 for CO.

4 Result and Discussion

This chapter firstly looks into the obtained data in order to gain a better understand of the real world emission and how well the model input was able to depict it. The result of in situ concentration observations is also briefed. Such evaluations are then used to guide the model analysis and validation. Openair, an R package designed specifically for performance evaluation of dispersion models, was employed to compare modelled results against observations.

Due to the time and data limitation, even though the dispersion of 5 pollutants including NO_x, VOC, $PM_{2.5}$, PM_{10} and CO were modelled, only CO results would be discussed. Not only that CO measurements are available for all stations (Table 7), it is also the most chemically stable species among those modelled. This decision was made under the consideration that the project could not include chemical reactions, which have been discussed in (Section 2.6).

4.1 Input Data

Since Hanoi and HCMC have different socio-economic functions and travelling demands, merging vehicle data is expected to inherit certain drawbacks. For instance, Hanoi's short-distance travelling need, which is mostly described by vehicle count for Secondary and Tertiary roads, may increase. As a result, concentrations at points closer to the city's core may be overestimated and vice versa for HCMC.

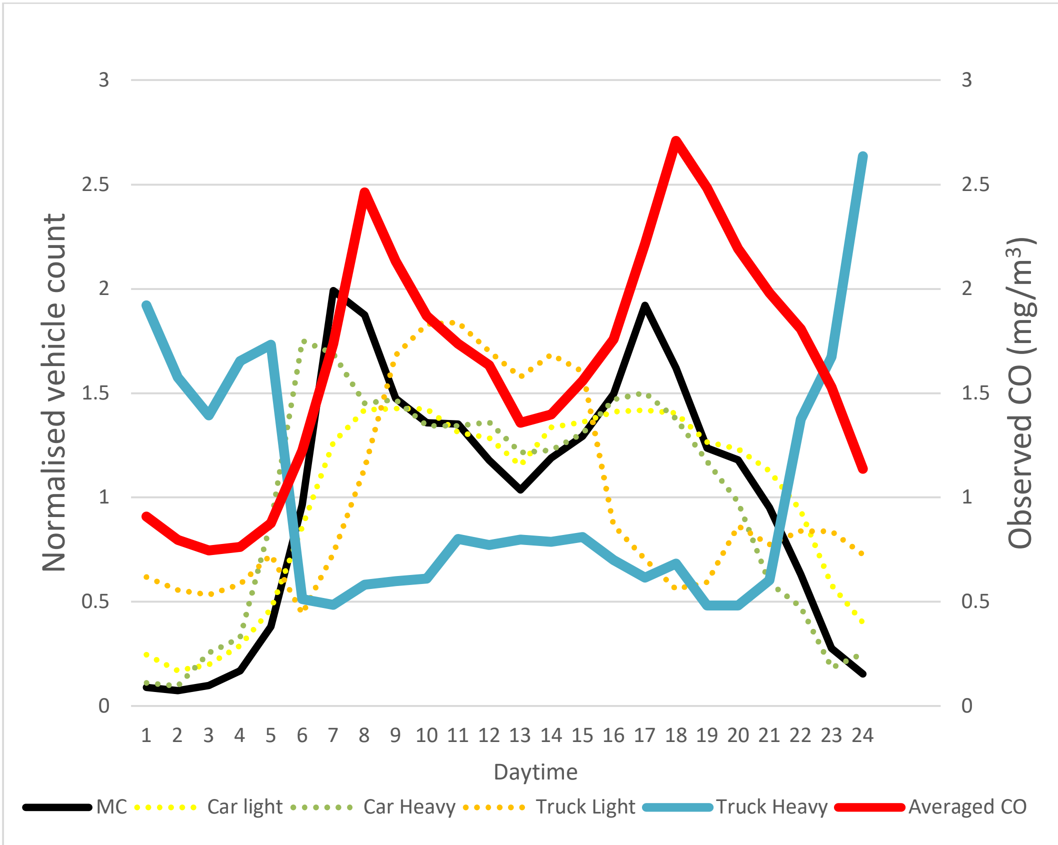


Figure 12. Diurnal variation of vehicle classes and averaged observed CO in Hanoi.

From Figure 12, all travelling patterns can be seen revolving around two rush periods at around 6 and 17. In detail, MC exhibits the sharpest peaks. This is because MC remains as the dominant vehicle fleet and is superior in short distance trips. Heavy Cars, most likely urban buses, have to follow their schedule. Their pattern is thus similar to which of MC but is also less drastic. Light Car avoids congestion by spreading its activities throughout the day. Regarding to logistics, Light Truck operates mainly in the mid-day calm period (around 9 and 15) whereas Heavy Truck works almost exclusively during the night. This unique inverted variation pattern is created because lorries are banned from running in urban areas during sensitive hours (i.e. susceptive to congestion). Lunch break around 11 and 14 are clearly visible with private and semi-private vehicles (i.e. MC and Light Car), whilst public Heavy Car has much smoother decline. All vehicle classes also display a slight increase from 23 to 3. This may be an indication of the long-distance travelling needs.

Aligned with the wealth growth, instead of buying new MCs, people seem to prefer cars and logistic vehicles. As such, increased demand for other vehicles has caused the dominance of the MC fleet to gradually decline. Similar trend has been observed in China where car overpopulation is currently an issue. Even though this work's input ignored the shift in vehicle composition, plotting data for each year might provide some insights. Additionally, from an urban management standpoint, this is an interesting topic for future works.

For Hanoi, the wind data used for the modelling (i.e. GFS) is compared against observations from NVC, TY and MK (Figure 13). Whilst the differences during summer and autumn are negligible, in winter and spring, northern winds in GFS present much more frequently and have higher velocity. Unfortunately HCMC does not have any other wind data for reference.

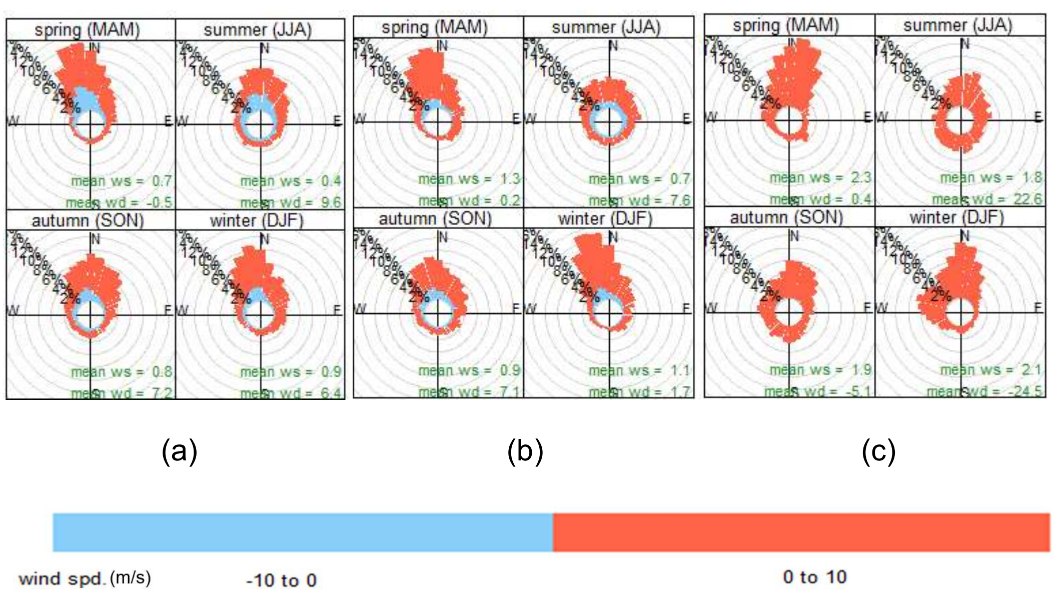


Figure 13. Frequency of counts by wind direction (%): GFS wind at Hanoi core compared to wind at: A – NVC, B – TY and C – MK.

4.2 In Situ Concentration Data

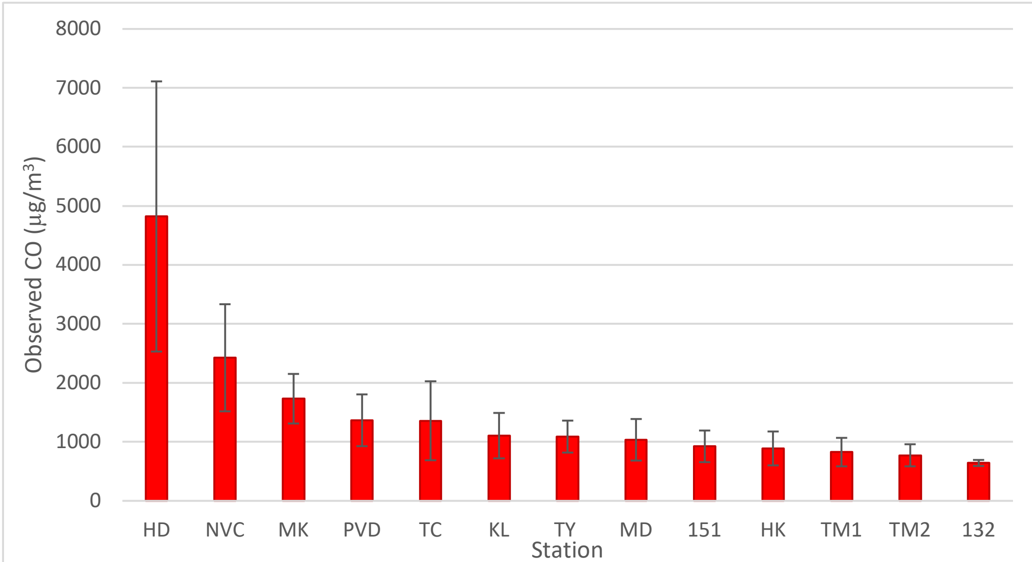


Figure 14. Mean and standard deviation of CO observation at monitoring stations in Hanoi.

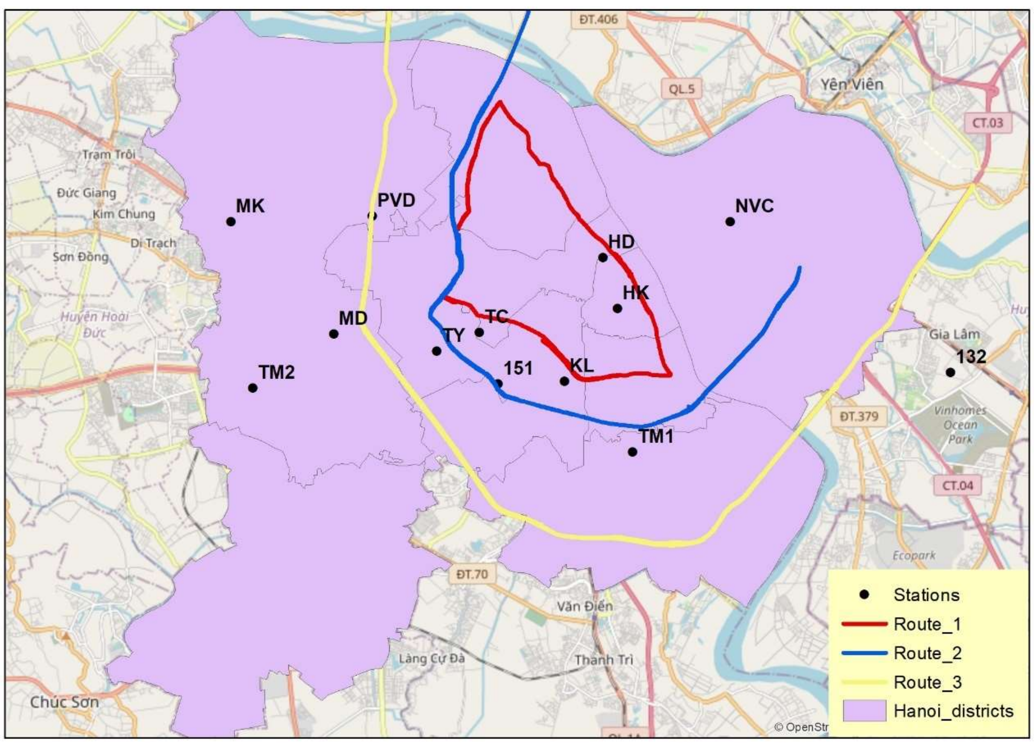


Figure 15. Location of monitoring stations in Hanoi in relation with the city's zones of development.

Figure 12 shows close correlations between the diurnal variation of MC and measured CO concentration. This suggests that MC is the main contributor of traffic derived pollution. With free flow traffic, pollutants are dispersed considerably easier thanks to vehicle induced turbulence. During peak times however, congestion allows emission to accumulate, slowing its release to the surrounding areas. The slight lags between MC and observations at 8 and 18 can be an indication of this effect. The increased concentrations during the night may originate from other vehicle classes, most likely Light and Heavy Truck.

As for measurements at each station, HD has distinctively higher CO observations compared to the remaining 12 stations (i.e. 1.6 to 4.3 times higher than the average). HD is followed, with a large margin, by NVC, MK, PVD and TC (Figure 14). This roughly aligns with Hanoi's urban expansion strategy in which there are three concentric layers of development. The first, most inner ring being the Old Quarter, contains HD and HK and is densely occupied by tourism businesses (Figure 15). Because of its historical value, this zone restricts most infrastructure developments and focuses on services. It is therefore suspected that elevated CO concentrations at HD is a result of extreme MC traffic and domestic activities, most likely cooking using coal briquettes, which is widely referred as the second major emission source in urban areas of Vietnam (MONRE, 2013). The second, marked by the circular route next to TY, TC, KL and NVC, has the fastest population growth rate (JICA, 2007). It therefore allows small developments, mostly for domestic uses. PVD and MD situate next to the circular route that bounds the third zone, where the construction of skyscrapers and industries is encouraged.

In general, the obtained measured data agree with (Hien, et al., 2020), which found pollution level (albeit not CO) decreased from the urban centre outward and concentrations spiked at around 9km away from HD due to the newly established manufacturing facilities and their resident labours, e.g. the areas containing MK. Transportation remains as the dominant source, however industrial processing and domestic cooking can have a notable contribution, e.g. 15% CO, 70% SO_x and 50% NO_x (Hoang, et al., 2017). Similarly, (Truc & Oanh,

2007) suggested that dinner time may intensify the pollution level during evening peaks as it collides with the traffic rush hour.

Other than HD and NVC, measured CO from other stations are very comparable, i.e. peak at around 2mg/m^3. Similar observations were also found in (Sakamoto, et al., 2018), in which CO concentration reaches 2ppm (2.3mg/m^3) at 8:00 and 18:00, and drops to 0.8ppm (0.93mg/m^3) at 3:00. These values are well under the Vietnamese regulatory standard for ambient CO (i.e. hourly average = 10mg/m^3). This aligns with official findings that CO and SO_2 are not yet a notable concern for the cities, whilst NO_x and PMs should be the focus (MONRE, 2013) (MONRE, 2016). Those with the lowest concentrations are Cranfield's 132, TM1 and TM2 (Figure 14). No clear explanation could be given to low CO concentrations at HK.

(Hoang, et al., 2017) and (Bang, et al., 2019) estimated the CO emission of anthropogenic sources other than road-traffic, including thermoelectricity, industrial production, services, and domestic activities, to be 15% and 1.2% respectively. Combined with another 40% originated from natural processes (World Health Organization, 2000), the background CO concentration in Hanoi, as estimated using the obtained measured data, would be around 800µg/m^3 (based on (Hoang, et al., 2017)) and 600µg/m^3 (based on (Bang, et al., 2019)). As such, both values are fairly close to CO concentration shown at station 132 as well as to the hourly minimum concentration (represented by the red coloured lower bound in Figure 19). For instance, station 132, with its remote location, exhibits a very consistent CO level throughout the day at 642.7 ± 51.4µg/m^3, whereas the hourly minimum observation is 602.2 ± 59.5µg/m^3. It is therefore believed that either dataset would be a great fit for background pollution in future models. Additionally, the use of minimum concentration as background pollution in ADMS has been carried out by (Biggart, et al., 2019), with very positive results.

4.3 Model Result and Validation

4.3.1 Experiment 1

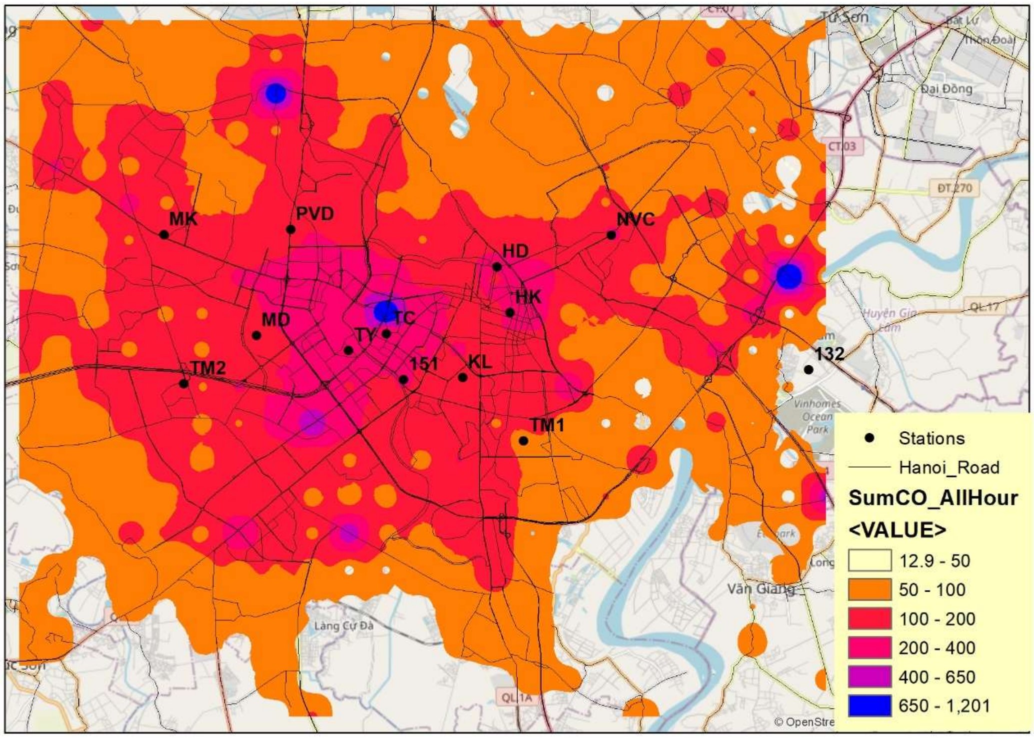


Figure 16. Long term contour of CO concentration in Hanoi core, without background pollution.

As shown in Figure 16, the long-term averaged CO concentration in Hanoi core is simulated to reach 1.2mg/m^3. There are three hotspots with over 1mg/m^3, which are navy blue coloured, they are all located at junctions between two or more large roads (i.e. Motorways, Trunks and Primaries). During evening peak time (18 o'clock), concentration at those spots exceed 2mg/m^3.

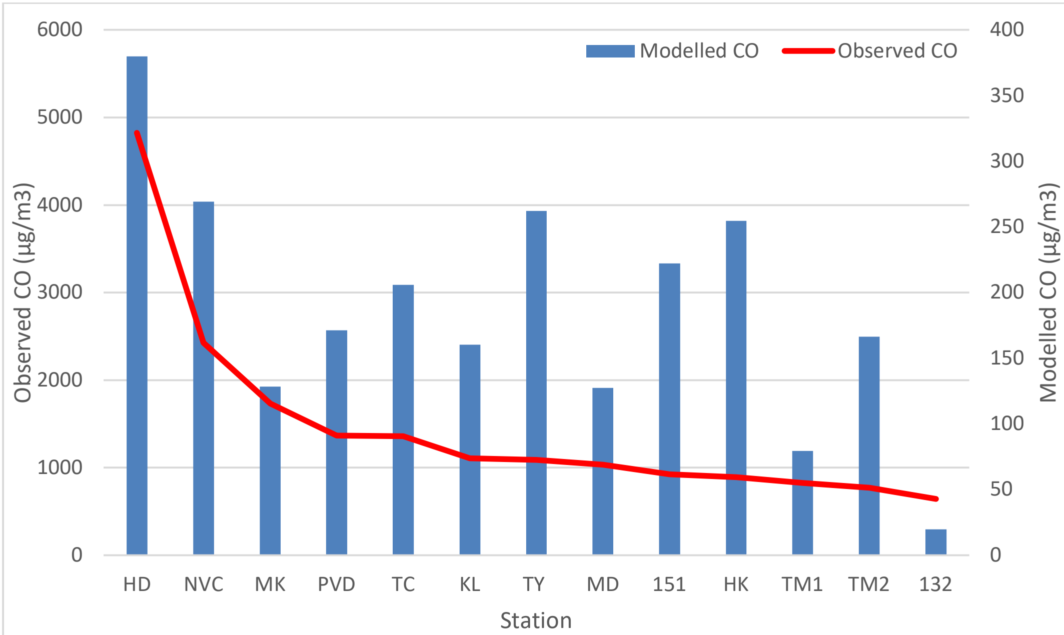


Figure 17. Mean modelled CO concentration, without background pollution, compared to mean observed CO, as plotted by monitoring sites in Hanoi.

Similar to observation, HD and NVC exhibit the highest averaged modelled result, whilst TM1 and 132 have the lowest (Figure 17), which is rational considering the road network (Figure 16). The distribution of modelled emission also follows Hanoi's development pattern (i.e. from the city centre decreasing outward). There are differences between observed and modelled CO concentration that suggest the contribution of other emission sources namely natural, industrial, and domestic, e.g. MK, from the 3rd highest, drops to the 4th lowest. Except for HD, NVC and MK, which might need special treatment due to their surroundings, the delta between observed and modelled CO is around 800µg/m^3. As such, the currently chosen background concentration (i.e. CO = 1940.41 µg/m^3) is likely to cause overestimations. Contrarily, as discussed in Section 4.2, the use of station 132's measurements or hourly minimum concentration would be a better fit for future comparisons.

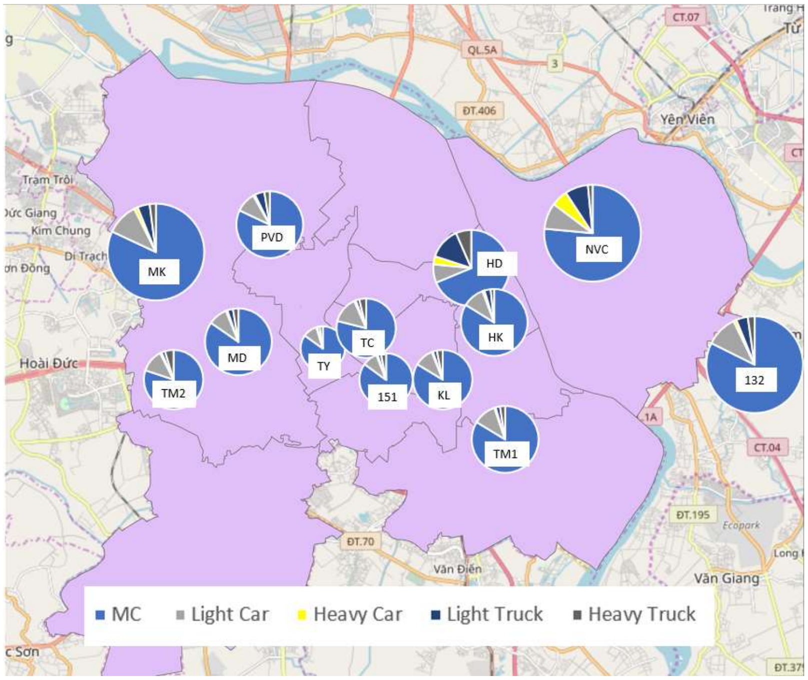


Figure 18. Contribution of each vehicle type in the modelled CO concentrations at stations in Hanoi.

As anticipated, MC takes the largest share of the modelled CO concentration at all stations (Figure 18). The contribution of Heavy Car and Light Truck at HD and NVC is significantly higher compared to which at the remaining stations. Whilst this result is plausible for NVC as it situates very close to a major Trunk route, and the Old Quarter has many bus hubs, it is odd for emissions of Light Truck to be high at HD. It is therefore suspected that the random generating process (Section 3.1.3.2) could have allocated some roads near HD with extreme counts of Light Truck, thus enhances CO emissions at this station. Reinvestigating the input data shows that HD is surrounded with several roads containing 3 times the Light Truck mean (e.g. tertiary roads, each with 180 – 210 trucks/hour). As a result, future works must find a method to refine the vehicle random generation, e.g. cap the vehicle count at certain values depending on the city's development zone.

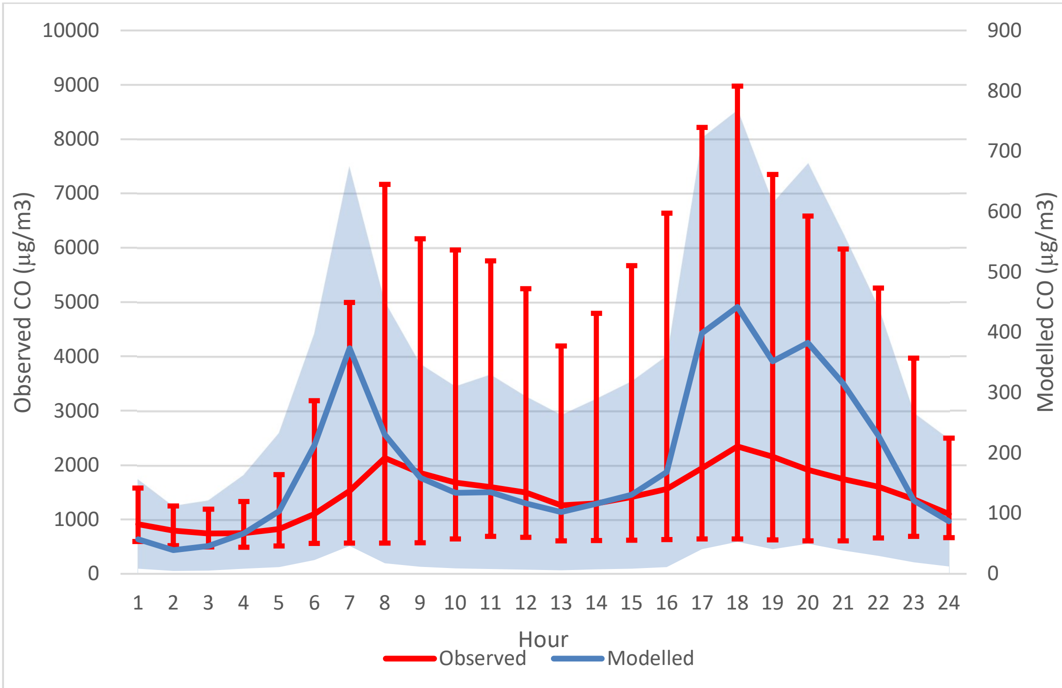


Figure 19. Mean, maximum and minimum of modelled and observed CO concentration at monitoring stations in Hanoi, as plotted by daytime.

The diurnal variation of modelled concentration follows the pattern of MC, in which there are 2 main peaks as well as high concentrations during the evening (Figure 19). There is also a third peak at 20, probably due to the increase in number of MC and Light Truck. Compared to in situ observations, the evening peak at 18 o'clock suggests that the model has simulated the effect of congestion. High concentrations during the night are however not as noticeable with the modelled outcome.

According to Figure 20, the modelled result has an overall weak to moderate correlation with all sensors in Hanoi for CO. However, this relationship stays constant except for Light and Heavy Truck. As a result, instead of rejecting the model completely, this indicates that there remain many background variables which need to be evaluated. There are not many differences between stations of the Vietnamese monitoring network, suggesting they have indeed been improved through centralised communication.

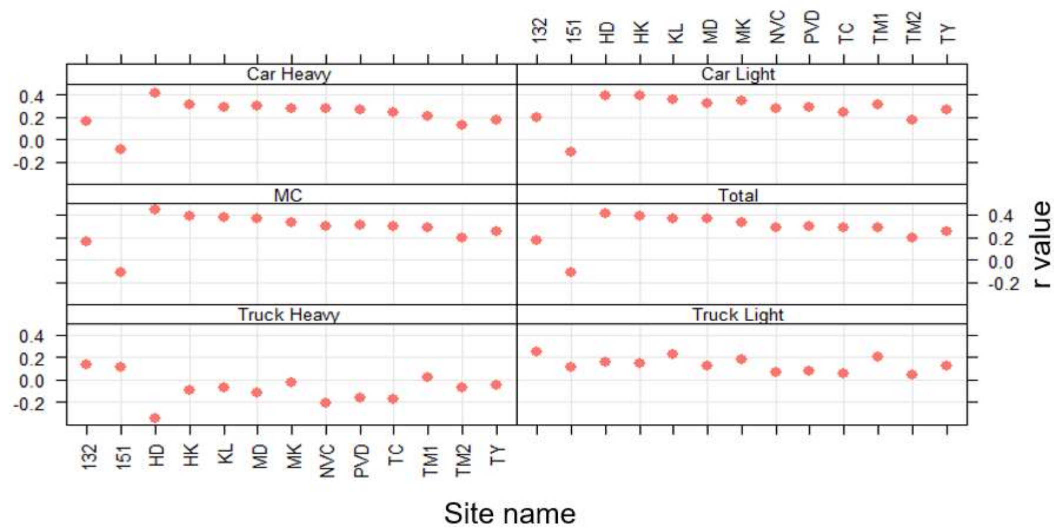


Figure 20. r value between modelled and observed CO, plotted by monitoring sites in Hanoi (see Figure 15).

HD has the highest correlation with all vehicle classes except trucks, both light and heavy. This confirms the above discussion regarding HD being in the Old Quarter thus its elevated CO concentration is caused by extreme traffic but strictly not from duty vehicles.

Station 151 has the lowest r value for the main vehicle classes (e.g. MC, cars, buses) yet the relationship improves with trucks. No obvious reason is suitable to explain this result. However, perhaps because 151 is placed in the Le Hong Phong political school, i.e. an institute only teaches government officials with a relatively large campus, the observed concentrations are already a poor representation of the actual road-side emission. As for trucks, both light and heavy, their diurnal variations specifically target non-congested periods, i.e. lunchtime and late evening. It is thus possible that r value at 151 is enhanced simply because truck's schedule collide with the peaks of domestic activities.

Looking into the stations' location, the closer into Hanoi's core, the more r appears to increase. For instance, HD has an overall r of 0.41, which is followed by HK (0.39) and KL (0.37). Even though this increasement is very slight and is not completely linear, it statistically helps the above discussion regarding the relationship between emission and Hanoi's pattern of expansion. Similarly, these values suggest that further geographical profiling can be done to Hanoi to improve the prediction of traffic derived emission.

The contribution of Heavy Truck to modelled CO concentration is very insignificant, ranging between 1.5 to 6.5% (Figure 18). (Bang, et al., 2019) even estimated Heavy Truck's contribution to be lower, at 0.11% of the total CO emission in HCMC. It is therefore very difficult to evaluate the impact brought by changes in Heavy Truck's emission. Subsequently, it explains why this vehicle class, with the least vehicle count, exhibits almost no correlation with observation, as plotted by monitoring site (Figure 20). Evaluating the contribution of lorries thus requires the simulation of diurnal variation.

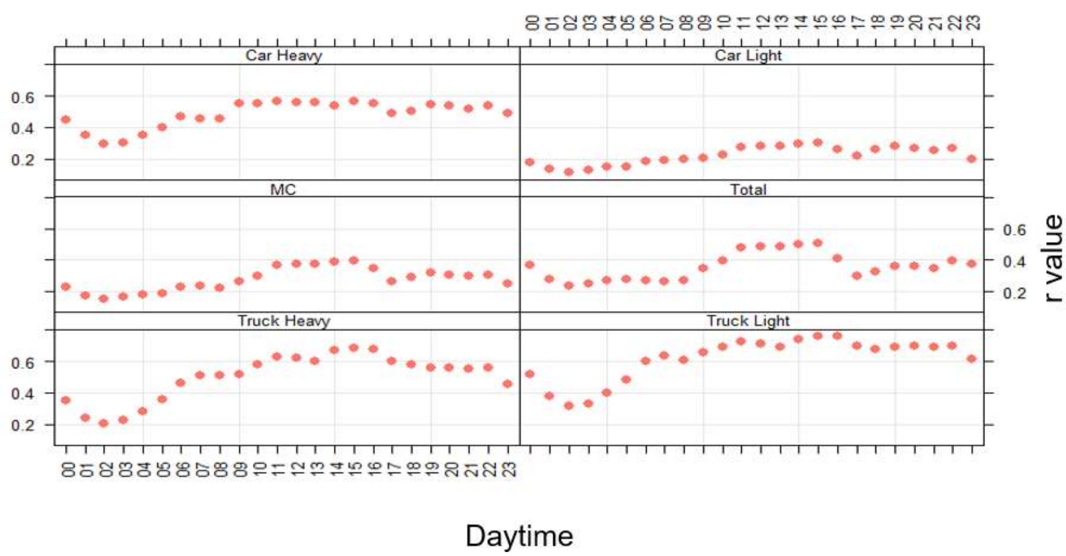


Figure 21. r value between modelled and observed CO in Hanoi, plotted by daytime.

Prediction for large vehicles (i.e. Light and Heavy Trucks and Heavy Cars) performs well (i.e. r > 0.5), especially after the morning rush (Figure 21). The overall correlation is however reduced by Light Car and MC. The lowest r values are observed at 2 o'clock, meaning whilst the model suggests reduced vehicle count and emission, observed concentrations stay mostly unchanged (Figure 19) shows a consistent measured CO concentration from 1:00 to 5:00. This indicates that concentration during the night is being maintained by the underlying conditions uninfluenced by the modelled sources (Section 4.2).

As for the shift in vehicle fleet composition, Figure 22 shows an overall negligible difference through the years. By adding a yearly increment rate of 0.3% and 3.38% for Light Car and MC respectively (Section 3.1.3), the correlation gradually drops. Additionally, whilst the city's development and economic growth can be seen through the r value of Heavy Car and Light Truck, personal vehicles have a rather neutral relationship. This suggests a very slow transition from MC to other vehicle classes in the urban core, contrasting to which estimated by previous studies.

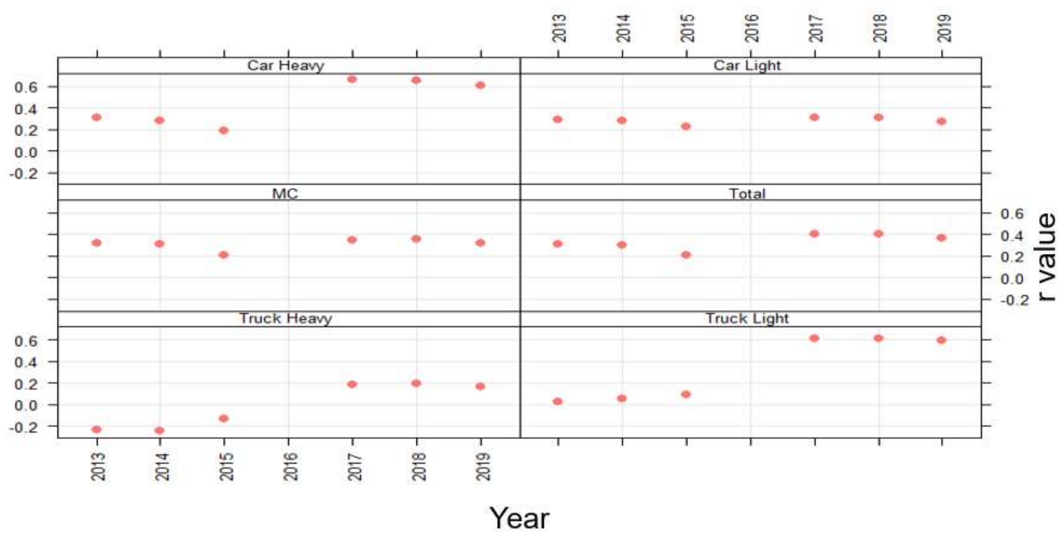


Figure 22. r value between modelled and observed CO in Hanoi, plotted by year.

For seasonal variation, as suggested by the above wind roses, summer has the highest correlation with a r value of 0.214, followed by autumn, winter and spring.

However, the differences in correlation are extremely minimum. In this project, seasons therefore should not have any significant impact upon the concentration result.

The prediction error is now evaluated. Without the addition of a fixed background concentration, underestimation can clearly be seen from all stations (Figure 17). Through a conditional quantile plot, which splits the data into evenly spaced bins of value and shows the Median (red line), 25/75th and 10/90th percentile for each bin (Figure 23), it is apparent that a value of 1940.41 µg/m^3 for background CO is however too high and would cause overestimation. MK and NVC for example, show very good agreement with observation during peak times and in the evening (i.e. from 18 to 23) (Figure 24). But because their result reaches minimum at the background value, one third of the modelled concentrations (i.e. during the night and early morning from 0 to 6) are almost guaranteed to be too high.

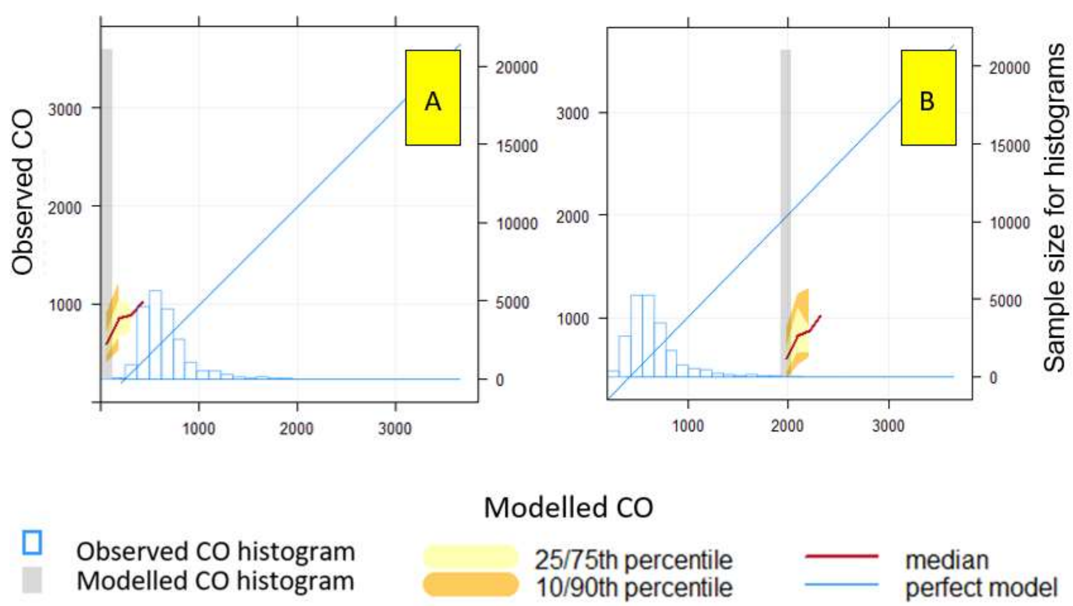


Figure 23. Conditional quantile plot for hourly CO at 132, without (A) and with (B) background.

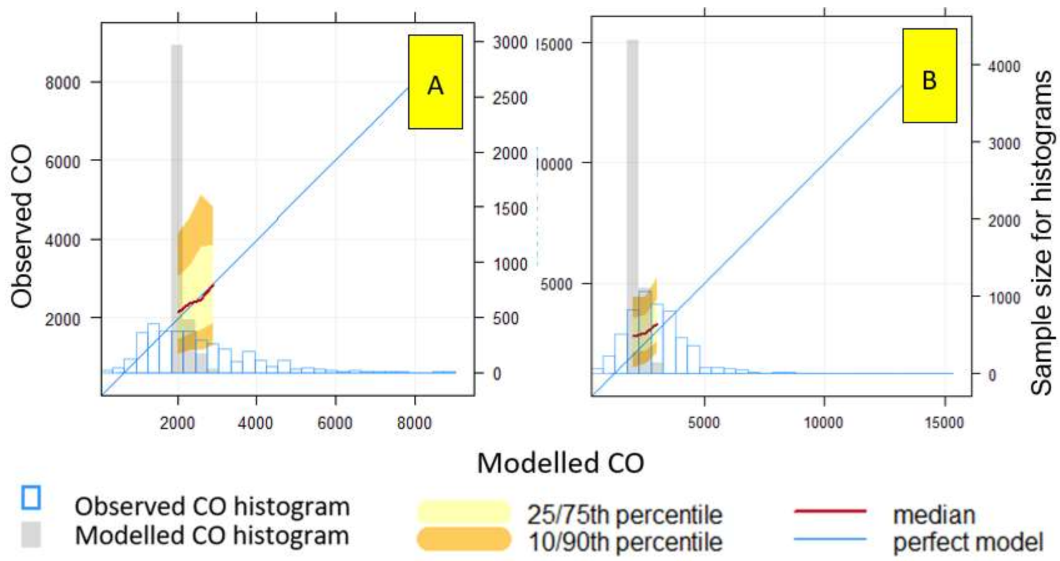


Figure 24. Conditional quantile plot for hourly CO at: A – MK; B – NVC, at 18 o'clock (evening peak), with background.

Looking into other statistics, when compared against TM1 (i.e. the lowest observed set of concentrations), the fraction of predictions within a factor of two (FAC2) of model with background CO is only 0.09. FAC2 for other stations ranges from 0.25 (TM1) to 0.77 (MK). This is very plausible, considering that 132 is located in the suburban area. Similarly, result at 132 has the highest mean bias (MB) of 1256 μg/m^3, i.e. the most overpredicted, whilst HD (-2750 μg/m^3) and NVC (-260 μg/m^3) till manage to achieve negative MB, i.e. underestimated. It is also worth mentioning that NVC situates closely to a military airport, which can potentially have an impact on this comparison. MK appears to be the best fit with a MB of only 231 μg/m^3. These values suggest that for a city as diverse and rapid changing as Hanoi, the use of one background concentration dataset alone is not enough. Similar conclusion has been made in Biggart, et.al. (2019) for Beijing.

In overall, even with a fixed background value, good agreements and correlations were achieved for most of the day (i.e. from 7 to 23). A simple 24-hour time sequential background dataset can therefore significantly improve the prediction.

4.3.2 Experiment 2

4.3.2.1 Hanoi

Experiment 2 simulates the dispersion of road traffic emission of all 10 domains for the duration between May 2017 to May 2019. Observation data from NVC, which covers 2010 to 2015, therefore is not discussed in these models.

Figure 25 shows the emission of Hanoi core at 18 o'clock moving toward the North West. This is anticipated since the prevailing wind direction in the region is East South East. Similarly, Figure 26 demonstrates how emission from the suburban areas can affect CO concentrations in the populated districts. The East domain can be seen having the most influence whereas the impacts from West and South are minimum. As such, the inclusion of the East domain may affect the agreement with observation at stations 132, HD and HK. With some Northern winds observed during winter, the North domain also has a slight contribution. However, it is unlikely to be enough to have an impact on any of the 12 available stations.

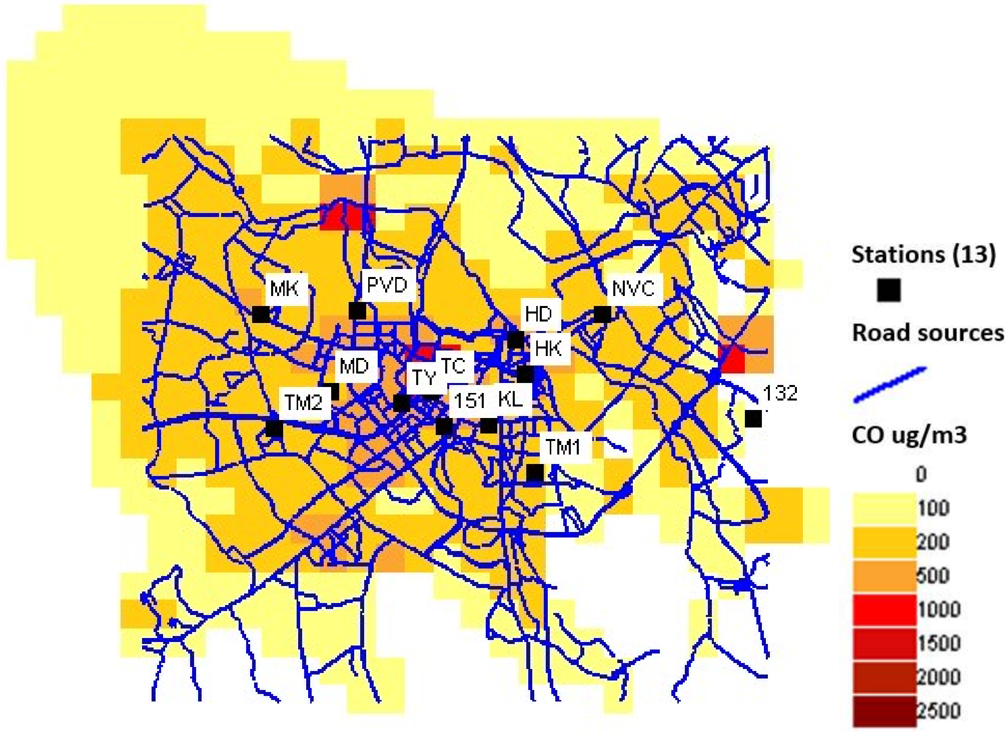


Figure 25. Long term contour of CO in Hanoi core at 18 o'clock, without background.

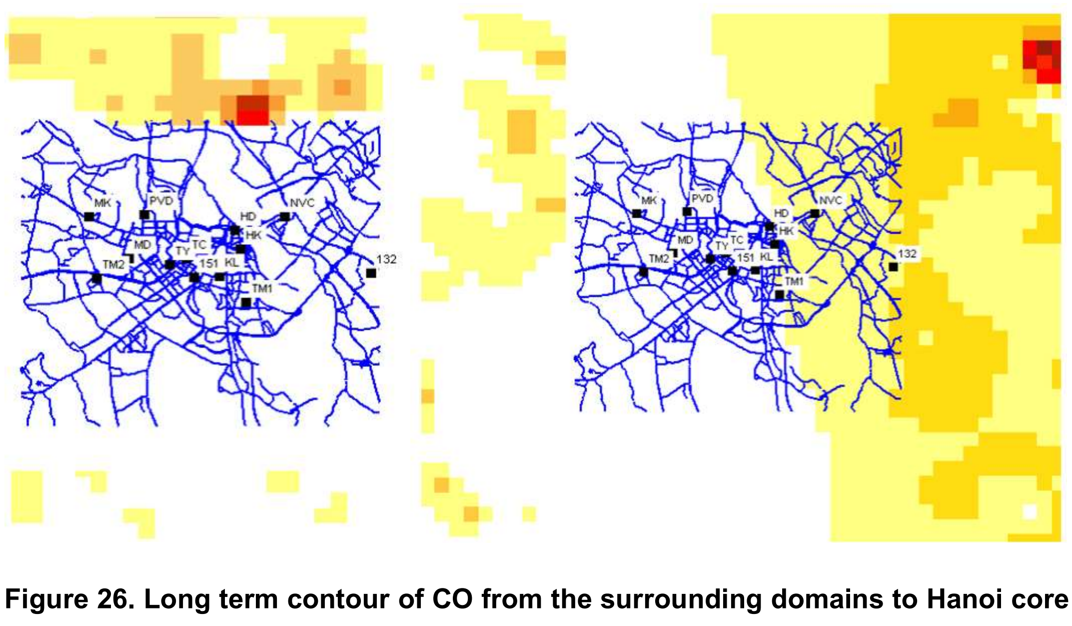


Figure 26. Long term contour of CO from the surrounding domains to Hanoi core at 18 o'clock, without background.

Because both experiments use meteorological data from the same source, when evaluate against winds observed by stations, the results are very similar to which of Experiment 1, i.e. Northern winds appear more frequently with higher velocity in Winter and Spring, minimum differences in Summer and Autumn. Compared directly to Experiment 1 however, Experiment 2's winds, which were obtained by averaging multiple GFS grid points, are more neutral (Figure 27). Looking into the seasonal variation of concentration, to Experiment 1, Experiment 2 has a negligible reduction in correlation between modelled and observed data. Perhaps with more detailed data in the future, the effect of winds in Hanoi can be further studied. However, within the scope of this project, from this point, both experiments are now considered to have indifferent wind input.

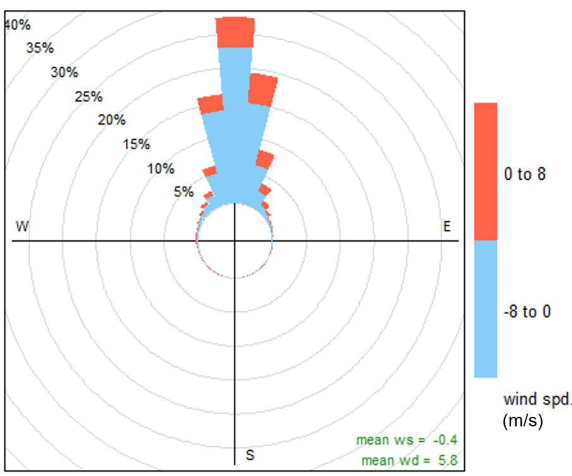


Figure 27. Frequency of counts by wind direction (%) of Exp2's winds compared to Exp1's winds.

As suspected with the contours, the inclusion of suburban domains does have an impact, most significantly at 132 and HK, i.e. the Eastern sites. Whilst concentrations can be seen being increased, the overall agreement with observation is not actually improved (Figure 28). As such, it seems there is threshold at which point emission from other domains becomes excess.

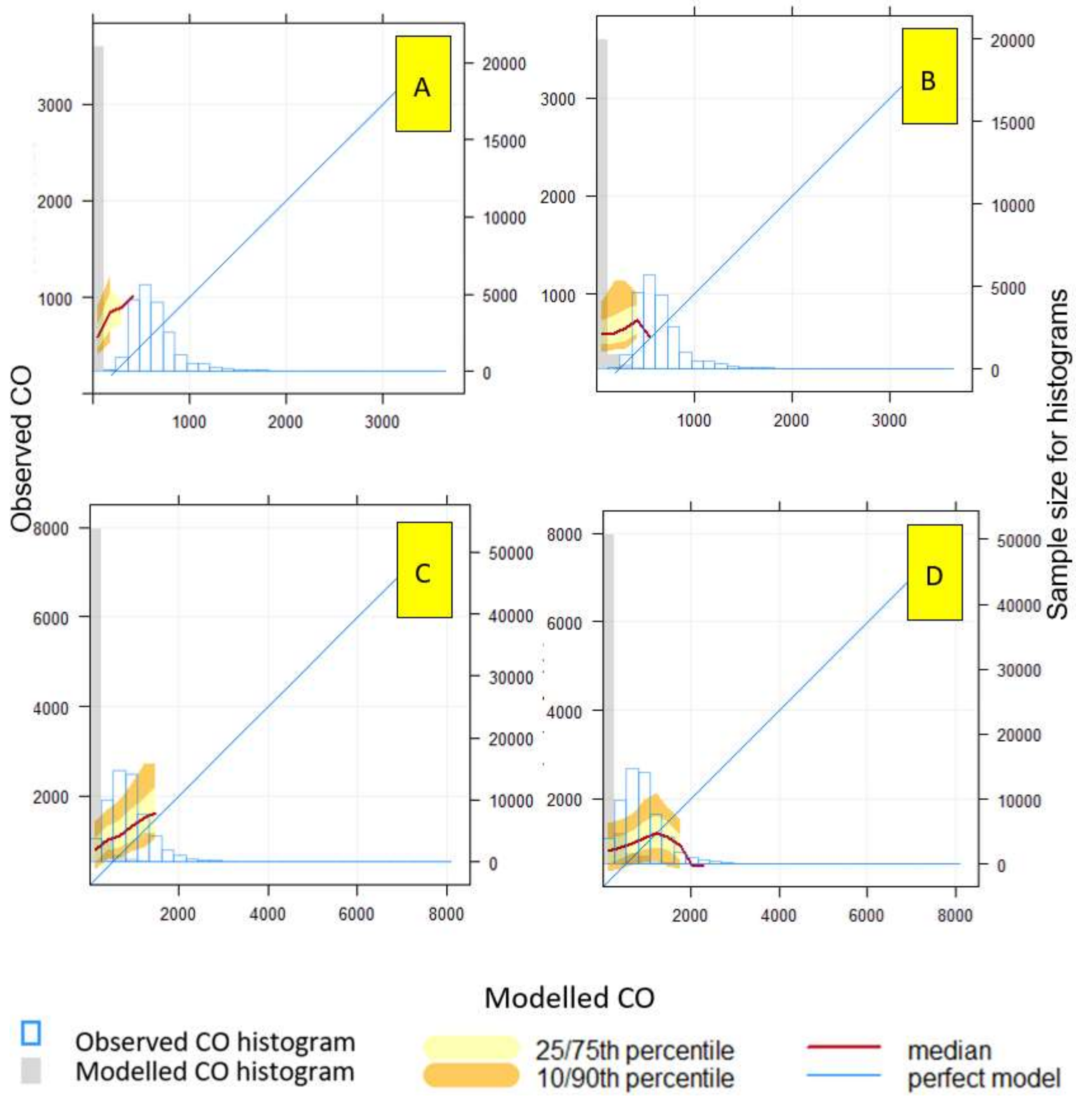


Figure 28. Conditional quantile plot for hourly CO at: A - MK(Exp1); B – MK(Exp2); C - HK(Exp1); D – HK(Exp2), without background.

The diurnal variation of concentration derived from heavy vehicles however received a significant improvement. Based on Figure 29, except for during the night, the r value stays relatively constant at positive 0.8. Since this experiment aims specifically to better understanding the impact of long-distance travelling, this is a very good result which indicates that Hanoi's input is undermining the fleet of heavy vehicles from out site of the city core. On the other hand, Car Light and MC appear to have their correlation slightly reduced, i.e., personal vehicles operate mostly in the core, adding further emission of these classes from other domains causes overestimation.

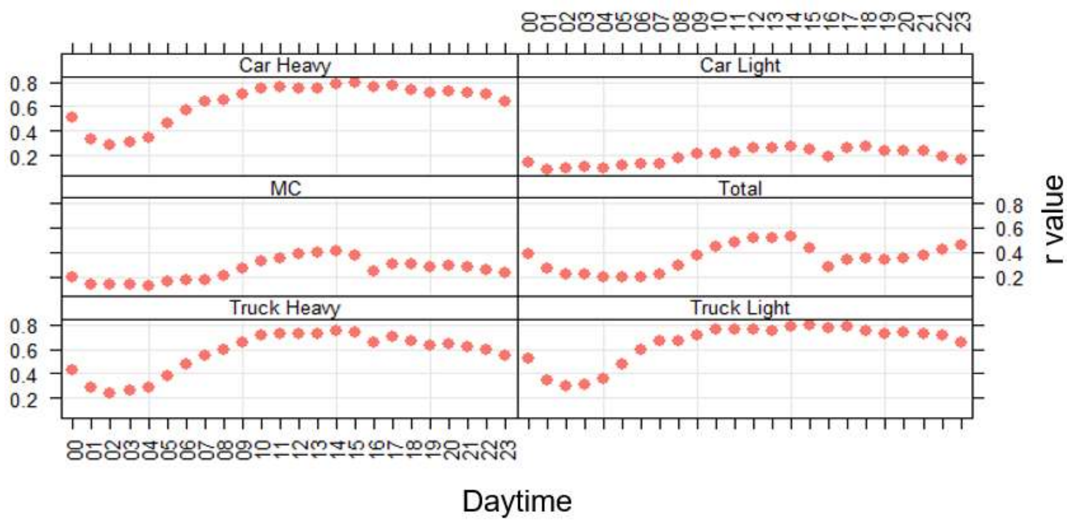


Figure 29. r value between modelled and observed CO in Hanoi, plotted by daytime.

4.3.2.2 Ho Chi Minh City

The above discussion suggests that HCMC, with its centralising manner of urban planning, would likely to have CO concentrations following MC and Light Car closely, whereas the contribution of heavy vehicles is negligible. However, Figure 30 shows that the measured concentrations actually resemble all vehicle classes and that similar to Hanoi, high CO concentrations are found during the night. The modelled result however follows MC too closely and is unable to produce such features.

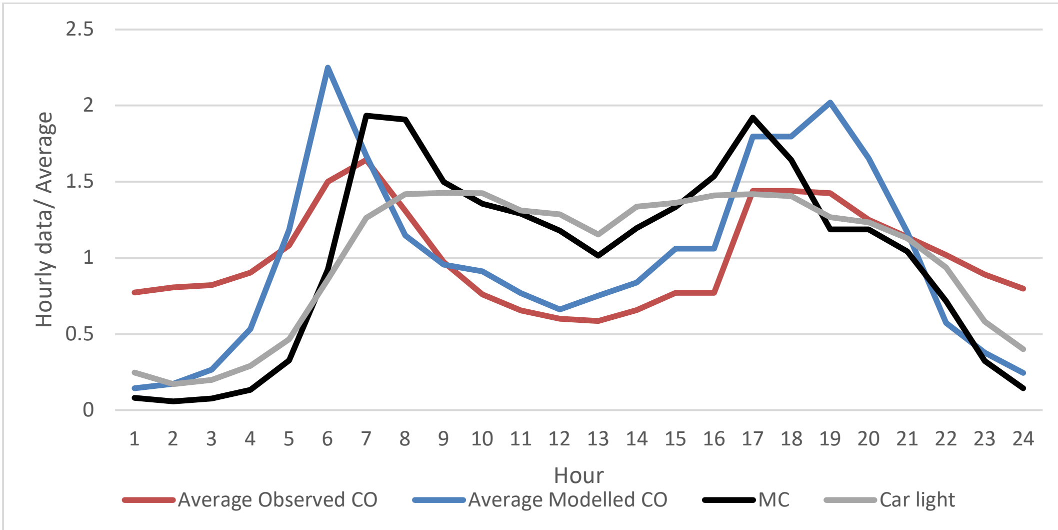


Figure 30. Normalised diurnal variation of mean modelled and observed CO in HCMC, without background pollution, and of MC and Car Light

Figure 31 shows that Light Car and MC have a much better relationship with observed data in HCMC, as compared to Hanoi. This is expected since the largest set of input data was sourced from HCMC. Both cities' models have poor correlation with observed data during the night. Again, this indicates another emission source beside traffic.

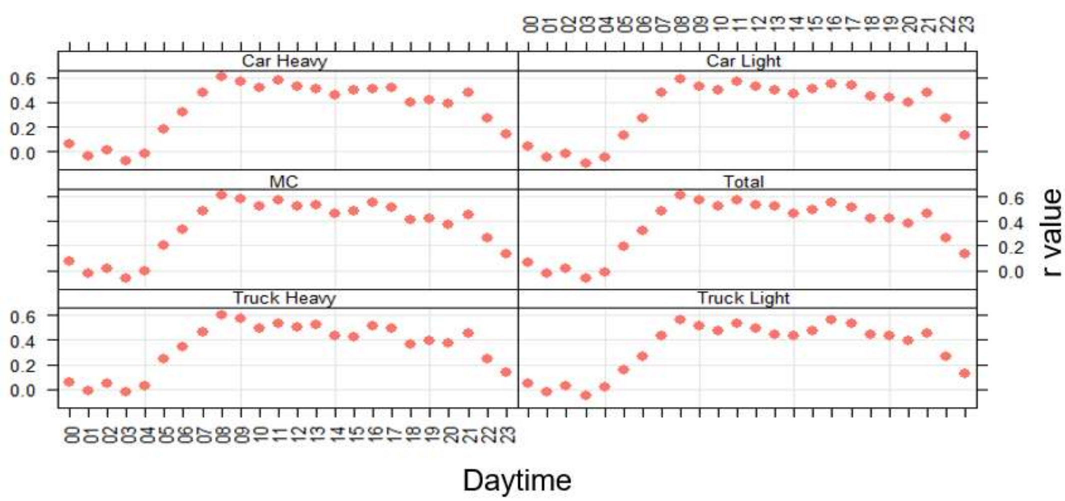


Figure 31. r value between modelled and observed CO in HCMC, plotted by daytime.

As for the prediction accuracy, 1940.41μg/m^3 is again too high for CO background concentration (Figure 32). The result stays relatively parallel with the perfect model until reaching 1000μg/m^3 and exhibits overestimations caused by the emission from external domains.

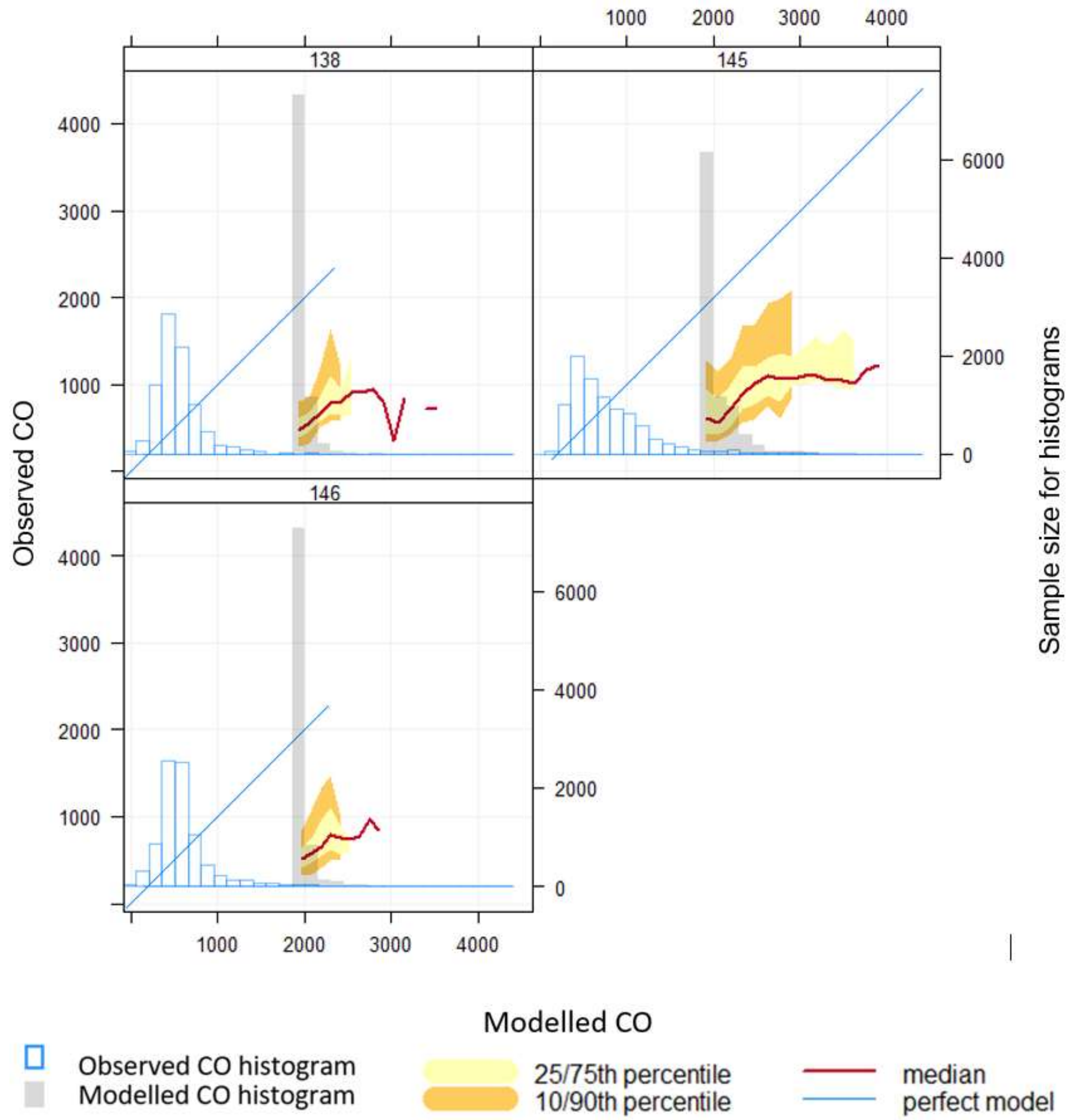


Figure 32. Conditional quantile plot for hourly CO at stations in HCMC, with background pollution.

4.4 Summary

There are currently too many variables to be considered, namely meteorological data, diurnal variation, vehicle classes and randomly generated vehicle count. Combined with that data for in situ concentration and background pollution are limited, the model validation is deemed cluttered. As a result, further investigations regarding to e.g. the impact of road canyons, road elevation/

tunnels, could not be made. Still, this chapter has managed to evaluate positive correlations between the modelled results and the available monitoring network in Hanoi and HCMC. Encouraging agreements have been achieved and can be easily improved with the addition of simple, more appropriate background concentration. In situ observations whilst agree with models that MC is the dominant urban polluting source, also indicates the contribution of other vehicle classes as well as of emission sources other than traffic.

5 Conclusion

ADMS was employed to simulate the dispersion of road traffic derived air pollution in 2 cities: Hanoi and HCMC and their respective surrounding areas. In order to satisfy ADMS's input requirement however, a wide range of issues were addressed with the current state of data in Vietnam, most namely the lack of effectively enforcement and quality standards which usually results in data inconsistency and unreliability. Other constrains such as inefficiency in data collection and handling were also discussed. A semi-automated modelling framework was therefore designed to transfer quality data from previous studies to produce a comprehensive input dataset. This also allowed the use of various ADMS's optional modules.

With the aim of demonstrating the applicability of dispersion modelling in developing countries, this work emphasises on adaptability and future potentials rather than in-depth investigations of air pollution in Hanoi and HCMC.

Multiple model configurations were initially idealised. However due to various constrains, namely data limitations and time/ computational resources, only 2 configurations were used by this thesis. They were thoroughly referred as Experiment 1 and Experiment 2. Concentration data from in-situ observations were sourced as to validate the model outcomes.

Experiment 1 modelled a relatively small spatial area (i.e. Hanoi's main districts) for 7 years. Since most air pollution and traffic related studies in Vietnam remain focused on the inner urban regions, this approach utilised the obtained data most effectively. Experiment 2 demonstrated another potential application, in which the dispersion of air pollutants was investigated on a much larger scale.

Without a suitable set of background concentration, the result validation experienced some difficulties. However, the model does show encouraging relationships with in-situ observations. In specific, stations in both cities achieved strong positive correlation was with modelled emission derived from bus and trucks. Result from HCMC suggests that the framework was able to produce a very good representation of the real-world traffic activities. Similarly, even with a suboptimal background concentration, good agreements were obtained during traffic active hours (i.e. 7 to 23). This shows that there remain many potentials to improve the prediction.

This project has made three main contributions:

- Based on the actual requirement of a developing country, proposes a new suitable tool to the air quality management system.

- By successfully using ADMS, this work has managed to demonstrate the applicability of advanced modelling software in a developing country.

- A modelling framework has been developed to create a fairly comprehensive dataset that satisfies the input requirement of dispersion modelling.

5.1 Limitations

There are three major limitations derived from this work's methodology, those are: 1. The scale of models, 2. The emission random generating process, and 3. The lack of an appropriate background concentration dataset. This section will briefly discuss each, their impact on the study and provide potential improvements.

The scale of models in this work was initially intended as to evaluate the overall pollution level for both urban and suburban areas of Hanoi and HCMC. Consideration was also given to the differences in cities' socio-economic functions, in hope that comparisons can be made to study the impact of those differences. Inadequate data however have proven this to be too ambitious. As such, not only that model output could not be evaluate fully (e.g. no CO concentration available outside the core domains), the scale was too large and thus introduced extra workload (e.g. main domains had to be divided into several layers of nested domains). As a result, rather than fitting every agendas in a domain, the planning of future models should be more concise, well optimised for a singular type of emission source (e.g. urban traffic). If required, models can then be improved onward.

The random generating process of vehicle count was critically an innovative technique, as it explicitly designates each road with individual emission data, in turn creating a full road network containing both real-world randomisation and controlled diurnal variation profiles. However the technique is not fully matured, evidently producing extreme values that caused uncertainties in the model output. Further refinements can be made including adding the consideration of city's development zones, or integrating data from traffic congestion models, which is similar to what was done in (Yang, et al., 2019).

The lack of suitable information on background pollution was evident during the model performance evaluation. By using a constant CO concentration that was too high for the underlying condition, most stations experienced overestimation. Similarly, the evaluation had to rely heavily on correlation (r – values) rather than employing other statistics such as error (NMSE) and bias (FB). Fortunately, Chapter 4 had manged to identify the contribution of background CO and gave recommendations on more suitable sets of background information for future uses.

5.2 Future Applications

This section will discuss the five most potential applications of this work and of ADMS in general with regards to the conditions and needs of urban areas in Vietnam, as an example of a developing country. Those applications are: 1. evaluate traffic NO_x and PMs emissions, 2. Fast, periodic predictions, 3. Scenario/ urban planning prediction, 4. Study of large scale, difficult to define sources, 5. Uses in health and exposure researches.

As discussed thoroughly in this work, CO was focused as a pilot species for dispersion model because of the data and time limitation. CO is not the most pressing concern for the cities in Vietnam, as compared to PMs and NO_x. It is therefore important to have the models refined and applied on NO_x and PMs and produce in-depth evaluations that is previously not available, e.g. diurnal variation, chemical reactions. Perhaps, with those models the contribution of Heavy Truck and Heavy Car can be further investigated since according to (Bang, et al., 2019), their contribution to the total anthropogenic NO_x are 9% and 11% respectively, as compared to MC's 29% contribution.

This work has produced a semi-automated modelling framework, that utilises various large mega datasets. It is therefore able to re-use data, add customisations and new inputs to serve other modelling purposes. As such, being versatile, reasonably fast/ computational demanding whilst able to maintain appropriate output resolution and accuracy, the framework is very suitable for prediction and scenarios testing. For instance, a periodic modelling scheme can be designed to predict air quality for the next few hours for the city centre. Not only provide a useful tool for the protection of public health, this would also encourage the departure from the use of Air Quality Index toward conventional air quality standards, benefiting both researches and enforcements.

Similarly, the framework can exploit various ADMS's integrated functions to test urban planning scenarios. For example, survey showed that only 32% of MC in HCMC met the EURO III standard (Bang, et al., 2019), using the UK's Emission Factor Toolkit supplied by ADMS, it is possible to promptly quantify the impact on air pollution if the above value is doubled. Similar prediction can be done with scenarios such as MC being replaced entirely by public transport or outgrown by other vehicles within 10 years.

With the ability to explicitly specify emission source of various types, including line, point, area, volume and grid, the framework can be applied to study the emission of sources that have rarely been digitised before. Figure 33 is an example of this work being used to simulate the dispersion of pollution produced by crop/ biomass burning in the area around Hanoi.

Lastly, thanks to its road canyon functions and the ability to produce output at high spatial/ temporal resolutions, this work can be adapted to street scale dispersion modelling, thus allows fine evaluations regarding human exposure/ health – related impact.

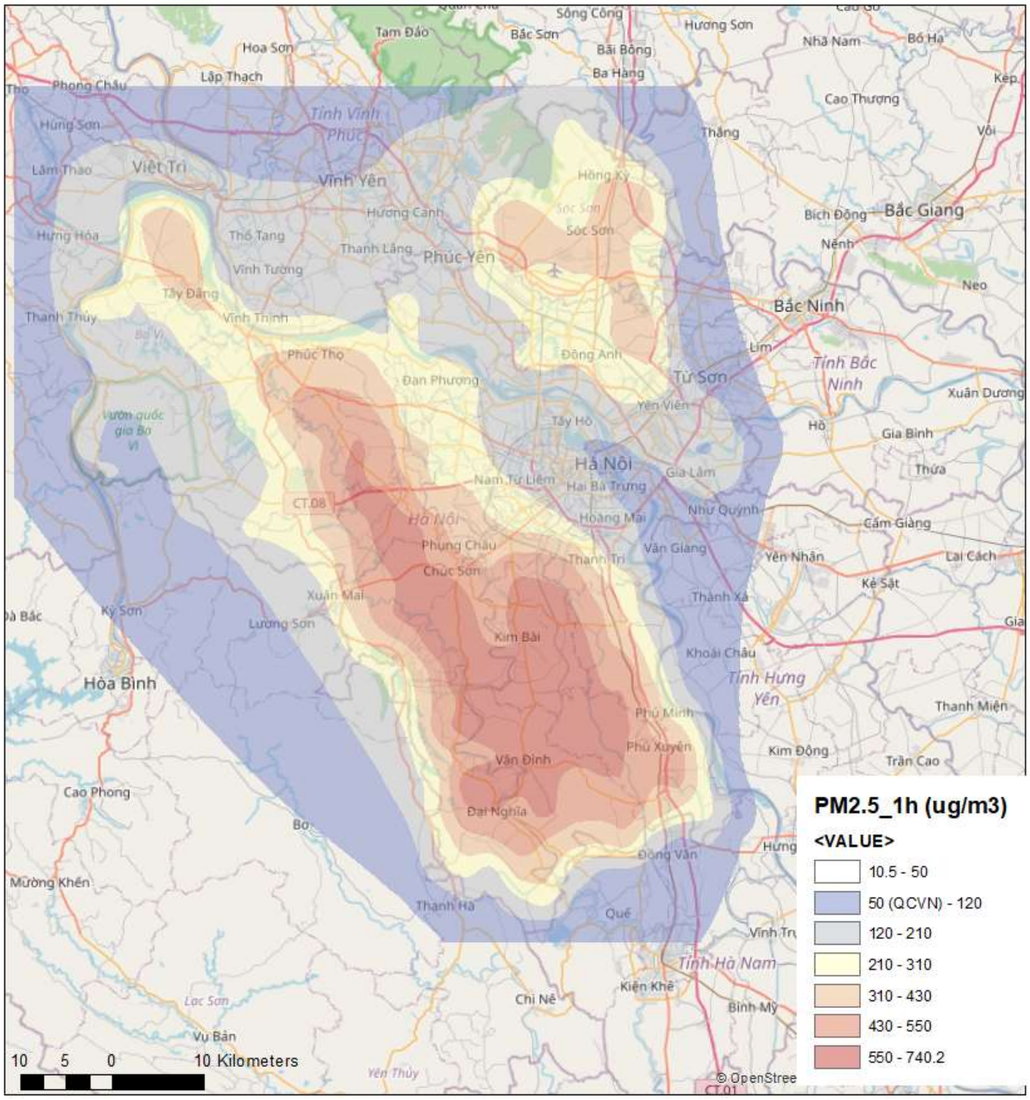


Figure 33. Contour of the dispersion of $PM_{2.5}$ emitted from rice straw burning in Northern Vietnam, as modelled using ADMS.

6 References

Amorim, J. et al., 2014. Modelling the Effects of Urban Morphology, Traffic and Pedestrian Dynamics on Students' Exposure to Air Pollution. In: D. Steyn & R. Mathur, eds. *Air Pollution Modelling and its Application XXIII.* s.l.:Springer.

Anfossi, D. & Physick, W., 2004. Lagrangian Particle Models. In: P. Zannetti, ed. *Air Quality Modeling – Theories, Methodologies, Computational Techniques, and Available Databases and Software. Vol. II – Advanced Topics.* Pittsburgh: EnviroComp Institute & Air & Waste Management Association.

Antoniou, A., Montazeri, H., Neophytou, M. & Blocken, B., 2019. CFD Simulation of Urban Microclimate: Validation Using High-resolution Field Measurements. *Science of The Total Environment,* Volume 695, p. 133743.

Bang, H., 2010. *Optimal Methodology to Generate Road Traffic Emissions for Air Quality Modeling: Application to Ho Chi Minh City.* Suisse: École Polytechnique Fédérale de Lausanne.

Bang, H., 2017. Photochemical Smog Modelling Using the Air Pollution Chemical Transport Model (TAPM-CTM) in Ho Chi Minh City, Vietnam. *Environmental Modelling & Assessment,* Volume 24, pp. 295-310.

Bang, H., 2018. *Bao Cao Nghiem Thu Ket Qua Khoa Hoc Cong Nghe: Nhiem Vu Thiet Lap Ban Do Lan Truyen O Nhiem Khong Khi Doi Voi Hoat Dong Giao Thong, San Xuat Cong Nghiep Tai Thanh Pho Ho Chi Minh.* Ho Chi Minh City: Vietnam National University.

Bang, H. et al., 2019. A Combination of Bottom-up and Top-down Approaches for Calculating of Air Emission for Developing Dountries: A Case of Ho Chi Minh City, Vietnam. *Air Quality, Atmosphere & Health,* Volume 12, p. 1059–1072.

Berkowicz, R. et al., 2008. Evaluation and Application of OSPM for Traffic Pollution Assessment for a Large Number of Street Locations. *Environmental Modelling & Software,* 23(3), pp. 296-303.

Biggart, M. et al., 2019. Street-scale air quality modelling for Beijing during a winter 2016. *Atmospheric Chemistry and Physics Discussions, in review.*

Brulfert, G., 2004. *Modélisation des circulations atmosphériques pour l'étude de la pollution des vallées alpines.* Grenoble: University Joseph Fourier.

Cai, H. & Xie, S., 2011. Traffic-related Air Polution Modelling During the 2008 Beijing Olympic Games: The Effects of an Odd-Even Day Traffic Restriction Scheme. *Science of The Total Environment,* Volume 408, pp. 1935- 1984.

Carruthers, D. J. et al., 1997. ADMS Urban – An Integrated Air Quality Modelling System for Local Government. *Transaction in Ecology and Environment,* Volume 15.

Carruthers, D. & Stidworthy, A., 2019. Urban Emission Inventory Optimisation Using Sensor Data, An Urban Air Quality Model and Inversion Techniques. *International Journal Environmental and Pollution,* 66(4), pp. 252-266.

Carslaw, D., 2011. *Defra Urban Model Evaluation Analysis – Phase 1.* London: Defra.

CERC, 2017a. *ADMS-Urban Urban Air Quality Management System: User Guide, Version 4.1.1.* Cambridge: Cambridge Environmental Research Consultants Ltd..

CERC, 2017b. The Met Input Module. *ADMS Technical Specifications,* p. P05/01S/17.

CERC, 2017c. Road Sources. *ADMS Technical Specifications,* p. P31/01B/17.

CERC, 2017d. Advanced Street Canyon Specification. *ADMS Technical Specifications,* p. P28/02A/17.

CERC, 2017e. Road Tunnel Specification. *ADMS Technical Specifications,* p. P35/01A/17.

CERC, 2020. *ADMS – Roads.* [Online] Available at: https://www.cerc.co.uk/environmental-software/ADMS-Roads-

model.html
[Accessed 08 January 2020].

Chatzimichailidis, A., Argyropoulos, C., Assael, M. & Kakosimos, K., 2019. Implicit Definition of Flow Patterns in Street Canyons—Recirculation Zone—Using Exploratory Quantitative and Qualitative Methods. *Atmosphere,* 10(12).

Chen, C. H. et al., 2009. Impact of Ambient Air Pollution on Public Health Under Various Traffic Policies in Shanghai, China. *Biomedical and Environmental Sciences,* Volume 22, pp. 210-215.

Chu, X. & Thi, H., 2017. Actual Situation and Solutions for Reducing the Traffic Jams and Congestion in Vietnam. *Advances in Natural and Applied Sciences, Vol. 11(12),* pp. 26-34.

Costabile, F., Allegrini, I. & Wang, F., 2006. CDF Modelling of Traffic-related Air Pollutants Around an Urban Street-canyon in Suzhou. *Transactions on Ecology and the Environment,* Volume 86.

Dixon, N. S., Boddy, J. W. D., Smalley, R. J. & Tomlin, A. S., 2006. Evaluation of a Turbulent flow and Dispersion Model in a Typical Streey Canyon in York, UK. *Atmospheric Environment,* 40(5), pp. 958-972.

Douglas, P. et al., 2016. Sensitivity of Predicted Bioaerosol Exposure from Open Windrow Composting Facilities to ADMS Dispersion Model Parameters. *Journal of Environmental Management,* Volume 184, pp. 448-455.

Fallah-Shorshani, M., Shekarrizfard, M. & Hatzopoulou, M., 2017. Intergrating a Street-canyon Model with a Regional Gaussian Dispersion Model for Improved Characterisation of Near-road Air Pollution. *Atmospheric Environment,* Volume 153, pp. 21-31.

Fenger, J., 2009. Urban Air Pollution. In: C. Hewitt & A. Jackson, eds. *Atmospheric Science for Environmental Scientists.* Chichester: Blackwell Publishing.

Fernandez-Lopez, J. & Schliep, K., 2019. *Package ‘rWind’.* [Online].

Gulliver, J. & Briggs, D., 2011. STEM-Air: A Simple GIS-based Air Pollution Dispersion Model for City-wide Exposure Assessment. *Science of the Total Environment,* Volume 409, pp. 2491-2429.

Hanoi People's Committee , 2012. *QUYẾT ĐỊNH: VỀ VIỆC PHÊ DUYỆT "QUY HOẠCH MẠNG LƯỚI QUAN TRẮC KHÔNG KHÍ CỐ ĐỊNH TRÊN ĐỊA BÀN THÀNH PHỐ HÀ NỘI ĐẾN NĂM 2020".* Hanoi: Hanoi People's Committee .

Harmanto, D., Setiawan, M. I. & Wiwoho, S., 2015. *CFD Simulation for Predicting the Wind Effect on the High Rise Building: NET Tower Surabaya.* s.l., Narotama International Conference on Civil Engineering.

Hien, P. D., Men, N. T., Tan, P. M. & Hangartner, M., 2020. Impact of Urban Expansion on the Air Pollution Landscape: A Case Study of Hanoi, Vietnam. *Science of The Total Environment,* Volume 702, p. 134635.

Ho Chi Minh City People's Committee, 2009. *QUYẾT ĐINH: VỀ SỬA ĐỔI, BỔ SUNG MỘT SỐ ĐIỀU CỦA QUYẾT ĐỊNH SỐ 135/2007/QĐ-UBND NGÀY 08 THÁNG 12 NĂM 2007 CỦA ỦY BAN NHÂN DÂN THÀNH PHỐ QUY ĐỊNH VỀ KIẾN TRÚC NHÀ LIÊN KẾ TRONG KHU ĐÔ THỊ HIỆN HỮU TRÊN ĐỊA BÀN THÀNH PHỐ HỒ CHÍ MINH.* Ho Chi Minh City : Ho Chi Minh City People's Committee.

Hoai, N., 2019. *Không thể có chuyện thuỷ ngân bằng 0 trong môi trường.* [Online] Available at: https://www.tienphong.vn/cong-nghe/khong-the-co-chuyen-thuy-ngan-bang-0-trong-moi-truong-1459560.tpo
[Accessed 08 January 2020].

Hoang, T. A., Chu, N. X. & Tran, T. V., 2017. The Environmental Pollution In Vietnam: Source, Impact and Remedies. *International Journal of Scientific & Technology Research,* 6(02), pp. 249-253.

Holmes, N. S. & Morawska, L., 2006. A Review of Dispersion Modelling and its Application to the Dispersion of Particles: An Overview of Different Dispersion Models Available. *Atmospheric Environment,* Volume 40, p. 5902–5928.

Hood, C. et al., 2018. Air Quality Simulations for London Using a Coupled Regional-to-Local Modelling System. *Atmospheric Chemistry and Physics,* Volume 18, pp. 11221-11245.

Hung, N., 2010. *Urban Air quality Modelling and Management in Hanoi, Vietnam.* Denmark: Aarhus University.

Huynh, D. & Gomez-Ibanez, J., 2017. Vietnam. In: D. Pojani & D. Stead, eds. *The Urban Transport Crisis in Emerging Economics.* s.l.:Springer International Publishing Switzerland.

JICA, 2007. *The Comprehensive Urban Development Programme in Hanoi Capital City of the Social Republic of Vietnam (HAIDEP).* s.l.:JICA.

JICA, 2010. *The Comprehensive Study on the Sustainable Development of Transport System in Vietnam.* s.l.:Ministry of Transport.

Jorquera, H., 2002. Air Quality at Santiago, Chile: A Box Modelling Approach – I. Carbon Monoxide, Nitrogen Oxides and Sulfur Dioxide. *Atmospheric Environment,* 36(2), pp. 315-330.

Jorquera, H., 2002. Air Quality at Santiago, Chile: A Box Modelling Approach – II. PM2.5, Coarse and PM10 Particulate Matter Fractions. *Atmospheric Environment,* 36(2), pp. 331-344.

Kakosimos, K. E., Hertel, O., Ketzel, M. & Berkowicz, R., 2010. Operational Street Pollution Model (OSPM) – A Review of Performed Application and Validation Studies, and Future Prospect. *Environmental Chemistry,* Volume 7, pp. 485-503.

Kaseb, Z., Hafezi, M., Tahbaz, M. & Delfani, S., 2020. A Framework for Pedestrian-level Wind Conditions Improvement in Urban Areas: CFD Simulation and Optimization. *Building and Environment,* Volume 184, p. 107191.

Khanh, P., 2019. *Sa bàn 3D về ô nhiễm không khí ở Hà Nội hoạt động như thế nào?.* [Online]
Available at: https://baomoi.com/sa-ban-3d-ve-o-nhiem-khong-khi-o-ha-noi-

hoat-dong-nhu-the-nao/c/31582555.epi
[Accessed 08 January 2020].

Lazić, L. et al., 2016. Traffic Contribution to Air Pollution in Urban Street Canyons: Integrated Application of the OSPM, Moss Biomonitoring and Spectral Analysis. *Atmospheric Environment,* Volume 141, pp. 347-360.

Leadbetter, S. J. & Hort, M. C., 2011. Volcanic ash hazard climatology for an eruption of Hekla Volcano, Iceland. *Journal of Volcanology and Geothermal Research,* Volume 199, p. 230–241.

Metcalfe, S. & Derwent, D., 2005. *Atmospheric Pollution and Environmental Change.* London: Hodder Arnold.

Ministry of Construction, 2007. *Construction Standards: Urban Roads.* Hanoi: Ministry of Construction.

Misra, A., Roorda, M. J. & MacLean, H. L., 2013. An Intergrated Modelling Approach to Estimate Urban Traffic Emission. *Atmospheric Environment,* Volume 73, pp. 81-91.

MONRE, 2013. *National Environmental Report 2013, Atmospheric Environment.* Hanoi: Minister of Natural Resources and Environment.

MONRE, 2016. *National Environmental Report 2016, Urban Environment.* Hanoi: Minister of Natural Resources and Environment.

Moradpour, M. & Hosseini, V., 2020. An Investigation into the Effects of Green Space on Air Quality of an Urban Area Using CFD Modeling. *Urban Climate,* Volume 34, p. 100686.

Nam, D. T., Anh, L. H. & Luan, V. N., 2018. A Study of Ambient Air Quality Status in Hanoi, By Applying Air Quality Index (AQI). *Vietnam Environmental Adminstration Magazine,* Volume 4, pp. 31-36.

Oanh, N. T. K., Phuong, M. T. T. & Permadi, D. A., 2008. Analysis of Motorcycle Fleet in Hanoi for Estimation of Air Pollution Emission and Climate Mitigation Co-benefit of Technology Implementation. *Atmospheric Environment,* Volume 59, pp. 438-448.

Oettl, D. et al., 2012. Lagrangian Dispersion Modeling of Vehicular Emissions from a Highway in Complex Terrain. *Journal of the Air & Waste Management Association,* Volume 53, p. 1233–1240.

Palmer, A., 2007. *The Development of an Integrated Routing and Carbon Dioxide Emissions Model for Goods Vehicle.* Cranfield: Cranfield University.

Phong, Q., 2019. *Hanoi Neu Lo Trinh Cam Xe May Trong Noi Thanh.* [Online] Available at: https://dantri.com.vn/xa-hoi/ha-noi-neu-lo-trinh-cam-xe-may-trong-noi-thanh-20191024150018332.htm
[Accessed 20 07 2020].

Phuc, N. H. & Oanh, N. T. K., 2018. Determining Factors for Levels of Volatile Organic Compounds Measured in Different Microenvironments of a Heavy Traffic Urban Area. *Science of the Total Environment,* Volume 627, p. 290–303.

Pojani, D. & Stead, D., 2017. The Urban Transport Crisis in Emerging Economies: An Introduction. In: D. Pojani & D. Stead, eds. *The Urban Transport Crisis in Emerging Economics.* s.l.:Springer International Publishing Switzerland.

Pucher, J., 2017. The Urban Transport Crisis in Emerging Economies. *Transport Reviews,* 38(2), pp. 271-273.

Rafiei, M. & Sturm, P. J., 2018. Modelling of Carbon Monoxide Dispersion Around the Urban Tunnel Portal. *Global Journal Environmental Science Management,* 4(3), pp. 359 -372.

Ramm, F., 2017. *OpenStreetMap Data in Layered GIS Format.* [Online] [Accessed 08 January 2020].

Redington, A. L. & Derwent, R. G., 2002. Calculation of Sulphate and Nitrate Aerosol Concentrations over Europe Using a Lagrangian Dispersion Model. *Atmospheric Environment,* 36(28), pp. 4425-4439.

Roser, M., 2019. *Two Centuries of Rapid Global Population Growth Will Come to an End.* [Online] Available at: https://ourworldindata.org/world-population-growth-past-future [Accessed 10 September 2020].

Russell, A. et al., 2014. Use of Air Quality Modelling Results in Health Effects Research. In: D. Steyn & R. Mathur, eds. *Air Pollution Modelling and its Application XXIII.* London: Springer.

Sakamoto, Y. et al., 2018. Air Quality Study in Hanoi, Vietnam in 2015–2016 Based on a One-year Observation of NOx, O3, CO and a One-week Observation of VOCs. *Atmospheric Pollution Research,* 9(3), pp. 544-551.

Samuel, P., 2019. *How Air Pollution in Vietnam Affects the Business environment. Vietnam Briefing.* [Online] Available at: https://www.vietnam-briefing.com/news/air-pollution-vietnam-affects-business-environment.html/. [Accessed 07 February 2020].

Sanchez, B. et al., 2017. Modelling NOX Concentrations Through CFD-RANS in an Urban Hot-spot Using High Resolution Traffic Emissions and Meteorology from a Mesoscale Model. *Atmospheric Environment,* Volume 163, pp. 155-165.

Santiago, J. L. & Martin, F., 2015. Use of CFD Modeling for Estimating Spatial Representativeness of Urban Air Pollution Monitoring Sites and Suitability of their Locations. *Física de la Tierra,* Volume 27, pp. 191-221.

So Tai Nguyen va Moi Truong, 2019. *Bao Cao: Ket Qua Quan Trac Hien Trang Moi Truong Tinh Ha Giang Dot 1 nam 2019.* Ha Giang: So Tai Nguyen va Moi Truong.

Sternar, L. et al., 1997. *Urban Air Quality Management Strategy in Asia – Guidebook.* s.l.:The World Bank Press.

Stocker, J. et al., 2013. *Road Source Model Intercomparison Study Using New and Existing Datasets.* Spain: 15th International Conference Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purpose.

Thuy, H., 2019. *Thủy ngân quanh kho Rạng Đông vượt ngưỡng 10-30 lần VNExpress.* [Online]
Available at: https://vnexpressnet/thoi-su/thuy-ngan-quanh-kho-rang-dong-vuot-nguong-10-30-lan-3977486html
[Accessed 2020 January 08].

Tiwary, A. & Colls, J., 2010. *Air Pollution: Measurement, Modelling and Mitigation.* 3rd ed. Abingdon: Routledge.

Tominaga, Y. & Stathopoulos, T., 2011. CFD Modeling of Pollution Dispersion in a Street Canyon: Comparison between LES and RANS. *Journal of Wind Engineering and Industrial Aerodynamics,* Volume 99, pp. 340-348.

Trang, T., Van, H. & Oanh, N., 2015. Traffic Emission Inventory for Estimation of Air Quality and Climate Co-benefits of Faster Vehicle Technology Intrusion in Hanoi, Vietnam. *Carbon Management,* 6(3-4), pp. 117-128.

Truc, V. T. Q. & Oanh, N. T. K., 2007. Roadside BTEX and Other Gaseous Air Pollutants in Relation to Emission Sources. *Atmospheric Environment,* 41(36), pp. 7685-7697.

UN-Habitat, 2006. *State of the World’s Cities 2006/7.* Nairobi: Earthscan.

United Nations, 2019. *World Urbanization Prospects: The 2018 Revision.* New York: United Nations.

US Environmental Protection Agency (EPA), 2008. *EPA’s 2008 Report on the Environment.* Washington DC: Natural Center for Environment Assessment.

Vardoulakis, S., Fisher, B. E. A., Pericleous, K. & Gonzalez-Flescs, N., 2003. Modelling Air Quality in Street Canyons: A Review. *Atmospheric Environment,* Volume 37, pp. 155-182.

Vardoulakis, S., Valiantis, M., Milner, J. & Apsimon, H., 2007. Operational Air Pollution Modelling in the UK-Street Canyon Application and Challenges. *Atmospheric Environment,* 41(22), pp. 4622-4637 .

Vietnam General Statistical Office, 2016. *National Population and Statistics.* [Online]
[Accessed 07 January 2020].

Williams, M. L., 2014. Atmospheric Dispersal of Pollutants and the Modelling of Air Pollution. In: R. M. Harrison, ed. *Pollution: Causes, Effects and Control.* 5th ed. Cambridge: The Royal Society of Chemistry.

World Bank, 2011. *Vietnam Urbanization Review.* s.l.:World Bank.

World Health Organization, 1992. *Urban air pollution in megacities of the world.* Oxford: Blackwell Reference.

World Health Organization, 2000. Carbon monoxide. In: *Air Quality Guidelines - Second Edition.* Copenhagen: WHO Regional Office for Europe.

World Health Organization, n.d. *Air Pollution in Vietnam.* [Online] Available at: https://www.who.int/vietnam/health-topics/air-pollution [Accessed 07 January 2020].

Yang, D. et al., 2019. High-resolution Mapping of Vehicle Emissions of Atmospheric Pollutants Based on Large-scale, Real-world Traffic Datasets. *Atmospheric Chemistry and Physics,* 19(13), p. 8831–8843.