Paper number ITS-AN-SP0363

Integrating a mesoscopic traffic simulation model and a simplified NO_2 estimation model for predicting the impact of air pollution

Adriana Simona Mihăiţă[1*], Mirian Benavides Ortiz[2], Mauricio Camargo[2]

1. DATA61, 13 Garden Str., Eveleigh, 2015, Australia, simona.mihaita@nicta.com.au, (+61) 94905615

2. ERPI laboratory EA6737, 8 Rue Bastien Lepage, 54010, Nancy, France

Abstract

Continuous growth in traffic demand has led to a decrease in the air quality in various urban areas. Local authorities in environmental protection and urban planners are therefore interested in performing traffic and air pollution simulations for testing various urban scenarios and raising citizen awareness where necessary. This article is focused on the traffic and air pollution in the eco-neighbourhood "Nancy Grand Cœur", located in a medium-size city from north-eastern France. The main objective of this work is to build an integrated simulation model which would help predicting and visualizing various environmental changes inside the neighbourhood, such as air pollution, traffic flow or meteorological data. Firstly, we build the 3D mesoscopic traffic simulation model around the central train station of the eco-neighbourhood, using real data sets from the local traffic management centre. Secondly, by using reliable data sets from the local air-quality management centre, we build a regression model for predicting the evolution of NO_2 concentrations, as a function of the simulated traffic flow and meteorological data. Lastly, we validate the estimated results by comparisons with the real data sets, with the purpose to support the traffic engineering decision-making and the smart city sustainability.

Keywords:

Mesoscopic traffic simulation, air pollution, emission estimation, eco-neighbourhood

Introduction

As the urban areas have known a rapid increase, it is expected that by 2050, almost 70% of the population will live in cities (UN, 2014). According to Tom Tom[1], some of the cities with the highest congestion level around the globe are: Istanbul (58%), followed by Mexico City (55%), Rio de Janeiro (51%), Moscow (50%) and most of the major European cities. The large amount of vehicles in densely populated areas contributes to an increasing deterioration of the air

[11] https://www.tomtom.com/en_au/trafficindex/#/list

quality due to motor vehicle emissions. In 2012, the U.S. Environmental Protection Agency (US EPA) has shown that 61% of the total emissions of carbon monoxide (CO) and 35% of total emissions of nitrogen oxide were produced by highway vehicles (US EPA, 2016). Many of these pollutants not only cause adverse effects on the human health, but they also entail high economical costs. For example, in France, the financial and economic cost of air pollution amounts to 101.3 billion Euros each year (AÏCHI, 2015). The complexity of the air pollution lies in its extent and the large amount of factors changing its behaviour, making it even more difficult to implement measures for protecting the citizens. According to the 2012 air quality assessment (MEDE, 2012), air pollution is caused by various industrial, commercial, domestic, agricultural activities, but the traffic congestion is the major cause.

Therefore a systematic evaluation of efficient traffic management strategies that would reduce traffic congestion requires effective transportation simulation in order to efficiently assess operational traffic performance and emission impacts at different spatial and temporal resolutions (e.g. network, corridor and segment levels, second-by-second, peak hours, 24-h, multiple days) (Xuesong, et al. 2015). Mesoscopic and microscopic traffic simulation tools are widely used for obtaining very detailed traffic analysis and provide traffic dynamic second-by-second. Although they require a high amount of data and can be computationally intensive, they can be efficiently combined with various emission tools; a recent comparison study and analysis of various integrated traffic simulation and emission tools is provided in (Fontes, et al. 2015). For detailed insights regarding all traffic, emissions and air quality modelling approaches, the reader is redirected to (Shorshani, et al. 2015).

While the construction of the traffic simulation or air pollution models can follow standard modelling techniques, integrating both modules together requires an increased complexity and data analysis. In this article we build a simplified nitrogen dioxide ($NO_{2)}$ estimation model from traffic and meteorological data which helps us validate the integrated mesoscopic traffic simulation model. In the next section we introduce the context of our project and the case study we present in this paper. We further conduct a data profiling analysis and build the 3D mesoscopic traffic simulator using traffic volume real data which we receive from the Grand Nancy local authority. Altough various research studies consider other factors, such as speed, acceleration, etc. for moelling and predicting air emissions, these data sets are not always available in many urban areas with limited traffic detectors or monitoring capabilites, this being the case as well of Nancy Grand Cœur. The results of the simulation are used to implement the simplified estimation model for NO_2, by considering as well meteorological data received from the local air monitoring association Air Lorraine. The last subsection presents the seasonality study we conduct over the traffic volume and we validate the estimation results over real data.

Integrated traffic and air pollution simulation model

Context

Maintaining sustainable development in congested cities has become a priority for local authorities, in order to provide flexibility, intermodal transportation systems and green mobility for its citizens. As a response to these problems, the "eco-neighbourhoods" have become the perfect testbed for new technologies in the context of a smart city. With the urban project "Nancy Grand Cœur"[2](NGC), the Grand Nancy local community wants to rehabilitate the boundaries of the historic train station of Nancy (hosting almost 9 million passengers every year), and the surrounding belonging to the city centre, as represented in Figure 1a).

(a) Urban project for NGC by 2020, Source: Arep Ville – J.M. Duthilleul.

(b) Urban area in 2013.

(c) Focus on the "Viaduc Kennedy", around the fixed Air Lorraine pollution station.

Figure 1. Case study of the eco-neighbourhood "Nancy Grand Cœur".

This ecological urban project is intended to be delivered by 2020, and the objectives for the central train station of the city are manifold: new green mobility, traffic regulation, reconciliation between historical and modern neighbourhoods of the city, environment quality and green public spaces, reduced energy consumption, comfortable homes and offices. As the neighbourhood is currently suffering structural reconfigurations that are meant to reduce congestion and increase fluidity, no study has been undertaken so far to analyse the impact of these traffic reconfigurations on the air pollution inside the eco-neighbourhood.

2 www.grand-nancy.org/grands-projets/nancy-grand-coeur/

The traffic analysis in this paper is a continuation of our previous study of the neighbourhood presented in (Mihaita, et al. 2014) and currently submitted for review in (Mihaita, et al. 2015), in which we build the 3D mesoscopic traffic simulation model of the eco-neighbourhood in FlexSim, and propose an evolutionary algorithm for optimising the traffic plan in the most congested intersection of NGC (as marked in Figure 1b)). In this paper, we only focus on the *Viaduc Kennedy*, located near the train station area, and which contains the fixed air pollution station of the local air monitoring association Air Lorraine (Figure 1c)). The main objective is to conduct an initial study in order to simulate, predict and analyse simultaneously the traffic flow and the pollution emissions in this highly congested area. The choice of focusing on this area is also connected to the air quality workstation of Air Lorraine, which provides accurate and real time information for model validation and testing.

Organisational framework

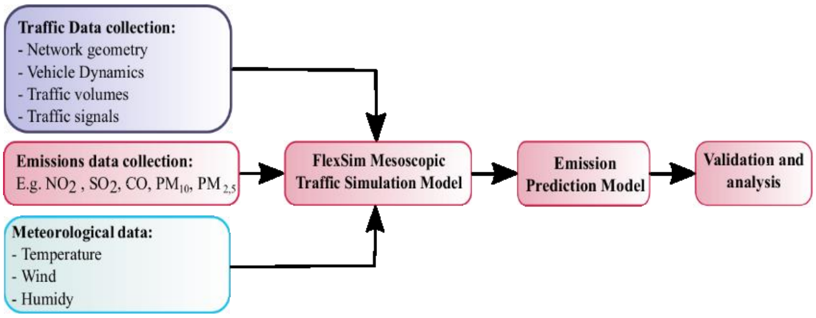


Figure 2. Organisational framework of the article.

Figure 2 summarises the organisational framework of the work proposed in this paper, which contains the traffic simulation model and the emission prediction model. We start by collecting and analysing the real data sets available for our study. For building the 3D traffic simulation model in FlexSim, the local community of Grand Nancy provided: a) the network geometry of the Viaduc Kennedy, b) the hourly traffic volumes during one month period (January 2015) and c) the traffic signals of the intersection. For analysing the emissions registered in the neighbourhood, the local air quality association Air Lorraine provided for us the hourly emissions registered by the fixed air monitoring station (as marked on Figure 1c)), as well as meteorological indicators during the chosen study period, such as wind, temperature and humdity. As 56% of the nitrogen dioxide in the air is caused by road transportation (MEDE, 2012), for this initial study we focus only on the NO_2 concentrations. Using the provided real data sets we build the 3D mesoscopic traffic simulation model, which allows us to obtain statistical measures of the traffic model, such as mean number of cars and average stay-time, and also graphically visualise the evolution of meteorological conditions and NO_2 emissions. Using the outputs of the simulation model, we construct a simplified prediction model for the NO_2, which is later validated through analysis and comparison tests with real-data sets.

Data profiling

The first step in understanding the traffic and air pollution behaviour is to analyse the initial data we have received from the Grand Nancy local community and Air Lorraine association, during the month of January 2015. Figure 3a) presents the evolution of the NO_2 concentrations during 24 hours; we represent every working day of the month in blue, weekends in yellow, and in red the evolution of the daily average NO_2 concentrations. As the behaviour of the pollutant is highly influenced by the traffic in the city, in Figure 3b) we represent as well the hourly evolution of the traffic volume, which we denote Nr_{cars}. As the two graphics show, during the morning and afternoon traffic rush hours (07:00-09:00 and 16:00-19:00), high levels of NO_2 emissions are dispersed in the air and registered by the air pollution station. An interestingly remark is that in two of the NO_2 series corresponding to weekend days, we observe higher NO_2 levels (80 μg/m^3) than in working days (70 μg/m^3).

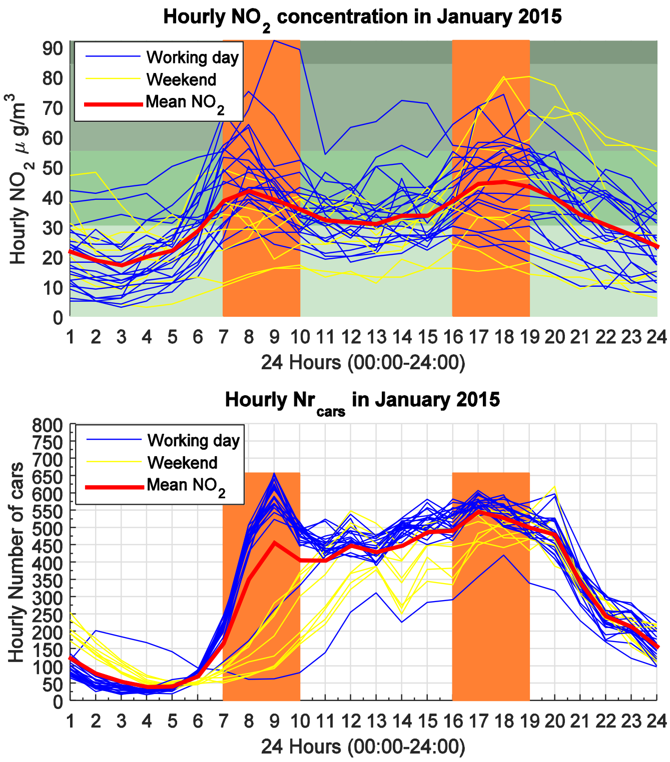
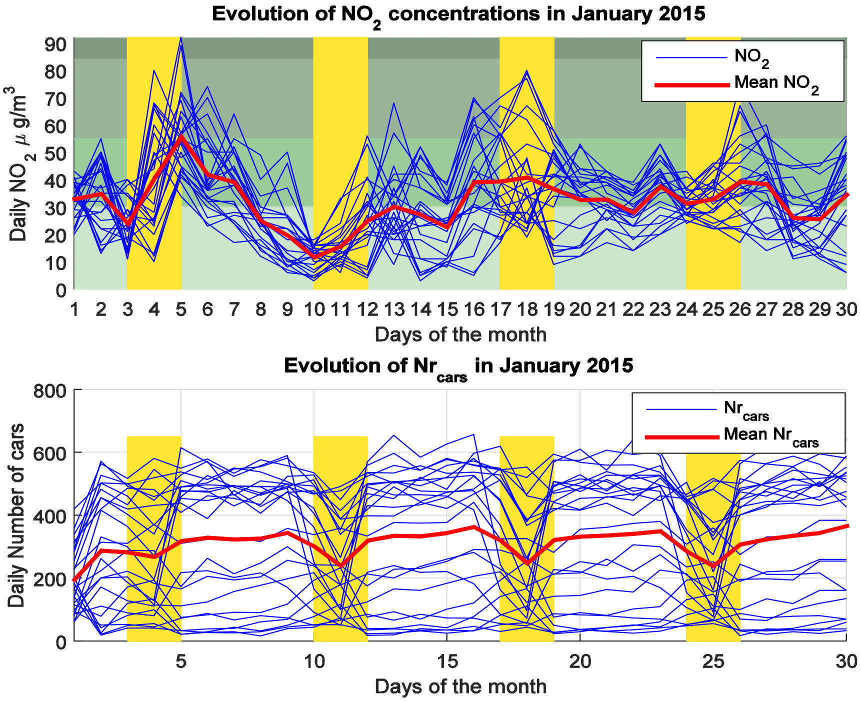


Figure 3. Daily evolution of the NO_2 concentrations and traffic volume in January 2015.

Figure 4a) and 4b) presents the evolution of traffic and NO_2 concentrations over the whole month. The maximal NO_2 concentration took place on the 5th of January 2015, when it reached 90 μg/m^3, an acceptable value according to the air quality index of the European Union (van den Elshout, 2012). By looking at the weekend 10th to 12th of January 2015, one would remark a dramatically NO_2 reduction compared to the previous weekend, and initially conclude to a similar trend during other weekends; the trend does not seem to repeat itself, as the weekend 17th to 19th of January 2015 signalised another increase in NO_2, although the traffic volume was similar to the previous weekend.



g

Figure 4. Monthly data evolution of NO_2 concentrations and traffic volume.

The evolution of $NO_2[\mu g/m^3]$ seems to be as well influenced by other external factors, such as wind$[km/h]$, humidity[%] and temperature [°C], which we consider for the current analysis, as shown in Figure 5. For easing the visualisation of variables with different unit measures, we plot only the normalized values (ex. for temperature: $(T - T_{min})/(T_{max} - T_{min})$).

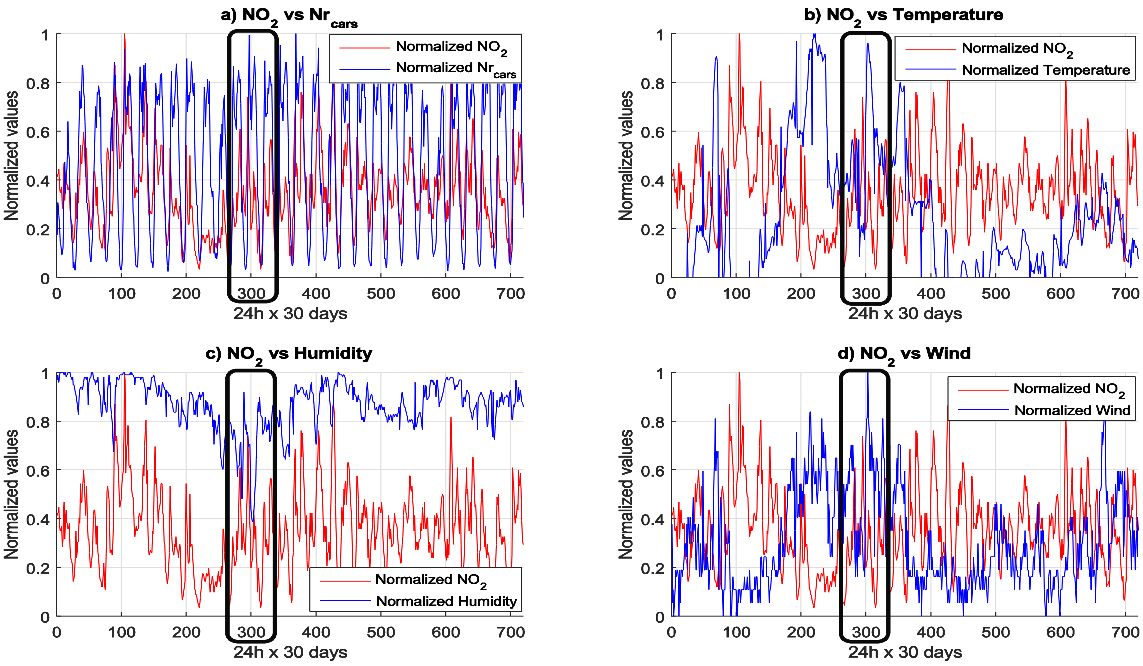


Figure 5. Normalized concentrations of NO_2 versus Nr_{cars}, Temperature, Humidity and Wind.

By closely looking at the highlighted region in Figure 5d) we observe that low NO_2 concentrations usually occur when either the wind is strong, humidity is high or there are few cars on the streets. Conversely, high NO_2 emissions appear when either wind and humidity are low, or there is a higher number of cars on the streets (compared data fields 100 in Figure 5).

To further investigate the stochastic influence of the climate and human factors over the air pollution, we conduct a correlation analysis between these four parameters and the NO_2 concentration, for January 2015, as presented in Table 1. The loose Spearman correlation (0.509) between the NO_2 concentrations and the number of cars, indicates that other factors (apart from the number of cars) seem to inluence pollution, as previously mentioned. Although Wind and Temperature are highly correlated (0.767), they both have a negative correlation with nitrogen dioxide. This confirms to the complexity of the model and the need to further build a regression model which would help predict the changes in the air pollution and help traffic planners and local authorities to test various reconfiguration traffic scenarios.

Indicator	Nr_{cars}	Temperature	Wind	Humidity	NO_2
Nr_{cars}	1	0.12	0.194	-0.185	0.509
Temperature	-	1	0.767	-0.55	-0.441
Wind	-	-	1	-0.55	-0.492
Humidity	-	-	-	1	0.281
NO_2	-	-	-	-	1

Table 1. Spearman correlations between various traffic and air quality indicators.

Simulation model

Traffic flow inside the NGC neighbourhood is highly impacted by the central train station of the city and its particular suspended structure as previously shown in Figure 1c). The study zone we are considering for the traffic simulation contains the junction made by *Viaduc Kennedy* with *Rue Saint Léon*, *Avenue Foch* and *Rue de la Commanderie*. In order to construct the traffic simulation in FlexSim, we import the geometry of the network (Autocad files), the 3D environment build on this geometry, the real number of cars passing through the network in January 2015, as well as the traffic signals conceived for giving priority to the tramway (received from Grand Nancy). Figure 6 shows a snapshot of the 3D mesoscopic traffic simulation model of the study area in FlexSim.

Figure 6. FlexSim 3D simulation model of the Viaduc Kennedy in FlexSim.

The simulation model uses conveyors for representing the streets, sources and sinks for generating and disposing the simulating vehicles in the model, processors for random vehicle insertion in the model every 15 minutes, FlowBin items for pedestrians, and Visual Tools for representing the traffic lights adapted for the tramway stop, as shown in Figure 7a) and 7b).

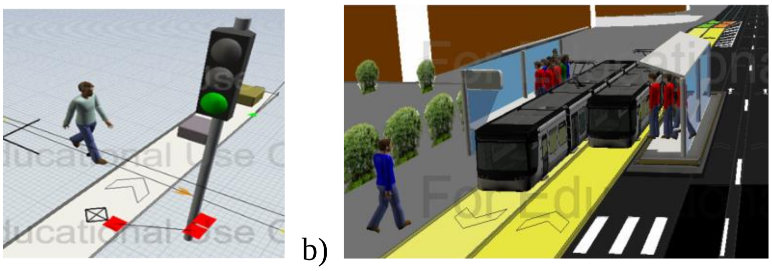
a) b)


Figure 7. a) Traffic light plans representation b) Tramway stop in FlexSim.

Apart from constructing a clear insight of the traffic conditions in this area during congested periods, our efforts have been also oriented towards the integration of emission data and meteorological information in the simulation model, which can be a complementary source of information in the process of traffic and air quality dynamic monitoring. Various air pollution studies use dedicated tools for representing the concentrations of various pollutants in a specific area, such as MISKAM (Microscale Flow and Dispersion Model), which has been used for the air pollution study in the eco-neighbourhood Danube, from Strasbourg, France (ASPA, 2012). MISKAM is a three-dimensional miscroscopic simulation tool which integrates fluid dynamics equations to simulate pollutant concentration as a 3D mesh. Although it can offer precise estimations of various pollutant concentrations, it needs large amount of input data, such as the topography and height of the buildings, the annual traffic emissions computed with Circul'Air (Galineau, 2012), meteorological data including the direction and wind speed, etc. The main limitation we encountered for integrating MISKAM and FlexSim outputs was mainly connected to the differences in the static and dynamic behaviour of the tools. As our initial study is to build a rapid and dynamic simulation model which offers an immediate visual insight on the hourly variation of the NO_2 emissions during January 2015, we integrated the air pollution emissions as entities circulating in the model (as shown in Figure 8a), while the wind, temperature and humidity as visual indicators (Figure 8b). The final traffic simulation model has been validated through manual field measurements. The outputs provide average traffic volumes to be used in the following step of the application.

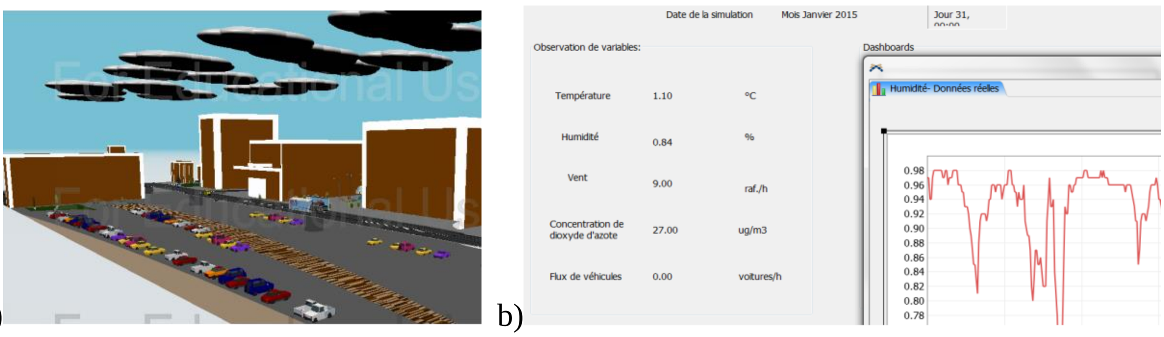
a) b)


Figure 8. a) Visual representation of NO_2 b) Climate indicators dashboard.

Regression model of the NO_2

After validating and building the traffic simulation model, we use the simulated average traffic volumes to build the estimation model of the NO_2 concentrations. We make the hypothesis that all vehicles considered here are contributing to NO_2 concentrations, as no data regarding different types of vehicles circulating in the area is currently available. We start by training a least squares multiple regression model (LSMRM) on the first 20 days of our data set from January 2015, and we further test the results on the temporally holdout 10 days. As an observation, we have chosen LSMRM after first fitting a simple linear regression model that only slightly improved the performance over a baseline. Using the statistical tool Minitab, we obtain the following regression equation, showing the evolution of the NO_2 concentration as a function of the average number of cars on the Viaduc Kennedy, the temperature, the wind and the humidity, which we denote as predictors in this section:

$$NO_2 = 28.43 + 0.05Nr_{cars} - 1.04Temperature - 0.91Wind + 2.04Humidity \qquad (1)$$

As an example from our winter data set, if we consider that no cars are on the road (Nr_{cars} =0), and the Temperature is 2.98°C, Wind =11.28 km/h and Humidity = 85.47%, then we obtain an NO_2 concentration of 29.74$[\mu g/m^3]$, which falls under the average concentration which was 32.17 $[\mu g/m^3]$. Table 2 indicates whether the predictors we have used are independent variables, as shown by the variance inflation factor (*VIF*) (Liao & Valliant, 2012). A predictor with a *VIF* indicator superior to 5 would indicate that it is highly correlated to another predictor. For our study the current results show that the considered predictors are non-correlated, as their *VIF* values are inferior to 5. In order to analyse the weight and influence of each predictor over the NO_2 concentration, we analyse the *P-value* which is based on the *T-test*. *P* determines the appropriateness of rejecting the null hypothesis in a hypothesis test and ranges from 0 to 1. According to (Schlotzhauer, 2007), a predictor with a a *P-value* superior to 0.05 has a high chance of not influencing the initial equation, which is the case of the Humidity variable, with a P value of 0.68. This result indicates that we can stop considering the Humidity variable for further studies. The intercept value (28.4) in equation (1) indicates the presence of other external factors which we haven't included in this study and which might have a strong influence over the NO_2 variation, such as topography, height of the buildings, wind direction, etc. The regression test also returnes a value of $R^2 = 0.67$ for the coefficient of determination, which indicates that the data fits in a good proportion to the regression model.

Predictor	Coefficient	T	P	VIF(Variance inflation factor)	[Min; Max]
Constant	28.43	5.63	0.00	-	-
Nr_{cars}	0.05	21.96	0.00	1.06	[0;657]
Temperature[°C]	-1.04	-5.88	0.00	3.19	[-4.7;13.1]
Wind [km/h]	-0.91	-8.82	0.00	3.30	[0;37]
Humidity[%]	2.04	0.40	0.68	1.71	[38;98]

Table 2. Regression analysis results for the first 20 days of January.

Validation of the prediction model

Using equation (1) we forecast the NO_2 concentrations over the last 10 days of the data set, and compare them to the initial real data set, as shown in Figure 9. According to (Makridakis, et al. 1998), there are various indicators for measuring the accuracy of a forecast, such as the Mean Absolute Percentage Error (MAPE), Mean absolute Deviation (MAD), Mean squared deviation (MSD) or Mean Squared Errors (MSE). For the error accuracy and ease of calculation we use the MAPE, which is defined as:

$$MAPE = \frac{1}{n}\sum_{k=1}^{n}\left(\frac{F_k - Y_k}{Y_k}\right) \tag{2}$$

where n is the number of observation points, Y_k is the actual observation of the studied variable and F_k is the forecasted value. For our study we obtained a MAPE of 23.61%, which validates the forecast model. According to (Barlas, 1994) the MAPE needs to be inferior to 30% for model validation.

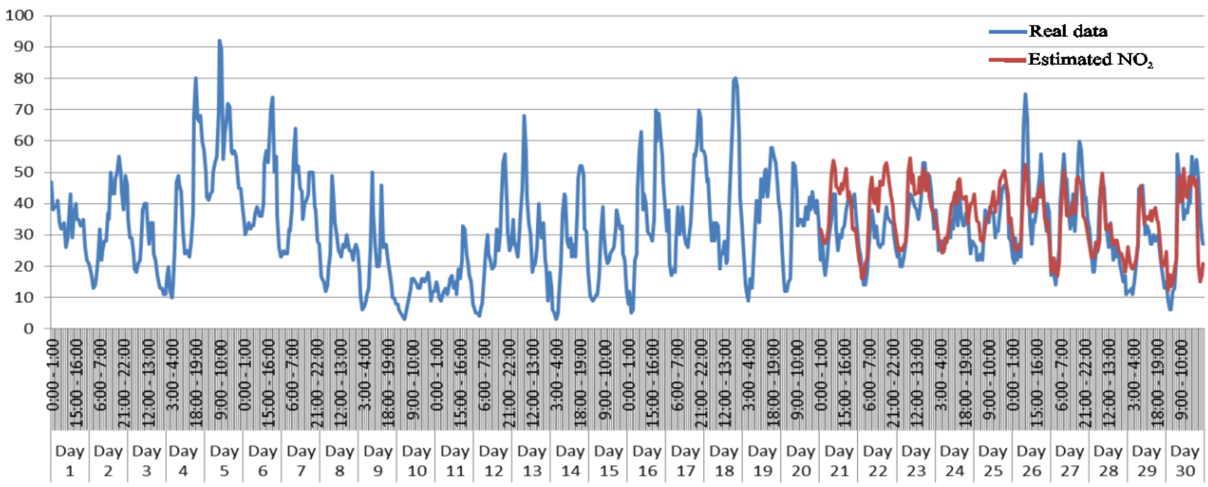


Figure 9. Real versus estimated NO_2 concentrations during the last 10 days of January 2015.

Seasonality study

By analyzing the current results, we observe that one of the independent predictors (traffic volume) is sensitive to time changes during the whole data set we are testing. As previously shown in Figure 3b), the traffic volume has a daily seasonality, with increasingly number of vehicles during the morning and afternon rush hours, and low volumes at night. By conducting a trend analysis and seasonal decomposition on the first 20 days of the data set, we predict as well the number of cars during the last 10 days of January 2015. The best estimations we have obtained is by using a multiplicative model (trend and seasonal components are multiplied and then added to the error component), which returned a MAPE of 21.15%, compared to the aditive model (the effects of individuals factors are differentiated and added together to model the data) which returned a MAPE of 43.6%. After integrating the fitted trend equation of the multiplicative model in equation (1), we obtain the a new regression equation:

$$NO_2 = 28.43 + 0.05((337 + 0,0026 \cdot H) \cdot i) - 1.04 Temperature - 0.91 Wind + 2.04 Humidity \quad (3),$$

where *H* represents the daily hours and *i* the seasonality index per hour. When validating again the NO_2 concentrations over the last 10 days of the data set using equation (3), we obtain a MAPE of 25.39%, which is still below the 30% validation threshold.

Although less accurate than the initial estimation, this result could provide good NO_2 estimations when access to traffic data is limited. The MAPE can be further improved if longer data sets would be available for future studies and comparison. Validating the model on longer periods needs a more detailed analysis and verification, especially if other predictors would be available for the study. The main limitations of this work are: a) restrained data sets we have received concerning the air pollutant emissions and traffic counts; having access to longer periods of data would help building up traffic demand scenarios and analyse for example how traffic control calming measures would improve the air quality in the neighbourhood, b) integrating the emissions impact from the train station would also improve the model.

Conclusions and future perspectives

In this paper, we propose an integrated air pollution and traffic simulation model for building a simplified NO_2 estimation model, inside the eco-neighbourhood Nancy Grand Coeur. We start by a data profiling analysis, before constructing the 3D mesoscopic traffic simulation model. We further use simulated traffic volumes to build the regression model and estimate the NO_2 concentrations over the last 10 days of the model. The results show a good forecasting and are subject to improve if more data would be available. The integrated platform represents a good support for testing various traffic scenarios, such as the limitation of traffic access during rush hours for regular cars, in order to encourage the use of electric vehicles. The simulation platform is a perfect tool for testing as well fire plan optimization methods or reconfiguration scenarios that would help improve the air quality in this highly circulated neighbourhood. We are currently working on extending the emission model to the whole NGC eco-neighbourhood, and build more detailed air pollution emission models which would take into consideration all the types of vehicles on the roads. We are also testing wireless air pollution sensors in the neighbourhood, which offer real-time information of the air pollution at the human level. These sensors can be used on a daily basis for home-to-work journeys and represent an accurate supplementary source of information regarding the air pollution emissions.

Acknowledgements

This work has been developed in the ERPI laboratory, from Nancy France, under the Chaire REVES project funding. The final writing and submission of the paper has been done in the DATA61|CSIRO research laboratory from Sydney, Australia. The authors of this work are grateful for the data and support provided by Grand Nancy, Air Lorraine and FlexSim Conseil.

References

AÏCHI, L. (2015). *Rapport fait au nom de la commission d'enquête sur le coût économique et* *financier de la pollution de l'air.* Senat. Paris: Journal Officiel – Édition des Lois et Décrets du 9 juillet 2015.

ASPA. (2012). *Modélisation de la qualité de l'air sur le futur éco-quartier Danube.* Alsace: Association pour la surveillance et l'étude de la Pollution Athmosphérique en Alsace.

Barlas, Y. (1994). Model Validation in System Dynamics. *International Systems Dynamics Conference,* (pp. 1-10).

Fontes, T., Pereira, S., Fernandes, P., Bandeira, J., & Coelho, M. (2015). How to combine different microsimulation tools to assess the environmental impacts of road traffic? Lessons and directions. *Transportation Research Part D, 34*, 293-306.

Galineau, J. (2012). *Bilan des émissions atmosphériques du transport routier en Lorraine.* Nancy: Air Lorraine.

Liao, D., & Valliant, R. (2012). Variance inflation factors in the analysis of complex survey data. *Survey Methodology, 38*(1), 53-62.

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting Methods and Applications.*

MEDE. (2012). *Bilan de la qualite de l'air en France en 2012.* Paris: Ministère de l'écologie et du développement durable, Direction Générale de l'Energie et du Climat.

Mihaita, A., Camargo, M., & Lhoste, P. (2014). Optimization of a complex urban intersection using discrete event simulation and evolutionary algorithms. *International Federation of Automatic Control World Congress (IFAC WC 2014)* (pp. 1-8). Cape Town, Africa: IFAC Proceedings.

Mihaita, A., Dupont, L., & Camargo, M. (2015, December, submitted). An urban traffic signal optimization using a 3D mesoscopic simulation approach and evolutionary algorithms. *Transportmetrica B: Transport Dynamics*, 1-22.

Schlotzhauer, S. D. (2007). *Elementary Statistics Using JMP.* Cary: NC: SAS Institute Inc.

Shorshani, M. F., André, M., Bonhomme, C., & Seigneur, C. (2015). Modelling chain for the effect of road traffic on air and water quality: Techniques, current status and future prospects. *Environmental Modelling & Software, 64*, 102-123.

UN. (2014). *World Urbanizatin Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352).* Depa7rtment oof Economic and Social Affairs, Poulation Division. New York: United Nations.

US EPA. (2016). *What Are the Six Common Air Pollutants?* Retrieved from http://www3.epa.gov/airquality/urbanair/

van den Elshout, S. (2012). *Citeair II, Common Information to European Air.* European Regional Development Fund, Regional Initiative Project. Bruxelles: European Union.

Xuesong, Z., Shams, T., Hao, L., Taylor, J., Bin, L., Nagui M., R., et al. (2015). Integrating a simplified emission estimation model and mesoscopic dynamic traffic simulator to efficiently evaluate emission impacts pf traffic manageent strategies. *Transportation Research Part D, 37*, 123-136.