18th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
9-12 October 2017, Bologna, Italy

TOWARDS AIRGIS RE-DEVELOPMENT – A GIS BASED AIR POLLUTION AND HUMAN EXPOSURE MODELLING SYSTEM

Jibran Khan[1,2],Matthias Ketzel[1], Konstantinos Kakosimos[2], Ulla Arthur Hvidtfeldt[3], Mette Sørensen[3], Ole Raaschou Nielsen[1,3], Jørgen Brandt[1], Thomas Ellermann[1] and Steen Solvang Jensen[1]

[1]Department of Environmental Science, Aarhus University, Roskilde, Denmark
[2] Department of Chemical Engineering, Texas A&M *University at Qatar, Doha, Qatar*
[3]The Danish Cancer Society Research Centre, Copenhagen, Denmark

Abstract: AirGIS and the semi-parametrized Operational Street Pollution Model (OSPM®) have been frequently used across the globe to assess local or street-scale air pollution. Recent developments in OSPM and new technological advances in geospatial tools have led towards an inception of a substantially revised version of the AirGIS exposure modelling system (AirGIS 2). The new prototype system uses open source software as the geospatial programming tools in a PostgreSQL/PostGIS database in conjunction with R scripts for pre-processing and post-processing of the datasets. This study provides an overview of the re-development, initial testing phase together with strengths, limitations and future prospects of the new system. AirGIS has been re-evaluated against various long-term and short-term field measurements. The present study gives a brief overview of the evaluation and shows the main results and conclusions derived from this validation.

Key words: *AirGIS, OSPM, exposure assessment, GIS, PostgreSQL/PostGIS.*

INTRODUCTION

The Danish AirGIS – a GIS based air pollution and human exposure modelling system – was developed and validated for the assessment of local- or street-scale air pollution at many address locations (Jensen et al., 2001; Ketzel et al., 2011). AirGIS is routinely used in conjunction with the integrated air quality modelling system THOR (Brandt et al., 2001) for mapping of air pollution across Denmark (http://luftenpaadinvej.au.dk/) (Jensen et al., 2017) and exposure assessment in epidemiological studies (e.g., Andersen et al., 2016). Recently, the OSPM has been updated e.g. for inhomogeneous emissions which are not considered in the former AirGIS implementation (Ottosen, 2016). Moreover, recent advances in Geographic Information System (GIS) tools allow for a more effective handling of large spatial data sets compared to the older version of AirGIS based on the no longer supported programming language "Avenue". Thus, we developed a substantially revised version of the AirGIS system to reflect upon these latest developments.

The new AirGIS has been implemented in open source software PostgreSQL/PostGIS in conjunction with R scripts for pre- and post-processing of the datasets. The main objectives of using open source software and geospatial programming have been to (1) optimize the model performance over a large geographical region with limited computing resources (2) comply with the latest developments as well as future adjustments of the model as per need/requirements (3) replace the outdated programming language Avenue and (4) become independent of commercial GIS software. The present study describes the overall structure of the new system together with its strengths, limitations and future prospects. Furthermore, the newly developed AirGIS has been evaluated against various measurements and the results are summarized herein.

METHODOLOGY

AirGIS generates input parameters for OSPM based on national GIS data sets to assess air pollution at any address location in Denmark. AirGIS integrates in the latest version three dispersion models of the THOR system operating at three pollution levels or spatial scales: (1) local scale air pollution – modelled with OSPM (2) urban background concentrations represented in a 1 x 1km^2 grid – modelled with UBM and (3) regional background concentrations – modelled with DEHM. Air pollution concentrations are estimates for each specific location by merging the contributions from the relevant scales.

The re-developed AirGIS makes use of GIS programming tools in a PostgreSQL database (https://www.postgresql.org/) with GIS functions provided by its spatial extension PostGIS (https://postgis.net) in conjunction with scripting language R (https://cran.r-project.org/) for pre-processing and post-processing of the datasets. In terms of processing speed for spatial operations, PostGIS has greater performance than other common desktop GIS applications due to its efficient use of spatial indexing (Gulliver et al., 2015). Figure 1 shows an overall structure and dataflow of the new system.

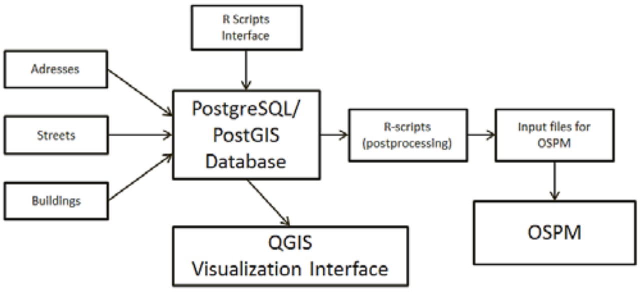

Figure 1. An overall structure and dataflow of the new AirGIS system.

In Denmark, road traffic information is based on a national traffic database in form of a GIS multi-line string shapefile while information on building foot prints is available as a multipolygon shapefile from the Danish Geodata Agency (gst.dk) and geocodes of address locations are usually obtained via the Central Person Registry. AirGIS imports this available GIS information (addresses, streets and buildings) into a PostgreSQL database with its spatial extension, PostGIS that is performing all GIS calculations (Figure 1). R scripts (via R-Studio) provide a relatively user-friendly interface to query data from PostgreSQL database as well as to execute PostGIS commands. The main aim of using R software has been (1) to strengthen the new prototype by its significant statistical computing capabilities (2) to perform preliminary data analyses as soon as PostGIS commands execution is completed (3) to provide a single flexible environment to perform all data handling from preprocessing via running GIS queries to post-processing and export to OSPM and UBM model runs. Finally AirGIS generates the output which is further processed through R scripts (post-processing) to produce input files for OSPM calculations in required format. QGIS software serves as the visualization interface during the whole process i.e., both input and output files can be readily visualized on the fly.

Model evaluation is indispensable for reliable air pollution exposure estimates. AirGIS has been re-evaluated against various available measurements datasets, shown as overview in Table 1. In this paper the evaluation against dataset 1 (at selected stations) and dataset 2 (Figure 2) is presented and briefly discussed. The oral presentation will give an overview of further results.

Table 1. An overview of various measured air pollution datasets used for AirGIS evaluation

Dataset	Name	Pollutant	Measurement site	Location	Time Period
1	Danish Air Quality Monitoring Programme (NOVANA)	NOx, NO_2, PM_{10}, $PM_{2.5}$	6 permanent urban street stations	Copenhagen, Odense, Aarhus, Aalborg	1994 - 2015
2	ESCAPE-EU Danish measurement campaign	NOx, NO_2, PM_{10}, $PM_{2.5}$	41 streets (20 streets for PM)	Copenhagen	November 2009 – October 2010
3	Five weeks passive sampling campaign	NO_2	10 major streets	Copenhagen	October 24 – November 28, 2011

Modelled concentrations of air pollution were compared with measured. The receptor point for modelling was at 2 meters height at the façade of the building closest to the address point. All calculations were performed on an hourly basis and concentrations were averaged over the time period corresponding to the measured data. It should be noted that the modelling of PM concentrations is still under development within the AirGIS system. New components of the PM mass as e.g. secondary organic aerosol (SOA) have recently been added. However a comparison of modelled PM_{10} and $PM_{2.5}$ concentrations against measurements using the EU-reference method to determine PM at the permanent stations of the Danish Monitoring Program (Ellermann et al., 2016) reveals that the model still underestimates the measured PM. The reason for this is most likely a remaining underestimation for some of the PM constituents such as primary organic PM or water content in PM. In order to compensate for the underestimation by the model we applied correction factors of 1.46 and 1.26 to our modelled PM_{10}, $PM_{2.5}$ values in this study. Furthermore, to determine the strength of linear relationship between modelled and measured values, Pearson's correlation coefficients (r) were computed. All statistical analyses were performed in R version 3.4.0 (https://www.r-project.org/). Results have been summarized and discussed in the following section.

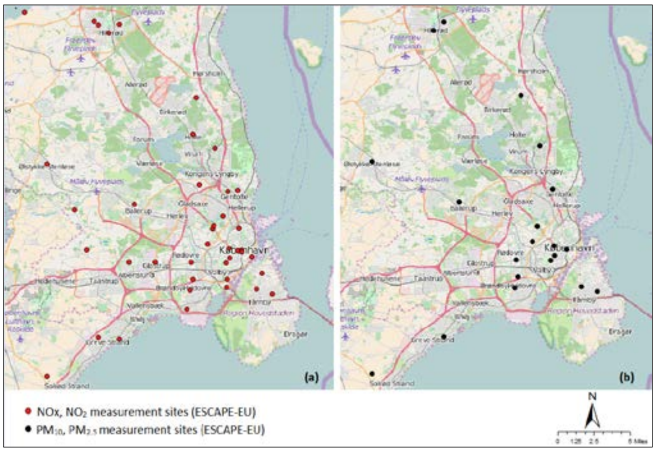

Figure 2. ESCAPE-EU measurement sites in Copenhagen, Denmark. (a) 41 sites for NOx, NO_2 measurement. (b) 20 sites for $PM_{2.5}$ and PM_{10} measurement (Background map ©OpenStreetMap)

RESULTS AND DISCUSSIONS

EVALUATION AGAINST MEASUREMENTS DATASET 1

In this section the evaluation against measured NOx and NO_2 (µg/m^3) at one of the permanent measurement stations i.e., Jagtvej (JGTV) is presented. Results for model runs at JGTV over 22 years are given as annual averages in Figure 3. A significant decrease in NOx levels at JGTV can clearly be observed. In general, this trend is well reproduced by AirGIS over a significant period of time.

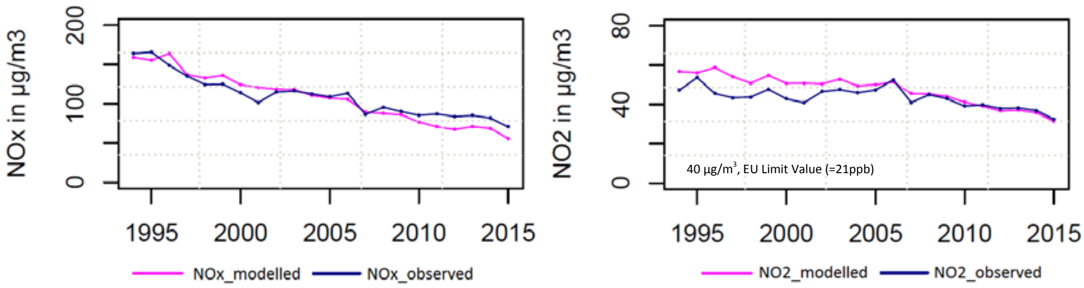

Figure 3. Trends of annual averages for observed and modelled NOx (left) and NO_2 (right) (µg/m^3) at JGTV (1994 – 2015)

For some years e.g., 1997 – 2002, the model overestimated NOx concentrations while for 2007 – 2015 the model underestimated the observed values. It should be noted that, in general, long-term trends in NOx concentrations at street level are mainly influenced by the changes in traffic emissions from light and heavy duty vehicles and uncertainty in this information might result in significant deviation between modelled and measured concentrations considering single years. The Pearson's correlation coefficient between long-term annual average modelled and measured concentrations (1994 – 2015) of NOx at JGTV is 0.95 indicating a very good agreement. For NO_2 the observed concentrations have been nearly constant over the past two decades, with a slight tendency for a reduction (Figure 3). Likewise NOx, long-term NO_2 trends are also well reproduced by AirGIS. A good match between measured and modelled NO_2 values is observed for 2008 – 2015 while the model overestimated for 1996 – 2003. The correlation coefficient between modelled and measured NO_2 values (1994 – 2015) was found to be 0.83.

EVALUATION AGAINST MEASUREMENTS DATASET 2

Figure 4 shows the scatterplots of modelled NOx and NO_2 against their measured values while Figure 5 shows the same plots for PM_{10} and $PM_{2.5}$. Moreover, in Table 2 descriptive statistics on observed (obs.) and modelled (mod.) values of air pollutants is presented. As observed, the modelled concentrations of NOx and NO_2 (Figure 4) appear more scattered with a few key outliers as compared to the plots of PM_{10} and $PM_{2.5}$ (Figure 5). The correlation coefficients between modelled and measured NOx and NO_2 concentrations were found to be 0.74 and 0.73 (Table 2). There seems to be a better agreement between measured and modelled values of NO_2 as compared to NOx. On the other side, correlation coefficient values between modelled and measured PM_{10} and $PM_{2.5}$ (in µg/m^3) were found to be 0.73 and 0.80, respectively. For all compounds, the model overestimates the concentrations compared to the observed values. For NO_2 and NOx the overestimation is quite large which is in contrast to the comparison with dataset 1, where the agreement was much better. The reason for the overestimation with dataset 2 has to be further investigated and might be caused by measurement uncertainty or bias due to the applied low-cost portable samplers.

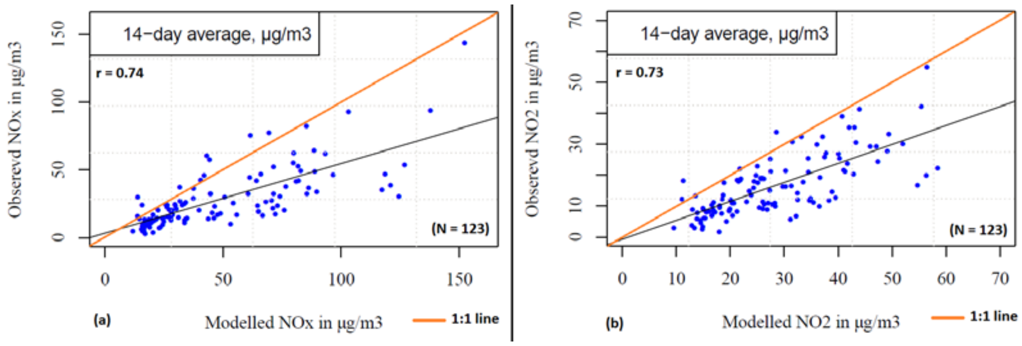

Figure 4. Evaluation of AirGIS against dataset 2. (a) Modelled average concentrations of NOx (in µg/m3) plotted against observed NOx concentrations (N=123). (b) Same plot for NO_2 (in µg/m3) (N=123).

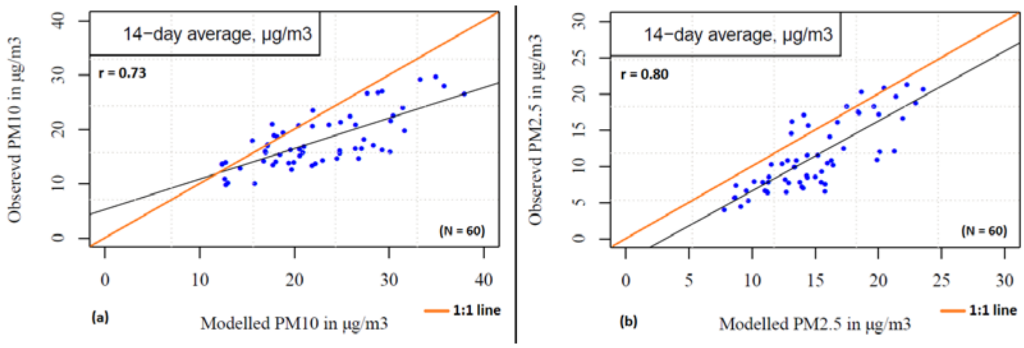

Figure 5. Evaluation of AirGIS against dataset 2. (a) Modelled average concentrations of PM_{10} (in µg/m3) plotted against observed PM_{10} concentrations (N=60). (b) Same plot for $PM_{2.5}$ (in µg/m3) (N=60).

Table 2. Descriptive statistics on observed (obs.) and modelled (mod.) values of air pollutants: NOx, NO_2, $PM_{2.5}$, PM_{10} (in µg/m^3) for Dataset 2, ESCAPE-EU measurements period: November 2009 – October 2010. Av = average, Min = minimum, Max = maximum, CoV = coefficient of variation within the dataset, MB = mean bias, RMSE = root mean squared error, r = Pearson's correlation coefficient (r)

Pollutant	Method	Av	Min	Max	CoV	MB	RMSE	r
NOx	obs.	28.1	3.1	143.8	0.77			
	mod.	48.3	11.7	152.2	0.65	19.7	29.1	0.74
NO_2	obs.	16.4	1.7	54.8	0.59			
	mod.	28.1	9.5	58.4	0.41	11.3	14.1	0.73
PM_{10}	obs.	17.8	9.9	29.7	0.27			
	mod.	22.5	12.3	37.9	0.28	4.6	6.3	0.73
$PM_{2.5}$	obs.	11.3	4.0	21.3	0.42			
	mod.	14.9	7.8	23.6	0.27	3.5	4.6	0.80

The strengths and limitations of AirGIS should also be considered. While modelling exposure to air pollution for a large number of locations, there will remain some exposure misclassifications, due to uncertainties in the input data and the model itself. On the positive side, implementation of AirGIS in PostgreSQL/PostGIS offers excellent efficiency in terms of processing time i.e., the new system is 200 times faster than the former AirGIS system in generation of traffic and street configuration data for OSPM. Moreover, due to the use of open source GIS programming tools the new AirGIS can be easily adjusted for future requirements and further developments.

CONCLUSION

The re-developed GIS-based air pollution and human exposure modelling system, AirGIS, has been evaluated against various measured datasets. The strength and limitations of the new system have been highlighted. Being in its initial evaluation phase, AirGIS showed satisfying performance in terms of correlation coefficient against measured values and gave a reasonable good description of the spatial variation of ambient air pollution (e.g., NOx, NO_2, $PM_{2.5}$, PM_{10} in this study) within a region of interest. It is therefore well suited for air pollution modelling and health related studies based on air pollution exposure assessment. However, the bias between measurements and model is still relatively large (especially with dataset 2) and needs to be reduced in further model developments.

More work has to be done in the future (i) to evaluate the new system for other pollutants (CO, Black Carbon, O_3 etc.) (ii) to further test street configuration generation of the new system (iii) to implement inhomogeneous emissions scenario once relevant datasets are available.

ACKNOWLEDGEMENT

This work is based on a PhD study funded by the Graduate School of Science and Technology (GSST), Aarhus University, Denmark and NPRP award [NPRP 7 - 649 - 2 - 241] from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.

REFERENCES

Andersen, Z. J., Ravnskjaer, L., Andersen, K. K., Loft, S., Brandt, J., Becker, T., . . . Brauner, E. V. (2016). Long-Term Exposure to Fine Particulate Matter and Breast Cancer Incidence in the Danish Nurse Cohort Study. Cancer Epidemiology and Prevention Biomarkers, cebp. 0578.2016.

Brandt, J., J. H. Christensen, L. M. Frohn and R. Berkowicz, 2001. Operational air pollution forecast from regional scale to urban street scale. Part 1: system description. Physics and Chemistry of the Earth (B), Vol. 26, No. 10, pp. 781-786, 2001.

Gulliver, J., Morley, D., Vienneau, D., Fabbri, F., Bell, M., Goodman, P., . . . Fecht, D. (2015). Development of an open-source road traffic noise model for exposure assessment. Environmental Modelling & Software, 74, 183-193. doi:http://dx.doi.org/10.1016/j.envsoft.2014.12.022

Jensen, S. S., Berkowicz, R., Hansen, H. S., & Hertel, O. (2001). A Danish decision-support GIS tool for management of urban air quality and human exposures. Transportation Research Part D: Transport and Environment, 6(4), 229-241.

Jensen, S. S., Ketzel, M., Becker, T., Christensen, J., Brandt, J., Plejdrup, M., . . . Ellermann, T. (2017). High resolution multi-scale air quality modelling for all streets in Denmark. Transportation Research Part D: Transport and Environment, 52, Part A, 322-339. doi:https://doi.org/10.1016/j.trd.2017.02.019

Ketzel, M., Berkowicz, R., Hvidberg, M., Jensen, S. S., & Raaschou-Nielsen, O. (2011). Evaluation of AirGIS: a GIS-based air pollution and human exposure modelling system. International Journal of Environment and Pollution, 47(1-4), 226-238.

Ottosen, T.-B. O. (2016). Improved Local Air Quality Modelling for Complex Urban Landscapes. Syddansk Universitet. Institut for kemi-, bio- og miljøteknologi. Retrieved from http://findresearcher.sdu.dk/portal/files/118859509/Thor_Bj_rn_Ottosen.pdf

Ellermann, T., Nygaard, J., Nøjgaard, J.K., Nordstrøm, C., Brandt, J., Christensen, J., Ketzel, M., Massling, A. & Jensen, S.S. (2016) The Danish air quality monitoring programme. Annual summary for 2015. Aarhus University, DCE – Danish Centre for Environment and Energy, 65 pp. Scientific report from DCE – Danish Centre for Environment and Energy No. 201.