The influence of meteorology of indoor $PM_{2.5}$ concentration: application of advanced modelling techniques

L. Falcão[1], S. Clifford[2], M. Mazaheri[2], D. Broday[3], L. Morawska[2]

[1] Federal University of Paraná, Curitiba 80060-000, Brazil. [2] Queensland University of Technology, Brisbane 4000, Australia. [3] Technion – Israel Institute of Technology, Haifa 32000, Israel.

SUMMARY

Analysis are presented of indoor/outdoor air pollution measurements in an office located in Haifa, Israel, and the influence of meteorology on the concentrations of measured indoor air quality parameters. Generalized Additive Models (GAMs) and Classification and Regression Trees (CARTs) were applied to examine temporal trends in concentrations of several pollutants, the impacts of meteorology on the concentrations, and to determine which parameters are most relevant to explaining the indoor concentrations of outdoor generated pollutants. Wind direction was found to have a significant influence on indoor concentrations of all the air pollutants, while temperature variation was linked to increased levels of PM_{10}, NO_x, NO_2 and SO_2. CART's and GAMs explained the influence of the meteorological parameters, as well as computed the smooth, periodic annual trend.

PRACTICAL IMPLICATIONS

The approach developed can be utilized in analyzing the indoor/outdoor relationship in other settings, and the information obtained on the impact of methodology on the trends in the I/O relationship – towards building design and management of indoor air pollution.

KEYWORDS

GAM, Decision in tree, meteorology, Air pollution, Indoor air quality

1 INTRODUCTION

GAMs and CART are two semi-/non-parametric statistical modelling techniques which are useful for prediction of air pollution. CARTs are most commonly used to model complex variation in large data sets and maximize predictive power rather than describing the functional relationship between an outcome variable and its predictors [1, 3]. In contrast, GAMs are used to describe a relationship whose functional form is not known [2].

2 MATERIALS/METHODS

Measurements of indoor and outdoor NO_2, NOx, NO, O_3, SO_2 and PM_{10} were collected simultaneously half-hourly during the calendar years 2004-2007 at an office building in Haifa, Israel, along with outdoor meteorological parameters: wind speed and direction, precipitation and relative humidity. Wind speed and direction were transformed from polar to Cartesian coordinates, *Vx* and *Vy*, wind velocity in the *x* (East to West) and *y* (North to South) directions. Temperature was transformed to be the deviation from the monthly average temperature across the four years of data, to decouple temperature and day of the year to allow temporal variation to be accounted for with the season of the year (derived from month of the year). For each pollutant, the outdoor concentration of that pollutant was used as an explanatory variable. To determine which parameters were most relevant to explaining differences in indoor air quality, prediction from both GAMs and CARTs were performed. The tree with minimal deviance was found and pruned to reduce complexity while retaining predictive power. GAMs were fitted with the option for variables to be de-selected from fitting.

3 RESULTS

A CART was fit for each indoor pollutant (Table 1).

Table 1: Tree resulting from CART analysis illustrating the predictors, its respective level of importance for each Indoor air pollutants and the cross-validation R^2 value for each model.

Indoor pollutant	*Season*	*Vy*	*Vx*	*Temp. deviation*	*RH*	*Precipitation*	R^2
O_3	29	39	12		11		0.47
PM_{10}	29	14	12	34	10		0.29
NO_2	31	21	19	19			0.13
NO	16	39	37		8		0.34
NOx	31	29	28	13			0.22
SO_2	46	13		41			0.08

4 DISCUSSION

Using CART, *Season* and *Vy* were present in all models. GAM showed *Season* as a very important predictor when plotting indoor concentrations against date, presenting variability in concentrations mainly between summers and winters. Wind direction was already expected to be an important predictor for indoor pollutant concentrations since there were no sources of pollutants in the office building, and wind is one of the significant ways of air pollutant transport. Both *Vy* and *Vx* were considered important predictors for the models, however *Vy* has a larger influence for all of the compounds and *Vx* was not found as a predictor for SO_2. PM_{10} was less influenced by Wind than the other air pollutants. Temperature was found as a strong contributor for indoor levels of PM_{10}, NO_2, and SO_2. Relative Humidity was considered a predictor for O_3, PM_{10} and NO models. With cross-validated R^2 values, the O_3 model had the highest degree of goodness of fit and the most predictors. NO and PM_{10} R^2 values indicated a moderate amount of fit. SO_2 was not fit well by the CART model.

5 CONCLUSIONS

GAM and CART offer an alternative to regression approaches and can be well-suited for identifying complex patterns of joint effects in the big data.

6 REFERENCES

[1] Breiman L., Freidman J., Olshen R., and Stone C. 1984. Classification and Regression Trees, Monterey:Wadsworth and Brooks.

[2] Clifford, S., Low Choy, S., Hussein, T., Mengersen, K., Morawska, L. 2011. Using the Generalised Additive Model to model the particle number count of ultrafine particles, Atmospheric Environment 45: 5934-5945,

[3] Quinlan J. 1986. Introduction of decision tree, Machine Learning Journal, 1:.81-106.