1. INTRODUCTION

In many urban centres around the world, air pollution is an important environmental concern. Given that vehicles are a significant source of emissions into the atmosphere, pollution concentrations can be reduced by changing the patterns of vehicle use, particularly in urban areas where concentrations tend to be high. A reliable method of being able to forecast periods when air quality would otherwise be poor would allow regulatory authorities and individuals to change their daily activity patterns (i.e. choose lower emitting forms of transport) to minimise emissions and hence the concentrations of air pollutants and environmental impacts.

Our present ability to model pollution concentrations in urban areas on days of poor air quality using theoretical modelling is limited, primarily due to the complexity of surface wind flows in urban areas resulting from the complicated topography (buildings, etc.) and emission patterns. This is particularly the case near intersections where the spatial variability in concentrations is high and the vehicle flow kinematics are complicated (Hoydysh and Dabberdt, 1994). However, it has been shown that there is a simple one-to-one relationship between wind speed and direction at particular sites (but which changes from site to site) and the resulting air pollution concentrations (Dirks et al., 2002) since wind speed is one of the most significant factors influencing pollutant concentrations (Comrie and Diem, 1999). Previous studies have shown that, if the surface winds are known for a particular site, the pollution concentrations can be estimated reliably. Dirks et al., 2002, and Dirks et al., submitted demonstrated this using archived CO, PM_{10}, NO and NO_{2} data, as well as local wind measurements for both a site in a flat urban area in New Zealand and for an urban site in mountainous terrain in Northern Italy. In the present paper, we refer to this model as the Site-Optimised Semi-Empirical (SOSE) model.

This one-to-one relationship between the surface wind and air pollution levels could be used to forecast pollution concentrations if a reliable method could be found to “down scale” the mesoscale model predicted surface winds to the local winds at pollution monitoring sites. This is the subject of the present paper. The paper begins with an evaluation of the statistics used to optimise the SOSE model parameters, and evaluate model performance. A comparison of different model statistics is presented, along with the rationale for choosing to use the RMSE for model optimisation. With the SOSE model optimised and evaluated, forecasting using the SOSE-MM5 combinations of models is carried out and the results presented.

1.1 Study location

The current study location is the Aosta Valley, located in the north-west of Italy. A map showing the location of Aosta within northern Italy is given in Figure 1. The data used were collected at a site in the centre of the town of Aosta, the largest in the region and located at the intersection of the Aosta valley and one of its major tributaries.

![Figure 1: Map of northern Italy showing the location of the monitoring site at Aosta.](image)

The data set includes meteorological information (wind speed, wind direction, ambient temperature, rainfall, solar radiation, atmospheric pressure and humidity) as well as a variety of commonly monitored air pollutants (CO, NO, NO_{2}, PM_{10}, O_{3}, SO_{2} and several organic compounds).
2. MODELS

This project involves combining two models: the Site-Optimised Semi-Empirical (SOSE) air pollution model developed by Dirks et al. (2002) to relate the surface winds to the pollutant concentrations, and the PSU/NCAR fifth generation mesoscale model (MM5) for providing mesoscale surface wind forecasts from the synoptic wind patterns (Grell et al. 1994). The SOSE-MM5 combination is used to investigate the feasibility of predicting site-specific air pollution concentrations based on forecasted surface winds provided by the mesoscale model.

2.1 Site-Optimised Semi-Empirical Air Pollution Model

SOSE is based on a simple box model and the assumption that for monitoring stations located next to busy roads, most of the pollutants originate from this particular road with only minor contributions from other roads in the vicinity. The emission rate $Q$ is assumed to be constant along the road, and the pollutants well-mixed within a two-dimensional box of height $\Delta z$ (Hanna et al., 1982). The wind speed $u$ is assumed to be uniform within the box, and at an angle $\theta$ to the road. The background concentration from sources other than the road is given by $C_B$. The concentration of the pollutant in the box, $C$, is given by:

$$C = \frac{Q}{(u \Delta z \sin \theta)} + C_B$$

It has been shown (Dirks et al. 2002) that for such roadside sites, the exact direction of the wind has only a negligible impact on the pollutant concentrations. The important consideration is whether the direction causes the monitoring site to be downwind (wind passes over the road adjacent to the monitor before passing by the monitoring station) or upwind of the road (the wind passes the monitor before passing over the road).

In order to optimise the SOSE model, data are separated by time of year (spring, summer, autumn, early winter and late winter) as there were found to be seasonal differences in the relationship between meteorological conditions and pollutant concentrations. Weekends and weekdays are treated separately, since the commuting behaviours (and hence emission patterns) are significantly different between weekends and weekdays. To determine the relationship between wind and pollutant concentration, a linear regression is carried out on the function

$$C = \frac{Q}{(\Delta z(u + u_0))} + C_B$$

separately for each time of day, for both upwind and downwind cases. The regression parameters are $Q/\Delta z$ and $C_B$. A small factor $u_0$ is added to the denominator to prevent the prediction of infinite concentrations at low wind speeds, and is optimised by the minimisation of the RMSE of the fit to the measured data.

This way, the model can be trained to provide estimates of pollutant concentrations at specific sites using only wind speed and direction information as input (as well as the time of day, the season and whether it is a weekday or weekend).

2.2 MM5

MM5 is used to provide forecasted surface wind information to the SOSE model. The extreme topography of the Aosta Valley region requires MM5 to be run at high spatial resolution. Four nests are used with resolutions of 90km, 30km, 10km and 3.33km. The model is initialised with forecast data from the UK Meteorological Office Unified Model and SAA global SST analysis data sets for 2000 and 2001, which coincide with the available air pollution data archive.

The 10m $u$ and $v$ winds forecast by MM5 at the location of Aosta are combined to give wind speed and direction data which are then input into the SOSE model. The performance of the SOSE model using MM5-predicted winds can then be compared with model output using measured winds, and observed concentration data.

It is also believed that the existence of temperature inversions may have a significant effect – limiting the effective box height in the SOSE model. Such conditions are of particular interest as they are often associated with high air pollution levels. We believe that MM5 is capable of identifying the existence of inversions providing appropriate radiation code and sufficient vertical levels are used.

Once a statistical evaluation of the MM5-SOSE model has been completed for the Italian case study, it is intended that the combined model be trialled using data from a different location, such as Auckland, New Zealand, where effects such as scavenging of pollutants by precipitation may affect the performance of the model (Peace et al., submitted). If scavenging by precipitation is found to have a significant effect on the accuracy of the SOSE model, it may be necessary to add an appropriate function. In this case the precipitation field from MM5 may be used in addition to the predicted 10 m winds to allow the SOSE model to run predictively.
3. EVALUATION OF MODEL PERFORMANCE

Since the SOSE model requires optimisation to fit the parameters, particularly the wind speed offset parameter, it is important to choose an appropriate model evaluation statistic (such as the RMSE) with which to optimise the model, as this can have a significant effect on the resulting model predictions.

For the prediction of air pollutant concentrations, it is most important that peak values are predicted correctly. For this reason, the RMSE statistic has been found to be appropriate.

Figure 2 shows the daily average patterns of measured and SOSE predicted CO concentrations, and the daily average RMSE (and the mean absolute error (MAE), mean bias error (MBE), and fractional RMSE (FRMSE) for comparison). It can be seen that the FRMSE is very high during the period of low concentration between days 25 and 30 and for this reason is a poor measure.

The MBE is inconvenient for model optimisation because it may be positive or negative. The RMSE was preferred over the MAE because a comparison of times of high RMSE or MAE with a qualitative assessment of ‘bad’ model performance (large discrepancy at high concentration) showed best agreement for the RMSE.

4. RESULTS

An example of one week of observed and SOSE-predicted CO concentrations is given in Figure 3. The average RMSE for the season (Autumn 2001) was found to be 0.233. Similar results were found for the other seasons with average RMSE values being higher during the winter months compared to the summer months, due to the higher average concentrations at this time of the year. The average RMSE values for CO for each season in the two years 2000 and 2001 are shown in Table 1.

5. CONCLUSIONS

The RMSE was found to be a suitable model performance statistic with which to optimize the SOSE model. All optimisations were carried out using this statistic.

The SOSE model is able to estimate pollutant concentrations based solely on wind speed and direction information. Using the 10m winds from the MM5 model as the input wind information, it is possible to predict pollutant concentrations with no measured data from the site. Therefore, if the concentration data for a single site exists over a representative period of time, including weekends and weekdays for each of the five seasonal periods used by the model, the MM5-SOSE system can predict pollutant concentrations.
concentrations without continued monitoring at the site. A single air pollution monitor could be moved annually to obtain year-long data sets for several sites. Forecasts could then be made for multiple sites, using only one monitor.

It is believed that the combination of the SOSE air pollution model and the MM5 mesoscale model will allow the short term forecast of bad pollution days.

6. ACKNOWLEDGEMENTS

We would like to acknowledge APRA VdA, the Aosta Valley Regional Environmental Authority, and in particular G. Pession and M. Zublena for providing the data required for this study. The authors would also like to thank A. Nanni for his useful discussions and ongoing support of our work.

7. REFERENCES


Dirks, K. N., Nanni, A., Dirks, V.I., Modelling and Predicting Urban Atmospheric Pollutants in the Aosta Valley Region of Italy using a Site-Optimised Model, Submitted to Atmospheric Science Letters, April 2005.


