2.08 AIR POLLUTION FORECAST IN PORTUGAL: A DEMAND FROM THE NEW AIR QUALITY FRAMEWORK DIRECTIVE

Alexandra Monteiro[1]; Robert Vautard[2]; Myriam Lopes[1]; Ana Miranda[1] and Carlos Borrego[1]
[1]Departamento de Ambiente, Universidade de Aveiro (UA), Aveiro, Portugal
[2]Laboratoire de Meteorologie Dynamique (LMD), Ecole Polytechnique, Palaiseau, France

INTRODUCTION

Air quality forecasting is both a challenge and a scientific problem, which has recently emerged as a major priority in many urbanised and industrialised countries due to the increasing consciousness of the effect, on health and environment, caused by airborne pollutants emissions. Furthermore, is one of the requirements of the new Air Quality Framework Directive (96/62/CE) and a key issue of the CAFE Programme. The goals of reliable air quality forecasts are obvious: population exposure can be more efficiently reduced and protected by means of information and real-time emission abatement measures. So far, the numerical tools used operationally for air quality forecasting are essentially statistical models, based on relations between a set of meteorological predictors and concentration values at different monitoring stations. Statistical forecast uses a variety of mathematical models such as regression methods, time series filtering, cluster analysis and artificial neural networks. These probability approaches proven to be very efficient in many cities (*Simpson and Layton,* 1983, *Gardner and Dorling,* 1998, etc). However, statistical forecasting presents several limitations such as the lack of diagnosis capabilities, map representation, and the requirement of large stationary training data sets for tuning the model's coefficients, which is a limiting constraint for the specific case of Portugal. In fact, the source of pollution cannot be identified by sensitivity experiments, which is a major drawback when concrete actions are to be taken in real time. By contrast, models based on physical equations (called "deterministic models"), driven by the chemistry and the transport of pollutants, require numerous accurate input data (emissions, meteorology, land cover), which are difficult to collect in real time. Also, fine-tuning of physical model parameters and validation of the model are long processes, due to the large computer time required. These latter problems should become less significant with time, due to increases in computer capabilities and to improvements in the data bases required for the deterministic approach of forecasting. The application in real time and the validation of the CHIMERE numerical chemistry-transport model (CTM), for Portugal domain, during Summer 2003 (1 June to 30 September) are presented and analysed. This CTM has been already tested and validated for the Paris region and for Europe (*Vautard et al.*, 2000; *Schmidt et al.*, 2001). The validation is performed by a direct comparison with measurements from the Portuguese air quality monitoring network.

THE FORECASTING SYSTEM

The operational design

The forecasting system is designed to be as simple as possible in order to fit the real-time constraints and to deliver forecasts in the early morning for the same and the next 2 days. Meteorological forecasts are obtained at date D+0 using the MM5 mesoscale meteorological model forced by the AVN/NCEP global forecasts. Every day (D+0) the forecasts start the day before (D-1) at 00UT and runs until Day+2. These meteorological forecasts were carried out with a spectral resolution, meaning an effective resolution of 9 Km, and 25 sigma levels. Processed meteorological variables (horizontal wind, temperature, specific humidity, surface pressure and temperature, heat fluxes, friction velocity and liquid water) are then provided to the chemistry-transport model, as well as the emissions. The model then produces the ozone

forecasts, for 4 different lead times: Day+0 (forecast for the day), Day+1, Day+2, and Day-1. This latter lead time is based on "first guess meteorological fields" until 12 UT for the first simulation day and on forecasts thereafter itself. The photochemical model is initialised at 00 UT on Day-1, using the previous 24 hour forecast, without any use of observations. The model thus performs a 4-day integration. Once all these calculations are achieved, the meteorological and air quality outputs are delivered on a web server (http://euler.polytechnique.fr/portugal), in the form of graphics.

The CTM system application

The model system was applied using a simple one-way nesting technique. A first continental-scale run is performed with CHIMERE over a regional area from 10.5W to 22.5E and from 35N to 57.5N with a 50 km grid resolution (Figure 1a), followed by a nested simulation over the Portuguese domain of 290×580 km, with a 10 km grid resolution (Figure 1b).

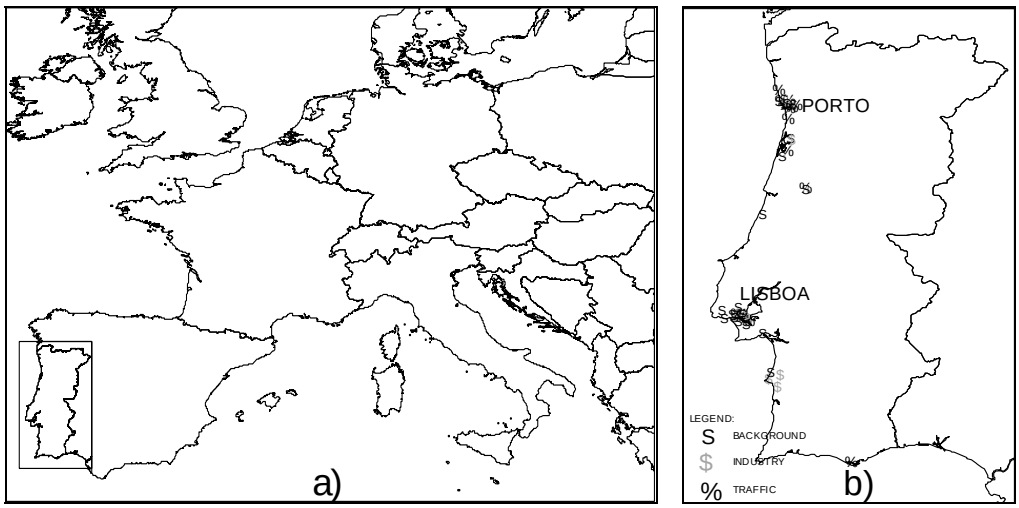

Figure 1. Geographical domains used by the CHIMERE model. Dots in Figure b) represent the locations of the Portuguese air quality network stations with ozone monitors.

The CHIMERE model uses the same physics approach for the two domains simulations. Six vertical layers above topography are considered, the first one (the *surface layer)* having a depth of 50 m. The model top lies at 700 hPa. The continental run is forced at the lateral and top boundaries by monthly climatology of species concentration issued by the MOZART second-generation model (*Horowitz et al.,* 2003). The lateral boundaries of the small-scale domain (Portugal) are taken from the large-scale run species concentrations (including ozone and its precursors) and the top boundaries as the MOZART monthly climatological values. The model version used here is primarily described in *Schmidt et al*. (2001) and further updates, especially for the smaller-scale version, can be found in *Vautard et al.* (2003). At European scale, emissions were derived from the annual totals of the EMEP database for 1999, in a methodology similar to that described in Schmidt et al. (2001). Over the Portuguese domain, area-sources annual emission data are obtained from the Portuguese EMEP database for the year 2001, by different pollutant source activity. The emissions were then spatially downscaled to the sub-municipality level for each activity sector. Large point sources annual emissions were obtained directly from the each industrial plant database. Emissions time disaggregation is obtained by application of monthly, weekly and hourly profiles from the GENEMIS project (1994). The NMVOCs are disaggregated into 227 individual VOCs according to the U.K. speciation (Passant, 2002). The methodology for biogenic emissions of isoprene and terpenes is described in *Schmidt et al.* (2001).

OZONE FORECAST ERROR STATISTCS

The main aim of this work was to estimate the model skill in forecasting, at various lead times, the ozone concentrations at the study area. The Portuguese air quality network includes 36 ozone monitoring stations, with different typology (background, traffic, etc.) according to location and environmental criteria. For validation purposes, the maximal forecast and the maximal observed peaks were compared, as well as, the hourly and octo-hourly averages, for each type of monitoring station. The exceedances of the information and alert thresholds values defined by the Ozone European Directive 2002/3/CE were also analysed. Exceptionally long-lasting and spatially extensive episodes of high ozone concentrations occurred during this 2003 summer, mainly in the first half of August, in all Europe (*EEA*, 2003). In Portugal, some of these episodes were related to the strong forest fire episodes, recorded in several parts of the country, and related to the heat wave verified in this summer in all Europe. In order to evaluate this influence on the ozone concentrations, the model performance was also validated excluding this specific days.

Daily peak forecast

Figure 2 shows some error statistics, as a function of lead time, for the two situations: with and without considering the forest fire days. As expected, model performance improves when fire days are excluded from the validation exercise, since the chemistry transport model did not consider specific forest fire emissions. For background stations, the forecasts of the daily peak concentrations correlate fairly well with observations at Day-1, and correlation decreases with lead time, indicating a clear influence of the meteorological forecast errors. This is confirmed by the linear increase of the root mean square (RMS) error from Day-1 to Day+2. The forecast for industrial stations exhibits higher errors and inferior correlation, showing that probably emissions/representativeness errors are higher than meteorological effects. At the background stations, forecasts are slightly underestimated on average, with a bias in the range 0–3 $\mu g.m^{-3}$, by contrast with the traffic and industrial stations where the ozone peaks are overestimated.

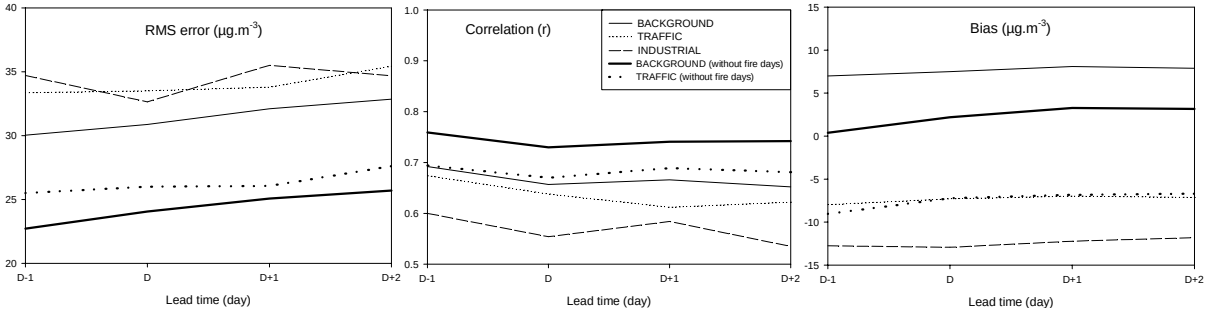

Figure 2. Model average skill for the forecasts of the ozone daily peaks for background, traffic and industrial monitoring stations, as a function of lead time: the RMS error, the correlation factor, and the mean systematic error (bias).

1-h and 8-h averages forecast

Figure 3 shows the validation analysis for 1-h and 8-h averages forecast using the same statistical indices as in Figure 2, for background and traffic stations and excluding the episode fire days. Errors are significantly lower (and the correlation higher) when considering the 8-h moving averages. In general, RMS errors lie above 22 $\mu g.m^{-3}$ for 8-h average data analysis and reach 26 $\mu g.m^{-3}$ for hourly data. The correlation is superior for background stations, although the errors (RMS and bias) are similar to traffic sites. Quite surprisingly there is no

obvious trend for model performance decreasing with the lead time between the day D and D+2.

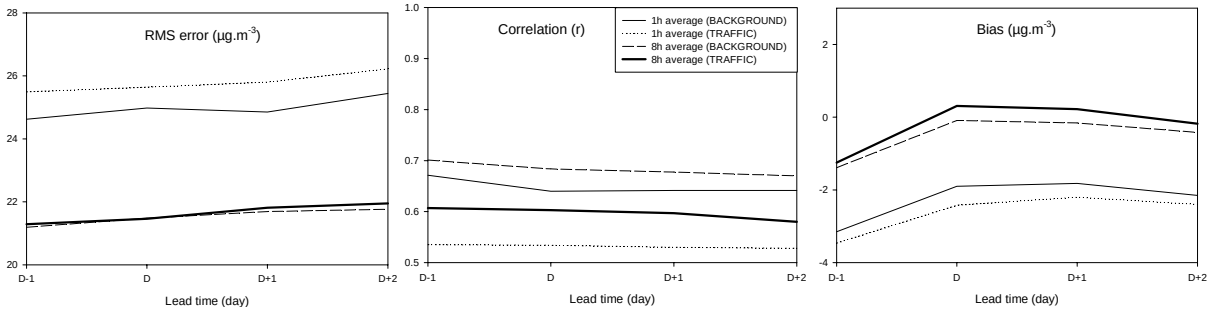

Figure 3. Model average skill for the forecasts of 1h and 8h-averages ozone concentration for background and traffic monitoring stations, as a function of lead time: the RMS error, the correlation factor, and also the mean systematic error (bias).

Thresholds exceedance forecast

An important feature of a forecast model is its ability to correctly predict peak concentrations above 180 µg.m^{-3} and 240 µg.m^{-3} since these are the thresholds values for which an information and an alert, respectively, of bad air quality is sent to the public and for which emission reductions are recommended. Table 1 presents the skill parameters defined by *van Aalst et al*. (1997), considering only the model performance for background monitoring stations and excluding the days with forest fires.

Table 1. Number of observed/predicted exceedance of the information (180 µg.m^{-3}) and alert (240 µg.m^{-3}) threshold levels, and of the target value for health protection (120 µg.m^{-3})

	180 µg.m^{-3} (1h average)				240 µg.m^{-3} (1h average)				120 µg.m^{-3} (8h average)			
	D-1	D	D+1	D+2	D-1	D	D+1	D+2	D-1	D	D+1	D+2
Successful forecast exceedances	23	18	13	10	0	0	0	0	596	565	452	421
False alarms	9	21	18	25	0	0	1	1	102	134	153	168
Non predicted events	14	19	24	27	1	1	1	1	268	299	412	443

This Table shows that the score parameters strongly vary from D-1 to D+2. The first two information days are more correctly forecast, but with several "false alarms". As expected, the 8-h average target value is more correctly forecasted (50-70%) then the information threshold (30-60%). The model was not able to forecast the only alert threshold exceedance. Local emission effects can possibly explain this high value, which was measured in a background station with strong industrial influence.

SUMMARY AND CONCLUSION

Results of deterministic forecasts of ozone pollution for the mainland Portugal were presented. The system is based on a chemistry-transport model (CHIMERE), directly forced by MM5 meteorological model forecasts. At the boundaries, the model is forced by a large-scale, less resolved version of the same model, using a "one-way nesting procedure". The system provides forecasts at lead times up to +2 days. The validation procedure consists in comparing the forecasts with observations during the 2003 summer season, over a total of 36 monitoring sites. The model has some skill in forecasting 180 ug.m^{-3} exceedances as well as 1-h and 8-h averages. Skill scores clearly improve when the forest fire days are excluded of the statistical analysis, indicating that unaccounted sources due to these fires may be responsible for unskilful forecasts. The skill of the model slightly decreases with the lead time, with correlations ranging between 0.7 and 0.8 for short lead times, and between 0.65 and 0.75 for longer lead times. The occurrences of exceedances of the 180 µg.m^{-3} threshold

are forecast in about 40% of the cases, with a rate of false alarms of about 40%, but this skill increases significantly if we consider the exceedances of the 8-h average target value. This demonstrates that the model properly represents the important physical, meteorological and chemical processes, which makes this tool enough reliable for helping operational forecasters. The best forecasts were achieved for background sites located in rural and suburban areas, which are more representative of a coarse grid mesh. The worst forecasts are for industrial sites that are probably affected by highly variable (industries) emissions. The results of this validation exercise calls for improving the forecasting system in several aspects such as the model grid resolution, industrial area emissions estimation and background ozone prediction.

ACKNOWLEDGEMENTS

The authors wish to thank to the Portuguese Institute of Environment for giving access to observational data, to the financial support of the 3rd EU Framework Program and the Portuguese Ministério da Ciência e do Ensino Superior, for the PhD grant of A. Monteiro (SFRH/BD/10922/2002).

REFERENCES

Van Aalst, R.M., de Leeuw, FAAM 1997. European Topic Centre on Air Quality (RIVM, NILU, NOA, DNMI) National Ozone Forecasting Systems and International Data Exchange in Northwest Europe, Report of the Technical Working Group on Data Exchange and Forecasting for Ozone Episodes in Northwest Europe (TWG-DFO).

EEA (European Environment Agency), 2003. Air pollution by ozone in Europe in summer 2003. Report to the European Commission based on data provided in the framework of Council Directive 92/72/EEC on air pollution by ozone.

Gardner, M.W. and Dorling, S.R., 1998. Artificial neural networks – a review of applications in the atmospheric sciences. *Atmospheric Environement* **32** 14–15, pp. 2627–2636.

GENEMIS (Generation of European Emission Data for Episodes) Project, 1994. EUROTRAC Annual Report, 1993, Part 5. EUROTRAC International Scientific Secretariat, Garmisch-Partenkirchen.

Horowitz, L. W., Walters, S., Mauzerall, D., Emmons, L., Rasch, P., Granier, C., Tie, X., Lamarque, J., Schultz, M. and Brasseur, G., 2003. A global simulation of tropospheric ozone and related tracers: Description and evaluation of MOZART, version 2, *J. Geophys. Res.,(in press).*

Passant, N., R., 2002. Speciation of U.K. emissions of non-methane VOC, AEAT/ENV/0545 Issue 1.

Schmidt, H., Derognat, C., Vautard, R., and Beekmann, M., 2001. A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in Western Europe. *Atmospheric Environment* **35**, 2449-2461.

Simpson, R.W. and Layton, A.P., 1983. Forecasting peak ozone levels. *Atmospheric Environment* **17**, pp. 1649–1654.

Simpson, D., 1992. Long period modelling of photochemical oxidants in Europe Calculations for July 1985. *Atmospheric Environement* **26**, pp. 1609–1634.

Vautard, R., Beekmann, M., Roux, J. and Gombert, D., 2000. Validation of a deterministic forecasting system for the ozone concentrations over the Paris area. *Atmospheric Environment* **35**, 2449-2461.

Vautard, R., Martin, D., Beekmann, M., Dropbinski, P., Friedrich, R., Jaubertie, A., Kley, D., Lattuati, M., Moral, P., Neininger, B., Theloke, J., 2003. Paris emission inventory diagnostics from the ESQUIF airborne measurements and a chemistry-transport model, *J. Geophys. Res.*, in press.