1.12 A METHODOLOGY FOR SEASONAL PHOTOCHEMICAL MODEL SIMULATION ASSESSMENT

Veronica Gabusi and Marialuisa Volta
D.E.A. - University of Brescia. Brescia, Italy

INTRODUCTION

Atmospheric dispersion models are widely used to assess the impact and to develop emission claiming strategies; however, no standard performance procedure exists to assess the reliability and the accuracy of these models (*Russell, A. and R. Dennis*, 2000). In order to perform a carefully analysis of the model results and to evaluate in a critical and constructive perspective the model skill in simulating the pollutants dynamics, it is necessary to develop a rigorous model evaluation methodology, as pointed out by several authors (e.g. *Hogrefe, C. et al.*, 2001, *Kasibhatla, P. et al.*, 2001).

This work is addressed to propose a methodology for seasonal photochemical modelling simulation assessment. That tries to solve two characteristic aspects of the proposed problem: (1) the comparison between the measured and the computed values is intrinsically distorted by the space scale parameter: in fact, the measure at a specific point is compared with a computed value, representing a wider area; (2) for modelling application dealing with long-term period each run of the models results in a large amount of output data, with hourly concentrations of pollutants for each grid point over this extensive period. Thus, it can be helpful to focus the analysis for a restricted number of typical concentration patterns and to summarize the simulation results over long time period by means of proper performance indexes.

THE MODEL EVALUATION METHODOLOGY

The presented methodology involves three main steps:

(1) a cluster analysis application to the observed data, in order to detect similarities in terms of concentration levels and temporal variability (as the frequency and peak distribution and the daily shape);
(2) the identification of a restricted set of monitoring sites to be used during the comparison with the computed results;
(3) the evaluation of performance indicators, in term of statistical or graphical indexes.

Clustering process

Cluster analysis technique is used for classifying patterns (*Lavecchia, C. et al.*, 1996, *Ludwig, F.L. et al.*, 1995). In this paper the variables considered are the concentration data, registered in different monitoring sites, geographical spread in the investigated area. As for the methodology (1) the set of concentration measured data for a specific time period is considered to be a point in a space of many dimensions (this space has many dimensions as there are variables under studies); (2) points classified “close together”, referring to a specific distance, are grouped into the same category; the distance measure can highlights similarities in a quantitative point of view, as the Euclidean distance, or focalises similarities among temporal trend phases (*Lavecchia, C. et al.*, 1996).

Station selection

Each cluster can be represented by a single station, selected as representative of the group, or it is possible to define a *virtual station* averaging, hour by hour, the measurements recorded in the stations belonging to the group. In order to identify the representative station, a new cluster process is performed, taking into account only the concentrations of the station in the set, and the station more closed to cluster centre is selected as representative monitoring site. Such methodology allows to go beyond the inadequate observed-computed comparison: in fact, the representative station reflects pollution peculiarities of a monitoring station typology, that can be used to characterized a portion of the domain, allowing to perform a more appropriate model results analysis.

Model performance assessment

Model performance evaluation takes into account both the US EPA recommendations for an acceptable model performance and the new European Directive; moreover, an extensive literature analysis has been performed in order to extrapolate statistical and graphical indicators, as well as methodological approach.

The evaluation methodology, processing measurements and simulated concentrations, provides: (a) the computation of US-EPA indicators, as the Mean Normalised Bias Error (MNBE) and the Mean Normalised Gross Error (MNGE), and “Directive 2002/3/EC” statistical indexes; (b) other widely used indicators as the correlation coefficient between simulated and measured concentrations (R), the Root Mean Square Error (RMSE), the mean concentrations for the simulated period; (c) AOT40, AOT60 ozone exposure indexes; (d) the 25th, 50th, 75th and 95th percentile values for AOTx, 1-h and 8-h daily maximum concentrations.

SEASONAL APPLICATION CASE

The proposed methodology has been applied to analyse seasonal simulations performed by GAMES modelling system (*Volta and Finzi*, 2003) over Northern Italy domains. The preliminary application (*Gabusi et al.*, 2003), performed as contribution to EUROTRAC2-SATURN project, refers to 1996 summer season.

In this paper, the results obtained in the framework of CityDelta Modelling Exercise (*http://rea.ei.jrc.it/netshare/thunis/citydelta/)* are presented and discussed. The selected simulation domain (300×300 km^2) includes the whole Lombardia Region, a complex terrain area located in the Po Valley and one of the most industrialized and populated area in Northern Italy. Air quality simulations, for the 6-months period April-September 1999, have been carried out by means of GAMES system, including the 3D meteorological processor CALMET *(Scire et al.*, 1990) and the photochemical transport model CALGRID (*Yamartino, R.J. et al.*, 1992)

Ozone pattern classification

For this modelling exercise, the considered air pollution data sets are the monitored concentrations at 16 stations, provided by Regione Lombardia and Regione Emilia Romagna air quality network. As described above, a clustering procedure is applied to ozone observed values, and the similarity among monitoring sites is assessed by means of the squared Euclidean distance.

In Figure 1 is reported the obtained cluster tree for the analysed monitoring sites and the geographical distribution inside the domain of the grouped stations. As can be easily noted, the clustering process identified station groups with homogenous spatial dislocation. The first set (Cossato, Varese and Chiavenna) is composed of monitoring stations located in the northern part of the domain, near the Alps; the second one consists of four stations (Crema, Piacenza, Parma and Reggio Emilia), located in the Po Valley and mainly classified as rural sites. Moreover, the performed statistical analysis pointed out three different station clusters inside the Milan metropolitan area, mainly recognized as sub-urban stations:

(a) the four stations (Agrate, Vimercate, Limito e Meda) constituting the first metropolitan cluster are situated in the northern part of this urbanized area;
(b) the second one, composed by only two stations (Arconate e Motta Visconti), is the south-west region;
(c) the third one, again composed of only two stations (Castellanza e Magenta) is dislocated along the north-south axis.

Finally, the sixth cluster is represented by only one site (Torino), located near the western boundary of the considered domain.

Besides, the mean day featured for the virtual station, representing the respective belonging cluster is shown (Figure 1). Analysing the patterns, it can be observed that the Cluster 1, composed of rural stations, exhibits high ozone value, also during night-time, phenomenon due to the advection of polluted air towards the rural area. Highest values are observed in the Cluster 4, constituted by suburban stations in the western metropolitan urban area, where moderated NOx concentrations are not be able to reduce the produced ozone.

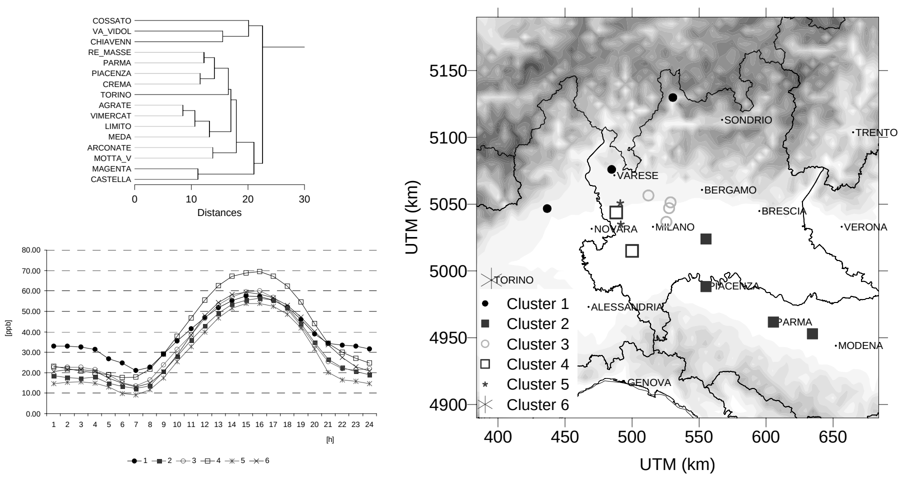

Figure 9. The cluster tree (upper left), the mean day features of the identified cluster (bottom left) and the monitoring sites with their cluster grouping (right).

To gain further insight into the virtual average station definition, in Figure 2, for three different clusters (urban, suburban and rural), ozone mean day pattern graphs for the representative and the average station, both for the observed and for the calculated values, are presented. As we can see, for all the clusters, the difference between the virtual and the representative station are very negligible. This fact can be interpreted as a robustness of the proposed methodology: the clusters are well identified, properly representing the chemical and meteorological condition characterizing homogenous domain area.

Performance assessment

1-hourly and 8-hourly daily maximum ozone measured and simulated concentrations have been processed to assess the model performances (Table 1), as suggested in (*Schmidt, H. et al.*, 2001, *US EPA*, 1991). The indexes have been computed for the virtual. They provide a immediately information about the model behaviour, in terms of a mean overestimation or underestimation (MNBE) and overall performance (MNGE).

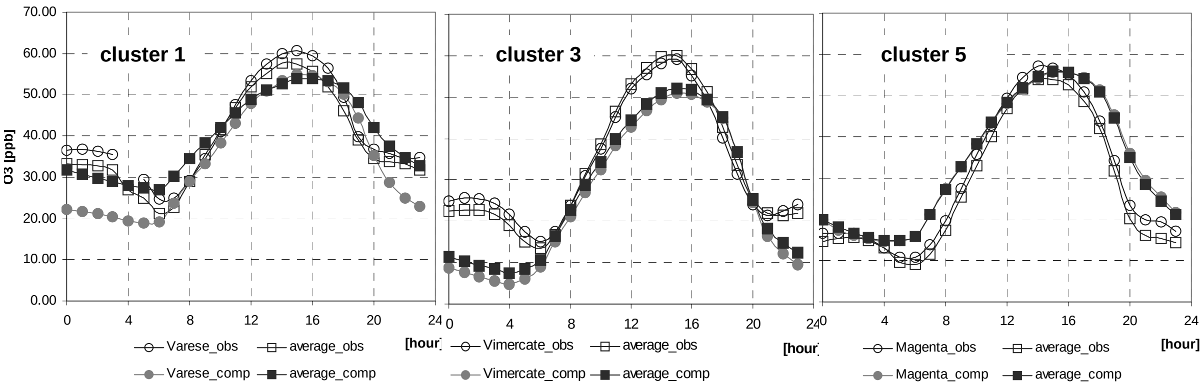

Figure 2. Mean ozone pattern for the representative and the virtual station.

Estimated 25th, 50th, 75th, and 95th percentile values for the 1-h and 8-h ozone peak at each measurement site are shown in Figure 3; again, it should be observed that the information obtained analysing each single monitoring site constituting the cluster can be appropriately summarized by the virtual station.

Table 1. Statistics for daily maximum O3 concentrations; the confidence level criteria suggested by EPA are respectively ±5÷15% for the MNBE and 30÷35% for the MNGE.

	Max 1h				Max 8h			
Cluster	**MNBE**	**MNGE**	**R**	**RMSE**	**MNBE**	**MNGE**	**R**	**RMSE**
1	-7	22	0.37	17.52	-2	11	0.41	14.05
2	13	26	0.60	14.48	28	13	0.61	15.34
3	-10	25	0.63	20.11	-3	15	0.64	16.47
4	-18	23	0.62	22	-14	12	0.65	17.66
5	15	33	0.59	17.24	19	16	0.69	14.15
6	-46	50	0.23	69.62	-44	23	0.43	58.67

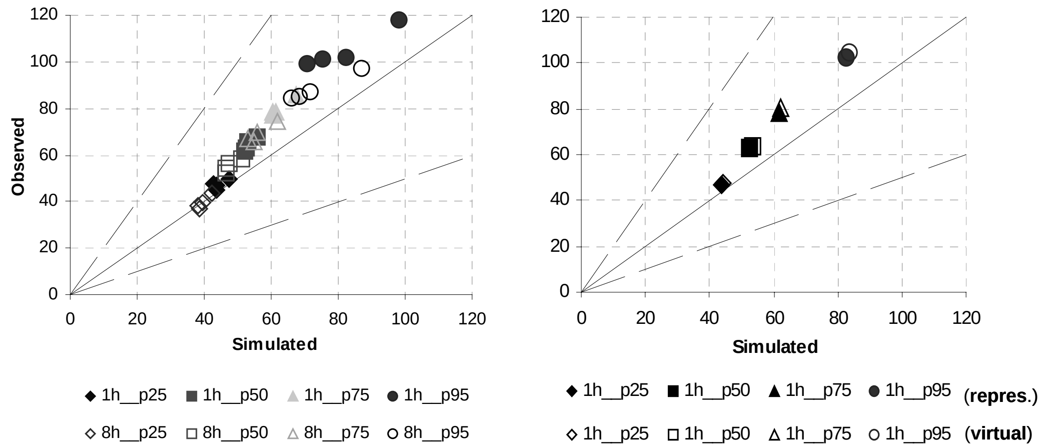

Figure 3. Observed versus modelled ozone peak concentrations at various percentiles: for all the sites composing the set (left), both for 1-h and 8-h O3 daily maximum, and for the virtual and the representative station (right).

CONCLUSIONS

In this paper a methodological approach to evaluate seasonal photochemical simulation has been presented. This methodology has been applied to analyse seasonal simulations performed by GAMES modelling system over Northern Italy domain in the framework of EUROTRAC2-SATURN and CityDelta Projects. The main issues of such approach can be summarized as follows: (a) the model evaluation is performed by using statistical and graphical methods; (b) these indicators can be obtained for each measurement station or for a limited number of representative stations of the domain, selected by means of a clustering process; (c) the evaluation of representative stations should be preferred with long-term simulation assessments.

ACKNOWLEDGEMENTS

The authors are grateful to Guido Pirovano (CESI), Marco Bedogni (AMA), and Enrico Minguzzi (ARPA Emila Romagna) for their kind and valuable cooperation in the framework of CityDelta Project. This work has been supported by the MIUR (Italian Ministry of the University and Research) and University of Brescia (Fondo Giovani Ricercatori).

REFERENCES

Gabusi, V., C. Pertot, and G. Finzi, 2003: Performance assessment of a long-term photochemical modelling system. Int. J. Environment and Pollution, Vol. 20, 64-74.

Hogrefe, C., S.T. Rao, P. Kasibhatla, W. Hao, G. Sistla, R. Mathur and J. McHenry, 2001: Evaluating the performances of regional-scale photochemical modeling systems: Part II – Ozone predictions. *Atmospheric Environment*, 35, 4175-2188.

Kasibhatla, P. and Chameides, W.L., 2000: Seasonal modeling of regional ozone pollution in the Eastern United States, Geophysical Research Letters, Vol. 27, pp. 1415-1418.

Lavecchia, C., E. Angelino, M. Bedogni, E. Brevetti, R. Gualdi, G. Lanzani, A. Musitelli and M. Valentini, 1996: The ozone patterns in the aerological basin of Milan (Italy), Environmental Software, 11, 73-80.

Ludwig, F.L., J. Jiang and J. Chen, 1995: Classification of ozone and wheather patterns associated with high ozone concentrations in the San Francisco and Monterey Bay areas, Atmospheric Environment, 29, 2915–2928.

Russell, A. and R. Dennis, 2000: NARSTO critical review of photochemical models and modeling. Atmospheric Environment, 34, 2283–2324.

Schmidt, H., C. Derognat, R. Vautard and M. Beekmann, 2001: A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in Western Europe. Atmospheric Environment, 35, 6277-6297.

Scire, J.S., E.M. Insley and R.J. Yamartino, 1990: Model formulation and user's guide for the CALMET meteorological model. Report A025-1, California Air Resources Board.

Sistla, G., W. Hao, J. Ku, G. Kallos, K. Zhang, H. Mao, and S.T. Rao, 2001: An operational evaluation of two regional-scale ozone air quality modeling systems over eastern United States. Bulletin of American Meteorogical Society.

Volta, M. and G. Finzi, 2003: GAMES, a new comprehensive gas aerosol modeling system. In: Fourth international conference on Urban Air Quality -Measurement, Modelling & Management, 14–17.

US Environmental Protection Agency, 1991: Guideline for regulatory application of the Urban Airshed Model. EPA-450/4-91-013, Research Triangle Park, NC 27711.

Yamartino, R.J., J.S. Scire, G.R. Carmichael and Y.S. Chang, 1992: The CALGRID mesoscale photochemical grid model - I. Model formulation. Atmospheric Environment, 26A, 1493-1512.