H13-86

COMBINATION OF MEASURED AND MODELLING DATA IN AIR QUALITY ASSESSMENT IN SPAIN

Fernando Martín, Inmaculada Palomino, Marta G. Vivanco

División de Contaminación Atmosférica, Departamento de Medio Ambiente, CIEMAT, Madrid

Abstract: In this presentation, a methodology to combine measurements from air quality stations and estimates from the CHIMERE model for air quality assessment in Spain is described. The methodology consists of using linear regression and kriging interpolation to correct the model results improving the fit to the observations. It was separately applied to rural and urban conditions, yielding to maps for each case, which were then combined by taking into account the distribution of rural and urban areas in the domain. The results for several pollutants and its application to air quality assessment in Spain are shown and discussed.

Key words: *Air quality assessment, air pollution modelling, kriging, linear regression, measurements and model combination.*

INTRODUCTION

The European and Spanish laws oblige to the Governments to provide reliable information about the air quality in Spain every year regarding concentration levels and exceedances of air quality standards. The use of just air quality measurements can provide an incomplete picture of the air quality, as monitoring sites can not cover all the territory. Thus, the use of complementary techniques, such as modelling, is allowed and recommended in many cases. The combination of air quality measurements at stations and validated model estimates is a good choice, due to the accuracy of measurements and the good spatial cover of models.

In this presentation, a methodology to combine measurements from air quality stations and estimates from the CHIMERE model for air quality assessment in Spain is described. The methodology consists of using linear regression and kriging interpolation to correct the model results improving the fit to the observations. It was separately applied to rural and urban conditions, yielding to maps for each case, which were then combined by taking into account the distribution of rural and urban areas in the domain. The results for several pollutants and its application to air quality assessment in Spain are shown and discussed.

METHODOLOGY

Measurements-model combination methodology

Combining model results and observations has been widely applied in meteorological modelling, but fewer studies have been done to combine air quality observations and model outputs. Among others, several works can be pointed out: Tarrasón *et al.* (1998) for measurements and EMEP model data combination, Wiegand and Diegmann, (2000) to develop the German system "FLADIS", Denby *et al.* (2005) reviewing different methodologies to combine and assimilate observations and models, Denby *et al.* (2008) discussing the uncertainty sources in air quality mapping and Fiala (2009) applying a methodology for ozone and PM10 assessment in Europe.

In Spain, the first studies are from Martín *et al.* (2005), who proposed to use a methodology based on the assimilation techniques (Benjamin and Seaman, 1985) used in meteorological models. An influence area was defined for each observation depending on the station type, the distance between the grid point and the measurement point and the wind flow. It was used in some annual air quality assessments in Spain, but it shows some shortcomings yielding less realistic air quality maps.

The methodology used in this study is based on the idea that the real concentration of an atmospheric pollutant *C* in a station *k* can be expressed as

$$C_k = M_k + e_k + s_k \tag{1}$$

where M_k is a concentration estimate (i.e., by a dispersion model), e_k is the systematic error of the estimate (i.e., modelling error) and s_k is the inherent error or measurement error. The question is how to reduce the model error e_k, that is, how to correct the model results to provide a best fit to observations and to get a more realistic map of the spatial distribution of pollutant concentrations. Among the several options, the linear regression and the kriging interpolation methods are the most interesting (Fiala, 2009).

The linear regression technique assumes that a better estimate of the concentration $C_k^{'}$ can be obtained by

$$C_k^{'} = aM_k + b + r_k \tag{2}$$

where *a* and *b* are the regression coefficients and r_k is the residual error which includes the measurement error and the non-solved part of the modelling error. This method corrects the concentration estimates by taking into account any influence of the concentration values on them.

The kriging interpolation technique assumes:

$$C_k^{'} = \sum_{i=1}^{n} \lambda_i M_k + r_k \qquad \sum_{i=1}^{n} \lambda_i = 1 \tag{3}$$

being λ_i the weights assumed on the basis of a variogram in order to minimize the mean-square-error, they range between 0 and 1. The variogram is a function representing how a measured variable varies with distance:

$$\gamma(h) = \frac{1}{2}\frac{1}{n}\sum[C(x) - C(x+h)]^2 \qquad (4)$$

where n is the number of stations pairs located to a same distance h between them.

In our study, the variogram can be computed by plotting the values of the concentration differences (or the model residuals) between pairs of stations against the distances between them. The resulted scatter plot can be fitted to simple functions, such as logarithm, exponentials, etc. This method corrects the concentration estimates by taking into account any influence of the distance or spatial representativeness of the air quality stations on the concentration estimates.

In a former study of Martín *et al.* (2009), several possibilities to apply this methodology were analyzed and as a conclusion the authors recommended:

1. to apply the methodology to urban and rural stations separately in order to take into account the different spatial distribution patterns of air pollution concentrations for rural and urban areas obtaining different maps for rural and urban patterns.
2. to use linear regression and kriging in the case of model residuals for rural stations, and only kriging for urban areas.
3. to use spherical variogram for kriging
4. to use population density as surrogate indicator for merging urban and rural air pollution maps.

A summary of the methodology is shown in the figure 1. This methodology is applied to the residuals of the CHIMERE model (observation minus model estimation). More details about the combination methodology can be seen in Martín *et al.* (2009).

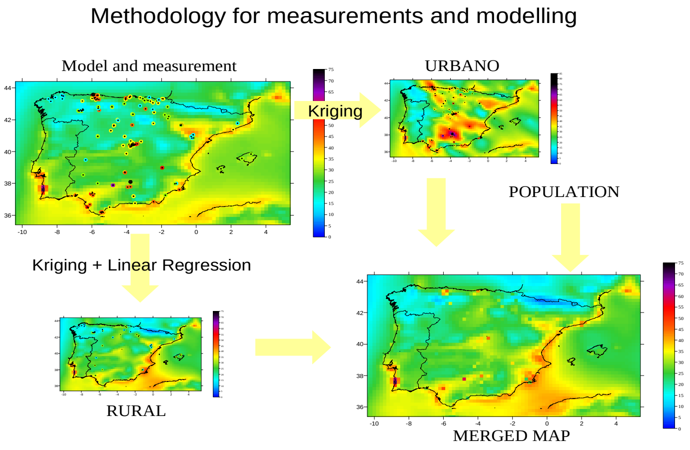

Figure 1. Scheme of the methodology to combine measurements and modelling.

Model setup

Simulations of photochemical compounds were carried out using the CHIMERE chemistry-transport model (Bessagnet *et al.*, 2004; Hodzic *et al.*, 2005), version 2008c. This model is being used for the annual simulations of air quality in Spain since 2004 (Martín *et al.*, 2004, Vivanco *et al.*, 2007). It has been evaluated using measured data of ambient pollutant concentrations from a large number of Spanish stations (Vivanco *et al.*, 2009a and b) and compared with other models such as CMAQ (Baldasano *et al.*, 2008). The model was shown to be suitable for air quality assessment as the uncertainty statistics were lower than the maxima established by the EU directives and the EPA criteria. The impact of the spatial computing resolution was also discussed in Vivanco *et al.* (2008). The MM5 model was the meteorological processor used to feed the CHIMERE model. The models were applied to a European domain and then, to an Iberian Peninsula one. (Figure 2). In Figure 3, the scheme of the model system, boundary conditions, inputs, grid resolution, etc is shown. More details can be found in Vivanco *et al.* (2009b).

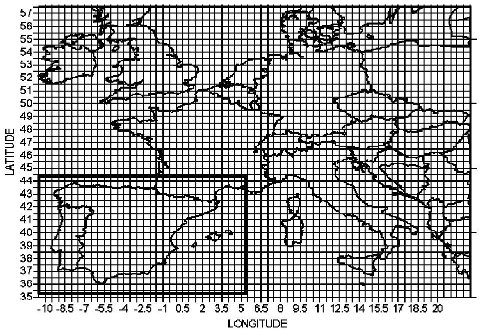

Figure 2. Map showing the computing domains used with the CHIMERE model.

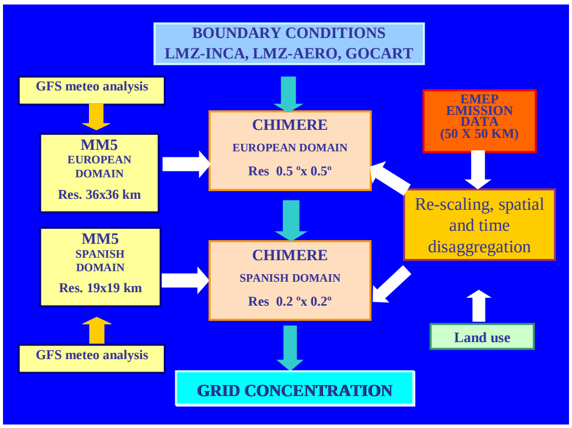

Figure 3. Scheme showing the model setup

RESULTS AND DISCUSSION

The CHIMERE model was run for 2007 in order to provide concentrations of SO_2, O_3, NO_2 and PM10 in the Iberian Peninsula and the Balearic Islands. The described methodology to combine measurements and modelling was applied to the residuals of the CHIMERE model computed for the set of air quality stations used for air quality assessment except to the traffic stations.

Maps of average concentrations and the N^{th} higher value at every grid cell, such as $N=Np+1$, where Np is the number of exceedances allowed by the European directives for each pollutant. In addition, maps showing the uncertainty of the combination methodology were computed based on the uncertainty of the kriging interpolation, which is:

$$\delta_c(x,y)=\sqrt{2\gamma(h)} \qquad (5)$$

This uncertainty estimate was used to compute the probability of having more exceedances of limit or target values than allowed by legislation using the approach of Fiala *et al.* (2009). Maps showing the exceedance probability are the main output of the described methodology.

How does the combination methodology improve the air quality assessment?

As the objective of the described combination methodology is to provide more reliable information about the air quality in a territory, it is needed to check whether the resulted air quality maps are better than those using only the CHIMERE model. The Relative Directive Error (RDE) as defined and used in Denby *et al.* (2010) and computing the Maximum Relative Directive Error (MRDE) for the entire domain as the maximum of the RDE values found at 90% of the available stations. As shown in Table 1, in all the cases the results of the methodology are much better than the model results complying the legal requirements of allowed uncertainty for model techniques used in air quality assessment.

Table 1. Maximum Relative Directive Error (MRDE) for the entire domain for all the limit and target values obtained with CHIMERE model results and with the combination methodology for modelling and measurements for SO2, O3, NO2 and PM10 in the Iberian Peninsula and the Balearic Islands.

Reference value	MRDE Combination methodology	MRDE CHIMERE Model	Pollutant
Target value 120 µg m^{-3} (eight-hour average)	0.1196	0.1570	O_3
Information value 180 µg m^{-3} (hourly average)	0.2056	0.2510	
Alert value 240 µg m^{-3} (hourly average)	0.1542	0.2064	
Limit value 200µg m^{-3} (hourly average)	0.2315	0.3268	NO_2
Limit value 40 µg m^{-3} (annual average)	0.0549	0.3272	
Limit value 350 µg m^{-3} (hourly average)	0.3288	0.5282	SO_2
Limit value 125 µg m^{-3} (daily average)	0.0804	0.2394	
Limit value 50 µg m^{-3} (daily average)	0.2311	0.6217	PM10
Limit value 40 µg m^{-3} (annual average)	0.1045	0.5224	

Maps for air quality assessment

Maps of air pollutant concentrations and probability of having more exceedances than the legally allowed are shown in figures 4, 5 and 6 for O_3, PM10 and NO_2, respectively for 2007.

For ozone, main problems are in the Mediterranean coast, western Andalucia (Guadalquivir valley), some areas in the Cantabric Coast and close to large urban areas (Madrid and Barcelona). In the case of PM10, the risk of exceedances is high in all the Mediterranean Coast, Guadalquivir and Ebro Valleys, Madrid and Asturias (in the north of the Iberian Peninsula). Respect to NO_2, the areas of high probability of exceedances are in large urban areas such as Madrid and Barcelona.

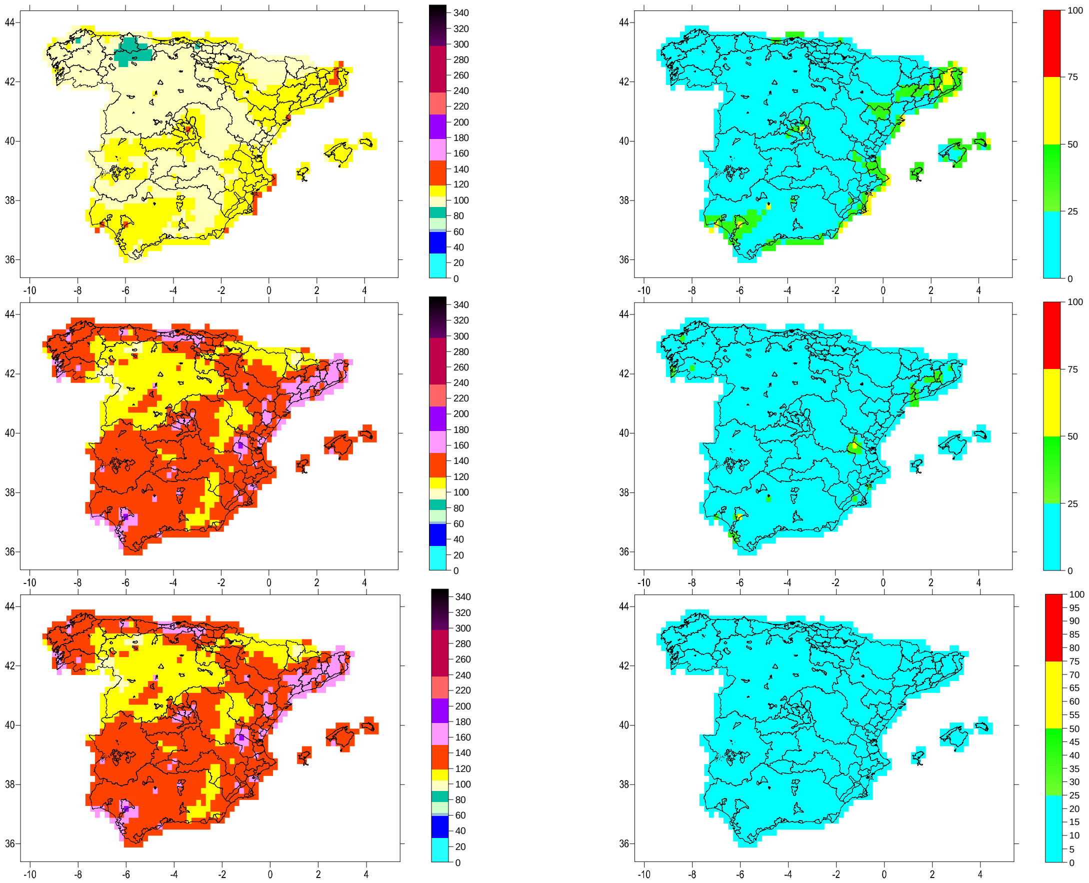

Figure 4. Maps of O_3 concentrations (left) and probability of having more exceedances than the legally allowed (right) for target value (above), information value (middle) and alert value (below).

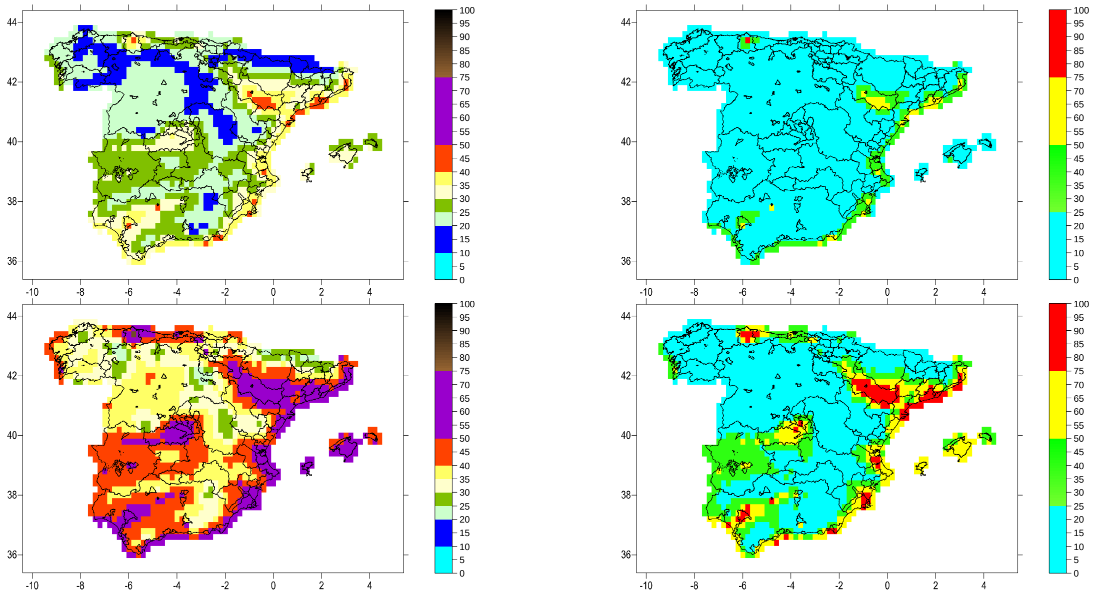

Figure 5. Maps of PM10 concentrations (left) and probability of having more exceedances than the legally allowed (right) for the annual limit value (above) and the daily limit value (below).

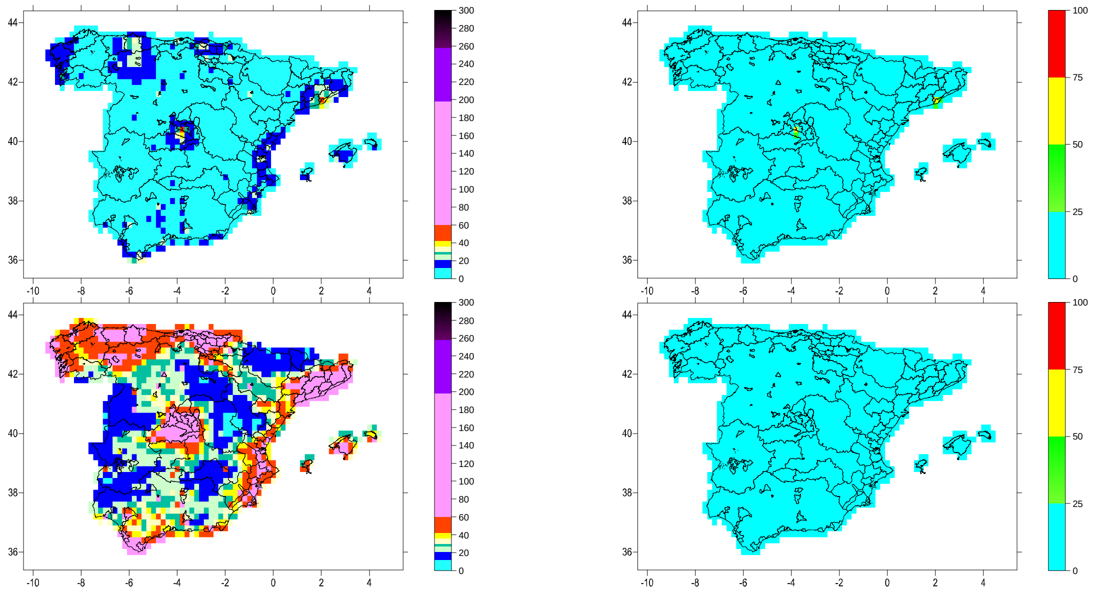

Figure 6. Maps of NO_2 concentrations (left) and probability of having more exceedances than the legally allowed (right) for the annual limit value (above) and the hourly limit value (below).

ACKNOWLEDGEMENTS

This work is being financially supported by the Spanish Ministry of Environment, Marine and Rural Affairs under a collaboration agreement for the use of atmospheric modelling for air quality assessment in Spain.

REFERENCES

Baldasano J. M., *et al.*, 2008: Caliope: an operational air quality forecasting system for the Iberian Peninsula, Balearic Islands and Canary Islands – first annual evaluation and ongoing Developments. *Adv. Sci. Res*., 2, 89–98

Benjamin, S. G. and N. L. Seaman, 1985: A simple scheme for objective analysis in curved flow. *Mon. Wea. Rev.*, 113, 1184-1198.

Bessagnet, B., A. et al, 2004: Aerosol modeling with CHIMERE –first evaluation at continental scale. *Atmos. Environ*.38, 2803-2817.

Denby, B. *et al.*, 2005: Interpolation and Assimilation Methods for European Scale Air Quality Assessment and Mapping Part I: Review and Recommendations, ETC/ACC Technical Paper 2005/7.

Denby *et al.*, 2009: Sources of uncertainty and their assessment in spatial mapping. ETC/AAC Technical Paper 2008/20.

Denby *et al.*, 2010: Guidance on the use of models for the European Air Quality Directive. A working document of the Forum for Air Quality Modelling in Europe FAIRMODE. ETC/ACC report Draft.Version 5.1.

Fiala J., 2009: Spatial assessment of PM10 and ozone concentrations in Europe (2005). EEA Technical report No 1/2009. ISSN 1725-2237. 52 p.

Hodzic, A. et al, 2005: On the quality of long-term urban particulate matter simulation with the CHIMERE model. *Atmos. Environ*, 39, 5851-5864.

Martin F, I. Palomino and M. Garcia, 2004: Aplicación de un modelo de dispersión para la evaluación de la calidad del aire en España. Año 2003 y actividades del 2º Semestre de 2004. Informe para la Dirección General de Calidad y Evaluación Ambiental. Ministerio de Medio Ambiente.19 Noviembre 2004. Ref: 05/2004

Martin F, I. Palomino and M. Garcia, 2005: Aplicación de un modelo de dispersión para la evaluación de la calidad del aire en España. Año 2004 y actividades del 2º Semestre de 2005. Informe para la Dirección General de Calidad y Evaluación Ambiental. Ministerio de Medio Ambiente. Ref: 07/2005

Martín F. I. Palomino and M. G. Vivanco, 2009: Application of a method for combining measured data and modelling results in air quality assessment in Spain. Física de la Tierra. Recent Advances in Meteorology and Climatology, Vol. 21, 65-78

Tarrasón L. *et al.*, 1998: Geographical distribution of sulphur and nitrogen compounds in Europe derived both from modelled and observed concentrations. The Norwegian Meteorological Institute, Oslo, Norway, EMEP/MSC-W Note 4/98).

Vivanco M.G., I. Palomino and F. Martín, 2007: Air quality assessment in Spain for 2002-2004 using modelling techniques. Proceedings of the 6th International Conference on Urban Air Quality, Cyprus, 27-29 March 2007.

Vivanco M.G. *et al.*, 2008: Influence of model resolution on ozone predictions over Madrid area (Spain). International Conference on Computational Science and Its Applications. Italy.

Vivanco M.G. *et al.*, 2009a: Evaluación de la calidad del aire en España utilizando modelización combinada con mediciones, Re-evaluación 2006. Encomienda de Gestión MMA-CIEMAT para evaluación de la calidad del aire mediante modelización. Unidad de Contaminación Atmosférica. CIEMAT. Madrid.15 Abril 2008, Ref: 17/2009

Vivanco M.G. et al, 2009b: Multi-year assessment of photochemical air quality simulation over Spain. *Environmental Software & Modelling*. 24, 63–73.

Wiegand, G. and Diegmann, V., 2000: FLADIS - A system for extending air pollution point data to continuous spatial information; In: *Air Pollution VIII*, WIT Press, pp. 191-200, ISBN 1-85312-822-8.