FORECASTING OZONE AT EUROPEAN SCALE WITH ENSEMBLES

Vivien Mallet[1,*], Bruno Sportisse[2]

[1] INRIA, CLIME project (common project ENPC - INRIA) [France]

[2] ENPC, laboratory CEREA (joint research laboratory ENPC - EDF R&D), CLIME project [France]

INTRODUCTION

A major limitation of air quality forecasts is the numerous uncertainties in the chemistry-transport models (CTMs) and their input data. The physical formulation of a CTM is uncertain. The numerical discretization introduces further uncertainties in the computed concentrations. The large set of input data to the models (emissions, meteorological fields, ...) shows high uncertainties, often ranging from 30% to 50%, sometimes reaching 100% or more. Considering practical issues, one may mention the uncertainties lying in the computer code (approximations, bugs) and in the use of the code (choices left to the user, mistakes). All these uncertainties result in uncertain output concentrations of observed and unobserved pollutants.

The state (i.e., the concentrations vector) of a CTM is of high dimension: its size is usually 10^6 or 10^7. Meanwhile about a few hundreds of components are observed with monitoring stations or with dedicated instrument during intensive observation periods. Consequently, the models are slightly constrained by the observations and the uncertainties in the forecasts remain, especially for unobserved components.

In order to improve the forecasts, the current developments mainly rely on advances in physics and chemistry. Meanwhile the model performances slightly increase, probably because the uncertainties shadow the modeling efforts and because the models abilities are based on tuned configurations.

As a consequence, the models should be considered as stochastic models. In this context, ensemble approaches, based on a set of simulations instead of a single simulation, are suited to estimate the uncertainties (Section 2) and to improve the forecasts (Section 3). Multimodel ensembles are generated within the framework of the Polyphemus air quality modeling system (Section 1).

1 THE AIR QUALITY MODELING SYSTEM POLYPHEMUS

1.1 Purpose

Polyphemus [Mallet et al., 2007] is a rather new air quality modeling system, developed by the École Nationale des Ponts et Chaussées (ENPC), the French National Institute for Research in Computer Science and Control (INRIA) and Électricité de France (EDF R&D), with support from Institut de Radioprotection et de Sûreté Nucléaire (IRSN) and the French National Institute for Industrial Environment and Risks (INERIS).

It was built to cover the scope and the abilities of modern air quality systems, notably ensemble forecasting and data assimilation. It was designed to share developments inside and outside the system, and to host several models. A strong flexibility at preprocessing stages and simulation stage enables to deal with multiple model configurations. High-level methods that embed one or more CTMs, such as data assimilation, are implemented independently of the models, so that they may be applied to several models available on the platform.

Polyphemus deals with applications at different scales (from local to continental scale) with two Gaussian models and two Eulerian models (Castor and Polair3D). Its target pollutants are currently passive tracers, radionuclides, photochemical species and aerosols.

1.2 Structure

The system is made of four independent levels: data management, physical parameterizations, numerical solvers and high-level methods such as data assimilation, ensemble forecast or model coupling.

At preprocessing stage, data from several sources are managed, and several alternative parameterizations are available to compute the main fields (deposition velocities, vertical diffusion coefficients, ...). At simulation stage, the model is essentially a numerical solver: most physical calculations have been performed in preprocessing steps. Within a model, the numerical schemes for advection, diffusion and chemistry may be changed.

On top of the model(s), a driver implements a method like data assimilation or Monte Carlo simulations. In a driver, the model (or the models) is seen as a black box with a light interface. The drivers are implemented independently of the models so that any driver may be used in combination with any model (that has the light interface required by the driver).

The development of such a system requires advanced technical features. Hence Polyphemus is mostly written in C++.

[*] *Corresponding author:* Vivien.Mallet@inria.fr

1.3 Uncertainty Estimation and Ensemble Forecasts in Polyphemus

In Polyphemus, estimation of uncertainties due to input data may be carried out with drivers: a driver for Monte Carlo simulations (perturbation of input data fields), a driver to perform adjoint simulations (sensitivity analysis). Note that the adjoint of one CTM (Polair3D, gas phase version) is mainly obtained through automatic differentiation.

In order to estimate the uncertainties coming from the physical formulation (physical parameterizations) and the numerical schemes, the flexibility of Polyphemus is used to build multimodel ensembles. Each model of the ensemble is built with a different set of physical parameterizations and numerical schemes. The raw data sources can also be changed from one model to another (e.g., land use data, or meteorological data).

2 Uncertainty Estimation

2.1 Experimental Setup

In this abstract, all base simulations have essentially the same configuration. They are primarily set up to estimate ozone concentrations over Europe in summer 2001. In short, the reference configuration is

1. domain: $[40.25°\mathrm{N}, 10.25°\mathrm{W}] \times [56.75°\mathrm{N}, 22.25°\mathrm{E}]$ with $0.5°$ resolution on the horizontal and five vertical levels;

2. meteorological data: ECMWF[1] fields (resolution of $0.36° \times 0.36°$, TL511 spectral resolution in the horizontal, 60 levels, time step of 3 hours, 12 hours forecast-cycles starting from analyzed fields);

3. chemical mechanism: RACM [Stockwell et al., 1997];

4. emissions: the EMEP[2] inventory, converted according to Middleton et al. [1990];

5. biogenic emissions: computed as proposed in Simpson et al. [1999];

6. deposition velocities: the revised parameterization from Zhang et al. [2003];

7. vertical diffusion: within the boundary layer, the Troen and Mahrt parameterization described in Troen and Mahrt [1986]; above the boundary layer, the Louis parameterization found in Louis [1979];

8. boundary conditions: output of the global chemistry-transport model Mozart 2 [Horowitz et al., 2003];

9. numerical schemes: a first-order operator splitting; a direct space-time third-order advection scheme with a Koren flux limiter; a second-order order Rosenbrock method for diffusion and chemistry [Verwer et al., 2002].

[1] European Centre for Medium-Range Weather Forecasts

[2] Co-operative Programme for Monitoring and Evaluation of the Long-range Transmission of Air Pollutants in Europe

Input data	Uncertainty
Cloud attenuation	$\pm 30\%$
Deposition velocities (O_3 et NO_2)	$\pm 30\%$
Boundary conditions (O_3)	$\pm 20\%$
Anthropogenic emissions	$\pm 50\%$
Biogenic emissions	$\pm 100\%$
Photolysis rates	$\pm 30\%$

Table 1: Assumed uncertainties in input data.

The simulated concentrations are compared to observations from a set of monitoring stations in Europe (mainly in France and Germany) in summer 2001. On ozone peaks, the root mean square error is 22.4 $\mu g\, m^{-3}$ and the correlation is 0.78.

2.2 Uncertainties Due to Input Data

In order to assess the uncertainties due to input data (except base meteorological fields), we carried out Monte Carlo simulations. We used the full configuration introduced in Section 2.1 over 7 days (plus 4 days of spin-up). The perturbations in the input data were devised from several sources – mainly Hanna et al. [1998, 2001]. See Table 1.

800 Monte Carlo simulations were performed. About 200 to 400 simulations were needed for convergence of highly averaged outputs such as the spatio-temporal average of hourly ozone concentrations (Figure 1). To summarize, the uncertainty (relative standard deviation) on mean ozone peaks was about 8%. Details may be found in Mallet [2005].

2.3 Uncertainties Due to the Model Formulation

The numerical model formulation is defined by a set of physical parameterization and a set of numerical schemes. In order to account for uncertainties in the model formulation, a multimodel ensemble was derived from the reference configuration with changes in the parameterizations and in the model discretization (resolution, numerical schemes). For instance, two chemical mechanisms (RACM and RADM [Stockwell et al., 1990]) were used and three vertical diffusion estimates were considered. Changes in the input-data sources were also included (e.g., land use categories). About 20

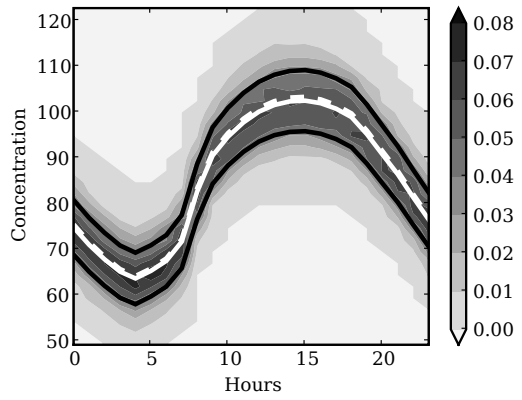

Figure 1: Mean daily profiles for ozone (concentrations in $\mu g\,m^{-3}$). The probability density (shaded) is shown in (b), together with the expectation (continuous white line), the expectation plus ou minus the standard deviation (black lines) and the profile of the reference simulation (discontinuous white line).

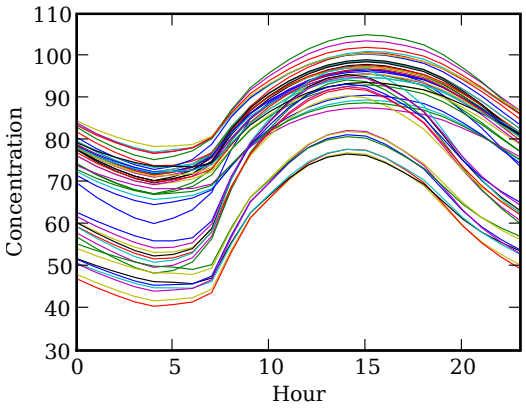

Figure 2: Ozone daily profiles of 48 models built in Polyphemus. The concentrations are in $\mu g\,m^{-3}$ and are averaged over Europe (at ground level) and over four months (mostly summer 2001).

changes were available, which enabled to build an ensemble with 48 members.

The ensemble shows a wide spread – see Figure 2. The estimated uncertainty (relative standard deviation) is over 15%. Details may be found in Mallet and Sportisse [2006b].

3 Ensemble Methods: Combining Models

3.1 Introduction

Ensemble simulations bring more information than a single simulation, which may be used to improve the forecasts. As an illustration of the information brought by the ensemble of 48 members, Figure 3 shows the index of the best model over Europe for

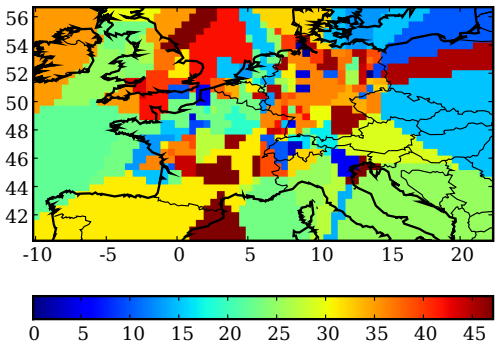

Figure 3: Map of best model indices. In each cell of the domain, the color shows which model (marked with its index, in $[\![0, 47]\!]$) gives the best ozone peak forecast on 7 May 2001 at the closest station to the cell center. It shows that many models can deliver the best forecast at some point.

ozone peaks on 7 May 2001.

In order to overtake the limitations of uncertainties in forecasts, we tried to linearly combine the models of our multimodel ensemble (Section 2.3). At a given forecast date, the weight associated with a member of the ensemble may depend on past concentrations (of all ensemble members) and on past observations. The methods are applied to ozone peaks in the sequel.

The output of model m at time t (or day t as we focus on ozone peaks) and position x (or station x) is denoted $M_{m,t,x}$. The ensemble method computes a linear combination with weights $\alpha_{m,t}$: $E_{t,x} = \sum_m \alpha_{m,t} M_{m,t,x}$. Note that the weights do not depend on the position. This way, it is still possible to compute 2D fields of concentrations as output of the modeling effort – although this point needs further investigation. Moreover, if the weights were dependent of the position, the ensemble method would compete with purely statistical methods which are highly efficient and less computationally expensive. The available observations are $O_{t,x}$. The performance of a combination E is measured by the root mean square error $\mathrm{RMSE}(E,O) = \sqrt{\frac{1}{N_o}\sum_{t,x}(E_{t,x}-O_{t,x})^2}$, where N_o is the total number of observations.

The reference to be improved is the best model EB in the ensemble, which satisfies $EB_{t,x} = M_{\widehat{m},t,x}$ where $\widehat{m}$ minimizes $\mathrm{RMSE}(M_m, O)$. The reference performance for ozone peaks is $\mathrm{RMSE}(EB,O) = 22.4\ \mu g\,m^{-3}$.

The ensemble mean is defined as $EM_{t,x} = \overline{M_{m,t,x}}^m$, that is, with $\alpha_{m,t} = \frac{1}{48}$ since there are 48 models in the ensemble. It shows poor performances $\mathrm{RMSE}(EM,O) = 23.9\ \mu g\,m^{-3}$. Hence methods that take into account past observations

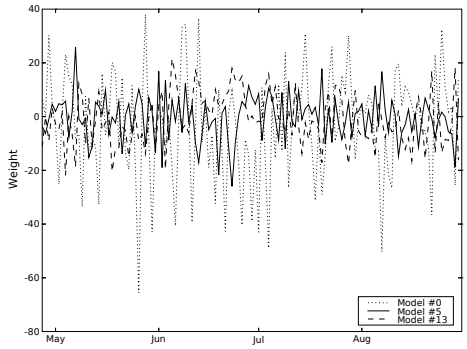

Figure 4: Optimal weights in the least-squares sense (ELS) against time, for three models.

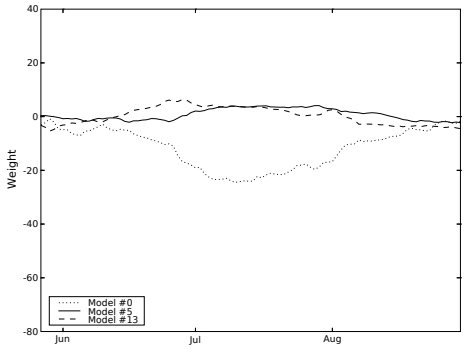

Figure 5: Weights associated with ELS^{30} against time, for three models (same models and same ordinate range as in Figure 4).

are required.

3.2 Least-Squares Methods

Least-squares methods, called "superensembles" in [Krishnamurti et al., 2000], minimize a squared error: $ELS_{t,x} = \sum_m \alpha_{m,t} M_{m,t,x}$ where $\forall t \quad \alpha_{\cdot,t} = \operatorname{argmin} \sum_x \left(O_{t,x} - \sum_m \alpha_{m,t} M_{m,t,x}\right)^2$. In this case, all observations are taken into account (in the past and in the future). Hence ELS shows the potential of least-squares methods: $\mathrm{RMSE}(ELS, O) = 12.0\ \mu\mathrm{g\,m}^{-3}$. Unfortunately, the optimal weights show rapid variations in time, which makes them hard to forecast (Figure 4).

Hence forecasting weights in this framework requires some smoothing. Then the optimal weights are computed over a learning period. This learning period is a moving window of 30 days preceding the day to be forecast. The combination is $ELS^{30}_{t,x} = \sum_m \alpha^{30}_{m,t} M_{m,t,x}$ where $\forall t \quad \alpha^{30}_{\cdot,t} = \operatorname{argmin} \sum_{t-30 \leq T < t,x} \left(O_{T,x} - \sum_m \alpha_{m,T} M_{m,T,x}\right)^2$. This method successfully computes improved forecasts: $\mathrm{RMSE}(ELS^{30}, O) = 20.2\ \mu\mathrm{g\,m}^{-3}$ (against $22.4\ \mu\mathrm{g\,m}^{-3}$ for the best model). This time, the weights are much smoother, which explains the good performances – see Figure 5

For further details, refer to Mallet and Sportisse [2006a].

3.3 Machine Learning Algorithms

Least-squares methods can be efficient, but – as far as we know – there is no theoretical background to support them. Other algorithms have been developed in the machine learning community. These algorithms come with theoretical bounds on the discrepancy between the performances of the forecast combination and the performance of the best combination (with constant weights). Preliminary tests in Mallet and Sportisse [2006a] and undergoing work (in collaboration with Gilles Stoltz, École Normale Supérieure de Paris) show that good performances may be reached with these algorithms: the root mean square error may decrease down to $19.5\ \mu\mathrm{g\,m}^{-3}$.

Conclusion

The air quality modeling system Polyphemus enables ensemble forecasting thanks to its flexible structure. Monte Carlo simulations were carried to assess the uncertainties due to input data (except base meteorological fields). The uncertainties due to the model formulation were also studied through a multimode ensemble, and a strong impact on the output uncertainties was found. The multimodel ensemble brings useful information. Its members can be efficiently combined to improve the forecasts.

Among the next steps, one may find the use of ensembles in daily operational forecasts. Polyphemus has already run for several months in ensemble mode (with 43 models and daily forecasts) for a test on the operational platform Prév'air (operated by INERIS), but no online combination of the members was performed.

References

Hanna, S. R., Chang, J. C., and Fernau, M. E. (1998). Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables. *Atmos. Env.*, 32(21):3,619–3,628.

Hanna, S. R., Lu, Z., Frey, H. C., Wheeler, N., Vukovich, J., Arunachalam, S., Fernau, M., and Hansen, D. A. (2001). Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied

to the July 1995 OTAG domain. *Atmos. Env.*, 35(5):891–903.

Horowitz, L. W., Walters, S., Mauzerall, D. L., Emmons, L. K., Rasch, P. J., Granier, C., Tie, X., Lamarque, J.-F., Schultz, M. G., Tyndall, G. S., Orlando, J. J., and Brasseur, G. P. (2003). A global simulation of tropospheric ozone and related tracers: description and evaluation of MOZART, version 2. *J. Geophys. Res.*, 108(D24).

Krishnamurti, T. N., Kishtawal, C. M., Zhang, Z., T. LaRow, D. B., and Williford, E. (2000). Multimodel ensemble forecasts for weather and seasonal climate. *J. Climate*, 13:4,196–4,216.

Louis, J.-F. (1979). A parametric model of vertical eddy fluxes in the atmosphere. *Boundary-Layer Meteor.*, 17:187–202.

Mallet, V. (2005). *Estimation de l'incertitude et prévision d'ensemble avec un modèle de chimie-transport – Application à la simulation numérique de la qualité de l'air*. PhD thesis, École nationale des ponts et chaussées.

Mallet, V., Quélo, D., Sportisse, B., Ahmed de Biasi, M., Debry, É., Korsakissok, I., Wu, L., Roustan, Y., Sartelet, K., Tombette, M., and Foudhil, H. (2007). Technical Note: The air quality modeling system Polyphemus. *Atmos. Chem. Phys. Discuss.*, 7(3):6,459–6,486.

Mallet, V. and Sportisse, B. (2006a). Ensemble-based air quality forecasts: A multimodel approach applied to ozone. *J. Geophys. Res.*, 111.

Mallet, V. and Sportisse, B. (2006b). Uncertainty in a chemistry-transport model due to physical parameterizations and numerical approximations: An ensemble approach applied to ozone modeling. *J. Geophys. Res.*, 111(D1).

Middleton, P., Stockwell, W. R., and Carter, W. P. L. (1990). Aggregation and analysis of volatile organic compound emissions for regional modeling. *Atmos. Env.*, 24A(5):1,107–1,133.

Simpson, D., Winiwarter, W., Börjesson, G., Cinderby, S., Ferreiro, A., Guenther, A., Hewitt, C. N., Janson, R., Khalil, M. A. K., Owen, S., Pierce, T. E., Puxbaum, H., Shearer, M., Skiba, U., Steinbrecher, R., Tarrasón, L., and Öquist, M. G. (1999). Inventorying emissions from nature in Europe. *J. Geophys. Res.*, 104(D7):8,113–8,152.

Stockwell, W. R., Kirchner, F., Kuhn, M., and Seefeld, S. (1997). A new mechanism for regional atmospheric chemistry modeling. *J. Geophys. Res.*, 102(D22):25,847–25,879.

Stockwell, W. R., Middleton, P., Chang, J. S., and Tang, X. (1990). The second generation regional acid deposition model chemical mechanism for regional air quality modeling. *J. Geophys. Res.*, 95(D10):16,343–16,367.

Troen, I. and Mahrt, L. (1986). A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. *Boundary-Layer Meteor.*, 37:129–148.

Verwer, J. G., Hundsdorfer, W., and Blom, J. G. (2002). Numerical time integration for air pollution models. *Surveys on Math. for Indus.*, 10:107–174.

Zhang, L., Brook, J. R., and Vet, R. (2003). A revised parameterization for gaseous dry deposition in air-quality models. *Atmos. Chem. Phys.*, 3:2,067–2,082.