PhD Thesis
Coupling Ensemble Forecast and Data Assimilation

Application to Air Quality Simulation

Vivien Mallet (INRIA)

March 2011

Scientific Issues

Nowadays, air quality forecasts are carried out using chemistry-transport models. Based on meteorological forecasts, these models compute pollutant concentrations (like ozone over Europe – see Figure 1) for a few days ahead. Shortcomings in the forecasts originate from the high uncertainties in the input data to the models (meteorological fields, emissions, ...) and in the physical formulation of the models (turbulence, chemistry, ...). In such a context, forecasting should not rely on a single model. Instead, a forecast should be based on an *ensemble of models* that should account for all uncertainty sources.

In order to reduce the uncertainties, *data assimilation* methods take advantage of ground-based and satellite observations. These assimilation methods actually merge the information contained in a numerical model and the information brought by the observations, so as to produce the estimate of the model state that minimizes the error variance. Several such methods are appropriate for high-dimensional systems like air quality models [1,4]. They naturally apply to a single model.

Meanwhile, better forecasts have been produced by ensemble methods in which the forecasts of several models are linearly combined [2]. The weights of the linear combination may be determined by *machine learning* algorithms, based on past observations and forecasts. This approach is often referred to as *ensemble forecasting*.

The objective of the PhD thesis is to develop methods that combine the two approaches: data assimilation and ensemble forecasting. A proper theoretical framework will be needed for these new methods. Application to air quality forecast will probe their efficiency. The first research advances in this direction are extremely promising [3].

Related publications:

[1] *A comparison study of data assimilation algorithms for ozone forecasts.* Wu, Mallet, Bocquet and Sportisse. Journal of Geophysical Research, 2008. [download the preprint]
[2] *Ozone ensemble forecast with machine learning algorithms.* Mallet, Stoltz and Mauricette. Journal of Geophysical Research, 2009. [download the preprint]
[3] *Ensemble forecast of analyses: coupling data assimilation and sequential aggregation.* Mallet. Journal of Geophysical Research, 2010. [download the preprint]
[4] *Reduced minimax state estimation.* Mallet and Zhuk. INRIA report, 2010. [download]

Application Context

The research team has developed powerful software, unique in several technical aspects, for data assimilation and ensemble forecasting in air pollution modeling. This software allows to quickly implement advanced methods, even though the involved numerical models are complex.

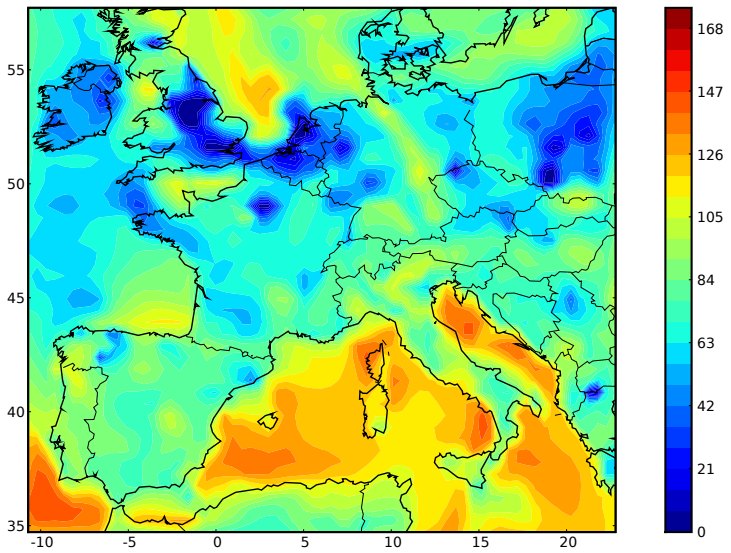

Figure 1: A forecast ozone map (ground concentration, in $\mu g\,m^{-3}$).

The air quality modeling system Polyphemus (http://cerea.enpc.fr/polyphemus/) will be employed. This system is at the same time a research tool and an operational tool (on a daily basis, it carries out operational air quality forecasts over Europe). In addition, the team is currently developing the data assimilation library Verdandi (http://verdandi.gforge.inria.fr/) that will serve as a reliable framework during the thesis.

Institutional Context

INRIA, the French national institute for research in computer science and control, operating under the dual authority of the Ministry of Research and the Ministry of Industry, is dedicated to fundamental and applied research in information and communication science and technology. The Institute also plays a major role in technology transfer by fostering training through research, diffusion of scientific and technical information, development, as well as providing expert advice and participating in international programs.

The student will be hosted in the project-team CLIME[1] which is located in the INRIA research center Paris-Rocquencourt. This team is joint with CEREA[2].

Further Information and Contact

Research field: applied mathematics in numerical modeling
Starting in: October 2011 (this can be adjusted)
Duration: 3 years (which is the normal duration in France)
Net pay: 1596 euros per month[3], INRIA support
Location: INRIA Paris-Rocquencourt (in Rocquencourt, close to Versailles, also accessible from Paris), in the project-team CLIME
Supervisor: Vivien Mallet (research scientist at INRIA)
Contact: Vivien.Mallet@inria.fr

[1] http://www-rocq.inria.fr/clime/index.en.html

[2] Research and teaching centre in atmospheric environment, http://cerea.enpc.fr/en/index.html – itself joint École des ponts / EDF R&D laboratory

[3] In the first two years. Third year: 1679 euros per month.