Data assimilation for smart cities

PhD thesis

1 Context and objectives

In their move toward more efficient management of infrastructures, communication and services, cities need better tools for **monitoring** and **decision making** in various areas. These tools may rely on in situ **observations** and **numerical simulations**. The two information sources are complementary, and future urban management systems should optimally combine both.

Data assimilation methods are precisely developed in order to couple numerical models and observations. They aim to estimate the state of a system, to improve forecasts and to quantify the associated uncertainties. Data assimilation has originally been developed for meteorology, where it made a huge contribution to the quality of the forecasts. Nowadays, many geosciences make use of data assimilation. The assimilation techniques are even breaking through medicine and biology.

At city scale, several models are already available to simulate the road traffic, the noise exposure, the sanitation system, ... Observations, often gathered from sensors distributed across the city, provide accurate but local information. The objective of this PhD program is to optimally combine these models and the relevance observations for a better management of the city. A consistent and complete urban modeling system will be designed, using one or several numerical models and all kinds of observations. The scientific objectives will mainly be **state estimation**, a priori and a posteriori (i.e., before and after assimilation) **uncertainty quantification** and **optimal design of an observation network**.

The developments will at least be applied to **air quality**. The optimal design of the observation network will be a key part of the project. There will be a special focus on the use of mobile sensors (e.g., installed on public buses, or embedded in smartphones) for which the data-gathering requires an optimized strategy. Other applications may be tackled, like noise exposure, which is a new issue investigated by two teams involved in this PhD program.

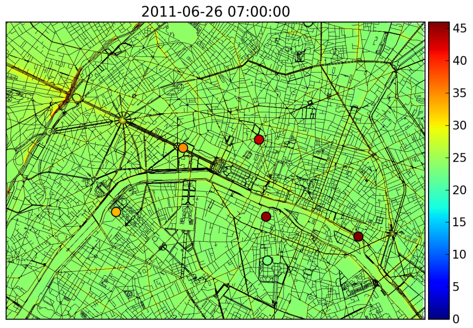

Figure 1: Map of NO_2 concentrations ($\mu g\,m^{-3}$) over a part of Paris, on 26 June 2011 at 07:00; before assimilation of the observations (marked with the disks), on the left, and after the assimilation, on the right. In this case, the assimilation makes large corrections to the initial simulated map. This result is from the project "Votre Air", with SME Numtech and the association Airparif. An important question of the PhD program is the assimilation of a large number of observations from mobile sensors.

2 Strategy

The methods for uncertainty quantification are computationally intensive to the point they could not be applied to the models currently available. Consequently, the research should rely on model reduction, using **statistical emulation** (Gaussian processes). A reduced model should reproduce the main features of the corresponding complete model. In particular, its forecast skills have to be equivalent along a trajectory followed by a mobile sensor.

The uncertainties in the inputs (in the case of air quality: mainly meteorological variables and emissions) can be propagated using Monte Carlo simulations. These simuations need to be calibrated using observations, so that the spread of the simulations represents the actual uncertainty of the model. The **calibration** may be carried out to quantify the **a priori uncertainty** as well as the **a posteriori uncertainty** (after assimilation of observational data). A promising approach to compute the a posteriori uncertainty is **Bayesian inference**, relying on Markov chain Monte Carlo and the reduced model.

Data assimilation should be employed to **reduce the uncertainties**. An important question is the **assimilation of observations from mobile sensors**. Various questions are absolutely open: how many mobile sensors are necessary, what data should be retrieved and when (to minimize the costly communications on the network), what instrumental accuracy is required, what is the impact of potential error correlations between the mobile sensors, which fixed sensors would be the best addition to the mobile sensors, ... The relevant mathematical methods will be identified and adapted to address these questions.

All methods employed in this research should be highly consistent with each other. In the long term, we aim to integrate them all in a **decision-making software** that will be accessible to operational actors and will show an overall consistency. It is very likely that model reduction will be a key component around which the system can be built.

3 Involved research teams

The hosting Inria project-team, Clime, works on data assimilation for environmental problems. It is part of the laboratory "Centre d'enseignement et de recherche en environnement atmosphérique" (CEREA) whose main activity is air quality simulation (see figure 1). Among other collaborations, Clime works with INERIS, the air quality agency Airparif and the SME Numtech for air quality at urban scale. Clime develops the data assimilation library Verdandi.

The Inria project-team Urbanet works on interconnected communication devices in urban capillary networks. Urbanet deals with mobile sensors and especially the optimal retrieval of their data to minimize communication costs. Urbanet exploits mobility data which will be helpful in the study of potential sensors installed on buses or cars.

The Inria project-team ARLES-MiMove designs software architectures for distributed systems on hybrid networks (e.g., hosting different kinds of sensors). ARLES-MiMove is especially interested in citizen involvement in these networks, e.g., measuring ambient noise in a city using smartphones.

4 Additional information and contact

PhD position starting: October 2014 (negotiable)
Duration: 3 years
Salary: 1596 euros net per month during the first two years; 1679 euros net per month during the third year
Location: Inria Paris-Rocquencourt (in Rocquencourt, near Versailles and accessible with a shuttle from Paris), in project-team Clime
Supervision: Vivien Mallet (Clime), Hervé Rivano (Urbanet), Valérie Issarny (ARLES-MiMove)
Contact:

- `Vivien.Mallet@inria.fr` — +33 1 39 63 55 76
- `Herve.Rivano@inria.fr`
- `Valerie.Issarny@inria.fr`