IN PARTNERSHIP WITH:
Ecole des Ponts ParisTech

Activity Report 2012

Project-Team CLIME

Coupling environmental data and simulation models for software integration

IN COLLABORATION WITH: Centre d'Enseignement et de Recherche en Environnement Atmosphérique

RESEARCH CENTER
Paris - Rocquencourt

THEME
Observation and Modeling for Environmental Sciences

Table of contents

Project-Team CLIME

Keywords: Data Assimilation, Geophysics, Image Processing, Inverse Problem, Simulation

Creation of the Project-Team: September 01, 2005 .

1. Members

Research Scientists
- Marc Bocquet [Deputy Director of CEREA, École des Ponts ParisTech, HdR]
- Isabelle Herlin [Team Leader, Senior Researcher, Inria]
- Vivien Mallet [Junior Researcher, Inria]

Faculty Member
- Etienne Huot [Associate Professor, University of Versailles Saint-Quentin-en-Yvelines (UVSQ)]

External Collaborators
- Dominique Béréziat [Associate Professor, University Pierre and Marie Curie (UPMC), HdR]
- Sergiy Zhuk [Research scientist, IBM Research Dublin, Ireland]

Engineers
- Raphaël Périllat [since December, Inria]
- Kévin Charpentier [Inria]
- Anne Tilloy [Inria]
- Sylvain Doré [since October, École des Ponts ParisTech]

PhD Students
- Karim Drifi [University Paris-Centre, Inria]
- Mohammad Reza Koohkan [University Paris-Est, École des Ponts ParisTech]
- Yann Lepoittevin [University Paris-Centre, since October, Inria]
- Paul Baudin [Ecole Polytechnique, joint to Classic and Clime, since September, Inria]
- Victor Winiarek [University Paris-Est, École des Ponts ParisTech]

Post-Doctoral Fellows
- Lin Wu [until February, École des Ponts ParisTech]
- Guiseppe Papari [until July, ERCIM]

Administrative Assistants
- Véronique Dehlinger [École des Ponts ParisTech]
- Nathalie Gaudechoux [joint to Axis, Cascade and Clime, Inria]

2. Overall Objectives

2.1. Clime in short

The international politic, economic and scientific contexts are pointing out the role that is played by models and observation systems for forecasting and evaluating environmental risks.

The complexity of the environmental phenomena as well as the operational objectives of risk mitigation necessitate an intensive interweaving between geophysical models, data processing, simulation, visualization and database tools.

For illustration purpose, we observe that this situation is met in the domain of atmospheric pollution, whose modeling is gaining an ever-increasing significance and impact, either at local (air quality), regional (transboundary pollution) or global scale (greenhouse effect). In this domain, numerical modeling systems are used for operational forecasts (short or long term), detailed case studies, impact studies for industrial sites, as well as coupled modeling, such as pollution and health or pollution and economy. These scientific subjects strongly require linking/coupling the models with all available data either of physical origin (e.g., models outputs), coming from raw observations (satellite acquisitions and/or information measured *in situ* by an observation network) or obtained by processing and analysis of these observations (e.g., chemical concentrations retrieved by inversion of a radiative transfer model).

Clime has been jointly created, by Inria and École des Ponts ParisTech, for studying these questions with researchers in data assimilation, image processing, and modeling.

Clime carries out research activities in three main areas:

- Data assimilation methods: inverse modeling, network design, ensemble methods, uncertainties estimation, uncertainties propagation.
- Image assimilation: assimilation of structures in environmental forecasting models, study of ill-posed image processing problems with data assimilation technics, definition of dynamic models from images, reduction of models.
- Development of integrated chains for data/models/outputs (system architecture, workflow, database, visualization).

3. Scientific Foundations

3.1. Data assimilation and inverse modeling

This activity is one major concern of environmental sciences. It matches up the setting and the use of data assimilation methods, for instance variational methods (such as the 4D-Var method). An emerging issue lies in the propagation of uncertainties by models, notably through ensemble forecasting methods.

Although modeling is not part of the scientific objectives of Clime, the project-team has complete access to models developed by CEREA: the models from Polyphemus (pollution forecasting from local to regional scales) and Code_Saturne (urban scale). In regard to other modeling domains, such as meteorology and oceanography, Clime accesses models through co-operation initiatives.

The research activities of Clime tackle scientific issues such as:

- Within a family of models (differing by their physical formulations and numerical approximations), which is the optimal model for a given set of observations?
- How to reduce dimensionality of problems by Galerkin projection of equations on subspaces? How to define these subspaces in order to keep the main properties of systems?
- How to assess the quality of a forecast and its uncertainty? How do data quality, missing data, data obtained from sub-optimal locations, affect the forecast? How to better include information on uncertainties (of data, of models) within the data assimilation system?
- How to make a forecast (and a better forecast!) by using several models corresponding to different physical formulations? It also raises the question: how should data be assimilated in this context?
- Which observational network should be set up to perform a better forecast, while taking into account additional criteria such as observation cost? What are the optimal location, type and mode of deployment of sensors? How should trajectories of mobile sensors be operated, while the studied phenomenon is evolving in time? This issue is usually referred as “network design”.

3.2. Satellite acquisitions and image assimilation

In geosciences, the issue of coupling data, in particular satellite acquisitions, and models is extensively studied for meteorology, oceanography, chemistry-transport and land surface models. However, satellite images are mostly assimilated on a point-wise basis. Three major approaches arise if taking into account the spatial structures, whose displacement is visualized on image sequences:

- Image approach. Image assimilation allows the extraction of features from image sequences, for instance motion field or structures' trajectory. A model of the dynamics is considered (obtained by simplification of a geophysical model such as Navier-Stokes equations). An observation operator is defined to express the links between the model state and the pixel value. In the simplest case, the pixel value corresponds to one coordinate of the model state and the observation operator is reduced to a projection. However, in most cases, this operator is highly complex, implicit and non-linear. Data assimilation techniques are developed to control the initial state or the whole assimilation window. Image assimilation is also applied to learn reduced models from image data and estimate a reliable and small-size reconstruction of the dynamics, which is observed on the sequence.
- Model approach. Image assimilation is used to control an environmental model and obtain improved forecasts. In order to take into account the spatial and temporal coherency of structures, specific image characteristics are considered, and dedicated norms and observation error covariances are defined.
- Correcting a model. Another topic, mainly described for meteorology in the literature, concerns the location of structures. How to force the existence and to correct the location of structures in the model state using image information? Most of the operational meteorological forecasting institutes, such as MétéoFrance, UK-met, KNMI (in Netherlands), ZAMG (in Austria) and Met-No (in Norway), study this issue because operational forecasters often modify their forecasts based on visual comparisons between the model outputs and the structures displayed on satellite images.

3.3. Software chains for environmental applications

An objective of Clime is to participate in the design and creation of software chains for impact assessment and environmental crisis management. Such software chains bring together static or dynamic databases, data assimilation systems, forecast models, processing methods for environmental data and images, complex visualization tools, scientific workflows, ...

Clime is currently building, in partnership with École des Ponts ParisTech and EDF R&D, such a system for air pollution modeling: Polyphemus (see the web site http://cerea.enpc.fr/polyphemus/), whose architecture is specified to satisfy data requirements (e.g., various raw data natures and sources, data preprocessing) and to support different uses of an air quality model (e.g., forecasting, data assimilation, ensemble runs).

4. Application Domains

4.1. Introduction

The central application domain of the project-team is atmospheric chemistry. We develop and maintain the air quality modeling system Polyphemus, which includes several numerical models (Gaussian models, Lagrangian model, two 3D Eulerian models including Polair3D) and their adjoints, and different high level methods: ensemble forecast, sequential and variational data assimilation algorithms. Advanced data assimilation methods, network design, inverse modeling, ensemble forecast are studied in the context of air chemistry. Note that addressing these high level issues requires controlling the full software chain (models and data assimilation algorithms).

The activity on assimilation of satellite data is mainly carried out for meteorology and oceanography. This is addressed in cooperation with external partners who provide numerical models. Concerning oceanography, the aim is to improve the forecast of ocean circulation, by assimilation of fronts and vortices displayed on image data. Concerning meteorology, the focus is on correcting the model location of structures related to high-impact weather events (cyclones, convective storms, *etc*.) by assimilating images.

4.2. Air quality

Air quality modeling implies studying the interactions between meteorology and atmospheric chemistry in the various phases of matter, which leads to the development of highly complex models. The different usages of these models comprise operational forecasting, case studies, impact studies, *etc*, with both societal (e.g., public information on pollution forecast) and economical impacts (e.g., impact studies for dangerous industrial sites). Models lack some appropriate data, for instance better emissions, to perform an accurate forecast and data assimilation techniques are recognized as a major key point for the improvement of forecast's quality.

In this context, Clime is interested in various problems, the following being the crucial ones:

- The development of ensemble forecast methods for estimating the quality of the prediction, in relation with the quality of the model and the observations. Sensitivity analysis with respect to the model's parameters so as to identify physical and chemical processes, whose modeling must be improved.
- The development of methodologies for sequential aggregation of ensemble simulations. What ensembles should be generated for that purpose, how spatialized forecasts can be generated with aggregation, how can the different approaches be coupled with data assimilation?
- The definition of second-order data assimilation methods for the design of optimal observation networks. Management of combinations of sensor types and deployment modes. Dynamic management of mobile sensors' trajectories.
- How to estimate the emission rate of an accidental release of a pollutant, using observations and a dispersion model (from the near-field to the continental scale)? How to optimally predict the evolution of a plume? Hence, how to help people in charge of risk evaluation for the population?
- The definition of non-Gaussian approaches for data assimilation.
- The assimilation of satellite measurements of troposphere chemistry.

The activities of Clime in air quality are supported by the development of the Polyphemus air quality modeling system. This system has a modular design, which makes it easier to manage high level applications such as inverse modeling, data assimilation and ensemble forecast.

4.3. Oceanography

The capacity of performing a high quality forecast of the state of the ocean, from the regional to the global scales, is of major interest. Such a forecast can only be obtained by systematically coupling numerical models and observations (*in situ* and satellite data). In this context, being able to assimilate image structures becomes a key point. Examples of such image structures are:

- apparent motion linked to surface velocity;
- trajectories, obtained either from tracking of features or from integration of the velocity field;
- spatial objects, such as fronts, eddies or filaments.

Image Models for these structures are developed and take into account the underlying physical processes. Image data are assimilated in Image Models to derive pseudo-observations of state variables, which are further assimilated in numerical ocean forecast models.

4.4. Meteorology

Meteorological forecasting constitutes a major applicative challenge for Image Assimilation. Although satellite data are operationally assimilated within models, this is mainly done on an independent pixel basis: the observed radiance is linked to the state variables via a radiative transfer model, that plays the role of an observation operator. Indeed, because of their limited spatial and temporal resolutions, numerical weather forecast models fail to exploit image structures, such as precursors of high impact weather:

- cyclogenesis related to the intrusion of dry stratospheric air in the troposphere (a precursor of cyclones),
- convective systems (supercells) leading to heavy winter time storms,
- low-level temperature inversion leading to fog and ice formation, *etc*.

To date, there is no available method for assimilating such data, which are characterized by a strong coherence in space and time. Meteorologists have developed qualitative Conceptual Models (CMs), for describing the high impact weathers and their signature on images, and tools to detect CMs on image data. The result of this detection is used for correcting the numerical models, for instance by modifying the initialization. The aim is therefore to develop a methodological framework allowing to assimilate the detected CMs within numerical forecast models. This is a challenging issue given the considerable impact of the related meteorological events.

5. Software

5.1. Polyphemus

Participants: Vivien Mallet, Anne Tilloy.

Polyphemus (see the web site http://cerea.enpc.fr/polyphemus/) is a modeling system for air quality. As such, it is designed to yield up-to-date simulations in a reliable framework: data assimilation, ensemble forecast and daily forecasts. Its completeness makes it suitable for use in many applications: photochemistry, aerosols, radionuclides, *etc*. It is able to handle simulations from local to continental scales, with several physical models. It is divided into three main parts:

- libraries that gather data processing tools (SeldonData), physical parameterizations (AtmoData) and postprocessing abilities (AtmoPy);
- programs for physical preprocessing and chemistry-transport models (Polair3D, Castor, two Gaussian models, a Lagrangian model);
- drivers on top of the models in order to implement advanced simulation methods such as data assimilation algorithms.

Figure 1 depicts a typical result produced by Polyphemus. Clime is involved in the overall design of the system and in the development of advanced methods in model coupling, data assimilation and ensemble forecast (through drivers and post-processing).

In 2012, Polyphemus received several physical developments on secondary organic aerosols, modeling of pollution due to traffic emissions and coupling of local and regional scale models. Further integration with the data assimilation library Verdandi has been carried out. A Python interface to the Eulerian model Polair3D has been introduced.

5.2. Data assimilation library: Verdandi

Participants: Kévin Charpentier, Marc Fragu [MACS], Vivien Mallet, Dominique Chapelle [MACS], Philippe Moireau [MACS], Sergiy Zhuk, Anne Tilloy.

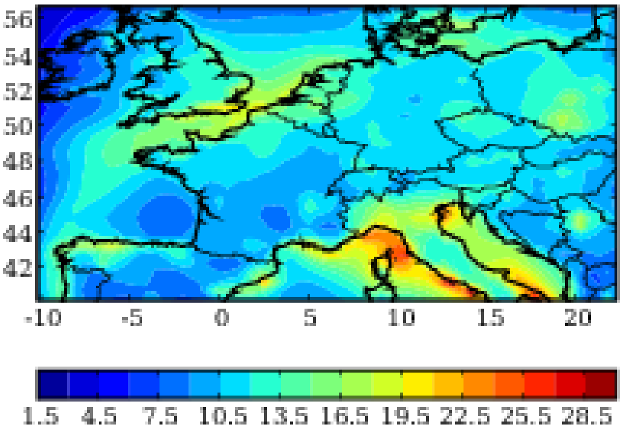

Figure 1. Map of the relative standard deviation (or spread, %) of an ensemble built with Polyphemus (ozone simulations, $\mu g\ m^{-3}$). The standard deviations are averaged over the summer of 2001. They provide an estimation of the simulation uncertainties.

The leading idea is to develop a data assimilation library intended to be generic, at least for high-dimensional systems. Data assimilation methods, developed and used by several teams at Inria, are generic enough to be coded independently of the system to which they are applied. Therefore these methods can be put together in a library aiming at:

- making easier the application of methods to a great number of problems,
- making the developments perennial and sharing them,
- improving the broadcast of data assimilation works.

An object-oriented language (C++) has been chosen for the core of the library. A high-level interface to Python is automatically built. The design study raised many questions, related to high dimensional scientific computing, the limits of the object contents and their interfaces. The chosen object-oriented design is mainly based on three class hierarchies: the methods, the observation managers and the models. Several base facilities have also been included, for message exchanges between the objects, output saves, logging capabilities, computing with sparse matrices.

In 2012, versions 1.2, 1.3 and 1.4 were released. The design of the library has been improved for optimal performance and is now stable. It is possible to write models and observation managers in Python. Efficient support for parallel computations has been introduced. The documentation has been improved.

5.3. Urban air quality analysis

Participants: Anne Tilloy, Vivien Mallet.

“Urban Air Quality Analysis” carries out data assimilation at urban scale. It merges the outputs of a numerical model (maps of pollutant concentrations) with observations from an air quality monitoring network, in order to produce the so-called analyses, that is, corrected concentration maps. The data assimilation computes the Best Linear Unbiased Estimator (BLUE), with a call to the data assimilation library Verdandi. The error covariance matrices are parameterized for both model simulations and observations. For the model state error covariances, the parameterization primarily relies on the road network. The software handles ADMS Urban output files, for a posteriori analyses or in an operational context.

6. New Results

6.1. New methods for data assimilation

Since the beginning, Clime is focused on developing new techniques for data assimilation in geophysical sciences. Clime is active on several of the most challenging theoretical aspects of data assimilation: data assimilation methods based on non-Gaussian assumptions, methods for estimating errors, ensemble filtering techniques, 4D variational assimilation approaches, ensemble-variational methods, etc. This year, we revisited several of these topics. A dual algorithm has been developed for the finite-size ensemble Kalman filter, that shows how to estimate optimal inflation that counteracts sampling errors. A variational method coupled to a subgrid scale statistical model has been introduced and validated to quantify the representativeness errors. We also started to work on ensemble variational methods that are brand new techniques emerging in the meteorological data assimilation field.

6.1.1. Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems

Participants: Marc Bocquet, Pavel Sakov [NERSC, Norway].

The finite-size ensemble Kalman filter (EnKF-N) is an ensemble Kalman filter (EnKF) which, in perfect model condition, does not require inflation because it partially accounts for the ensemble sampling errors. For the Lorenz '63 and '95 toy-models, it was so far shown to perform as well or better than the EnKF with an optimally tuned inflation. The iterative ensemble Kalman filter (IEnKF) is an EnKF which was shown to perform much better than the EnKF in strongly nonlinear conditions, such as with the Lorenz '63 and '95 models, at the cost of iteratively updating the trajectories of the ensemble members. This study aims at further exploring the two filters, and at combining both into an EnKF that does not require inflation in perfect model condition and which is as efficient as the IEnKF in very nonlinear conditions.

In this study EnKF-N is first introduced and a new implementation is developed. It decomposes EnKF-N into a cheap two-step algorithm that amounts to computing an optimal inflation factor. This offers a justification of the use of the inflation technique in the traditional EnKF and why it can often be efficient. Secondly, the IEnKF is introduced following a new implementation based on the Levenberg-Marquardt optimization algorithm. Then, the two approaches are combined to obtain the finite-size iterative ensemble Kalman filter (IEnKF-N). Several numerical experiments are performed on IEnKF-N with the Lorenz '95 model. These experiments demonstrate its numerical efficiency as well as its performance that offer, at least, the best of both filters.

6.1.2. Accounting for representativeness errors in the inversion of atmospheric constituent emissions: Application to the retrieval of regional carbon monoxide fluxes

Participants: Mohammad Reza Koohkan, Marc Bocquet.

A four-dimensional variational data assimilation system (4D-Var) is developed to retrieve carbon monoxide (CO) fluxes at regional scale, using an air quality network. The air quality stations that monitor CO are proximity stations located close to industrial, urban or traffic sources. The mismatch between the coarsely discretized Eulerian transport model and the observations, inferred to be mainly due to representativeness errors in this context, leads to a bias (averaged simulated concentrations minus observed concentrations) of the same order of magnitude as the concentrations. 4D-Var leads to a mild improvement in the bias because it does not adequately handle the representativeness issue. For this reason, a simple statistical subgrid model is introduced and is coupled to 4D-Var. In addition to CO fluxes, the optimization seeks to jointly retrieve *influence coefficients*, which quantify each station's representativeness. The method leads to a much better representation of the CO concentration variability, with a significant improvement of statistical indicators. The resulting increase in the total inventory estimate is close to the one obtained from remote sensing data assimilation. This methodology and experiments suggest that information useful at coarse scales can be better extracted from atmospheric constituent observations strongly impacted by representativeness errors.

6.1.3. *Real-time data assimilation*

Participants: Vivien Mallet, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech], Cécile Honoré [Airparif], Édouard Debry [INERIS].

Based on Verdandi, Polyphemus and the "Urban Air Quality Analysis" software, real-time data assimilation was carried out at urban scale. The Best Linear Unbiased Estimator (BLUE) was computed for every hourly concentration map that the ADMS model computed. A posteriori tests were conducted over Clermont-Ferrand and Paris. We addressed the key issue of the covariance of the state error. The form of the error covariance between two points was determined based on the road network, considering the distance between points along the road and the distance of each point to the road. A few parameters (primarily two decorrelation lengths) were determined thanks to cross validation with several months of simulations and observations. The results showed strong improvements even at locations where no data was assimilated.

At larger scale, the data assimilation library Verdandi was used to apply data assimilation (optimal interpolation) with the air quality model Chimere. This preliminary work will help INERIS to apply optimal interpolation for ozone and particulate matter in the operational platform Prev'air.

6.2. Inverse modeling

Many of this year's studies have focused on inverse modeling, including the reconstruction of the Fukushima radionuclide atmospheric and marine source terms. All were targeted to a particular application. However most of them include new methodological developments, in particular non-Gaussian data assimilation schemes.

6.2.1. *Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the Fukushima Daiichi source term*

Participants: Victor Winiarek, Marc Bocquet, Olivier Saunier [IRSN], Anne Mathieu [IRSN].

The aim of this research activity is the implementation of data assimilation methods, particularly inverse modeling methods, in the context of an accidental radiological release from a nuclear power plant and their application in the specific case of the Fukushima Daiichi accident. The particular methodological focus is the a posteriori estimation of the prior errors statistics. In the case of the Fukushima Daiichi accident, the number of available observations is small compared to the number of source parameters to retrieve and the reconstructed source is highly sensitive to the prior errors. That is the why they need to be well established and justified. In this aim, three methods have been proposed: one method relies on a L-curve estimation technique, another one on the Desroziers' iterative scheme and the last method, assumed to be the most robust, relies on the maximum likelihood principle, generalised to a non-Gaussian context. These three methods have been applied to the reconstruction of cesium-137 and iodine-131 source terms from the Fukushima Daiichi accident. Because of the poor observability of the Fukushima Daiichi emissions, these methods provide lower-bounds for cesium-137 and iodine-131 reconstructed activities. Nevertheless, with the new method based on semi-Gaussian statistics for the background errors, the lower-bound estimates for cesium-137, $1.2 - 4.0\,10^{16}$ Bq with an estimated standard deviation range of $15 - 20\%$, and for iodine-131, $1.9 - 3.8\,10^{17}$ Bq with an estimated standard deviation range of $5 - 10\%$, are of the same order of magnitude as those provided by the Japanese Nuclear and Industrial Safety Agency, and about 5 to 10 times less than the Chernobyl atmospheric releases.

6.2.2. *Assessment of the amount of Cesium-137 released into the Pacific Ocean after the Fukushima accident and analysis of its dispersion in Japanese coastal waters*

Participants: Claude Estournel [LA], Emmanuel Bosc [IAEA], Marc Bocquet, Caroline Ulses [LA], Patrick Marsailex [LA], Victor Winiarek, Iolanda Osvath [IAEA], Cyril Nguyen [LA,LEGOS], Thomas Duhaut [LA], Florent Lyard [LEGOS], Héloïse Michaud [LA], Francis Auclair [LA].

Numerical modeling was used to provide a new estimate of the amount of cesium-137 released directly into the ocean from the Fukushima Daiichi nuclear power plant (NPP) after the accident in March 2011 and to gain insights into the physical processes that led to its dispersion in the marine environment during the months following the accident. An inverse method was used to determine the time-dependent cesium-137 input responsible for the observed concentrations. The method was then validated through comparisons of the simulated concentrations with concentrations measured in seawater at different points in the neighborhood of the nuclear power plant. An underestimation was noticed for stations located 30 km offshore. The resulting bias in the release inventory was estimated. Finally, the maximum cesium-137 activity released directly to the ocean was estimated to lie between 5.1 and 5.5 PBq (Peta Becquerel = 1015 Bq) but uncertainties remain on the amount of radionuclides released during the first few days after the accident. This estimate was compared to previous ones and differences were further analysed. The temporal and spatial variations of the cesium-137 concentration present in the coastal waters were shown to be strongly related to the wind intensity and direction. During the first month after the accident, winds blowing toward the south confined the radionuclides directly released into the ocean to a narrow coastal band. Afterwards, frequent northward wind events increased the dispersion over the whole continental shelf, leading to strongly reduced concentrations.

6.2.3. *What eddy-covariance measurements tell us about prior land flux errors in CO2-flux inversion schemes?*

Participants: Frédéric Chevallier [LSCE], Tao Wang [LSCE], Philippe Ciais [LSCE], Marc Bocquet, Altaf Arain [McMaster University, Canada], Alessandro Cescatti [Joint Research Centre, Italy], Jiquan Chen [University of Toledo, USA], Johannes Dolman [Vrije Universiteit, the Netherlands], Beverly Law [Oregon State University, USA], Hank Margolis [Université Laval, Canada], Leonardo Montagnani [University of Bolzano, Italy].

To guide the future development of CO2-atmospheric inversion modeling systems, we analysed the errors arising from prior information about terrestrial ecosystem fluxes. We compared the surface fluxes calculated by a process-based terrestrial ecosystem model with daily averages of CO2 flux measurements at 156 sites across the world in the FLUXNET network. At the daily scale, the standard deviation of the model-data fit was 2.5 gC·m-2·d-1; temporal autocorrelations were significant at the weekly scale (> 0.3 for lags less than four weeks), while spatial correlations were confined to within the first few hundred kilometers (< 0.2 after 200 km). Separating out the plant functional types did not increase the spatial correlations, except for the deciduous broad-leaved forests. Using the statistics of the flux measurements as a proxy for the statistics of the prior flux errors was shown not to be a viable approach. A statistical model allowed us to upscale the site-level flux error statistics to the coarser spatial and temporal resolutions used in regional or global models. This approach allowed us to quantify how aggregation reduces error variances, while increasing correlations. As an example, for a typical inversion of grid point (300 km × 300 km) monthly fluxes, we found that the prior flux error follows an approximate e-folding correlation length of 500 km only, with correlations from one month to the next as large as 0.6.

6.3. Monitoring network design

In this section, we report studies that are related to the evaluation of monitoring networks and to new monitoring strategies. This year, network designs techniques have been applied to the inverse modeling of CO2 fluxes.

6.3.1. *Network design for mesoscale inversions of CO2 sources and sinks*

Participants: Thomas Lauvaux [Pennsylvania State University, USA], Andy Schuh [Colorado State University, USA], Marc Bocquet, Lin Wu, Scott Richardson [Pennsylvania State University, USA], Natasha Miles [Pennsylvania State University, USA], Ken Davis [Pennsylvania State University, USA].

Recent instrumental deployments of regional observation networks of atmospheric CO2 mixing ratios have been used to constrain carbon sources and sinks using inversion methodologies. In this study, we performed sensitivity experiments using observation sites from the Mid Continent Intensive experiment to evaluate the required spatial density and locations of CO2 concentration towers based on flux corrections and error reduction analysis. In addition, we investigated the impact of prior flux error structures with different correlation lengths and biome information. We show that, while the regional carbon balance converged to similar annual estimates using only two concentration towers over the region, additional sites were necessary to retrieve the spatial flux distribution of our reference case (using the entire network of eight towers). Local flux corrections required the presence of observation sites in their vicinity, suggesting that each tower was only able to retrieve major corrections within a hundred of kilometers around, despite the introduction of spatial correlation lengths (100 to 300 km) in the prior flux errors. We then quantified and evaluated the impact of the spatial correlations in the prior flux errors by estimating the improvement in the CO2 model-data mismatch of the towers not included in the inversion. The overall gain across the domain increased with the correlation length, up to 300 km, including both biome-related and non-biome-related structures. However, the spatial variability at smaller scales was not improved. We conclude that the placement of observation towers around major sources and sinks is critical for regional-scale inversions in order to obtain reliable flux distributions in space. Sparser networks seem sufficient to assess the overall regional carbon budget with the support of flux error correlations, indicating that regional signals can be recovered using hourly mixing ratios. However, the smaller spatial structures in the posterior fluxes are highly constrained by assumed prior flux error correlation lengths, with no significant improvement at only a few hundreds of kilometers away from the observation sites.

6.3.2. Potential of the International Monitoring System radionuclide network for inverse modeling

Participants: Mohammad Reza Koohkan, Marc Bocquet, Lin Wu, Monika Krysta [The Preparatory Commission for the Comprehensive Nuclear Test-Ban Treaty Organization, UNO].

The International Monitoring System (IMS) radionuclide network enforces the Comprehensive Nuclear-Test-Ban Treaty, which bans nuclear explosions. We have evaluated the potential of the IMS radionuclide network for inverse modeling of the source, whereas it is usually assessed by its detection capability. To do so, we have chosen the *degrees of freedom for the signal* (DFS), a well established criterion in remote sensing, in order to assess the performance of an inverse modeling system. Using a multiscale data assimilation technique, we have computed optimal adaptive grids of the source parameter space by maximizing the DFS. This optimization takes into account the monitoring network, the meteorology over one year (2009) and the relationships between the source parameters and the observations derived from the FLEXPART Lagrangian transport model. Areas of the domain, where the grid-cells of the optimal adaptive grid are large, emphasize zones where the retrieval is more uncertain, whereas areas, where the grid-cells are smaller and denser, stress regions where more source variables can be resolved. The observability of the globe through inverse modeling is studied in strong, realistic and small model error cases. The strong error and realistic error cases yield heterogeneous adaptive grids, indicating that information does not propagate far from the monitoring stations, whereas in the small error case, the grid is much more homogeneous.

In all cases, several specific continental regions remain poorly observed such as Africa as well as the tropics, because of the trade winds.

The northern hemisphere is better observed through inverse modeling (more than 60% of the total DFS), mostly because it contains more IMS stations. This unbalance leads to a better performance of inverse modeling in the northern hemisphere winter. The methodology is also applied to the subnetwork composed of the stations of the IMS network that measure noble gases.

6.4. Reduction and emulation

The use of environmental models raise a number of problems due to:

- the dimension of their inputs, which can easily be $10^5 - -10^8$ at every time step;

- the dimension of their state vector, which is usually $10^5 - -10^7$;
- their high computational cost.

In particular, the application of data assimilation methods and uncertainty quantification techniques may require dimension reduction and cost reduction. The dimension reduction consists in projecting the inputs and the state vector to low-dimensional subspaces. The cost reduction can be carried out by emulation, i.e., the replacement of costly components with fast surrogates.

6.4.1. Reduction and emulation of a chemistry-transport model

Participants: Vivien Mallet, Serge Guillas [University College London].

Both reduction and emulation were applied to the dynamic air quality model Polair3D from Polyphemus. The reduction relied on proper orthogonal decomposition on the input data and on the state vector. The dimension of the reduced subspace for the input data is about 80, while the dimension of the reduced state vector is less than 10. The projection of the state vector on its reduced subspace can be carried out before every integration time step, so that one can reproduce a full state trajectory (in time) using the reduced model.

Significant advances were made to emulate the reduced model, which requires about 90 inputs (reduced input data and reduced state vector) and computes about 10 outputs (reduced state vector). 90 inputs is however a large number to build an emulator using a classical approach like krigging. Promising results were however obtained with an interpolation method based on inverse distance weighting.

6.4.2. Reduction and emulation of a static air quality model

Participants: Vivien Mallet, Anne Tilloy, Fabien Brocheton [Numtech], David Poulet [Numtech].

The dimension reduction was applied to the outputs of the urban air quality model ADMS Urban, which is a static model with low-dimensional inputs and high-dimensional outputs. A proper orthogonal decomposition on the outputs allowed us to drastically reduce their dimension, from 10^4 to just a few scalars. First attempts of emulation of the reduced model rely on Gaussian process emulation.

6.4.3. Motion estimation from images with a wavelets reduced model

Participants: Giuseppe Papari, Isabelle Herlin, Etienne Huot, Karim Drifi.

The dimension reduction was applied to an image model, composed of Lagrangian constancy of velocity and transport of image brightness. Wavelets basis have been computed on the image domain for subspaces of images, motion fields and divergence-free motion fields. Image assimilation with this reduced model allows to estimate smooth velocity fields with properties defined by user. This also solves the issue of complex geographical domains and the difficulty of applying boundary conditions on these domains. First results are obtained with a reduced dimension of motion to a few scalars, to be compared with the original problem that has the size of image domain.

6.5. Ensemble forecasting with sequential aggregation

The aggregation of an ensemble of forecasts is an approach where the members of an ensemble are given a weight before every forecast time, and where the corresponding weighted linear combination of the forecasts provides an improved forecast. A robust aggregation can be carried out so as to guarantee that the aggregated forecast performs better, in the long run, than any linear combination of the ensemble members with time-independent weights. The approaches are then based on machine learning. The aggregation algorithms can be applied to forecast analyses (generated from a data assimilation system), so that the aggregated forecasts are naturally multivariate fields.

6.5.1. Application of sequential aggregation to meteorology and air quality

Participants: Anne Tilloy, Vivien Mallet, Fabien Brocheton [Numtech], David Poulet [Numtech].

Nowadays it is standard procedure to generate an ensemble of simulations for a meteorological forecast. Usually, meteorological centers produce a single forecast, out of the ensemble forecasts, computing the ensemble mean (where every model receives an equal weight). It is however possible to apply aggregation methods. Each time new observations are available, new weights for the linear combination are computed and applied for the next forecast. We applied the discounted ridge regression algorithm, which we previously introduced for sequential aggregation of air quality forecasts, to forecast wind and temperature at given observation stations. The ensemble was generated with forecasts at different range from two models. The aggregation proved to be efficient for one-day forecasts at least.

The discounted ridge regression was also applied to the simulations of the Air Quality Modeling Evaluation International Initiative (AQMEII) over Europe and North America, for different pollutants (gases and particulate matter).

6.5.2. Sequential aggregation with uncertainty estimation

Participants: Vivien Mallet, Sergiy Zhuk [IBM research], Paul Baudin, Gilles Stoltz [CLASSIC], Karine Sartelet [CEREA].

A new issue is the estimation of the uncertainties associated with the aggregated forecasts. One investigated direction relies on the framework of machine learning, with the aggregation of an ensemble of probability density functions instead of the point forecasts of the ensemble.

Another direction, which led to finalized results in 2012, is to reformulate the aggregation problem in a filtering problem for the weights. The weights are supposed to satisfy some dynamics with unknown model error, which defines the state equation of a filter. An observation equation compares the aggregated forecast with the observations (or analyses) with known observational error variance. The filter finally computes estimates for the weights and quantifies their uncertainties. We applied a Kalman filter and a minimax filter for air quality forecasting.

6.6. Image assimilation

Sequences of images, such as satellite acquisitions, display structures evolving in time. This information is recognized of major interest by forecasters (meteorologists, oceanographers, *etc*) in order to improve the information provided by numerical models. However, these satellite images are mostly assimilated in geophysical models on a point-wise basis, discarding the space-time coherence visualized by the evolution of structures such as clouds. Assimilating in an optimal way image data is of major interest and this issue should be considered in two ways:

- from the model's viewpoint, the problem is to control the location of structures using the observations,
- from the image's viewpoint, a model of the dynamics and structures has to be built from the observations.

6.6.1. Divergence-free motion estimation

Participants: Dominique Béréziat [UPMC], Isabelle Herlin, Sergiy Zhuk [IBM Research, Ireland].

This research addresses the issue of divergence-free motion estimation on an image sequence, acquired over a given temporal window. Unlike most state-of-the-art technics, which constrain the divergence to be small thanks to Tikhonov regularization terms, a method that imposes a null value of divergence of the estimated motion is defined.

Motion is characterized by its vorticity value and assumed to satisfy the Lagragian constancy hypothesis. An image model is then defined: the state vector includes the vorticity, whose evolution equation is derived from that of motion, and a pseudo-image that is transported by motion. An image assimilation method, based on the 4D-Var technics, is defined and developed that estimates motion as a compromise between the evolution equations of vorticity and pseudo-image and the observed sequence of images.

The method is applied on Sea Surface Temperature (SST) images acquired over Black Sea by NOAA-AVHRR sensors. The divergence-free assumption is roughly valid on these acquisitions, due to the small values of vertical velocity at the surface. Fig. 2 displays data and results. As no ground truth of motion is available, the method is quantified by the value of correlation between the pseudo-images and real acquisitions [28].

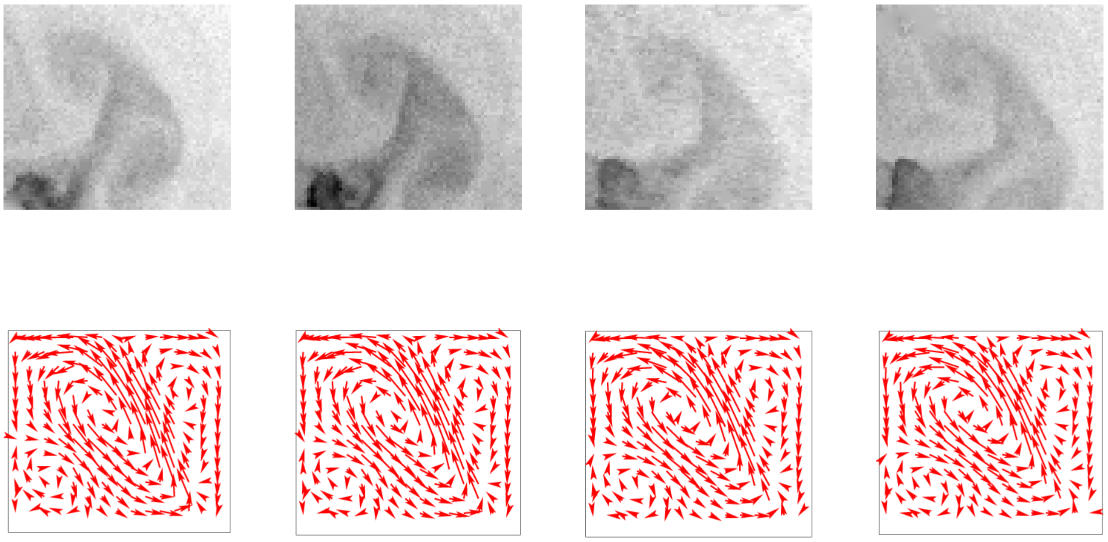

Figure 2. SST image observations and motion results.

6.6.2. Improvement of motion estimation by assessing errors on the dynamics

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

Data assimilation technics are used to retrieve motion from image sequences. These methods require a model of the underlying dynamics, displayed by the evolution of image data. In order to quantify the approximation linked to the chosen dynamic model, we consider adding a model error term in the evolution equation of motion and design a weak formulation of 4D-Var data assimilation. The cost function to be minimized simultaneously depends on the initial motion field, at the begining of the studied temporal window, and on the error value at each time step. The result allows to assess the model error and analyze its impact on motion estimation [27].

This error assessment method is evaluated and quantified on twin experiments, as no ground truth would be available for real image data. Fig. 3 shows four frames of a series of observations obtained by integrating the evolution model from an initial condition on image and velocity field (the ground truth $\mathbf{w}_{\mathrm{ref}}(0)$ displayed on the left of Fig. 4). An error value is added at each time step on the motion value, when integrating the simulation model. This error is a constant bias.

We performed two data assimilation experiments. The first one considers the evolution model as perfect, with no error in the evolution equation. It is denoted PM (for Perfect Model). The second one, denoted IM (for Imperfect Model) involves an error in the motion evolution equation. In Fig.4 are displayed the motion fields retrieved by PM and IM at the beginning of the temporal window.

As it can be seen, IM computes a correct velocity field while PM completely fails.

6.6.3. Nonlinear Observation Equation For Motion Estimation

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

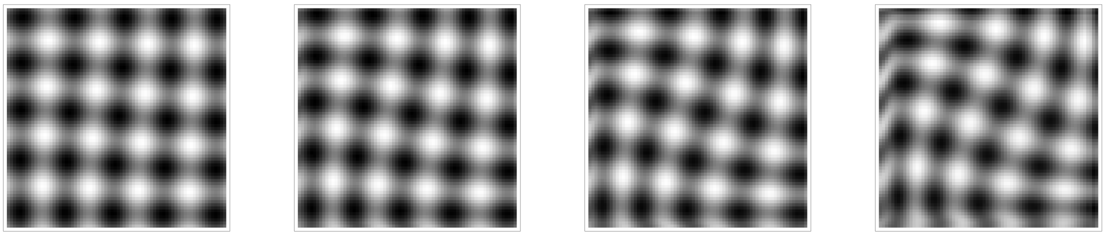

Figure 3. Observations Images.

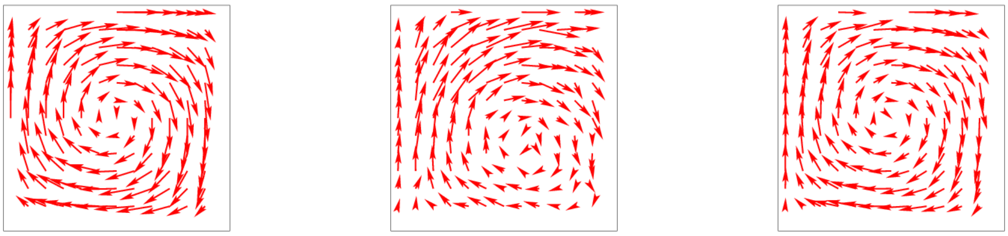

Figure 4. Left: ground-truth, middle: PM, right: IM.

In the image processing literature, the optical flow equation is usually chosen to assess motion from an image sequence. However, it corresponds to an approximation that is no more valid in case of large displacements. We evaluated improvements obtained when using the non linear transport equation of the image brightness by the velocity field [25]. A 4D-Var data assimilation method is designed that simultaneously solves the evolution equation and the observation equation, in its non linear and linearized form. The comparison of results obtained with both observation equations is quantified on synthetic data and discussed on oceanographic Sea Surface Temperature (SST) images. We show that the non linear model outperforms the linear one, which underestimates the motion norm. Fig.5 illustrates this on SST images (motion vectors are displayed by arrows).

The aim of this research is to achieve a correct estimation of motion when the object displacement is greater than its size. Howerver, in this case, coarse-to-fine incremental methods as well as the non linear data assimilation method fail to retrieve a correct value. The perspective is then to include, in the state vector, a variable describing the trajectory of pixels. The observation operator will then measure the effective displacement of pixels, according to their trajectories, and allow a better estimation of motion value.

6.6.4. Sliding windows method for motion estimation on long temporal image sequences

Participants: Karim Drifi, Isabelle Herlin.

This study concerns the estimation of motion fields from satellite images on long temporal sequences. The huge computational cost and memory required by data assimilation methods on the pixel grid makes impossible to use these techniques on long temporal intervals. For a given dynamic model (named full model), on the pixel grid, the Galerkin projection on subspaces provides a reduced model, that allows image assimilation at low cost. The definition of this reduced model however requires defining an optimal subspace of motion. A **sliding windows** method is thus designed:

- The long image sequence is split into small temporal windows that half overlap in time.
- Data assimilation in the full model is applied on the first window to retrieve the motion field.

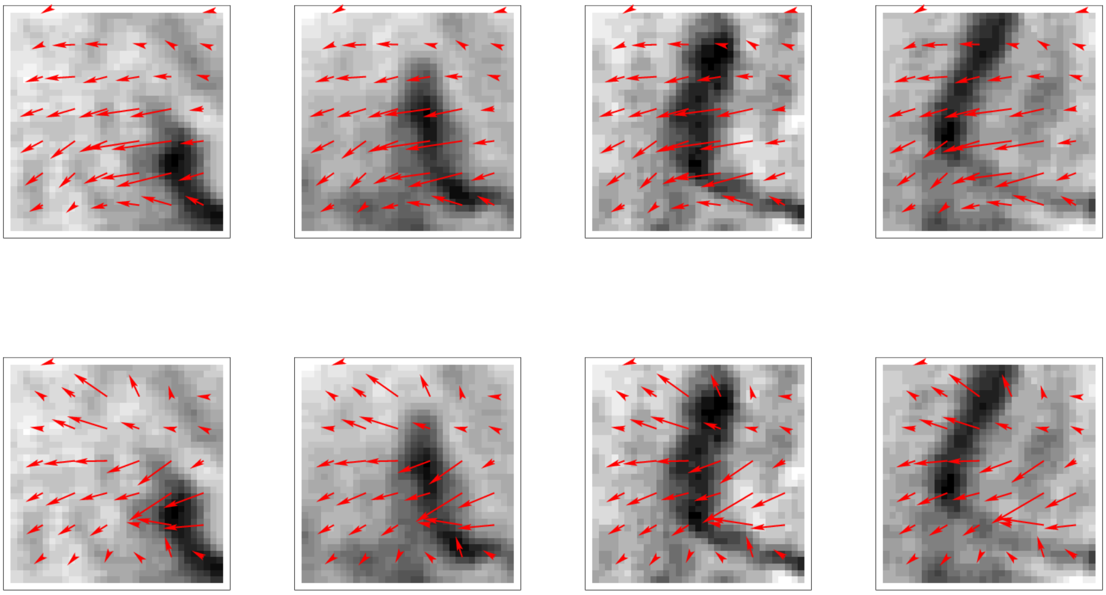

Figure 5. Top: Non-linear observation equation. Bottom: linear.

- The estimate of motion field at the beginning of the second window makes it possible to define the subspace for motion and a reduced model is obtained by Galerkin projection.
- Data assimilation in the reduced model is applied for this second window.
- The process is then iterated for the next window until the end of the whole image sequence.

Experiments were designed to quantify the results of this sliding windows method with base obtained by Principal Orthogonal Decomposition or computed as bi-sine functions [29].

6.6.5. Tracking of structures from an image sequence

Participants: Yann Lepoittevin, Isabelle Herlin, Dominique Béréziat [UPMC].

The research concerns an approach to estimate velocity on an image sequence and simultaneously segment and track a given structure. It relies on the underlying dynamics' equations of the studied physical system. A data assimilation method is designed to solve evolution equations of image brightness, those of motion's dynamics, and those of distance map modeling the tracked structures. The method is applied on meteorological satellite data, in order to track tropical clouds on image sequences and estimate their motion.

Part of research was concerned on the numerical schemes applied for advecting the distance map and designing its adjoint.

6.7. Minimax filtering

In minimax filtering for state estimation, the initial state error, the model error and the observational errors are classically supposed to belong to one joint ellipsoid. In this case, it is only assumed that the errors, stochastic or deterministic, are bounded. For each assimilation experiment, the filter computes an ellipsoid where one will find at least all states compatible with observations and errors description. The state estimate is taken as the center of the ellipsoid. No assumption on the actual distribution of the errors in needed and the state estimate minimizes the worst-case error, which makes the filter robust.

6.7.1. A posteriori minimax motion estimation

Participants: Sergiy Zhuk [IBM Research, Ireland], Isabelle Herlin, Olexander Nakonechnyi [Taras Shevchenko National University of Kyiv], Jason Frank [CWI, the Netherlands].

Data assimilation algorithms based on the 4D-Var formulation look for the so-called conditional mode estimate. The latter maximizes the conditional probability density function, provided the initial condition, model error and observation noise are realizations of independent Gaussian random variables. However this Gaussian assumption is often not satisfied for geophysical flows. Moreover, the estimation error of the conditional mode estimate is not a first-hand result of these methods. The issues above can be addressed by means of the Minimax State Estimation (MSE) approach. It allows to filter out any random (with bounded correlation operator) or deterministic (with bounded energy) noise and assess the worst-case estimation error.

The iterative MSE algorithm was developed for the problem of optical flow estimation from a sequence of 2D images. The main idea of the algorithm is to use the "bi-linear" structure of the Navier-Stokes equations and optical flow constraint in order to iteratively estimate the velocity. The algorithm consists of the following parts:

1) we construct pseudo-observations $\widehat{I}$ as the estimate of the image brightness function $I(x, y, t)$ solving the optical flow constraint such that $\widehat{I}$ fits (in the sense of least-squares) the observed sequence of images. To do so, we set the velocity field in the optical flow constraint to be the current minimax estimate of the velocity field $\mathbf{w}$, obtained at the previous iteration of the algorithm, and construct the minimax estimate $\widehat{I}$ of the solution of the resulting linear advection equation using the observed image sequence as discrete measurements of the brightness function;

2) we plug the estimate of the image gradient, obtained out of pseudo-observations $\widehat{I}$ in 1), into the optical flow constraint and the current minimax estimate $\mathbf{w}$ of the velocity field into the non linear part of Navier-Stokes equations so that we end up with a system of linear PDEs, which represents an extended state equation: it contains a linear parabolic equation for the velocity field and linear advection equation for the image brightness function; we construct the minimax estimate of the velocity field from the extended state equation using again the observed image sequence as discrete measurements of the brightness function;

3) we use the minimax estimate of the velocity field obtained in 2) in order to start 1) again.

Point 1) has been implemented and tested. As Point 2) is currently under development, it is replaced by one of our motion estimation method in order to be plugged in Point 3).

6.8. Fire application

6.8.1. Model evaluation for fire propagation

Participants: Vivien Mallet, Jean-Baptiste Fillipi [CNRS], Bahaa Nader [University of Corsica].

In the field of forest fires risk management, important challenges exist in terms of people and goods preservation. Answering to strong needs from different actors (firefighters, foresters), researchers focus their efforts to develop operational decision support system tools that may forecast wildfire behavior. This requires the evaluation of models performance, but currently, simulation errors are not sufficiently qualified and quantified. As the main objective is to realize a *decision support system*, it is required to establish robust forecast evaluations. In the context of the ANR project IDEA, the evaluation of model simulations has led to the definition and implementation of a series of forecast scores. The merits and shortcomings of the scores were evaluated on synthetic cases. This demonstrated the efficiency of scores that take into account the fire dynamics, where some classical scores may fail. This was also found on real fires, using field observations.

In addition, we consider that the proper evaluation of a model requires to apply it to a large number of fires – instead of carrying out a fine tuning on just one fire. We implemented a software to simulate a large number of fires (from the Prométhée database, http://www.promethee.com/) with the simulation model ForeFire (CNRS/University of Corsica) and evaluate the results with error measures. One simulation requires mainly the following data: the ignition point, the ground elevation, the vegetation cover and the wind field. See illustration on Fig. 6.

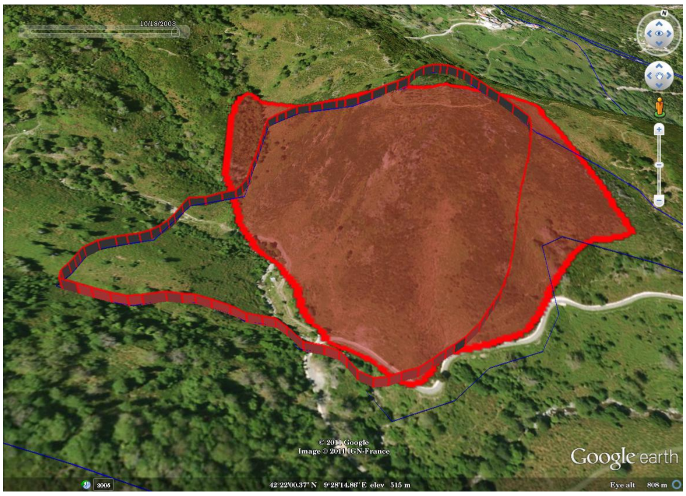

Figure 6. Fire simulation (using ForeFire) in red elevated contour, and observation (from Prométhée) of the burned area in filled red contour, for a 2003 fire near San-Giovanni-di-Moriani (Corsica).

7. Bilateral Contracts and Grants with Industry

7.1. Bilateral Contracts with Industry

- Clime is partner with INERIS (National Institute for Environmental and Industrial Risks) in a joint cooperation devoted to air quality forecast. This includes research topics in uncertainty estimation, data assimilation and ensemble modeling.

  Clime also provides support to INERIS in order to operate the Polyphemus system for ensemble forecasting, uncertainty estimations and operational data assimilation at continental scale.
- Clime is partner with IRSN, the French national institute for radioprotection and nuclear safety, for inverse modeling of emission sources and uncertainty estimation of dispersion simulations. The collaboration aims at better estimating emission sources, at improving operational forecasts for crisis situations and at estimating the reliability of forecasts. The work is derived at large scale (continental scale) and small scale (a few kilometers around a nuclear power plant).
- Clime takes part to a joint Ilab with the group SETH (Numtech). The objective is to (1) transfer Clime work in data assimilation, ensemble forecasting and uncertainty estimation, with application to urban air quality, (2) identify the specific problems encountered at urban scale in order to determine new research directions.

8. Partnerships and Cooperations

8.1. Regional Initiatives

- Clime is involved in the project PREQUALIF–IZNOGOUD–BARC, with many partners including the leading partner LSCE ("Laboratoire des Sciences du Climat et l'Environnement"), which aims at designing methods for the evaluation of the measures to be taken in the ZAPA areas ("Priority Areas for Air Quality Measures"). Clime focuses on the assimilation of observations to better evaluate the actual air quality.

8.2. National Initiatives

8.2.1. ANR

- Clime is one partner of the ANR project GeoFluids. It focuses on the specification of tools to analyse geophysical fluid flows from image sequences. Clime objectives concern the definition of reduced models from image data.
- Clime takes part to the ANR project IDEA that addresses the propagation of wildland fires. Clime is in charge of the estimation of the uncertainties, based on sensitivity studies and ensemble simulations.
- The MSDAG project (Multiscale Data Assimilation in Geophysics) is an ANR project. Fours partners are in the project: CEREA (Clime project-team, Marc Bocquet, PI of the whole project), Fluminance (Étienne Mémin), Moise Project-team (Laurent Debreu), LSCE (Frédéric Chevallier). It has ended the 30th of September 2012.

8.2.2. INSU

- Clime is running the project MIDAR “Inverse modeling of deposition measurements in case of a radiological release”, under the framework of the LEFE-ASSIM program of INSU. This includes a cooperation with the Institute for Safety Problems of Nuclear Power Plants (National Academy of Sciences of Ukraine). This project has ended in summer 2012.
- Clime is part of the INSU/LEFE project ADOMOCA-2, with about ten French teams working in atmospheric chemistry data assimilation.

8.3. European Initiatives

8.3.1. Collaborations in European Programs, except FP7

Program: COST Action ES104.

Project acronym: EuMetChem.

Project title: European framework for online integrated air quality and meteorology modeling.

Duration: January 2011 - December 2014.

Coordinator: Alexander Baklanov, Danish Meteorological Institute (DMI) Danemark.

Other partners: around 14 european laboratories, experts from United States, ECMWF.

Abstract: European framework for online integrated air quality and meteorology modeling (EuMetChem) will focus on a new generation of online integrated Atmospheric Chemical Transport (ACT) and Meteorology (Numerical Weather Prediction and Climate) modeling with two-way interactions between different atmospheric processes including chemistry (both gases and aerosols), clouds, radiation, boundary layer, emissions, meteorology and climate. At least, two application areas of the integrated modeling are planned to be considered: (i) improved numerical weather prediction (NWP) and chemical weather forecasting (CWF) with short-term feedbacks of aerosols and chemistry on meteorological variables, and (ii) two-way interactions between atmospheric pollution/ composition and climate variability/change. The framework will consist of four working groups namely: 1) Strategy and framework for online integrated modeling; 2) Interactions, parameterizations and feedback mechanisms; 3) Chemical data assimilation in integrated models; and finally 4) Evaluation, validation, and applications. Establishment of such a European framework (involving also key American experts) will enable the EU to develop world class capabilities in integrated ACT/NWP-Climate modeling systems, including research, forecasting and education.

8.3.2. Collaborations with Major European Organizations

Partner: ERCIM working group “Environmental Modeling”.

The working group gathers laboratories working on developing models, processing environmental data or data assimilation. In 2012, the working group organized sessions during IEMSs conference in Leipzig, Germany.

8.4. International Initiatives

8.4.1. Inria International Partners

Partner: Chilean meteorological office (Dirección Meteorológica de Chile)

The partner produces its operational air quality forecasts with Polyphemus. The 3-day forecasts essentially cover Santiago. The forecasts are accessible online in the form of maps, time series and video (http://www.meteochile.cl/modelos.html).

Partner: Marine Hydrophysical Institute, Ukraine.

The collaboration concerns the study of the Black Sea surface circulation and the issue of image assimilation in forecasting models.

Partner: Institute of Numerical Mathematics, Russia.

The collaboration concerns the estimation of uncertainty of the motion field derived from image data with data assimilation technics.

8.4.2. Participation In International Programs

- Clime is running a two-year project under the PHC-DNIPRO program with Taras Shevchenko University of Kyiv, Ukraine. The subject concerns a posteriori minimax motion estimation from images.

8.5. International Research Visitors

8.5.1. Visits of International Scientists

- Sergii Demydenko, Taras Shevchenko University of Kyiv, Ukraine,July 2012.
- Andrii Filipenkov, Taras Shevchenko University of Kyiv, December 2012.
- Takemasa Myoshi, University of Maryland, USA, June 2012.
- Oleksandr Nakonechnyi, Taras Shevchenko University of Kyiv, December 2012.
- Sergiy Zhuk, IBM, Dublin Research Lab, Ireland, December 2012.

8.5.2. Visits to International Teams

- Vivien Mallet took part in June to a HARVEST project, funded by Pascal2. He visited the Department of Statistical Science at University College London. The project dealt with uncertainty quantification using statistical emulation of geophysical models, mainly for climate modeling.

9. Dissemination

9.1. Scientific Animation

- Marc Bocquet is co-chair of the INSU/LEFE MANU scientific commitee.
- Marc Bocquet is a member of the Scientific Council of the CERFACS institute in Toulouse, France.
- Marc Bocquet is Associate Editor of the Quaterly Journal of the Royal Meteorological Society.
- Marc Bocquet has co-organised the Les Houches international summer school “Advanced data assimialtion for geosciences”, Les Houches, 28 May 2012 - 15 June 2012.
- Marc Bocquet has co-organised the 4th “Colloque national d’assimilation de données”, Nice, 17-19 December 2012.

- Isabelle Herlin is member of the Scientific Council of CSFRS (High Council for Strategic Education and Research in France).
- Isabelle Herlin organised sessions on "Analysis of data of different scales and sources for mesoscale environmental models" for the International Congress on Environmental Modeling and Software (IEMSs2012).
- Isabelle Herlin is a member of Evaluation Committee at Inria.
- Isabelle Herlin is a member of the Scientific Council of OSU-EFLUVE.

9.2. Teaching - Supervision - Juries

9.2.1. Teaching

Master OACOS: Marc Bocquet, Vivien Mallet; Introduction to Data Assimilation for Geophysics; 30 hours; M2; UPMC, X, ENS, ENSTA ParisTech, École des Ponts ParisTech; France.

Master "Nuclear Energy": Marc Bocquet, Vivien Mallet, Victor Winiarek; 12 hours; M2; École des Ponts ParisTech; France.

Master SGE and École des Ponts ParisTech: Vivien Mallet; Air Pollution; 6 hours; M2; École des Ponts ParisTech , Paris 7-Diderot, Paris Est; France.

9.2.2. Supervision

PhD : Mohammad Koohkan, “Modélisation inverse et assimilation de données en qualité de l’air”, University Paris Est, December 20th, 2012, Marc Bocquet.

PhD in progress : Paul Baudin, “Agrégation séquentielle de prédicteurs appliquée à la prévision de la qualité de l’air”, September 2012, Vivien Mallet and Gilles Stoltz.

PhD in progress : Karim Drifi, “Reduced models for image assimilation”, University Paris Centre, October 2009, Isabelle Herlin.

PhD in progress : Yann Lepoittevin, “Tracking of image structures”, University Paris Centre, October 2012, Isabelle Herlin.

Yiguo Wang, "Lidar data assimilation”, October 2009, Karine Sartelet, Patrick Chazette, Marc Bocquet.

PhD in progress : Victor Winiarek, “Dispersion atmosphérique en milieu urbain et modélisation inverse pour la reconstruction de sources”, University Paris Est, October 2009, Marc Bocquet.

9.2.3. Juries

- Bocquet, M., HdR, Thibaut Montmerle, “Assimilation des données à moyenne échelle pour l’étude des systèmes précipitants”, 27 January 2012, Toulouse, France.
- Bocquet, M. HdR, Olivier Pannekoucke, “Dynamique et modélisation de l’information dans les modèles météorologiques”, 15 November 2012, Toulouse, France.

9.3. Popularization

- Vivien Mallet introduced mathematical modeling for environmental sciences at the award ceremony for Paris mathematical olympiads.
- Anne Tilloy contributed to La Fête de la Science, which is a national initiative for science popularization to which Inria Paris-Rocquencourt took part

10. Bibliography

Major publications by the team in recent years

[1] M. BOCQUET, P. SAKOV. *Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems*, in "Nonlinear Processes in Geophysics", June 2012, vol. 19, n^{o} 3, p. 383-399 [*DOI :* 10.5194/NPG-19-383-2012], http://hal.inria.fr/hal-00714384, http://hal.inria.fr/hal-00714384.

[2] D. BÉRÉZIAT, I. HERLIN. *Solving ill-posed Image Processing problems using Data Assimilation*, in "Numerical Algorithms", February 2011, vol. 56, n^{o} 2, p. 219-252 [*DOI :* 10.1007/S11075-010-9383-Z], http://hal.inria.fr/inria-00538510, http://hal.inria.fr/inria-00538510.

[3] D. GARAUD, V. MALLET. *Automatic calibration of an ensemble for uncertainty estimation and probabilistic forecast: Application to air quality*, in "Journal of Geophysical Research", October 2011, vol. 116, n^{o} D19304, http://hal.inria.fr/hal-00655771/en.

[4] I. HERLIN, D. BÉRÉZIAT, N. MERCIER, S. ZHUK. *Divergence-free motion estimation*, in "Proceedings of European Conference on Computer Vision (ECCV)", Firenze, Italy, Lecture Notes in Computer Science, October 2012, vol. 7575, p. 15-27 [*DOI :* 10.1007/978-3-642-33765-9_2], http://hal.inria.fr/hal-00742021, http://hal.inria.fr/hal-00742021.

[5] M. R. KOOHKAN, M. BOCQUET. *Accounting for representativeness errors in the inversion of atmospheric constituent emissions: Application to the retrieval of regional carbon monoxide fluxes*, in "Tellus B", July 2012, vol. 64, n^{o} 19047 [*DOI :* 10.3402/TELLUSB.V64I0.19047], http://hal.inria.fr/hal-00741930, http://hal.inria.fr/hal-00741930.

[6] G. K. KOROTAEV, E. HUOT, F.-X. LE DIMET, I. HERLIN, S. V. STANICHNY, D. M. SOLOVYEV, L. WU. *Retrieving ocean surface current by 4-D variational assimilation of sea surface temperature images*, in "Remote Sensing of Environment", April 2008, vol. 112, n^{o} 4, p. 1464-1475, Special issue on data assimilation [*DOI :* 10.1016/J.RSE.2007.04.020], http://hal.archives-ouvertes.fr/hal-00283896, http://hal.archives-ouvertes.fr/hal-00283896.

[7] V. MALLET. *Ensemble forecast of analyses: Coupling data assimilation and sequential aggregation*, in "Journal of Geophysical Research", December 2010, vol. 115, n^{o} D24303, http://hal.inria.fr/inria-00547903.

[8] B. SPORTISSE. *Pollution atmosphérique: des processus á la modélisation*, Springer-Verlag France, 2008, 350 pages, ISBN : 978-2-287-74961-2, http://hal.inria.fr/inria-00581172/en.

[9] V. WINIAREK, M. BOCQUET, O. SAUNIER, A. MATHIEU. *Estimation of Errors in the Inverse Modeling of Accidental Release of Atmospheric Pollutant: Application to the Reconstruction of the Cesium-137 and Iodine-131 Source Terms from the Fukushima Daiichi Power Plant*, in "Journal of Geophysical Research Atmospheres", March 2012, vol. 117, n^{o} D05122 [*DOI :* 10.1029/2011JD016932], http://hal.inria.fr/hal-00704999, http://hal.inria.fr/hal-00704999.

[10] L. WU, V. MALLET, M. BOCQUET, B. SPORTISSE. *A Comparison Study of Data Assimilation Algorithms for Ozone Forecasts*, in "Journal of Geophysical Research", October 2008, vol. 113, n^{o} D20310, PP. 17, http://hal.inria.fr/inria-00582376/en.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] M. R. KOOHKAN. *Modélisation inverse et assimilation de données en qualité de l'air*, Université Paris-Est Marne-la-Vallée, December 2012.

Articles in International Peer-Reviewed Journals

[12] M. Bocquet. *Parameter-field estimation for atmospheric dispersion: application to the Chernobyl accident using 4D-Var*, in "Quarterly Journal of the Royal Meteorological Society", April 2012, vol. 138, n^{O} 664, p. 664-681 [*DOI :* 10.1002/qj.961], http://hal.inria.fr/hal-00705045.

[13] M. Bocquet, P. Sakov. *Combining inflation-free and iterative ensemble Kalman filters for strongly nonlinear systems*, in "Nonlinear Processes in Geophysics", June 2012, vol. 19, n^{O} 3, p. 383-399 [*DOI :* 10.5194/npg-19-383-2012], http://hal.inria.fr/hal-00714384.

[14] P. Chazette, M. Bocquet, P. Royer, V. Winiarek, J.-C. Raut, P. Labazuy, M. Gouhier, M. Lardier, J.-P. Cariou. *Eyjafjallajökull ash concentrations derived from both lidar and modeling*, in "Journal of Geophysical Research", 2012, vol. 117, D00U14 [*DOI :* 10.1029/2011JD015755], http://hal.inria.fr/hal-00643869.

[15] F. Chevallier, T. Wang, P. Ciais, F. Maignan, M. Bocquet, M. A. Arain, A. Cescatti, J. Chen, A. J. Dolman, B. E. Law, H. A. Margolis, L. Montagnani, E. J. Moors. *What eddy-covariance measurements tell us about prior land flux errors in CO2-flux inversion schemes*, in "Global Biogeochemical Cycles", March 2012, vol. 26, n^{O} GB1021 [*DOI :* 10.1029/2010GB003974], http://hal.inria.fr/hal-00705026.

[16] C. Estournel, E. Bosc, M. Bocquet, C. Ulses, P. Marsaleix, V. Winiarek, I. Osvath, C. Nguyen, T. Duhaut, F. Lyard, H. Michaud, F. Auclair. *Assessment of the amount of Cesium-137 released into the Pacific Ocean after the Fukushima accident and analysis of its dispersion in Japanese coastal waters*, in "Journal Of Geophysical Research Oceans", November 2012, vol. 117, n^{O} C11014 [*DOI :* 10.1029/2012JC007933], http://hal.inria.fr/hal-00761366.

[17] M. R. Koohkan, M. Bocquet. *Accounting for representativeness errors in the inversion of atmospheric constituent emissions: application to the retrieval of regional carbon monoxide fluxes*, in "Tellus B", July 2012, vol. 64, n^{O} 19047 [*DOI :* 10.3402/tellusb.v64i0.19047], http://hal.inria.fr/hal-00741930.

[18] M. R. Koohkan, M. Bocquet, L. Wu, M. Krysta. *Potential of the International Monitoring System radionuclide network for inverse modelling*, in "Atmospheric Environment", July 2012, vol. 54, p. 557-567 [*DOI :* 10.1016/j.atmosenv.2012.02.044], http://hal.inria.fr/hal-00704981.

[19] T. Lauvaux, A. E. Schuh, M. Bocquet, L. Wu, S. Richardson, N. Miles, K. J. Davis. *Network design for mesoscale inversions of CO2 sources and sinks*, in "Tellus B", June 2012, vol. 64 [*DOI :* 10.3402/tellusb.v64i0.17980], http://hal.inria.fr/hal-00714386.

[20] V. Winiarek, M. Bocquet, O. Saunier, A. Mathieu. *Correction to "Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the cesium-137 and iodine-131 source terms from the Fukushima Daiichi power plant"*, in "Journal of Geophysical Research Atmospheres", September 2012, vol. 117, n^{O} D18118 [*DOI :* 10.1029/2012JD018107], http://hal.inria.fr/hal-00761575.

[21] V. Winiarek, M. Bocquet, O. Saunier, A. Mathieu. *Estimation of errors in the inverse modeling of accidental release of atmospheric pollutant: Application to the reconstruction of the cesium-137 and iodine-131 source terms from the Fukushima Daiichi power plant*, in "Journal of Geophysical Research Atmospheres", March 2012, vol. 117, n^{O} D05122 [*DOI :* 10.1029/2011JD016932], http://hal.inria.fr/hal-00704999.

[22] L. Wu, F.-X. Le Dimet, P. De Reffye, B.-G. Hu, P.-H. Cournède, M.-Z. Kang. *An optimal control methodology for plant growth–Case study of a water supply problem of sunflower*, in "Mathematics and Computers in Simulation", January 2012, vol. 82, n^o 5, p. 909-923 [*DOI :* 10.1016/j.matcom.2011.12.007], http://hal.inria.fr/hal-00705066.

[23] Y. Zhang, M. Bocquet, V. Mallet, C. Seigneur, A. Baklanov. *Real-time air quality forecasting, part I: History, techniques, and current status*, in "Atmospheric Environment", December 2012, vol. 60, p. 632-655 [*DOI :* 10.1016/j.atmosenv.2012.06.031], http://hal.inria.fr/hal-00761344.

[24] Y. Zhang, M. Bocquet, V. Mallet, C. Seigneur, A. Baklanov. *Real-Time Air Quality Forecasting, Part II: State of the Science, Current Research Needs, and Future Prospects*, in "Atmospheric Environment", December 2012, vol. 60, p. 656-676 [*DOI :* 10.1016/j.atmosenv.2012.02.041], http://hal.inria.fr/hal-00761347.

International Conferences with Proceedings

[25] D. Béréziat, I. Herlin. *Non Linear Observation Equation For Motion Estimation*, in "International Conference Inverse Problems: Identification, Design and Control (ICIP)", Orlando, United States, October 2012, http://hal.inria.fr/hal-00761332.

[26] I. Herlin, D. Béréziat, E. Huot. *Image Assimilation and Motion Estimation of Geophysical Fluids*, in "International Environmental Modelling and Software Society (iEMSs)", Leipzig, Germany, July 2012, http://hal.inria.fr/hal-00741977.

[27] I. Herlin, D. Béréziat, N. Mercier. *Improvement of motion estimation by assessing the errors on the evolution equation*, in "VISAPP 2012 : International Conference on Computer Vision Theory and Applications", Rome, Italy, February 2012, vol. 2, p. 235-240, http://hal.inria.fr/hal-00677662.

[28] I. Herlin, D. Béréziat, N. Mercier, S. Zhuk. *Divergence-Free Motion Estimation*, in "ECCV 2012 - European Conference on Computer Vision", Florence, Italy, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, C. Schmid (editors), Lecture Notes in Computer Science, Springer, October 2012, vol. 7575, p. 15-27 [*DOI :* 10.1007/978-3-642-33765-9_2], http://hal.inria.fr/hal-00742021.

[29] I. Herlin, K. Drifi. *Learning reduced models for motion estimation on long temporal image sequences*, in "IEEE International Geoscience and Remote Sensing Symposium (IGARSS)", Munich, Germany, 2012, http://hal.inria.fr/hal-00730515.

[30] I. Herlin, E. Huot. *Monitoring surface currents from uncertain image observations*, in "XX International Conference on Problems of Decision Making under Uncertainties (PDMU)", Brno, Czech Republic, September 2012, http://hal.inria.fr/hal-00739091.

[31] I. Herlin, O. Nakonechnyi, S. Zhuk. *Minimax optical flow estimation from a sequence of 2D images*, in "XX International Conference on Problems of Decision Making under Uncertainties (PDMU)", Brno, Czech Republic, September 2012, http://hal.inria.fr/hal-00739089.

Research Reports

[32] D. Garaud, V. Mallet. *Uncertainty Estimation and Decomposition based on Monte Carlo and Multimodel Photochemical Simulations*, Inria, March 2012, n^o RR-7903, 33, http://hal.inria.fr/hal-00678306.