INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Team clime

Couplage de la donnée environnementale et des modèles de simulation numérique pour une intégration logicielle

Paris - Rocquencourt

THEME NUM

Activity Report

2008

Table of contents

1. Team

Research Scientist

Isabelle Herlin [Team Leader, Research Director, INRIA]
Marc Bocquet [Deputy Director of CEREA, Ecole des Ponts ParisTech, HdR]
Jean-Paul Berroir [Research scientist, CR, INRIA]
Vivien Mallet [Research scientist, CR, INRIA]

Faculty Member

Dominique Béréziat [MC, UPMC, until August 2008]
Etienne Huot [MC, University of Versailles]

Technical Staff

Meryem Ahmed de Biasi [Technical Staff, INRIA, until June 2008]
Claire Mouton [Technical Staff, INRIA, from October 2008]
Nicolas Mercier [Technical Staff, INRIA, from November 2008]
Pierre Tran [Technical Staff, Ecole des Ponts ParisTech]

PhD Student

Ahmed Abdel Meguid [PhD Student, INRIA, from November 2008]
Rachid Abida [PhD Student, Ecole des Ponts ParisTech]
Damien Garaud [PhD Student, EdF R&D, Ecole des Ponts ParisTech]
Irène Korsakissok [PhD Student, Ecole des Ponts ParisTech]
Marilyne Tombette [PhD Student, Ecole des Ponts ParisTech, until March 2008]

Post-Doctoral Fellow

Xiaoni Faivre-Wang [Post-Doctoral Fellow, INRIA]
Lin Wu [Post-Doctoral Fellow, Ecole des Ponts ParisTech]

Visiting Scientist

Maithili Sharan [Indian Institute of Technology, New Dehli, India, from June to July 2008]

Administrative Assistant

Christine Anocq [Administrative Assistant, INRIA, until March 2008]
Véronique Dehlinger [Administrative Assistant, Ecole des Ponts ParisTech]
Nathalie Gaudechoux [Administrative Assistant, INRIA, from March 2008]

Other

Carlos Dantas [Internship, INRIA, from January to June 2008]
Sébastien Gerchinovitz [Internship, CNRS, from May 2008 to September 2008]
Nicolas Mercier [Internship, INRIA, until October 2008]

2. Overall Objectives

2.1. Overall Objectives

The international political and scientific context is indicating the serious potential risks related to environmental problems and is pointing out the role that can be played by models and observation systems for the evaluation and forecasting of these risks. At the political level, agreements, such as the Kyoto protocol, European directives on air quality or on major accident hazards involving dangerous substances (Seveso directive) and the French Grenelle de l'Environnement establish objectives for the mitigation of environmental risks. These objectives are supported at a scientific level by international initiatives, like the European GMES program (Global Monitoring of Environment and Security), or national programs, such as the Air Chemistry program, which give a long term structure to environmental research. These initiatives emphasize the importance of observational data and also the potential of satellite acquisitions.

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

This situation is met for instance in atmospheric pollution, an environmental domain 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, modeling systems are used for operational forecasts (short or long term), detailed case studies, impact studies for industrial sites, as well as coupled modeling (e.g., pollution and health, pollution and economy). These scientific subjects strongly require linking 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 area:

- Environmental data processing: solving ill-posed image processing problems by assimilating image data within Image Evolution Models ; assimilation of image structures (characterized by a strong spatio-temporal coherency) within environmental forecasting models.
- Data and model coupling, by means of data assimilation techniques and related issues (optimization problems, targetting observation, uncertainties propagation, ...).
- Development of integrated chains for data/models/outputs (system architecture, workflow, database, visualization, ...).

3. Scientific Foundations

3.1. Environmental image and data processing

Keywords: *fluid flow*, *image assimilation*, *image processing*.

Data with image nature, and especially satellite data, represent a huge amount of observations which is up to now largely unexploited by the environmental numerical forecast models. The operational state-of-the-art is mainly the assimilation of satellite data on a pixel basis: each pixel constitutes an independent information, expressed as a more or less simple function of the model's state variables. The objective is to exploit the structure of the image observation by defining Image Assimilation methodologies: how to assimilate data with spatial and temporal coherency, such as observations of evolving fronts or eddies? Three alternative approaches are considered:

- Extension of image processing techniques by accounting for all available physical information, either on the observed phenomena (evolution law) or on the process of image formation (radiative transfer), to derive estimates of the forecasting model's state variables, that can be directly assimilated. A typical example is the assessment of apparent motion from image observations of turbulent fluids, the approach taking advantage of the fluid mass conservation principle as well as numerical regularization techniques (div-curl) preserving the vorticity of turbulent motion.
- Definition of Image Evolution Models to solve ill-posed image processing problems, usually addressed using numerical regularization techniques. The Image Model describes the dynamics of the image sequence and makes it possible to formulate a data assimilation problem, where image observations are assimilated within the Image Model. This approach constitutes a relevant way to solve image processing problems, in which difficult issues such as occlusions or missing data are easily formulated. The usual spatial regularization is replaced by the temporal evolution laws for solving the underdetermination issue.

- Definition of Physical Image Models coupling variables from the image domain and from the forecasting model (in the same spirit than the qualitative Conceptual Models developed by meteorologists to describe specific meteorological phenomena and their signature on image data). The assimilation is then performed in two steps: first, in the Physical Image Model to yield “bogus” observations of the forecasting model’s state variables, then directly in the forecasting model.

3.2. Data assimilation and inverse modeling

Keywords: *data assimilation*, *ensemble forecast*, *inverse modeling*, *network design*.

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

Although modeling is not part of the scientific objectives of Clime, we have complete access to models developed by CEREA (joint ENPC/EDF R&D laboratory): Polair3D (photochemical pollution forecasting at continental and regional scales) and Mercure-Saturne (urban scale) for air quality problems. In regard to other modeling domains, Clime accesses models through co-operation initiatives either directly (for instance the shallow water model developed at MHI, Ukrain, has been provided to the team), or indirectly (for instance, image assimilation methodologies for meteorology are tested by partners on their models).

The research activities tackle scientific issues such as:

- 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 the trajectories of mobile sensors be operated, while the studied phenomenon is evolving in time? This issue is usually referred as “network design”.
- How to assess the quality of the prediction? 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?
- Within a family of models (differing by their physical approximations or their discretization parameters) which is the optimal model for a given set of observations?
- 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?

3.3. Software chains for environmental applications

Keywords: *database*, *system architecture*, *visualization*, *workflow*.

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 CEREA and EDF R&D, such a system for air pollution modeling: Polyphemus (see web site http://cerea.enpc.fr/polyphemus/), whose architecture is specified to satisfy the 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. Panorama

The central application of the project-team is atmospheric chemistry, to which the majority of resources are allocated. We develop and maintain the air quality modeling system Polyphemus, which includes several numerical models (Gaussian, local scale, the 3D Eulerian model Polair3D) and their adjoints, and different high level methods: ensemble forecast, sequential and variational data assimilation algorithms. Advanced data assimilation, 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).

An important activity on assimilation of satellite data is also carried out for meteorology and oceanography. This is addressed in cooperation with external partners who provide the numerical models. Concerning meteorology, the focus is on the forecast of high-impact weather events (cyclones, convective storms, *etc*.). Concerning oceanography, the aim is to improve the forecast of ocean circulation, in relation with global warming issues. The underlying researches concern the assimilation of structures observed in satellite images.

Finally, Clime carries out applied remote sensing activities in close cooperation with environmental experts. These studies concern the potential use of satellite data for monitoring agriculture and soil degradation. An image model of the observed phenomenon is first established, then tracking and change detection techniques are applied in the model space.

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 key point for the improvement of forecast's quality. These techniques, and notably the variational ones, are progressively surfacing in atmospheric chemistry.

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

- 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.
- 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?
- 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 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. Meteorology

Meteorological forecasting constitutes a major application domain for Image Assimilation. Although satellite data are operationally assimilated within models, this is done on an independent pixel basis: the observed radiance is linked to the state variable 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 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 initialisation. The challenge is therefore to develop a methodological framework allowing the assimilation of the detected CMs within numerical forecast models, a very important issue considering the considerable impact of the related meteorological events.

4.4. Oceanography

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

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

Image Models for these structures are developed taking into account the underlying physical processes. Image data are assimilated within the Image Models to derive pseudo-observations of the state variables which are further assimilated within the numerical ocean forecast model.

4.5. Remote sensing for natural risks and agricultural monitoring

Clime is involved in several research actions related to environmental monitoring which are interrelated from a methodological point of view. These studies are based on the analysis of satellite image time-series at low (AVHRR, MODIS) or high (SPOT) spatial resolution for monitoring changes in the environment. The concern is monitoring land degradation.

5. Software

5.1. Polyphemus

Participants: Vivien Mallet, Irène Korsakissok, Meryem Ahmed de Biasi, Pierre Tran, Marilyne Tombette, Lin Wu.

Polyphemus (see 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 fields: 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, Mercure-Saturne and two Gaussian models);
- 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).

A workshop was organized in Paris on 27 and 28 October. The workshop was open to French core users only. Fourteen talks were given, addressing many different topics: simulations at different scales, data assimilation and ensemble forecast, development of physics, applications (e.g., impact of EDF power plants).

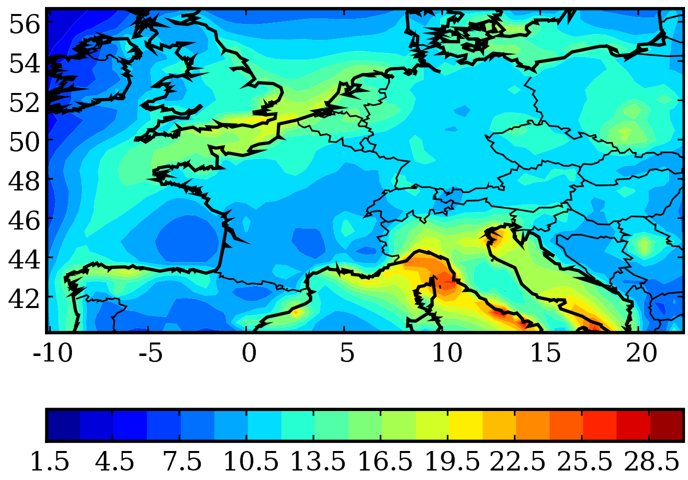

Figure 1. Map of the standard deviation (or spread) of an ensemble built with Polyphemus (ozone simulations, $\mu g\ m^{-3}$). The standard deviations are averaged over the summer of 2001. They provide an estimation of the simulation uncertainties.

5.2. ENVIAIR

Participants: Nicolas Mercier, Isabelle Herlin, Jean-Paul Berroir.

Enviair is an operational processing platform for analyzing multi-temporal data from the MODIS sensor and extracting land use information. It includes software libraries for:

- Manipulation of satellite images (I/O, management of acquisition dates and geographical coordinates, extraction of temporal data).

- Interpretation of multitemporal information (temporal filtering, computation of temporal features, classification programmes).
- Interpretation of MODIS metadata (pixel quality).

This software has been applied to deforestation monitoring in the high Taquari basin, Brazil (see figure 2). It is currently being registered commonly by INRIA and Embrapa (Brazilian agricultural research organization).

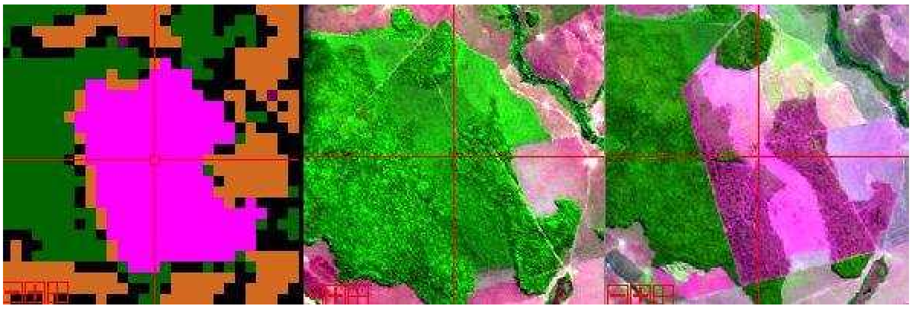

Figure 2. Left image: classification of a MODIS NDVI sequence with the ENVIAIR software (forest in green, pasture in orange, deforestation in pink). Middle and right images: control using two high resolution Landsat images, acquired before and after the MODIS sequence. The deforested parcel is clearly visible on the right image.

6. New Results

6.1. Solving Ill-posed Image Processing problems with missing acquisitions

Keywords: *data assimilation*, *ill-posed problem*, *image processing*, *spatial regularization*.

Participants: Dominique Béréziat, Isabelle Herlin.

In this study, the framework of Data Assimilation is used as a generic tool to solve ill-posed Image Processing problems: the image data or related parameters constitute observations which are assimilated within a dynamic model. A data assimilation system has then to be defined with the evolution equation of the state vector (the quantity to be estimated), the observation equation linking this state vector and the image data, and an initial condition. These three components include error terms, quantified as error covariance matrices.

A specific effort has been done concerning the theoritical properties of these matrices and their relations with image processing. We first proved that modeling the evolution model error with a Gaussian covariance is equivalent to perform a Tikhonov regularization of the state vector at any order. Second, a generic formulation of the covariance matrix associated to the observation error has been proposed to consider the issue of missing or low quality observations based on a confidence measure (potentially included in metadata associated to the acquisition). On pixels with low confidence, the error value becomes high and the observation will not be taken into account in the computation.

We applied this to the problem of estimating optical flow on a sequence of image data. In a first attempt simple dynamics assumptions have been considered: (i) the observation model is chosen as the transport of image brightness by velocity; (ii) the evolution model is written as the transport of velocity by itself. Experiments are conducted on image sequences displaying missing frames (the sensor has not acquired any data) or missing data (Fig 3). The resulting motion is then mainly obtained by the evolution model and is still correctly estimated on these pixels if the model accurately describes the dynamics of the state vector on the image sequence.

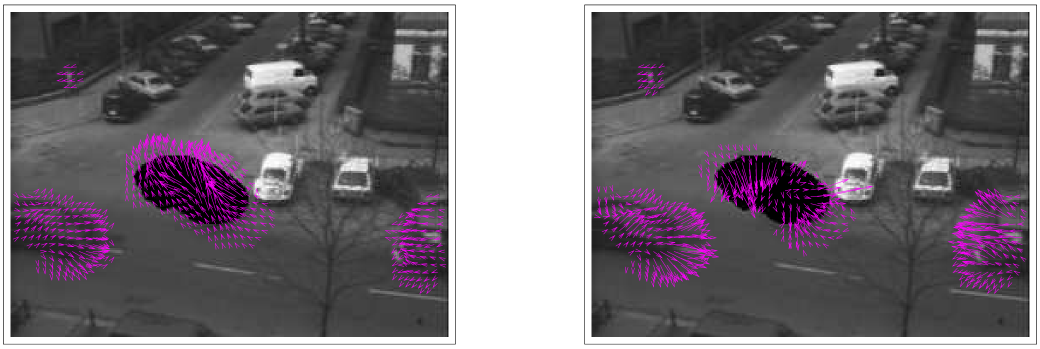

Figure 3. Handling missing data (black region in the middle) (Left: With data assimilation ; Right:Without assimilation).

Moreover, if the dynamics of the sequence is known and represented by the evolution equation, it has been proven that the Data Assimilation framework allows to better estimate the motion field on the sequence (Fig 4), because avoiding the usual regularisation which spatially smoothes the result.

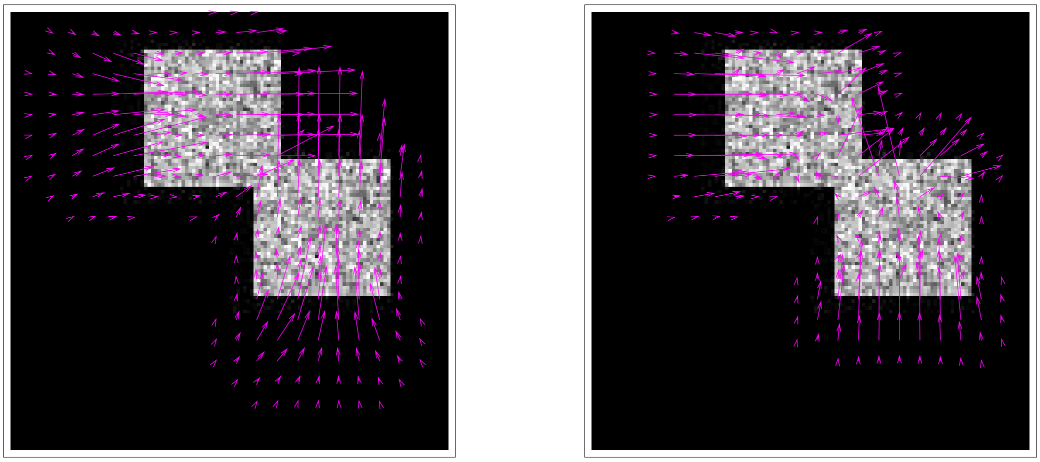

Figure 4. Use of temporal dynamics (Left: With data assimilation ; Right:Without assimilation).

6.2. 3D+t reconstruction from 2D+t images sequence using Data Assimilation

Keywords: *3D reconstruction*, *data assimilation*.

Participants: Dominique Béréziat, Carlos W. Dantas de Almeida.

The objective is to reconstruct a 3D sequence from (1) and initial 3D image and (2) a temporal sequence of 2D images, each one being the projection of the 3D object of interest. This is in principle similar to tomography. The aim is to reduce the number of 2D projections, required for the 3D reconstruction at each time step, by introducing dynamical models.

The problem is formulated as finding a 3D warping that makes it possible to reconstruct the 3D sequence from the initial condition and the temporal 2D sequence. The warping is such that the discrepancy between one 2D image and the projection of the corresponding 3D warped image is minimized. As the projection from the 3D to 2D domains is not invertible, this problem is ill-posed. We use the framework of Data Assimilation by formulating the dynamics of the 3D sequence. The observation are constituted by the 2D images; the evolution model is the transport of 3D velocity by itself.

The method has been tested on an ideal experiment in which the parameters of the projection are exactly known: artificial spherical cells are observed by an actual biological imaging system. The objective is the reconstruction in time of the 3D spheres. The first results show that it is possible to recover the 3D sequence.

6.3. Ocean surface velocity estimation using an Image Model

Keywords: *data assimilation*, *motion estimation*, *oceanography*, *remote sensing*.

Participants: Etienne Huot, Isabelle Herlin, Gennady Korotaev [Marine Hydrophysical Institute, Ukrain].

The objective of the study is to assess, from a sequence of images, a motion field which is coherent with the dynamics of the observed phenomenon. The application is the estimation of surface velocity from a sequence of oceanographic satellite images.

For that purpose, we define an *Image Evolution Model* describing the temporal evolution of image information and associated parameters. The state vector includes the measured temperature, the two components of velocity and the thickness of the surface layer. The evolution equations are: a simplified advection-diffusion scheme for temperature, the shallow-water equations for velocity and thickness. A data assimilation system is then defined by adding an observation equation describing the transport of temperature by the motion field. The initial conditions are chosen according to image processing softwares. Sea Surface Temperature data are then assimilated in the Image Model using a 4D-variational scheme for optimizing the initial conditions and deriving the flow field along the sequence. This field is further assimilatesd in an oceanic simulation model for validation as illustrated on Fig. 5.

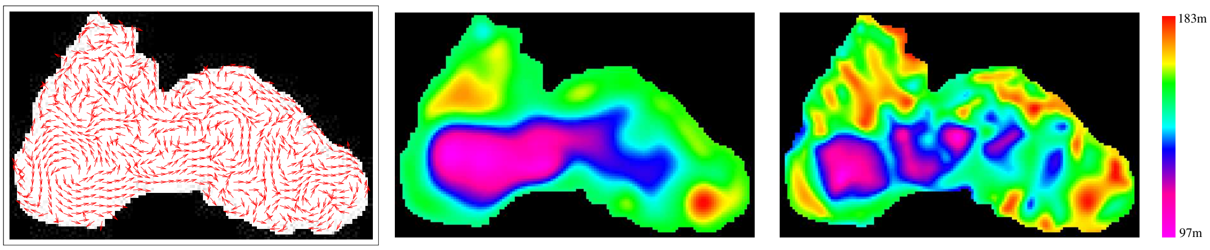

Figure 5. Exemple of velocity field estimation on the Black Sea surface (left) and the impact of its assimilation on the sea surface elevation: without assimilation (center), with assimilation (right).

6.4. Assimilation of aerosol ground observations

Keywords: *aerosol*, *air quality*, *chemistry-transport model*, *data assimilation*.

Participants: Vivien Mallet, Marilyne Tombette.

Data assimilation of PM_{10} ground observations was carried out over Europe and for January 2001. The numerical model was a usual configuration of the aerosol variant of Polair3D, part of the Polyphemus modeling system whose implementation of the optimal interpolation method was applied here. The method was notably applied in operational-like conditions.

The assimilation is beneficial to the forecast, but it has a limited impact in time: after a few hours only, the effect of assimilation faints (figure 6). Sensitivity tests investigated the spatial and temporal effects of data assimilation in order to better describe this limitation.

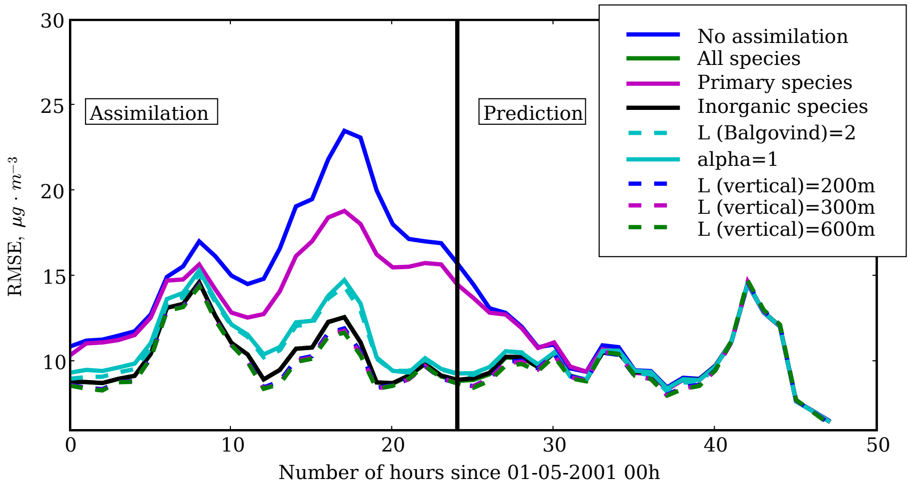

Figure 6. Time evolution of the RMSE for the PM_{10} *forecast. Each curve concerns one configuration of the assimilation procedure. The vertical line delimits the assimilation period from the prediction period. Whatever the configuration, the effect of assimilation disappears rapidly.*

6.5. Assimilation of NO_2 columns for air quality forecast

Keywords: *air quality*, *data assimilation*, *satellite observation*.

Participants: Xiaoni Wang, Jean-Paul Berroir, Vivien Mallet, Isabelle Herlin, Marc Bocquet.

This study concerns the data assimilation of satellite observations for improving the air quality forecast, performed by the Polyphemus air quality system.

Nitrogen dioxide (NO_2) plays an important role in the tropospheric chemistry and has a direct impact on the public health. A better knowledge and forecast of NO_2 concentration are important to all issues related to air quality. In this research work available satellite data are considered: the Ozone Monitoring Instrument (OMI), aboard NASA Aura satellite, provides NO_2 column data with a good spatial resolution (13 by 24 km^2) and daily global coverage.

First satellite data have been compared to Polyphemus simulations: the OMI column data and the Polyphemus simulations have both been averaged over November-December 2005 in Europe, demonstrating a good consistency in Spain, Italy and North Europe (see figure 7).

The satellite observations are then assimilated in Polyphemus. The forecast obtained with and without assimilation are compared with ground observations for validation. It is found that assimilation of these satellite data improve the NO_2 forecast, with the RMSE between model results and ground observations reduced after assimilation.

6.6. Ensemble forecasting with machine learning algorithms

Keywords: *air quality*, *chemistry-transport model*, *data assimilation*, *ensemble forecast*.

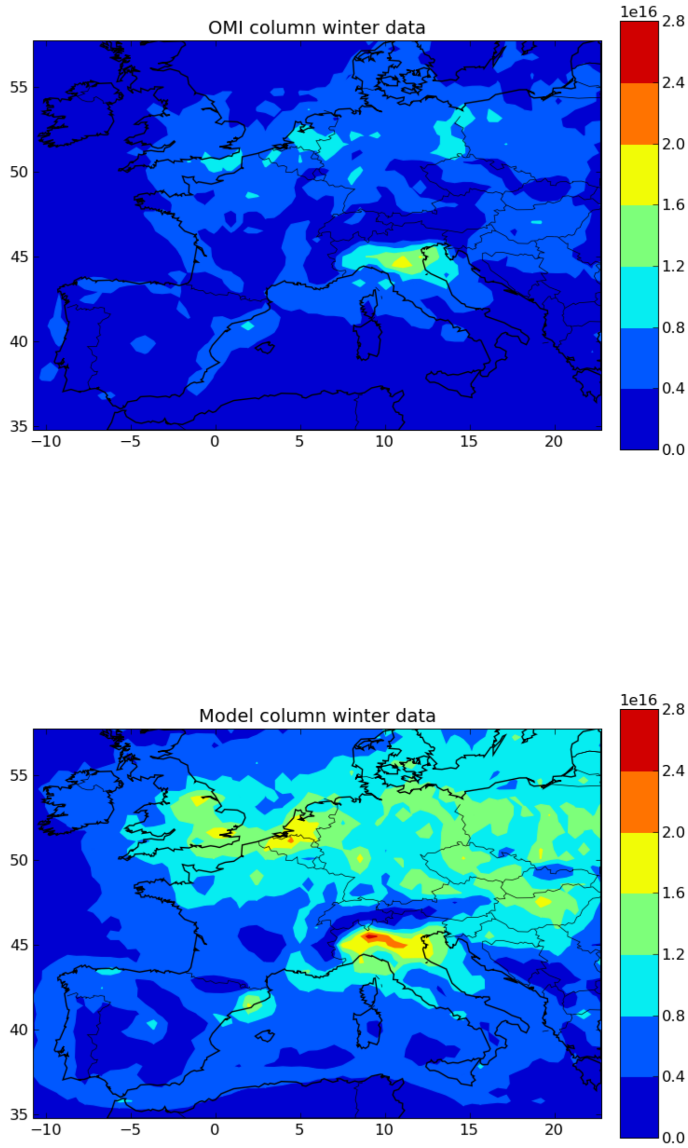

Figure 7. Top: the OMI column data averaged in Europe over November and December of 2005. Down: the Polyphemus simulation of model columns averaged in Europe in the same time period. Good consistency is seen in Spain, Italy, Great Britain, North Europe. Model simulations have higher values than satellite observations.

Participants: Vivien Mallet, Gilles Stoltz [CNRS], Sébastien Gerchinovitz, Édouard Debry [INERIS].

Based on ensemble simulations, improved forecasts can be generated by means of linear combinations of the individual-model forecasts. A weight is associated to each model, depending on past observations and simulations (figure 8). New machine learning algorithms (sequential aggregation) were developed and used for this purpose. Most of these methods provide theoretical bounds on the performance (relatively to the optimal constant model combination) and deliver significantly improved forecasts in all configurations.

The practical performance of the methods is very satisfactory. The theoretical bounds are always reached: the potential of the ensemble is well exploited. This was checked for large ensembles (dozens of models) as well as for small ensembles (a few models). The methods were successfully applied to forecast ozone, nitrogen dioxide and aerorols in operational mode, on the Prév'air platform (test mode) managed by INERIS.

The new results focus on the calibration of the parameters used in the learning algorithms: automatic calibration, renormalization of processed data, adaptative penalization when relevant, sparsity. First steps were also made to reduce the computational cost through a priori selection of the ensemble members.

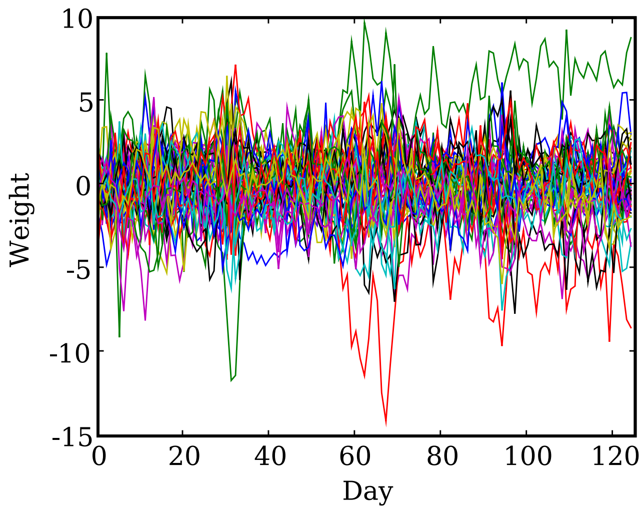

Figure 8. Weights associated by a machine learning algorithm to the 48 members of an ensemble, against time. The weights are not constrained (they can be negative).

6.7. Uncertainty estimation based on multimodel ensembles

Keywords: *air quality*, *chemistry-transport model*, *ensemble forecast*.

Participants: Damien Garaud, Vivien Mallet.

Air quality forecasts are limited by strong uncertainties especially in the input data and in the physical formulation of the models. There is a need to estimate these uncertainties for: the evaluation of the forecasts, the production of probabilistic forecasts and a more accurate estimation of the error statistics required by data assimilation.

Because a large part of the uncertainty in the forecast originates from uncertainties in the model formulation (primarily the physical parameterizations), multimodel ensemble seems to be the adequate tool for uncertainty estimation. A large ensemble with 100 members over year 2001 was generated and analyzed with criteria like the Brier score. Preliminary work on the calibration of the ensemble was carried out (figure 9): the ensemble members were selected so as to optimize the evaluation criteria. This may be formulated as a combinatorial optimization problem where one searches for an optimal combination of models out of a huge space of acceptable models.

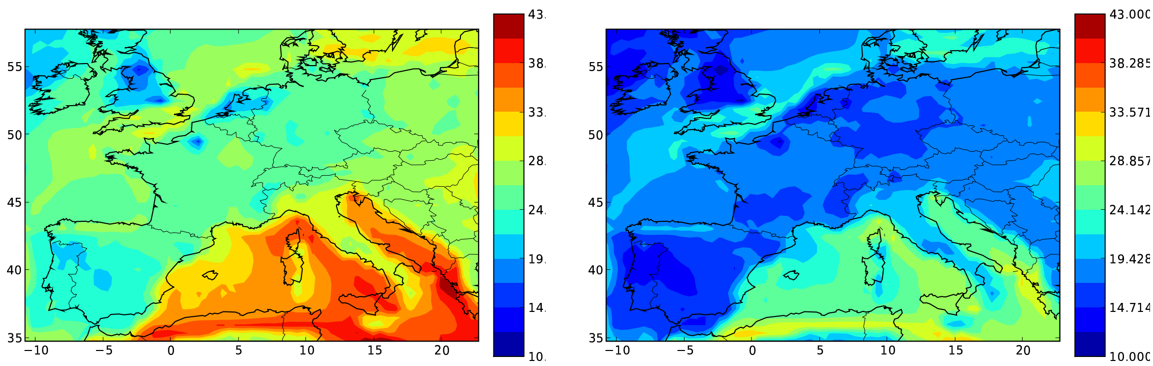

Figure 9. Uncertainty estimate (standard deviation) of two ensembles: a raw ensemble on the left, a calibrated ensemble on the right. The uncertainty estimation is significantly reduced after calibration.

6.8. Design of a monitoring network over France in case of a radiological accidental release

Participants: Marc Bocquet, Olivier Saunier [Ecole des Ponts ParisTech/IRSN].

The Institute of Radiation Protection and Nuclear Safety (France) is planning the set-up of an automatic nuclear aerosol monitoring network over the French territory (Descartes network), which complements the Teleray network. Each of the stations will be able to automatically sample the air aerosol content and to provide with activity concentration measurements on several radionuclides. This should help monitor the French and neighbouring countries nuclear power plant park. It would help evaluate the impact of a radiological incident on this park.

After the completion of the first phase (2006 and 2007), the second stage of the study has started in March 2008. The resolution has increased from 0.36 $^\circ\times$ 0.36 $^\circ$ to 0.25 $^\circ\times$ 0.25°, which doubles the number of potential sites, and hence the complexity of the optimisation. Meteorological fields have been generated with MM5 instead using ECMWF fields. New considerations are taken into account: the inclusion of foreign nuclear power plants, the validation of the optimal network on new cost functions that have not been considered yet, or taking into account the population density as a weighting factor. Because the Descartes network might be deployed sequentially, we are also considering sub-optimal network design algorithms. The computational time which was an important issue in the first stage is now a decisive issue because of the resolution increase. In order to accelerate optimization, we are considering and testing reduction of the accident database.

6.9. Reduction of an air quality monitoring network over France and Région Ile-de-France

Participants: Lin Wu, Marc Bocquet.

Ozone is an important air pollutant and observational networks are constructed for its estimation at the ground level. Due to the heterogeneous nature of the ozone field, the way ozone is observed does matter in the estimation of the concentrations. The evaluation of the network is thus of both theoretical and practical interests. In this study, we assess the efficiency of the BDQA (Base de Donnée sur la Qualité de l'Air) network, by investigating a network reduction problem. We examine how well a subset of this network can represent the full network. The performance of a subnetwork is taken to be the root mean square error of the spatial estimations of ozone concentrations over the whole network based on the observations from that subnetwork. Spatial interpolations are conducted for the ozone estimation taking into account the spatial correlations. Several interpolation methods, namely ordinary kriging, simple kriging about means, kriging with means as external drifts, are compared for a reliable estimation. It is found that the statistical information about the means improves significantly the kriging results. We employ a translated exponential model for the spatial correlations. We show that it is necessary to consider the correlation model to be hourly-varying but daily stationary. The network reduction problem is solved using the simulated annealing algorithm. We obtain considerable improvements for the subnetworks with different sizes. The redundant stations can thus be neglected to save maintenance costs.

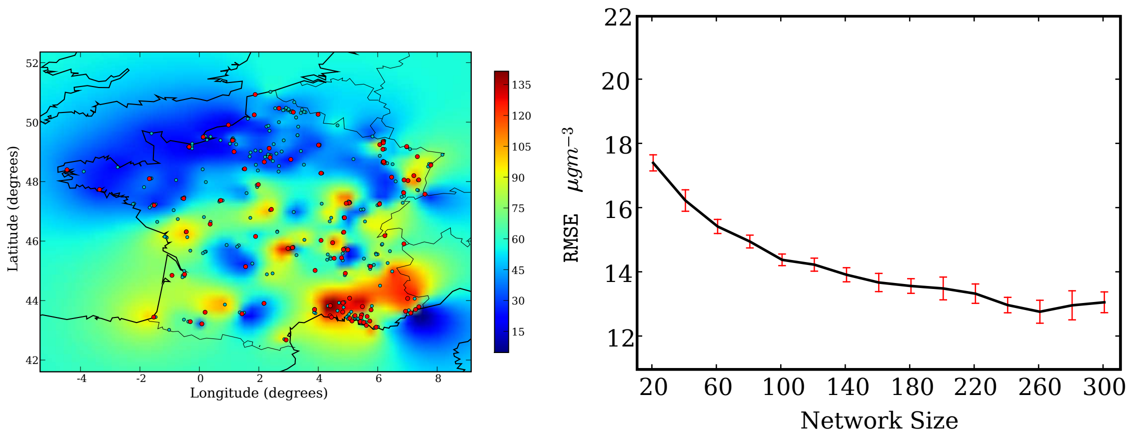

Figure 10. On the left: example of an optimal design and subsequent nowcasting. 100 ozone stations are kept out of the 351 BDQA stations. The average kriging rms error is 12.30 $\mu g.m^3$. On the right: in comparison, average rms error from a random network depending on the size of the network [R2DS project].

6.10. Targeting of observations in case of a nuclear accidental release

Participants: Rachid Abida, Marc Bocquet.

In the event of an accidental atmospheric release from a nuclear power plant, high resolution and accurate information on the spread of the radioactive plume around the accident site constitute a major key point, acutely required by decision makers in order to evaluate early countermeasure actions and consequences. Therefore, deploying mobile measuring devices constitutes an adequate monitoring strategy that allows to follow the real-time evolution of the radioactive plume. In fact, the collected measurements from the mobile network could be assimilated conjointly with data derived from the fixed monitoring network, so that to enhance knowledge on the state of the radioactive cloud. The targeting design consists in seeking the optimal spatial locations of the mobile stations at a certain time that satisfy some design criterion based on all available previous information. To illustrate how much a targeting strategy could improve the available information on the state of the radioactive plume, we considered an hypothetical accident release occurring at the Bugey power plant and a sequential data assimilation scheme based on inverse modeling to reconstruct the accident event. This assimilation scheme was coupled with a targeting strategy. Our first results are promising and show that our approach is strongly sensitive to the errors modeling, so that an important effort have to be made on this point in order to make the used data assimilation scheme more realistic.

6.11. Modeling wildland fire propagation

Keywords: *Firespread*, *Multivac*, *level set method*.

Participants: Vivien Mallet, David Keyes [Columbia University], Frank Fendell [Northrop Grumman Space Technology].

Modeling wildland fires consists in simulating the propagation of a fire front idealized as a curve on a non-flat terrain. The curve may be closed or not. The front may also be split into several unrelated curves. Starting from an ignition location, the front expands with a velocity determined by the wind and the fuel load.

Based on an empirical physical model developed by Fendell and Wolff, simulations of idealized fires and of one real fire (figure 11) were carried out with the level set method implemented in the package Multivac. This method comes with a well-developed mathematical framework that may be missing in alternative approaches. The study demonstrated its practical value for firespread simulations.

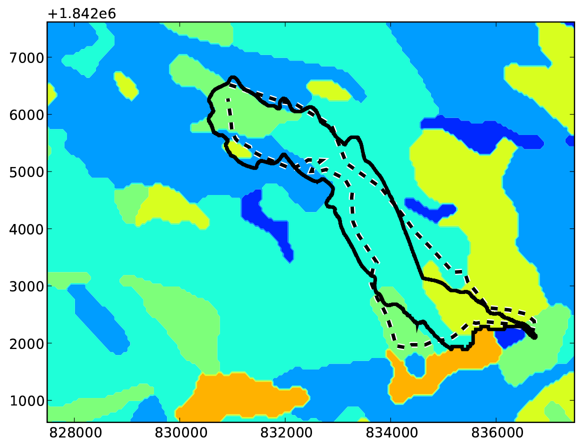

Figure 11. Simulated firefront (solid line) and observed front (dashed line) for a fire that occurred in 2005 in Lançon de Provence. Background colors delimit vegetation types.

6.12. Automating the learning process of a large scale classification chain

Keywords: *classification*, *deforestation*, *environmental monitoring*, *learning*, *satellite image*.

Participants: Nicolas Mercier, Isabelle Herlin, Jean-Paul Berroir.

This study concerns the automation of the learning process required for operating a large scale classification, from time-series of low resolution satellite data, and monitoring deforestation. The satellite sensor acquires temporal profiles (time-series of vegetation indices) which are represented in a specific feature space and classified. Each use of this large scale classification requires an adaptation to the sensor, the site, the observed classes, and hence a learning must be performed. The objective is to automate as much as possible this learning phase, which includes:

- Determination of the low resolution classes;
- Characterization of the profiles by high-level features;
- Selection of the relevant features allowing class discrimination;
- Selection and training of a statistical or fuzzy classifier.

The robustness of the processing chain and fuzzy classification schemes have been improved. An operational classification software, derived from this learning process and applied to deforestation monitoring in Brazil using MODIS images, is currently being registered commonly by INRIA and Embrapa (Brazilian agricultural research organization).

7. Contracts and Grants with Industry

7.1. IRSN

A convention has been established between CEREA (Clime is simultaneously a team of CEREA and a project-team of INRIA) and IRSN. This convention notably includes the installation of the Polyphemus software at the technical crisis centre of IRSN. A research contract is furthermore underway concerning network design: its objective is the optimal definition of an aerosol monitoring network for detection and diagnosis in the event of an accident/event in a French or neighbouring nuclear power plant.

7.2. INERIS

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, primarily for local-scale applications (dispersion of water plumes).

8. Other Grants and Activities

8.1. National initiatives

- Clime is running the project “Application of advanced inverse modelling and data assimilation methods to accidental dispersion of pollutants in an emergency situation”, under the framework of the LEFE-ASSIM program of INSU. This includes a cooperation with the Finnish Meteorological Institute. A postdoctoral grant of IRSN has been labelled by this program.
- Clime is member of the ADDISA contract (Ministry Grant, started at the beginning of 2007), with MOISE (INRIA Grenoble Rhône-Alpes), LEGI, the CNRM/GAME laboratory of Météo-France and the MIP laboratory of Université Paul Sabatier in Toulouse. This contract concerns image assimilation with application to forecast of extreme meteorological events and marine circulation.
- Clime is leading ADDISAAF (*Assimilation de Données Distribuées et Images SAtellite pour l'AFrique*) funded by IRD in the Corus program framework, in collaboration with ENIT (Tunisia), the Yaoundé University, and MOISE.

- A project aiming to the optimal design of an air quality monitoring network at regional and national scale has been accepted by the Ile-de-France region in the framework of the R2DS program, granting two years of postdoctoral fellowship.
- Clime takes part to the ANR project ATLAS ("From Applications to Theory in Learning and Adaptive Statistics"). Clime collaborates with Gilles Stoltz, co-leader of ATLAS, on the application of machine learning to ozone forecasting.
- Clime is participating to the TRAQ proposal (TRophospheric composition and Air Quality) for the definition of the future ESA troposphere chemistry mission (before 2015), currently pre-selected in a short list. Clime is therefore involved in the “GST TRAQ” (TRAQ scientific group), involving the French community of atmospheric remote sensing and coordinated by CNES.
- The three-year project Multiscale Data Assimilation in Geophysics [MSDAG] has been accepted by the ANR SYSCOMM. Fours partners are in the project: CEREA (Clime project-team, Marc Bocquet, PI of the whole project), INRIA Rennes (Vista Project-team, Etienne Mémin), LSCE (Peter Rayner), and INRIA Rhône Alpes (MOISE project-team, Laurent Debreu). The preparatory work has led to the definition of a document where an overview of state-of-the-art methodological approaches for multiscale data assimilation is presented. The project is due to start January 2009.

8.2. European initiatives

- Clime is a Principal Investigator of the ESA-Eumetsat “EPS-Metop” Research Announcement of Opportunity on assimilation of space-borne chemical measurements of the troposphere with application to air quality forecast. The EPS-Metop satellite is officially operational since May 2007 and the “level 2” chemical products will be soon available. Clime has performed preliminary studies and conducts researches on the assimilation of EPS data.
- Clime is member of the ERCIM working group “Environmental Modeling”. Within this working group, Clime cooperates with FORTH-IACM on remote sensing methodologies, GIS and definition of ontologies for complex applications.

8.3. International initiatives

- Clime has cooperations with CMM (Chile) and University of Cordoba (Argentina) on establishing air quality forecast systems and data assimilation capacities in Chile and Argentina. This cooperation is currently supported by the research project STIC-AmSud. The objective is the forecast of air quality using data assimilation techniques in South America. This project involves CMM (Chile), the Chilean weather office (DMC, Chile), the University of Cordoba (Argentina), and the environmental monitoring group of CNEA (Argentina). The Chilean meteorological office (Dirección Meteorológica de Chile) now 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).
- A new collaboration with the Nansen-Zhu International Research Center (Institute of Atmospheric Physics, Chinese Academy of Sciences) started with two objectives: performing operational forecast with Polyphemus for the Beijing Olympic Games (2008) and a long-term collaboration for air quality forecast, data assimilation and ensemble forecast.
- Clime leads, in cooperation with Moise, the associated team ADAMS (*Advanced Data Assimilation for the Sea*). The team is composed of Clime and Moise, the Marine Hydrophysical Institute (Ukrain), the Institute of Numerical Mathematics (Russia) and the Nodia Institute of Geophysics (Georgia).
- An ECO-NET project, ADOMENO, started in 2008, in collaboration with Georgia, Russia and Ukrain. The objectives of ADOMENO are the enhancement of data assimilation techniques by the use of high level data (such as image or Lagrangian data) and advanced assimilation methods. The application domain is the Black Sea circulation.

- A Safeti project (cooperation with South Africa in Information Technologies) is running with F'SATIE, MERAKA Institute (RSA), IRD and ESIEE (France) on the detection, recognition, tracking and characterization of satellite image features for environmental forecast and monitoring, with applications to ocean circulation and fire detection and forecast.
- An INRIA P3+3 Méditerranée project, DESMED, started in 2006 and was completed in 2008 on the study of desertification in Northern Africa and Southern Italy using long term series of vegetation indices. This project involved CNR-ISAC in Italy, INSAT in Tunisia and University Ibn Tofail in Morocco.
- A research project, ENVIAIR, under the framework of the INRIA-CNPQ programme, (with the State University of Rio de Janeiro, Federal University of Rio de Janeiro and Embrapa Solos) has been completed. This aims at monitoring deforestation and sustainable agricultural practices in the Taquari basin (Brasil).

9. Dissemination

9.1. Leadership within scientific community

- Dominique Béréziat is member of the commission of specialists for UPMC.
- Marc Bocquet is a permament guest of the scientific commitee of the INSU/LEFE action Assimilation.
- Isabelle Herlin is member of the commission of specialists for University Paris 12.
- Clime is involved in the national and European scientific community of satellite date for atmosphere chemistry. Clime is PI of ESA-EumetSat project for the exploitation of EPS-MetOp data to air quality forecast and is member of the French scientific group coordinated by CNES in the framework of the TRAQ proposal.

9.2. Teaching

- Computer Vision: 36 hours, Master EDITE de Paris / UPMC (Dominique Béréziat).
- Data Assimilation for Geophysics (ENSTA/ENPC): 30 hours (Marc Bocquet, Vivien Mallet).
- Algorithmics: 30 hours, ESIEE Management (Isabelle Herlin).
- Computational Physics for Environmental Applications (ENSTA/ENPC): 12 hours (Irène Korsakissok, Vivien Mallet).
- Introduction to chemistry-transport models (Paris VII): 3 hours (Vivien Mallet).
- Air Pollution (ENPC): 1h30 (Vivien Mallet).

9.3. Conference and workshop committees, invited conferences

- Marc Bocquet:
  - Invited seminar, LJK/MOISE, Grenoble, February 2008.
  - Invited seminar, UTC/LMAC, Compiègne, March 2008.
  - Workshop (invitation), "Mathematical Advancement in Geophysical Data Assimilation", BIRS, Banff, Canada, February 2008.
  - Oral presentation, EGU 2008, Vienna, Austria, April 2008.
  - Invited talk, "Statistical modeling of extremes in data assimilation and filtering approaches", Assimilex, Strasbourg, June 2008.

  - Invited talk, EMEP workshop, TFMM, Oslo, Norway, October 2008.
  - Invited seminar and visiting scientist, CMM, Santiago of Chile, Chile, November 2008.
  - Invited talk, "Workshop on 4DVar and Ensemble Kalman Filter Inter-Comparisons", Buenos-Aires, Argentina, November 2008.
- Isabelle Herlin:
  - Invited seminar and visiting scientist, EMBRAPA, Brasil, February 2008.
  - Invited talk, MHI, Ukrain, March 2008.
  - Invited talk, INSAT, Tunisia, May 2008.
  - Invited seminar, KNMI, Netherland, July 2008.
  - Invited seminar, LJK/MOISE, Grenoble, July and Septembre 2008.
  - Invited talk, EADS, Paris, December 2008.
- Vivien Mallet:
  - Invited seminar and visiting scientist, IAP, Chinese Academy of Science, February 2008.

9.4. Visiting scientists

- Pr. Gennady Korotaev from the Marine Hydrophysical Institute of Sebastopol, Ukraine: from October 13th to 17th.
- Pr. Maithili Sharan from Indian Institute of Technology, New Delhi, India : from June 28th to July 31th.

10. Bibliography

Major publications by the team in recent years

[1] M. BOCQUET. *Reconstruction of an atmospheric tracer source using the principle of maximum entropy*, in "Journal of the Royal Meteorological Society", vol. 131, n^{o} part B, 2005, 2191.

[2] M. BOCQUET. *High resolution reconstruction of a tracer dispersion event: application to ETEX*, in "Q. J. R. Meteorol. Soc.", vol. 133, 2007, p. 1013–1026.

[3] I. HERLIN, F.-X. L. DIMET, E. HUOT, J.-P. BERROIR. *e-Environement: progress and challenges*, in "Coupling models and data: which possibilities for remotely-sensed images?", P. PRASTACOS, J. MURILLO, J. L. DÍAZ DE LEÓN, U. CORTÉS (editors), Instituto Politécnico Nacional, México, 2004, p. 365–383.

[4] I. HERLIN, E. HUOT, J.-P. BERROIR, F.-X. L. DIMET, G. KOROTAEV. *Estimation of a motion field on satellite images from a simplified ocean circulation model*, in "ICIP International Conference on Image Processing, Atlanta, USA", October, 8-11 2006.

[5] M. KRYSTA, M. BOCQUET. *Source reconstruction of an accidental radionuclide release at European scale*, in "Quarterly Journal of the Royal Meteorological Society", vol. 133, 2007, p. 529-544.

[6] V. MALLET, B. SPORTISSE. *Ensemble-based air quality forecasts: A multimodel approach applied to ozone*, in "Journal of Geophysical Research", vol. 111, n^{o} D18, 2006.

[7] V. MALLET, B. SPORTISSE. *Uncertainty in chemistry-transport model due to physical parameterizations and numerical approximations: An ensemble approach applied to ozone modeling*, in "Journal of Geophysical Research", vol. 111, n^{o} D1, 2006.

[8] B. SPORTISSE. *Pollution atmosphérique: des processus à la modélisation*, 350 pages, Springer France, 2007.

[9] L. WU, V. MALLET, M. BOCQUET, B. SPORTISSE. *A Comparison Study of Data Assimilation Algorithms for Ozone Forecasts*, in "J. Geophys. Res.", vol. 113, n^{o} D20310, 2008.

Year Publications

Articles in International Peer-Reviewed Journal

[10] R. ABIDA, M. BOCQUET, N. VERCAUTEREN, O. ISNARD. *Design of a monitoring network over France in case of a radiological accidental release*, in "Atmospheric Environment", vol. 42, 2008, p. 5205-5219.

[11] M. BOCQUET. *Inverse Modelling of Atmospheric Tracers: non Gaussian methods and second-order sensitivity analysis*, in "Non-linear processes in Geophysics", vol. 15, 2008, p. 127-143.

[12] G. KOROTAEV, E. HUOT, F.-X. L. DIMET, I. HERLIN, S. STANICHNY, D. SOLOVYEV, L. WU. *Retrieving Ocean Surface Current by 4-D Variational Assimilation of Sea Surface Temperature Images*, in "Remote Sensing of Environment, Special issue on data assimilation", DOI: 10.1016/j.rse.2007.04.020, vol. 112, April 2008, p. 1464-1475.

[13] M. KRYSTA, M. BOCQUET, J. BRANDT. *Probing ETEX-II data set with inverse modelling*, in "Atmospheric Chemistry and Physics Discussion", vol. 8, 2008, p. 2795-2819.

[14] V. MALLET, D. KEYES, F. FENDELL. *Modeling Wildland Fire Propagation with Level Set Methods*, in "Computers & Mathematics with Applications", accepted, 2009.

[15] V. MALLET, B. SPORTISSE. *Air quality modeling: From deterministic to stochastic approaches*, in "Computers & Mathematics with Applications", vol. 55, n^{o} 10, 2008, p. 2,329–2,337.

[16] V. MALLET, G. STOLTZ, B. MAURICETTE. *Ozone ensemble forecast with machine learning algorithms*, in "Journal of Geophysical Research", accepted, 2009.

[17] E. TOMAI, I. HERLIN, J.-P. BERROIR, P. PRASTACOS. *Ontology-based land degradation assessment from satellite images*, in "International Journal of Remote Sensing", 2008.

[18] M. TOMBETTE, V. MALLET, B. SPORTISSE. PM_{10} *data assimilation over Europe with the optimal interpolation method*, in "Atmospheric Chemistry and Physics", accepted, vol. 8, n^{o} 3, 2009, p. 9,607–9,640.

[19] L. WU, V. MALLET, M. BOCQUET, B. SPORTISSE. *A Comparison Study of Data Assimilation Algorithms for Ozone Forecasts*, in "J. Geophys. Res.", vol. 113, n^{o} D20310, 2008.

[20] H. YAHIA, A. TURIEL, N. CHRYSOULAKIS, J. GRAZZINI, P. PRASTACOS, I. HERLIN. *Application of the Multifractal Microcanonical Formalism to the Detection of Fire Plumes in NOAA-AVHRR Data*, in "International Journal Of Remote Sensing", 2008.

Articles in National Peer-Reviewed Journal

[21] M. BOCQUET, B. SPORTISSE. *Modélisation inverse pour la qualité de l'air : éléments de méthodologie et exemples*, in "Pollution Atmosphérique", 2008.

International Peer-Reviewed Conference/Proceedings

[22] E. HUOT, I. HERLIN, G. KOROTAEV. *Assimilation of SST Satellite Images For Estimation Of Ocean Circulation Velocity*, in "Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Boston, Massachusetts, U.S.A.", July 2008.

[23] T. ISAMBERT, J.-P. BERROIR, I. HERLIN. *A Multiscale Vector Spline Method for Estimating the Fluids Motion on Satellite Images*, in "ECCV'08: European Conference on Computer Vision, Marseille, France", Springer, October 2008.

[24] R. TEINA, D. BÉRÉZIAT, B. STOLL, S. CHABRIER. *Toward A Global Tuamotu Archipelago Trees Sensing Using High Resolution Optical Data*, in "IGARSS, Boston, USA", July 2008.

[25] R. TEINA, D. BÉRÉZIAT, B. STOLL. *A Spatial Poisson Process to classify coconut fields on Ikonos pansharpened images*, in "SPIE Asia Pacific Remote Sensing, Noumea (New Caledonia)", November 2008.

[26] E. TOMAI, I. HERLIN, J.-P. BERROIR, P. PRASTACOS. *Ontology-based land degradation assessment from satellite images*, in "International Society for Photogrammetry and Remote Sensing, Beijing, China", July 2008.

National Peer-Reviewed Conference/Proceedings

[27] D. BÉRÉZIAT, I. HERLIN. *Calcul du flot optique par assimilation variationnelle de données*, in "16ème congrès de Reconnaissance des Formes et Intelligence Artificielle, RFIA'08, Amiens, France", 22-25 January 2008, p. 341–350.

Research Reports

[28] D. BÉRÉZIAT, I. HERLIN. *Solving ill-posed Image Processing problems using Data Assimilation. Application to optical flow*, Research Report, n^{o} 6477, INRIA, March 2008, http://hal.inria.fr/inria-00264661.

[29] T. ISAMBERT, J.-P. BERROIR, I. HERLIN. *A multi-scale vector spline method for estimating the fluids motion on satellite images*, Research Report, n^{o} 6478, INRIA, 03 2008, http://hal.inria.fr/inria-00264727.

[30] W. REKIK, D. BÉRÉZIAT, S. DUBUISSON. *Fusion de données spatiales (3D) et temporelles (2D+t) pour la reconstruction de scènes 3D+t : deux approches*, Technical report, Laboratoire d'Informatique de Paris 6, juin 2008, http://www.lip6.fr/lip6/reports/2008/lip6-2008-001.pdf.