Activity Report 2016

Project-Team CLIME

Coupling environmental data and simulation models for software integration

RESEARCH CENTER
Paris

THEME
Earth, Environmental and Energy Sciences

Table of contents

Project-Team CLIME

Creation of the Project-Team: 2005 September 01, end of the Project-Team: 2016 December 31

Keywords:

Computer Science and Digital Science:

3.3. - Data and knowledge analysis
3.4.1. - Supervised learning
3.4.5. - Bayesian methods
5.3. - Image processing and analysis
5.9. - Signal processing
6.2.1. - Numerical analysis of PDE and ODE
6.2.3. - Probabilistic methods
6.2.7. - High performance computing
6.3. - Computation-data interaction
6.3.1. - Inverse problems
6.3.2. - Data assimilation
6.3.3. - Data processing
6.3.4. - Model reduction
6.3.5. - Uncertainty Quantification

Other Research Topics and Application Domains:

3.3. - Geosciences
3.4. - Risks
4.3.3. - Wind energy
4.3.4. - Solar Energy
8.2. - Connected city

1. Members

Research Scientists

Isabelle Herlin [Team leader, Inria, Senior Researcher, HDR]
Vivien Mallet [Inria, Researcher]

Faculty Members

Julien Brajard [Univ. Paris VI, Associate Professor, until Aug 2016]
Étienne Huot [Univ. Versailles Saint-Quentin-en-Yvelines, Associate Professor]

Engineers

Guillaume Chérel [Inria, from Apr 2016]
Nicolas Claude [Inria, until Feb 2016]
Marius Guérard [Inria, from Nov 2016]

PhD Students

Jean Thorey [EDF]
Raphaël Ventura [Inria]

Administrative Assistant

Nathalie Gaudechoux [Inria]

Others

Dominique Béréziat [Univ. Paris VI, Associate Professor, external collaborator, HDR]
Timothée Maziol [Inria, Internship, from Apr 2016 until Aug 2016]
Louis Philippe [Inria, Internship, from Jul 2016]

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 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. All 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 created for studying these questions with researchers specialized in data assimilation and image processing.

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. Research Program

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 air quality models through collaborations with École des Ponts ParisTech and EDF R&D: the models from Polyphemus (pollution forecasting from local to regional scales) and Code_Saturne (urban scale). In regard to other modeling domains, such as oceanography and meteorology, Clime accesses models through co-operation with LOCEAN (Laboratoire d'OCEANographie et du climat, UPMC) and Météo-France.

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 values or some image features. 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 Météo-France (in France), UK-met (in United Kingdom), 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 first 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 assess ocean surface circulation, by assimilating fronts and vortices displayed on image acquisitions. Concerning meteorology, the focus is on correcting the 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 improving 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. The ensemble methods allow 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. The two main objectives are: management of combinations of sensor types, and deployment modes and 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 field that represents 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 of these structures are developed and take into account the underlying physical processes. Image acquisitions are assimilated into these 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.

4.5. Smartcity

There is a growing interest for environmental problems at city scale, where a large part of the population is concentrated and where major pollutions can develop. Numerical simulation is well established to study the urban environment, e.g., for road traffic modeling. As part of the smartcity movement, an increasing number of sensors collect measurements, at traditional fixed observation stations, but also on mobile devices, like smartphones. A number of research issues can be raised:

- How to properly take into account the city geometry that makes the data assimilation problems unique?
- How to make use of the various sensors, sometimes mobile, of low quality but numerous?
- How to couple all the systems that are intricated at urban scale?

Pratical applications include air pollution and noise pollution. These directly relate to road traffic. Data assimilation and uncertainty propagation are key topics in these applications.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. Awards

Inria and Paris City were awarded a Décibel d'Argent 2016 in research category for the mobile application Ambiciti. The award was attributed by the Conseil National du Bruit, which depends on the Ministry of Ecology, Sustainable Development and Energy, and is a national organization in charge of noise. The selection committee pointed out the Ambiciti articulation between research, citizen involvement, city or government actions and the operational development of a rich and perennial mobile application.

6. New Software and Platforms

6.1. Heimdali

- Participants: Isabelle Herlin, Dominique Bereziat and David Froger
- Contact: Isabelle Herlin

SCIENTIFIC DESCRIPTION

The main components of Heimdali concern:

- the pre/post processing of image sequences,
- the image assimilation with numerical models,
- the visualization of image sequences.

FUNCTIONAL DESCRIPTION

The initial aim of the image processing library Heimdali was to replace an internal Inria library (named Inrimage) by a library based on standard and open source tools, and mostly dedicated to satellite acquisitions.

The leading idea of the library is to allow the following issues:
- making easier the sharing and development of image assimilation softwares. For that purpose, the installation is easily achieved with the package manager Conda.
- developing generic tools for image processing and assimilation based on ITK (Insight Segmentation and Registration Toolkit http://www.itk.org). In reverse provide tools to ITK and contribute to the ITK community. Our software corresponds to issues related to satellite acquisitions but could be of interest for processing medical image sequences.

6.2. Image Forecast

- Authors: Isabelle Herlin and Yann Lepoittevin
- Contact: Isabelle Herlin

SCIENTIFIC DESCRIPTION

From a given number of images, Image Forecast synthetizes the future images at a given and short temporal horizon.

FUNCTIONAL DESCRIPTION

Image forecast includes two components:

- it estimates the dynamics from an image sequence. Various options are available in the software: stationarity, Lagrangian conservation, description of structures. The result is the motion field explaining the temporal evolution of image data.
- the estimated dynamics is applied for forecasting future images at short temporal horizon.

6.3. Polyphemus

- Participants: Sylvain Doré (CEREA, École des Pont ParisTech) and Vivien Mallet
- Contact: Vivien Mallet
- URL: http://cerea.enpc.fr/polyphemus/

FUNCTIONAL DESCRIPTION

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 post-processing abilities (AtmoPy),
- programs for physical pre-processing and chemistry-transport models (Polair3D, Castor, two Gaussian models, a Lagrangian model),
- model drivers and observation modules for model coupling, ensemble forecasting and data assimilation.

Fig. 1 depicts a typical result produced by Polyphemus.

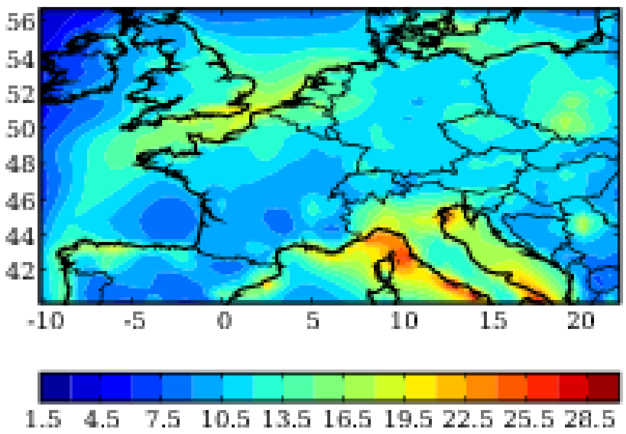

Figure 1. Map of the relative standard deviation (or spread, %) of an ensemble built with Polyphemus (ozone simulations, $\mu g\ m^{-3}$). The standard deviations are averaged over the summer of 2001. They provide an estimation of the simulation uncertainties.

Clime is involved in the overall design of the system and in the development of advanced methods in model coupling, data assimilation and uncertainty quantification (through model drivers and post-processing).

6.4. SoundCity - Ambiciti

- Authors: Pierre-Guillaume Raverdy (SED), Fadwa Rebhi (Mimove), Cong Kinh Nguyen (Mimove), Rajiv Bhatia (TheCivicEngine), Vivien Mallet and Valerie Issarny (Mimove)
- Contact: Valerie Issarny (Mimove)

FUNCTIONAL DESCRIPTION

Ambiciti measures the actual noise levels to which you are exposed. It can monitor noise levels throughout the day and inform you about your instantaneous, hourly and daily exposures.
Ambiciti also computes the air quality index in your region or at the exact location where you stand. You can also access to forecasts.
Ambiciti includes a lot of features:

- Measuring noise level, anytime on demand or automatically during the day,
- Air quality indexes, in the past, present and future hours or days,
- Pollution levels for nitrogen dioxide, fine particulate matter and ozone,
- Statistics on exposure to pollutions, hourly, daily, during daytime and nighttime,
- Maps with your noise measurements,
- Hourly air quality maps, at street resolution in Paris, San Francisco, Oakland, Richmond (California), at present time,
- The recommendation of pedestrian routes which minimize the exposure to noise pollution or to air pollution,
- The ability to take pictures with pollution levels on top.

6.5. Urban noise analysis

- Authors: Vivien Mallet, Raphael Ventura and Guillaume Cherel
- Contact: Vivien Mallet

FUNCTIONAL DESCRIPTION

This software merges noise simulations and mobile observations. It can extract a given region of a noise map and filter out the buildings. It extends a previous software for data assimilation of air pollution observations at city scale. This prior software computes the so-called best linear unbiased estimator (BLUE), with a special background error covariance model that depends on the city geometry. The extension for noise introduces special treatments for the errors in mobile observations, and includes more statistical verifications.

The software also comes with a Python module for the management of a large database of mobile noise measurements, especially with many filters relying the observations metadata.

The software finally includes the automatic generation of a report based on intensive measurements in a city district. This report targets participants of crowdsensing experiments.

6.6. Verdandi

- Participants: Vivien Mallet, Gautier Bureau (Medisim), Dominique Chapelle (Medisim), Sébastien Gilles (Medisim) and Philippe Moireau (Medisim)
- Contact: Vivien Mallet
- URL: http://verdandi.gforge.inria.fr/

FUNCTIONAL DESCRIPTION

Verdandi is a free and open-source (LGPL) library for data assimilation. It includes various methods for coupling one or several numerical models and observational data. Mainly targeted at large systems arising from the discretization of partial differential equations, the library is devised as generic, which allows for applications in a wide range of problems (biology and medicine, environment, image processing, etc.). Verdandi also includes tools to ease the application of data assimilation, in particular in the management of observations or for a priori uncertainty quantification. Implemented in C++, the library may be used with models implemented in Fortran, C, C++ or Python.

7. New Results

7.1. State estimation: analysis and forecast

One major objective of Clime is the conception of new methods of data assimilation in geophysical sciences. Clime is active on several challenging aspects: non-Gaussian assumptions, multiscale assimilation, minimax filtering, etc.

7.1.1. Assimilation of drifter data in the East Mediterranean Sea

Participants: Julien Brajard, Isabelle Herlin, Leila Issa [Lebanese American University, Lebanon], Laurent Mortier [LOCEAN], Daniel Hayes [Oceanography Centre, Cyprus], Milad Fakhri [CNRS, Lebanon], Pierre-Marie Poulain [Oceanography Institute of Trieste, Italy].

Surface velocity fields of the ocean in the Eastern Levantine Mediterranean are estimated by blending altimetry and surface drifters data. The method is based on a variational assimilation approach for which the velocity is corrected by matching real drifters positions with those predicted by a simple advection model, while taking into account the wind effect. The velocity correction is done in a time-continuous fashion by assimilating at once a whole trajectory of drifters with a temporal sliding window. Except for the wind component, a divergence-free regularization term was added to constrain the velocity field. Results show that, with few drifters, our method improves the estimated velocity in two typical situations: an eddy between the Lebanese coast and Cyprus, and velocities along the Lebanese coast. A description of these results is published in the Ocean Modelling journal.

7.1.2. State estimation for noise pollution

Participants: Raphaël Ventura, Vivien Mallet, Valérie Issarny [Mimove], Pierre-Guillaume Raverdy [SED], Fadwa Rebhi [Mimove], Cong Kinh Nguyen [Mimove].

70 million observations of ambient noise have been collected with the mobile application Ambiciti (previously, SoundCity). An important work was carried out on the calibration of the measurements. Over 100 mobile phones were calibrated against a sound level meter, at various noise intensities and frequencies, in order to test their response and devise a calibration strategy.

A data assimilation procedure has been put in place in order to select and assimilate the most reliable observations. Simulated noise maps have been improved with the observations, by computing the so-called best linear unbiased estimator (BLUE) with error covariance models suitable for noise pollution. The assimilation of mobile observation introduces new errors, like location errors, compared to the assimilation of the more common observations from fixed monitoring stations.

7.2. 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, the 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 image data in an optimal way is of major interest. This issue is twofold:

- from the model's viewpoint, the location of structures on the observations is used to control the state vector.
- from the image's viewpoint, a model of the dynamics and structures is built from the observations.

7.2.1. Estimation of motion and acceleration from image data

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

Image sequences allow visualizing dynamic systems and understanding their intrinsic characteristics. One first component of this dynamics is obtained by retrieving the velocity of the structures displayed on the sequence. This motion estimation issue has been extensively studied in the literature of image processing and computer vision. In this research, we step beyond the traditional optical flow methods and address the problem of recovering the acceleration from the whole temporal sequence, which has been poorly investigated, even if this is of major importance for some data types, such as fluid flow images. Acceleration is here viewed as the space-time function resulting from the forces applied to the studied system. To solve this issue, we propose a variational approach where a specific energy is designed to model both the motion and the acceleration fields. The contributions are twofold: first, we introduce a unified variational formulation of motion and acceleration under space-time constraints; second, we define the minimization scheme, which allows retrieving the estimations, and provide the full information on the discretization schemes. Experiments are conducted on synthetic and real image sequences, visualizing fluid-like flows, where direct and precise calculation of acceleration is of primary importance.

7.2.2. *Rain nowcasting from radar image acquisitions*

Participants: Isabelle Herlin, Étienne Huot.

This research concerns the design of an operational method for rainfall nowcasting that aims at prevention of flash floods. The nowcasting method includes two main components:

- a data assimilation method, based on radar images, estimates the state of the atmosphere: this is the estimation phase.
- a forecast method uses this estimation to extrapolate the state of the atmosphere in the future: this is the forecast phase.

Results are analyzed on space-time neighborhoods in order to prevent consequences of flash floods on previously defined zone.

Current research concerns the following issues:

- the use of object components in the state vector. The objective is to improve the description of the image data in order to get a better motion estimation and a more accurate localization of endangered regions.
- the extension of the estimation phase to a multiscale process.
- the merging with measures acquired by a network of pluviometers.

7.2.3. *Ensemble Kalman filter based on the image structures*

Participants: Dominique Béréziat [UPMC], Isabelle Herlin.

One major limitation of the motion estimation methods that are available in the literature concerns the availability of the uncertainty on the result. This is however assessed by a number of filtering methods, such as the ensemble Kalman filter (EnKF). Our research consequently concerns the use of a description of the displayed structures in an ensemble Kalman filter, which is applied for estimating motion on image acquisitions. Compared to the Kalman filter, EnKF does not require propagating in time the error covariance matrix associated to the estimation, resulting in reduced computational requirements. However, EnKF is also known for exhibiting a shrinking effect when taking into account the observations on the studied system at the analysis step. Various methods are available in the literature for correcting this effect, but they do not involve the structures displayed on the image sequence. We defined two alternative solutions to that shrinking effect: a dedicated localization function and an adaptive decomposition domain. These methods are both well suited for fluid flows images and applied on satellite images of the atmosphere.

7.3. Uncertainty quantification and risk assessment

The uncertainty quantification of environmental models raises a number of problems due to:

- the dimension of the inputs, which can easily be 10^5-10^8 at every time step;
- the dimension of the state vector, which is usually 10^5-10^7;
- the high computational cost required when integrating the model in time.

While uncertainty quantification is a very active field in general, its implementation and development for geosciences requires specific approaches that are investigated by Clime. The project-team tries to determine the best strategies for the generation of ensembles of simulations. In particular, this requires addressing the generation of large multimodel ensembles and the issue of dimension reduction and cost reduction. The dimension reduction consists in projecting the inputs and the state vector to low-dimensional subspaces. The cost reduction is carried out by emulation, i.e., the replacement of costly components with fast surrogates.

7.3.1. Sequential aggregation with uncertainty estimation

Participants: Jean Thorey, Vivien Mallet, Christophe Chaussin [EdF R&D].

In the context of ensemble forecasting, one goal is to combine an ensemble of forecasts in order to produce an improved probabilistic forecast. We previously designed a new approach to predict a probability density function or cumulative distribution function, from a weighted ensemble of forecasts. The procedure aims at forecasting the cumulative distribution function of the observation which is simply a Heaviside function centered at the observed value. Our forecast is the weighted empirical cumulative distribution function based on the ensemble of forecasts. Each forecast of the ensemble is attributed a weight which is updated whenever new observations become available. The performance of the forecast is given by the continuous ranked probability score (CRPS), which is the square of the two-norm of the discrepancy between the forecast and the observed cumulative distribution functions. The method guarantees that, in the long run, the forecast cumulative distribution function has a continuous ranked probability score at least as good as the best weighted empirical cumulative function with weights constant in time.

The CRPS computed from an ensemble of forecasts is subject to a bias. We proposed a new way to compute the CRPS in order to mitigate the bias and obtain better aggregation performance.

The work was applied to the forecast of photovoltaics production, both at EDF production sites and for global France production.

7.3.2. Sensitivity analysis of air quality simulations at urban scale

Participants: Vivien Mallet, Louis Philippe, Fabien Brocheton [Numtech], David Poulet [Numtech].

We carried out a sensitivity analysis of the urban air quality model Sirane. We carried out dimension reduction on both inputs and outputs of the air quality model. This designed a reduced-order model, which we then emulated. We sampled the (reduced) inputs to the reduced model, and emulated the response surface of the reduced outputs. A metamodel was derived by the combination of the dimension reduction and the statistical emulation. This metamodel performs as well as the original model, compared to field observations. It is also extremely fast, which allowed us to compute Sobol' indices and carry out a complete sensitivity analysis.

7.3.3. Sensitivity analysis of road traffic simulations and corresponding emissions

Participants: Ruiwei Chen [École des Ponts ParisTech], Vivien Mallet, Vincent Aguiléra [Cerema], Fabien Brocheton [Numtech], David Poulet [Numtech], Florian Cohn [Numtech].

This work deals with the simulation of road traffic at metropolitan scale. We compared state-of-the-art static traffic assignment and dynamic traffic assignment, which better represents congestion. The work was applied in Clermont-Ferrand and its surrounding region, for a time period of two years, and using about 400 traffic loop counters for evaluation. The dynamic model showed similar overall performance as the static model.

We developed an open source software for the computation of the emissions of traffic. It computes the emissions of the main air pollutants, according the vehicle fleet.

For both traffic assignment and pollutant emissions, we carry out sensitivity tests with respect to limit speed, roads capacities or fleet composition. A complete sensitivity analysis is out of reach with the complete, computational intensive, traffic assignment model. Hence further work has been engaged with the metamodeling of the traffic assignment model. Preliminary results are encouraging and tend to show that a very fast metamodel can perform as well as the complete model.

7.3.4. *Ensemble variational data assimilation*

Participants: Julien Brajard, Isabelle Herlin, Marc Bocquet [CEREA], Jérôme Sirven [LOCEAN], Olivier Talagrand [LMD, ENS], Sylvie Thiria [LOCEAN].

The general objective of ensemble data assimilation is to produce an ensemble of analysis from observations and a numerical model which is representative of the uncertainty of the system. In a bayesian framework, the ensemble represents a sampling of the state vector probability distribution conditioned to the available knowledge of the system, denoted the a-posteriori probability distribution.

Ensemble variational data assimilation (EnsVar) consists in producing such an ensemble by perturbating N times the observations according to their error law, and run a standard variationnal assimilation for each perturbation. An ensemble of N members is then produced. In the case of linear models, there is a theoretical guarantee that this ensemble is a sampling of the a-posteriori probability. But there is no theoretical result in the non-linear case.

The objective of this work is to study the ability of EnsVar to produce "good" ensemble (i.e. that sampled the a posteriori probablility) on a shallow-water model. Statistical properties of the ensemble are evaluated, and the sensitivity to the main features of the assimilation system (number, distribution of observations, size of the assimilation window, ...) are also studied.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Grants with Industry

A new Iilab, Rain_Water, has been accepted in 2016. It concerns joint research with the company Weather Measures. Rain_Water aims to define a platform of local meteorology. Users are mainly farmers that will use the platform for monitoring the agricultural practices at the parcel level.

9. Partnerships and Cooperations

9.1. National Initiatives

- The ANR project Estimair aims at quantifying the uncertainties of air quality simulations at urban scale. The propagation of uncertainties requires the use of model reduction and emulation. A key uncertainty source lies in the traffic emissions, which are generated using a dynamic traffic assignment model. Ensembles of traffic assignments are calibrated and used in the uncertainty quantification. Estimair is led by Clime.
- The IPSL project "AVES" (Ensemble Variational Assimilation applied to a shallow-water model) aims at estimating the quality of an ensemble produced by a variational ensemble algorithm applied on a shallow-water numerical model. A focus is made on the bayesian properties of the ensemble, i.e. its capacity to sample the a-posteriori probability law of the model state.
- Two new ANR projects have been accepted in 2016 and will begin in January 2017.
  FireCaster aims at fire forecasting and risk mitigation.
  Cense aims at the estimation of urban noise, using numerical simulation and a dense monitoring network.

9.2. European Initiatives

9.2.1. *FP7 & H2020 Projects*

9.2.1.1. EoCoE

Title: Energy oriented Centre of Excellence for computer applications

Program: H2020

Duration: October 2015 - October 2018

Coordinator: CEA

Partners:

CEA, Commissariat à l'Énergie Atomique et aux Énergies Alternatives (France)

Forschungszentrum Julich (Germany)

Max Planck Gesellschaft (Germany)

ENEA, Agenzia Nazionale Per le Nuove Tecnologie, l'energia E Lo Sviluppo Economico Sostenibile (Italy)

CERFACS, European Centre for Research and Advanced Training in Scientific Computing (France)

Instytut Chemii Bioorganicznej Polskiej Akademii Nauk (Poland)

Universita Degli Studi di Trento (Italy)

Fraunhofer Gesellschaft (Germany)

University of Bath (United Kingdom)

CYL, The Cyprus Institute (Cyprus)

CNR, National Research Council of Italy (Italy)

Université Libre de Bruxelles (Belgium)

BSC, Barcelona Supercomputing Center - Centro Nacional de Supercomputacion (Spain)

Inria contact: Michel Kern (Serena team)

The aim of the projevt is to establish an Energy Oriented Centre of Excellence for computing applications (EoCoE). EoCoE (pronounce "Echo") will use the prodigious potential offered by the ever-growing computing infrastructure to foster and accelerate the European transition to a reliable and low carbon energy supply. To achieve this goal, we believe that the present revolution in hardware technology calls for a similar paradigm change in the way application codes are designed. EoCoE will assist the energy transition via targeted support to four renewable energy pillars: Meteo, Materials, Water and Fusion, each with a heavy reliance on numerical modelling. These four pillars will be anchored within a strong transversal multidisciplinary basis providing high-end expertise in applied mathematics and HPC. EoCoE is structured around a central Franco-German hub coordinating a pan-European network, gathering a total of 8 countries and 23 teams. Its partners are strongly engaged in both the HPC and energy fields; a prerequisite for the long-term sustainability of EoCoE and also ensuring that it is deeply integrated in the overall European strategy for HPC. The primary goal of EoCoE is to create a new, long lasting and sustainable community around computational energy science. At the same time, EoCoE is committed to deliver high-impact results within the first three years. It will resolve current bottlenecks in application codes, leading to new modelling capabilities and scientific advances among the four user communities; it will develop cutting-edge mathematical and numerical methods, and tools to foster the usage of Exascale computing. Dedicated services for laboratories and industries will be established to leverage this expertise and to foster an ecosystem around HPC for energy. EoCoE will give birth to new collaborations and working methods and will encourage widely spread best practices.

9.2.1.2. Env&You 2016

Title: Env&You

Program: EIT Digital

Duration: January 2016 - December 2016

Coordinator: Inria (MiMove)

Partners and subgrantees:

Inria

NUMTECH

Cap Digital

Forum Virium (Finland)

TheCivicEngine (United States)

Ambientic

Inria contact: Valérie Issarny (Mimove project-team)

Env&You aims at delivering the whole picture of urban pollution, from the individual exposure to neighborhood-by-neighborhood and day-to-day variation, to citizens and governments, informing their decisions for healthy urban living. There is a clear, and probably increasing, desire from the citizens to better know their individual exposure to pollution. Partial solutions exist to the exposure data problem but each focuses on one or another domain of information—crowdsourcing exposure, translating governmental open data to usable consumer information, harnessing social media information, harnessing biometrics—what is unique about Env&You is that it will assimilate a multi-dimensional picture of exposure and provide the integrated information to citizen, government, and business use.

9.3. International Initiatives

9.3.1. Inria International Partners

9.3.1.1. Informal International Partners

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: IBM Research, Dublin, Ireland

The collaboration addresses the assimilation of classical observations as well as images, with application to geophysics. New assimilation methods are developed, mainly based on minimax filtering.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Selection

10.1.1.1. Reviewer

- Isabelle Herlin: International Conference on Computer Vision and Pattern Recognition (CVPR), International Conference on Computer Vision (ECCV), Asian Conference on Comuter Vision (ACCV), International Conference on Image Processing (ICIP).

10.1.2. Journal

10.1.2.1. Reviewer - Reviewing Activities

- Julien Brajard: IEEE Transactions on Geoscience and Remote Sensing (TGRS).
- Isabelle Herlin: IEEE Transactions on Geoscience and Remote Sensing (TGRS), IEEE Geoscience and Remote Sensing Letters, Mathematical Methods in Applied Sciences.

10.1.3. Invited Talks

- Vivien Mallet: Processing environmental simulations and observations for smart cities; Perspectives and New Challenges in Data Science, École des Ponts ParisTech; February 2016.
- Vivien Mallet: Assimilation de données et prévision d'ensemble appliquées à la qualité de l'air; CEA seminar; June 2016.

10.1.4. Scientific Expertise

- Isabelle Herlin is a member of the Scientific Council of ANDRA (French national radioactive waste management agency)
- Isabelle Herlin is a member of the Scientific Council of CSFRS (High Council for Strategic Education and Research in France).
- Isabelle Herlin and Vivien Mallet reviewed several research proposals, especially for ANR (France).

10.1.5. Research Administration

- Isabelle Herlin is a member of the program committee of DIGITEO, French research cluster in science and technology of information.
- Isabelle Herlin is a member of the Evaluation Committee at Inria.

10.2. Teaching - Supervision - Juries

10.2.1. Teaching

Master OACOS/WAPE: Marc Bocquet, Vivien Mallet, Jean-Matthieu Haussaire; Introduction to Data Assimilation for Geophysics; 12 hours; M2; UPMC, X, ENS, ENSTA ParisTech, École des Ponts ParisTech; France.

Master SGE and 3rd-year class at École des Ponts ParisTech: Vivien Mallet; Air quality modeling; 4.5 hours; M2; Universities Paris Diderot- Paris 7, Paris 12 and École des Ponts ParisTech, France.

Training: Vivien Mallet; Uncertainty Quantification: Ensembles and Data Assimilation – Application to Climate and Geosciences; 5.25 hours; CERFACS; France.

Training: Vivien Mallet; Introduction to data assimilation: Kalman filters and ensembles; 4.5 hours; CEMRACS (summer school).

10.2.2. Supervision

- PhD in progress : Pacôme Eberhart, "Génération automatique de codes performants et fiables pour l'assimilation de données", September 2013, Fabienne Jezequel, Pierre Fortin and Julien Brajard
- PhD in progress : Jean Thorey, "Prévision d'ensemble du rayonnement solaire pour la production photovoltaïque du parc EDF", November 2013, Vivien Mallet.
- PhD in progress: Ruiwei Chen, "Quantification d'incertitude en simulation des émissions du trafic routier", November 2014, Vivien Mallet.
- PhD in progress: Raphaël Ventura, "Simulation numérique de la ville par couplage entre la modélisation et l'observation", September 2014, Vivien Mallet.
- PhD in progress: Ngoc Bao Tran Le, "Quantification d'incertitude par réduction de modèle de dispersion atmosphérique", November 2016, Vivien Mallet.

10.2.3. Juries

- Vivien Mallet for the PhD defense of Michaël Zamo, "Statistical Post-processing of Deterministic and Ensemble Windspeed Forecasts on a Grid", December 2016, Météo France.

10.3. Popularization

Vivien Mallet and Raphaël Ventura, together with Cong Kinh Nguyen (MiMove), Pierre-Guillaume Raverdy (SED), Fadwa Rebhi (MiMove), Fabienne Giboudeaux (Paris city), Awa Ndiaye (Paris city), Gilles Plattner (Particitae) and Laure Turcati (Particitae), took part to collaborative measuring of noise pollution with volunteers from the general public. Using the mobile application Ambiciti, the volunteers carried out a journey in a city district in order to measure noise pollution with their smartphones, and therefore help to map noise pollution in the area. At the end of the journeys, the observations were merged with an existing noise map so as to produce an improved map. Discussion about the strengths and limitations of mobile observation were engaged, and the activity helped to inform about noise pollution. This was carried out in Paris at the opening of the "Canopée des Halles", at Futur en Seine festival and during the "journée sans voiture".

11. Bibliography

Major publications by the team in recent years

[1] M. BOCQUET, P. SAKOV. *An iterative ensemble Kalman smoother*, in "Quarterly Journal of the Royal Meteorological Society", October 2013 [*DOI :* 10.1002/QJ.2236], http://hal.inria.fr/hal-00918488

[2] D. BÉRÉZIAT, I. HERLIN. *Solving ill-posed Image Processing problems using Data Assimilation*, in "Numerical Algorithms", February 2011, vol. 56, n° 2, pp. 219-252 [*DOI :* 10.1007/S11075-010-9383-Z], 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 [*DOI :* 10.1029/2011JD015780], http://hal.inria.fr/hal-00655771

[4] I. HERLIN, D. BÉRÉZIAT, N. MERCIER, S. ZHUK. *Divergence-Free Motion Estimation*, in "ECCV 2012 - European Conference on Computer Vision", Florence, Italie, A. FITZGIBBON, S. LAZEBNIK, P. PERONA, Y. SATO, C. SCHMID (editors), Lecture Notes in Computer Science, Springer, October 2012, vol. 7575, pp. 15-27 [*DOI :* 10.1007/978-3-642-33765-9_2], 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° 19047 [*DOI :* 10.3402/TELLUSB.V64I0.19047], http://hal.inria.fr/hal-00741930

[6] G. K. KOROTAEV, E. HUOT, F.-X. LE 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", April 2008, vol. 112, n° 4, pp. 1464-1475, Remote Sensing Data Assimilation Special Issue [*DOI :* 10.1016/J.RSE.2007.04.020], http://hal.inria.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 [*DOI :* 10.1029/2010JD014259], http://hal.inria.fr/inria-00547903

[8] B. SPORTISSE. *Pollution atmosphérique. Des processus à la modélisation*, Ingénierie et développement durable, Springer-Verlag France, 2008, 350 p. , http://hal.inria.fr/inria-00581172

[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 [*DOI :* 10.1029/2011JD016932], 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 [*DOI :* 10.1029/2008JD009991], http://hal.inria.fr/inria-00582376

Publications of the year

Articles in International Peer-Reviewed Journals

[11] L. ISSA, J. BRAJARD, M. FAKHRI, D. HAYES, L. MORTIER, P.-M. POULAIN. *Modelling surface currents in the Eastern Levantine Mediterranean using surface drifters and satellite altimetry*, in "Ocean Modelling", 2016, vol. 104, pp. 1-14 [*DOI :* 10.1016/J.OCEMOD.2016.05.006], https://hal.inria.fr/hal-01357669

International Conferences with Proceedings

[12] D. BÉRÉZIAT, I. HERLIN, Y. LEPOITTEVIN. *Ensemble Kalman Filter based on the image structures*, in "International Conference on Computer Vision Theory and Applications", Porto, Portugal, February 2017, http://hal.upmc.fr/hal-01414786

[13] P. EBERHART, J. BRAJARD, P. FORTIN, F. JÉZÉQUEL. *Estimation of Round-off Errors in OpenMP Codes*, in "IWOMP 2016 - 12th International Workshop on OpenMP", Nara, Japan, Lecture Notes in Computer Science, Springer International Publishing, October 2016, vol. 9903, pp. 3-16 [*DOI :* 10.1007/978-3-319-45550-1_1], https://hal.inria.fr/hal-01380131

[14] V. ISSARNY, V. MALLET, K. NGUYEN, P.-G. RAVERDY, F. REBHI, R. VENTURA. *Dos and Don'ts in Mobile Phone Sensing Middleware: Learning from a Large-Scale Experiment*, in "ACM/IFIP/USENIX Middleware 2016", Trento, Italy, Proceedings of the 2016 International Middleware Conference, December 2016 [*DOI :* 10.1145/2988336.2988353], https://hal.inria.fr/hal-01366610

[15] Y. LEPOITTEVIN, I. HERLIN. *Regularization terms for motion estimation. Links with spatial correlations*, in "VISAPP - International Conference on Computer Vision Theory and Applications", Rome, Italy, Springer Verlag, February 2016, vol. 3, pp. 458-466 [*DOI :* 10.5220/0005712104560464], https://hal.inria.fr/hal-01235718

[16] V. MALLET, I. HERLIN. *Quantification of uncertainties from ensembles of simulations*, in "International Meeting "Foreknowledge Assessment Series"", Paris, France, December 2016, https://hal.inria.fr/hal-01399175

Other Publications

[17] I. HERLIN, E. HUOT. *Image-Based Forecast of the Surface Circulation of Black Sea*, September 2016, Black Sea from Space Workshop, Poster, https://hal.inria.fr/hal-01349391