Team clime

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

Rocquencourt
# Table of contents

1. Team
2. Overall Objectives
   2.1. Overall Objectives
3. Scientific Foundations
   3.1. Environmental images and data processing
   3.2. Data assimilation and inverse modelling
   3.3. Software chains for environmental applications
4. Application Domains
   4.1. Panorama
   4.2. Air quality
   4.3. Oceanography
   4.4. Remote Sensing for natural risks and agricultural monitoring
5. Software
   5.1. Polyphemus
6. New Results
   6.1. Estimation of hail damages from SPOT images.
   6.2. Large scale classification and change detection using MODIS data
   6.3. Plume analysis and detection in satellite images
   6.4. Fluid flow estimation using vector splines
   6.5. Image data assimilation in an oceanographic circulation model
   6.6. An optimal control approach to compute velocity field from SST image sequence
   6.7. Microcanonical multifractal formalism for the analysis of geophysical flows.
   6.8. Geophysical flows motion in the framework of microcanonical multifractal formalism
   6.9. A multifractal analysis software: JFluid
   6.10. Information theory-based inverse modelling of tracers
   6.11. Inverse modelling of mercury over Europe
   6.12. Inverse modelling of accidental sources of radioactivity at European scale
   6.13. A priori uncertainty in chemistry-transport models
   6.14. Sensitivity of ozone forecasts to emissions over Europe
   6.15. Ensemble forecast applied to ozone
7. Contracts and Grants with Industry
   7.1. LYNX
   7.2. IRSN
   7.3. INERIS
8. Other Grants and Activities
   8.1. National initiatives
   8.2. European initiatives
   8.3. International initiatives.
   8.4. Visiting scientists
9. Dissemination
   9.1. Leadership within scientific community.
   9.2. Teaching.
   9.3. Conference and workshop committees, invited conferences.
10. Bibliography
1. Team

Head of project-team
Isabelle Herlin [Research Director (DR) INRIA]

Vice-head of project team
Bruno Sportisse [ENPC]

Administrative assistant
Christine Anocq [Secretary (SAR) INRIA]
Karine Barrès [ENPC]

Staff member
Jean-Paul Berroir [Research Associate (CR) INRIA]
Marc Bocquet [ENPC]
Jean-Pierre Issartel [ENPC]
Hussein Yahia [Research Associate (CR) INRIA]

Research scientist
Dominique Béréziat [University of Paris 6]
Jaouad Boutahar [EHTP, Casablanca, Morocco]
Etienne Huot [University of Versailles]

Post-doctoral Fellow
Lin Wu [INRIA, from december 2005]

Technical staff
Zulfukar Arslan [INRIA, from September 2005]
Vincent Picavet [INRIA, until October 2005]
Denis Quélo [ENPC]

Ph.D. student
Till Isambert [INRIA]
Milton Jonathan [University of Estado do Rio de Janeiro, from March 2005]
Monika Krysta [ENPC]
Vivien Mallet [ENPC]
Yelva Roustan [ENPC]

Student intern
Milton Jonathan [University of Estado do Rio de Janeiro, until March 2005]

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 also 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) 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 programmes such as the Air Chemistry programme, which will give a long term structure to environmental research. These initiatives emphasize the importance of observational data and also the exploitation of satellite acquisitions.

The complexity of the environmental phenomena, as well as the operational objectives, necessitate a growing interweaving between physical models, data processing, simulation and database tools.
This situation is met for instance in atmospheric pollution, an environmental domain whose modelling is gaining a widening importance, either at small (air quality), regional (transboundary pollution) or global scale (greenhouse effect). In this domain, modelling systems are used for operational forecast (short or long term), detailed case studies, impact studies for industrial sites, management of different spatial and temporal scales, coupled modelling (e.g. pollution and health, pollution and economy). These scientific subjects strongly require coupling the model with all available data; these data being either of numerical origin (e.g. models outputs), or coming from raw observations (e.g. satellite acquisitions or information measured on a spatial grid), or obtained by processing and analysis of these observations (e.g. chemical concentrations retrieved by inversion of a radiative transfer model).

The Clime team has been created for studying these questions by joining researchers in data assimilation and modelling from the CEREA laboratory (ENPC, Ecole Nationale des Ponts et Chaussées) and INRIA researchers in environmental data and image processing. The Clime team carries out research in three directions:

1. Environmental data processing, notably satellite data, by means of computer vision techniques and by accounting for the physical information on the acquisition process and on the dynamic of the observed phenomena.
2. Data and model coupling, by means of data assimilation techniques and related issues (optimization problems, targeting observation, uncertainties propagation, ...).
3. Integrated chains of data/models/outputs (system architecture, workflows, databases, visualisation, ...).

3. Scientific Foundations

3.1. Environmental images and data processing

Keywords: image assimilation, inverse problems, matching, monitoring, motion, tracking.

The overall objective is the extraction of dynamic information from sequences of images -and notably satellite images-, for provision to environmental forecast or monitoring systems. Extraction of information must satisfy the following constraints:

1. All the available physical information should be accounted for. This information concerns the specifically observed phenomena (e.g. evolution laws) and the process of image formation (e.g. radiative transfer and sensor models).
2. The information is extracted from images in view of data assimilation within a numerical model: the nature of information, as well as the process used to extract it, must therefore be compatible with data assimilation.

The dynamic information to be extracted can be classified into two main groups:

1. Dynamic structures: structures moving on image sequences. This is for instance the case of fluid structures in meteorological and oceanic image sequences (clouds, fronts, eddies, filaments). These structures are tracers of the fluid flow, and hence their apparent motion is an information that can be assimilated in an environmental forecast system. The fluid and turbulent nature of the involved flows necessitates the design of specific image processing methodologies for apparent motion estimation, tracking of deformable structures, trajectory estimation, turbulence characterization. The investigations concern the nature of the physical information to be used for dynamic structures extraction. The two following alternative approaches are studied: establishment of visual description models of images, use of the physical evolution and transport equations, expressed in the model space or in an image space to be defined.
• Parameters: evolution in time of a physical parameter. Environmental models frequently require a physical parameterisation: for instance, chemistry-transport and radiative transfer models require energy budget parameters (emissivity, albedo), agronomical models require vegetation indices and leaf area index, hydrological models require the knowledge of evapotranspiration. This research falls in the general framework of inverse problems, where a model (either physical or empirical) of image formation is inverted to retrieve the parameter. Addressing this problem in the context of image data permits the design of image-based regularisation techniques, where the information coming from neighboring pixels is used to constraint the inversion process. Finally, monitoring the temporal evolution of the retrieved parameter, and being able to characterise and detect normal and abnormal evolutions, is a strong requirement of environmental monitoring systems.

The effective use of the extracted information by environmental models arises the general problem of image assimilation. Existing approaches are restricted to the assimilation of images, considered as a collection of individual measurements, directly related to state variables. The challenge is the design of methods for assimilating the structured information present in image data and sequences. This requires tackling the following issues:

• Image processing methodologies for extracting structured information, as discussed above.
• Definition of mathematical spaces for characterising the structured information. Different levels of abstraction (e.g. a curve can be seen as a collection of points, or as a parameterised curve, or as a level set) must be considered, each carrying a different amount of information.
• Definition of observation operators mapping the model’s state space to the image space.
• Definition of a norm, or of a family of norms, in the image space. This norm is used to assess the discrepancy between model’s results and image observations.

3.2. Data assimilation and inverse modelling

Keywords: Data assimilation, inverse modelling.

This activity is one of the present major stake in environmental sciences. It matches up the setting and the use of data assimilation methods, notably variational methods (4D-var). An emerging point lies in uncertainties propagation in models, notably through ensemble prevision methods.

Although modeling is not part of the scientific objectives of the Clime team, we have access to models developed by the CEREA (joint ENPC/EDF R&D laboratory): Polair3D (photochemical pollution forecasting at continental and regional scales) and MERCURE (urban scale).

The research activities tackle scientific issues such as:

• Which observational network must be set up for performing a better forecast, taking into account additional criteria such as observation cost? What are the optimal location, type and mode of deployment of sensors? How to operate the trajectories of mobile sensors, 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 does 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?
• Among a family of models (differing by their physical approximations or their discretization parameters), what is the optimal model for a given set of observations?
• How to perform forecast (and a better forecast!) by using several models coming from different institutes, or different parameterizations (corresponding to different physical configurations) of the same model? In both cases this is a set of models and it raises the question: how to assimilate data in this context?
3.3. Software chains for environmental applications

**Keywords:** Database, system architecture, visualisation, workflow.

Building on the scientific capabilities developed in the previous directions, the ambition of the Clime project lies in the participation to the design and realization of software chains for impact assessment and environmental crisis management. Such software chains put together static or dynamic databases, data assimilation systems, forecast models, processing methods for images and environmental data, complex visualization tools, scientific workflows...

The Clime project is currently building such a system for air pollution modelling: Polyphemus (see Internet page [http://www.enpc.fr/cerea/polyphemus](http://www.enpc.fr/cerea/polyphemus)), which architecture is specified to account for the data requirements (e.g. various raw data nature and sources, data preprocessing) and the different usages of an air quality model (e.g. forecast, data assimilation, ensemble runs).

4. Application Domains

4.1. Panorama

The priority application domain is atmospheric chemistry, since this is the research topic of CEREA. We have at our disposal the 3D Eulerian chemistry transport model Polair3D and its adjoint, thus capable of data assimilation. Atmospheric data processing will aim at providing Polair3D with relevant input data.

A second application domain is oceanography (supported by the ASSIMAGE ACI). The methodology is comparable: providing oceanic circulation models with relevant information, extracted from satellite images. This information is ingested within models using data assimilation techniques. As numerous sources of satellite images are relevant, image assimilation becomes potentially interesting for oceanography.

Finally, Clime carries out applied studies in close cooperation with environmental experts. These studies concern applications of remote sensing for monitoring agriculture and natural risks. Application domains are monitoring of land cover changes, assessment of hail damages, soil degradation and fire plumes. For these studies, an image model of the observed phenomenon is established, then tracking and change detection techniques are applied in the state space of this model.

4.2. Air quality

Air quality modelling 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 forecast, 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). A model lacks the appropriate data, notably emissions, for performing an accurate forecast and data assimilation techniques are recognised as crucial for the improvement of forecast’s quality. These techniques, and notably the variational ones, are barely surfacing in atmospheric chemistry, and they are used for non linear systems of very large dimension.

In this context, the Clime team is interested in different problems:

- Definition of second order data assimilation for the design of optimal observation networks. Management of combinations of sensor types and deployment modes. Dynamic management of mobile sensors trajectories.
- Development of ensemble forecast methods for estimating the quality of the prediction in comparison to the quality of the model and of the observations. Sensitivity analysis with respect to model’s parameters so as to identify physical and chemical processes, whose modelling must be improved.
- Development of methodologies for super-ensemble forecast (different models, or different configurations of the same model). Investigation on how super-ensembles must be generated, with how many members and with which constraints.
The activities of the Clime team in air quality are supported by the development of the Polyphemus air quality modelling system. This system has a modular design, which makes it easier to manage high level applications such as inverse modelling, data assimilation, ensemble forecast.

4.3. Oceanography

This application domain is studied in the framework of the ASSIMAGE ACI involving the INRIA teams IDOPT, Clime and VISTA, the CNRS laboratories LGGE and LEGI, and CEMAGREF. Research works in the continuation of the former ARC Thalweg also support this application through strong collaborations with the Institute of Marine Sciences in Barcelona, Spain.

The main motivation is to extract from satellite images structured (in space and time) information that can be related to the state variables of oceanic circulation models, hence allowing for data assimilation. We are currently investigating the following issues:

- surface motion estimation from Sea Surface Temperature and ocean color measurements, directly related to a state variable (current),
- trajectories estimation by tracking specific structures in satellite sequences, requiring the use of Lagrangian data assimilation techniques,
- turbulence characterization.

4.4. Remote Sensing for natural risks and agricultural monitoring

This domain concerns specific case studies on crucial environmental topics, involving strong cooperation with environmental experts and access to multiple data sources (satellite images, ground truth, validation data, models). Currently the Clime team is involved in four activities:

- Soil degradation: this study is led in cooperation with the State University of Rio de Janeiro and Embrapa Solos (Brazil), in the framework of the INRIA-CNPq project ENVIAIR. AVHRR and MODIS image sequences are used to assess the changes of land use and land cover in the Pantanal area (Brazil), for monitoring of soil degradation, erosion, deforestation and sustainable agricultural practices. The Clime team will also be involved as external partner to a wide project funded by the Inter American Institute for Global Change Research, which aims at developing land use monitoring systems for the Rio de la Plata basin, the largest South American basin (3,000,000 square kilometres).
- Land cover changes and early classification: this study is led in cooperation with CESBIO. It aims at assessing the potential of SPOT sequences for the detection of land use changes in agricultural areas, and for obtaining an early land use classification with respect to the vegetation cycle.
- Hail damage assessment: this study is led for the LYNX company. It aims at defining a method for the observation and quantification, from satellite images, of damages caused by hail to vineyards. The potential of the method is assessed by comparison to insurance expertises.
- Fire plume detection: this study is led in cooperation with FORTH-IACM (Greece), and aims at the definition of methods for fire plumes detection from multispectral satellite images (NOAA-AVHRR, METEOSAT). The investigation concerns the definition of spatial and temporal constraints for improving the robustness of pixel-based detection methods.
5. Software

5.1. Polyphemus

Polyphemus is a complete simulation system for air quality modeling. It is designed to yield up-to-date simulations in a reliable framework: aerosols, data assimilation, ensemble forecast and daily forecast. Its completeness allows to deal with many fields: photochemistry, heavy metals, radionuclides, etc. It is able to handle simulation from regional to continental scale. Its main components are:

- A flexible chemistry-transport model, Polair3D.
- A library that brings up-to-date parameterization, AtmoData.

Polyphemus can be downloaded from the web: http://www.enpc.fr/cerea/polyphemus.

6. New Results

6.1. Estimation of hail damages from SPOT images.

**Keywords:** change detection, remote sensing.

**Participants:** Isabelle Herlin, Jean-Paul Berroir, Zulfukar Arslan.

This study is an expertise led for the LYNX company. Its objective is the evaluation of the ability of satellite images to assess damages caused to vineyards by hail events, these damages being quantified as the production loss in percentage per parcel.

A pair of SPOT images, acquired approximately one month before and after a hail event, are used for the investigation. Validation data consist in insurance expertises after the hail event, occured in May 2004 in the area of Lezignan, in South-West of France.

We have first investigated the adequation of various vegetation indices for the observation of hail damages: ratio-based index (NDVI), soil corrected indices (SAVI and derived indices), atmospheric resistant index (GEMI). The average of the relative variation for the chosen vegetation index is computed for each parcel and compared to insurance expertise.

This study proved that parameters computed from SPOT images allow distinguishing hail damaged parcels from parcels outside of the hail event’s surface. A good linear correlation is observed between the relative variation of the vegetation index over one parcel and the insurance expertise, enabling the establishment of an algorithm for the assessment of hail damages. Further studies are being carried out to analyse additional events occuring at different dates and corresponding to different phenological stages.

6.2. Large scale classification and change detection using MODIS data

**Keywords:** change detection, classification, deforestation, erosion.

**Participants:** Jean-Paul Berroir, Isabelle Herlin, Milton Jonathan, Margareth Mereilles [State University of Rio de Janeiro].

This study is carried out in the framework of the ENVIAIR project (INRIA-CNPq project in cooperation with UERJ and EMBRAPA, Brazil) and was supported by an European Commission’s Alβan master scholarship.

The objective is to assess the potential of low resolution satellite data (temporal sequences of MODIS visible data) for detecting land use changes and degradation on potentially very large areas. The test site is the Taquari basin in Mato Grosso (Brazil), for which various validation data are available (field campains, high resolution satellite images and their classification).

The study builds on two phases: training, and simulation of operational use. Training consists in the following three steps:
Data preprocessing. We have at our disposal a classification of the test site performed on a Landsat-TM image (July 2001), and a sequence of MODIS data acquired from August 2000 to July 2001. This sequence contains one image per day at resolution of 250 meters. Due to cloud coverage and to important viewing angle, only 60 images are usable. This new sequence is co-registered with the Landsat classification and filtered over time.

Learning. The temporal behavior of the main land cover types is learned on a part of the Taquari basin, which is used as learning area. The Landsat classification serves to identify pure MODIS pixels and extract their temporal profiles. These profiles are then fitted by polynomial curves in order to filter the noise affecting the signal (due to acquisition and atmosphere contribution). The coefficients of these polynomials define a vector space \( V \), where temporal profiles are represented.

Test. Learning results are then validated on the remaining part of the basin: classification is performed in the vector space \( V \), yielding an overall accuracy of 84%, suitable for operational large scale classification. Furthermore, the analysis of results shows that deforestation is characterised by an abnormal decrease of NDVI throughout the MODIS sequence. This quantitative effect is then used to detect deforestation (see figure 1) and is usable as an environmental alert criteria.

A potential operational use consists in applying the previous low resolution classification method on large areas bearing the same biomes as the Taquari basin: a follow-up of this study actually consists in applying the technique to the Rio de la Plata basin, which superficy (3,000,000 square kilometers) is approximately 6 times the superficy of France.


6.3. Plume analysis and detection in satellite images

Keywords: fire plumes, multispectral image, natural risk.

Participants: Isabelle Herlin, Hussein Yahia, Vincent Picavet, Nektarios Chrysoulakis [FORTH-IACM], Jacopo Grazzini [FORTH-IACM].

This study concerns the detection on NOAA-AVHRR images of plumes, caused by fires affecting forests or industrial installations. It is led in the framework of the PLUMESA T project (PLATON cooperation programme between France and Greece) in collaboration with FORTH-IACM, Heraklio, Greece.

A plume detection program, JPlume, has been developed. It integrates methods devised by the researchers at FORTH, which combines two indices for characterizing plumes: NDVI (Normalised Differential Vegetation Index), and CLD, defined as:
CLD = \frac{T_5 - R_1}{T_5 + R_1}

where \( T_5 \) is the brightness temperature in the fifth AVHRR channel, and \( R_1 \) the reflectance in the first AVHRR channel. A plume is detected as having low NDVI and high CLD, i.e., as being not vegetated, hot, and poor scatterer of visible light. Detection is illustrated on figure 2, where CLD is assigned to the red channel, NDVI to the green channel, and 0 to the blue channel. With such a color combination, plumes appear red.

![Figure 2. Plume detection, forest fire in the area of Madrid, Spain, July 16th 2005.](image)

The plume detection method has been improved in order to minimize the rate of false alarms. Additional features include:

- The brightness temperature acquired in the third AVHRR channel (mid infrared).
- Spatial indicator: for example the local variance of \(|R_2 - R_1|\), difference between the reflectances in the two visible AVHRR channels.

### 6.4. Fluid flow estimation using vector splines

**Keywords:** fluid motion, multiscale, optical flow, quasi-interpolation, radial basis functions, reproducing kernels, turbulence, vector splines.

**Participants:** Till Isambert, Jean-Paul Berroir, Isabelle Herlin, Christine Graffigne [Université Paris5].

This study concerns estimation of apparent motion on sequences of oceanographic and meteorological satellite images. The challenge is to establish methods adapted to turbulent flows, characterised by strong rotational patterns and by the presence of structures with very different spatial and temporal scales.

Our approach is based on the use of two models for thin-plate splines: interpolating or approximating vector splines. Both models allow controlling the rotational and the divergence of the motion field. However the first type of spline exactly interpolate motion information at selected control points, whereas the second realises a compromise between regularity and confidence in motion information at these control points. In both cases, thin plate splines rely on a reproducing kernel, which is based on the derivatives of \( K(r) = r^4 \log r \) (\( r \) being the distance to control points). This kernel grows arbitrary large with \( r \), and this causes two main inconveniences: first, the modification of a single control point impacts the whole spline, and second, the kernel is not suitable for a multiscale representation of the motion field.

We have reformulated these models in the framework of quasi-interpolation, a technique realising an approximation of the thin-plate vector spline, by considering a finite difference implementation of the underlying Partial Differential Equation. Quasi-interpolation has the advantage of introducing bell-shaped reproducing kernels \( \Psi \), suitable for multiscale representation. The vector spline \( w \) is expressed as a linear
combination of the kernels $\Psi$ translated around control points $r_i$, where the vector spline $w$ equals $w_i$. A scale factor $h$ is introduced to control the dilation of the kernel $\Psi$. $h$ is related to the spacing between control points.

$$w(x) = \sum_i w_i \Psi(\frac{x - r_i}{h})$$  \hspace{1cm} (1)

A methodology has then been devised for using the quasi-interpolation technique to estimate the apparent motion $w$. This methodology requires first computing the motion vectors $w_i$ at control points $r_i$. The classical aperture problem explains that only the projection of motion is available from image data. Expressing these projections for all the control points results in a linear system. This system is inverted to yield the motion vectors $w_i$. Then the quasi-interpolation method estimates motion using a regular grid of control points, which spacing is controlled by the scale parameter $h$. At each scale, the spline is expressed as in equation 1, allowing to estimate apparent motion at the same scale, as seen in figure 3. Current work investigates the adequacy between this multiscale representation and the different spatial scales of the turbulent structure.

![Figure 3. Motion estimation on a synthetic oceanographic sequence, using quasi-interpolant technique with div-curl regularity. Left: scale factor=100. Right: scale factor=40.](image)

6.5. Image data assimilation in an oceanographic circulation model

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

**Participants**: Etienne Huot, Isabelle Herlin, Gennady Korotaev [MHI], Jean-Paul Berroir, Till Isambert.

Image assimilation concerns the assimilation of data, obtained from remote sensing acquisitions, within geophysical fluid forecast models. A good prediction requires simultaneously a model to describe the evolution of a state variable (generally it is a non-linear system of Partial Differential Equations), and observation distributed in space and time. Data assimilation provides a mathematical solution to combine these data and the model.

In this study, we are interested in oceanographic circulation forecasting. We tackle the problem of coupling an oceanographic circulation model with satellite images. The model is a shallow-water oceanographic circulation model developed at the Marine Hydrophysical Institute of Sevastopol (Ukraine). Image data are Sea Surface Temperature (SST) image sequences, acquired by NOAA/AVHRR sensors over the Black Sea. The coupling is performed by an optimal processing of the images, that defines an image observation space, and a data assimilation scheme relying on an operator, that maps the image space into the model’s state space.
The image processing phase consists in two steps: - preprocessing to remove different artefacts due to clouds, sensor calibration, missing data, ..., - apparent motion estimation on SST images. This apparent motion is considered equivalent to the shallow-water velocity in order to use it as velocity observation at each point of the surface. The data assimilation phase, based on a simplified Kalman filtering method, improves the model’s forecast by assimilating the velocity observation.

Figure 4 illustrates the difference of forecast results with and without assimilation of the velocity observations. Assimilation allows the forecast of significant small-size structures: broad currents on the left image are replaced by narrow meandering jets. The actual size of mesoscale structures is also better represented: the gyre in the western part of the right image has a smaller dimension with respect to the left.

6.6. An optimal control approach to compute velocity field from SST image sequence

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

**Participants:** Etienne Huot, Isabelle Herlin, Gennady Korotaev [MHI], François-Xavier Le Dimet [INRIA-IDOPT].

Oceanographic models forecast the motion fields on a daily basis. On the other hand, surface temperature satellite data allow the observation of the surface motion. The objective is to estimate motion from these image sequences for assimilating it within the model.

Classical velocity estimation methods rely on a conservation equation such as gray level value conservation. This single equation is not sufficient to compute velocity and an additional equation is required to close the system. This equation is generally obtained by a regularity assumption on the velocity field.

The main idea is to define the regularity constraint from the equations of the oceanographic model. First, a simplified transport model of the temperature is introduced: 

\[
\begin{align*}
\frac{\partial T}{\partial t} + \nabla T \cdot \mathbf{w} &= K_T \Delta T \\
\frac{\partial \mathbf{w}}{\partial t} &= 0.
\end{align*}
\]

The first equation expresses the advection/diffusion of the temperature, while the second corresponds to the frozen velocity assumption: a particle of water is assumed to move with a constant velocity. The state vector of this model is \((T, \mathbf{w})\) where \(\mathbf{w}\) is the motion vector and \(T\) the Sea Surface Temperature. Observations consist in Sea Surface Temperature images acquired by NOAA-AVHRR. Variational data assimilation of these temperature observations is performed within model 2 to compute the motion field \(\mathbf{w}\) and assimilate it within...
the oceanographic model. This approach is able to manage the missing data problem, which is often occurring when clouds prevent the observation of the sea surface by satellite.

6.7. Microcanonical multifractal formalism for the analysis of geophysical flows.

**Keywords:** multifractal, oceanography, stream function, turbulence.

**Participants:** Hussein Yahia, Antonio Turiel [Institute of Marine Science], Jordi Font [Institute of Marine Science].

The canonical framework has been the most widely used approach to study multifractal signals. It is of statistical nature and in that formalism, each signal \( s \) must be regarded as a particular realization of a random variable, which makes sense when analyzing, for instance, the fully developed turbulence, where all the invariances (translational, scale) are restored only in a statistical sense. In addition, systems studied under the canonical framework are usually ergodic or, as a weaker requirement, they have stationary expectation values, allowing to exchange evaluating averages over different realizations with evaluating averages over different positions in the same realization. However, the introduction of such stationarity hypothesis may cause some problems (mainly, because the system is not stationary) in practical situations and so stationarity should be carefully checked as a pre-requisite. In this context, we are currently working on a geometrical approach, called microcanonical multifractality.

Microcanonical multifractality is observed if there exists local power-law scaling behaviour for each point in the signal domain. The signal is processed by means of an appropriate local functional, which serves to unveil the correlation degree of the signal at a given point and across variable distances. We will say that a signal \( s(\vec{x}) \) has microcanonical multifractality if, for a given family of functionals \( T_r \), the following relation holds for any point \( \vec{x} \):

\[
T_r s(\vec{x}) = \alpha_T(\vec{x}) r^{h(\vec{x})} + o\left(r^{h(\vec{x})}\right)
\]

for some functions \( \alpha_T(\vec{x}) \) and \( h(\vec{x}) \). Although the multiplicative term \( \alpha_T(\vec{x}) \) depends on the particular functional chosen, the exponent \( h(\vec{x}) \) should be independent of it. Like the source vector field, the stream function is recovered from the reduced signal, which is the signal reconstructed from the Most Singular Manifold (MSM) with a normalized gradient, perpendicular to the MSM.

Flows under fully developed turbulence advect (i.e., parallel transport) their multifractal structure. Under such hypothesis we can conclude that each fractal manifold is composed by instantaneous stream lines. This property has been used to propose a method to derive the quasi-geostrophic instantaneous stream function from the indirect measure of an easy-to-measure scalar, namely the Sea Surface Temperature (SST). From these results, we are further investigating the properties of the recovered stream function, notably by examining datasets coming from different sensors (e.g.: altimetry).

In Figure 5 we show some examples on the application of the Maximal Singular Stream (MSS) methodology to real SST data. The method can be applied to other tracers as for instance the chlorophyll concentration. The results must be independent of the tracer, so the combined information of several tracers could be used to reduce errors.

6.8. Geophysical flows motion in the framework of microcanonical multifractal formalism

**Keywords:** fluid flow, motion, multifractal formalism, turbulence.

**Participants:** Hussein Yahia, Antonio Turiel [Institute of Marine Science].

The microcanonical multifractal formalism opens new ways of studying motion in acquired time series of geophysical flows, like the ocean or the atmosphere. This is very interesting because motion estimation and
Figure 5. **Top left:** Monthly average, August 2005, of Sea Surface Temperature, measured by MODIS AQUA satellite. **Top right:** Associated Maximum Singular Stream-functions. **Bottom:** Velocities evaluated from the MSS.
analysis in the case of geophysical fluid flows, and more generally in the case of fully developed turbulence is very challenging. Since singularity exponents are related to the thermodynamical properties of the flow, and since the microcanonical multifractal formalism, being a geometrical formulation, allows the determination of physically meaningful geometric structures, new ways of inferring motion properties become available, for instance by matching similar singularity exponents in consecutive frames. The key point is the relation of the Most Singular Manifold with the stream function.

In Figure 6 we show how the singularity exponents allow the possibility of inferring interesting motion information in turbulent signals. The image on the left displays an infrared Meteosat acquisition of a vortex. Singularity exponents are computed on three consecutive images in the sequence: red at time $t - 1$, green at time $t$, blue at time $t + 1$. The result, displayed in the image on the right, clearly displays the winding structures inside the vortex. Near the vortex center, "blue" exponents (which are the most "recent" in the time acquisition) are clearly distinguishable from their red counterparts, these latter exponents being also located farther, hence making a clear description of the motion inside the vortex.

![Figure 6. Left: An infrared Meteosat acquisition. The red rectangle shows the selected area, centered on a large vortex. Right: Result of the singular exponents computation. These two images are displayed with the Jfluid software.](image)

6.9. A multifractal analysis software: Jfluid

**Keywords:** Java programming language, microcanonical multifractal formalism, user interface.

**Participants:** Hussein Yahia, Antonio Turiel [Institute of Marine Science].

The microcanonical multifractal formalism is currently implemented in a set of software tools, called **Jfluid**. The software tools are written in Java for platform independence. The planned features of **Jfluid** are the following:

- It has an intuitive User Interface (UI), allowing the computation of various quantities and objects associated with the microcanonical multifractal formalism (singularity exponents, manifolds, stream function, etc.).
- Core classes are separated from the UI and are callable from outside by third party-developers.
- It allows to work on multiple files in a same session, and to read data in the different formats used in the geophysical community.
- Easy development of third-party modules.
• Distributed computation on large datasets.

**Jfluid** is intended to be distributed in the scientific community to promote and diffuse the microcanonical multifractal formalism as a new tool for studying datasets in geophysical sciences. The images displayed in Figure 6 were generated using **Jfluid**.

### 6.10. Information theory-based inverse modelling of tracers

**Keywords:** atmospheric dispersion, grid resolution, inverse problem, lagrangian duality, maximum entropy.

**Participant:** Marc Bocquet.

Over the past years, it turned out to be of considerable importance to trace back the source of chemical species dispersed through the atmosphere, with increasing precision in the source resolution. We have studied the high-resolution retrieval at continental scale of the source of an atmospheric passive tracer, given a set of concentration measurements. Theoretical grounds have been developed for this reconstruction. The approach is based on information theory, and in particular the principle of maximum entropy on the mean. It offers a general framework in which the information input prior to the inversion is used in a flexible and controlled way. The inversion is shown to be equivalent to the minimisation of an optimal cost function, expressed in the dual space of observations. Examples of such cost functions are given for different priors of interest to the retrieval of an atmospheric tracer. Also this framework has been enlarged to incorporate noisy data in the inversion scheme.

A numerical score has been proposed to quantify the quality of inversions. This indicator generalises the root mean square indicator used in air quality models to priors that are not necessarily gaussian.

Using this score, the influence of the grid resolution on the source reconstruction in an inversion has been investigated both analytically and numerically.

Generalisations of the method to problems with thresholds, which make use of a Lagrangian duality (nonlinear convex analysis), have been proposed but not tested yet.

### 6.11. Inverse modelling of mercury over Europe

**Keywords:** boundary conditions, inverse problem, mercury dispersion.

**Participants:** Marc Bocquet, Yelva Roustan.

Adjoint techniques have been used to perform a sensitivity analysis on a regional Eulerian transport model devoted to mercury. We have shown how to relate rigorously and explicitly the modelled measurements to the forcing fields. Because gaseous elemental mercury is a long-lived specy in the atmosphere, boundary conditions must be properly taken into account. Ground measurements of gaseous mercury are very sensitive to the uncertainties attached to those forcing conditions.

Since the inverse problem is clearly defined, inverse modelling has been employed to constrain the forcing fields and improve the predicted mercury concentrations. More generally, it allows to reduce the weaknesses of a regional model against a global or hemispherical model. Using EMEP (Cooperative Program for Monitoring and Evaluation of the Long Range Transmission of Air Pollutants in Europe) measurements of gaseous mercury and performing the inversion, it has been shown that boundary conditions can be improved significantly as well as the forecasted concentrations. Using inverse modelling to improve the emission inventory is however much more difficult since there are currently not enough mercury monitoring stations, and their locations are far from Europe centre.

### 6.12. Inverse modelling of accidental sources of radioactivity at European scale

**Keywords:** inverse modelling, maximum entropy, radionuclide dispersion.

**Participants:** Marc Bocquet, Monika Krysta.
New source reconstruction techniques have been put forward, developed and applied to experimenting upon fictive releases of radioactivity to the atmosphere. The method has been tested on a set of observing system simulation experiments, involving all the European nuclear sites as candidate sources. Prior knowledge that one might possess in an accidental situation has been identified. Information theory based methods allow to construct purposeful cost functions consistent with this prior knowledge. The newly designed cost functions generalise the classical variational assimilation techniques. The solutions to the adjoint of the numerical dispersion model Polair3D have been employed to construct those cost functions efficiently. Several experiments have been performed to demonstrate source retrieval sensitivity to various factors and among them the prior knowledge, meteorological conditions, source profile shape and errors. Finally, a reconstruction is attempted on the Algeciras incident that took place in the south of Spain in May 1998. Although a truly successful inversion seems out of reach due to scarcity of the available observations, model error reduction might lead to satisfying reconstructions.

6.13. A priori uncertainty in chemistry-transport models

**Keywords:** air quality, input data, model formulation, numerical approximation, ozone, uncertainty.

**Participants:** Vivien Mallet, Bruno Sportisse.

Little work has been performed with respect to uncertainty estimates in full air-quality models. Models are usually assessed according to their discrepancy to observations. We investigate the a priori uncertainties of ozone forecasts, that is, the spread of possible computed-concentrations by air-quality models.

The uncertainties in air-quality forecasts come from the uncertain input data, the numerical approximations and the choices in the model formulation. With 800 Monte Carlo simulations, we propagate the main uncertainties in the input data, except meteorological data. The relative uncertainty due to the input data is about 8% over the studied period.

Comparisons between different numerical schemes allow us to assess the impact of numerical schemes and therefore their potential resulting uncertainties. Although the advection scheme and the time step may have a strong impact on the results, the numerical schemes appear to have a low impact on ozone concentrations.

Based on the flexibility of the modelling system Polyphemus, we perform ensemble simulations with changes in the model formulation (notably physical parameterisations). Output concentrations appear to be highly dependent upon the model formulation, with a relative uncertainty above 17% (relative standard deviation). Such a result shows that most air-quality applications should include ensemble simulations.

6.14. Sensitivity of ozone forecasts to emissions over Europe

**Keywords:** Monte Carlo, adjoint, sensitivity, tangent linear.

**Participants:** Vivien Mallet, Denis Quélo, Bruno Sportisse.

In order to prepare inverse modelling of emissions at European scale, with ground observations of ozone, a detailed sensitivity study has been performed. The primary goal is to find out the best control parameters for an inverse modelling experiment based on ozone ground-observations.

Three techniques are used to accurately analyse the sensitivities. The tangent linear model, obtained with the automatic differentiation tool Odyssée, allows us to describe the sensitivity of ozone concentrations with respect to the main European sources.

The adjoint model, still built with Odyssée, is used to compute the sensitivity of ozone concentrations in given cells, e.g. at monitoring stations, with respect to all European emissions. The highest sensitivities are found close to the sources, in regions with small extent. The time evolution of the sensitivities confirms the very local impact of emissions.

Monte Carlo simulations have been performed to quantify these sensitivities, and to enforce the conclusions of the previous studies. The conclusion is a low sensitivity of ozone concentrations to emissions, even to NO emissions which show the highest sensitivities.
6.15. Ensemble forecast applied to ozone

**Keywords:** ensemble forecast, model combination, ozone forecast.

**Participants:** Vivien Mallet, Bruno Sportisse.

Model developments (physical parameterisations, additional computational resources, etc.) have led to little improvements in ozone forecast because of the high uncertainty in the model. To overtake these uncertainties, data assimilation techniques have been designed. We investigate the potential of ensemble forecast and the linear combinations of members.

Up to 48 members, generated with the system Polyphemus, are included in ensemble simulations during four months over Europe. Several linear combinations of the models show strong potential performances in model-to-data comparisons.

Optimal weights associated with each model are not robust in space and time. In forecast mode, the weights are therefore smoothed over a learning period. Significant improvements are reached with these weights, especially on hourly ozone forecasts. In addition, algorithms from machine learning show promising performances.

7. Contracts and Grants with Industry

7.1. **LYNX**

A research contract is established with LYNX (a company that will be created in 2006). It is a feasibility study, with the objective of assessing hail damages to vineyards from sequences of satellite images.

7.2. **IRSN**

The Clime team is involved in the MIRA project (Inverse Modelling of Atmospheric Wastes) with IRSN (national Institute of Radiological protection and Nuclear Safety) and Ecole Centrale de Lyon. This project deals with small scale inverse modelling of radionuclides sources.

A research contract with IRSN is furthermore established on the coupling of meteorological data and the Chemistry-Transport Model Polair3D for the mesoscale transport of radionuclides.

7.3. **INERIS**

The Clime team is involved in a joint project with INERIS (National Institute for Environmental and Industrial Risks) devoted to air quality forecast. This includes research topics in data assimilation and ensemble modeling.

8. Other Grants and Activities

8.1. National initiatives

The Clime team is involved in the ASSIMAGE project (Ministry Grant), dealing with image data assimilation, in collaboration with two other INRIA teams (IDOPT and VISTA), CEMAGREF and CNRS.

The Clime team is also involved in the ADOQA project (INRIA Investigation Grant) dealing with advanced data assimilation techniques for air quality forecast: non linear assimilation with particle filtering and assimilation of raw satellite measurements. The ADOQA project involves the INRIA teams IDOPT (Rhône-Alpes) and ASPI (Rennes), and INERIS.

We have academic collaborations with the following laboratories:

- Medical Imaging team from the computer science laboratory (LIP6) of the University of Paris 6.
- CETP (Centre d’études des Environnements Terrestre et Planétaires) in Vélizy.
8.2. European initiatives

The Clime team participates to the ERCIM working group “Environmental Modelling”. With members of this group, the Clime team is involved in the following activities: a PAI-Platon research project, PLUMESAT, has been initiated with FORTH-IACM (Heraklio, Greece). This project is co-ordinated by Hussein Yahia. It aims at developing algorithms for fire plumes detection from satellite data, and improving these algorithms by using spatial (texture) and temporal (tracking) indicators. A PAI-Procope project, AIRGRID, has been submitted with Fraunhofer-FIRST (Berlin, Germany). Its objective concerns the porting of air quality models to Grids environments.

A ECO-NET project, ADIMO, has been accepted and will start in 2006. This project is led in cooperation with the IDOPT project (INRIA Rhône-Alpes), the Institute for Numerical Analysis in Moscow (Russia), and the Institute of Oceanography in Sevastopol (Ukraine). The objectives of ADIMO concern the assimilation of images within ocean circulation models.

Hussein Yahia is currently a visiting scientist for one year at the Institute of Marine Science (ICM, Barcelona, Spain), on theoretical advances of the microcanonical multifractal formalism and on the development of the Jfluid toolbox.

8.3. International initiatives.

A research project, ENVIAIR, was accepted 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. This project aims at monitoring soil degradation and sustainable agricultural practices in the Taquari basin (Brasil). The Clime team is also involved, as external partner, in the Rio de la Plata project, funded by the Inter American Institute for Global Change Research. The objective is the monitoring of land use changes in the Rio de la Plata basin, a region large as six times the superficy of France.

A research project, named AIRPOL, which was risen in the INRIA/CONICYT collaboration (Chile), is established with CMM (Centre for Mathematical Modelling, Santiago, Chile). This project aims at studying data assimilation methodologies (ground measures) for inverse modelling of static sources of arsenic. Following this project, the Chilean Meteorological Office will use the Polyphemus system for operational air quality forecasts.

A research project, ADOQA-Conesud, was initiated under the France-Conesud programme, with the University of Cordoba (Argentina), the National Centre for Nuclear Energy (Argentina), CMM (Chile) and the Chilean Meteorological Office (Chile). The objective are inverse modelling of air pollution sources by means of sequential data assimilation procedures.

8.4. Visiting scientists

- Margareth Simões Meirelles - UERJ Brazil - from 02/18 to 02/28/05
- Gennady Korotaev - MHI Ukrain - from 01/10 to 01/12/05 and from 11/13 to 11/21/05
- Ricardo Alcafuz - DMC, Meteorological Office, Chile - from 11/28 to 12/09/05

9. Dissemination

9.1. Leadership within scientific community.

- Isabelle Herlin is member of the commission of specialists for University Paris 12.
- Bruno Sportisse is member of Vivien Mallet’s PhD Jury: “Assessment of uncertainty and ensemble forecast with a Chemistry-Transport Model - Application to numerical simulation of air quality”, defended on Dec 6th, 2005 at ENPC.
- Marc Bocquet is member of Yelva Roustan’s PhD Jury: “Modelling of atmospheric dispersion of mercury, lead and cadmium at the European scale”, defended on Dec 12th, 2005 at ENPC.
9.2. Teaching.
- Data assimilation and inverse modeling: ENSTA, 42 hrs (Marc Bocquet and Bruno Sportisse and Vivien Mallet).
- Man-Machine interfaces: Leonard De Vinci University, 20 hrs (Jean-Paul Berroir).
- C/C++ programming: courses and tutorial class: ISTM engineering school, 35 hrs (Jean-Paul Berroir and Till Isambert).
- Multimedia: ISTM engineering school, 36 hrs (Isabelle Herlin).
- Algorithms: Leonard De Vinci University, 20 hrs (Isabelle Herlin).

9.3. Conference and workshop committees, invited conferences.
- Isabelle Herlin is member of the scientific program committee of the conferences: IGARSS’05, ICIP’05, RFIA’06, ICASSP’06.
- During the STIC-AmSud seminar, organised in Santiago de Chile in December 2005, Isabelle Herlin presented land degradation assessment from satellite data, and Jean-Paul Berroir presented inverse modelling of arsenic sources.
- Etienne Huot: Participation to the PaRISTIC days. Presentation of the studies done in the ACI (Ministry Grant) ASSIMAGE (Image Data Assimilation within Simulation Models for Geophysical Fluids).

10. Bibliography

Major publications by the team in recent years


**Doctoral dissertations and Habilitation theses**


**Articles in refereed journals and book chapters**


**Publications in Conferences and Workshops**


Internal Reports