Chapter 8

Applications in environmental sciences

8.1 ▪ Physical oceanography

8.1.1 ▪ Presentation of the domain

After numerical weather forecasting, oceanography is the second most active application domain for DA. It started in the late 1980s, following the atmosphere community. Specific difficulties caused a 10- to 20-year delay between atmosphere and ocean assimilation, in particular: cost and complexity of data acquisition (especially at depth), presence of complex boundaries, flux modeling at the air–sea boundary, interactions with chemistry and biology, model errors, and strong turbulence. The arrival of satellites produced a rapid increase in data quantity and quality and drastically improved both models and assimilation systems for the ocean. The dynamism of the field is such that an extensive review is out of reach—we therefore propose a selection of significant questions and contributions.

8.1.1.1 ▪ Global assimilation and operational forecasting

In the late 1990s, the Global Ocean Data Assimilation Experiment (GODAE) was established. The idea was to use state-of-the art ocean models and combine them with similarly advanced DA methods to demonstrate the feasibility of global ocean DA. This 10-year experiment had two objectives. First, provide short-term ocean forecasts (or initial conditions for climate models). Second, provide global (re)analyses to improve the oceanographic knowledge base and develop tools to improve the observing system.

Associated with this experiment, national operational forecasting agencies were created, aiming to provide real-time and accurate ocean forecasts for coastal and offshore industry, fishing, and other marine activities.

Difficulties for real-time or global ocean DA are similar to the weather forecasting problem, in particular the trade-off between accuracy and computing power.

8.1.1.2 ▪ Coastal Lagrangian search and rescue

Because of human activities and presence along the coast, coastal assimilation (including forecast for search and rescue) is of course of great interest. It is related to Lagrangian assimilation, which aims to retrieve the ocean state (currents in particular) from drifting buoys' location information, which is a convenient way to obtain additional local data at a reasonable cost.

8.1.1.3 ▪ Ocean and atmosphere or biogeochemistry coupling

Another hot topic is DA for coupled models. Coupling may vary depending on the application. For climate (or tropical storm), ocean–atmosphere models are considered, whereas for marine biology, models coupling the physical processes with chemistry and/or (phyto)plankton are developed. The latter are related to ocean color images, which give information about chlorophyll content.

8.1.1.4 ▪ Error and bias management

As for the atmosphere, and even more so because of the inaccessible ocean depths, managing model biases as well as correlated observation errors is an extremely difficult task, which is kept quiet in most of the DA literature. However, as DA systems improve, it is becoming more and more difficult to ignore.

8.1.2 ▪ Examples of DA problems in this context

8.1.2.1 ▪ Global assimilation and operational forecasting

The GODAE [Smith, 2000; Bell et al., 2009], as well as other initiatives, such as ECCO (Estimating the Circulation and Climate of the Oceans) and WOCE (World Ocean Circulation Experiment), instigated many global DA experiments. Among the significant developments, MIT developed a global ocean DA model based on automatic differentiation [Wunsch and Heimbach, 2007], for example, in Stammer [2004] it provided improved estimates of air–sea fluxes for climate models.

On the operational side, we can cite two representative works, which take advantage of two other ocean models: Brasseur et al. [2005] describe the French operational forecasting project MERCATOR, based on the European ocean model NEMO; Chao et al. [2009] rely on the Regional Ocean Modeling System (ROMS) to design a forecasting system for the coast of central California.

8.1.2.2 ▪ Coastal Lagrangian search and rescue

Another interesting avenue of oceanographic applications is the assimilation of Lagrangian data (drifting float position information); see, e.g., Nodet [2006], Kuznetsov et al. [2003], Fan et al. [2004], Castellari et al. [2001]. These observations are not straightforward to use, because the observation operator involved is nonlinear and quite sensitive to both initial perturbations and small scale phenomena, but they provide "cheap" (compared to satellite data) observations. They are sometimes used for search and rescue operations [Davidson et al., 2009] and/or along the coast, as well as for oil spill monitoring [Jordi et al., 2006]. There are also a few attempts at coastal morphodynamics forecasting [Smith et al., 2009], which is obviously extremely difficult even from a direct modeling point of view.

8.1.2.3 ▪ Ocean and atmosphere or biogeochemistry coupling

Model coupling is a hot topic in geoscience modeling. Indeed, as computing power and modeling progress, more realism is gained by correctly representing external forcing and possible interactions between a system and connected systems, and model coupling is an accurate way to do so. For example, for El Niño forecasting, it is crucial to correctly represent the ocean–atmosphere interaction [Behringer et al., 1998; Kleeman et al., 1995]. DA in this framework can be complicated, as technical implementations of coupling can complicate the DA algorithms (e.g., adjoint methods need to be able to differentiate the coupling interaction and communication). Other than ocean–atmosphere coupling, coupled models are used in marine biogeochemistry [Brasseur et al., 2009; Carmillet et al., 2001; Gregg, 2008; Natvik and Evensen, 2003]. DA in these models usually relies on ocean color observations, as the ocean color is strongly related to its chlorophyll and phytoplankton content.

8.1.2.4 ▪ Error and bias management

DA with errors (systematic bias in the model or correlated observation errors) is quite common. Regarding systematic error and model bias correction, we can mention Bell et al. [2004] and Huddleston et al. [2004]. Regarding observation errors, the usual assumption implies that the observation error covariance matrix **R** is diagonal; hence the error is uncorrelated. This causes most satellite images to be thinned before use, so that in the end only a small percentage of the data is actually used. There have been some attempts at proposing nondiagonal **R** matrices; e.g., Chabot et al. [2015] use a change of variable into wavelet space to put correlation information back into **R**.

8.1.3 ▪ Focus: Operational oceanography

Operational oceanography aims to provide monthly and seasonal forecasts of the ocean state. These ocean forecasts can have multiple uses: for coastal industries, off-shore activities, fisheries, and sailing races for example. They can also be used as forcings for monthly and seasonal operational weather forecasting. Indeed, at this range, the ocean impact on the weather must be precisely taken into account. This is the object of this focus. The work of Vidard et al. [2009] deals with the operational ocean system analysis of ECMWF, which provides initial ocean states for the weather forecasts. This work presents the first implementation of the assimilation of altimetric sea level satellite data into the ocean analysis system.

In this work, the assimilation scheme is quite simple—we use OI. OI is a simplified version of the BLUE algorithm, where only relevant (local) observations are taken into account for the computation of the analysis at a given point.

Figure 8.1 presents correlation maps between currents computed by the ECMWF ocean analysis system and the ones produced by the OSCAR project (Ocean Surface Current Analysis), which have been proven to be of good quality, in particular in the tropical ocean. The closer to one the correlation is, the better the results. As we can see, the combination of both in situ and satellite assimilation yields the best correlation improvement.

8.2 ▪ Glaciology

8.2.1 ▪ Presentation of the domain

Glaciology is a multifaceted domain whose various subfields are crucial for the understanding and forecasting of current climate change. We can cite many reasons why this is so.

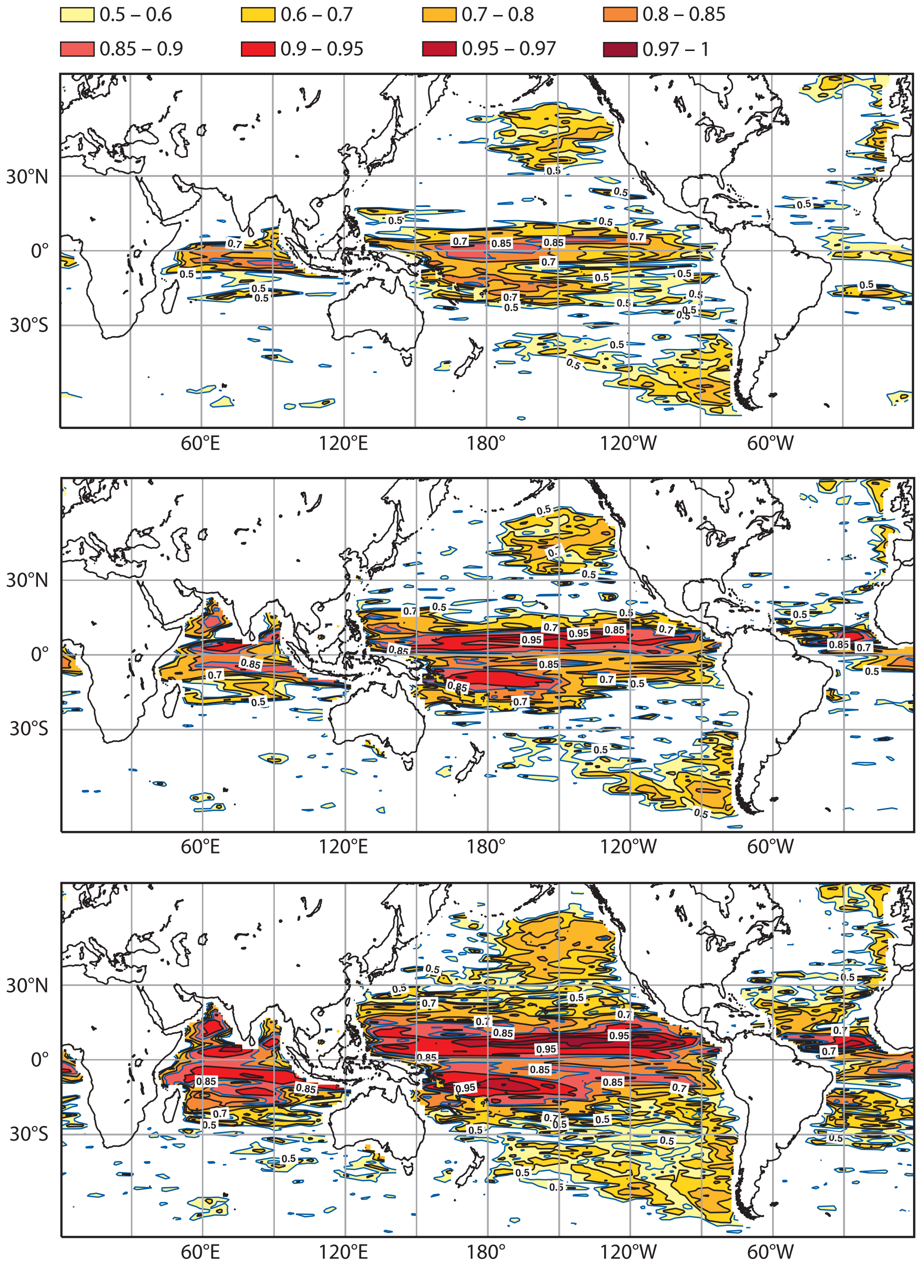

Figure 8.1. *Correlation map between currents produced by ECMWF ocean analysis and the OSCAR surface currents. Top panel: control run (without assimilation). Middle panel: assimilation of in situ temperature and salinity observations only. Bottom panel: combined assimilation of in situ (temperature and salinity) and altimetric (sea level height) observations. The darker the color, the closer to 1 the correlation is, and the better the performance. Reprinted with permission from American Meteorological Society [Vidard et al., 2009].*

8.2.1.1 ▪ Paleoclimatology

First, glaciology, through ice core modeling, gives plenty of information about the climate of the past. Indeed, when ice forms, air bubbles are confined inside and act as a snapshot of the current atmosphere content. Deep ice cores then carry information about past atmosphere gas composition, and isotopic analysis allows us to some extent to reconstruct temperature history. Ice cores also contain ash, pollen, and dust, and all of these combined with the gas bubbles and the ice volume give information about many past phenomena (such as volcanic eruptions, the sun's activity, ocean volume, precipitation, atmospheric chemistry, fires, and plants).

8.2.1.2 ▪ Climate 1: Albedo

Second, the ice on the planet (ice sheets, ice caps, glaciers, sea ice) is a crucial component of climate models because of its link to the earth albedo and the associated positive feedback loop. Indeed, the ice albedo, though variable, is generally high (meaning that a lot of sunlight energy is reflected away). If the ice starts to melt, then the albedo in that area will decrease, leading to more heat being absorbed, which in turn will accelerate the ice melt and the albedo decrease. The reverse is also true (at least locally in time and space, under some temperature and humidity assumptions), as a snowfall would increase the albedo and lower the temperature, allowing more snow to fall, etc.

8.2.1.3 ▪ Climate 2: Sea level rise

Third, terrestrial ice (everything but sea ice) makes a significant contribution to sea level change and is also associated with interactions with the ocean (and the climate in general) in this regard. Global warming affects the ice sheets, ice caps, and glaciers and causes them to contribute to half of the current total sea level rise (which is about 3 mm/year). More dramatically, the melting of the two ice sheets of the planet (Antarctica and Greenland) would cause a sea level rise of 61.1 and 7.2 meters, respectively. Such a massive melt would of course take time, but a positive feedback loop also exists here, called the "small ice sheet instability," caused by the link between altitude and temperature: if the ice melts so that the ice sheet elevation goes beyond a given altitude, then its surface temperature will reach melting temperature, which in turn will increase the melt and decrease the altitude again, accelerating the phenomenon and causing the ice sheet to disappear completely. Due to the massive impact of sea level rise on human population, forecasting the contribution of Antarctica and Greenland is therefore a very important topic in glaciology.

8.2.2 ▪ Examples of DA problems in this context

Glaciology's specific difficulties are numerous: coexistence of very small and very large scales in both time and space; strong nonlinearities in the fluid behavior (non-Newtonian viscosity); complex thermomechanical behavior of the ice structure and ice material properties; difficulty of observation acquisition (inside and below the ice); etc. Still, the progress of both modeling and observing systems allows the use of DA methods to answer a couple of questions.

8.2.2.1 ▪ Paleoclimatology

Providing a good reconstruction of past temperatures is a key ingredient to understanding climate variations. Many proxies (indirect observations) give information about it: atmospheric composition from ice cores, oxygen isotopes from ocean sediment cores, pollen from soil cores, and so on. DA aims to combine these indirect observations with ice sheet dynamics models to reconstruct paleotemperatures.

For example, Bintanja et al. [2004] used the nudging method combined with observations of sea level into a large-scale 3D model of ice dynamics to reconstruct northern hemisphere temperature from 120 000 years ago until now. More recently, Bonan et al. [2012] used the adjoint method on an academic test case to study the assimilation of ice volume data to recover past temperature.

Another example is the dating of ice cores, which aims to combine ice and gas stratigraphic observations and glaciological modeling to reconstruct the chronology of the ice. Lemieux-Dudon et al. [2008] and Lemieux-Dudon et al. [2010] developed a variational assimilation scheme to consistently date multiple Antarctic ice cores at once and provide confidence intervals on the inferred dating.

8.2.2.2 ▪ Sea level rise forecast

The Intergovernmental Panel on Climate Change (IPCC) deemed it urgent in its fourth assessment report in 2007 to develop robust forecasting methods to predict Antarctica's and Greenland's contribution to sea level rise, in particular the part due to ice calving (ice discharge to the ocean, happening at the boundary of the ice sheets). This is a complex problem, as it depends on the ice velocity in ice streams (or outlet glaciers) and ice shelves, which in turn is highly sensitive to poorly observed parameters, such as the elevation of the bedrock and the basal boundary conditions (which can vary greatly between two extremes: grounded frozen ice and melting fast sliding ice).

Various DA methods were implemented to infer these parameters from surface observations, for example Arthern and Gudmundsson [2010], Arthern and Hindmarsh [2006], Arthern [2003], and Gillet-Chaulet et al. [2012] (approximate adjoint method); Heimbach and Bugnion [2009] (automatic differentiation adjoint); and Bonan et al. [2014] (ETKF).

8.2.2.3 ▪ Local glacier and ice stream monitoring

On smaller scales the problem of inferring basal properties of the ice from satellite surface data to forecast the evolution of glaciers and ice shelves is also of strong interest for the glaciological community.

Many studies used DA methods to do so, e.g., Jay-Allemand et al. [2011], Larour [2005], MacAyeal [1992], Morlighem et al. [2010], Petra et al. [2012], and Vieli et al. [2007], who used diverse variations of the adjoint method to study ice streams, glaciers, and ice shelves for various applications: study of a glacier surge, reconstruction of the ice shelf rheology, reconstruction of the basal condition of an ice stream, etc.

8.2.3 ▪ Focus: Sea level rise

The first implementation of DA methods to provide a forecast of Greenland's contribution to sea level rise was done in Gillet-Chaulet et al. [2012]. The main issue was to estimate the basal parameter in the friction law at the bedrock interface. The ice discharge (volume loss) is indeed highly sensitive to this parameter, as it governs the sliding velocity, which can therefore vary strongly with this parameter. Then the

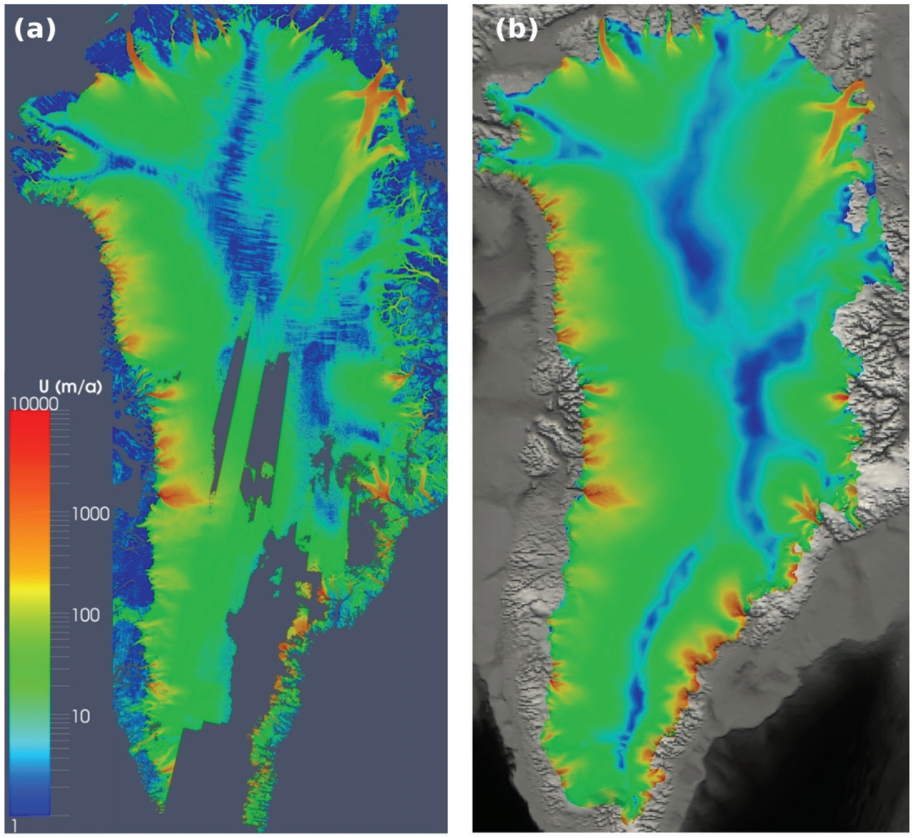

Figure 8.2. *Greenland surface velocities, on the left* (a) *observed velocities; on the right* (b) *velocities after reconstruction of the basal coefficient field by an approximate adjoint method. Reprinted with permission from European Geosciences Union [Gillet-Chaulet et al.,* 2012*].*

sliding velocity in turn influences the discharge of ice toward the sea. As this parameter (which is not a single number but a high-dimensional variable, with one value for every grid point on all the bedrock) concerns basal boundary conditions, it is very difficult to estimate directly because measurements are simply impossible.

DA methods allow us to use surface observations to recover this parameter field. An approximate adjoint method was used, in other words a variational assimilation method in which the adjoint was simplified. The simplification was necessary at that point because the ice sheet model that was used is a state-of-the-art finite element model with unstructured and adaptive mesh (using classical error estimates), for which no full adjoint is yet available.

For this experiment, the surface observations were the surface height of the ice and the surface velocity. In this framework, with real data, exact validation is impossible as the true state of the unknown parameter is unreachable. Therefore, the validation was made by comparing the simulated surface velocities after assimilation and the observed velocities; see Figure 8.2. As we can see, the main features of the flow are well reproduced: low velocities in the central areas, fast-flowing ice streams individualized and well localized, and good rendering of the largest outlet glaciers and their watersheds, all of which are crucial for sea level rise estimation.

8.3 ▪ Fluid–biology coupling; marine biology

8.3.1 ▪ Presentation of the domain

In this section we consider applications in biology and ecology coupled with fluid dynamics (be it the ocean or a tank of water). This includes the study of marine ecosystems as well as the engineering use of algae or bacteria for ecological purposes. Both these questions are of significant importance for the future and are strongly related to climate change and human activities.

8.3.1.1 ▪ Marine ecosystems

One such problem is the evolution of fish populations. Because millions of people depend on fisheries, this question is clearly important. The fish population evolution depends on multiple factors: fishing industry catches, predator health/number/location, food availability (other smaller fishes, plankton, or phytoplankton), ocean temperature/currents/chemistry, climate change, etc. This is a typical ecology problem, and the complete answer would require the modeling of the whole food chain and a complete ecosystem. Similarly, the study of the response of marine ecosystems (not only fish populations) and food chain equilibriums to global change is of ecological interest.

8.3.1.2 ▪ Microbiology and ecological remediation

Another problem related to fluid–biology coupling is the use of microbiology for ecological engineering. For example, knowing that a given bacterium is known to consume a given synthetic molecule contributing to pollution, how could we monitor, forecast, and control the evolution of the bacteria population so that we achieve bioremediation/pollution control? In the same area, people are also interested in using algae to produce biofuels, and the same questions about monitoring and controlling the total biomass hold.

8.3.2 ▪ Examples of DA problems in this context

8.3.2.1 ▪ Fishing industry

Due to the arrival of factory ships (e.g., purse super seiners, factory ships that can fish hundreds of tons in one catch), overfishing has been a growing concern. In this framework, modeling and forecasting of the fish population evolution can be a huge help for decision-makers. The fish population depends not only on fishing catch but also on various geophysical parameters (climate, currents, food availability, etc.), so that coupled and complex marine ecosystem models are required to provide accurate trends. Increased computational power also allows an increase of complexity and improved modeling capacity. However, and this is a common problem in ecology, it also causes the multiplication of unknown parameters in the models.

DA, more specifically the adjoint method, has been used to calibrate model parameters in this context, e.g., Lawson et al. [1995] for an academic predator–prey model, McGillicuddy et al. [1998] for a coupled plankton–ocean model, and Dueri et al. [2012a] and Faugeras and Maury [2005] for skipjack tuna population evolution in the Indian ocean.

8.3.2.2 ▪ Marine ecology

Closely linked to fish population ecology, there has been an increase of interest in marine ecology. In particular, models were developed for the coupling between physical oceanography, biogeochemistry, and plankton biology, as key ingredients for fishing ecology studies. As mentioned above, despite the upgraded complexity, these models still rely on strong parameterization of unresolved scales and phenomena, involving many unknown parameters. Brasseur et al. [2009] propose a short review of the domain, and more comprehensive studies can be found in Dowd [2007] (Bayesian statistical assimilation) and Faugeras et al. [2004, 2003] (variational assimilation). The review Luo et al. [2011] offers a broader view of DA of ecosystems modeling, not only in the marine framework.

8.3.2.3 ▪ Microbiology

Biological remediation is a tool to achieve depollution of large water bodies, such as reservoirs, lakes, and rivers. It relies on the simple principle that putting (well-chosen) bacteria into polluted water will lead the bacteria to consume the contaminant, grow, and produce biomass (dead bacteria containing the contaminant). Extracting the biomass while renewing the polluted water ensures that the bacteria still live and gradually clean the water. This process can be done in tanks that pump the contaminated water, remove the biomass, and pump clean water back into the water body. The question is then to optimize the flow output into the tank to achieve maximal depollution in minimal time.

This problem is modeled by coupling the biology (bacteria/contaminant population evolution) and fluid dynamics (contaminant transport versus clean water diffusion into the water body). Control methods were successfully applied, both with simple ODE-based models [Gajardo et al., 2011] and more complex spatialized PDE-based models [Barbier et al., 2016].

8.3.3 ▪ Focus: Tuna population ecology

A model of tuna dynamics has been developed and presented in Dueri et al. [2012a] and Faugeras and Maury [2005]. This model is structured in 3D space and fish size, and couples the fish population, the marine ecology, and the physical oceanography. It is based on PDEs representing the evolution of the density of the fish at every given size. Of course, this kind of model is highly complex and depends on many unknown parameters that must be finely tuned to reproduce the real tuna dynamics well. Dueri et al. [2012a] proposed computing a cost function measuring the negative loglikelihood of various observations (fish catches, size frequencies) and minimizing it to calibrate the model parameter. The minimization required the gradient of the cost function, which was obtained using an automatically differentiated TLM.

After assimilation, the model proves to be satisfactory; e.g., when comparing the vertically integrated exploitable biomass to the catch distribution in Figure 8.3, we can see that it is able to properly represent the main features of the tuna distribution at different periods of the year and under different environmental conditions.

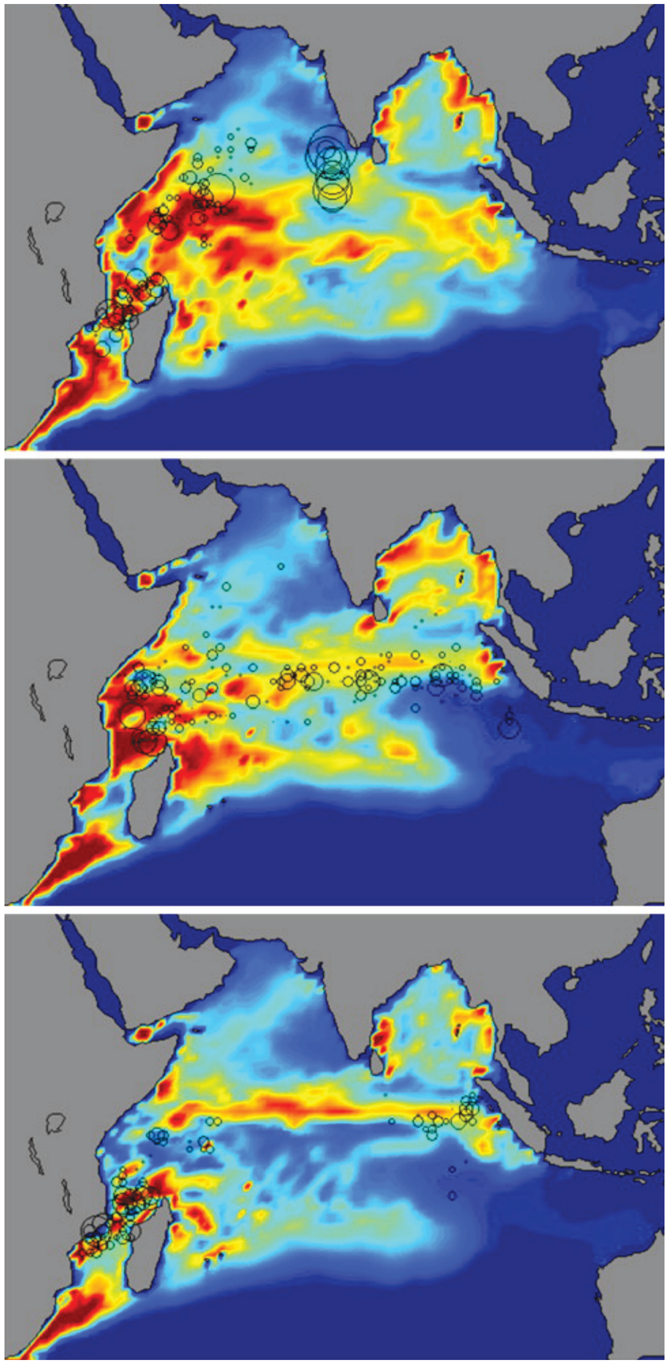

Figure 8.3. *Exploitable population of skipjack tuna computed by the model versus observations (circles) in the Indian Ocean. Top: April* 1993, *middle: February* 1998, *and bottom: April* 1998. *Reprinted with permission from Elsevier [Dueri et al.,* 2012b*].*

8.4 ▪ Land surface modeling and agroecology

8.4.1 ▪ Presentation of the domain

This section groups applications covering plant, water, and carbon cycles, e.g., agronomy, agriculture, forestry, soil sciences, carbon pools, and associated climate retroactions. This vast group of domains is equally crucial for the current and future well-being of human populations and the planet in general.

8.4.1.1 ▪ Agroecology for water quality and availability

Issues related to water and agriculture have an immediate impact on human lives. Supply of both water and plants (in terms of quantity as well as quality) is changing fast because of human activities and climate change. The understanding of their processes, as well as their monitoring and forecasting, is therefore of utmost importance. In this area, scientific issues are numerous; we can cite a few: evolution and monitoring of subsurface water reservoirs due to agricultural irrigation and precipitation changes, pollution of freshwater by pesticides and bioremediation using trees or buffer zones, and crop modeling and plant growth modeling for water saving optimization. It is clear that each of these questions directly impacts the quantity and quality of the water supply of the local population. They are closely related to agroecology and agriculture questions (use of pesticides, irrigation, crop management) and therefore food security (also in terms of both quantity and quality).

8.4.1.2 ▪ Vegetation, continental water, and carbon cycle modeling for climate

Agriculture, forestry, and surface and subsurface hydrology, as well as the carbon cycle, have many links and feedback loops with climate change, which are and will indirectly affect human, animal, and plant populations as well. A few of the questions of interest here are contribution of the forests/soil/plants to the carbon cycle and modeling of ecosystems; modeling of the water exchange and humidity on continental surfaces, and feedback with the atmosphere and climate; modeling of the ground cover distribution and evolution, link with the albedo and humidity modeling, and feedback into the climate models; and so on. These questions are important for our future. We will see that DA provides tools to look for answers and further our knowledge and understanding of them.

8.4.2 ▪ Examples of DA problems in this context

Land surface and agroecology modeling are complex for many reasons: scarcity of observations below the surface of the earth, strong spatial heterogeneity, multiple scales, reactive chemical species, nonlinearities, threshold phenomena, etc. Up until recently, most models were limited to conceptual models (in distributed hydrology, for example) or to empirical modeling (in crop modeling, for example). However, the increase of computer power and the availability of high-resolution satellite observations allow both the spatialization of models and the implementation of DA methods.

8.4.2.1 ▪ Agroecology

For example, Wu et al. [2012] implemented the adjoint method to optimize the water supply for sunflower growth. They identified that plant growth is sensitive to the supply frequency and total during each development phase of the plant and were able to propose an optimal strategy for fruit filling.

Another example is crop modeling, which provides useful tools for testing of agronomical strategies and decision-making to both minimize environmental consequences and maximize crop production. The review by Dorigo et al. [2007] studies the combined use of remote sensing observations and DA methods developed to improve agroecosystem modeling. For example, as has been said before, such models rely on poorly known soil parameters. Lauvernet et al. [2002] developed a variational method to assimilate vegetation properties from remote sensing images into a crop model, to estimate soil and crop characteristics, and to better forecast the production. Varella et al. [2010] have implemented an inverse model (importance sampling method that consists of one analysis step of the particle filter) to infer these parameters from real observations.

8.4.2.2 ▪ Vegetation, water, and climate

Land surface modeling is a key component of climate models, as it impacts albedo, humidity, carbon cycle, and more. Because of vegetation and evapotranspiration, feedback loops exist between (agricultural) land use and local climate and water resources. To study this interaction, Courault et al. [2014] developed a coupled model (microclimate/land surface), for which they calibrated the (unknown) soil parameters using variational DA of remote sensing observations. Other applications were performed on the same principle: use satellite data to retrieve poorly known vegetation/land surface biophysical parameters. For example, Pellenq and Boulet [2004] implemented the EnKF for a vegetation model, and Lauvernet et al. [2002, 2008] developed a variational method on a radiative transfer model to estimate vegetation properties at the top of the canopy from temporal series of upper atmosphere data. Compared to classical inversion methods used in radiative transfer, the adjoint model combined with temporal and spatial constraints from remote sensing images drastically improved the results on twin experiments.

8.4.2.3 ▪ Carbon cycle dynamics

Modeling of the carbon cycle on the land surface is a difficult task because it requires accurate modeling of a complex ecosystem (called the terrestrial carbon ecosystem) with many interactions and uncertainties. The terrestrial carbon ecosystem is a carbon sink (i.e., the land biosphere carbon absorption exceeds its losses). Compared to the other components of climate models, it presents the largest uncertainties, up to the point that predictions for the future remain largely unknown. There have therefore been recent attempts to constrain these models using DA; see e.g., Bloom and Williams [2015] (focus on the carbon pool ecosystem) and Delahaies et al. [2013] (focus on the variational assimilation system). At a smaller scale, Williams et al. [2005] studied forest carbon dynamics using the EnKF, and Ribbens et al. [1994] calibrated a model of tree seedling dispersion using an inverse method (direct likelihood maximization).

8.4.3 ▪ Focus: Crop modeling

In crop modeling, most of the input parameters are either empirical (since the functioning of vegetation is not a priori described by exact equations) or difficult to estimate, due to their large variability in time and space. LAI (leaf area index) is a key canopy variable that directly quantifies green vegetation biomass. Though it is the most observable canopy parameter (using remote sensing), assimilation of LAI is usually performed without capitalizing on the information from temporal and spatial dependencies of the vegetation. Lauvernet et al. [2014] performed such a study and assumed that the model parameters are governed by spatial dependencies.

Figure 8.4 (top) depicts observations of LAI (generated by their Bonsai model) on a whole-wheat growth cycle, from sowing to harvest, on 100 different pixels of the landscape: x-axis is time in days; y-axis is the pixels; and z-axis is the LAI value on each pixel, at each time step. One distinguishes the spatial levels considered: the two big groups of LAI represent two different varieties of wheat (from pixel 1 to 50, then 51 to 100), where the varietal parameters are equal in the model. For each variety, there are 5 plots with similar plot parameters (pixel 1 to 10, 11 to 20, etc., ...), and 10 pixels per plot, with independent parameters.

In Figure 8.4 (middle), LAI was assimilated without using spatial constraints, i.e., on each pixel independently. Results are quite satisfying in this example due to a very large set of observations of LAI (one image each day). If the number of observations decreases (e.g., one image every 20 days, like a satellite), the estimation quality decreases drastically. However, using spatial constraints (as has been done on Figure 8.4 (bottom)) allows us to get better estimates with a large dataset, but also to keep stability when the observation frequency decreases.

8.5 ▪ Natural hazards

8.5.1 ▪ Presentation of the domain

Climate change is most likely increasing, and will continue to do so for at least the next couple of decades, according to the probability of extreme event occurrence. This section deals with the modeling and forecasting of such events to mitigate the risk for human populations and activities.

8.5.1.1 ▪ Floods

As population increases in already dense areas, a certain disregard for the risk of rare events (e.g., floods with long return periods such as 25, 50, or 100 years) is not uncommon. Therefore, floods caused by rivers or sea storms impacting housing or commercial buildings are regular natural hazards. In parallel to the population increase, there is also a development of the observation network for rivers (e.g., satellite remote sensing or in situ monitoring systems), which allows for operational flood forecasting. Still, some scientific issues remain: hydrological modeling of extreme events, data sparsity, or, on the contrary super high resolution data versus large-scale models, etc.

8.5.1.2 ▪ Wildfires

As was said above, climate change will most likely increase the risk of natural hazards. Wildfires are also affected, as changes in precipitation and temperature will probably cause severe drought in some areas and increase the risk of large and destructive wildfires. Therefore, real-time forecasting of wildfire propagation is a crucial issue for both wildfire hazard prevention and emergency management. This is a complex problem, in terms of both modeling and DA, because of numerous difficulties: nonlinear and complex front propagation, strong dependency on atmospheric conditions, need of real-time observations and predictions, feedback loops between fire and atmosphere, etc.

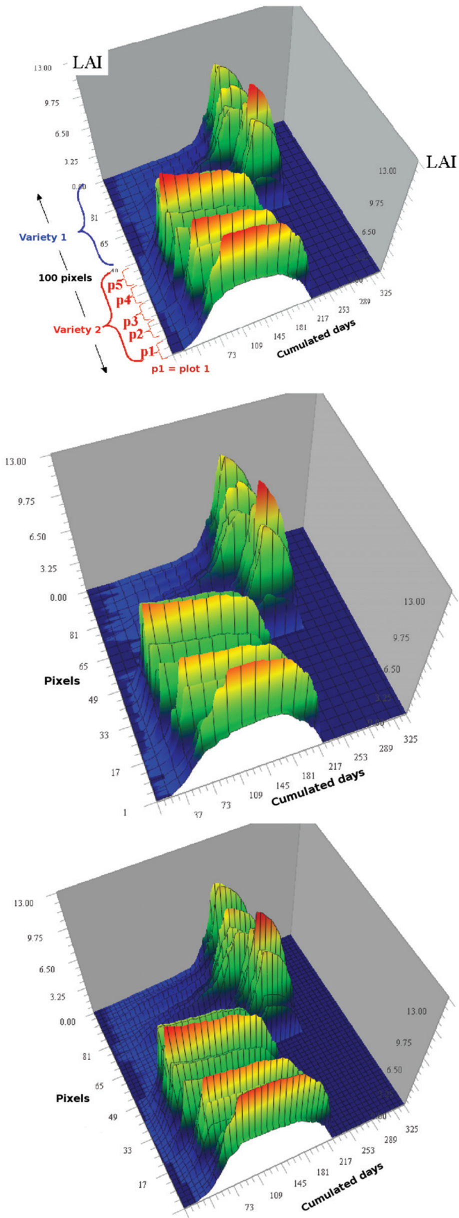

Figure 8.4. *LAI maps. Top: LAI observed in time and space (represented here without noise). Middle: LAI estimated after classical DA of noisy LAI (RMSE = 0.39623). Bottom: LAI estimated after DA of noisy LAI using spatial constraints (RMSE = 0.36573). Reprinted with permission from RNTI [Lauvernet et al., 2014].*

8.5.1.3 ▪ Hurricanes

Hurricanes or tropical storms are extreme storms that develop over warm ocean water and lose intensity over land, so that they cause huge damage over coastal tropical areas. According to IPCC [Field, 2012], “attribution of single extreme events to anthropogenic climate change is challenging.” Experts still disagree on the question, and the evolution of losses due to tropical storms is also unclear, because of two conflicting factors: population and development increase in coastal and sensitive areas, versus forecasting and prevention progress. From a scientific point of view, hurricane track forecasting has much improved over the last decades, but hurricane intensity forecasting is still lacking. Indeed, intensity is closely related to the inner core structure of the storm, which is both insufficiently modeled (as it requires very high resolution and/or fine parameterization of small-scale physics) and insufficiently observed.

8.5.2 ▪ Examples of DA problems in this context

8.5.2.1 ▪ Hydrology and floods

Many recent studies have tackled the problem of forecasting flash flood extreme events. Some of them focus on river hydraulics (using PDE models such as shallow-water models), while others study complete river catchments (using distributed models of independent grid cells connected by transfer functions).

For PDE-based models, real-time flash flood modeling requires accurate input variables. For example, Ricci et al. [2011] implemented the KF to retrieve the upstream inflow from river water level observations, while Habert et al. [2014] estimated the lateral inflow using the EKF.

On the other hand, Harader et al. [2012] studied the retrieval of the rainfall input in a rainfall-runoff model of a river catchment, which is a complex task because of numerous uncertainties and heterogeneities in the model. On this subject, the difficulties have been studied and highlighted by Coustau et al. [2013], who implemented the BLUE to estimate the peak discharge on the same catchment for real-time flood forecasting.

8.5.2.2 ▪ Wildfires

Wildfire modeling has recently improved, with complex and precise models for solving the flame structure (combustion coupled with computational fluid dynamics (CFD)). However, these models are too computationally expensive to be used for real-time forecasting, and simplifications need to be made for DA experiments.

Mandel et al. [2008] proposed a coupled PDE model of energy (temperature) and fuel balance and successfully performed synthetic experiments with the EnKF to provide temperature corrections.

However, as even this type of model is deemed too expensive to be used for real-time forecasting at the regional scale, recent approaches have developed simplified models based on a front-tracking solver in which the fire is considered as an interface (i.e., a front) propagating between the burning area and the unburnt area. Errors in the model inputs and equations translate into errors in the simulated characteristics of the fire, and thus should be reduced. For example, Rochoux et al. [2015, 2014] proposed a two-part comprehensive DA study to first retrieve biomass fuel and wind parameters and second provide real-time forecasts of the fire front.

8.5.2.3 ▪ Hurricanes

Real-time forecasting of tropical cyclones has of course been of major interest for decades. Earlier DA experiments used the nudging method; see e.g., Anthes [1974]. With the development of NWP, more sophisticated methods were implemented. The first implementations of variational methods gave poor results because fixed covariance information did not correspond well to the hurricanes' strong variability at their core. To overcome this problem, Xiao et al. [2000] developed a two-step variational method: the first step is dedicated to the correct positioning of the hurricane, and the second step is for the classical assimilation of the other variables. More recently, Torn and Hakim [2009] implemented the EnKF to make the most of its handling of error covariances. As the EnKF naturally evolves and adapts forecast error covariances, the results were indeed improved.

8.5.3 ▪ Focus: Wildfire modeling and real-time forecasting

In the context of wildfire modeling, Rochoux et al. [2014] performed DA on a front-tracking model to retrieve some poorly known model parameters and state. The biomass fuel properties and the near-surface wind as well as the position of the front are sequentially corrected as new observations become available. In Rochoux et al. [2015] they used the calibrated model to provide a real-time forecast of the fire front.

The DA ingredients are as follows: the control variable is either the input parameters of the front-tracking model [Rochoux et al., 2014] or the positions of markers along the fire front [Rochoux et al., 2015]; observations of this front are available (as geotracking tools now make it possible to obtain such data from mid-infrared imaging), and the EnKF is used to combine the data with the front-tracking model.

Figure 8.5 presents the assimilation results of a controlled grassland fire experiment with four EnKF assimilation cycles from initial time 50 s to final time 106 s. In this figure the benefit of the DA can be clearly seen, as it really provides a precise tracking of the fire front. This ability of DA to provide real-time analysis of the fire front over a simple real-life (albeit controlled) grassland fire shows the maturity of both the model and the method.

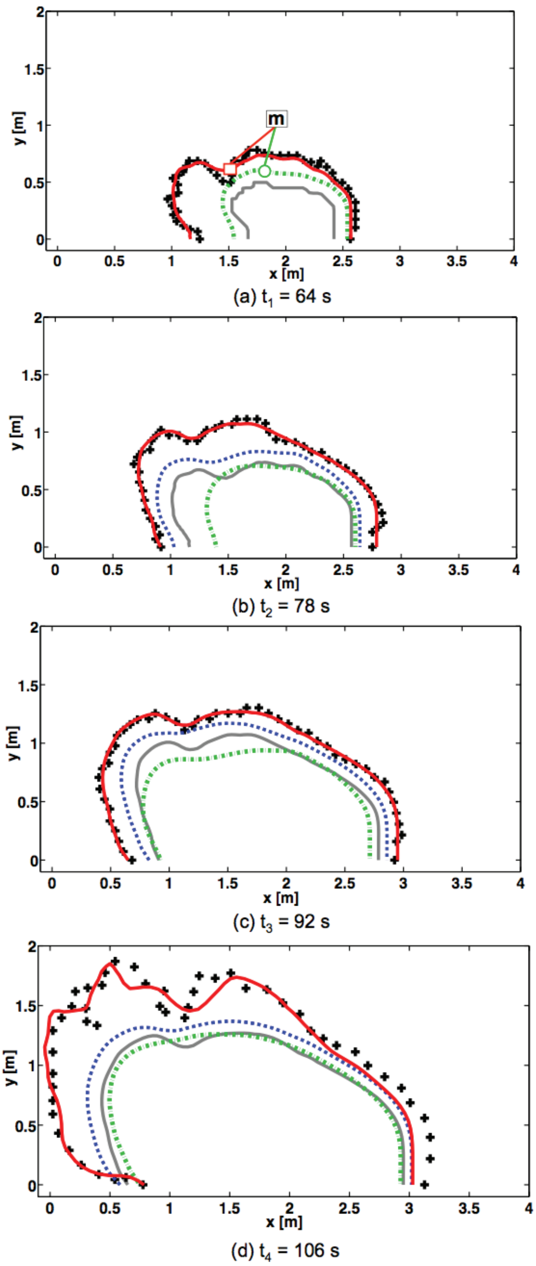

Figure 8.5. *Controlled grassland fire experiment (data from King's College London) with multiple assimilation cycles from* $t_0 = 50$ s *to* $t_4 = 106$ s *with fire front position estimation. Black crosses correspond to observations, the gray solid line corresponds to the initial condition of the assimilation cycle, the green dashed-dotted line corresponds to the free run (without DA), the blue dashed line corresponds to the mean forecast estimate (without DA for the first observation time or with DA at the previous observation time), and the red solid line corresponds to the mean analysis estimate (with a DA update at the current observation time). Assimilation cycles* (a) *[50; 64 s];* (b) *[64; 78 s];* (c) *[78; 92 s]; and* (d) *[92; 106 s]. Reprinted with permission from European Geoscience Union [Rochoux et al., 2015].*