Quantification of uncertainties from ensembles of simulations

Vivien Mallet, Isabelle Herlin

Quantification of uncertainties from ensembles of simulations

Vivien Mallet* and Isabelle Herlin
INRIA, Paris, France
E-mails: Vivien.Mallet@inria.fr, Isabelle.Herlin@inria.fr

INTRODUCTION

Decision making for environmental issues increasingly relies on numerical simulations and various observational data. However, the numerical models are limited by strong uncertainties because of poor input data and inaccurate physical, chemical, biological or mathematical modeling. Moreover, measurement instruments do not allow for a complete observation of environmental systems, and they often acquire noisy observations. Nevertheless, there is a strong need to optimally and jointly exploit numerical simulations and field observations for an objective assessment of risks on present and future times.

In this context, it is critical to quantify the uncertainties of all information sources (numerical models, empirical rules, fixed observations, mobile observations, qualitative observations) and to evaluate the best estimates that are derived from all the information. The final scientific products that may help decision-making are the probability distribution of the target quantities, confidence intervals or probabilistic forecasts.

These various products can be derived from ensembles of simulations possibly combined with observations by the so-called data assimilation methods. The ensembles can be calibrated, including for the forecasts, in order to approximate the distribution of simulation errors. Such methods are for instance operationally applied for weather forecasting. The distribution of ensembles can be processed for a better quantification of the uncertainty. It is for instance possible to derive risk maps. Practical applications, like the protection of populations after a nuclear disaster as in Fukushima, can benefit from such risk maps, e.g., to determine an evacuation zone.

GENERATION OF ENSEMBLES

The accurate evaluation of the state of the environment tends to rely on numerical simulations and observational data. On the one hand, field observations often offer quality evaluations, but only at a limited number of locations and for a few variables. On the other hand, numerical models produce rich data including all state variables, at any location, for past, present and future times. Numerical models have reached a wide range of applications, from weather forecast to earthquake risk assessment, from urban noise pollution to flood cartography. However, a major issue for the use of these models for any decision-making lies in the high uncertainties they may be subject to.

The uncertainties in the numerical models come from various sources. The lack of understanding of the modeled phenomena leads to an approximate formulation of the processes. For example, not all chemical reactions in the atmosphere are accurately described. Even when the phenomena are perfectly known, they may not be all described by the model because of lack of data or unbearable computational costs. As a consequence, many models rely on subgrid parameterizations that approximately represent processes occurring at unresolved scales. For example, turbulence is not explicitly computed in large-scale dispersion models. Simulation models are also subject to the approximations of their numerical schemes, which can be high, especially for non-smooth states or fast varying phenomena. Finally, a well-known source of uncertainty lies in the input data to the models. Only part of the input data can be observed, and the rest is provided by uncertain numerical models or human expertise.

The combination of all the uncertainties sources often leads to high uncertainties in environmental applications. The uncertainty levels, which may easily reach 20% to 60%, are so high that the stochastic approach appears more suitable than the deterministic approach. Ideally, the complete probability distribution for the state of the target system should be computed. In practice, it is approximated by the sampling of an ensemble of simulations. This ensemble is generated so that it accounts for all uncertainties sources and properly represents the uncertainties on the computed state. In the best case, the ensemble is generated from multiple numerical models that provide plausible descriptions of the studied processes. The models should differ in their physical, chemical or biological formulation, they could rely on different mathematical formulations, and they may use various numerical schemes. The input data to these models are perturbed according to their attributed uncertainties, just like in Monte Carlo simulations. Also, alternative data sources may be used. Sensitivity analysis is sometimes applied to identify which input data are potentially the main uncertainty sources (See left image of Fig. 1).

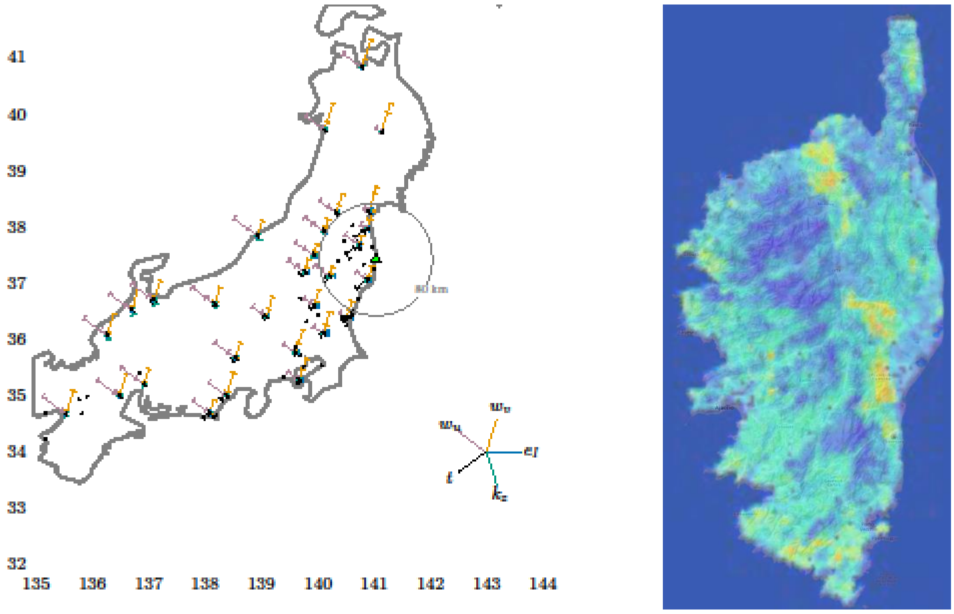

Figure 1: Left: Simulation of the dispersion of radionuclides after Fukushima disaster: representation of the input data whose uncertainties can influence the most the gamma dose rates. wu and wv are the zonal and meridional winds, eI is the emission of iodine, kz is the vertical diffusion, t is the emission time. Right: Risk map of fire over Corsica, for a given date in July 2009. This map is computed based on the burned-area distribution for the potential fires.

The computational cost of ensemble simulations may be too high in practical applications. The typical size of an ensemble is 50 to 100 members, although larger samples might prove their usefulness. Two main approaches are applied for cost reduction. In weather forecasting for instance, the models used to generate the ensemble are run at lower resolution than the traditional deterministic simulation. Another approach is meta-modeling where a model is replaced by an approximate model that runs much faster. The surrogate model can be automatically built by dimension reduction and statistical emulation [1].

CALIBRATION OF ENSEMBLES

Once an ensemble of simulations is generated, a key question is how well it represents the actual uncertainties. The only additional source of information that is available to decide on this question is observational data. However, an ensemble of simulations represents a probability density function while the observations can be seen as a realization of the simulated phenomena. Despite the difference in nature between the ensemble and the observations, there are scores that evaluate an ensemble using the observations. These scores can evaluate the resolution of the ensemble, that is, its ability to make distinct predictions for different subsets of events. They also evaluate its reliability, that is, the ability of the ensemble to produce a probability for an event that matches the observed frequency for that event.

It is rare that an ensemble of simulations shows good scores before a considerable amount of work has been invested in its calibration. Typically an initial ensemble is usually underspread because not all uncertainty sources are included and the actual uncertainties tend to be underestimated by modelers. The calibration can be carried out by optimization of an ensemble score. In the process, the distributions of the perturbations on the input data can be optimized. The models in the ensemble can also be given different weights depending on how likely they are assumed to be. Another approach is to generate a large ensemble and select a sub-ensemble that shows good performance.

Operational ensemble forecasting usually relies on the field observations, such as point measurements or satellite images, in order to improve the forecasts and reduce the uncertainties. The data assimilation methods merge the information included in the numerical model and the information given by the observations in order to compute an improved estimate of the system state and consequently an improved forecast. The Ensemble Kalman Filter (EnKF)[2] carries out the time integration of the ensemble of states, with the numerical model, and computes a new estimate of the states when observational data are available, according to the Bayes' rule. In the update equation, EnKF approximates the error variance of the state with the ensemble spread, and assumes Gaussian distributions on the simulated-state error and the observational error. This may result in spurious covariance values between distant locations of the spatial domain. Localization methods [3] are then usually applied in order to correct these effects.

Another approach to improve the forecast with observational data is the sequential aggregation of the forecasts ensemble. An aggregated forecast is computed as the weighted linear combination of the forecasts of the ensemble. These weights are derived from the past observations and ensemble predictions. Cesa-Bianchi and Lugosi [4] describe the use of machine learning algorithms for computing the weights values. An extension [5] was designed to produce a probability distribution instead of a single deterministic forecast.

OUTPUT PRODUCTS AND INTERPRETATION

Ensemble prediction systems provide a rich additional information compared to deterministic forecasting. This should support an improved decision-making. Various scientific products are computed from processing the ensemble forecasts in order to communicate on the uncertainty. Their significance should be accurately understood for an optimal use of these ensemble methods and a correct interpretation of their results.

The ensemble mean is obtained as the simple average of each variable across all ensemble members. It cannot be used as such, since it will not be able to describe the low probability events.

The discrete distribution of each variable, at each forecasting time, estimated from the ensemble of forecasts, is interpreted as the probability distribution of this variable. Its accuracy is usually evaluated with scoring rules such as the Brier score [6].

The ensemble of forecasts allows estimating the probability of a given event. For instance, if 80 out of 100 members forecast that the variable x will be greater than a threshold v, the probability of exceeding v is usually taken as 80%. These probabilities should be verified on a large set of past cases, in order to validate the adequacy between forecast probabilities and observed frequencies.

Risk maps (see right image of Fig. 1) are additionally produced by processing the forecasts ensemble. The probability of an event, usually a threshold exceedence, is multiplied by its cost for estimating the risk and supporting emergency measures such as evacuation for instance.

REFERENCES

[1] Girard, S., et al. (2016) Emulation and Sobol' sensitivity analysis of an atmospheric dispersion model applied to the Fukushima nuclear accident. J. of Geophys. Res. 121.

[2] Evensen G. (2003) The ensemble Kalman filter: Theoretical formulation and practical implementation. Ocean Dynamics. 53, 343-367.

[3] Gaspari G., and Cohn, S. E. (1999) Construction of correlation functions in two and three dimensions. QJRMS. 125, 723-757.

[4] Cesa-Bianchi N., and Lugosi G. (2006) Prediction, learning, and games. Cambridge University Press.

[5] Thorey, J. et al. (2016) Online learning with the CRPS for ensemble forecasting. QJRMS. accepted.

[6] Brier (1950) Verification of forecasts expressed in terms of probability. Monthly Weather Review. 78.