A METHODOLOGY TO CHARACTERISE THE SOURCES OF UNCERTAINTIES IN ATMOSPHERIC TRANSPORT MODELLING

Fabien Brocheton[1], Patrick Armand[2], Lionel Soulhac[3] and Emmanuel Buisson[1]

[1]NUMTECH, 6 allée Alan Turing, BP 30242, F-63175 Aubière, France.
[2]Commissariat à l'Energie Atomique, Département Analyse, Surveillance, Environnement Service Radioanalyse, Chimie, Environnement, Bruyères-le-Châtel, F-91297 Arpajon CEDEX, France
[3]Université de Lyon, Lyon, F-69003, France; Ecole Centrale de Lyon, Ecully, F-69134, France; CNRS, UMR5509, Laboratoire de Mécanique des Fluides et d'Acoustique, Ecully, F-69134, France; Université Lyon 1, Villeurbanne, F-69622, France; INSA de Lyon, Villeurbanne, F-69622, France.

Abstract: The atmospheric dispersion modelling of pollutants is based on models, but also on data and users, who lead to uncertainties, i.e. to differences between the results of the models and the physical reality to describe. The question of the uncertainty of dispersion models is a subject of increasing interest for primarily two reasons:

- In spite of the significant number of research works on atmospheric dispersion in the last 30 years, results of simulations preserve an important level of uncertainty. Since the reduction of this uncertainty will be more and more difficult, it thus seems today necessary to characterize it and, if possible, to quantify it.
- The development of computer performances allows today the use of models in real time for decision-making aid, e.g. for the control of an industrial facility or for the management of an accident. In this context, where important decisions must be made quickly, it becomes extremely important to provide to the results of models an information on the uncertainty associated with these results.

For better characterising uncertainties associated with atmospheric dispersion models, we propose a methodology of analysis which led us to make an inventory of all the sources of uncertainties of a model, considering all kind of models (Gaussian, Lagrangian, Eulerian). This methodology is based on a decomposition of the modelling process in four steps:

- The collection of input data (sources, meteorology, topography, land cover…), with uncertainties in the measurement of these data, in their treatment and in their representativeness.
- The modelling of the meteorological field, with uncertainties associated with intrinsic nature of atmospheric phenomena and with the quality of the models/approaches used to describe these phenomena.
- The modelling of dispersion processes and of the physicochemical transformations.
- The statistical treatment of results: it is shown that the choice of the output parameters (average annual or hourly maximum, maximum ground value or position of this maximum, etc.) modify, in an important way, the uncertainty of the results and how to evaluate it.

The methodology proposed in this work should make it possible to progress towards a more systematic quantification of uncertainties associated with modelling of atmospheric dispersion.

Key words: *atmospheric dispersion modelling, uncertainty inventory.*

1. INTRODUCTION

Atmospheric dispersion models are used to characterize the environmental impact of releases of effluents (gas or particulate) under normal operation or to evaluate the risk related to a scenario of accidental release. Computational results of atmospheric dispersion are currently generally provided without associated uncertainties, whereas many interlocutors wish to have such information, in particular the operational services and the regulation authorities. Indeed, uncertainty related to the input data used in the models or related to internal parameters, induces variability on simulations, which could be necessary to be estimated. Such a problem is not new and in the literature, there are many articles or reports dealing with this subject in all the fields of modelling. In particular we can quote the old review of Baird (1989) and the recent ones of De Rocquigny et al. (2008) about the quantification of uncertainties in industrial practice. In addition, many works tried to list and to classify the sources of uncertainties, and/or to quantify their impacts on the simulated results, but generally focusing on a particular modelling approach of atmospheric dispersion or on a particular application (see for example Roache, 1997; Hanna *et al.*, 1998; Dabberdt and Miller, 2000; Irwin *et al.*, 2001; Colville *et al.*, 2002; Stein and Wyngaard, 2002; Levy *et al.*, 2002; Perkins *et al*, 2002 ; Mallet and Sportisse, 2006; Godoy *et al.*, 2006) and seldom in an exhaustive way.

This exhaustive approach is the aim of this project, which remains an exploratory work whose finality is to produce a report, which is at the same time teaching and as exhaustive as possible. Nevertheless, this work has been limited to numerical tool for atmospheric dispersion of gaseous or particulate pollutants, from local scale to regional scale. The focus for outputs is limited to concentration fields and deposition fluxes (dry and wet) in terms of [average, percentiles, geographical location] according to several integration durations [instantaneous, few minutes, hour, day, month, year].

This article presents a synthesis of the full report of this project (Brocheton *et al*, 2008). In the first section, we define the concepts of uncertainty, variability and error associated with a modelling tool. In the second section, we present the classification used to carry out the inventory of the sources of uncertainty. In the third section, we illustrate our methodology on the particular case of the wind velocity and the advection process.

2. CONCEPT OF UNCERTAINTY, VARIABILITY AND ERROR

To define the concepts of uncertainty, variability and error, we illustrated on figure 1 how the physical reality is described in a modelling tool. Schematically, the physical reality of the atmospheric dispersion process corresponds to a set of parameters, which interact on various spatial and temporal scales, to lead to concentrations (and deposition fluxes) in 4 dimensions (3D for position x, y, z and 1D for time t). The real processes F are then approximated by means of a simplified representation (see Fig. 1) via a model f (choice of parameterisations, simplifications of processes…), which, starting from a reduced set of input data (x'), gives a representation (y') of the reality.

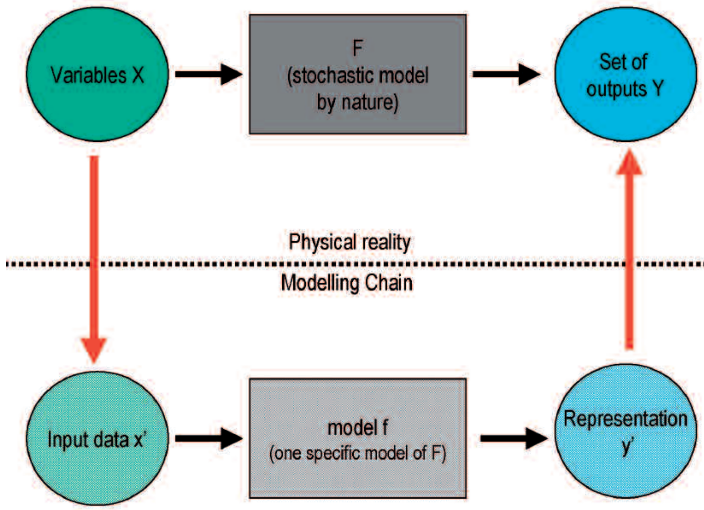

Figure 1. Representation of the physical reality by the mathematical modelling.

This representation highlights that there are differences between the reality and the outputs of the model, through the 3 concepts that are uncertainty, variability and error. To illustrate these concepts, let us take a practical example: the realization of a survey (equivalent to the model of dispersion) to determine the mean age of the population of a country (equivalent with the annual average concentration). The input data correspond only to the probed people:

- Variability: the age of each probed person is a variable (or stochastic) parameter since it corresponds to a single value in the interval 0-120 years, and since the age of each individual is independent of the age of the other probed people. Increasing the number of probed people will not reduce the range of the possible values for the age of an individual.
- Uncertainty: the average age of the country corresponds to an uncertain parameter since the range of the possible values of this average age will be reduced if the number of probed people increases.
- Error: the method of survey to determine the average age of the population can induces a systematic error if, for example, the survey is made only with adult people.

For uncertainties, some authors consider the following classification:

- The natural uncertainty associated to the stochastic nature of turbulence. One can also speak about “variability” or “external uncertainty”.
- The uncertainty of formulation of the model such as simplifications in the mathematical description of the physical and chemical processes, parameterisations, numerical methods, discretization in time and in space… One can also speak about “internal uncertainty” or “model uncertainty”.
- Uncertainty of the input data such as meteorological data, emission data…
- Uncertainty in the validation of the models related to the stochastic nature of the observations, i.e. the representativeness of the result y' compared to the physical reality. One can indeed compare y' only with a measurement of Y which is in general a sample or averaged value of the true value of Y.

In addition to uncertainties, some errors can be introduced by the user during modelling. They are generally of two types:

- the error of use related to the ignorance of the modelling tool or to the lack of experience of the user,
- the error of decision related to the choice made by the user to take into account a parameter or process in the modelling methodology (for example, the choice to use a space or temporal averaged data to reduce the computing time).

3. SOURCES OF UNCERTAINTIES

The aim of this work is mainly:

1. to identify the possible sources of uncertainties,
2. to characterize, and if possible to quantify and classify these sources of uncertainties,
3. to inventory methodologies for evaluation of the models against observations, methodologies of sensitivity analysis (necessary for item 2) and methodologies of uncertainty analysis (necessary also for item 2), in order to develop a general methodology of analysis and quantification of uncertainties.

In order to carry out an exhaustive inventory of the sources of uncertainties, the modelling process was divided into four principal steps, which are (see figure 2):

- Input data with uncertainties in the measurement of data, in their treatment, and their representativeness. One distinguishes here:
  - Geographical data (relief, land use, buildings)
  - Meteorological data at various spatial and temporal scale
  - Emission data (pollutant released, location and geometry, kinetics of release, mass rate…)
- The modelling of the meteorological field, with uncertainties associated with the intrinsic nature of the processes involved and with the quality of the models used to describe these phenomena. One also highlights the variable sensitivity of the models to the input data, which results in amplifying or in attenuating uncertainty on the final result.
- The modelling of atmospheric dispersion and the physicochemical transformations: one distinguishes the Gaussian, Lagrangian and Eulerian models because the difference in the formulation of these modelling approaches induces different kind of uncertainties.
- Statistical processing of the results: it is shown that the choice of the output variables (annual average or maximum value, value of the maximum on the ground or position of this maximum…) modify in an important way the uncertainty of the results and how to evaluate it.

4. EXAMPLE OF UNCERTAINTY ASSOCIATED TO THE WIND VELOCITY

In order to illustrate the adopted methodology to carry out the inventory of the sources of uncertainty, we choose to describe the case of the wind velocity and of the process of advection. For that, we will present the four steps described previously:

Step 1: input data

The wind velocity is always an input data in an atmospheric dispersion model. The first source of uncertainty is the space and temporal representativeness of the wind velocity: is the provided value representative of the site of interest or is it measured on a distant meteorological station several kilometres away? The frequency of the provided data (1h, 3h, 6h…) does it make it possible to represent the variability of the processes on the site? For example, data from national weather centres are averages over the 10 last minutes of the hour whereas a radio-sounding is available only all the 12 hours.

The second source of uncertainty is the precision itself of the velocity value. Obtained by measurement or resulting from a weather model on a large scale, velocity value is always affected by error and uncertainty. For example, the wind velocity provided by national weather centres is generally given with a resolution of 1 ms^{-1}.

Step 2: modelling of the meteorological field

The velocity data provided as input can be or not used in a meteorological model to reconstitute a more complete wind field. One distinguishes:

- 0D models, when the velocity is used such as it.
- 1D models, which reconstitute a vertical velocity profile.
- 3D models, which reconstitute all the complexity of the wind field:
  - Diagnostic models, of type “mass-consistent” or linearized.
  - Prognostic models, for forecast or specific application.

In this case, uncertainty depends on the quality of the selected model, but also of the adequacy between the model and the situation to describe (presence of relief or flat ground? Existence or not of a thermal flow? Etc…).

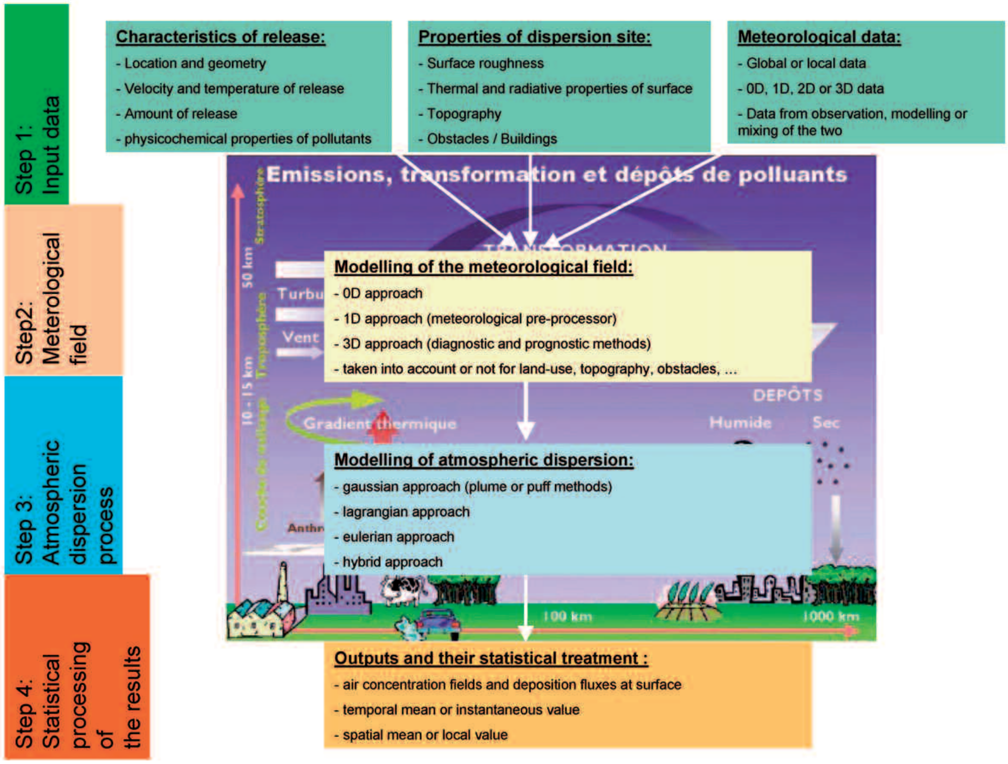

Figure 2. Modelling process divided into four principal steps.

Step 3: modelling of the dispersion process

Influence of the velocity on dispersion is due to the process of advection. One generally considers that for a stationary release, the pollutants concentration is inversely proportional to the wind velocity, so that uncertainties of steps 1 and 2 are reflected directly on the concentration. In the case of a non stationary release, the wind velocity influences mainly the time of advection of the cloud of pollutant: an uncertainty on velocity is transposed directly on the time of arrival of a toxic cloud and can lead to very important differences in concentration at a given time. Lastly, the velocity shear generates a more important spreading of the cloud of pollutant. If this shear is not described by the data (if there is a unique velocity data) or is not taken into account in the model of dispersion (for example in a simple Gaussian puff model), a complementary uncertainty will appear in the values of concentration.

Step 4: statistical processing of the results

Various uncertainties relating to steps 1, 2 and 3 can be modulated by the choice of the output variable or by the choice of the performed statistical treatment. For example, it is obvious that the time of arrival of a non stationary toxic cloud is directly affected by the velocity uncertainty but in another hand, the maximum values observed on the ground are rather independent to this uncertainty. Another example is the attenuation of uncertainty when the duration of average is increased.

5. TREATMENT OF UNCERTAINTIES

The example of the wind velocity illustrates the complexity and the interdependence between the physical data, the processes, and their representation in the models. Thus, it is often difficult to evaluate a global uncertainty for the whole modelling chain. Because of this difficulty, it is necessary to use an exhaustive and rigorous methodology of inventory of the sources of uncertainties before to apply specific methods to treat uncertainties.
Basically, there are two approaches:

- Sensitivity analysis which consist to study how a modelling chain (or only one part of this chain, such as the meteorological modelling for example) react to a variation of one or several uncertain parameters. This approach allows to identify and to order the most influential parameters (including branch of the atmospheric dispersion modelling) for which, it is then important to focus and reduce uncertainty.
- Uncertainty propagation, which consist to quantify the effect of uncertainties (generally considering only input data) on the outputs. The outputs are then associated to some indicators such as intervals or percentage of confidence. Many methods can be used (probabilistic or deterministic methods, see De Rocquigny *et al.*, 2008) which depends on the measure of the uncertainty we want, but also to the studied case (composed of the

modelling chain, the uncertain inputs and the quantity and variable of interest for output) and the objective of the studies (quantification of the uncertainties is not necessary the same in an environmental impact study compared to a risk study, or in an operational real-time application). We can quote the well-known Monte Carlo approach, which is probably the most complete approach, but also the most CPU-time consuming and most difficult to apply (need a full identification and quantification of the uncertain parameters).

6. CONCLUSION

This article presents a synthesis of an inventory of the sources of uncertainties in the modelling of atmospheric dispersion (Brocheton *et al*, 2008).

At the end of this work, it is right now possible to identify 4 axes of improvements of the modelling process:

- Improvement of the expertise of the users in the knowledge of the processes and of the models. It requires particularly to generalize the redaction of guidelines of good modelling practices.
- Reduction of identified uncertainties, in particular for the input data and by the improvement of models quality.
- Quantification of residual uncertainties, in order to be able to bring this information in complement of the model results.
- Integration of the concept of uncertainty in the strategy of analysis of the results, in order to use output variables, which are least sensitive to uncertainties.

REFERENCES

Baird B.F., 1989: Managerial Decisions Under Uncertainty, New-York, John Wiley & Sons Ltd.

Brocheton, F., P. Armand, L. Soulhac and E. Buisson, 2008: Identification des sources d'incertitudes dans les simulations du transport atmosphérique. *CEA/DAM Technical Report* n°**R/252.0707/MAE**, 116 pp.

Colville, R.N., N.K. Woodfield, D.J. Carruthers, B.E.A. Fisher, A. Rickard, S. Neville and A. Hughes, 2002: Uncertainty in dispersion modelling and urban air quality mapping. *Environmental Science and Policy*, **5**(3), 207-220.

Dabberdt, W.F. et E. Miller, 2000: Uncertainties, ensembles and air quality dispersion modeling: applications and challenges. *Atm. Env.*, **34**(27), 4667-4673.

De Rocquigny, E., N, Devictor, N. Perot, S. Tarantola, W. Castaings, F. Mangeant, C. Schwob, R. Bolado-Lavin, J-R. Massé, P. Limbourg, W. Kanning and P.H.A.M. Van Gelder, 2008: Uncertainty in industrial practice: A guide to quantitative uncertainty management. Editors Etienne de Rocquigny, Nicols Devictor and Stefano Tarantola, John Wiley & Sons publishers, 339 pp..

Godoy, S.M., A.S.M. Santa Cruz and N.J. Scenna, 2006: Risk analysis and safe distances calculation considering atmospheric parameters uncertainty. *Latin American Applied Research*, **36**: 263-268.

Hanna, S.R., J.C. Chang, and M.E. Fernau, 1998: Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables. *Atm. Env.,* **32**, 3619-3628.

Irwin, J.S., C. Kakkarinen, and H. Felman, 2001: Uncertainty in airquality modelling for risk calculations. Guideline on Air Quality Models: A New Beginning. *Air& Waste Management Association*, Newport, RhodeIsland.

Levy, J.I., J.D. Spengler, D. Hlinka, D. Sullivan and D. Moon, 2002: Using CALPUFF to evaluate the impacts of power plant emissions in Illinois: model sensitivity and implications. *Atm. Environment*, **36**, 1063-1075.

Mallet, V. and B. Sportisse, 2006: Uncertainty in a chemistry-transport model due to physical parameterisations and numerical approximations: An ensemble approach applied to ozone modelling. *J. Geophys. Res.*, **111**(D01302), 538-543.

Perkins, R., L. Soulhac and P. Mejean, 2002: Modélisation de la dispersion des émissions atmosphériques d'un site industriel – Vers un guide de l'utilisateur. 1ère Partie: Etat de l'art, Association R.E.C.O.R.D. edn, Laboratoire de Mécaniques des Fluides et d'Acoustique, UMR CNRS 5509 – UCBL1 – ECL.

Roache, P.J., 1997: Quantification of uncertainty in computational fluid dynamics. *Annu. Rev. Fluid. Mech.*, **29**, 123-160.

Stein, A. and J. Wyngaard, 2002: Fluid Modelling and the Evaluation of Inherent Uncertainty. *J. Applied. Meteor.*, **40** (10), 1769-1774.