# Screening sensitivity analysis of a radionuclides atmospheric dispersion model applied to the Fukushima disaster

Sylvain Girard[a,*], Irène Korsakissok[a], Vivien Mallet[b,c]

[a] *Institut de radioprotection et de sûreté nucléaire, 31, avenue de la Division Leclerc, 92260, Fontenay-aux-Roses, France*
[b] *Inria, Domaine de Voluceau, BP 105, 78153, Le Chesnay cedex, France*
[c] *CEREA, Joint Laboratory École des Ponts ParisTech/EDF R&D, Université Paris Est, Marne-la-Vallée, France*

**Abstract**

Numerical models used to forecast the atmospheric dispersion of radionuclides following nuclear accidents are subject to substantial uncertainties. Input data, such as meteorological forecasts or source term estimations, as well as poorly known model parameters contribute for a large part to this uncertainty.

A sensitivity analysis with the method of Morris was carried out in the case of the Fukushima disaster as a first step towards the uncertainty analysis of the Polyphemus/Polair3D model. The main difficulties stemmed from the high dimension of the model's input and output. Simple perturbations whose magnitudes were devised from a thorough literature review were applied to 19 uncertain inputs. Several outputs related to atmospheric activity and ground deposition were aggregated, revealing different inputs rankings. Other inputs based on gamma dose rates measurements were used to question the possibility of calibrating the inputs uncertainties.

Some inputs, such as the cloud layer thickness, were found to have little influence on most considered outputs and could therefore be safely discarded from further studies. On the contrary, wind perturbations and emission factors for iodine and caesium are predominant. The performance indicators derived from dose rates observations displayed strong sensitivities. This emphasises the share of the overall uncertainty due to input uncertainties and asserts the relevance of the simple perturbation scheme that was employed in this work.

*Keywords:* sensitivity analysis, atmospheric dispersion, Fukushima, Polyphemus/Polair3D, Morris method

*Corresponding author.
*Email addresses:* sylvain.girard@irsn.fr (Sylvain Girard), irene.korsakissok@irsn.fr (Irène Korsakissok), vivien.mallet@inria.fr (Vivien Mallet)

**Contents**



## 1. Introduction

Numerical simulations of the atmospheric dispersion of radionuclides are used during the early stages of nuclear accidents as input to the decision making. They also provide a valuable complement to field measurements for the long term assessment of environmental and sanitary impact, as illustrated by the cases of the Chernobyl and Fukushima disasters.

The meteorological fields fed into the model are typically issued from operational forecasts by meteorological models and involve substantial uncertainties. The source term itself is also subject to high uncertainties, even several years after the accident. For instance, several estimations of the atmospheric release induced by the Fukushima Daiichi power plant have been proposed after the

crisis with the help of environmental data (see for instance Terada et al., 2012; Stohl et al., 2012; Saunier et al., 2013; Winiarek et al., 2014). Despite the amount of field measurements, and the better understanding of the installation events, the range of variation in these source terms show that the knowledge of the release rate and kinetics is still partial and uncertain. Other important sources of uncertainty lie in the dry deposition, the wet scavenging, the computation of the vertical diffusion coefficient, and possibly the numerical schemes for the integration of the transport equations.

All these elements have an influence on the output of the model and induce uncertainties which undermine predictions solely based on a deterministic approach. The present study is a first step in an effort to account for uncertainties of the Polyphemus/Polair3D model in predicting the dispersion of an accidental release of radionuclides in the atmosphere. It is difficult to study the model in a fully generic context because its input include complex spatio-temporal fields. The case studied here is the atmospheric release of radionuclides following the Fukushima Daiichi disaster.

Below is a rough outline of how the uncertainty characterisation could be carried out:

1. Determine the main sources of uncertainty and select the input variables to the model that adequately represent them.

2. Define model output variables relevant to crisis management or long-term impact evaluation.

3. Model the uncertainty of each input variable by a random variable with given probability distribution.

4. Propagate the uncertainty with a Monte Carlo scheme by sampling from the probability distributions built at step 3.

5. Use available observations to assess the choice of input variables and calibrate the associated uncertainty models.

This process may be iterated until the output uncertainty is consistent with available observations.

There are several issues that arise when dealing with detailed environmental models which are often of high dimensionality and computationally demanding. The raw outputs of the dispersion model are spatio-temporal fields of radionuclides concentrations or gamma dose rates. Simply constructing confidence intervals for each species at each time step and location would be fastidious and weakly informative. In addition, ignoring spatio-temporal correlations is likely to deteriorate the uncertainty estimates, a fact that geostatisticians or practitioners of data assimilation are familiar to. Hence, step 2 of the procedure above can be seen as a problem of dimension reduction. The objective of this step is to derive new model outputs of sufficiently low dimension to allow for computation and interpretation while preserving most of the information carried by a spatio-temporal analysis.

Step 3 is particularly challenging when complex inputs, such as meteorological fields, are involved. High dimensional inputs are indeed difficult to handle, especially when they display spatial correlation, temporal correlation or singularities that are structurally characteristic of the physical phenomenon at hand. Precipitation fields for instance are made of patches of varying shape that appear, deform and move over time, which cannot be modelled by a simple probability distribution. The emitted amount of a given species seen as a time series displays strong auto-correlation but also very temporally localised peaks. Additionally, several fields are constrained by physical relations, such as wind fields that need to satisfy the continuity equation. The choice of input variables and their uncertainty description are set out in section 4.

Given these difficulties, the observations mentioned in step 5 are an invaluable assessment tool. They may intervene for instance to ensure that no major source of uncertainty was left aside or to appreciate the quality of the input uncertainty descriptions.

The details of step 4 will be relevant when the actual problem of uncertainty analysis will be tackled. For now, the present paper deals with *sensitivity* analysis, an approach differing in its objectives, but related to uncertainty quantification (Saltelli et al., 2008). The rationale for this preliminary step is that undertaking the issues evoked above all at a time seemed too complicated. The generic motive of sensitivity analysis is to quantify the relative influence of a set of inputs on the output of a model. The method employed here and detailed in section 3 belongs to the *screening methods* category which aims at classifying input variables into influential and negligible with a view of reducing the computational burden for further studies by setting aside those of smaller influence. While the focus is clearly on step 1 of the procedure given above, this work constitute a starting point in the reflection upon the subsequent problem of uncertainty quantification, especially steps 2 and 3 but also step 5, as will be seen in section 5.4. The results of the sensitivity analysis are presented in section 5.

## 2. Polyphemus/Polair3D

The atmospheric dispersion of the radionuclides is carried out with the air quality modeling system Polyphemus (Mallet et al., 2007) and its Eulerian transport model Polair3D. Polair3D is essentially a numerical solver for a system of 3D advection-diffusion equations. The equation of this system for a given radionuclide denoted by a subscript $r$ reads

$$\frac{\partial c_r}{\partial t} + \operatorname{div}(\boldsymbol{w} c_r) = \operatorname{div}\left(\rho \mathbf{K} \nabla \frac{c_r}{\rho}\right) - \mathbf{F}\boldsymbol{c} + E_r - \Lambda c_r, \tag{1}$$

where $c_r$ is the concentration in the air, $\boldsymbol{c}$ the vector of the concentrations of all considered radionuclides linked a matrix $\mathbf{F}$ of decay coefficients, $\boldsymbol{w} = (w_u, w_v, w_z)^T$ the wind velocity, $\rho$ the air density, $\mathbf{K}$ the turbulent diffusion matrix assumed to be a diagonal matrix with diagonal $(K_u, K_v, K_z)$, $E_r$ is the emission source term and $\Lambda$ the scavenging coefficient. On the ground, the

boundary condition reads $\rho \mathbf{K} \nabla \frac{c_r}{\rho} \cdot \boldsymbol{n} = v_d c_r$, where $\boldsymbol{n}$ is the normal to ground oriented towards higher altitudes and $v_d$ is the deposition velocity.

The equation is solved using first-order operator splitting, with diffusion integrated after advection. The advection scheme is a third-order direct-space-time scheme with flux limiting (Verwer et al., 2002). The spatial resolution is 0.125° and the numerical time step is 10 min. The simulations are carried out with 10 vertical layers, whose center altitudes are 20 m, 100 m, 220 m, 340 m, 500 m, 700 m, 1000 m, 1500 m, 2200 m and 3000 m.

## 3. Morris method for sensitivity analysis

Sensitivity analysis is the study of how variations in the inputs of a model affect its outputs. Here, the word *model* refers to any deterministic process that can be associated to a mathematical application mapping a set of input variables to one output value. The case of multivariate outputs is usually handled one variable at a time.

*Local* sensitivity analysis is concerned with the response of the model in the vicinity of a reference point. In this respect, it pertains to Taylor expansion and derivatives approximation. Should the model response be resolutely non-linear, extrapolation of the local sensitivity measures to regions far from the reference point are likely to be seriously flawed (Saltelli and Annoni, 2010). By contrast, *global* sensitivity analysis aims at estimating the relative importance of the inputs over their whole domain of variation.

Another desirable feature of a sensitivity analysis method is its ability to estimate *interactions*. Interactions are effects that appear when two or more inputs vary simultaneously. For instance, variations in the wind direction or delays in emissions can induce the plume to avoid a rain event at some location, which may remove any sensitivity to the rain intensity at the given location. The rain intensity is therefore in interaction with the wind and the emissions. Our purpose here is to sieve the inputs and eliminate the least influential from further studies. In this context, estimating interactions is a required safeguard against type II error, namely classifying a variable as non-influential when it actually has a non negligible impact on the output (Saltelli et al., 2008, p. 15 and 110).

The method of Morris (1991) allows for global sensitivities and interactions estimations while requiring relatively few model evaluations to be robust.

### *3.1. From elementary effects to global sensitivities*

Thereafter, the model response to given inputs, represented by the vector $\boldsymbol{x}$ of size $m$, will be denoted $y(\boldsymbol{x})$. Let $\boldsymbol{x} = (x_i)_i$ be a $m$-dimensional vector representing a reference set of inputs and $\boldsymbol{x}_{\neg i}$ the vector constituted of all $\boldsymbol{x}$ components but $x_i$. Morris (1991) names *elementary effect caused by the $i$-th input* the following ratio obtained by perturbing only the $i$-th component of $\boldsymbol{x}$ with a quantity $\delta$:

$$d^{(i)} = \frac{y(\boldsymbol{x}_{\neg i}, x_i + \delta) - y(\boldsymbol{x}_{\neg i}, x_i)}{\delta}. \tag{2}$$

The first step in computing sensitivities is to build a *design of experiment*. An *input space* is first defined by choosing the input variables and the values they are allowed to take. Then, a set of points in the input space where the model will be evaluated are selected. The commonly used one-at-a-time design of experiment consists in choosing a reference point and computing one elementary effect per input variable. This is a local sensitivity analysis method which ignores interactions. In order to overcome these shortcomings, Morris (1991) suggested to randomise the reference point and to examine the distribution of elementary effects by sampling it. He proposed to use the first two moments of the distribution as sensitivity measures, which allows to classify the input variables into three groups: 1. those of negligible influence are characterised by a low mean of elementary effects; 2. those of strong linear influence without interaction have important mean and low standard deviation; and 3. those of strong influence, either non-linear or strongly interacting, have high mean and high standard deviation. Hereafter, the empirical mean and standard deviation of the elementary effects caused by the $i$-th variable will be denoted $\mu_i$ and $\sigma_i$ respectively.

The mean of elementary effects as a measure of sensitivity can be deceiving when the response is not monotone. Campolongo et al. (2007) have devised an alternative measure, the average of the absolute value of elementary effects denoted $\mu_i^\star$, that addresses this issue:

$$\mu_i^\star = \frac{1}{n}\sum_{k=1}^{n}|d_k^{(i)}|, \tag{3}$$

where the summation is carried over the sample of elementary effects. Input variables of substantial influence necessarily have strong $\mu^\star$ while their $\mu$ can be low due to elementary effects of opposing signs cancelling out. When both $\mu^\star$ and $\mu$ are high, the sign of $\mu$ indicates the direction of the effect.

### *3.2. Algorithm*

The input variables can be assumed here to be all uniformly distributed on $[0, 1]$ without loss of generality. Indeed, transformations for scaling each input to its proper range of variation and unit can be included in the model. Similarly, the inverse transform method can be used to sample another distribution using the uniform samples. It is almost always preferable to compute elementary effects with the unscaled uniformly distributed denominator in equation (2), because it allows to compare elementary effects of variables with very different ranges of variation.

While elementary effects associated to the same variables need to be independent, elementary effects of different variables can be correlated without compromising the sensitivity measures. Taking advantage of this, more sparing designs have been proposed by Morris (1991) and further improved by Campolongo et al. (2007). The sampling method used in this paper requires a total $(m+1)\times r$ model evaluations to achieve a sample size of $r$ elementary effects per

input instead of the $2m \times r$ required when simply stringing together one-at-a-time design. This is described in further details in AppendixA.

Each input can take $k$ different values, called *levels* in the following. The first level corresponds to the lower bound of the variation interval of the input. The last level coincides with the upper bound of this interval. In between, the levels are regularly distributed in the interval. The number of levels should be selected consistently with the computational budget and the expected degree of irregularity of the model. It is no use increasing the number of levels without concurrently increasing the sample size because this would result in a poor coverage of the input space. A restricted number of levels is usually sufficient when the response surface is monotone, smooth and not highly non-linear. In order to ensure convergence of the sensitivity measures, the number of levels was set to $m = 8$ and the sample size to $r = 100$ per input variable. The results presented in section 5 were obtained with this safe but somewhat costly design of experiment. Simulations with the more common choice $(m, r) = (4, 10)$ (Saltelli et al., 2008, p. 119) were conducted too. Section 5.5 provides a brief comparison of the results obtained with the two designs of experiment. Finally, the trustworthiness of the estimators was monitored with bootstrap confidence intervals.

## 4. Design of experiment

The coefficients of the transport equations of the form (1) are obtained either directly from input data or from physical parametrisations available in Polyphemus. Meteorological input data are provided by the European Centre for Medium-Range Weather Forecasts (ECMWF). The fields are from the 12-hour forecast cycles starting from analysed fields at 00:00 UTC. They have a resolution of 0.36° horizontally, 60 sigma-levels vertically and a timestep of 3 h. Sections 4.1 to 4.5 describe 12 of the 19 inputs of the sensitivity analysis that represent uncertainty in the parametrisation coefficients and meteorological input data of Polyphemus/Polair3D. The remaining 7 input variables relate to the source term and are dealt with in section 4.6

All input variables were assigned a uniform distribution. At this stage, the aim is not to accurately model the actual uncertainty of the inputs, but rather to cover evenly the input space, in order to get a global understanding of the model sensitivity. The inputs ranges of variations are summarised in table 1.

### *4.1. Wet scavenging*

The parametrisation of the scavenging coefficient is of the form $\Lambda_r = ap^b$, where $p$ is the rain intensity in $\mathrm{mm\,h^{-1}}$. Different coefficients describe the scavenging below and inside the clouds. The choice of simple exponential models with only two parameters to represent both below-cloud and in-cloud scavenging is driven by the poor knowledge of micro-physical parameters, such as the particulate size distribution, that are required by more detailed models. These uncertainties are reflected in the wide intervals presented in table 1 which

encompass the values found in the literature for comparable settings (Sportisse, 2007). All species were attributed the same $a$ and $b$ coefficients. We also applied the same perturbation on those coefficients for all species in order to limit the number of input variables. The sensitivity to wet scavenging with this clustering is assumed be greater than the aggregation of sensitivities with individual perturbations. This is a safeguard against the risk of screening out an important input.

#### *4.1.1. Cloud layer*

The cloud layer base height affects below-cloud scavenging, and the cloud layer thickness affects in-cloud scavenging. Both were perturbed by random multiplicative factors, which induce an increase in variability with altitude and thickness. Indeed, the upper part of the atmosphere was deemed less well characterised, a fact already reflected by the increase of the vertical layers thickness with the altitude. Since not many references to these uncertainties were found in the literature, these inputs were conservatively attributed wide ranges of variation.

#### *4.1.2. Precipitation*

Hanna et al. (2001) elicited from a panel of experts some estimates of uncertainty ranges for several meteorological input parameters in the context of photochemical modelling. They suggest to model rain intensity with a log-normal distribution. The idea of a multiplicative perturbation was retained in the present study but a uniform distribution was preferred to a log-normal to ensure a sufficient sampling of the extreme levels. Indeed, using another distribution, especially one with a strong mode, could result in degenerate cases in practice due to the limited sample size.

The bounds of the support of the uniformly distributed perturbation factor were set in agreement with the recommendations from Hanna et al. (2001) but rounded off to one significant digit.

### *4.2. Dry deposition*

The range of variation of dry deposition velocity is derived from a literature review (Maryon et al., 1991; Thykier-Nielsen et al., 1999; Baklanov and Sørensen, 2001; Brandt et al., 2002). All species were assumed to have the same dry deposition velocity. As for wet scavenging, a single perturbation for all species was used.

### *4.3. Horizontal diffusion*

The horizontal diffusion coefficients, $K_u$ and $K_v$ (first and second diagonal elements of $\mathbf{K}$), are constant and homogeneous. Not many references on the determination of horizontal diffusion coefficients were found in the literature. The range displayed in table 1 encompasses typical values found in the literature (Brandt et al., 1998; Ryall and Maryon, 1998; Sørensen, 1998; Yamartino, 2000)

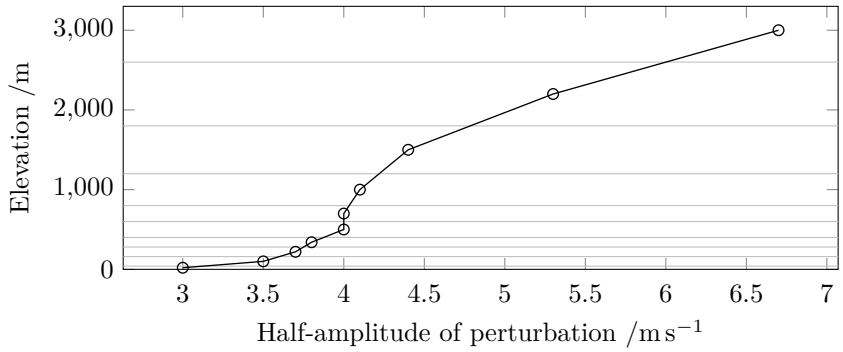


Figure 1: Vertical profile of the half amplitude of wind components perturbation. The grey lines indicates the boundaries of the vertical layers of the model.

### *4.4. Vertical diffusion*

The vertical diffusion coefficient, $K_z$, is computed by Louis (1979) parametrisation above the boundary layer and inside the stable boundary layer. It is computed with the Troen and Mahrt (1986) parametrisation inside the unstable boundary layer. Hanna et al. (2001) suggest an uncertainty factor of 3 with a log-normal hypothesis to encompass 95 % of the possible values. This leads, after rounding off to one significant digit, to the bounds in table 1.

### *4.5. Wind*

The horizontal wind is directly derived from the data, but the vertical wind is computed so that it satisfies $\mathrm{div}(\rho \boldsymbol{w}) = 0$. Wind perturbations have to meet this requirement too. Hanna et al. (2001) dealt with this constraint by perturbing the speed of the wind field by a factor and its direction by an angular increment, both homogeneously throughout the spatial domain. They suggest to model the wind speed by a log-normal distribution with an uncertainty range of plus or minus a factor of 1.5 and the direction by a normal distribution with an uncertainty range of plus or minus 40°, both encompassing 95 % of the possible values. In the present study, the zonal and meridional components of the wind were instead additively perturbed by random increments uniformly distributed on symmetric intervals of equal length. In order to account for the mean wind speed while using additive perturbations, different interval lengths were attributed to each vertical level. The vertical profile of the upper bound of the perturbation intervals is displayed in figure 1. Details about the methodology used to determine these values in relation with the recommendations from Hanna et al. (2001) are given in AppendixB.

### *4.6. Source term*

The source term devised by Mathieu et al. (2012) was used here. It contains emission rates for 73 species that were grouped into four families. The caesium family includes $^{134}$Cs, $^{136}$Cs, $^{137}$Cs, $^{138}$Cs as well as $^{137}$Ba because it is in secular

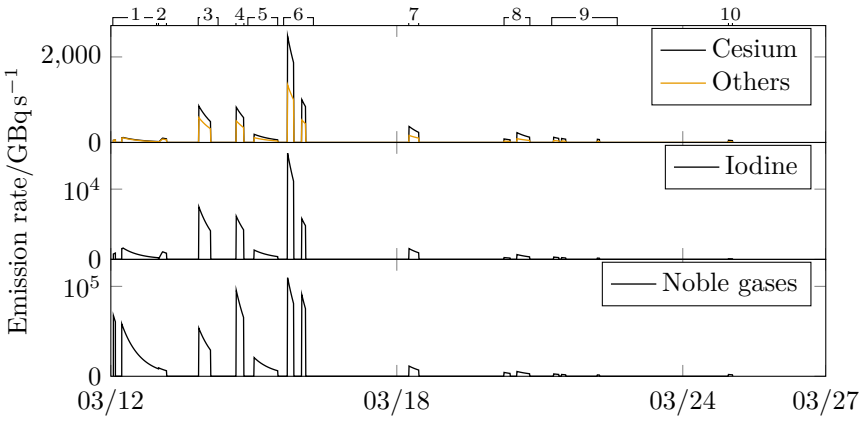


Figure 2: Emission rates of the four families of emitted species as functions of time. The numbered brackets at the top delimit the emission events identified by Mathieu et al. (2012).

equilibrium with $^{137}$Cs. The iodine family includes $^{131}$I, $^{132}$I, $^{133}$I, $^{134}$I and $^{135}$I, both in molecular and aerosol forms, as well as $^{132}$Te because it is in secular equilibrium with $^{132}$I. The noble gases family includes $^{85}$Kr, $^{87}$Kr, $^{88}$Kr, $^{133}$Xe, $^{135}$Xe and $^{138}$Xe. The last family includes all other remaining species that are supposed to have been emitted during the Fukushima disaster.

The emission rates of the four families are displayed in figure 2. The emission events were derived from plant measurements (water level and pressure in the reactor vessel, pressure in the containment) and the chronology provided by the Tokyo Electric Power Company (TEPCO) for events such as containment venting or onset of smoke. This timeline was further modified to comply with the most significant gamma dose rate peaks measured by on-site monitoring devices. A detailed breakdown of these events was proposed by Korsakissok et al. (2013). They are indicated in figure 2 by the numbered brackets at the top of the plots.

The source term is perturbed by an additive time shift, a shift of the model layer in which the radionuclides are released, and multiplicative factor on the emission rate of each of the four families. The time and vertical shifts are the same for all species. Since Polair3D is an Eulerian model, only the emission layer matters, not the exact emission altitude. Each of the four levels in our application of Morris method corresponds to a different vertical layer. The range of variation of the four emission factors, time shift, and emission altitude were set to high values so as to reflect the inherent inaccuracy of available data, as emphasised by Mathieu et al. (2012).

## 5. Results

The method described in section 3 applies to scalar output models. However, Polyphemus/Polair3D outputs are spatio-temporal fields of gamma dose rates whose dimensions are very high. Handling values at each location and each

| Variable | Lower bound | Upper bound | Symbol |
|---|---|---|---|
| Scavenging factors/$\mathrm{h\,s^{-1}\,m^{-1}}$ | 0.05 | 0.5 | $a_b$ (below-cloud), $a_i$ (in-cloud) |
| Scavenging exponent | 0.6 | 1 | $b_b$ (below-cloud), $b_i$ (in-cloud) |
| Precipitation$^{(\times)}$ | 0.5 | 2 | $p$ |
| Clouds base height$^{(\times)}$ | 0.667 | 2 | $C_h$ |
| Clouds thickness$^{(\times)}$ | 0.5 | 2 | $C_t$ |
| Dry deposition velocity/$\mathrm{mm\,s^{-1}}$ | 0.5 | 5 | $v_d$ |
| Horizontal diffusion/$10^4\,\mathrm{m^2\,s^{-1}}$ | 0 | 1.5 | $K_u$ (zonal), $K_v$ (meridional) |
| Vertical diffusion$^{(\times)}$ | 0.333 | 3 | $K_z$ |
| Winds$^{(+)}$/$\mathrm{m\,s^{-1}}$ | see figure 1 | | $w_u$ (zonal), $w_v$ (meridional) |
| Emission factors $^{(\times)}$ | 0.333 | 3 | $E_{Cs}$ (caesium), $E_I$ (iodine), $E_g$ (noble gases), $E_o$ (others) |
| Emission delay$^{(+)}$/h | $-6$ | 6 | $\Delta t$ |
| Source elevation | 1st layer | 4th layer | $z$ |

Table 1: Input variables, the bounds of their range of variation and the symbols that represent them later. Generally the levels indicate the actual values taken by the inputs. A variable whose name bears a $^{(\times)}$ superscript is multiplied by a perturbation inside the prescribed bounds. A variable whose name bears a $^{(+)}$ superscript is perturbed by adding a value inside the prescribed bounds. In the application of Morris method, the lowermost and uppermost levels correspond to the bounds displayed in this table. The intermediate levels are regularly distributed across the interval.

time step as independent outputs would yield too many sensitivity indices for convenient interpretation. Additionally, consecutive values at a given point are highly correlated and the output fields exhibit strong spatial correlation structures. It is thus necessary to derive new outputs of lower dimension, each one focusing on a feature of the initial data. The chosen model outputs are ambient dose rates in the atmosphere, restrained to the first vertical layer of the simulation domain, which is of interest for population exposure. Dose rates due to radionuclides deposed on the ground (hereafter called deposit dose rate) are also retained, being the main source of exposure to radiations after the end of the release. For each output, the most relevant sensitivity measures are displayed: the averaged of the elementary effects $\mu$ or its equivalent with absolute values $\mu^\star$, depending on whether the signed $\mu$ could be used without loss of information. The standard deviation of elementary effects $\sigma$ is also displayed to identify the non-linearity or interactions.

### *5.1. Aggregated outputs*

This section focuses on scalar outputs, aggregated both in time and space. They provide a synthetic view of the model sensitivity to variables, but they conceal any kinetic or geographic effect. For atmospheric dose rates, the spatial average of the temporal sums, noted $\langle \overline{\boldsymbol{y}}_a \rangle$, are used. The brackets refer to the spatial averaging, and the horizontal bar refers to time-integration. For deposit dose rate, already accumulated over time, the maximum value of the spatial average was selected and denoted $\max\langle \boldsymbol{y}_g \rangle$.

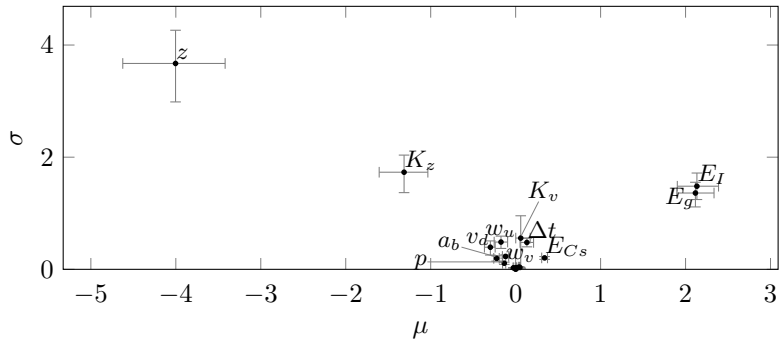


Figure 3: $\mu$ and $\sigma$ measures for the spatial average of the time-integrated atmospheric dose rate, $\langle \bar{\boldsymbol{y}}_a \rangle$, in $\mu\mathrm{Gy\,h^{-1}}$. Error bars represent 0.95 bootstrap confidence intervals.

Figures 3 and 4 represent the $\mu$ and $\sigma$ measures for these two outputs respectively. The relatively small 0.95 confidence intervals represented by the error bars show a good convergence of the estimations. In figure 3, the source altitude $z$, followed by the vertical diffusion coefficient $K_z$ and the emission factors for iodine $E_I$ and noble gases $E_g$, are the most influential variables since they have high $\mu$ in absolute value. The high value of $\sigma$, relatively to $\mu$, indicates strong interactions or non-linearity. Increasing the source altitude or the vertical diffusion coefficient tends to deplete the lower layer of the atmosphere, which decreases $\langle \bar{\boldsymbol{y}}_a \rangle$, resulting in a negative $\mu$. Obviously, emission factors have positive $\mu$ since they increase the dose rates.

Figure 4 shows a clear predominance of the emission factor for iodine on ground dose rate sensitivity. The $\mu$ for noble gases is null since they are not deposited. The source altitude and the vertical diffusion coefficient are of lesser importance for the deposit than the atmosphere, since radionuclides are scavenged by rain in all layers of the atmosphere. Rain intensity and scavenging factors have positive effects of similar magnitude, either non-linear or displaying interactions. Scavenging exponents have slight negative effects because the rainfalls that occurs near the source were of intensity below $1\,\mathrm{mm\,h^{-1}}$.

On both figures, iodine is predominant over caesium since it was emitted in higher quantities and has a higher dose coefficient. An overview of the least influential inputs for each considered output is provided in table 2.

### *5.2. Time-dependent outputs*

Figures 5 and 6 represent the $\mu$ computed at each time step for spatially averaged dose rates from the atmosphere, $\langle \boldsymbol{y}_a \rangle$, and deposit, $\langle \boldsymbol{y}_g \rangle$, respectively.

As expected, figure 5 presents positive peaks for emission factors at times corresponding to the release periods of figure 2. The most influential variables are the same as in figure 3, with the same signs for $\mu$. The time shift effect is negative before the peaks and positive after, since a delay in the release results in a forward time shift of the concentration peak. On figure 6, all $\mu$ values are very low until a substantial rain event occurred while the plume was widespread,

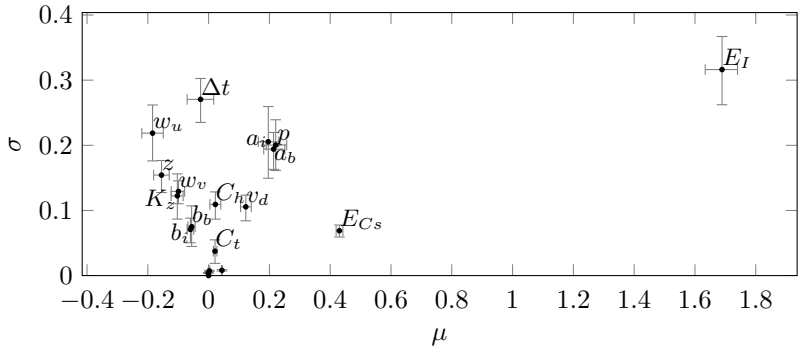


Figure 4: $\mu$ and $\sigma$ measures for the temporal maximum of spatially averaged ground dose rate, $\max\langle \boldsymbol{y}_g \rangle$, in µGy h$^{-1}$. Error bars represent 0.95 bootstrap confidence intervals.

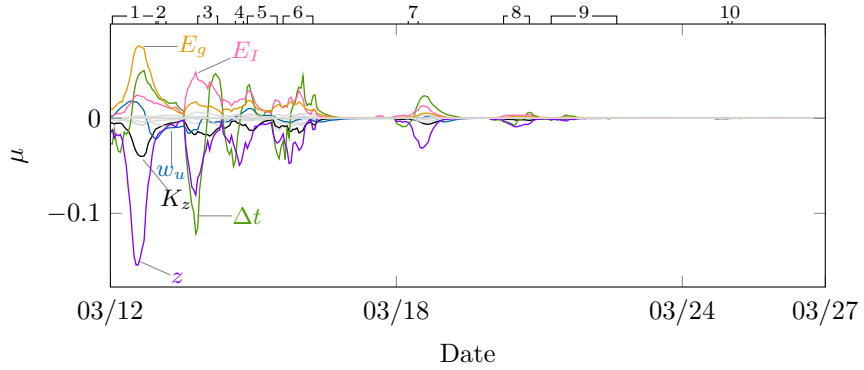


Figure 5: $\mu$ measures for the spatial average of atmospheric dose rate, $\langle \boldsymbol{y}_a \rangle$, in µGy h$^{-1}$. The numbered brackets on top correspond to the emission events from figure 2.

on March 15th. It is responsible for most of wet deposition, corresponding to the release event 5 (see figure 2) being scavenged within a few hours, as shown in Korsakissok et al. (2013). This explains the sharp negative peak for the time shift: delaying this emission makes the plume meet the rain event later. After this event, the influence of the emission factor for iodine and caesium both rise steadily before the former begins to drop due to radioactive decay. Having a much longer half-life, the latter remains almost constant until the end of the simulation. The variables found to have a minor influence for both the atmosphere and deposit are reported in table 2, and showed in gray in figures 5 and 6.

### *5.3. Spatial outputs*

Maps of sensitivity measures computed at each grid cell for time-aggregated outputs are given as supplementary material. Although some qualitative information could be derived from these maps, they cannot be easily interpreted

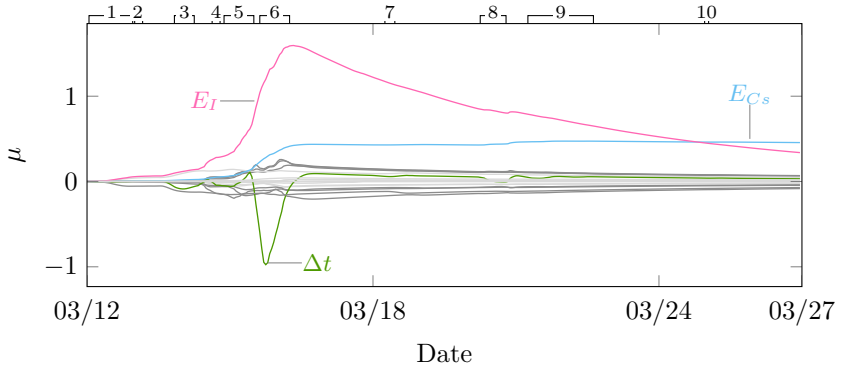


Figure 6: $\mu$ measures for the spatial average of the dose rate from the deposit, $\langle \boldsymbol{y}_g \rangle$, in µGy h$^{-1}$. The numbered brackets on top correspond to the emission events from figure 2.

quantitatively. Moreover, they carry redundant information because dose rates simulated at neighbouring locations are highly correlated. Thus, a criterion for input classification based on this kind of outputs would be difficult to construct and rather arbitrary. Principal component analysis proved to have some potential for efficient dimension reduction but the interpretation of the principal directions is not straightforward. This is further complicated by the very strong gradients that require a transformation or normalization of the data beforehand. For these reasons, a simpler approach was adopted. Indicator functions exploiting certain thresholds were considered in order to reduce the dimension without cancelling the information on spatial structures. Geometrical features such as the area of a region above a given threshold or the length of its contour may be considered. However, while these may be good indicators of the plume's spread, they fail to convey some valuable information about the shape or position of the region where the threshold is exceeded. In order to account for these, the following measure of dissimilarity, later referred to as *shape mismatch*, may be substituted to the notion of elementary effect given in equation (2):

$$\mathrm{sm}(y) = 1 - \frac{\mathrm{A}\left[\{y(\boldsymbol{x}) > \theta\} \cap \{y(\boldsymbol{x}_{\neg i}, x_i + \delta) > \theta\}\right]}{\mathrm{A}\left[\{y(\boldsymbol{x}) > \theta\} \cup \{y(\boldsymbol{x}_{\neg i}, x_i + \delta) > \theta\}\right]}, \tag{4}$$

where $\theta$ denotes the chosen threshold and $\mathrm{A}[\cdot]$ the area. The quantity defined in equation (4) is equal to 1 when the surface where the threshold is exceeded is unchanged by the input perturbation and equal to 0 if the changes are so big that there is no overlap between the reference and perturbed surfaces. As this measure of dissimilarity is always positive, the $\mu$ and $\mu^\star$ measures are equivalent.

Figures 7 and 8 display $\mu^\star$ and $\sigma$ measures for two thresholds for the atmospheric and deposit dose rates respectively. They were computed on time-aggregated outputs: the dose rates were time-integrated for the lower atmosphere, while the time maximum was used for the deposit, like in section 5.1. The thresholds were selected so that the proportion of the spatial domain where they are exceeded covers most possible values excluding very high and very low ones.

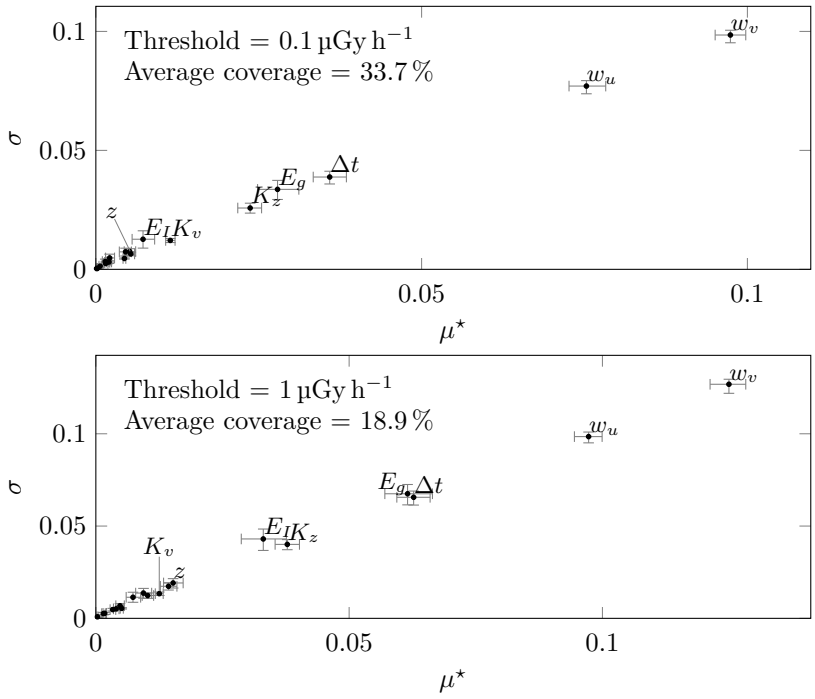


Figure 7: $\mu^\star$ and $\sigma$ measures based on shape mismatch of the exceedance zone of the time-integrated atmospheric dose rates, $\mathrm{ez}(\bar{\boldsymbol{y}}_a)$, for two thresholds, 1 µGy h$^{-1}$ and 0.1 µGy h$^{-1}$.

The average share of the simulation domain covered by the exceedance zone is indicated in the upper left corner of the figures. The winds and time shift are predominant for both output variables and thresholds. Emission factors are more important for the highest threshold while horizontal diffusion becomes noticeable for the lowest.

### *5.4. Involving observations*

Observations may be used in uncertainty analysis to calibrate the input probability distributions and assess the output uncertainty estimates. This paper focuses on a screening method whose aim is to classify the variables influence rather than precisely estimating their uncertainty. Still, the available samples of model evaluations can be exploited to get some insights on how to use observations later in the uncertainty analysis. Indeed, influential inputs where and when observational data are available are those whose uncertainty description can later be calibrated. Hence, particular attention should be devoted to variables that are preponderant for "classical" outputs but of subsidiary importance regarding model performance, and therefore, not subject to calibration. Possible responses to such a situation include using additional observations, altering the uncertainty descriptions, or, at least, attributing a wider range of variation to the problematic inputs so as to ensure conservative uncertainty estimates.

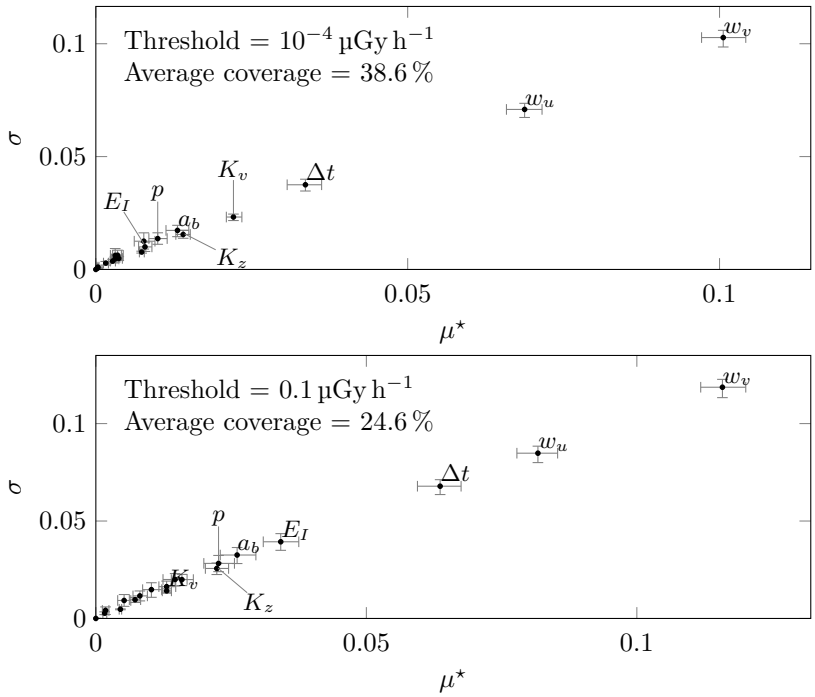


Figure 8: $\mu^\star$ and $\sigma$ measures based on shape mismatch of the exceedance zone of the time maximum deposit dose rates, ez(max($\boldsymbol{y}_g$)), for two thresholds, 0.1 µGy h$^{-1}$ and $10^{-4}$ µGy h$^{-1}$.

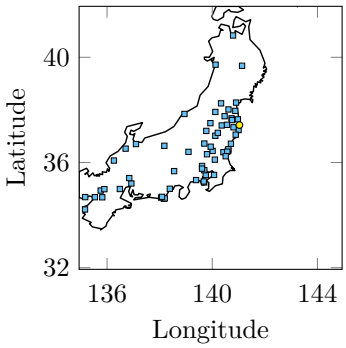


Figure 9: Spatial distribution of the observation stations where the dose rates measurements have been collected. Stations are figured by blue squares and the power plant by a yellow dot.

Sensitivity analysis focuses on the uncertainty of input variables or model parameters, as detailed in this paper. While these uncertainties play a key role in model-observations discrepancies, other sources of error were left aside: inadequacy of the model's physical equations, representativeness errors, measurement errors... The importance of inputs variability relatively to all these other sources of discrepancy can be appreciated by applying the sensitivity analysis methodology to outputs derived from the model assessment framework. If the considered inputs have no influence over model performance indicators, this may mean one of the following: the inputs ranges of variation are too small, the way they are perturbed is inadequate, some fixed input should be made random, or the model's predictive power is so low that uncertainty in the inputs is only marginally affecting its poor performance.

We used a total of 64 time series of ambient instantaneous dose rate measured with a period of 60 minutes or less, covering the release period with some gaps. Figure 9 shows that the spatial coverage over Japan is uneven but the spread of the measurement network is appreciable. These measurements have been analysed by Saunier et al. (2013), along with a model-to-data comparison to observations for ldX, IRSN's operational version of Polyphemus/Polair3D. Part of this data can now be accessed on the web database set up by the International Atomic Energy Agency (IAEA) (2012).

Two performance indicators were selected and computed for every simulation:

- the *factor* 2 *score* is the proportion of simulated values within a factor of 2 from the corresponding observations;
- the *figure of merit in time* is the average of the ratio of the minimum and maximum between the observations and simulations at each time step;

The factor 2 and figure of merit in time range respectively from 0.32 to 0.68 and from 0.34 to 0.62 and have a median of 0.53 and 0.51. These numbers indicate that the model is able to reproduce the observations reasonably well over the input ranges of variation. The performance of the model in a deterministic case

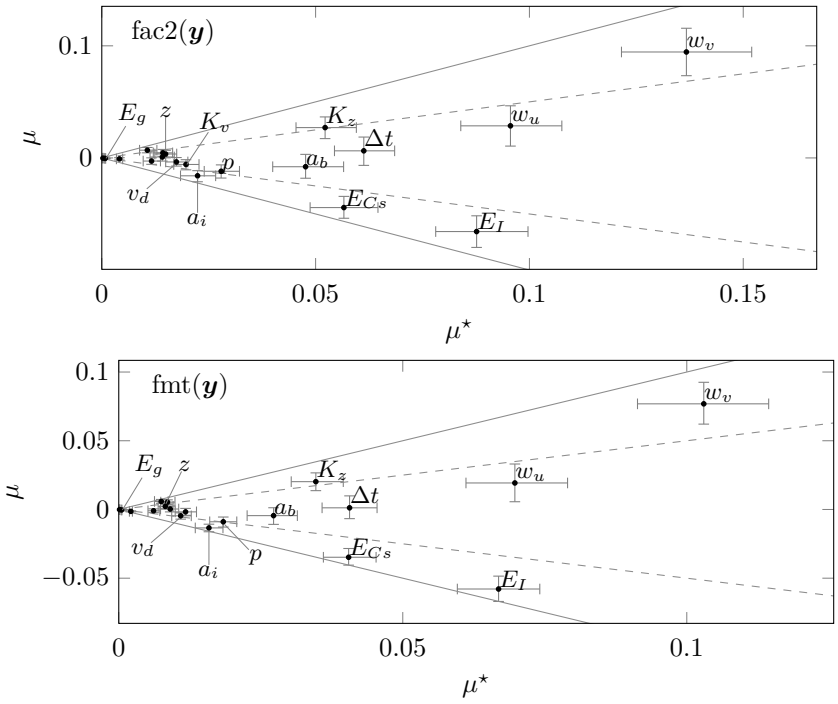


Figure 10: $\mu^\star$ and $\mu$ measures for the factor 2 score (fac2($\boldsymbol{y}$)) and figure of merit in time (fmt($\boldsymbol{y}$)) for ambient dose rate (from top to bottom). Error bars represent 0.95 bootstrap confidence intervals. The solid grey lines correspond to $|\mu| = \mu^\star$. The dashed grey lines correspond to $|\mu| = 0.5\,\mu^\star$.

was illustrated by Saunier et al. (2013). The extended ranges of variation of the statistical scores mean that input uncertainty has a strong impact on the output. They also confirm that dose rate measurements could be adequate for the calibration of the inputs uncertainty ranges.

The $\mu^\star$ and $\mu$ measures for these 2 scores are displayed in figure 10. In all 3 cases, the $\sigma$ measures, not displayed here, vary linearly with the associated $\mu^\star$ and are only slightly lower which denotes strong non-linearity or interactions. A few variables have $\mu$ much smaller than their $\mu^\star$. This is particularly true for the time shift whose influence is highly dependent on the location and time considered.

The sensitivity of the factor 2 score and figure of merit in time are almost the same. They are all dominated by the winds, the emission factors for the iodine and caesium families, the time shift. The emission factor for the noble gases family has almost no effect on these outputs. This can be related to the fact that these indicators are averages over the set of observations. Indeed, while a passage of the plume over a measurement station induces very high dose rates, these events rarely last longer than a few hours whereas the impact of deposited radionuclides lasts much longer. Thus, averaged indicators are mostly

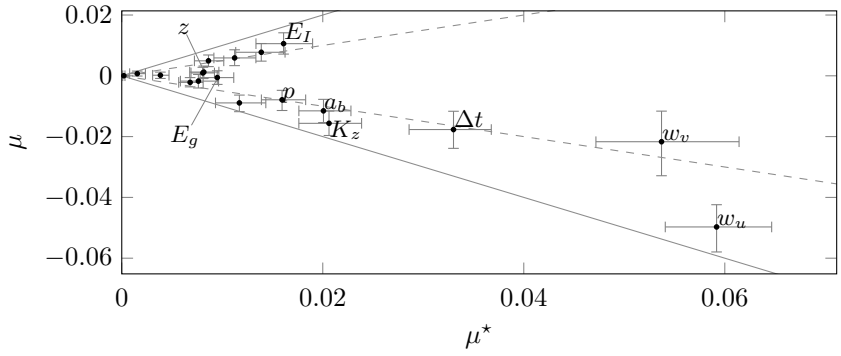


Figure 11: $\mu^\star$ and $\mu$ measures for the score based on peak time match between observations and simulations, ptm($\boldsymbol{y}$). Error bars represent 0.95 bootstrap confidence intervals. The solid grey lines correspond to $|\mu| = \mu^\star$. The dashed grey lines correspond to $|\mu| = 0.5\,\mu^\star$.

representative of the match of deposit dose rate and much less of atmospheric dose rate.

Another indicator based on the temporal match of simulated and observed peaks was designed to complement the previous scores. Peaks in both the simulated and observed signals were detected by recording change in the sign of their first derivative from positive to negative. Then, scores were computed from the observed and simulated peak timings with a procedure detailed in AppendixC. Each pair of matched peaks contribute a score which decreases exponentially with the time interval separating them: two peaks contribute a score of 1 when they are synchronised, a score of 0.5 when distant from 6 h and a score of 0 when distant from more than 18 h. The score of a given simulation was obtained by averaging the individual scores of matched peaks over all measurement stations.

The $\mu^\star$ and $\mu$ sensitivity measures derived from this score are displayed in figure 11. The low $\mu$ values indicate compensation effects which were expected, given that the scores associated to emission events temporally distant and measurement stations spatially distant were aggregated together. The most interesting feature here is that, contrary to classical statistical scores, this output is significantly affected by the emission factor for noble gases. This suggests that a calibration method for uncertainty quantification might need to rely on several indicators at once. Using other kinds of observations, such as deposition measurements or volume activities, may also provide invaluable information.

### *5.5. A remark on sampling*

The results detailed in section 5 were obtained with a sample unusually large for the Morris method: 8 levels and 100 trajectories, picked in a pool of 10 000 with the procedure described in AppendixA. The decision to augment the sample size was motivated mostly by the wide bootstrap confidence intervals associated to the sensitivity estimates obtained with the more common configuration of 4 levels and 10 trajectories, picked in a pool of 100.

The measures of sensitivity computed with the large sample are a little more spread out than those obtained with the small sample, and the associated confidence intervals are much smaller. While this gain in precision allows for a more accurate ranking of the input variables, there is no major different between the two rankings. The differences are more pronounced when observations are involved. There are small changes in the rankings for both the classical scores and the peak time match score. Yet, the swaps in positions are within what one would expect when looking at the bootstrap confidence intervals which are very wide for the small sample.

## 6. Synthesis

A set of 19 input variables were selected to represent potential sources of uncertainty to propagate in Polyphemus/Polair3D, which carried out the atmospheric dispersion of radionuclides after the Fukushima disaster. These variables were attributed probability distributions loosely representing their inherent uncertainty. More precise uncertainty descriptions are not achievable at this early stage of the uncertainty analysis and simple ones are sufficient for a preliminary screening sensitivity analysis. The ranges of variation of the inputs were fixed after a thorough review of the literature dealing with each aspect of the model.

The sensitivities were computed for different aggregated atmospheric and ground dose rates: spatio-temporal aggregated values, time series and spatial fields. Table 2 provides a classification of the inputs based on their $\mu^\star$. Each column corresponds to an output considered in a preceding section and, for each output, the maximal $\mu^\star$ over all inputs is taken as a reference. The "$--$", "$-$", "$=$", "$+$", and "$++$" denote inputs whose $\mu^\star$ values are respectively lower than 1 %, 5 %, 10 %, 50 % or above 50 % of the reference.

The most influential inputs are the winds, the emission factors for the caesium and iodine families, the time shift and, to a lesser extent, the source altitude. The cloud thickness and emission factor for other species weakly impact all the considered inputs so they could be safely discarded from further studies.

Noble gases are not deposited so their emission coefficient has no influence on deposition-related outputs but it is important for atmosphere-related outputs. The reverse can be noted for precipitation intensity which does not affect the outputs derived from atmospheric dose rates.

About the relative weakness of horizontal diffusion coefficients, it must be noted that the spatial resolution used here is 0.125°. They could have a more noticeable influence, should a finer spatial resolution be used.

Future work should focus on refining the uncertainty descriptions of the most influential variables, namely the wind, the precipitation fields and the source term. The emission factor for noble gases has no effect on traditional statistical scores because it only affects atmospheric dose rates which are short-lived. This highlights the necessity to use more temporally localised performance estimates, such as the score for peak time matching proposed here. More generally, refining

| In section | $\langle \bar{\boldsymbol{y}}_a \rangle$ 5.1 | $\bar{\boldsymbol{y}}_a$ 5.2 | ez($\bar{\boldsymbol{y}}_a$) 5.3 | max$\langle \boldsymbol{y}_g \rangle$ 5.1 | $\bar{\boldsymbol{y}}_g$ 5.2 | ez(max $\boldsymbol{y}_g$) 5.3 | fac2($\boldsymbol{y}$) 5.4 | fmt($\boldsymbol{y}$) 5.4 | ptm($\boldsymbol{y}$) 5.4 |
|---|---|---|---|---|---|---|---|---|---|
| $a_b$ | $=$ | $=$ | $+$ | $+$ | $+$ | $+$ | $+$ | $+$ | $+$ |
| $b_b$ | $--$ | $-$ | $-$ | $-$ | $-$ | $=$ | $+$ | $=$ | $+$ |
| $a_i$ | $--$ | $--$ | $-$ | $+$ | $+$ | $+$ | $+$ | $+$ | $+$ |
| $b_i$ | $--$ | $--$ | $-$ | $-$ | $-$ | $-$ | $=$ | $=$ | $=$ |
| $p$ | $-$ | $-$ | $=$ | $+$ | $+$ | $+$ | $+$ | $+$ | $+$ |
| $C_h$ | $--$ | $--$ | $-$ | $-$ | $-$ | $=$ | $=$ | $=$ | $+$ |
| $C_t$ | $--$ | $--$ | $--$ | $-$ | $-$ | $-$ | $--$ | $--$ | $--$ |
| $v_d$ | $=$ | $-$ | $=$ | $=$ | $=$ | $+$ | $+$ | $=$ | $+$ |
| $K_z$ | $+$ | $+$ | $+$ | $=$ | $=$ | $+$ | $+$ | $+$ | $+$ |
| $K_u$ | $--$ | $--$ | $-$ | $--$ | $--$ | $=$ | $=$ | $=$ | $+$ |
| $K_v$ | $-$ | $-$ | $+$ | $--$ | $-$ | $+$ | $+$ | $+$ | $+$ |
| $w_u$ | $=$ | $+$ | $++$ | $+$ | $+$ | $++$ | $++$ | $++$ | $++$ |
| $w_v$ | $-$ | $=$ | $++$ | $=$ | $+$ | $++$ | $++$ | $++$ | $++$ |
| $E_{Cs}$ | $=$ | $-$ | $=$ | $+$ | $+$ | $+$ | $+$ | $+$ | $+$ |
| $E_I$ | $++$ | $+$ | $+$ | $++$ | $++$ | $+$ | $++$ | $++$ | $+$ |
| $E_g$ | $++$ | $+$ | $+$ | $--$ | $--$ | $--$ | $--$ | $--$ | $+$ |
| $E_o$ | $-$ | $--$ | $-$ | $-$ | $-$ | $-$ | $-$ | $-$ | $-$ |
| $\Delta t$ | $=$ | $++$ | $++$ | $+$ | $++$ | $++$ | $+$ | $+$ | $++$ |
| $z$ | $++$ | $++$ | $+$ | $=$ | $+$ | $=$ | $+$ | $=$ | $+$ |

Table 2: Synoptic view of the sensitivities of all considered outputs. For each output, the inputs are classified by comparing their $\mu^\star$ to the maximal $\mu^\star$. The "$--$", "$-$", "$=$", "$++$", and "$++$" denote inputs whose $\mu^\star$ values are respectively lower than 1 %, 5 %, 10 %, 50 % or above 50 % of the maximal $\mu^\star$ for each output. The symbols ez, fac2, fmt, and ptm denote respectively the exceedance zone, factor 2 score, figure of merit in time and peak time match score. Multidimensional $\mu^\star$ were converted to scalars by taking a maximum: over time for $\bar{\boldsymbol{y}}_a$ and $\bar{\boldsymbol{y}}_g$, and over the set of thresholds for ez($\bar{\boldsymbol{y}}_a$) and ez(max($\boldsymbol{y}_g$)).

the input perturbations, for instance by taking into account the spatial structure of the meteorological fields or dividing the simulation time period into intervals, will necessitate calibration, and more precise and localised performance estimates.

Baklanov, A., Sørensen, J. H., 2001. Parameterisation of radionuclide deposition in atmospheric long-range transport modelling. Physics and Chemistry of the Earth 26 (10), 787–799.

Brandt, J., Bastrup-Birk, A., Christensen, J. H., Mikkelsen, T., Thykier-Nielsen, S., Zlatev, Z., 1998. Testing the importance of accurate meteorological input fields and parameterizations in atmospheric transport modelling using dream - validation against ETEX-1. Atmospheric Environment 32 (24), 4167–4186.

Brandt, J., Christensen, J., Frohn, L., 2002. Modelling transport and deposition of caesium and iodine from the Chernobyl accident using the DREAM model. Atmospheric Chemistry and Physics 2 (5), 397–417.

Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environmental Modelling and Software 22 (10), 1509–1518.

Hanna, S. R., Lu, Z., Frey, H. C., Wheeler, N., Vukovich, J., Arunachalam, S., Fernau, M., Hansen, D. A., 2001. Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the july 1995 OTAG domain. Atmospheric Environment 35 (5), 891–903.

International Atomic Energy Agency (IAEA), 2012. Fukushima monitoring database. `https://iec.iaea.org/fmd/search_by_dataset.aspx`, accessed on November 27, 2013.

Korsakissok, I., Mathieu, A., Didier, D., 2013. Atmospheric dispersion and ground deposition induced by the Fukushima nuclear power plant accident: a local-scale simulation and sensitivity study. Atmospheric Environment 70 (0), 267–279.

Louis, J.-F., 1979. A parametric model of vertical eddy fluxes in the atmosphere. Boundary-Layer Meteorology 17, 187–202.

Mallet, V., Quélo, D., Sportisse, B., Ahmed de Biasi, M., Debry, É., Korsakissok, I., Wu, L., Roustan, Y., Sartelet, K., Tombette, M., Foudhil, H., 2007. Technical Note: The air quality modeling system Polyphemus. Atmospheric Chemistry and Physics 7 (20), 5,479–5,487.

Maryon, R., Smith, F., Conway, B., Goddard, D., 1991. The U.K. nuclear accident model. Progress in Nuclear Energy 26 (2), 85–104.

Mathieu, A., Korsakissok, I., Quélo, D., Groëll, J., Tombette, M., Didier, D., Quentric, E., Saunier, O., Benoit, J.-P., Isnard, O., 2012. Atmospheric dispersion and deposition of radionuclides from the Fukushima Daiichi nuclear power plant accident. Elements 8 (3), 195–200.

Morris, M. D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics, 161–174.

Pujol, G., Iooss, B., Janon, A., 2013. sensitivity: Sensitivity Analysis. R package version 1.7.
URL http://CRAN.R-project.org/package=sensitivity

Ridders, C. J. F., 1979. A new algorithm for computing a single root of a real continuous function. Circuits and Systems, IEEE Transactions on circuits systems 26 (11), 979–980.

Ryall, D., Maryon, R., 1998. Validation of the UK met. office's name model against the ETEX dataset. Atmospheric Environment 32 (24), 4265–4276.

Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. Environmental Modelling & Software 25 (12), 1508–1517.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global sensitivity analysis: the primer. Wiley Online Library.

Saunier, O., Mathieu, A., Didier, D., Tombette, M., Quélo, D., Winiarek, V., Bocquet, M., 2013. An inverse modeling method to assess the source term of the Fukushima nuclear power plant accident using gamma dose rate observations. Atmospheric Chemistry and Physics Discussions 13 (6), 15567–15614.

Savitzky, A., Golay, M. J., 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical chemistry 36 (8), 1627–1639.

Sørensen, J. H., 1998. Sensitivity of the DERMA long-range gaussian dispersion model to meteorological input and diffusion parameters. Atmospheric Environment 32 (24), 4195–4206.

Sportisse, B., 2007. A review of parameterizations for modelling dry deposition and scavenging of radionuclides. Atmospheric Environment 41 (13), 2683–2698.

Stohl, A., Seibert, P., Wotawa, G., Arnold, D., Burkhart, J. F., Eckhardt, S., Tapia, C., Vargas, A., Yasunari, T. J., 2012. Xenon-133 and caesium-137 releases into the atmosphere from the Fukushima Dai-ichi nuclear power plant: determination of the source term, atmospheric dispersion, and deposition. Atmospheric Chemistry and Physics 12 (5), 2313–2343.
URL http://www.atmos-chem-phys.net/12/2313/2012/

Terada, H., Katata, G., Chino, M., Nagai, H., 2012. Atmospheric discharge and dispersion of radionuclides during the Fukushima Dai-ichi nuclear power plant accident. part II: verification of the source term and analysis of regional-scale atmospheric dispersion. Journal of Environmental Radioactivity 112 (0), 141–154.

Thykier-Nielsen, S., Deme, S., Mikkelsen, T., 1999. Description of the atmospheric dispersion module RIMPUFF. Tech. rep., Forschungszentrum Karlsruhe GMBH.

Troen, I., Mahrt, L., 1986. A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. Boundary-Layer Meteorology 37, 129–148.

Verwer, J. G., Hundsdorfer, W., Blom, J. G., 2002. Numerical time integration for air pollution models. Surveys on Mathematics for Industry 10, 107–174.

Winiarek, V., Bocquet, M., Duhanyan, N., Roustan, Y., Saunier, O., Mathieu, A., jan 2014. Estimation of the caesium-137 source term from the Fukushima Daiichi nuclear power plant using a consistent joint assimilation of air concentration and deposition observations. Atmospheric Environment 82 (0), 268–279.

Yamartino, R., 2000. Refinement of horizontal diffusion in photochemical grid models. In: $11^{\text{th}}$ conference on the application of air pollution meteorology. AMS/AWMA, Long Beach, CA.

## AppendixA. Sampling method

This section details the algorithm introduced in 3.2. In particular, it presents how to reuse model evaluations in order to compute more than one elementary effect. Morris (1991) proposed the following algorithm for generating samples of independent elementary effects sparingly. First, the model response is evaluated with a randomly selected vector of inputs. Then, each input is perturbed once in a random order and the model response is evaluated after each perturbation. This sequence draws a trajectory in the input space and every point of this trajectory is used to compute two elementary effects, except the first and last which are used only once. A number $r$ of such trajectories is produced, which constitutes $m$ samples of independent elementary effects corresponding to the $m$ input variables. The computational cost of the Morris method is then $(m+1) \times r$ model evaluations for a sample size of $r$. This *trajectory*-based design has the advantage of a slightly better coverage of the input space than the *radial* design in which perturbations are applied in each input direction of $r$ reference points.

So as to avoid generating border effects, the number of perturbation levels, $k$, is chosen even and the perturbation, $\delta$, fixed to twice the interval between two levels. Jumping only one step at a time would cause the extreme levels to be sampled less and choosing an odd integer for $k$ would cause uneven sampling, no matter the value $\delta$.

Finally, Campolongo et al. (2007) proposed an improvement of the sampling scheme described above. A larger sample of trajectories, say of 100 times the sample size, is generated and a sub-sample of size $n$ is then selected so as to maximise the distance between each trajectories. Here a maximin algorithm adapted from Pujol et al. (2013) was used to carry out this optimisation task.

First, pairwise Hausdorff distances are computed in the large sample using the following definition:

$$h\left(\mathbf{X_i}, \mathbf{X_j}\right) = \max\left[\max_{\boldsymbol{x}_i \in \mathbf{X_i}} \min_{\boldsymbol{x}_j \in \mathbf{X_j}} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|, \max_{\boldsymbol{x}_j \in \mathbf{X_j}} \min_{\boldsymbol{x}_i \in \mathbf{X_i}} \|\boldsymbol{x}_i - \boldsymbol{x}_j\|\right], \tag{A.1}$$

where $\mathbf{X_i}$ and $\mathbf{X_j}$ denote trajectories in the input space, that is sets of $m+1$ vectors of size $m$. Then the two most distant trajectories are selected as the first two trajectories of the optimised sample. Finally, trajectories from the large sample are added sequentially to the optimised sample using the Kenard-Stone algorithm, until the desired size is reached: the selected trajectory is the one that maximises the minimum Hausdorff distance between itself and the trajectories of the optimised sample.

## AppendixB. Derivation of the perturbation bounds of wind components

As stated in section 4.5, the two components of the wind field were perturbed independently and uniformly. As there is no reason to single out any direction, the intervals in which lie the perturbations were set symmetric and of equal length. Hence the wind perturbation range is controlled by a unique parameter, namely the length of the interval in which lies a wind component, denoted $2\beta$.

While it is not possible to establish an exact correspondence between this setting and the recommendations from Hanna et al. (2001), they can be used to estimate $\beta$. Hanna et al. (2001) suggest uncertainty ranges for the wind speed modelled by a log-normal distribution and for its direction modelled by a normal distribution. These confidence intervals are equivalent to the two following statements:

$$\mathrm{P}\left(\nu_m \in \left[\frac{\nu}{1.5}, 1.5\nu\right]\right) = 0.95, \tag{B.1}$$

and

$$\mathrm{P}\left(\theta_m \in [\theta - 40°, \theta + 40°]\right) = 0.95, \tag{B.2}$$

where $\nu$ and $\theta$ refer respectively to the speed and the direction of the wind and the $m$ subscript denote the median of these random variables.

As the uniform perturbation scheme of the wind components has only one parameter, the conjunction of these two statements should be considered:

$$\mathrm{P}\left(\nu_m \in \left[\frac{\nu}{1.5}, 1.5\nu\right],\ \theta_m \in [\theta - 40°, \theta + 40°]\right) = 0.95^2 = 0.9025. \tag{B.3}$$

This necessary condition stemming from the recommendations of Hanna et al. (2001) is sufficient to determine a unique value for the undetermined parameter. A geometric interpretation of the equation to solve is given in figure B.12. The domain sampled when perturbing each component uniformly is a square of side

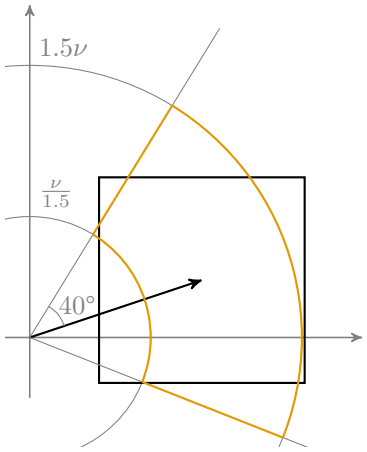


Figure B.12: Geometric interpretation of the comparison with the uncertainty range suggested by Hanna et al. (2001) for the wind field and the perturbation of its zonal and meridional components. The black arrow represents a wind vector of given speed and direction. The two circles delimit the domain $[\frac{\nu}{1.5}, 1.5\nu]$ that contains the median of the wind speed $\nu_m$ with a probability of 0.95. The two grey rays on either side of the wind vector delimit the domain $[\theta - 40°, \theta + 40°]$ that contains the median of the wind direction $\theta_m$ with a probability 0.95. The intersection of these two domains is highlighted in orange. The black square is the region uniformly sampled by the perturbation of the zonal and meridional components of the wind. The problem at hand amounts to finding $\beta$ so that the area of the orange portion of annulus is equal to 0.9025 of the area of the black square.

length $2\beta$ centred in $(\cos(\theta_m)\,\nu_m, \sin(\theta_m)\,\nu_m)$. The domain in equation (B.3) is a portion of annulus delimited by two circles and two rays. The condition stated in equation (B.3) is met for $\beta$ such that the area of the portion of annulus over the area of the square is equal to 0.9025. These areas can be computed analytically but the solution in $\beta$ is intricate and needs to be estimated numerically. Giving up on the exact solution, the intersection's area for a given $\beta$ were estimated with a Monte Carlo procedure and the root finding method devised by Ridders (1979) was used to solve in $\beta$.

One average value of the $\beta$ parameter per vertical layer was determined by averaging in each layer the $\beta$ values found for all the wind vectors in the meteorological data used for the simulations.

## AppendixC. Score based on peak time matching

The algorithm used in section 5.4 to compute scores representing the temporal match of the observed and simulated dose rate peaks comprehends three steps. First, peaks are identified using a first order derivative estimate. Then, a simple filter is applied in order to keep only the peaks that emerge from the signal noise. Finally, sets of simulated and observed peak times are compared at each station, and the resulting scores are averaged over all stations.

The peak detection step consists in locating the times when the first derivative of the signal changes sign from positive to negative. The first derivative is estimated using the smoothing filter proposed by Savitzky and Golay (1964) which relies on local least-square fit of polynomials. A window size of 5 points equivalent to 5 h and polynomials of order 2 were used for both the simulations and observations.

Simulated peaks with an amplitude smaller either than 10 times the measurement resolution, namely the lowest possible interval between two consecutive points in the signal, or 0.05 times the overall amplitude of the signal were set aside. For observations, these cutoffs were set to 10 and 0.1.

Given a pair of signals $A$ and $B$, the score of $A$ versus $B$ is the *proportion of peaks from $A$ matching a peak from $B$*. The quality of a match depends on the time interval separating the peaks and a temporal resolution parameter, $t_r$, which was set to 6 h by visual inspection of the signals. The procedure to compute the said proportion is as follows:

- First, an *individual score* is computed for each peak in $A$ by recording the time interval $\delta t$ separating it from the closest peak in $B$. It is equal to $\exp(-\frac{\delta t^2}{\ln(2)\, t_r})$ if $\delta t < 3t_r$, and 0 otherwise.
- Then, peaks from $A$ separated from each other by less than $3t_r$ and matched with the same peak in $B$ are grouped together. This aims to prevent short-timed series of matching peaks to override the influence of a missed peak elsewhere in the signal. *Group scores* are obtained by averaging the individual scores in each group.
- Finally, the average of the group scores yields the score of $A$ versus $B$.

The score of $A$ versus $B$ is usually different from the score of $B$ versus $A$ so they are averaged to yield the score of the pair $(A, B)$. The scores used in section 5.4 are the averages of the scores of all 82 simulation-observation pairs corresponding to the set of observation stations.