Geoscientific
Model Development

Development and technical paper

Accurate and fast prediction of radioactive pollution by kriging coupled with auto-associative models

Raphaël Périllat[1]**, Sylvain Girard**[1]**, and Irène Korsakissok**[2]

[1]Phimeca, Paris, France
[2]Autorité de sûreté nucléaire et de radioprotection, Paris, France

Correspondence: Raphaël Périllat (perillat@phimeca.com)

Abstract. Uncertainty estimation is a key issue in nuclear crisis situations. Probabilistic methods for taking uncertainties into account in assessments are often costly in terms of the number of simulations and computation time. This is why emulation methods, which enable rapid estimation of numerical model outputs, represent a promising solution. However, in the context of radioactive dispersion modeling, existing emulators are mostly limited to scalar outputs. In a crisis context, decisions are often based on dose maps, which are mathematically represented by high-dimensional data. In this study, we use the auto-associative model method to reduce the dimension of dose results and then predict these reduced representations using kriging. We also compare this prediction method with others used by the French Nuclear Safety and Radiation Protection Authority (ASNR) to predict the consequences of a nuclear accident.

1 Introduction

1.1 Context

In the event of a nuclear accident, numerical simulations of atmospheric dispersion are used to predict the territories potentially impacted by radioactive releases. The French Nuclear Safety and Radiation Protection Authority (ASNR) develops and uses atmospheric dispersion models embedded within its operational crisis platform called C3X to perform these calculations (Tombette et al., 2014). These simulations are used to infer operational indicators, such as the maximum distance from the source, where a dose threshold can be exceeded. The thresholds may be, for instance, regula-

tory protective action guide levels that could trigger protective actions such as population evacuation, sheltering, stable iodine prophylaxis, or food restrictions. Although the dose guide levels differ from one country to another, this approach based on zones of threshold exceedance is widely used, as described in IAEA's publications on emergency preparedness and response (International Atomic Energy Agency, 2021).

Such evaluations are subject to uncertainties due to a lack of information on the installation's status, meteorological forecast uncertainties, and models' approximations (Leadbetter et al., 2020; Le et al., 2021). Prediction errors can induce two kinds of wrong decisions: either insufficient population protection zones, where a threshold exceedance occurs but was not predicted, or unnecessary actions zones, where a threshold exceedance is forecast but does not come true. While the detriment to the population in the former case is obvious, leading to the use of conservative evaluations designed to avoid this situation at all costs, limiting evacuation and other restrictions where possible is also desirable as these actions may have a high and potentially long-term economical and health cost (Nomura et al., 2013, 2016). A better quantification of uncertainties may help refine the hypotheses and potentially reduce the margins of the conservative assumptions while ensuring sufficient population protection. In the approach applied by the ASNR's emergency center, the very first response generally relies on pre-calculated scenarios, the data of which are gathered in an "accident type sheet" (ATS). This database relies on calculations carried out in the preparedness phase for a number of accidental scenarios and for selected weather situations described by a few parameters (wind direction and speed, atmospheric stability, and rain) assumed to be constant in time and homogeneous over the sim-

ulation domain. In a second step, ASNR uses its local-scale Gaussian puff atmospheric dispersion model, called pX (Korsakissok et al., 2013; El-Ouartassy et al., 2022), to obtain predictions that correspond more closely to the actual accidental and meteorological situation.

Forecasting tools must be compatible with emergency response time constraints, when the first evaluations should be provided typically within 1 h after the alert. This timing includes not only the computational time required to set up and run the simulations themselves, but also the time required to gather meteorological forecasts and source term assessments, analyze the results, and communicate them to decision-makers. Thus, a numerical model such as pX requiring typically a few minutes to run can be used for a single, deterministic estimation. However, the computation time required to account for uncertainties using hundreds of simulations does not fit with these operational constraints.

1.2 Emulation and dimension reduction

An emulator is a surrogate model that approximates the output of a computationally intensive physical model while being much faster to evaluate. Emulators are often constructed by interpolating scalar outputs from a set of pre-computed simulations.

In radiological emergency contexts, emulators enable several practical applications:

– They enlarge the pre-calculated scenario database by including a variety of input parameters, allowing us to evaluate results for a larger range of meteorological situations than those considered in the ATS.

– They replace the original model in the case of uncertainty estimation (Le et al., 2018) or sensitivity analyses (Girard et al., 2016), where several hundred simulations are needed to perturb the model inputs in order to obtain a large number of outputs.

– Thanks to their speed, they allow interactive exploration of the input space with a graphical interface where it is possible to vary the model inputs to observe their influence on the output. Such a tool can be used for education and training purposes in order to demonstrate the influence of input parameters on the outputs and to help making "reasonably conservative" evaluations of uncertain parameter values.

However, the output of dispersion models is typically a spatial map – high-dimensional and structured – whereas most emulators are designed for scalar outputs. One naive solution would be to build an emulator for each grid point, but this approach ignores spatial correlations and is computationally inefficient.

In the field of atmospheric dispersion modeling, emulators have been used to predict spatio-temporal average quantities, values at a monitoring point (Le et al., 2018; Girard et al., 2016), or maximum exceedance distances (Périllat et al., 2020). Fitting an emulator for each grid point of a two-dimensional map would be difficult to both compute and implement as it would ignore the spatial correlations inherent in the data. For these reasons, emulating a spatial map requires a first step of dimension reduction.

The most widely used dimension reduction method is principal component analysis (PCA) (Jolliffe and Cadima, 2016; Jolliffe, 2002). It consists of projecting a set of points onto a vector subspace in a least-squares optimal way in order to obtain the most faithful representation of this set of points in a reduced dimensional subspace. It has been applied in the specific domain of atmospheric dispersion, sometimes in combination with emulation (Burgin et al., 2017; Le et al., 2018; Mallet et al., 2018; Swallow et al., 2017; Lumet et al., 2025). However, being a linear approximation, PCA fails to encode sets of maps that are too different from one another.

The auto-associative model (AAM) is an extension of PCA that captures nonlinear relationships in the dataset (Girard, 2000). Instead of projecting the data onto a linear subspace, the AAM approximates the dataset by a low-dimensional differentiable manifold. It does so through two complementary mappings: a projection function that reduces the dimensionality and a recovery function that reconstructs the original data from the reduced coordinates. These mappings are typically estimated using spline regression, allowing the model to learn smooth nonlinear variations in the dataset.

The AAM can be seen as a constrained form of an autoencoder, where both the architecture and the optimization strategy are adapted to small sample sizes and structured data, such as spatial fields. Compared to PCA, which assumes linear combinations of orthogonal basis vectors, the AAM provides better reconstruction when data exhibit spatial transformations such as shifts, deformations, or plume rotations – situations that frequently occur in dose maps produced by atmospheric dispersion models.

The AAM has already been applied once in this context to the dimension reduction of radiological dispersion maps (Girard et al., 2020). In that study, it was shown that the AAM could recover over 78 % of the variance using only two nonlinear coordinates, while PCA required six components to achieve the same reconstruction quality. The difference was particularly notable for simulations with varying wind directions, where PCA struggled to encode rotated structures. Visual comparisons confirmed that the AAM preserved spatial features more faithfully.

The present paper presents the first combination of the AAM with emulation applied to the prediction of dose maps in the event of an accidental release of radioactive materials into the atmosphere. We introduce the case study in Sect. 2, the methodology in Sect. 3, and the results in Sect. 4. We develop and validate the AAM independently in Sect. 4.1.1, the kriging model in Sect. 4.1.2, and finally the complete emulation framework in Sect. 4.1.3. Validation is performed on an operational scenario for predicting threshold exceedance

Table 1. Input parameters and ranges of variation for the construction of emulators.

Input variable	Range of variation	Unit
Wind module	$[0, 10]$	m s^{-1}
Rain intensity	$[0, 10]$	mm h^{-1}
Release height	$[0, 100]$	m
Source term amplitude	$[10, 100]$	%
Deposition velocity of iodine	$[1 \times 10^{-5}, 1 \times 10^{-2}]$	m s^{-1}
Deposition velocity of other elements	$[5 \times 10^{-4}, 5 \times 10^{-3}]$	m s^{-1}

areas. The performance of the proposed approach is compared to alternative prediction methods in Sect. 4.2. Finally, in Sect. 5, we illustrate the practical applications enabled by the emulator, particularly in terms of real-time response, probabilistic risk assessment, and decision support.

2 Case study

We simulated the result of a primary breach leading to a total core meltdown in 1 h of a 1300 MWe pressurized water reactor. This accidental scenario is one of the pre-calculated scenarios leading to the exceeding of protective action guide levels over significant distances. We used the pX Gaussian puff dispersion model with the Doury diffusion model in neutral atmospheric stability and with a meandering wind coefficient of 3. The meandering wind coefficient is a multiplicative factor applied to the diffusion model and designed to account for wind direction variations that occur during the time span of the release and are not taken into account by the meteorological inputs.

We focused here on two-dimensional maps of thyroid inhalation equivalent dose 24 h after the beginning of the releases. In France, stable iodine prophylaxis is related to a dose criteria of 50 mSv to the thyroid. We used a polar mesh with smaller cells close to the source, where there are strong spatial dose variations. The nodes of the mesh are distributed on 36 angles between 0 and 360° and on 61 different radii from 500 m, increasingly spaced from each other as we move away from the source until a distance of 30 km.

Thus, the output data have a dimension of 2196. The case study is stationary: inputs are assumed to be constant in time and space. This is not generally the case for meteorological variables but is consistent with the simple situations on which pre-calculated sheets are based. We considered six sources of uncertainty as inputs of the model, which are listed in Table 1. Two of them are related to the meteorological situation: the wind module and rainfall rate; two uncertain parameters, the source amplitude (a multiplicative factor applied to the source term computed for the chosen accidental scenario) and release height, describe the source term characteristics; and the two last parameters are used to define radionuclide deposition rates for iodine and others.

The choice of uncertain parameters and their ranges of variation is not representative of the full range of possible situations. They were chosen to be representative of the usual range of values encountered during emergency exercises to build and validate a proof of concept that can then be extended to other situations and scenarios. For instance, the source term computed for the accidental scenario comprises 224 radionuclides, each of them being associated with a release rate as a function of time. It was computed using conservative assumptions regarding the quantity of radioactive materials emitted into the atmosphere. Therefore, the multiplicative factor applied to these quantities is assumed to vary between 10 % and 100 % as the evaluation of the installation status at the time of the accident is more likely to lead to a downward revision of the source term.

The ranges of variation for each input parameter were selected based on operational knowledge and literature values, as well as their relevance in emergency scenarios:

- *Wind module*: varied between 0 and 10 m s^{-1}. This range covers typical near-surface wind speeds encountered in most meteorological conditions relevant for nuclear dispersion scenarios. Wind speeds above 10 m s^{-1} are less frequent and usually associated with lower concentrations and are therefore less likely to trigger threshold exceedance.

- *Rain intensity*: varied between 0 and 10 mm h^{-1}, covering the range from no precipitation to heavy rain. This interval includes typical rain intensities relevant for wet deposition modeling in nuclear dispersion scenarios. Rain intensities above 10 mm h^{-1} are less frequent and often short-lived, and are therefore not prioritized in standard emergency simulations.

- *Release height*: varies from 0 to 100 m. A value of 0 m corresponds to a ground-level release, which is relevant for some accident scenarios, such as leaks near the base of a reactor building. The upper bound of 100 m corresponds approximately to the height of the ventilation stacks or highest building points in most nuclear facilities and thus reflects realistic release heights for stack emissions or elevated plumes. Higher plume emissions would not lead to any threshold exceedance at the ground level for the accidental scenario considered.

- *Source term amplitude*: from 10 % to 100 %, reflecting the uncertainty in estimating the actual quantity of radioactive material released. It is common practice during an emergency to revise the source term amplitude downward as more information becomes available about the plant's condition and containment efficiency.

- *Deposition velocities*: for iodine, a range of 1×10^{-5}– $1 \times 10^{-2} \text{ m s}^{-1}$ is used, and for other elements from 5×10^{-4} to $5 \times 10^{-3} \text{ m s}^{-1}$. The range of variation was derived from a literature review (Baklanov and Sørensen,

2001) and is consistent with previous studies (Girard et al., 2014).

3 Emulation method

3.1 Auto-associative models

"Reducing the dimension" of an ensemble embedded in a high-dimensional vector space consists of building an associated ensemble with a lower-dimensional coordinate system. A rough definition of topological dimension would be "the minimum number of variables needed to represent a set" (Fukunaga and Olsen, 1971). More rigorously, we must choose the nature of the associated sets in order to have a precise definition of "coordinate system", for example, the one recalled by Milnor (1997) for differentiable manifold.

Given a set $G \subset \mathbb{R}^m$, with a large m, we try to construct the approximate set $A \subset \mathbb{R}^m$ in bijection with the vector space $C \subset \mathbb{R}^l$, with a small l:

$$G \subset \mathbb{R}^m \xrightarrow{\psi} C \subset \mathbb{R}^l \underset{\chi}{\overset{\chi^{-1}}{\leftrightarrows}} A \subset \mathbb{R}^m. \tag{1}$$

The auto-associative model (Girard and Iovleff, 2008) is a nonlinear method of dimension reduction. The term "linear" here means that the approximating space A would be a sub-vector space and $\chi \circ \psi$ an orthogonal projection. In contrast, in the "nonlinear" case, the approximating space A is a differentiable manifold and $\chi \circ \psi$ a composition of orthogonal projections by a nonlinear χ function.

3.2 Kriging

Kriging is a spatial interpolation method and the core of geostatistics. It was originally designed for optimizing gold mining (Chilès and Delfiner, 1999) by inferring from a few boreholes the spatial distribution of gold grades over the whole mining field. Kriging becomes an emulation method by replacing the spatial coordinates by the model inputs and the gold grade by the scalar model output.

The kriging emulator predicts the output at a new input point x as a linear combination of N known outputs $f(x_i)_{i=1}^N$ at previously simulated input points x_i. The weights w_j of this linear combination are chosen to minimize the variance of the prediction error under the assumption that the output is a realization of a Gaussian process. This Gaussian process is assumed to be second-order stationary, meaning its mean is constant and its covariance between two input points depends only on their relative position.

The weights are the solution to the following system:

$$\forall i \in \{1, \ldots, N\}, \quad \sum_{j=1}^N w_j K(x_i, x_j) + \lambda = K(x_i, x), \tag{2}$$

$$\sum_{j=1}^N w_j = 1, \tag{3}$$

where $K(\cdot, \cdot)$ is a positive definite covariance kernel, λ is a Lagrange multiplier enforcing unbiasedness, and x is the target point. The predicted centered value is then $\hat{f}(x) = \sum_{j=1}^N w_j f(x_j)$. The mean of the process can also be estimated by a generalized least-squares regression.

In this work, we use the `DiceKriging` R package (Roustant et al., 2012), which implements kriging metamodels for deterministic computer codes. We use a product of stationary covariance kernels, one for each input variable, and estimate the correlation lengths via maximum likelihood optimization.

A comprehensive introduction to kriging for emulation can be found in the textbook *Gaussian Processes for Machine Learning* (Rasmussen and Williams, 2006), and an application to atmospheric dispersion modeling is described in Girard et al. (2016).

3.3 Putting it into practice

We performed 2548 simulations uniformly sampling the five-dimensional input space of the release height, the wind module, the rain intensity, and the two deposition velocities (their range of variation is given in Table 1). The output dose being linear with the amplitude of the source term, we completed the simulation sample using two random amplitudes for each input vector, thus covering the whole six-dimensional input space with a total of 5096 points. 4096 of them were used to train the AAM, while the other 1000 were used to fit the model.

Given that the dose values vary by up to a factor of 10^5 depending on distance, we applied the method to their logarithm to better reveal relative variations. The AAM reduced the dimension of the results to 9 coordinates, whereas the initial data dimension was 2196, corresponding to a grid mesh 36×61 in size. Using fewer than nine coordinates resulted in larger errors, while higher-dimensional approximations increased computation times with only moderate improvements in results. Since the final goal of this parameterization is to focus on the zones where a dose threshold (th) is exceeded, we also truncated the logarithm of the doses: any values below log(th)/2 were set to log(th)/2. Thus, small dose variations do not disturb the parameterization by the AAM.

Kriging is then used to create emulators. The emulator combined with dimension reduction can then be used to predict an output map for any new input vector (see Fig. 2). The AAM can finally associate with these nine scalars a two-dimensional map of the inhalation dose. Each of the nine AAM coordinates is predicted independently using a separate kriging model, allowing us to capture the specific response behavior associated with each latent component.

The number of AAM coordinates was chosen based on a trade-off between reconstruction accuracy and model com-
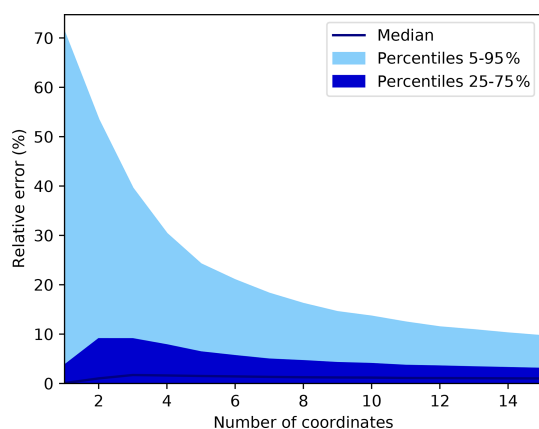
Figure 1. Evolution of the relative reconstruction error (in %) as a function of the number of AAM coordinates. The solid line represents the median error in the test set; the shaded areas correspond to the 25th–75th and 5th–95th percentiles.

plexity. We performed a systematic validation by varying the number of coordinates from 1 to 12 and computing the relative reconstruction error in a separate test set. As shown in Fig. 1, the relative error decreases sharply with the number of coordinates of up to eight to nine, beyond which additional coordinates yield only marginal improvements. The interquartile and 5 %–95 % ranges also stabilize at that point, indicating that the reconstruction quality no longer significantly benefits from added complexity. Based on this analysis, we retained nine coordinates as a good compromise.

4 Results

4.1 Validation

An additional simulation sample was built to test the reliability of the emulator. We drew $M = 1000$ random new points in our six-dimensional input space and then run the computational model M times with input parameters corresponding to these draws. We obtained M output results that were compared with the maps reconstructed after dimension reduction with the AAM (Sect. 4.1.1) and with emulator predictions (Sect. 4.1.2) and then with the combination of the two methods (Sect. 4.1.3).

4.1.1 Validation of the dimension reduction by the AAM

We assessed the validity of the dimension reduction step by comparing the maps two by two for our test sample, as in Fig. 3, and by calculating the figure of merit in space (FMS) of the dose criteria exceedance isolines for each of these maps. This score is calculated by dividing the area of the intersection of the two surfaces A and B by the area of the union of the two:

$$\text{FMS} = \frac{A \cap B}{A \cup B}.$$

The FMS of two very similar surfaces approaches 1, while when two isolines have little surface in common, it approaches 0. This metric has been used in previous studies of radioactive dispersion, such as the European CONFIDENCE project (Bedwell et al., 2020), where it helped quantify the spatial agreement between predicted and reference dose areas.

Figure 4 presents a histogram of the FMS values, showing that the areas enclosed by the isolines exceeding 50 mSv are well preserved by the AAM approximation. In 75.5 % of the cases of the testing samples, the FMS is greater than 0.8, indicating a strong agreement between the AAM prediction and the reference model. In 11.7 % of the cases, the FMS could not be computed because no threshold exceedance was observed, meaning that no isoline was formed. Overall, these results confirm that the predicted isolines are very close to those produced by the original model.

4.1.2 Validation of the interpolation by kriging

Figure 5 shows how well each scalar is predicted. The closer the set of points is to the line $y = x$, the better the prediction.

We quantified the prediction error with the standardized mean squared error (SMSE). For a set of N observed points $(x_i)_{i \in [1,N]}$ and a set $(\hat{x}_i)_{i \in [1,N]}$ of estimated points, the SMSE is defined as follows:

$$\text{SMSE} = \frac{\sum_{i=1}^{N} (x_i - \hat{x}_i)^2}{\sum_{i=1}^{N} (x_i - \bar{x})^2},$$

where $\bar{x}$ is the mean of $(x_i)_{i \in [1,N]}$.

We can observe in Table 2 that the first three coordinates are reconstructed with excellent accuracy (very low SMSE). The SMSE increases progressively with the score index, reaching high values for the last two scores. This degradation is largely due to one outlier point in the test sample, which heavily impact the SMSE because it is based on squared errors.

To provide a more robust evaluation of the prediction quality, we also computed the 95 % quantile of the relative absolute error. This metric gives an upper bound on the error affecting most points without being overly sensitive to rare
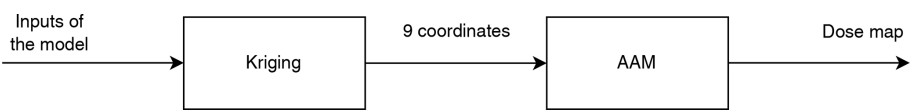
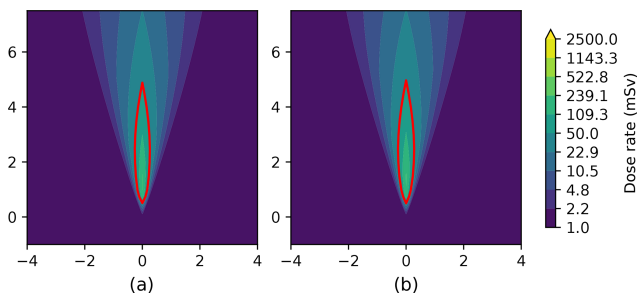**Figure 2.** Dose map prediction process using emulation coupled with the AAM.

Figure 3. Inhalation dose map and 50 mSv guide-level exceedance isolines for a model output in **(a)** and for the approximation of this output by the AAM in **(b)**.
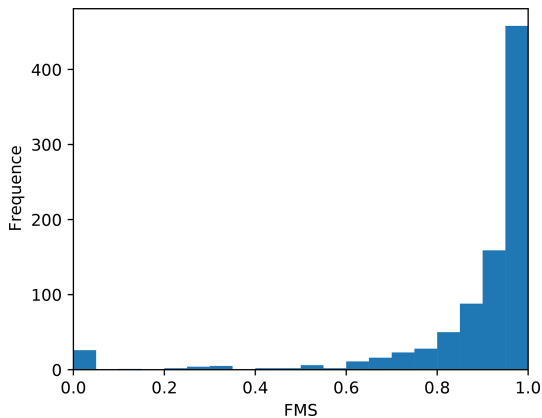

Figure 4. Histogram of the FMS that compares the threshold exceedance zones for the model outputs with AAM approximations across the testing samples. 75.5 % of the FMS is between 0.8 and 1, which implies that the projected results are very often similar to those obtained by simulation.

extreme values. As shown in Table 2, the Q_{95} scores remain acceptable even for the least accurate coordinates, confirming that the predictions remain generally reasonable despite isolated errors. These last coordinates bring fine details to the reconstruction and were therefore retained to preserve the quality of the final maps.

4.1.3   Validation of the emulator that combines the two methods

With the AAM, this prediction of the nine scalars can be transformed into a two-dimensional dose map, which can be used to determine a threshold exceedance isoline. The succession of the two methods thus allows us to convert the

Table 2. SMSE and the 95 % quantile of relative absolute error for each score.

Score	SMSE	Q_{95} rel. abs. error
1	9.939×10^{-04}	1.219×10^{-01}
2	6.482×10^{-03}	9.413×10^{-01}
3	1.466×10^{-02}	$1.811 \times 10^{+00}$
4	7.065×10^{-02}	$4.368 \times 10^{+00}$
5	1.228×10^{-01}	$4.044 \times 10^{+00}$
6	6.076×10^{-02}	$2.078 \times 10^{+00}$
7	1.604×10^{-01}	$3.231 \times 10^{+00}$
8	5.114×10^{-01}	$2.268 \times 10^{+00}$
9	$1.001 \times 10^{+00}$	$3.134 \times 10^{+00}$

model inputs into a decision-aiding map, defined by an isoline, to estimate whether or not a guide-level might be exceeded.

Comparisons between simulations and associated emulator predictions can be classified into four cases:

– *Case 1*: dose maps that are well reconstructed by the method, as shown in Fig. 6a and b. The shape of the isoline is preserved, and its size is similar to that given by the simulation, resulting in a well-estimated maximum distance.

– *Case 2:* dose maps that are well reconstructed, but the isolines can differ slightly in scale from what is expected, as illustrated in Fig. 6c and d.

– *Case 3*: dose maps that are less well reconstructed, where the general shape of the isoline is not accurately reproduced. This typically occurs under low-wind conditions, as shown in Fig. 6e and f.

– *Case 4*: dose maps that approach the threshold value with little or no exceedance, but the predicted and simulated isolines may differ significantly. Figure 6g and h illustrate this: although the two dose maps appear similar, a slight difference in intensity can cause a noticeable discrepancy in the exceedance isoline prediction.

We evaluated the adequacy of the simulated and predicted surfaces by calculating the FMS of our test sample simulations. Figure 7 represents the histograms of the FMS. We can notice a slight degradation compared to Fig. 4. However, 70.9 % of the FMS is between 0.8 and 1, which implies that the predicted isolines are very often similar to those obtained by simulation (Cases 1 and 2 described earlier). Also, 11.5 %
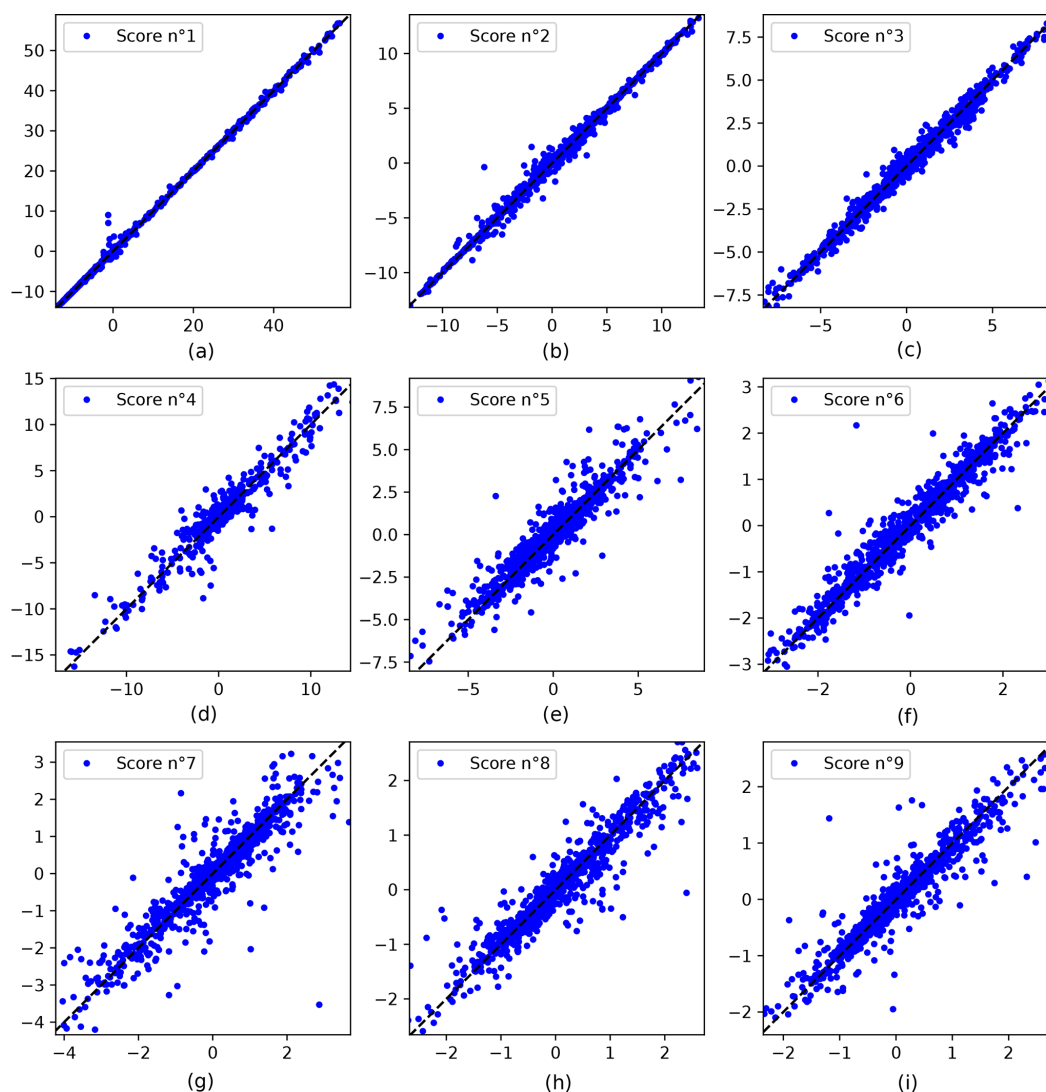
Figure 5. Comparison of the prediction of the emulator (on the y axis) with the target obtained by simulation (on the x axis). Each graph represents one coordinate of the AAM.

of the FMS cannot be calculated because both the simulation and emulator detect no threshold exceedance. Those cases correspond also to good prediction of the emulator. We note that 2.8 % of the FMS is equal to zero, which corresponds to the problem of the non-reached threshold mentioned previously (Case 4). The last 14.8 % of the intermediate FMS corresponds to badly reconstructed isolines (Case 3), for instance, when the wind module is low, which does not necessarily mean that the surface error is large because the FMS is a relative score. Excluding the cases where no threshold is exceeded, badly reconstructed areas amount to 17.6 %, which means that more than 80 % of the threshold exceedance areas are correctly forecast.

4.2 Comparison to other prediction methods

We benchmarked our new prediction method against two state of the art procedures:

– At the start of a crisis, a first estimate is derived from the ATS, yielding orders of magnitude and first results for distance and angular aperture of exceedance. Then simulations are performed with the pX model in order to predict a guide-level exceedance zone and estimate a maximum distance of threshold exceedance. An angular aperture is also estimated from pre-computed tables depending on the wind, atmospheric stability, and meandering wind factor. This simulated maximum distance, associated with this angular aperture obtained without emulation, allows us to deduce a portion of the circle
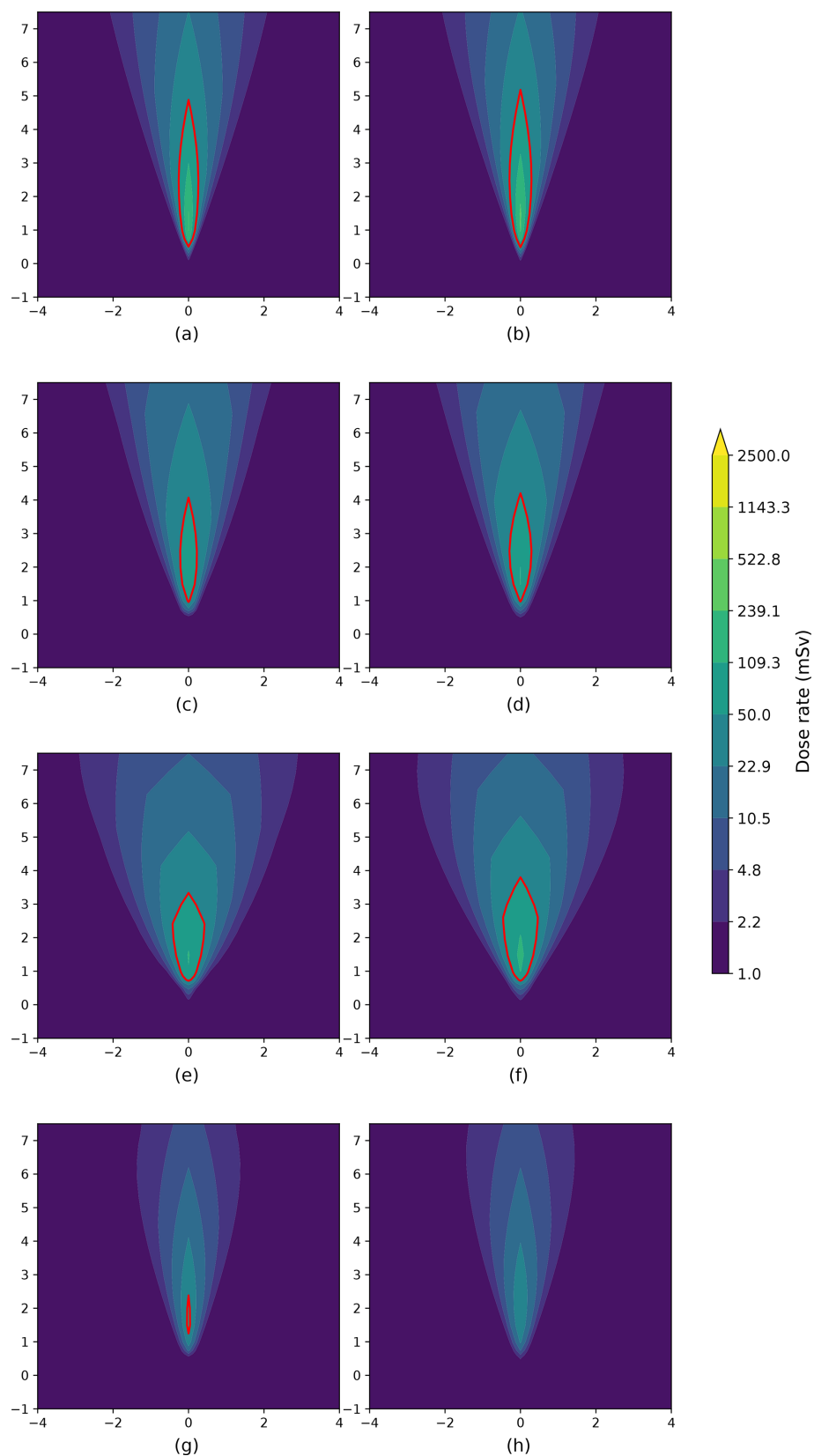
Figure 6. Examples of inhalation dose maps for simulated results by the original physical model (**a, c, e, g**) and emulator-predicted results (**b, d, f, h**).
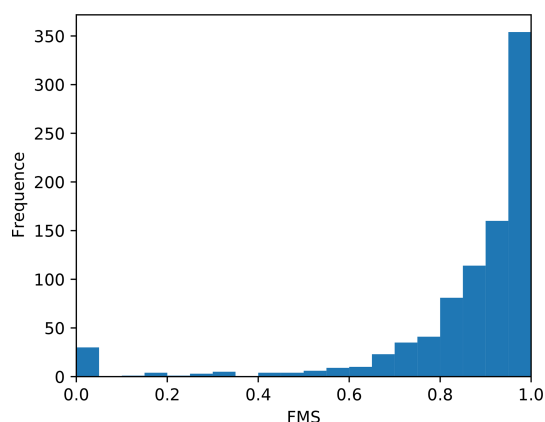
Figure 7. Histogram of the FMS that compares the threshold exceedance zones for the model outputs with the emulator outputs across the testing samples. 70.9 % of the FMS is between 0.8 and 1, which implies that the predicted results are very often similar to those obtained by simulation.

which corresponds to the zone for which decisions are recommended.

– In a previous study (Périllat et al., 2020), an emulator was created to directly predict, without using the AAM, the maximum distance of the threshold exceedance given by the model as well as the angular aperture of this zone. Kriging was used to estimate those two geometrical parameters from the original model, pX.

These two methods will be referred to as the "ATS estimator" and the "emulator of geometrical parameters" in the remainder of this paper.

Figure 8 shows the comparison between the AAM–kriging emulator described in Sect. 3 (Fig. 8c) and the two usual approaches (Fig. 8a and b). In addition, we tested a fourth method: the AAM–kriging emulator applied to a lower threshold exceedance than the actual guide-level value to take a margin on the results obtained by the emulator (Fig. 8d). To achieve that, instead of creating an isoline at $\log(th)$ on our logarithmic dose, we created an isoline at $\log(th)/1.1$.

We compared the isolines of these different prediction methods to the one given by the Gaussian puff model for the 1000 simulations of our test sample. Four kinds of areas may be defined:

– *True positives*: areas where both the predictor and the model forecast a threshold exceedance;

– *True negatives*: areas where the predictor and the model do not forecast a threshold exceedance;

– *False positives*: areas where the predictor forecasts threshold exceedance, while the model forecasts the opposite;
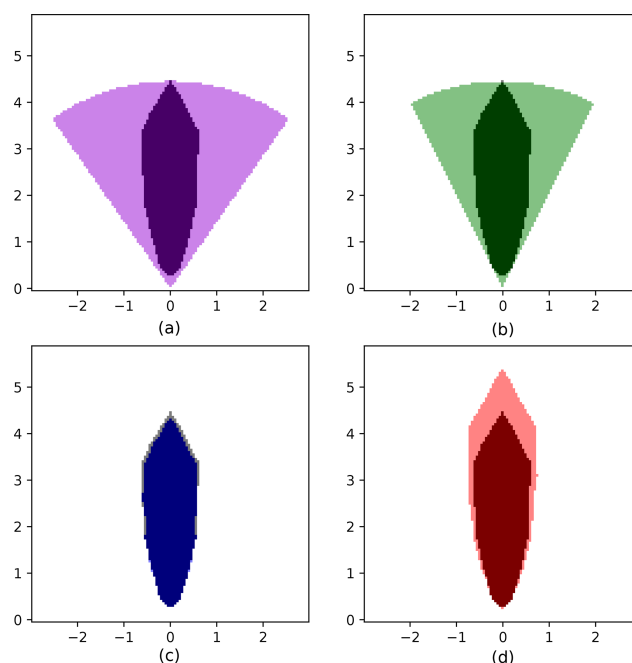

Figure 8. Guide-level exceedance zone for one simulation of the test sample obtained by the different estimation methods. Each subfigure corresponds to a different method: **(a)** the ATS estimator (in purple), **(b)** the emulator of geometrical parameters (in green), **(c)** the emulator coupling kriging to the AAM (in blue), and **(d)** the same emulator with a lower threshold (in red). In each plot, the estimated exceedance zone is shown in a light color, and the reference exceedance zone is given by the Gaussian puff model pX that is superimposed in grey. Colors are consistent with other figures to ease comparison between methods.

– *False negatives*: areas where the predictor does not forecasts a threshold exceedance, while the model forecasts the opposite.

These surfaces allow us to evaluate the performance of each method. Each output pair (model, predictor) is characterized by a certain number of true-positive, true-negative, false-positive, and false-negative areas.

Figure 9 presents the empirical cumulative distribution functions (CDFs) of the false-positive and false-negative surface areas (in km^2) over the 1000 test simulations. The x axis represents the area, while the y axis gives the proportion of simulations with a surface area less than or equal to that value. For both curves, a faster rise towards 1 on the y axis indicates better performance.

Table 3, in contrast, summarizes the total false-positive and false-negative areas accumulated across the 1000 simulations for each prediction method. It provides an aggregate view of each method's overall behavior.

We observe that the ATS estimator, used as a very first response, produces the largest false-positive surface ($10.665\,km^2$), with an extremely low false-negative area
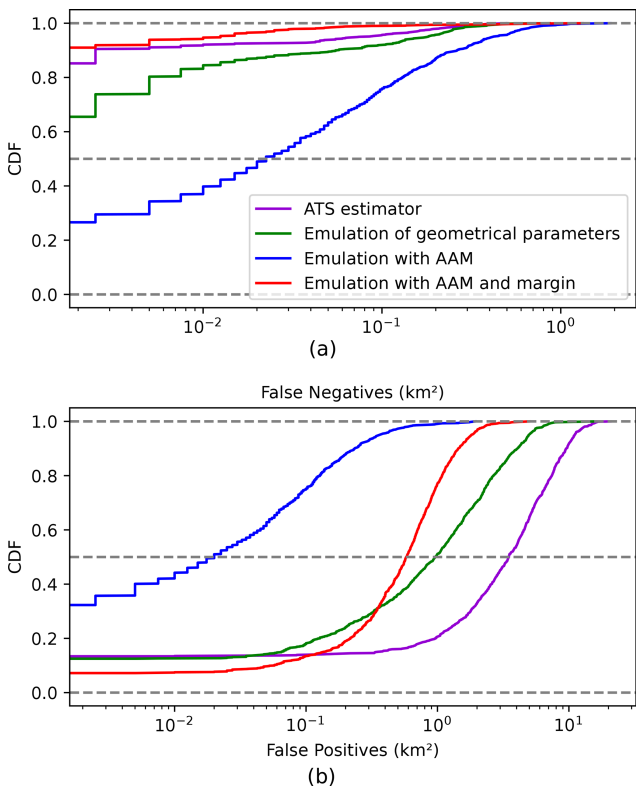
(a)

(b)

Figure 9. Cumulative distribution functions (CDFs) of the probability of obtaining false-positive and false-negative surfaces (in km^2) for the different predictors. These CDFs were estimated from a test sample of 1000 simulations. The plot in **(a)** corresponds to the distribution of false-negative surfaces, and the plot in **(b)** corresponds to the distribution of false-positive surfaces. For both plots, curves closer to the value 1 on the y axis indicate better performance as they correspond to a lower probability of significant false-positive or false-negative areas.

Table 3. Total surfaces (in km^2) of false-positive and false-negative areas over the 1000 test simulations for each method used to predict the zones of dose threshold exceedance. The total area of the studied domain is 94.25 km^2. A false-positive surface corresponds to an area predicted as exceeding the threshold by the method but not by the reference model. A false-negative surface corresponds to an area where the method failed to predict a threshold exceedance that was actually present according to the model. Lower values in both categories indicate better predictive performance.

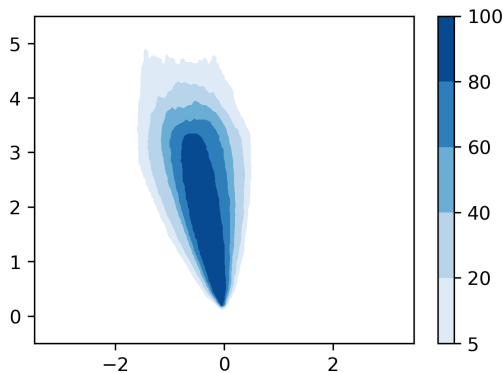| Estimator	Total surface (km^2)	
	False positive	False negative
ATS estimator	10.665	0.030
Emulator of geometrical parameters	3.846	0.056
AAM–kriging emulation	0.221	0.227
AAM–kriging emulation with margin	1.747	0.015

Figure 10. Example of a map which represent the estimation by the emulation with the AAM of the probability of threshold exceedance after a nuclear accident.

(0.030 km^2). This conservative behavior is expected as the method prioritizes avoiding any underestimation of risk.

The emulator of geometrical parameters developed in a previous study reduces the false-positive surface (3.846 km^2) but increases the false negatives (0.056 km^2), which are more critical from a public health perspective.

The AAM–kriging emulator achieves a balanced result with low errors on both sides: 0.221 km^2 of false positives and 0.227 km^2 of false negatives. This method aims to match the reference isoline closely without introducing conservative margins.

Finally, the margin-based AAM–kriging method slightly increases the false positives (1.747 km^2) but drastically reduces the false negatives (0.015 km^2). This trade-off leads to a false-negative performance similar to that of the ATS estimator while reducing false positives by more than a factor of 8.

5   Application

The combination of the AAM and kriging enables the construction of emulators capable of reproducing the model output in approximately 0.005 s compared to about 1 min for the original pX dispersion model. This drastic reduction in computation time opens up new possibilities for operational use, especially in situations where rapid decision-making is crucial, such as nuclear emergencies.

One immediate application of the emulator is to facilitate probabilistic risk assessments. The need to include uncertainties in experts' assessments provided to decision-makers in cases of radiological emergencies was documented in the European H2020 project CONFIDENCE (French et al., 2020), and the use of ensemble simulations was recommended provided the computational time was compatible with the operational time constraints of emergency situations (Bedwell et al., 2020). Since a single prediction now requires only a few milliseconds, it becomes possible to generate thousands
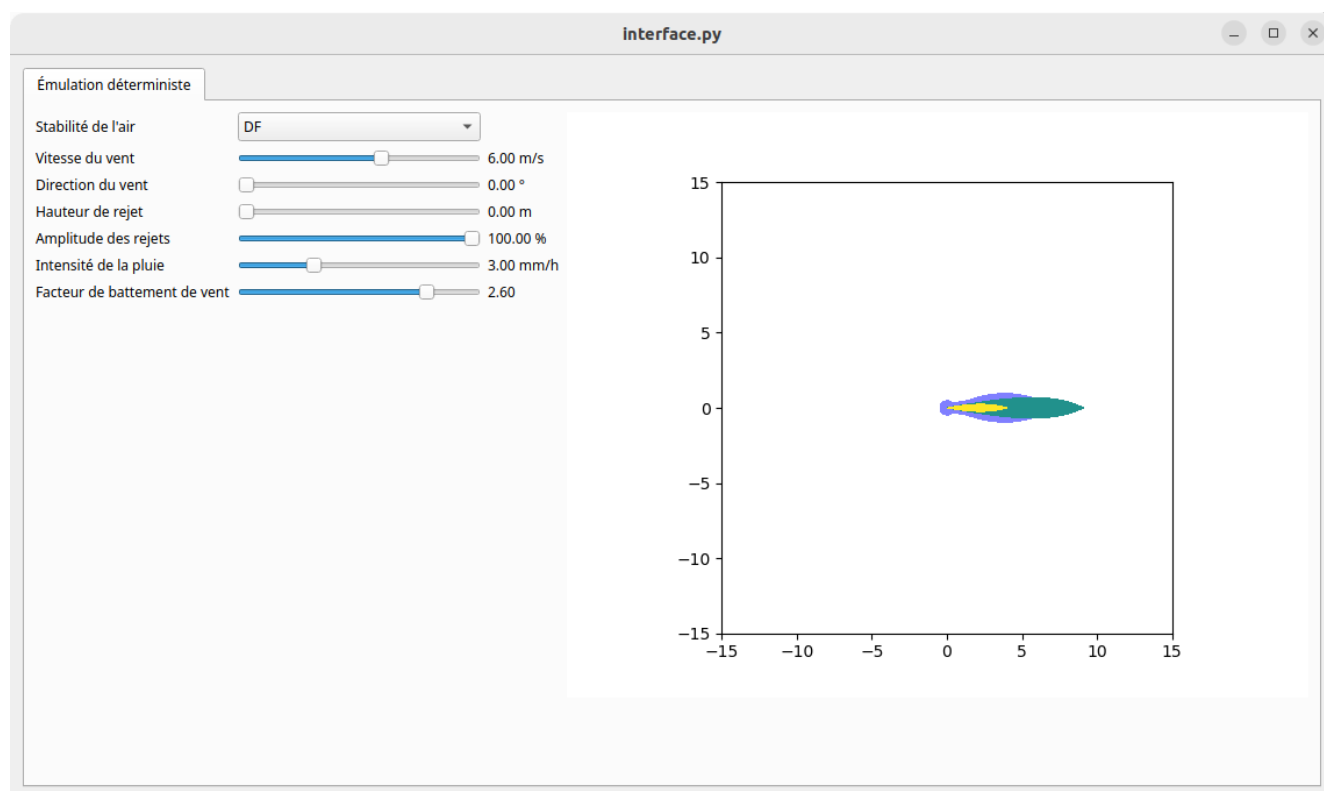
Figure 11. Example of a graphical interface created during the project that enables users to directly observe how each input impacts the dose map.

of simulations across the uncertainty space of the input parameters. This allows for the uncertainty inherent in the atmospheric dispersion process to be fully propagated through to the final dose predictions. As an illustration, Fig. 10 shows a map of the estimated probability of exceeding a given dose threshold. Such probabilistic maps provide emergency managers with a richer understanding of the risk zones compared to a single deterministic prediction (Nagy et al., 2020).

Beyond the probabilistic analysis, we also developed a dedicated graphical interface (see Fig. 11) to make the emulator easily accessible to non-expert users, such as emergency responders. This tool allows users to quickly adjust key input parameters (such as wind speed, release height, or source term amplitude) and immediately visualize the resulting dose map. It thus supports both real-time response activities and training sessions by offering an intuitive way to explore how changes in environmental or release conditions impact the dispersion of radioactive material.

Overall, the AAM–kriging emulator greatly enhances both the reactivity and the robustness of operational decision support systems in the event of an atmospheric radiological release.

6 Conclusions

The AAM coupled with kriging allows us to create emulators which can reproduce the model output with a drastic reduction in computational time. The guide-level dose exceedance isoline obtained with the emulator is very close to the one obtained with the original model: the FMS between the two is under 0.8 in less than 17.6 % of our test sample. For operational use, we recommend to take a margin by reducing the threshold exceedance of the dose. It will slightly increase the false positives but significantly decrease the false negatives.

We obtained very similar results with different guide-levels and with the Doury dispersion model in weak diffusion.

This method of creating an emulator is currently in an operationalization phase to reproduce it on several other source terms and to create a catalog of emulator which covers scenarios among the ATS of the ASNR.

This is the first time the AAM is used for a real and operational application. This study advances the state of the art in atmospheric dispersion by creating a new way to parameterize and predict quantity maps, which can be used in an operational context with probabilistic approaches where hundreds of results must be obtained.

The main limitation of our approach is that in some rare cases, mainly when the wind module is low, the emulator's ability to reconstruct the model's predicted map is lower. However, these cases are also badly forecast by the physical dispersion model itself, and the error of the emulator would not necessarily exceed that of the pX model, were they were compared to environmental observations. We think that these results would be improved by modifying the AAM construction method, which at some point in its process uses a Euclidean distance to compare maps. However, some mathematical distances, such as the Wasserstein distance (Kolouri et al., 2017), for example, could be more suitable for comparing two-dimensional dose maps within them. The AAM method could then benefit from a modification to use other distances in its algorithm to go one step further in improving the application cases.

Author contributions. RP developed the software, analyzed the data, contributed to the study design and methodology, and wrote the manuscript with input from all co-authors. SG and IK contributed to the study design and methodology. IK supervised the research.

References

Baklanov, A. and Sørensen, J.: Parameterisation of radionuclide deposition in atmospheric long-range transport modelling, Phys. Chem. Earth Pt. B, 26, 787–799, https://doi.org/10.1016/S1464-1909(01)00087-9, 2001.

Bedwell, P., Korsakissok, I., Leadbetter, S. J., Périllat, R., Rudas, C., Tomas, J., Wellings, J., Geertsema, G., and de Vries, H.: Operationalising an ensemble approach in the description of uncertainty in atmospheric dispersion modelling AND an emergency response, Radioprotection, 55, S75–S79, https://doi.org/10.1051/radiopro/2020015, 2020.

Burgin, L., Ekström, M., and Dessai, S.: Combining dispersion modelling with synoptic patterns to understand the wind-borne transport into the UK of the bluetongue disease vector, Int. J. Biometeorol., 61, 1233–1245, https://doi.org/10.1007/s00484-016-1301-1, 2017.

Chilès, J.-P. and Delfiner, P.: Geostatistics, Modeling Spatial Uncertainty, John Wiley and Sons, https://doi.org/10.1007/s11004-012-9429-y, 1999.

El-Ouartassy, Y., Korsakissok, I., Plu, M., Connan, O., Descamps, L., and Raynaud, L.: Combining short-range dispersion simulations with fine-scale meteorological ensembles: probabilistic indicators and evaluation during a ^{85}Kr field campaign, Atmos. Chem. Phys., 22, 15793–15816, https://doi.org/10.5194/acp-22-15793-2022, 2022.

French, S., Haywood, S., Oughton, D., and Turcanu, C.: Different types of uncertainty in nuclear emergency management, Radioprotection, 55, S175–S180, https://doi.org/10.1051/radiopro/2020029, 2020.

Fukunaga, K. and Olsen, D. R.: An Algorithm for Finding Intrinsic Dimensionality of Data, IEEE T. Comput., C-20, 176–183, https://doi.org/10.1109/t-c.1971.223208, 1971.

Girard, S.: A nonlinear PCA based on manifold approximation, Computation. Stat., 15, 145–167, https://hal.inria.fr/hal-00724764, 2000.

Girard, S. and Iovleff, S.: Auto-associative models, nonlinear Principal component analysis, manifolds and projection pursuit, in: Principal Manifolds for Data Visualisation and Dimension Reduction, edited by: Gorban, A. N., Kégl, B., Wunsch, D. C., and Zinovyev, A. Y., vol. 58 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, https://doi.org/10.1007/978-3-540-73750-6_8, 202–218, 2008.

Girard, S., Korsakissok, I., and Mallet, V.: Screening sensitivity analysis of a radionuclides atmospheric dispersion model applied to the Fukushima disaster, Atmos. Environ., 95, 490–500, https://doi.org/10.1016/j.atmosenv.2014.07.010, 2014.

Girard, S., Mallet, V., Korsakissok, I., and Mathieu, A.: Emulation and Sobol' sensitivity analysis of an atmospheric dispersion model applied to the Fukushima nuclear accident, J. Geophys. Res.-Atmos., 121, 3484–3496, https://doi.org/10.1002/2015jd023993, 2016.

Girard, S., Armand, P., Duchenne, C., and Yalamas, T.: Stochastic perturbations and dimension reduction for modelling uncertainty of atmospheric dispersion simulations, Atmos. Environ., 224, 117313, https://doi.org/10.1016/j.atmosenv.2020.117313, 2020.

International Atomic Energy Agency: Considerations in the Development of a Protection Strategy for a Nuclear or Radiological Emergency, Emergency Preparedness and Response, International Atomic Energy Agency, Vienna,

https://www.iaea.org/publications/14801/considerations-in-the-development-of-a-protection-strategy-for-a-nuclear-or-radiological-emergency (last access: 14 August 2025), 2021.

Jolliffe, I. T.: Principal Component Analysis, in: 2nd Edn., Springer, https://doi.org/10.1007/b98835, 2002.

Jolliffe, I. T. and Cadima, J.: Principal component analysis: a review and recent developments, Philos. T. Roy. Soc. A, 374, 20150202, https://doi.org/10.1098/rsta.2015.0202, 2016.

Kolouri, S., Park, S. R., Thorpe, M., Slepcev, D., and Rohde, G. K.: Optimal mass transport: signal processing and machine-learning applications, IEEE Signal Proc. Mag., 34, 43–59, https://doi.org/10.1109/MSP.2017.2695801, 2017.

Korsakissok, I., Mathieu, A., and Didier, D.: Atmospheric dispersion and ground deposition induced by the Fukushima Nuclear Power Plant accident: a local-scale simulation and sensitivity study, Atmos. Environ., 70, 267–279, https://doi.org/10.1016/j.atmosenv.2013.01.002, 2013.

Le, N. B. T., Mallet, V., Korsakissok, I., Mathieu, A., Perillat, R., and Didier, D.: Metamodeling and optimization of probabilistic scores for long-range atmospheric dispersion applied to the Fukushima nuclear disaster, in: vol. 20, EGU General Assembly Conference Abstracts, Vienna, Austria, 8–13 April 2018, 2018.

Le, N. B. T., Korsakissok, I., Mallet, V., Périllat, R., and Mathieu, A.: Uncertainty study on atmospheric dispersion simulations using meteorological ensembles with a Monte Carlo approach, applied to the Fukushima nuclear accident, Atmospheric Environment: X, 10, 100112, https://doi.org/10.1016/j.aeaoa.2021.100112, 2021.

Leadbetter, S. J., Andronopoulos, S., Bedwell, P., Chevalier-Jabet, K., Geertsema, G., Gering, F., Hamburger, T., Jones, A., Klein, H., Korsakissok, I., Mathieu, A., Pázmándi, T., Périllat, R., Rudas, C., Sogachev, A., Szántó, P., Tomas, J., Twenhöfel, C., de Vries, H., and Wellings, J.: Ranking uncertainties in atmospheric dispersion modelling following the accidental release of radioactive material, Radioprotection, 55, S51–S55, https://doi.org/10.1051/radiopro/2020012, 2020.

Lumet, E., Rochoux, M. C., Jaravel, T., and Lacroix, S.: Uncertainty-aware surrogate modeling for urban air pollutant dispersion prediction, Build. Environ., 267, 112287, https://doi.org/10.1016/j.buildenv.2024.112287, 2025.

Mallet, V., Tilloy, A., Poulet, D., Girard, S., and Brocheton, F.: Meta-modeling of ADMS-Urban by dimension reduction and emulation, Atmos. Environ., 184, 37–46, https://doi.org/10.1016/j.atmosenv.2018.04.009, 2018.

Milnor, J. W.: Topology from the Differentiable Viewpoint, Princeton University Press, ISBN 9780691048338, 1997.

Nagy, A., Perko, T., Müller, T., Raskob, W., and Benighaus, L.: Uncertainty visualization using maps for nuclear AND radiological emergencies, Radioprotection, 55, S197–S203, https://doi.org/10.1051/radiopro/2020033, 2020.

Nomura, S., Gilmour, S., Tsubokura, M., Yoneoka, D., Sugimoto, A., Oikawa, T., Kami, M., and Shibuya, K.: Mortality risk amongst nursing home residents evacuated after the Fukushima nuclear accident: a retrospective cohort study, PLoS One, 8, 1–9, https://doi.org/10.1371/journal.pone.0060192, 2013.

Nomura, S., Blangiardo, M., Tsubokura, M., Ozaki, M., and Hodgson, S.: Postnuclear disaster evacuation and chronic health in adults in Fukushima, Japan: a long-term retrospective analysis, BMJ Open, 6, 1–9, https://doi.org/10.1136/bmjopen-2015-010080, 2016.

Périllat, R.: perillat/emulation_aam: Metamodeling Scripts for Kriging and Auto-Associative Models, Zenodo [code], https://doi.org/10.5281/zenodo.14856799, 2024a.

Périllat, R.: Data for the paper "Accurate and fast prediction of radioactive pollution by Kriging coupled with Auto-Associative Models", Zenodo [data set], https://doi.org/10.5281/zenodo.14747261, 2024b.

Périllat, R., Korsakissok, I., Girard, S., and Quentric, E.: Emulators for the rapid prediction of consequences in case of nuclear hazards, in: 20th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, HARMO20, Tartu, Estonia, 14–18 June 2021, hal-04337764, 2020.

Rasmussen, C. E. and Williams, C. K. I.: Gaussian Process for Machine Learning, MIT Press, https://doi.org/10.7551/mitpress/3206.003.0019, 2006.

Roustant, O., Ginsbourger, D., and Deville, Y.: DiceKriging, DiceOptim: two R packages for the analysis of computer experiments by kriging-based metamodeling and optimization, J. Stat. Softw., 51, 1–55, https://doi.org/10.18637/jss.v051.i01, 2012.

Swallow, B., Rigby, M., Rougier, J. C., Manning, A. J., Lunt, M., and O'Doherty, S.: Parametric uncertainty in complex environmental models: a cheap emulation approach for models with high-dimensional output, arXiv [preprint], https://doi.org/10.48550/arXiv.1702.03696, 2017.

Tombette, M., Quentric, E., Quélo, D., Benoit, J.-P., Mathieu, A., Korsakissok, I., and Didier, D.: C3X: A software platform for assessing the consequences of an accidental release of radioactivity into the atmosphere, Poster presentation, presented at the 4th European IRPA Congress, 23–27 June 2014, Geneva, Switzerland, 2014.