17th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
9-12 May 2016, Budapest, Hungary

PROBABILISTIC ASSESSMENT OF DANGER ZONES ASSOCIATED WITH A HYPOTHETICAL ACCIDENT IN A MAJOR FRENCH PORT USING A SURROGATE MODEL OF CFD SIMULATIONS

Felipe Aguirre Martinez[1], Yann Caniou[1], Christophe Duchenne[2], Patrick Armand[2], and Thierry Yalamas[1]

[1]PHIMECA Engineering, F-75012 Paris, France
[2]CEA, DAM, DIF, F-91297 Arpajon, France

Abstract: This paper presents a probabilistic framework used to assess the risk of exceeding health thresholds after an accidental release of a pollutant. Probabilistic risk maps are produced that highlight the 95% confidence interval of the boundary of the danger zone. The methodology is applied to an accidental release of ammoniac simulated with PMSS. Uncertain release and atmospheric conditions are propagated through the model using a Monte-Carlo approach. Gaussian predictors are used to replace the long running PMSS model. It is shown that taking uncertainty into account prevents us from neglecting potential danger zones not identified by a simple deterministic approach.

Key words: *Probabilistic risk assessment, dispersion, Gaussian process predictors, Monte-Carlo simulations.*

IINTRODUCTION

Risk assessment studies are an invaluable support in the framework of emergency preparedness and response. They are of special importance when dealing with toxic industrial chemicals releases in complex built environments. Simulations on this kind of sites require 3D physical models capable of integrating the influence of the topography and buildings on the flow field and pollutant dispersion. The 3D simulations are most often carried out using a deterministic set of parameters describing the release and meteorological conditions (location, release height and rate, wind speed and direction, stability…). However, these parameters are highly variable or partially unknown. To tackle this, the inherent uncertainty can be propagated through the models using probabilistic methods. The interpretation of the danger zones gets associated to a probability of exceeding a critical dose conditioned by the uncertain parameters. However, this approach is very time consuming for large dimensions and high-resolution domains. An alternative method that reduces the computational cost is to resort to a surrogate model of the flow and dispersion models. Such a methodology has already been developed and tested on a local scale (Armand *et al*. 2014; Dubourg *et al*. 2013). This paper proposes the application of the methodology on a bigger scale, a fictitious accidental release on an industrial site in a French port. Gaussian process predictors are used in combination with a dimension reduction based on Principal Component Analysis. The results are validated by making use of the crude Monte-Carlo uncertainty propagation. Results are presented as probabilistic danger zones maps with a confidence interval of 95% and compared with the "deterministic" result obtained by simulation of the original models on the mode values of the input parameters. Finally, the interest in taking the uncertainties into account is commented on as well as the computational resources needed by the method.

CASE STUDY: FICTITIOUS ACCIDENTAL RELEASE OF AMMONIAC

The proposed method was applied to a fictitious situation of an accidental release on an industrial site in a French port. Following an accident in the port, ammoniac is released for a period of 45 minutes. The meteorological flows and atmospheric dispersion are simulated using the Parallel-Micro-SWIFT-SPRAY (PMSS) modelling system from the CEA and ARIA Technologies. The simulation extends for 75 mins after the release (120 mins in total), the time that the pollutant exits the domain of interest. Meteorological conditions fluctuate within two main wind directions impacting the lower and upper parts of the town (cf. **Figure 1**). The parameters that describe the release and meteorological conditions are considered to be uncertain. Table 1 summarizes these parameters as probability distributions whose parameters are based on expert judgement. The 9 km x 12 km domain is divided in 63 tiles (9 columns and 7 rows) with a horizontal mesh resolution of 3 meters. The cliffs relief and all buildings details are taken into account in the flow and dispersion models and their surrogate. The computational time is of approximately 1 hour and 40 minutes over 128 CPUS.

Table 1 - Probability model of the uncertain release and meteorological conditions

Variable	**Distribution**
Wind speed 1 (°)	$N(170, \sigma_1 = 10)$
Wind speed 1 (m/s)	$N(3, 0.3)$
Wind direction 2 (°)	$N(120, \sigma_2 = 10\ )$
Wind direction 2 (m/s)	$N(4, 0.4)$
Temperature gradient (°C/100 m)	$LN(\mu = -0.1, \sigma = 0.7, \gamma = -2)$
Rejection height (m)	$U[0,20]$
Amount of pollutant (kg/m^3/s)	$LN(\mu = 7.65\ E9, \sigma = 5.1\ E9, \gamma = 0)$

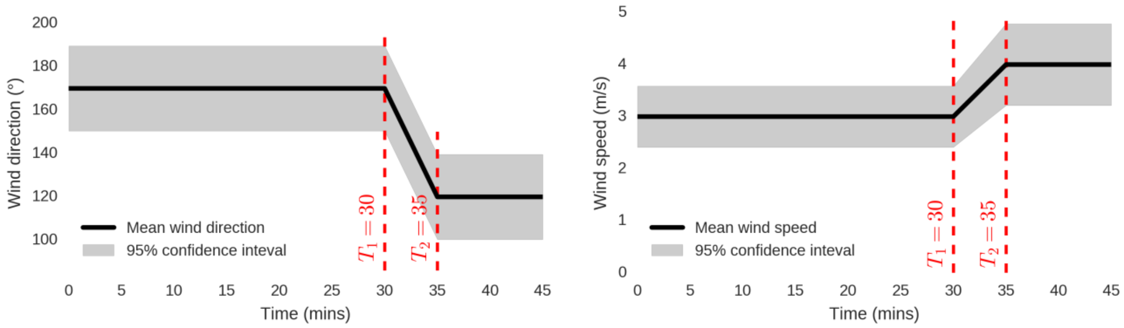
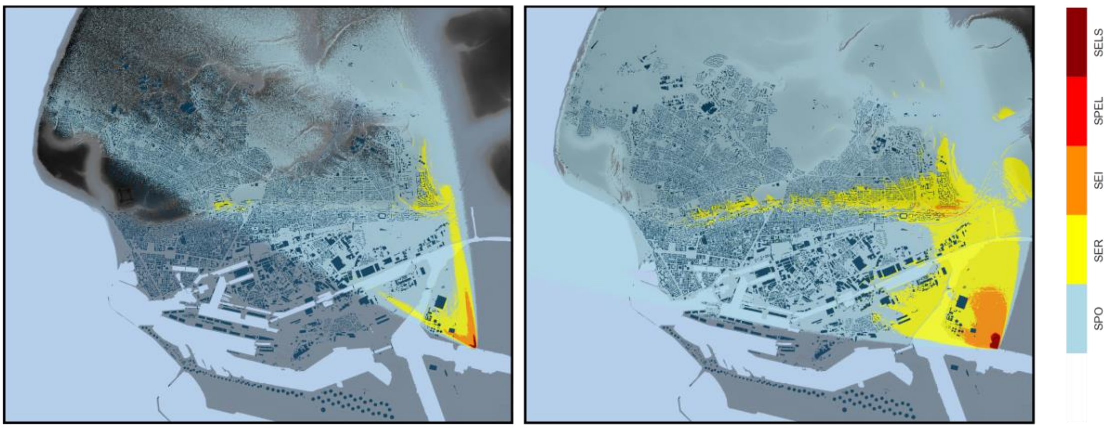


Figure 1. Wind conditions



a) Deterministic approach evaluated at the mode of the parameters

b) Monte-Carlo approach with $N = 1000$. The probability of exceeding the thresholds is greater than 0.025.

Figure 2. Comparison of risk maps obtained with deterministic and probabilistic approach.

The objective is to quantify the health risk through time and space following the accidental release of ammoniac. Figure 2.a shows the risk map of the proposed case study when applying a deterministic workflow using the mode of the uncertain parameters. This risk map is the reference upon which the probabilistic approach is compared. Five safety thresholds are depicted:

- SER: Reversible effects threshold
- SEI: Irreversible effects threshold
- SPEL: First lethal effects threshold
- SELS Significant lethal effects threshold

The SPO odor threshold is also considered.

DISPERSION MODELING

PMSS is the parallel version of the MSS modelling system integrating the diagnostic mass-consistent model MicroSwift with MicroSpray Lagrangian particle dispersion model. MicroSwift (Oldrini et al. 2013) interpolates the input wind data, coming from outputs of a larger scale model, from a dispersed meteorological network or both of them, on the simulation 3D domain through an objective analysis based on the mass conservation equation. Temperature and humidity can also be interpolated, such as some 2D parameters like the friction velocity or the mixing layer height, if these data are available. In MSS the total turbulence is obtained summing the local one, produced by the flow distortion around the obstacles, plus a background level obtained by standard boundary-layer parameterizations. The local turbulence is estimated on the basis of a mixing-length closure, with the mixing length being a function of the distance to the obstacle or the ground. MicroSpray (Tinarelli et al. 2013) is able to take into account the presence of obstacles. The dispersion of an airborne pollutant is simulated following the motion of a large number of fictitious particles. The mean ("transport") component of the particle velocity is provided by the meteorological driver. The stochastic ("turbulent") component of the particle motion is obtained by solving a 3-D form of the Langevin equation for the random velocity. The parallelization of MSS is based on the MPI message-passing system in order to deal with huge computation domains. Parallelization is implemented in ParallelMicroSwift both for splitting geographically huge domains, and for gaining speed-up based on code specific properties. ParallelMicroSpray has an elaborated load balancing to provide sub domains containing numerous particles with maximum core power, and is able to transition particles between sub domains.

UNCERTAINTY PROPAGATION FOR RISK ASSESSMENT

Risk is quantified as the probability that the ammoniac dose perceived by an individual in a given point of the city for a given period of time exceeds the critical health thresholds. Intuitively, this probability can be quantified with Monte-Carlo sampling. For example, the probability of exceeding the irreversible effects threshold dose SEI in a given point p, within a period of time t is estimated as follows:

$$P[D(X;p,t) > SEI] \approx \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{D(p,t)>SEI}(x^{(i)}) \qquad (1)$$

where $D(X;p,t)$ corresponds to the computational chain associated to the case study and evaluated for X, the random vector of the uncertain parameters. $\mathbf{1}_{D(p,t)>SEI}$ is the indicator function that is equal to one if the SEI threshold is exceeded for a given realization $x^{(i)}$ of X and zero otherwise. N is the size of the Monte-Carlo experiment.

As a rule of thumb, a confident estimate of a probability of the order 10^{-k} requires at least a Monte-Carlo experiment of size 10^{k+2} . For this study, the attention is focused on a 0.025 probability level; we then need a Monte-Carlo experiment of at least 3900 PMSS simulations in order to correctly estimate the risk map. Such an experiment would require 9 months on a sequential machine, which is incompatible with the urgency related to a potential real scenario. Moreover, each PMSS simulation produces ~10 GB of data, which means that ~39TB of available storage space are needed for such a design of experiments. To overcome this, we use the methodology introduced by (Armand *et al.* 2014). PMSS is replaced with vectorial Gaussian process predictors in combination with a dimension reduction based on Principal Component Analysis. Besides the fact that the scenario, the pollutant and the parameters taken into account have changed, the novelty of this study lies on the following factors:

- The models are implemented using the PMSS modelling system.
- The methodology was adapted so that it can be applied on a much larger scale with a division of the domain in tiles. The simulations are distributed on a High Performance Computing (HPC) cluster over 128cpus for each simulation.
- A change on the wind direction during the rejection period is considered.

RESULTS

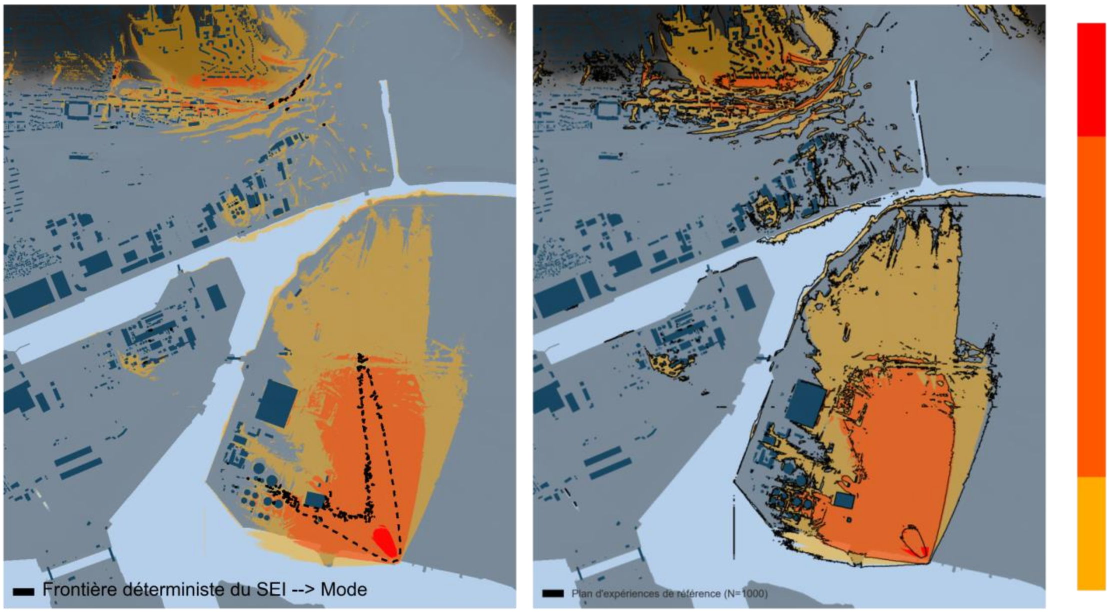
In order to validate the methodology, a reference probabilistic risk map was constructed with a Sobol design of experiments of size 1000. The Sobol design of experiments has the advantage that it is space-filling and guarantees that any of its sub-sequences is also space-filling (Santner et al. 2003). Thanks to this property, subsequences of the design of experiments can be used to fit the Gaussian process predictors without the need to run additional simulations. The simulations were distributed over an HPC cluster and obtained in 19 days. Due to storage limits, it was impossible to constitute the 3900 points. The predictors are then used to obtain the 3900 simulations needed to correctly estimate the risk maps. Figure 2.b shows the probabilistic risk map constructed from the 1000 simulations. The enclosed zones correspond to the upper bound (conservative stand) of the 95% confidence interval of the location of the boundary of the risk zones. For a closer comparison with the deterministic approach, Figure 3.a shows the 95% confidence interval of the location of the SEI boundary with the deterministic risk zone superposed on black. The orange zone (the 95% confidence interval) can be seen as the uncertainty of the location of the real boundary of the risk zone due to the uncertainty about the release and atmospheric conditions. The location of the boundary is highly dependent on the wind conditions, which explains why the confidence interval is so wide compared to the deterministic approach. Only near the source (red zone) it can be concluded with great certainty (probability greater than 0.975) that the SEI threshold is exceeded. The more we go further from the source, the more the risk will depend on the wind direction and speed. To the point where the yellow zone has a probability lower than 0.025 of exceeding the SEI threshold.

Vectorial Gaussian process predictors are used to predict the 3900 simulations needed to correctly estimate the 95% confidence interval. One predictor is fitted for each tile of the domain. For demonstration purposes, only the SEI threshold is considered. Due to memory limitations, some extra steps are required:

- Only the tiles / points of interest are taken into account in the Gaussian process predictors. A point is considered of interest if for the risk map in Figure 2.b it exceeds the safety threshold. A tile is considered of interest if it has at least one point of interest.
- For each tile of interest, a vectorial Gaussian process is fitted only over the points of interest.
- Tiles that have more than 10000 points of interest had to be divided further in sub-tiles.

One of the advantages of Gaussian process predictors is that they yield a response in the form of a probability distribution whose mean is the most probable prediction and whose standard deviation depends on an auto-evaluation of the prediction error (Santner et al. 2003). This information is used to compute confidence intervals of the predictions. In a conservative stand, we take into account the upper bound of such an interval for the dose predictions.

The methodology was tested with a learning design of experiments of 100, 200, 300 and 400 simulations. It was concluded that starting at 300 simulations the results were satisfactory. Below this number, there are not enough points to represent the whole range of the phenomena over the spatial grid, mostly due to the fact that the wind direction is uncertain. The Gaussian predictors fitted with 400 simulations managed to better predict low probability zones. The fitting time was of 40 minutes with a prediction time of 5 minutes for the 3900 simulations. The predictors were validated using the leave-one-out method to estimate the coefficient of determination R^2 (Santner *et al.* 2003). Only three tiles had an R^2 below 0.75, but they are far from the source and have low impact. **Figure 3**.b shows the risk map constructed with the Gaussian predictors approach. For the sake of comparison, the Monte-Carlo reference is superposed in black. It can be seen that the 0.025 probability boundaries are tightly closed together, which is the target from a conservative stand.



a) Deterministic approach superposed in black to the Monte-Carlo reference.

b) Monte-Carlo reference superposed in black to the Gaussian predictor approach.

Figure 3. 95% confidence interval of the boundary of the risk zone for the SEI threshold.

CONCLUSIONS

This study constitutes an improvement of the probabilistic risk assessment approach previously introduced by (Armand et al. 2014). It allowed the extension of the methodology to a larger scale divided in tiles and to the case where the wind conditions are considered uncertain and vary along the release period. It was concluded that the wind conditions have a great impact on the position of the boundary of the risk zone, which is why the 95% confidence interval is considerably wide. Taking uncertainty into account also revealed danger zones that are not identified by the deterministic approach (near the top at the level of a cliff). This reinforces the need to take into account uncertainty in this type of studies. Even when applying a worst-case-scenario deterministic approach, it is impossible to obtain a conservative risk map that reflects the uncertainty about the wind conditions and some potential danger zones may be neglected.

REFERENCES

Armand, P. et al., 2014. Probabilistic safety analysis for urgent situations following the accidental release of a pollutant in the atmosphere. *Atmospheric Environment*, **96**, .1–10. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1352231014005433.

Dubourg, V. et al., 2013. Uncertainty treatment in dispersion modelling of accidental releases. In *15th* International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes. Madrid, Spain.

Santner, T.J., Williams, B.J. & Notz, W., 2003. The design and analysis of computer experiments, Springer.

Tinarelli, G., L. Mortarini, S. Trini-Castelli, G. Carlino, J. Moussafir, C. Olry, P. Armand, and D. Anfossi. 2013. Review and validation of Micro-SPRAY, a Lagrangian particle model of turbulent dispersion. *Lagrangian Modeling of the Atmosphere, Geophysical Monograph*, Volume 200, AGU, 311-327.