Uncertainty propagation in a model for the estimation of the ground level concentration of dioxin/furans emitted from a waste gasification plant

G. Ripamonti, G. Lonati, Piero Baraldi, F. Cadini, Enrico Zio

To cite this version:


HAL Id: hal-00934518
https://hal-supelec.archives-ouvertes.fr/hal-00934518
Submitted on 22 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Uncertainty propagation in a model for the estimation of the
ground level concentration of dioxin/furans emitted from a waste
gasification plant

G. Ripamonti & G. Lonati
DIIAR-Environmental Section - Politecnico di Milano, Milano, Italy

P. Baraldi, F. Cadini,
Dipartimento di Energia – Politecnico di Milano, Milano, Italy

E. Zio
Chair on Systems Science and the Energetic challenge, European Foundation for New Energy-
Électricité de France

Dipartimento di Energia – Politecnico di Milano, Milano, Italy

ABSTRACT: In this paper we compare two approaches for uncertainty propagation in a model for
Environmental Impact Assessment (EIA). A purely Probabilistic (PMC) and a Hybrid probabilistic-
possibilistic Monte Carlo (HMC) method are considered in their application for the estimation of
the ground levels concentration of dioxin/furans emitted from a waste gasification plant. Under the
condition of insufficient information for calibrating the estimation model parameters, HMC is
shown to be a valid way for properly propagating parameters uncertainty to the model output, with-
out adopting arbitrary and subjective assumptions on the input probability distribution functions. In
this sense, HMC could improve the transparency of the EIA procedures with positive effects on the
communicability and credibility of its findings.

Keywords: Uncertainty propagation, Hybrid probabilistic-possibilistic Monte Carlo, waste gasifica-
tion plant, Environmental Impact Assessment.

1. INTRODUCTION

According to the current EU regulations (Directive 85/337/EEC as amended by Directives
97/11/EC and 2003/35/EC), an Environmental Impact Assessment (EIA) is required for public and
private projects likely to have significant impacts on the environment by virtue inter alia of their
nature, size or location. The scope of an EIA is the identification, description and assessment of a
project’s direct and indirect effects on human beings, fauna and flora, soil, water, air, climate, land-
scape, material assets and cultural heritage. The EIA is a part of the decisional process for the ap-
proval of a proposed project. The outcomes of the EIA help orient decision makers to the priorities,
thus ensuring that environmental considerations are taken into account in an effective way. Fur-
thermore, the EIA serves an important procedural role in the overall decision-making process by
promoting public information and facilitating participation. This entails that an EIA be carried out
taking into account that its findings will be handled not only by technical experts, but also commu-
nicated to decision makers and stakeholders of various nature, including the local population. Since
in most common settings the proposed project does not meet popular consensus *a priori*, communic-
ability of EIA results to the involved stakeholders plays a major and challenging role. To this pur-
pose the EIA procedure must be open and transparent, reviewable and leading to robust and repro-
ducible outcomes for informed decision-making and confident public awareness.

With reference to the proposal of new facilities responsible of atmospheric emissions (in the present
work we consider a waste thermal treatment plant as an example), the EIA procedure consists of
three main steps (Figure 1):

a) the characterization of the source: estimation of the atmospheric emissions resulting from the
operation of the plant, in terms of both the nature and the quantities of pollutants;

b) the estimation of the atmospheric pollutant concentration levels in the area close to the proposed
new installation, by means of proper atmospheric dispersion models;

c) the assessment of the pollutant concentration levels at the receptors, resulting from the superp-
osition of the estimated concentration to the existing background concentration levels.

The first objective of the EIA is then the demonstration that the resulting concentration levels at the
receptors comply with both long-term (i.e.: annual average) and short-term (i.e.: hourly and daily
averages) air quality limits for regulated pollutants. Secondly, as those facilities can be also respon-
sible for the emission of both carcinogenic and non-carcinogenic trace pollutants, as metals or diox-
ins and furans (PCDD/Fs) in the case of waste thermal treatment plants, additional evaluations may
also include human health risk assessment. This latter assessment is characterized by two main fea-
tures: i) a considerable complexity of the mechanisms involved and hence large related uncertain-
ties, ii) a general attitude of “aversion to risk”, with implications on the way uncertainties are
treated (Guyonnet et al., 2008).

The uncertainties that affect all the three steps of the EIA procedure may be originated from ran-
domness of the physical processes involved or from lack of precise knowledge of them. In the end,
regardless of the approaches and models adopted to estimate the plant emissions and their disper-
sion in the atmosphere, the final ambient concentrations estimates can be seriously affected by the
uncertainty of the values of the source term parameters and by the variability of meteorological pa-
rameters.

On the contrary, the above mentioned attitude of “aversion to risk” often leads to answering to the
question whether a given “acceptable” threshold may be exceeded, based on limited scenario calcu-
lations. In common practice, this translates in ignoring input parameter uncertainties to adopt a di-
rect, deterministic approach in which conservative assumptions on the scenarios and parameter val-
ues are made, e.g.: i) full-load plant operation and ii) concentrations of the emitted pollutants equal
to the regulatory limits. This approach leads to often excessively conservative results, not providing
a realistic and objective description of the plant impact on the environment (Lonati and Zanoni, 2012)

To overcome this problem, the EIA procedure can be framed within a probabilistic approach for i) describing the uncertainties of the input parameters and ii) propagating them to the models outputs, i.e.: the atmospheric pollutant concentrations. In the probabilistic approach, the uncertain variables are described by means of probability distribution functions (PDFs), which are then propagated through the impact assessment model by Monte-Carlo (MC) simulation according to a pure Probabilistic Monte Carlo (PMC) method (Schuhmacher et al. 2001, Sonnemann et al. 2002, Lonati et al. 2007). Nevertheless, even in the PMC approach uncertainty and variability in the atmospheric dispersion model are usually neglected.

In general, the construction by statistical analysis of the empirical PDFs representing the uncertainty in the input parameters may not be an easy task due to the scarce availability of data (Ferson and Ginzburg 1996, Baudrit et al. 2006). Within a probabilistic framework, the analyst traditionally solves this task forcing a subjective PDF on the scarce available data, within a Bayesian (subjective) view of probabilities. This presents two major drawbacks:

- doubts may be raised on the consistency and even conservativeness of the forced PDFs, in light of the incomplete/imprecise information at hand. The risk is to introduce artificial information that is in fact not there;
- the subjectivity in the analysis is often neither declared nor justified, causing a loss of transparency in the procedure.

These drawbacks are of particular relevance, considering the EIA role as decision aiding and informative tool. A lack of transparency in the procedure can lead to a loss of credibility by the decision makers and public.

To cope with this situation of limited information, while avoiding to introduce subjective assumptions, a number of alternative representations of uncertainty have been proposed (e.g.: fuzzy set theory, evidence theory, possibility theory and interval analysis) and applied in different fields of risk assessment (Klir and Yuan 1995; Aven and Zio 2011). Among them, possibility theory has received growing attention because of its representation power and its relative mathematical simplicity. In recent past years some applications of these alternative approaches have been exploited also in the environmental risk assessment field (Guyonnet et al. 2003, Kentel and Aral 2005, Li et al. 2007, Kumar et al. 2009).

In practice, different input parameters of the same model may be characterized by different amounts of information availability, thus possibly giving rise to both probabilistic and possibilistic uncertainty representations. In this regard, Hybrid probabilistic-possibilistic Monte Carlo methods
(HMC) have been recently proposed for combining both representation frameworks in the same uncertainty propagation analysis (Baudrit et al. 2006, Baraldi and Zio, 2008, Flage et al. 2010).

In this work, the uncertainty representation and propagation problems are considered with reference to the first two steps of the EIA procedure (i.e.: emission estimation and atmospheric dispersion).

For ease of illustration, we make reference to a case study concerning the project of a new waste gasification plant. The plant, designed for a daily waste throughput of 900 metric tons, has been recently proposed for realization in the South-East of the city of Milan, Italy and is subject to the requirement of EIA for its potential impact on air quality.

The focus of the analysis is on dioxins and furans (PCDD/Fs) since the thermal treatment of waste is a well-known source of PCDD/Fs to the atmosphere (Domingo et al. 2002, Schuhmacher and Domingo 2006) and these pollutants have a relevant role in health risk assessment due to their carcinogenic potential and to their persistency in the environment, once released into the atmosphere (Travis and Hattemer-Frey 1991, Mukerjee 1998).

The final objective of the work is the assessment of the contribution to the ambient air concentration $C_{\text{air}}$ (fg m$^{-3}$) on an annual average basis can be computed by multiplying an emission term $Q$ (ng s$^{-1}$), which quantifies the mass flow rate of...
PCDD/Fs released into the atmosphere, and an atmospheric dispersion factor DF (fg m$^{-3}$/ng s$^{-1}$), which quantifies the ambient concentration per unit of mass flow rate:

\[ C_{\text{air}} = Q \cdot DF \]  

(1)

Since the atmospheric dilution of the emission occurs at different extents depending on the distance from the plant stack, $C_{\text{air}}$ and DF are actually space-dependent and take different values, $C_{\text{air}}(x,y)$, DF(x,y), at the receptors points located in the study domain, usually described by a Cartesian grid centered on the plant stack.

Equation (1) follows a simplified approach for the assessment of $C_{\text{air}}$ since it assumes that the PCDD/Fs are emitted in the gas-phase only: this approach, although still providing realistic results, reduces the complexity of the calculation model thus enabling the application of the uncertainty propagation methods.

The emitted PCDD/Fs mass flow rate $Q$ in equation (1) has been computed on the basis of the plant daily throughput of waste, $P$ (Mg waste day$^{-1}$), the PCDD/Fs concentration in the emitted flue gas, $C_D$ (ng m$^{-3}$), and the specific gas production $V_F$ (m$^3$ Mg$^{-1}$ waste$^{-1}$):

\[ Q = \frac{P \cdot C_D \cdot V_F}{3600 \cdot 24} \]  

(2)

As usual for this kind of evaluations (Van den Berg et al. 1998, US EPA 2005), the PCDD/Fs concentration is expressed in terms of equivalent toxicity mass per unit gas volume at normal conditions (0° C, 101.3 kPa); coherently, the specific gas production refers to the same temperature and pressure conditions.

The dispersion term DF in equation (1) has been computed by simulating the atmospheric transport and dispersion of the emitted PCDD/Fs in the surroundings of the plant by means of the ISCST3 model (Industrial Source Complex – Short-term version 3). Formerly a US EPA regulatory model for air dispersion modeling, this model is still regarded as a screening model for the assessment of the impact on air quality of atmospheric emissions from industrial sources (US EPA, 1995).

Local meteorological conditions (wind speed and direction, ambient air temperature, atmospheric stability), source features (stack height, flue gas speed and temperature, pollutant mass flow rate) and geographical features of the area (terrain elevation, land use) are the basic input data for the dispersion model. Output of the ISCST3 model is the 2D field of the contribution to the ground-level concentration data in the study domain around the plant. Since in this model the output con-
centrations are linearly proportional to the emitted flow rate, the dispersion term DF can be evaluated running the model for a unit mass flow rate (1 ng s\(^{-1}\)).

The DF values to be used in equation (1) for the estimation of the PCDD/Fs annual average concentration have been obtained for the 1681 nodes of a Cartesian grid centered on the plant (250 m cell spacing) through one-year long model simulations based on the hourly time series of locally measured meteorological data.

3. UNCERTAINTY REPRESENTATION AND PROPAGATION

The input parameters of the emission model (equation (2)) have been characterized as follows:

- the plant daily throughput \(P\) is a constant parameter whose value is assigned during project design; for the plant of this case study, \(P\) has been set to the value of 900 Mg\(\text{waste day}^{-1}\);
- the PCDD/Fs concentration \(C_D\) is an uncertain parameter whose value varies during normal gasification operation due to the fluctuations of the process parameters and to the heterogeneous and variable composition of the fed waste;
- the specific gas production \(V_F\) is an uncertain parameter. Although its value is usually set in the plant design phase, actually \(V_F\) can present variations during plant operation caused by fluctuations of the energy content of the fed waste as a consequence of its heterogeneous and variable composition.

Due to the rather limited applications of the waste gasification process, extended time series data-sets containing \(C_D\) and \(V_F\) values collected during operation of similar plants are not available. However, different studies have recently investigated the pollutant emission from waste gasification plants (Klein 2002, Yamada et al. 2004, Porteous 2005, Arena 2012, University of California 2009).

From these studies, 35 \(C_D\) and four \(V_F\) values have been derived.

The availability of a statistically significant set of \(C_D\) average values, has allowed to represent the uncertainty on this model parameter by means of a probability distribution. A Kolmogorov–Smirnov’s test considering lognormal, Weibull, Beta, logistic PDFs has been performed in order to properly choose the PDF best representative of the literature data. Figure 2 shows the selected Beta-PDF of parameters \(\alpha = 0.36\) and \(\beta = 1.32\) and maximum value 0.07 ng m\(^{-3}\), that is 30% less than the regulatory limit of 0.1 ng m\(^{-3}\) (European Union, 2000).

Conversely, for \(V_F\) the available information (four literature values and the case study design value) is very scarce. Thus, two different alternative representations of the uncertainty affecting \(V_F\) have been considered:
a subjective triangular PDF based on analyst judgment. The PDF range has been chosen to be
[3360, 6670] (m$^3$ Mg$^{-1}$), which corresponds to the minimum and maximum values of the four
literature values, and its mode has been chosen to be 5420 m$^3$ Mg$^{-1}$, which corresponds to the
current case study design value (Figure 3);

- a triangular possibility distribution with the same range of the subjective PDF and most likely
value set equal to the case study design value (Figure 3).

Notice that, although the possibility and probability distributions used to represent the uncertainty
on $V_F$ have similar shapes, they convey very different information: the possibility distribution
summarizes a set of cumulative distributions, bounded by the so-called Necessity and Possibility
functions (see Appendix a.1), whereas the subjective PDF defines just one specific cumulative dis-
tribution of that set.

On the other hand, model predictions of the DF values in equation (1) are also uncertain quantities
affected by i) the natural variability in the input parameters (such as meteorological variables and
source features), ii) measurement errors and iii) modeling errors due to the difficulties in capturing
the atmospheric behaviors (Sax and Isakov 2003, Rao 2005). Neglecting for simplicity the model
uncertainties, it turns out that the DF values are mainly affected by the variability in the meteor-
ological input parameters, which largely exceeds the effect of the variability of flue gas temperature
and speed. In fact, the temperature is typically kept at a rather constant value during plant operation,
whereas the flue gas speed plays a negligible role on the atmospheric behavior of the emitted
plume. Here, the DF uncertainty has been estimated indirectly, by taking into account 10 different
years of local meteorological data and by separately running the dispersion model for each input da-
ta set: thus, 10 yearly DF values have been estimated for each grid node in the computation domain.

An example of the different wind conditions observed in the study area is given by the annual wind
roses reported in Figure 4 for four years out of the ten considered.

For illustration purposes, only one grid point has been considered for the uncertainty propagation,
chosen as the receptor most impacted by the plant emission ($x_M, y_M$), i.e. the grid node characterized
by the highest average of the 10 yearly DF values.

Due to the small dataset available, two alternative representations of the DF($x_M, y_M$) uncertainty
have been considered:

- a trapezoidal PDF based on analyst judgement with its minimum equal to 0, its core ranging be-
tween $5.69 \cdot 10^{-3}$ and $4.19 \cdot 10^{-2}$ fg m$^3$/ng s$^{-1}$, which corresponds to the minimum and maximum
values of the ten estimated DF values, and its maximum equal to 0.21 fg m$^3$/ng s$^{-1}$, which is the
DF obtained for the receptor point assuming the worst day conditions observed during the 10-
year period for atmospheric dispersion all year long (Figure 5);
a trapezoidal possibility distribution with the same features of the subjective PDF (Figure 5).

4. RESULTS

With respect to the uncertainty propagation, the PMC method has been applied to the case in which the uncertainty on $C_D$, $V_F$ and $DF(x_M,y_M)$ are represented by the PDFs, whereas the HMC method considers the PDF for $C_D$ and the possibility distributions for $V_F$ and $DF(x_M,y_M)$. The number $m$ of MC realizations of the parameters $C_D$, $V_F$ and $DF(x_M,y_M)$ used in the PMC method has been set to 1000. The HMC procedure, instead, has been run considering $m = 1000$ realizations of the parameter $C_D$ and, for each of these realizations, 21 $\alpha$-cut values (range 0-1, step 0.05) for the possibilistic variables $V_F$ and $DF(x_M,y_M)$.

The cumulative distributions of the model output variable, $C_{air}(x_M,y_M)$ obtained by applying the PMC and HMC uncertainty propagation methods are presented in Figure 6. The PMC method provides a single cumulative distribution (continuous line), whereas the HMC model provides the belief (lower dotted curve) and the plausibility (upper dashed curve) of the set $A=(-\infty,C_{air})$, which can be interpreted as the boundary cumulative distributions of $C_{air}$ at the receptor of coordinates $(x_M, y_M)$.

As expected, the cumulative distribution of the model output obtained by the PMC is within the belief and plausibility functions obtained by the hybrid approach.

Furthermore, the representation of the uncertainty on $C_{air}$ provided by the PMC method appears to be more concise and easy to be interpreted than the one provided by the HMC method. This is a consequence of the different representation of the uncertainty on the input parameters $V_F$ and $DF(x_M,y_M)$: while the HMC method considers a set of cumulative distributions, the PMC forces all the uncertainty on $V_F$ and $DF(x_M,y_M)$ to be represented by single distributions. However, given the scarcity of information available for the two parameters in this case study, the use of a specific PDF seems unjustified.

Figure 6 also shows that the HMC method allows to process separately the uncertainty on $C_D$, represented by a probability distribution, from the uncertainty on $V_F$ and $DF$, represented by possibility distributions. These two contributions are explicitly visible in the results: in fact, the uncertainty on $C_D$ affects the slope of the belief and plausibility functions, while the distance between the belief and plausibility functions reflects the lack of knowledge of the parameters $V_F$ and $DF$. On the other hand, in the PMC method the contributions of the input parameters uncertainties on the model output are merged, giving rise to the slope of the output cumulative distribution.
Within the EIA procedure, for communication purposes, it is more effective to lump all the information contained in the obtained cumulative distribution of $C_{\text{air}}$ into a single value, such as for example the $\beta$ percentile. To this aim, setting a degree of confidence $\beta = 0.95$, the PMC method provides a single value, $C_{\text{air}}(x_M,y_M)_{95^{\text{th}}} = 2.80 \cdot 10^{-1} \text{ fg}\cdot\text{m}^{-3}$, representative of the 95th percentile of the distribution of the PCDD/Fs annual average at the most impacted receptor point in the area. On the contrary, with the HMC the 95th percentile is an uncertain quantity itself, whose true value lies in the interval $[1.15 \cdot 10^{-2}, 5.65 \cdot 10^{-1}]$ fg·m$^{-3}$, i.e. the 95th percentile of the Belief and Plausibility functions. Table 2 reports the values of different percentiles for both PMC and HMC methods: PMC values are always in the range of HMC values and the distance between the plausibility and belief functions increases from lower to higher percentiles. Notice that for any percentile considered the gap between the plausibility and the belief functions is in the range of one/two orders of magnitude: the uncertainty on the VF and DF parameters appears to significantly affect the estimated ground level concentrations. In this respect, the capability of the HMC method to separately process the contribution of the different uncertainties gives rise to results that clearly show the effects of the lack of knowledge on the input parameters. This is a desirable quality leading to more informative and transparent outputs, also in the light of the subsequent calculations for health risk assessment.

It is worth noticing that the concentration estimates at the selected receptor for the most and the least favourable year for atmospheric dispersion obtained by means of the traditional deterministic approach are 0.03 fg·m$^{-3}$ and 0.24 fg·m$^{-3}$, respectively, thus varying by almost one order of magnitude. This means that the traditional deterministic approach based on just one single year simulation, as is done in the common practice, can lead to either too precautionary or non-conservative estimates based on the arbitrary choice of the analyst about the reference year to be used in the model simulation. In this respect, a preliminary assessment of the representativeness of the year chosen for the weather model simulation or multi-year modelling are recommended for deterministic approach calculations.

Finally, notice that both PMC and HMC methods have been applied under the simplifying assumption of independence of the parameters $C_D$ and $V_F$, although they are expected to be somehow correlated. Future developments of this work may thus include investigating the effects of the dependence between these parameters on the uncertainty propagation results, for example resorting to the approach suggested in (Pedroni and Zio 2012).
5. CONCLUSION

In this work the uncertainties in the first steps for the atmospheric emission and dispersion estimation of an Environmental Impact Assessment (EIA) procedure have been treated by both a Hybrid probabilistic and possibilistic Monte Carlo method (HMC) and a pure Probabilistic Monte-Carlo method (PMC). Calculations have been performed for the case study of the project of a new waste gasification plant, focusing on dioxin/furan emissions that are a primary concern for human health due to their carcinogenic potential. The expected impact of the plant emissions has been quantified in terms of the contribution to the annual average concentration of dioxins and furans in ambient air at the most impacted receptor in the neighbourhood of the plant. The comparison between the results obtained by the HMC and PMC methods points out that the former is more effective in propagating the input uncertainties through the model when little information is available for some input parameters. In particular, i) the information provided by the output of the HMC method is more consistent with that (little) available for the input parameters and ii) uncertainty is processed more “transparently” than with PMC methods, avoiding arbitrary and subjective assumptions by the analyst on the input probability distribution functions and separating the contributions to the output uncertainty due to the probabilistic and possibilistic input parameters.

This research has been conceived as a preliminary study to assess the potential for the implementation of HMC methods to the complete EIA procedure. In this regard, the satisfactory outcomes of this first analysis foster future works towards the extension of the HMC method to the remaining steps of an EIA procedure, i.e. the evaluation of the human exposure through the different impact pathways and the subsequent risk assessment.

Finally, it is worth noticing that the principal users of the EIA findings are decision-makers often responsible for communicating the results to the population living in the area of the planned installation. With respect to environmental and health-related issues, there is in general social aversion to accept information expressed in terms of probability. Therefore, future studies will also have to investigate the way of post-processing the results of a whole EIA procedure developed within an hybrid probabilistic-possibilistic framework to make them communicable to the stakeholders and to make them more easily understandable. To this purpose, the abilities of the HMC method to avoid arbitrary assumptions and to provide results that explicitly report both probabilistic and possibilistic uncertainties could be useful for a transparent and clear post-process.

APPENDIX

A.1. BASIC OF POSSIBILITY THEORY
In possibility theory, uncertainty is represented by a possibility function \( \pi(y) \). For each \( y \) in a set \( \Omega \), \( \pi(y) \) expresses the degree of possibility of \( y \). When \( \pi(y) = 0 \) for some \( y \), it means that the outcome \( y \) is considered an impossible situation. When \( \pi(y) = 1 \) for some \( y \), it means that the outcome \( y \) is possible, i.e., is just unsurprising, normal, usual (Dubois 2006). This is a much weaker statement than when probability is 1.

The possibility function gives rise to probability bounds, upper and lower probabilities, referred to as necessity and possibility measures \((N, \Pi)\). The possibility of an event \( A \), \( \Pi(A) \), is defined by

\[
\Pi(A) = \sup_{y \in A} \{\pi(y)\}
\]  

(a.1)

and the necessity measure \( N(A) \) is defined by:

\[
N(A) = 1 - \Pi(\neg A) = 1 - \sup_{y \in \neg A} \{\pi(y)\}
\]

9

Let \( \mathcal{P}(\pi) \) be a family of probability distributions such that for all events \( A \), \( N(A) \leq P(A) \leq \Pi(A) \).
Then,

\[
N(A) = \inf P(A) \text{ and } \Pi(A) = \sup P(A)
\]

(a.2)

where \( \inf \) and \( \sup \) are with respect to all probability measures in \( \mathcal{P} \). Hence the necessity measure is interpreted as a lower level for the probability and the possibility measure is interpreted as an upper limit. Referring to subjective probabilities, the bounds reflect that the analyst is not able or willing to precisely assign his/her probability, and the bounds are the best he/she can do given the information available; in other words, he or she can only describe a subset of \( \mathcal{P} \) which contains his/her probability (Dubois 2006).

**A.2. PROBABILISTIC-POSSIBILISTIC UNCERTAINTY REPRESENTATION AND PROPAGATION THROUGH A MODEL**

Let us consider a model whose output \( Z = f(Y_1, Y_2, \ldots, Y_n) \) is a function of \( n \) uncertain variables \( Y_i, i = 1, \ldots, n \). The uncertainties of the first \( k \) variables can be represented by probability distributions \( p_y(y) \) properly derived from data, whereas for the remaining \( n - k \) the scarcity of information provides a weaker statistical base for a specific probability assignment, making the construction of PDFs a critical operation. Two different approaches are here presented to represent and propagate such uncertainties.

**a. Pure probabilistic Monte-Carlo method (PMC)**

The uncertainties of all \( n \) input variables are treated in terms of probability distributions. Probability representations of the \( n - k \) variables can be assigned i.e.: based on expert judgment. For example, let us considered a typical case in which the available information on a model parameter \( x \) is only that its values are located somewhere between a value \( x_{\min} \) and a value \( x_{\max} \). In this case, a uniform probability distribution \( p(x) = 1/(x_{\max} - x_{\min}), \forall x \in [x_{\min}, x_{\max}] \) is typically assumed to represent the uncertainty on \( x \). This approach appeals to: i) Laplace principle of insufficient reason according to which all that is equally plausible is equally probable and to ii) the maximum entropy approach (Baraldi et al. 2010). However, doubts on the consistency of this uncertainty representation have been raised (Baudrit et al. 2006) as it seems that the insufficient knowledge may justify choices of specific functional probability distributions, like the uniform, but it should somehow account for the full set of possible probability distributions on \( X = [x_{\min}, x_{\max}] \), so that the probability of value \( x \in X \) is allowed to take any values in \([0,1]\) (Baraldi et al. 2010).
The PMC method then applies a single loop Monte Carlo simulation to propagate the uncertainties of the $n$ variables through the model. The operative steps of the propagation procedure are the following:

1. Sample the $i$-th realization $(y_1', \ldots, y_n')$ of the uncertain variables $(y_1, \ldots, y_n)$ from their respective PDFs $p_{y_i}(y_i), \ldots, p_{y_n}(y_n)$.

2. Compute the output of the model corresponding to the $i$-th realization of the uncertain variables: $Z_i = f(y_1', \ldots, y_n')$.

3. Repeat steps 1 and 2 for many times. Then, derive the cumulative distribution $F(Z)$ from the outputs $Z_i, i = 1, \ldots, m$.

The cumulative distribution $F(Z)$ obtained gives information on the output uncertainty. Setting a percentile $\beta$, that corresponds to the desired degree of confidence, a concise output value of $Z$ can be provided. The output realizations $Z_i, i = 1, \ldots, m$, found in step 2 can be directly used to compute other lumped indicators of the output distribution, such as its mean and standard deviation values.

### b. Hybrid probabilistic-possibilistic Monte-Carlo method (HMC)

In the Hybrid probabilistic-possibilistic method, the uncertainties of the $n - k$ variables are represented in terms of possibility distributions $(\pi^{y_1}, \ldots, \pi^{y_n})$.

Different methods have been developed to derive possibility distributions from the available information (Baudrit & Dubois 2006, Dubois 1993). Often the analyst only knows that an uncertain variable can take values in a given range $[a, b]$ wherein the most likely value is $c$. Triangular possibility function with the range $[a, b]$ taken as base and $c$ taken as vertex can be typically used to describe this information. It has been shown that the family of probability distributions defined by such possibility distribution with contains all the probability distributions with support $[a, b]$ and mode $c$ (Baudrit & Dubois 2006).

In the HMC method, the propagation of the uncertainties expressed in terms of both PDFs and possibility curves is performed by combining the Monte Carlo technique with fuzzy interval analysis (Baudrit et al. 2006). The operative steps of the propagation procedure are the following:

1. Sample the $i$-th realization $(y_1', \ldots, y_i')$ of the random variable vector $(y_1, \ldots, y_i)$.

2. Select a possibility value $\alpha \in [0,1]$ and the corresponding $\alpha$-cuts of the possibility distributions $(\pi^{y_1}, \ldots, \pi^{y_i})$, i.e. the intervals of values of the possibilistic variables $(y_1', \ldots, y_i')$ with associated possibility distributions greater or equal to $\alpha$.

3. Calculate the smallest and largest values of $f(y_1', \ldots, y_i', y_{i+1}, \ldots, y_n)$, denoted by $\underline{f_a}$ and $\overline{f_a}$ respectively, considering the fixed values $(y_1', \ldots, y_i')$ sampled in 1. for the random variables $(y_1, \ldots, y_i)$ and all values of the possibilistic variables $(y_{i+1}, \ldots, y_n)$ in the $\alpha$-cuts of their possibility distributions $(\pi^{y_{i+1}}, \ldots, \pi^{y_n})$ found in 2. Then, consider the extreme values $\underline{f_a}$ and $\overline{f_a}$ found in 3. as the lower and upper limit of the $\alpha$-cut of $f(y_1', \ldots, y_i', y_{i+1}, \ldots, y_n)$.

4. Return to step 2. and repeat for another $\alpha$-cut; the fuzzy random realization (fuzzy interval) $\pi'$ of $Z = f(y)$ is obtained as the collection of the values $\underline{f_a}$ and $\overline{f_a}$ for each $\alpha$-cut.

5. Return to step 1. to generate a new realization of the random variables.

The procedure is repeated for $i = 1, \ldots, m$; at the end of the procedure $m$ realizations of fuzzy intervals are obtained, i.e.: $(\pi_1', \ldots, \pi_m')$. 
For each set $A$ contained in the universe of discourse $U_z$ of the output variable $Z$, it is possible to obtain the possibility measure $\Pi_i^f(A)$ and the necessity measure $N_i^f(A)$ from the corresponding possibility distribution $\pi_i^f(z)$, by:

$$\Pi_i^f(A) = \max_{z \in A} \{ \pi_i^f(z) \}$$  
(a.3)

$$N_i^f(A) = \inf_{z \in A} \left\{ 1 - \pi_i^f(z) \right\} = 1 - \Pi_i^f(A)$$  
$$\forall A \subseteq U_z$$  
(a.4)

Finally, the $m$ different realizations of possibility and necessity measures can be combined to obtain the belief $Bel(A)$ and the plausibility $Pl(A)$ for any set $A$, respectively (Baudrit et al. 2006):

$$Bel(A) = \frac{\sum_{i=1}^{m} N_i^f(A)}{m}$$  
(a.5)

$$Pl(A) = \frac{\sum_{i=1}^{m} \Pi_i^f(A)}{m}$$  
(a.6)

For each set $A$, this technique computes the belief and plausibility as the average of the possibility measures associated with each output fuzzy interval.

The likelihood of the value $f(Y)$ passing a given threshold $z$ can then be computed by considering the belief and the plausibility of the set $A = (-\infty, z]$; in this respect, $Bel(f(Y) \in (-\infty, z])$ and $Pl(f(Y) \in (-\infty, z])$ can be interpreted as bounding, average cumulative distributions $F(z) = Bel(f(Y) \in (-\infty, z])$, $\overline{F}(z) = Pl(f(Y) \in (-\infty, z])$ (Baudrit et al. 2006). Thus, one way to estimate the total uncertainty on $f(Y)$ is to provide a confidence interval at a given level of confidence, taking the lower and upper bounds from $Pl(f(Y) \in (-\infty, z])$ and $Bel(f(Y) \in (-\infty, z])$, respectively (Baudrit et al. 2006). Notice, however, that it is not possible to directly obtain information on the mean and standard deviation of the output distribution from the plausibility and belief distributions.

REFERENCES


