Modeling incidental sequences in high environmental risk industrial plants: some simulation experiments

Matilde Trevisani, Antonio Capelo

Modeling incidental sequences in high environmental risk industrial plants: some simulation experiments

Matilde Trevisani†, Antonio Capelo‡

† Department of Statistical Sciences, University of Padova

‡ Department of Applied Statistics and Economics, University of Pavia

Abstract

Motivating problem of this work is consequence assessment of incidents in industries of relevant environmental risk. An incident consists generally of a chain of adverse events whose effects possibly produce serious damages to people, structures and environment. For consequence assessment mathematical models are utilized to simulate such complex processes. A feature of computer models like these is that they are deterministic; nevertheless they are affected by various sources of uncertainty. Objective of this work is planning computer experiments that produce a probabilistic evaluation of incidental effects. In particular, we consider the uncertainty associated with input variability in the models of interest. This specific source of variability is only rarely considered by risk analysts. However, it appears essential in assessing the quality of results especially in analysis of *domino* effects - consisting in causal connection of incidental subsequent occurrences - with serious error propagation.

In this paper, uncertainty analysis, primarily, and sensitivity analysis as well as optimization techniques, more incidentally, will be discussed in this relatively new context of computer models chains. Further, distinction between system-intrinsic

1

uncertainty and analyst subjective uncertainty, recovered from Helton extensive work (1997), is maintained. An efficient computational strategy of uncertainty analysis is outlined and, finally, applied to an illustrative example of a two-consequence model incidental sequence.

Key words: incidental sequence; computer experiment; domino effect; Monte Carlo simulation; importance sampling; Latin hypercube sampling; error propagation; uncertainty analysis; sensitivity analysis; optimization; subjective uncertainty; objective uncertainty; endogenous input.

1 Introduction

Motivating problem of this work is consequence assessment of incidental sequences in industries of relevant environmental risk. With relevant risk we mean the possibility of a negative uncertain occurrence which can produce serious damages to people, structures and environment. An incidental sequence is a chain of adverse events that typically starts from an uncontrolled release of dangerous substances or a short circuit, and, possibly, continues with other events such as fires, explosions or dispersions. In this context, *domino effect* analysis is carried out for evaluation of final incidental risk by concatenation of subsequent incidental chains.

Mathematical models are used to simulate these physical processes and derive damage outputs. Although these models are deterministic, consequence assessment is affected by several sources of uncertainty. Main object of the work is to carry out an uncertainty analysis specifically to quantify the effect of model input uncertainty on model output. Before dealing with the

2

tasks of this study, a brief introduction to the general statistical problems of *computer experiments* is given.

Scientific phenomena for which physical experimentation is too time- or cost-consuming or simply impossible, are investigated by *computer models*, *i.e.*, mathematical models for simulating these complex systems. Then, for computer experiment we mean a number of runs of the model code with various inputs. The following lines synthetically list the major features of this modeling area.

Our computer models are *black box*-like and deterministic (see table 1), *i.e.*, replicate observations from running the code with the same inputs will be identical. Further, they are typically characterized by mathematical complexity; high-dimensional and correlated inputs; multivariate output that is commonly a time-dependent function with presence of discontinuities; intensive computation. In summary, structural complexity compensates for the lack of random error. In more general terms, the issue of systematic error of a possible fitted response-curve, together with input uncertainty, substitutes for the traditional issue of random variability in experimental units.

From this premise, it is clear that non-standard statistical techniques are needed to solve the typical issues concerning computer experiments. In particular, given a general consequence model, the first step of a statistical analysis is the selection of an efficient experimental design, *i.e.*, the search for the minimum number of input configurations to obtain analysis results with the desired level of accuracy. To accomplish this task, the analyzer has to specify appropriately the joint distribution of input variables from which to generate the designed sample. Therefore, he has to choose the sampling

3

Table 1: A general consequence model

INPUT		OUTPUT
$\downarrow$		$\downarrow$
substances c. $\searrow$		$\nearrow$ measures
process c. $\rightarrow$	**deterministic**	$\rightarrow$ time
plant c. $\rightarrow$	**MODEL**	$\searrow$ damage
meteo c. $\nearrow$	$\uparrow$	$\uparrow$
$\uparrow$	**BLACK BOX!**	**random!**
fixed &	physic-chem.modeling	
random!	experimental param.	

method following various criteria as ease of implementation, flexibility, precision degree of the uncertainty estimation and adaptability to sensitivity analysis. Finally, he has to summarize in an opportune way the random output resulting from the input randomization. In the literature two main ways alternate: Monte Carlo integration or *meta model* fitting.

There are several analysis that a statistician may be required to face with by a computer experiment. *Validation* is aimed at assessing if the computer model is an adequate representation of the corresponding system in the real world. In the *pilot* phase of a simulation study, *screening* is the search for the few really important factors that controls most of the variability of the output (*Occam's razor* principle). *Sensitivity Analysis* (SA) is aimed at evaluating the contributions of individual influential inputs to the total uncertainty in the output. *Uncertainty Analysis* (UA) is aimed at quantifying

4

output uncertainty that derives from the total input uncertainty. *Optimization* (O) is the search for input regions that are transformed in maximum output.

In this paper, uncertainty analysis, primarily, and sensitivity analysis as well as optimization technique, but more incidentally, will be discussed in this relatively new context of computer models chains. Further, distinction between system-intrinsic uncertainty and analyst subjective uncertainty, recovered from Helton extensive work (1997, [1]), is maintained. An efficient computational strategy of uncertainty analysis is outlined and, finally, applied to an illustrative example of a two-consequence model incidental sequence.

A code in R language, executable with the free-software R (which can be downloaded at *http://cran.r-project.org/*), has been implemented to link together the statistical analysis with the deterministic one in a flexible way. The statistical utility accounts for the input sampling phase as well as for the output probabilistic analysis, leaving unaltered incidental C-codes (Carpignano, 1995, [2]) to elaborate the central deterministic part. Lastly, graphical summaries of output distributions are immediate tools for conveying the added value of this approach to uncertainty analysis.

This work is included in the I.T.E.R.E (Experts of Innovative Techniques for Environmental Risk) project joined by the Dipartimento di Ingegneria Chimica, Processi and Materiali of the University of Palermo, the Joint Research Center of Ispra (Va), the Dipartimento di Scienze Statistiche of the University of Padua, ARTES, the Assessorato dell'Agricoltura e Foreste of Sicily Region, the 2i map sud Consortium and financed by UE, the Ministero

del Lavoro e Previdenza Sociale and Sicily Region.

2 Formal Structure of UA

First object of this work is carrying out an appropriate UA in order to assess model output uncertainty that derives from model input uncertainty. In a wide current of computer models literature, a distinction is used to keep between *objective uncertainty* (otherwise said *stochastic* (Helton, 1997) or *simulation* uncertainty (Cheng and Holland, 1995, [3])) and *subjective uncertainty* (so called by Helton, otherwise said *parameter uncertainty*). The first one is system-intrinsic so as to be irreducible. In general it derives from oscillation of some input variables around their nominal values (or parameters) which are supposed to be definitely set in deterministic simulations. The other characterizes the degree of belief which is assigned to the analyst hypothesis. It generally depends on an imprecise knowledge of some parameters so as to be reducible as the state of knowledge improves.

Incidentally, this distinction does not match that one dividing input variables into noise factors and control factors. In fact, here, imprecisely-known parameters are a larger class of that of inputs which are set in a discretionary way by the analyst. However, sensitivity analysis concerns especially with subjective uncertain parameters within control factors are included as well

In general terms, UA is an aleatory transformation of the probability space, $(\mathbf{X}, A, P)$, characterizing input vector $\mathbf{x}$, via *black box* $f(\mathbf{x})$. Then, a complementary cumulative distribution function (*ccdf*), or, in equivalent way, a cumulative distribution function (*cdf*) or a probability density function

(pdf), is used to represent the uncertainty in f. However, after distinguishing the two types of input uncertainty, (X, A, P) may be interpreted as the product space built up from the probability space for irreducible noises, $(\mathbf{X}_{obj}, A_{obj}, P_{obj})$ and the probability space for imprecisely-known parameters, $(\mathbf{X}_{subj}, A_{subj}, P_{subj})$. Hence, UA turns into a family of aleatory variables, $\{t[f(\mathbf{X}_{obj}|\mathbf{x}_{subj})]\}_{\mathbf{X}_{subj}}$, arising from $(\mathbf{X}_{subj}, A_{subj}, P_{subj})$, where $t(\cdot)$ is an opportune function of the output value. Otherwise, UA can be interpreted as an aleatory transformation

$$\mathbf{X}_{subj} \longmapsto t[f(\mathbf{X}_{obj}|\mathbf{x}_{subj})]$$

defined on $(\mathbf{X}_{subj}, A_{subj}, P_{subj})$ and mapped onto the *probabilistic functional space* of aleatory functions $t[f(\mathbf{X}_{obj}|\mathbf{x}_{subj})]$ which are, in turn, defined on the conditional space $(\mathbf{X}_{obj}(\mathbf{x}_{subj}), A_{obj}(\mathbf{x}_{subj}), P_{obj}(\mathbf{x}_{subj}))$.

This formal structure synthesizes the UA framework of Helton extensive paper (1997) which outlines a sort of two-stage analysis. The first step involves the determination of the *ccdf* each one representing the system-intrinsic variability conditional to a specific parameter choice. This UA stage allows the analyst to explore the effect of varying subjective opinion on output uncertainty. A SA results by comparing either the total *ccdf*, associated to different $\mathbf{x}_{subj}$, or single synthetic values as $E(f|\mathbf{x}_{subj})$, $Q^{95^{o}}(f|x_{subj})$, *et caetera*. The second step summarizes the distribution of *ccdf* (as each *ccdf* has in concept a probability of zero) by, for instance, the distribution of the exceedance probabilities, $p(f > D)$, conditional to a certain damage output D, that derives from subjective uncertainty, or the expected value of this distribution, $\bar{p}(f > D)$. To accomplish these two steps some integration

7

problems need to be solved. They consist in

$$prob(f > D|\mathbf{x}_{subj}) = 1 - \int_{S_{obj}} \mathbf{1}[f(\mathbf{x}) \le D]\pi(\mathbf{x}_{obj}|\mathbf{x}_{subj})d\mathbf{x}_{obj}, \qquad (1)$$

for determining the *ccdf* of f associated with a specific $\mathbf{x}_{subj}$,

$$prob(p \le P|D) =$$

$$= \int_{S_{subj}} \mathbf{1}\left[\left(1 - \int_{S_{obj}} \mathbf{1}[f(\mathbf{x}) \le D]\pi(\mathbf{x}_{obj}|\mathbf{x}_{subj})d\mathbf{x}_{obj}\right) \le P\right]\pi(\mathbf{x}_{subj})d\mathbf{x}_{subj},$$

$$\qquad (2)$$

for determining the *cdf* of $p(f > D)$ conditional to a certain D, and

$$\bar{p}(f > D) = \int_{S_{subj}}\left[1 - \int_{S_{obj}} \mathbf{1}[f(\mathbf{x}) \le D]\pi(\mathbf{x}_{obj}|\mathbf{x}_{subj})d\mathbf{x}_{obj}\right]\pi(\mathbf{x}_{subj})d\mathbf{x}_{subj},$$

$$\qquad (3)$$

for, lastly, determining the marginal *ccdf* of f. As for notation, $\mathbf{1}[\cdot]$ is the usual indicator function and $\pi(\mathbf{z})$ denotes the *pdf* for the probability space of $\mathbf{Z}$.

Finally, the choice of *ccdf* in (1) as well as that of *cdf* in (2) reflects the interest in knowing how likely is to have a consequence worse than a certain damage or the aim of having low probabilities associated to high damages. Then, this preference is right suitable for incidental sequence issues.

3   An efficient strategy for UA of multimodel chains

The work outlines an efficient strategy to achieve UA goals (1), (2) and (3). It also involves, as by-products, some results of SA and O study. Besides it tackles the problem of endogenous input transmission in multimodel chains. The premise is that integrals like (1), (2) and (3) have to be approximated as their exact solution is generally precluded in real complex problems. Then,

8

MC method is the favored technique for approximate inference. Here, a particular procedure of MC sampling is worked out to produce an effective simulation design.

The plan requires basically the sampling from the *pdf*, $h(\mathbf{x})$, of the overall input probability space $(\mathbf{X}, A, P)$. Usually first chain model inputs are supposed to be statistically independent. Besides, the distribution form needs to be carefully specified, thus exploiting expert advice.

Directly, this basic sampling provides a MC estimate of (3). Yet, moreover, $h(\mathbf{x})$ is assumed to be the *importance sampling density* (*isd*) in the MC integration for (1). In fact, (1) will be rewritten as

$$I(\mathbf{x}_{subj}) = \int_{S_{obj}} t[f(\mathbf{x})] \frac{\pi(\mathbf{x}_{obj}|\mathbf{x}_{subj})}{h(\mathbf{x})} h(\mathbf{x}) d\mathbf{x}_{obj}. \tag{4}$$

A generic function $t(\cdot)$ of the output substitutes $\mathbf{1}[\cdot]$ in the original form (1) to stress that any interesting statistic of f distribution results typically from an integration.

For *importance sampling* integration (4) being effective, $\mathbf{x}_{obj}$ has to be statistically dependent on $\mathbf{x}_{subj}$ so that $\pi(\mathbf{x}_{obj}|\mathbf{x}_{subj})$ is varying with $\mathbf{x}_{subj}$. This is not fulfilled in hypothesis of stochastic independence between $\mathbf{x}_{obj}$ and $\mathbf{x}_{subj}$, as typically holds. Hence, $\mathbf{x}_{subj}$ input has to be in general interpreted both as uncertain parameter and as system variable randomly oscillating around its characterizing value. Section 4 will clarify how this condition naturally arises.

This expedient *importance sampling* involves several simulation gains.

It suffices only one loop of MC sampling in order to obtain the entire distribution of *ccdf* that arises from subjective uncertainty. MC estimate of

9

(4),

$$\hat{I}(\mathbf{x}_{subj}) = \frac{1}{N}\sum_{i=1}^{N} t[f(\mathbf{x}_i)]\frac{\pi(\mathbf{x}_i|\mathbf{x}_{subj})}{h(\mathbf{x}_i)} = \frac{1}{N}\sum_{i=1}^{N} t[f(\mathbf{x}_i)]w_i(\mathbf{x}_{subj}), \quad (5)$$

needs only to weight the same $t[f(\mathbf{x}_i)]$, for $i = 1, \ldots, N$, sample with varying coefficients $w_i(\mathbf{x}_{subj})$. The same applies for computing the inner integration in (2).

Then, it traditionally implies a possible *Variance reduction* of the MC estimate $\hat{I}(\mathbf{x}_{subj})$. Finally, it will yield as a by-product the gradient estimate

$$\frac{\partial \hat{I}(\mathbf{x}_{subj})}{\partial \mathbf{x}_{subj}} = \frac{1}{N}\sum_{i=1}^{N} t[f(\mathbf{x}_i)]\frac{\partial W_i(\mathbf{x}_{subj})}{\partial \mathbf{x}_{subj}}, \quad (6)$$

again without repeating the MC loop as only the w_i depend on $\mathbf{x}_{subj}$. This last result recovers the so-called *Parametric sampling* technique of Singhal and Pinel (1981, [4]).

Model chains have a further issue: transmission of output from a given model as input to a subsequent model. It will be called *endogenous input*. Two alternative solutions are considered in this setting.

The first one consists in making a *Latin hypercube sampling* (LHS) to simulate an endogenous input. That is, given the empirical distribution function of the output variable from some previous model, its range will be divided in, say, l nonoverlapping sets of equal marginal probability. Then, a sample of N/l dimension will be obtained from each set. It is worth noting that if ≥ 2 endogenous inputs are output of the same incidental model, they need to be correlated. In fact, they result from one set of deterministic relationships constituting the model box. If this is the case, input dependencies have to be incorporated: given a target rank correlation matrix (available from the

10

previous model), an appropriate pairing procedure of input variables will be performed (for instance, the *restricted pairing* procedure of Iman and Conover, 1982).

The alternative solution reduces the uncertainty flow that arises in linking models to single values. It drives the transmission choosing certain summary statistics (expected values, 95° percentiles, *et caetera*) of endogenous inputs, possibly conditional to sensitivity parameters of out-model.
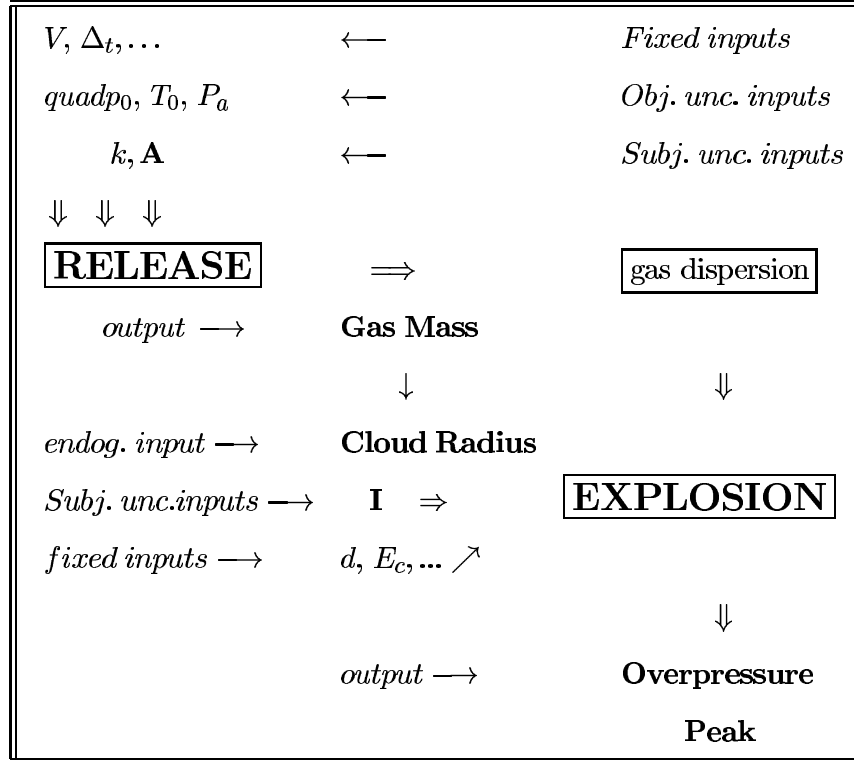
4   An illustrating example: a two model chain

A two model incidental sequence has been selected to describe the previous discussion on a UA effective strategy. The composition of a model for *gas release from vessel* with a model for *unconfined vapour cloud explosion* (UVCE) is synthetically illustrated in table 2.

Release model In *gas release from vessel* model input variables have been classified in certain or uncertain data on the basis of engineers expert opinion. In particular, fixed inputs are *gas code* (here, methane), thermodynamic hypothesis (here, adiabatic hypothesis), physical hypothesis (here, equation of perfect gas), internal vessel free volume ($V = 40\,m^3$) and, finally, time interval by which the release period is discretized for computing time-varying variables ($\Delta_t = 0.1\,s$). Aleatory inputs are internal pressure (p_0), internal temperature (T_0) and external pressure (p_a). They are assigned normal distributions with means $0.5MPa$, $360K$, and, respectively, $0.1MPa$, and standard deviations $0.1MPa$, $15K$, and, respectively, $0.05MPa$. Sensitivity inputs are the isentropic constant (k) and the area

11

Table 2: A two model chain

$V, \Delta_t, \ldots$	$\longleftarrow$	*Fixed inputs*
$quadp_0, T_0, P_a$	$\longleftarrow$	*Obj. unc. inputs*
$k, \mathbf{A}$	$\longleftarrow$	*Subj. unc. inputs*
$\Downarrow \ \Downarrow \ \Downarrow$		
$\boxed{\textbf{RELEASE}}$	$\Longrightarrow$	$\boxed{\text{gas dispersion}}$
output $\longrightarrow$	**Gas Mass**	
	$\downarrow$	$\Downarrow$
endog. input $\longrightarrow$	**Cloud Radius**	
Subj. unc.inputs $\longrightarrow$	$\mathbf{I} \ \Rightarrow$	$\boxed{\textbf{EXPLOSION}}$
fixed inputs $\longrightarrow$	$d, E_c, \ldots \nearrow$	
		$\Downarrow$
	output $\longrightarrow$	**Overpressure**
		Peak

of out-flow opening (A). According to experts opinion, k is a parameter whose value may be, in theory, exactly determined. Yet, in practice, it is imprecisely-known so that it is assigned a uniform distribution around its nominal value with extremes 1.1 and 1.2. Besides, model analysts are particularly concerned with the effects of an *irregular* variable such as A on output previsions. In fact, three differing categories can be associated to A: small (here, $0.05\,m^2$) as its most likely value, medium (here, $0.1\,m^2$) with probability, say, between 1/100 and 1/10, large (here, $0.30\,m^2$) with infinitesimal probability. Then, A is assigned a lognormal distribution that meets such requirements (figure 1 shows the density of a lognormal with $\log\mu = \log(0.054)$ and $\log\sigma = \log(1.859)$). Following the procedure which
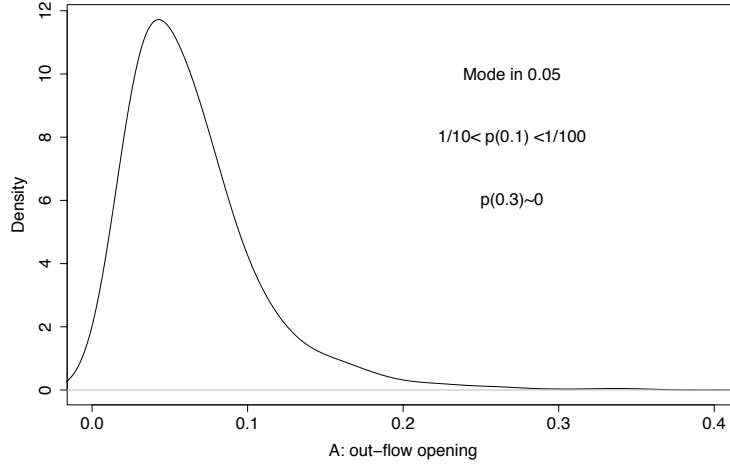

Figure 1: Lognormal density for out-flow opening A

has been described in section 3, the joint *pdf*, $h(\mathbf{x})$, is specified as product of the marginal densities for the random inputs, as the latter are assumed to be statistically independent. Moreover, $\pi(\mathbf{x}_{obj}|\mathbf{x}_{subj})$ in (4) consists of

13

the product of $N(a|\mu, \mu/2)$, representing the stochastic oscillation of A input around its nominal value μ, and the other normal distributions for the remaining noise factors. The weights in (5) and (6) are computed in the Appendix. Figure 2 shows *ccdf* distribution as results, on the basis of (4), from subjective uncertainty associated specifically with A. The graph shows, in particular, the *ccdf* related to $\mu = 0.01, 0.05, 0.1, 0.3$, the last three being from left-bottom to right-top, while the 0.01 one being amid the 0.05 and 0.1 ones. The bold line represents the mean *ccdf* computed as in (3). The vertical line set at 86.7 released mass, intersects the *ccdf* in points of height equal to $p(Release > 86.7|A = \mu)$ related to the inner integral in (2). The 86.7 number corresponds to the deterministic output computed at input nominal values (with A set equal to 0.05 mode value). It has been chosen just to point out how much more exhaustive is the information from a probabilistic evaluation of computer experiments.

Explosion model In *UVCE* model fixed inputs are distance (d) from explosion, energy of combustion (E_c) *et alia*. The only aleatory input is the endogenous input *cloud radius* (r) such as derived from released gas mass (m):

$$V_g = \frac{m}{0.7}, \quad V_c = V_g * 3, \quad V_c = \frac{4}{3}\pi r^3$$

where V_g, V_c (in m^3) refer to gas mixture volume and, respectively, cloud volume, m refers to (total) released mass of gas and r to cloud radius, whence $r = \sqrt[3]{V_n * 3/(4\pi)}$.

Sensitivity input is wave intensity (I) representing initial intensity level of explosion due to bordering space degree. It is assigned a uniform discrete
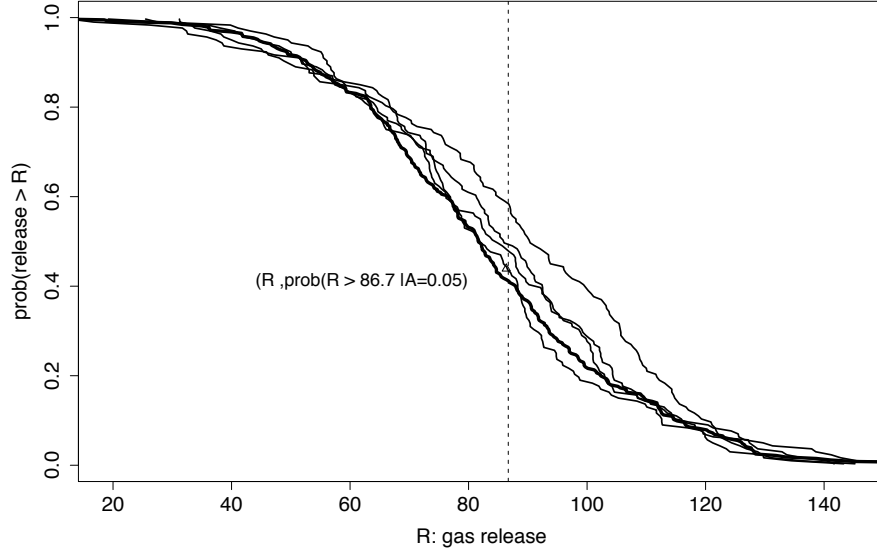
14

Figure 2: Distribution of release *ccdf*s conditional to out-flow opening A.

distribution on the integers from 1 to 10.

Before giving figure 3, we introduce a discussion on what kind of information a computer experiment may be required to provide. Indeed, computer experimenters have to account primarily for the type of questions which the *clients* are usually interested in. For instance, the users of the explosion-release computer sequence may wish to know the probability of a certain explosion degree happening, given certain initial conditions. Actually, an explosion may produce a pressure wave of differing intensity. In particular, the (static) over-pressure peak (P_s) output may be roughly classified into four categories: (1) $P_s < 0.03$ does provoke irrelevant damages; (2) $0.03 \leq P_s \leq 0.3$ does modest damages; (3) $0.3 \leq P_s \leq 0.7$ does very destructive damages; (4) $P_s \geq 0.7$ does disastrous damages. Then, examples

15

of questions are: (i) What is the probability distribution *in average* of P_s, given certain release and UVCE conditions?; (ii) What is its probability distribution conditionally to differing choices of sensitivity parameters?; (iii) What is the highest risk of an explosion of a certain gravity happening?; (iv) Or, what is the lowest risk? The interest is more focused on (iii) or, alternatively, (iv), depending on the users' *risk attitude*: risk aversion in one case, or, respectively, risk seeking in the other. Moreover, these last two questions can arise in a context of *domino effect* or optimization analysis.

Figure 3 answers to (i) providing the mean *ccdf* (the bold one) as well as to (ii) distinguishing the *ccdf* with respect to the sensitivity parameter I. More in detail, the *ccdf* move from left to right as I increases and are actually coincident for $I = 8, 9, 10$. The *ccdf* for $I = 1, 2$ (the vertical line at $P_s = 1$ and, respectively, the two-step curve with median approximately at $P_s = 2$) need probably a model check because of their differing behavior from the rest. Figure 3 also offers a qualitative answer to (iii) and (iv).

For instance, consider the critical $P_s = 0.7$ value. Then, the *ccdf* (alternatively, the *cdf* as in (2)) of the $prob(pressure > 0.7)$ - the intersection points with the *ccdf* on the vertical line at 0.7 of figure 2 - can be obtained. Figure 4 just provides the *ccdf* of the $prob(pressure > 0.7)$ variable with respect to subjective probability space of the explosion model, thus replying to (iv). For instance, the probability that $prob(pressure > 0.7) \geq 0.5$ is 0.5.

The imported values of gas mass input can eventually be certain synthetic values of the empirical distribution. In table 3, expected values or 95° percentiles of gas mass, conditional to $A = 0.05$, have been chosen for comparison with deterministic output. This last has been evaluated at mass= 86.7
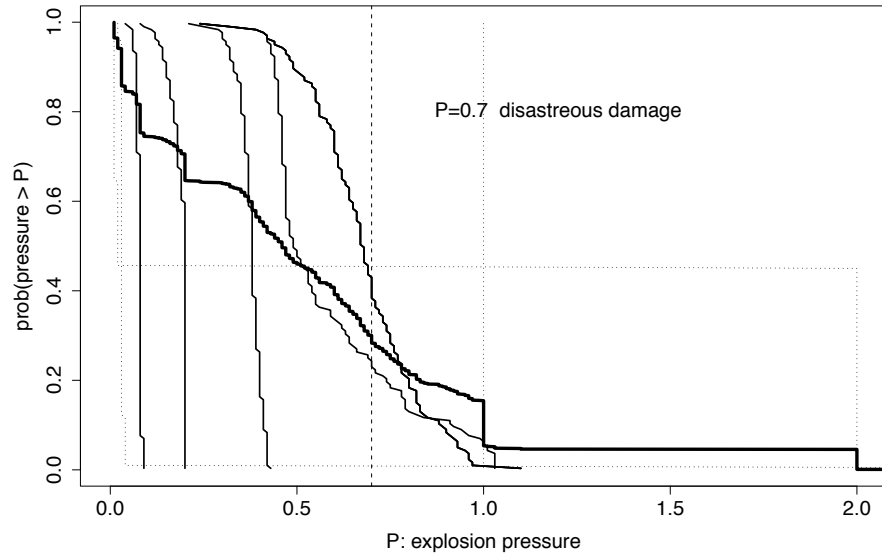
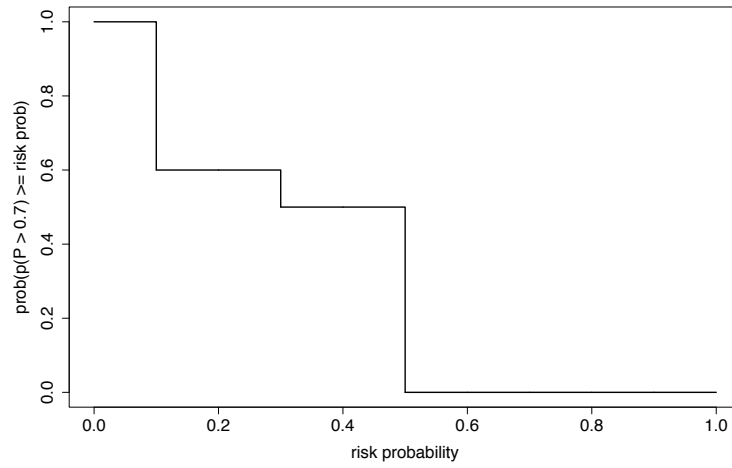Figure 3: Distribution of pressure *ccdf*s conditional to initial intensity I.

Figure 4: *ccdf* of $prob(pressure > 0.7)$ conditional to the predetermined 0.7 pressure output.

17

as results from the release model when it is exercised in a deterministic manner. Along this way, if the worst consequences are of greatest interest, distribution of 95° percentiles of gas mass, $Q^{95^o}(A)$, can be studied with respect to subjective uncertainty associated with A. So proceeding, a sort of stochastic process of the worst effects can be explored.

	Deterministic	mean($A = 0.05$)	$Q^{95^o}(A = 0.05)$
I=1	NA	NA	NA
I=2	0.01	0.01	0.02
I=3	0.03	0.03	0.04
I=4	0.08	0.08	0.09
I=5	0.2	0.2	0.2
I=6	0.38	0.38	0.42
I=7	0.54	0.49	0.99
I=8	0.7	0.68	0.92
I=9	0.7	0.68	0.92
I=10	0.7	0.68	0.92

Table 3: Comparison between deterministic and stochastic outputs

5   Appendix

In this Appendix the basic steps of the *importance sampling strategy* that has been used in the *release* model are reported.

18

Weights w_i

In the *release* model the weights $w_i(\mu) = \pi_i(\mu)/h_i$ result as following:

$$w_i(\mu) \; = \frac{\pi(\mathbf{x}_i|\mu)}{h(\mathbf{x}_i)}$$
$$= \frac{N(a_i|\mu,\mu/2)}{LogN(a_i|m,s)}$$

where m, s stay for log *mean* and log *sd* of the lognormal distribution.
After simple computations $w_i(\mu)$ are derived as

$$\frac{2\,s}{\mu}\, a_i \cdot \exp\{-2\,(\frac{a_i}{\mu} - 1)^2 + \frac{1}{2\,s^2}\,(\log a_i - m)^2\}.$$

Derivative of the weight function $w_i(\mu)$

The gradient estimate (6) can be easily computed using

$$\frac{\partial w_i}{\partial \mu} = w_i(\mu) \cdot \left[\frac{4a_i\,(a_i - \mu) - \mu^2}{\mu^3} \right]$$

which shows that weight derivative is a decreasing function of μ.

References

[1] Helton, J.C. Uncertainty and Sensitivity Analysis in the Presence of Stochastic and Subjective Uncertainty, *Journal of Statistical Computation and Simulation*, 1997, Vol. 57, pp.3-76.

[2] *STARS/FIREXP 1.1*, Modelli per l'analisi dei termini di sorgente, Andrea Carpignano, 1995.

[3] Cheng, R.C.H. and Holland, W. Sensitivity of Computer Simulation Experiments to Errors in Imput Data, *Journal of Statistical Computation and Simulation*, 1997, Vol. 57, 219-241.

[4] Singhal, K. and Pinel, J.F. Statistical Design Centering and Tolerancing Using Parametric Sampling, *IEEE Transactions on Circuits and Systems*, Vol. Cas-28, No. 7, 1981.

[5] Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn, H.P., Design and Analysis of Computer Experiments, *Statistical Science*, 1989, Vol. 4, No. 4, 409-435.

[6] McKay, M.D., Beckman, R.J. and Conover. W.J. A Comparison of Three Methods for Selecting Values of Imput Variables in the Analysis of Output from a Computer Code, *Technometrics*, Vol. 21, No. 2, 239-245.

[7] Iman, L.R. and Helton, J.C. An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models, *Risk Analysis*, Vol. 8, No. 1, 1988, 71-90.