Methodology for global sensitivity analysis of consequence models

Gant, S.E.[a]*, Kelsey, A.[a], McNally, K.[a], Witlox, H.W.M.[b], Bilio, M.[c]

[a] Health and Safety Laboratory, Harpur Hill, Buxton, SK17 9JN, UK

[b] DNV Software, Palace House, 3 Cathedral Street, London SE1 9DE, UK

[c] Health and Safety Executive, Redgrave Court, Merton Road, Bootle, L20 7HS, UK

* Corresponding author

simon.gant@hsl.gsi.gov.uk

Tel: +44(0)1298 218134

Fax: +44(0)1298 218840

adrian.kelsey@hsl.gsi.gov.uk

kevin.mcnally@hsl.gsi.gov.uk

henk.witlox@dnv.com

mike.bilio@hse.gsi.gov.uk

Research Highlights

- A methodology for global sensitivity analysis of consequence models is presented using a statistical emulator
- The methodology is demonstrated on the Phast consequence model for steady-state discharges of high-pressure carbon dioxide
- Dispersion model input parameters that have a significant effect on the extent of the plume are identified
- The study demonstrates that Bayesian analysis of model sensitivity can be conducted quickly and easily
- There is the potential for this to become a routine part of consequence modelling

Abstract

A methodology is presented for global sensitivity analysis of consequence models used in process safety applications. It involves running a consequence model around a hundred times and using the results to construct a statistical emulator, which is essentially a sophisticated curve fit to the data. The emulator is then used to undertake the sensitivity analysis and identify which input parameters (e.g. operating temperature and pressure, wind speed) have a

significant effect on the chosen output (e.g. vapour cloud size). Performing the sensitivity analysis using the emulator rather than the consequence model itself leads to significant savings in computing time.

To demonstrate the methodology, a global sensitivity analysis is performed on the Phast consequence model for discharge and dispersion. The scenarios studied consist of above-ground, horizontal, steady-state discharges of dense-phase carbon dioxide (CO_2), with orifices ranging in diameter from ½ to 2 inch and the liquid CO_2 stagnation conditions maintained at between 100 and 150 bar. These scenarios are relevant in scale to leaks from large diameter above-ground pipes or vessels.

Seven model input parameters are varied: the vessel temperature and pressure, orifice size, wind speed, humidity, ground surface roughness and height of the release. The input parameters that have a dominant effect on the dispersion distance of the CO_2 cloud are identified, both in terms of their direct effect on the dispersion distance and their indirect effect, through interactions with other varying input parameters.

The analysis, including the Phast simulations, runs on a standard office laptop computer in less than 30 minutes. Tests are performed to confirm that a hundred Phast runs are sufficient to produce an emulator with an acceptable degree of accuracy. Increasing the number of Phast runs is shown to have no effect on the conclusions of the sensitivity analysis.

The study demonstrates that Bayesian analysis of model sensitivity can be conducted rapidly and easily on consequence models such as Phast. There is the potential for this to become a routine part of consequence modelling.

Keywords

Sensitivity, Bayesian, emulator, Phast, consequence modelling

1.   Introduction

Consequence modelling is used in the process industries for many purposes, from plant design to risk assessment and incident investigation. In many applications, the inputs to the consequence model (e.g. operating temperature and pressure, wind speed) are either poorly defined or they feature a large degree of variability. It is important in these cases to know the effect of the range in input conditions on the model predictions. The results may be quite insensitive to certain inputs, but for some inputs a small difference may produce a critical change in the study outcome. With experience, modellers can often develop an understanding of the important factors in a given situation, but in complex multi-phase reacting flows this may be challenging, and model behaviour can sometimes be counter-intuitive.

The purpose of a sensitivity analysis is primarily to determine which input parameters have a significant effect on the model outputs. Knowing which factors are important can be useful in driving model refinement and in producing more reliable predictions. For example, in the analysis of dense gas dispersion in the Buncefield Incident, Gant & Atkinson (2011) initially found that the model predictions were sensitive to the slope of the ground and the presence of obstacles. As a consequence, to refine their model they used detailed topographical data from a site survey to construct the final Computational Fluid Dynamics (CFD) geometry. This type of uncertainty that can be reduced through improved knowledge of the system is known as *epistemic* uncertainty.

Another type of uncertainty that cannot be reduced in this way, known as *aleatoric* uncertainty, arises from the inherent variability in the physical system or environment that is being modelled. For instance, in modelling atmospheric dispersion there is a natural uncertainty in the wind speed due to the random nature of atmospheric turbulence. To account for this, the wind speed may be expressed in terms of a probability distribution that represents the likelihood of the wind speed taking a particular value over time. In a risk assessment, where the objective is to determine the cumulative risk over a year, the results from multiple simulations for a range of wind speeds may be combined and weighted using this distribution to account for the range in likely values.

In a sensitivity analysis, it is also beneficial to identify the inputs that have a negligible effect on the model output. This information can be used to limit the number of simulations required in a given study. For example, in a risk assessment involving a jet fire, if the ambient wind is demonstrated to have practically no effect on the thermal dose predictions, the risk assessment may not need to consider running multiple jet fire simulations for a range of different wind speeds, which may considerably reduce the total computing effort required.

The issue of sensitivity and uncertainty in consequence modelling has long been appreciated, and a number of examples can be found in the literature (Khoudja, 1988; Jahn et al., 2008; Witlox et al., 2011; Carpentieri et al., 2012). Often the approach used in these studies to examine model sensitivity has consisted of selecting a baseline case and then varying one input parameter at a time, i.e. local sensitivity analysis. This choice has often been taken due to the limitations of computing time and the ease of interpreting the results.

In recent years, a more rigorous approach to model sensitivity analysis has started to be applied to process safety applications, e.g. Brohus et al. (2007) and Pandya et al. (2012). In the latter study, a global sensitivity analysis was performed on the consequence model Phast (DNV, 2012), where multiple input parameters were varied at the same time in order to understand the interactions between the different inputs. The calculations involved running Monte-Carlo experiments on Phast directly, with sample sizes of 20,000 simulations and computing times of around 24 h, using several computers in parallel.

Despite these examples of global sensitivity analysis being applied to consequence models, such analyses have yet to become widely used by engineers in the chemical process safety industry. This has perhaps been due to the perception that such exercises are time-consuming and costly, and the fact that much of the literature describing sensitivity analysis is aimed at mathematicians rather than practising engineers.

The aim of the present work is to demonstrate an approach to global sensitivity analysis that is easy to use and can be applied routinely to consequence modelling for process safety applications. The approach involves running a consequence model around a hundred times and then using the results to construct a statistical model (essentially a curve fit, or response surface). This statistical model is then used to undertake the sensitivity analysis and identify important input parameters. The statistical analysis is undertaken here using the Gaussian Emulation Machine (GEM) software produced by Marc Kennedy and colleagues at Sheffield University (Kennedy, 2005). This software is freely-available for non-commercial use, and features an easy-to-use Graphical User Interface (GUI) and good documentation.

The process safety scenarios examined consist of horizontal, above-ground, steady-state discharges of high-pressure carbon dioxide (CO_2). Consequence model predictions have been obtained using the discharge and dispersion models contained in the hazard assessment software package Phast (DNV, 2012). Seven key Phast model input parameters have been varied and the results analysed for main effects and interactions.

2.   Methodology

2.1.   Phast

Phast is a hazard-assessment software package produced by DNV Software for modelling atmospheric releases of flammable or toxic chemicals (Witlox et al., 2008; Witlox, 2010). It includes methods for calculating discharge and dispersion, and toxic or flammable effects (see Figure 1). A principal component of Phast is the Unified Dispersion Model (UDM), which incorporates sub-models for two-phase jets, heavy and passive dispersion, droplet rainout and pool spreading/evaporation. The model can simulate both unpressurised and pressurised releases, time-dependent releases (steady-state, finite-duration, instantaneous or time-varying), buoyancy effects (buoyant rising cloud, passive dispersion or heavy-gas-dispersion), complex thermodynamic behaviour (multiple-phase or reacting plumes), ground effects (soil or water, flat terrain with uniform surface roughness), and different atmospheric conditions (stable, neutral or unstable).
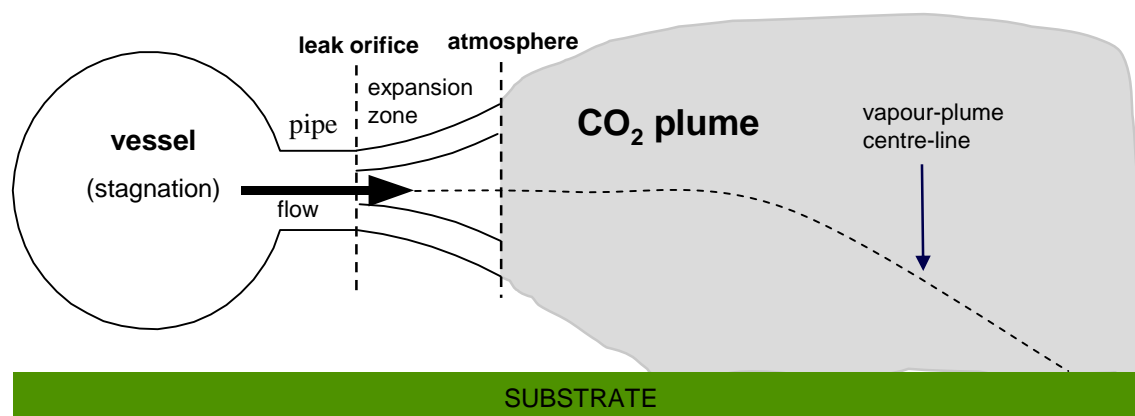

Figure 1 Phast discharge and dispersion model

Three key papers have been produced by DNV Software on CO_2 release and dispersion modelling using Phast. In the first of these, Witlox et al. (2009) described an extension to PHAST version 6.53.1 to account for the effects of solid CO_2. The modifications consisted principally of changing the way in which equilibrium conditions were calculated in the expansion of CO_2 to atmospheric pressure, to ensure that below the triple point the fluid conditions followed the sublimation curve in the phase diagram. Furthermore, two-phase vapour/solid effects instead of vapour/liquid effects were included downstream of the orifice following depressurisation of CO_2 to ambient pressure in the discharge and dispersion model. In the second paper, Witlox et al. (2011) reported the results of a sensitivity analysis for both liquid and supercritical CO_2 releases from vessels and pipes with the revised PHAST version 6.6 model. The sensitivity analysis was performed using a local "one-at-a-time" approach to model input variation, where each parameter was varied in turn whilst holding all other parameters fixed. In contrast, in the global sensitivity analysis presented here, all of the parameters have been varied simultaneously to calculate the effect of each parameter over the full range of other input parameters. In the third paper by Witlox et al. (2012), the results of a model validation study were published using measurements from a series of field-scale CO_2 tests originally commissioned by BP and Shell.

In the present work, Phast version 6.7 has been used, details of which can be found in the papers of Witlox et al. (2009, 2011, 2012). The guidance provided in the Phast version 6.6 release notes on the correct model configuration for CO_2 releases has been followed. For the

expansion from stagnation to orifice conditions with no pipe attached to the vessel, the fluid at the orifice has been assumed to be in a meta-stable liquid state, while for the expansion from orifice to ambient pressure, conservation of mass, momentum and energy has been assumed (see Figure 1).

Phast assumes that the two-phase flow in the jet is in homogeneous equilibrium, i.e. the solid CO_2 particles are assumed to have the same temperature and velocity as the surrounding vapour. Both Witlox et al. (2012) and Dixon et al. (2012) have shown that this is a valid approximation for pressurized releases of CO_2, for orifices up to 2 inches in diameter at least. Phast version 6.7 also assumes that the solid CO_2 particles remain within the dispersing jet and do not deposit on the ground. This appears to be a reasonable approximation for unobstructed jet releases of dense-phase CO_2, based on the results from the Shell and BP experiments.

2.2. Overview of Test Cases

The CO_2 releases simulated here consist of above-ground, unconfined, horizontal, steady-state orifice discharges from vessels in atmospheric conditions of neutral (D-class) stability. The range of conditions modelled using Phast is given in Table 1. For simplicity, uniform probability distributions have been used for each variable. This means, for example, that any wind speed between 0.5 and 50 m/s has been considered equally likely. The implications of this approximation are discussed later.

The model output that is considered to be of primary importance is the distance from the orifice to a particular limiting concentration of CO_2 (termed "dispersion distance" here). Initially, results are presented for a prescribed concentration limit of 6.9 % v/v, which for a steady exposure duration of 30 minutes corresponds to a Dangerous Toxic Load (DTL) of 1.5×10^{40} ppm.min^N (with N = 8). This is the Specified Level of Toxicity (SLOT) for CO_2 used by the UK Health and Safety Executive (HSE, 2012a). In Section 3, results are also reported for a range of other limiting CO_2 concentrations.

In steady-state atmospheric releases, the CO_2 plume concentrations are not constant over time but instead naturally vary about the mean concentration due to turbulence. Any such variations are important in the case of CO_2 since the DTL increases rapidly with concentration (to the power eight in HSE's model). Gant and Kelsey (2012) highlighted that predictions of the hazard range would be non-conservative if turbulent concentration fluctuations were ignored. Care should therefore be exercised in interpreting the 6.9 % v/v contour used here as the distance to SLOT.

Table 1 Parameters varied in Phast global sensitivity analysis

Number	Parameter	Minimum	Maximum
1	Vessel temperature	5 °C	30 °C
2	Vessel pressure	100 bar	150 bar
3	Orifice diameter	½ inch (12.6 mm)	2 inch (50.8 mm)
4	Wind speed	0.5 m/s	50 m/s
5	Relative humidity	0 %	100 %
6	Ground surface roughness	0.0001 m	1.0 m
7	Release height above ground	0.5 m	3 m

2.3.   Global Sensitivity Analysis

To identify the important model input parameters that affect the predicted dispersion distance, the statistical analysis examines the *variance* in the model predictions. The variance is the averaged squared distance between the model predictions and the average value of the model predictions. An input parameter that has a large effect on the output will account for a large proportion of the output variance. The total output variance is calculated by varying all of the seven input parameters given in Table 1 over the full range of conditions.

Two key parameters are used to identify which model inputs have a significant effect on the output: the *main effect* and *total effect*. The main (or first-order) effect of a given parameter *"A"* is the amount of variance that would be removed from the total output variance if the true value of *A* were known. The total effect, on the other hand, is the amount of variance that would remain unexplained if the values of all other parameters except *A* were known, and the variance was produced only by varying *A*. Formal mathematical definitions can be found in textbooks on sensitivity analysis, such as Saltelli et al. (2000).

In practical terms, the main effect quantifies the influence of just one parameter varying on its own, and the total effect comprises the main effect plus any variance due to interactions between that parameter and all of the other input parameters varying at the same time. The total effect is therefore always equal to or greater than the main effect. If there is not much interaction between the input parameters, the sum of the main effects for each of the input parameters will be close to the total output variance. The extent to which an input parameter interacts with other inputs is indicated by the difference between its main and total effects. A parameter that has a total effect that is much larger than its main effect is interacting strongly with other parameters.

The popularity of the main and total effects as a means of describing model sensitivity is largely due to the fact that these quantities are relatively easy to compute. Various methods have been developed for calculating them more efficiently than a typical brute force Monte-Carlo approach (Janssen et al., 1994; Saltelli et al., 1999; Sobol', 2001), although these still require many thousands of model evaluations. In principle, specific interaction variances can be calculated from numerical integration via Monte-Carlo sampling, but this requires sample sizes that are an order of magnitude larger than for the methods used to calculate just main and total effects variances.

Whilst a variance-based sensitivity analysis summarises the importance of each input with respect to the variability in model output, it conveys no information about precisely how the output responds with respect to each input or group of interacting inputs. Summaries that quantify how (on average) the model output changes in response to changes in the values of the inputs, displayed graphically, are an important tool for understanding the underlying model behaviour, although these are not frequently reported in the literature due to the computing time involved. In the current study, this kind of analysis allows us to identify the conditions that give rise to the greatest dispersion distance. Such summaries are related to variance-based methods, which Oakley and O'Hagan (2004) calculated via Monte-Carlo sampling.

2.4.   Gaussian Emulation Machine (GEM)

The global sensitivity analysis is conducted here using the Gaussian Emulation Machine (GEM) software (Kennedy, 2005). A useful introduction for non-specialists to the techniques employed by GEM is given in the paper by O'Hagan (2006), with further details provided by Oakley and O'Hagan (2004).

In essence, the emulator is a sophisticated curve-fit to a number of data points (known collectively as "training" data). In the present work, these data points are the dispersion distance results produced from a number of Phast simulations. Rather than just fitting a single line through these data points, the emulator fits a probability distribution, parameterised by a mean and standard deviation (Figure 2). The curve that passes through all of the data points represents the mean of the distribution. Between two neighbouring data points, there is some uncertainty in the output value due to interpolation and/or extrapolation errors. The emulator represents this uncertainty using the standard deviation of the probability distribution (i.e. the spread in possible values about the mean). At the data points, the emulator produces an output equal to the Phast results with zero uncertainty. Moving away from these points, the uncertainty gradually increases.

The principal underlying assumption used to derive the emulator is that the output is a homogeneously smooth, continuous function of the input parameters. The emulator is based upon a Gaussian Process (GP) regression model, which is specified in a Bayesian framework. The GP is parameterised by a mean function, which represents prior beliefs about how the output varies as the inputs are varied, and a correlation function, which represents beliefs about the smoothness of the output with respect to the inputs. The mean function and correlation functions are expressed in terms of further 'hyper-parameters', which are determined using the training data as the prior is updated using Bayes theorem. Both the posterior mean and correlations can be written as the prior expression plus a weighted linear combination of the observations, with weights determined by the location in parameter space where a prediction is sought, see Oakley and O'Hagan (2004) for details. The posterior surface can be viewed as a distortion of the original parametric approximation to the surface, such that it smoothly interpolates the observed data and the uncertainty pinches at the design points (Figure 2).
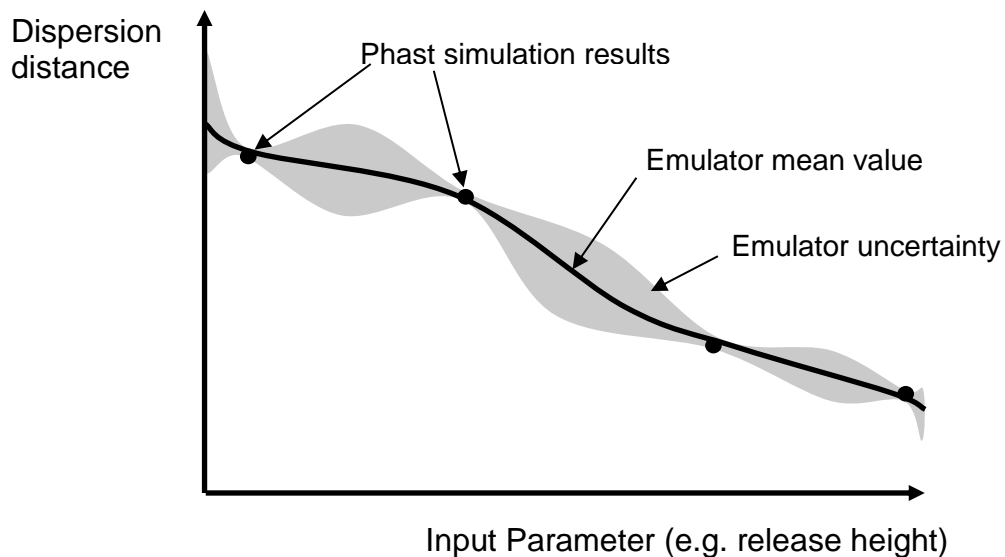
Figure 2 Illustration of the emulator fit to training data points

GEM assumes a squared exponential form for the correlation function and constant or linear forms (with respect to each of the inputs) for the mean function. Whilst these impact upon the flexibility of the approach, the resulting posterior can be analytically integrated with respect to the joint probability density function of the inputs (for Gaussian and uniform inputs), and all sensitivity measures described earlier (Section 2.3) can be derived without the need for numerical methods. The result is a highly efficient method, with computing times of typically around a minute on a standard office laptop computer. In the present work, the computer used for the calculations was equipped with a 2.53GHz Intel Core2 Duo P8700 CPU with 4Gb of RAM.

To specify the input conditions for the Phast simulations (i.e. the training data points), a maximin Latin hypercube algorithm is used. This is called from within the GEM software and the user simply needs to specify the ranges over which each of the input parameters are to be varied. The advantage of this sampling algorithm is that no values are repeated for each input so that if one or more of the input variables has no effect on the output, the training points provide greater coverage of the other input parameters.

This effect is illustrated in Figure 3, which shows the sample space for three hypothetical input parameters: *X*, *Y* and *Z*, which are each assumed to take any value between 0 and 1. Figure 3a shows how the maximin Latin hypercube algorithm distributes 30 samples across this space. The algorithm varies all three inputs simultaneously, rather than picking a baseline set of input values and varying each input one-at-a-time.
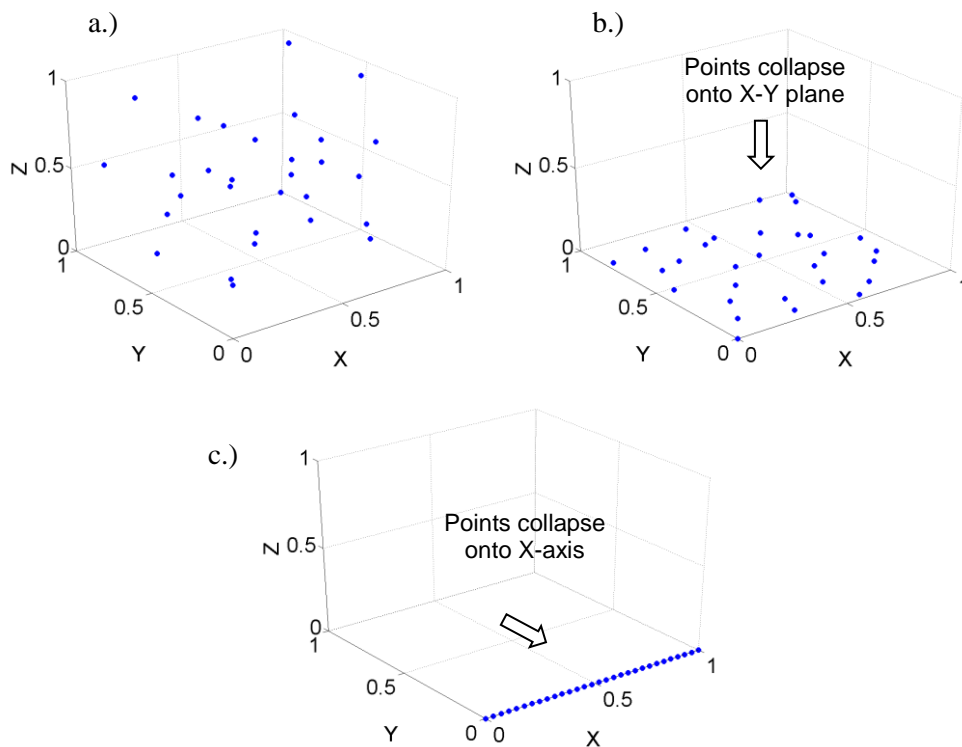

Figure 3 Maximin Latin hypercube sampling with 3 input parameters (*X*, *Y* and *Z*) and 30 samples, showing how samples are distributed a.) across all input parameters b.) when input parameter *Z* has no effect on the output c.) when both *Y* and *Z* input parameters have no effect on the output.

If input parameter *Z* has no effect on the model output, this means that all of the data points are in effect distributed across the plane of *X* and *Y* inputs, as shown in Figure 3b. None of the points lies on top of another and the points are fairly uniformly distributed across the plane, providing a good coverage of the range of input conditions. If both inputs *Y* and *Z* have no effect on the model output, the data points are effectively distributed along a single line in *X* (Figure 3c). The 30 samples are evenly distributed along this line, providing excellent coverage.

Since the maximin Latin hypercube algorithm varies all input parameters simultaneously, there is no pre-defined step size that is used to increment each input parameter between its specified minimum and maximum values. Instead, the distance between sampled points is a function of the number of input variables, their range (between specified minimum and maximum values) and the number of samples taken.

Often, in situations of practical interest, just one or two input parameters dominate the model predictions (Campolongo et al., 1999). As demonstrated by this simple example, the maximin Latin hypercube algorithm helps to maximise the coverage from a relatively small number of training data points. The uncertainty in the emulator predictions can be reduced by increasing the number of training data points, but in the present work it has been found using GEM that a hundred or so training points is sufficient to produce results with an acceptable level of accuracy. To obtain a similar level of accuracy using a standard "brute-force" Monte-Carlo method applied directly to Phast (without the emulator) would usually require many thousands of Phast runs, with much longer computing times, as demonstrated by Pandya et al. (2012).

Consequence models such as Phast calculate the dispersion behaviour by solving numerically a set of ordinary differential equations, using a variable step size in order to obtain the solution within a user-specified convergence tolerance. This means that there exists a small degree of error in the results (within the convergence tolerance), such that if the dispersion distance were plotted against a range of input conditions, the curve would not be perfectly smooth but would exhibit a small degree of scatter. Similar behaviour would be expected for CFD models, due to the limitations of numerical convergence and the approaches typically used to discretise time and space.

To avoid the emulator being forced to zigzag through every training data point, a small uncertainty bound can be introduced around each point during the emulator fitting process, using what is known in the statistical literature as a "nugget". The size of this error is estimated from the variance in the training data, and it enables a smoother fit of the emulator to the training data. Justification for use of a nugget has been provided by Gramacy & Lee (2012) and Andrianakis et al. (2012). A nugget has been used in all of the results presented here and the effect of not using a nugget is examined briefly in Section 3.3.

It is important to check that the emulator is constructed using a sufficient number of training data points in order to provide an accurate curve-fit to the underlying Phast model. To assess this, "cross-validation" tests are performed in which the emulator is fitted with one of the training data points left out. The emulator is then used to predict this missing training data point, and the emulator predictions are compared to the Phast predictions at that point. This process is repeated over all of the training data points, leaving each one out in turn, to obtain an overall picture of the emulator accuracy across the whole of the sample space. Further details on emulator validation can be found in the work of Bastos and O'Hagan (2008). The whole process of cross-validation is made easy in GEM and simply requires the user to tick a check-box and analyse the results. Results from cross-validation tests are reported in Section 3.4.

2.5.   Interfacing Models and Analysis Procedure

There are several steps involved in performing the sensitivity analysis of Phast using GEM. Firstly, the input parameters for the hundred Phast model runs are defined, using the maximin Latin hypercube algorithm in GEM. A short MATLAB script (MathWorks, 2011) is then used to take the text file produced by GEM and use this to modify a template Phast text-input (*.PSU) file to create a hundred separate Phast input files. The Phast simulations are then run in batch mode, typically taking a few seconds each. The text output (*.OUT) files from Phast are then post-processed using another MATLAB script to extract the desired output (in this case, the dispersion distance) and this data is written to another text output file. Finally, the input and output files are imported into GEM, which constructs the emulator and runs the sensitivity analysis in typically around a minute. If the results produced by GEM show that there are some interactions between model input parameters, it is straightforward to select those parameters to examine for joint effects using the GEM GUI, and then re-run the analysis.

The whole sensitivity analysis (including the Phast runs) typically requires less than 30 minutes of computing time on a standard office laptop computer. The most time-consuming aspect of the study, which may require a few days effort, is to write the MATLAB scripts to process the data and interface GEM and Phast. However, these scripts only need to be written once.

3.   Results

3.1.   Baseline Global Sensitivity Analysis

Sample dispersion predictions from Phast for three different orifice sizes are shown in Figure 4. These show that the horizontal jet with initial height of 1 m above the ground does not touch down for the smallest ½-inch orifice, but does so with the larger 1- and 2-inch orifices. For the larger diameter releases, the jet extends further into the ambient environment and, in doing so, the width of the jet expands due to entrainment of fresh air. The flow within the concentration contours shown is dominated by high-momentum jet behaviour rather than gravity slumping. Across the full range of conditions given in Table 1, the dispersion distances ranged from 12 m to 110 m, with a mean of 40 m and total variance of 393 m^2.

The contribution from each of the model input parameters to the total output variance is shown in Figure 5 in the form of a "Lowry plot". These results were produced from GEM using a sample size of a hundred Phast runs, where the output parameter used in the sensitivity analysis was the dispersion distance to 6.9 % v/v CO_2. The vertical bars in Figure 5 show the main and total effect for each parameter, ranked in order of main effect importance, whilst the lower and upper bounds of the curve show the cumulative sum of the main and total effects, respectively. The analysis shows that more than 70 % of the variance in the Phast results is due to the orifice diameter (Parameter 3 in Table 1). The second highest contribution comes from the release height above the ground. The remaining factors (or combinations of factors) account for less than 12 % of the total output variance.

The effect on the mean dispersion distance due to changes in the orifice diameter and the release height is shown in Figure 6. The dispersion distance increases with the orifice diameter, roughly proportional to the mass release rate (i.e. orifice diameter squared). In contrast, the dispersion distance decreases with the release height, due to the reduced air entrainment when the jet is closer to the ground. The error bars in these graphs show the uncertainty in the emulator predictions to two standard deviations either side of the mean (i.e.

a confidence interval of 95 %). The uncertainty increases towards the minima and maxima of the range of input values, since the emulator relies upon fewer training data points near these locations.

The size of the total effect relative to the main effect in Figure 5 indicates that there is some interaction between the orifice diameter and release height inputs. The interplay between these inputs is shown in Figure 7 in the form of a three-dimensional plot of the dispersion distance. The results show that for large orifices, the release height has a significant effect on the dispersion distance, but for smaller orifices the release height has practically no effect. This trend matches the flow behaviour shown in Figure 4, where the CO_2 jet did not touch down with the smallest ½-inch diameter orifice.

A notable finding from the sensitivity analysis is that the vessel pressure has little effect on the dispersion distance. At first, this may seem surprising, since an increase in pressure from 100 barg to 150 barg leads to an increase of around 20 % in mass release rate of CO_2. However, this increase in mass flow rate is also associated with an increase in the release velocity, which enhances the entrainment of fresh air into the jet. This additional dilution balances the increase in mass flow rate so that, overall, the dispersion distance is little affected. A similar effect is observed in subsonic gas jets, for details see Webber et al. (2011).

One of the benefits of GEM is that it produces the results shown in Figures 6 and 7, which provide a useful visual description of how the model responds to changes in the input variables. Similar results could not be produced using the Extended Fourier Amplitude Sensitivity Test (EFAST) global sensitivity analysis method employed used by Pandya et al. (2012), which calculates solely the main and total effects. In principle, Pandya et al. (2012) could have used direct Monte-Carlo sampling of Phast to produce similar graphs to those shown in Figures 6 and 7. To reproduce the Figure 6a would involve fixing the orifice diameter at a particular value and then sampling over the remaining six variable input parameters and then averaging over these results. The same process would be repeated for a range of different orifice diameters to produce the curve with its associated error bounds.

Using direct Monte Carlo sampling in this way is less efficient than EFAST and it would probably require an order of magnitude larger sample sizes in order to produce results with acceptable confidence intervals. In their original work, Pandya et al. (2012) used a sample size of 20,000 and ran their analysis on multiple computers to obtain results within around 24 h. This suggests that reproducing results such as Figure 6a would require prolonged computing times. The advantage of GEM in this respect is that the integration over the six variable input parameters for Figure 6a is performed analytically, making the process computationally very efficient. As a consequence, graphs like those shown in Figures 6 and 7 can be produced in a matter of minutes.
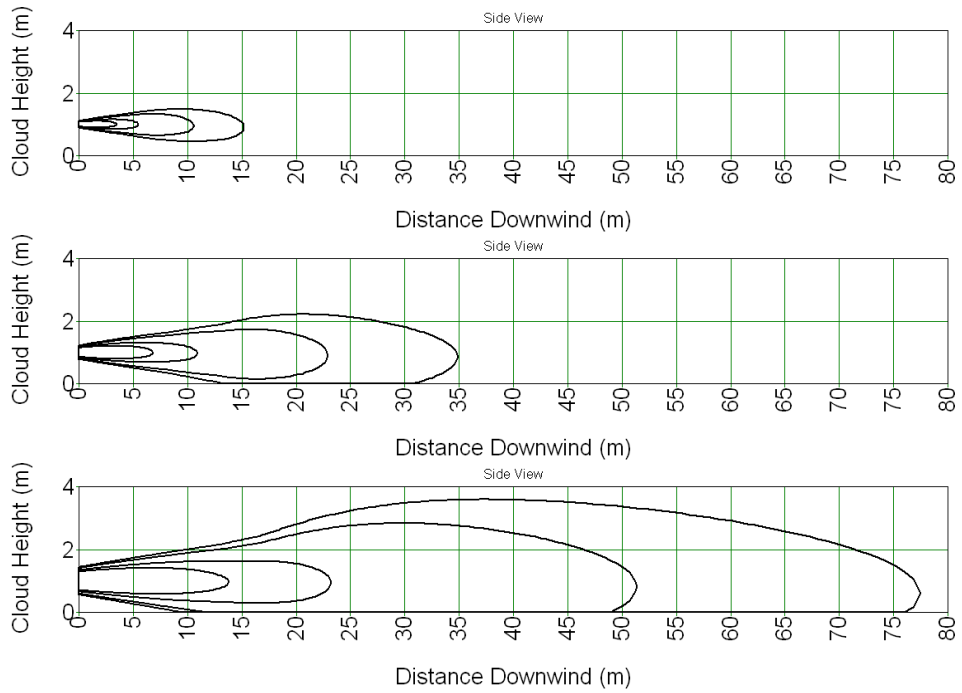
Figure 4 Sample output from Phast showing the plumes produced by orifice diameters: ½ inch (top), 1 inch (middle) and 2 inch (bottom). Contours shown are at 6.9, 10, 20 and 30 % v/v CO_2.
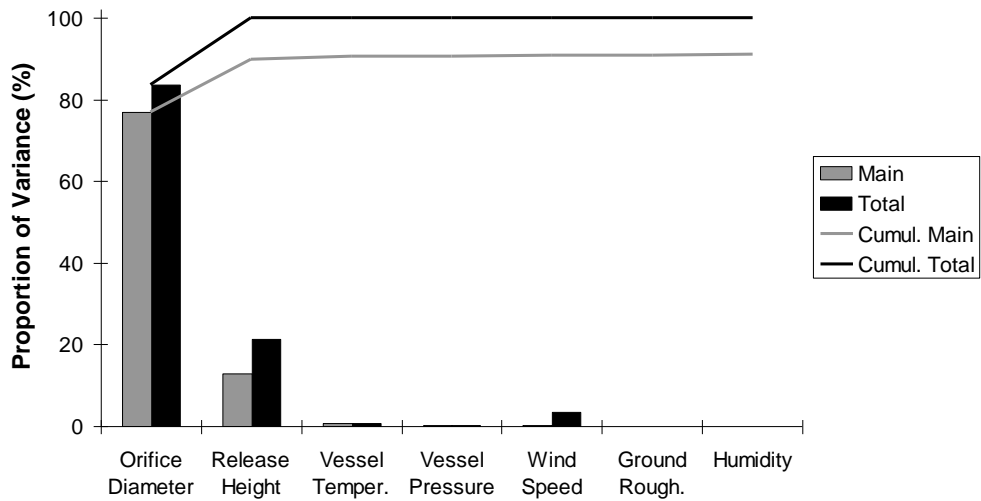

Figure 5 "Lowry" plot showing main and total effects for Phast model input parameters given in Table 1
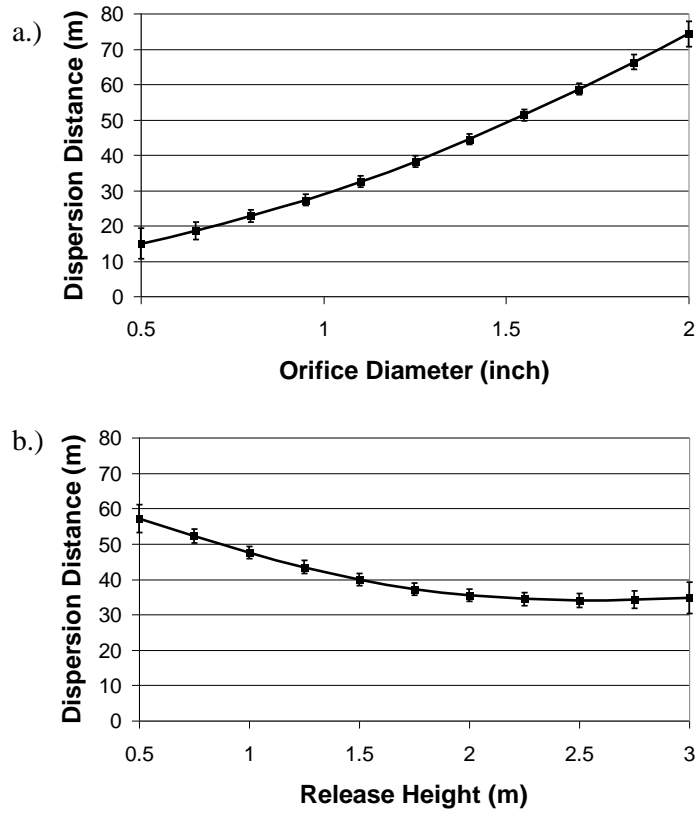
Figure 6 Mean dispersion distance predicted by GEM as a function of a) orifice diameter, and b) release height; error bars show 95% confidence interval
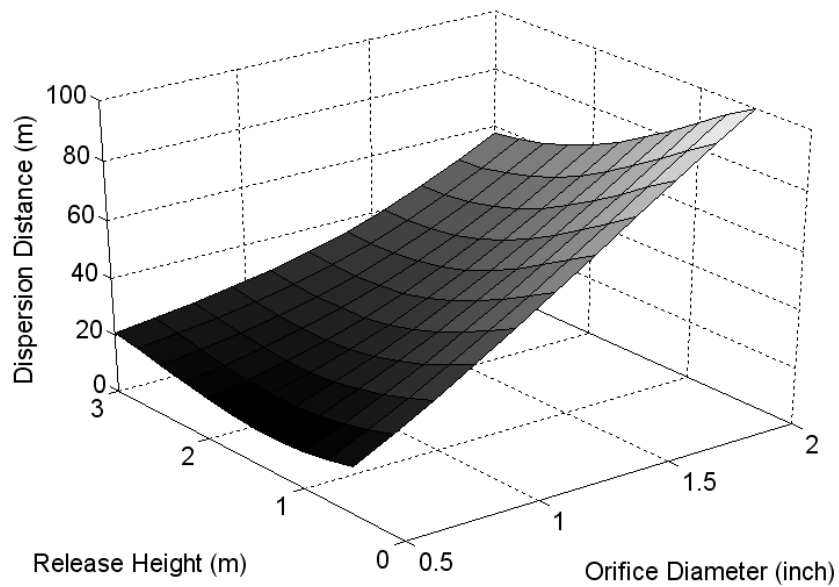

Figure 7 Joint effects resulting from varying the orifice diameter and release height simultaneously

3.2.   Effect of the Wind Speed

Another important finding from the sensitivity analysis shown in Figure 5 is that the ambient wind speed has very little effect on the dispersion distance, despite the sensitivity analysis being based on a wide range of wind speeds: from 0.5 m/s to 50 m/s. This result appears to be due to the dispersion behaviour of the CO_2 cloud above a concentration of 6.9 % v/v CO_2 being dominated by the initial momentum of the pressurised jet, rather than the ambient conditions. The Richardson number throughout this high-concentration part of the jet was well below a value of one, indicating that the flow was dominated by momentum rather than gravitational effects.

If a lower CO_2 concentration is used to define the edge of the cloud than 6.9 % v/v CO_2, the cloud extends further into the ambient environment and its velocity decreases. The wind speed then has a more significant effect. This is demonstrated in Figure 8, which shows the calculated main effects of the orifice size, release height and wind speed inputs for CO_2 clouds that are defined using a range of different concentrations, from 0.1 % to 6.9 % v/v CO_2. As the concentration defining the edge of the cloud is reduced, the flow behaviour is increasingly dominated by far-field passive dispersion. The results show that below a concentration of 1 % v/v, the effect of the release height decreases whilst the effect of the wind speed increases relative to the orifice size. If the edge of the cloud is defined using a low concentration of 0.1 % v/v CO_2, the wind speed is responsible for more than 50 % of the variance in the dispersion distance predictions.

There are two facts that should be borne in mind when interpreting the results shown in Figure 8. Firstly, the graph shows the main effects in terms of their contribution to the total output variance. This variance increases markedly as the concentration that is used to define the edge of the cloud decreases (and the CO_2 cloud generally becomes larger). Although in relative terms the main effect of the orifice diameter decreases as the concentration is decreased from 6.9 % to 0.1 % v/v CO_2, in absolute terms it actually increases.

Secondly, these main effects have been calculated using uniform probability distributions for the input variables, i.e. assuming that any wind speed between 0.5 and 50 m/s is equally likely. From a risk assessment perspective, wind speeds of between 1 m/s and 15 m/s are typically considered to be far more likely than those less than 1 m/s or greater than 15 m/s. If a more realistic distribution of wind speed were used in the sensitivity analysis, the wind speed would probably be less dominant. In order to undertake such analyses, meteorological data could be obtained from the Met Office for sites in the UK (http://www.metoffice.gov.uk, accessed 20.08.2012). Further guidance on realistic probability distributions for wind speeds was given by Ro and Hunt (2007).
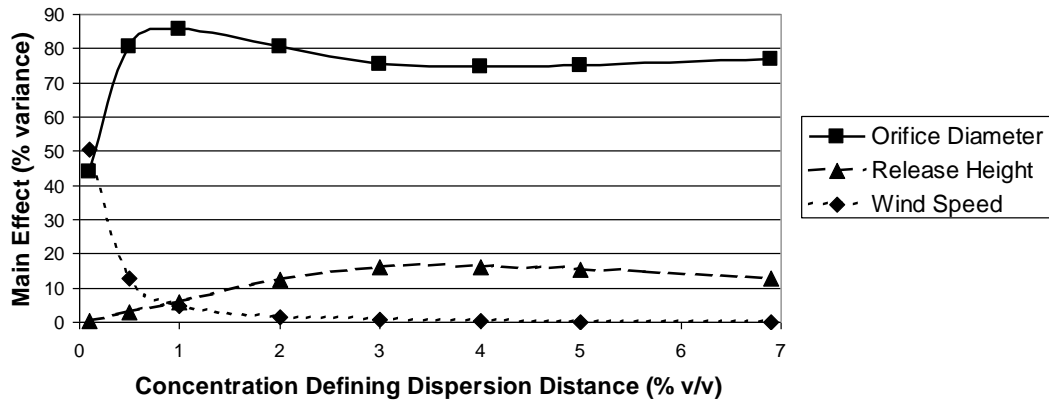
Figure 8 Main effects for the orifice diameter, release height and wind speed as a function of the CO_2 concentration used to define the dispersion distance.

3.3. Effect of Numerical Code Error

It was mentioned earlier that a "nugget" was used in the emulator to account for any model error. This in effect introduces a small uncertainty bound around each of the training data points so that the emulator is not forced to pass through each of the training data points, which therefore produces a smoother fit to the data. The size of the nugget is calculated automatically by GEM based on the variance in the output. Omitting the nugget leads to a decrease in the uncertainty in the emulator predictions (the error bars in Figure 6), with the average standard deviation falling from 1.2 m to 0.7 m. Similar behaviour was observed previously by Gramacy and Lee (2012).

Results from the cross-validation tests show that the predictive accuracy of the emulator is improved with the nugget. Only 80 % of the emulator's mean predictions of the dispersion distance are within 4 m of that predicted by Phast without the nugget (Figure 9), compared to 93 % with the nugget.

The overall effect of the nugget on the predicted main and total effects is fairly minor. Irrespective of whether or not the nugget is used, the same basic findings are obtained, i.e. the primary and secondary factors controlling the dispersion distance are still the orifice diameter and the release height (Figure 10). Use of the nugget had practically no effect on the computing time needed for the sensitivity analysis.
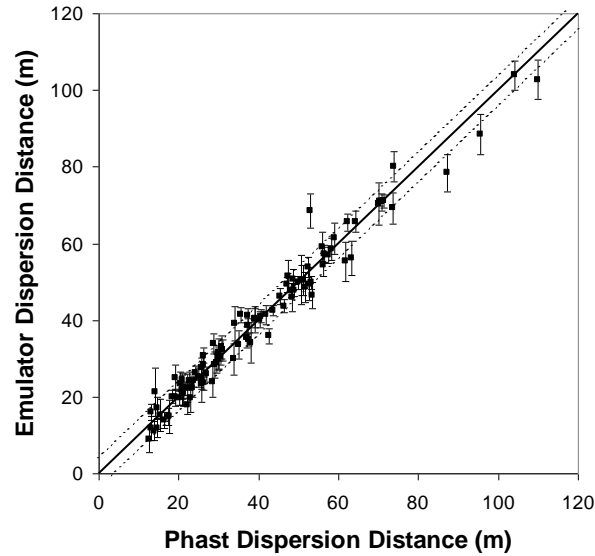
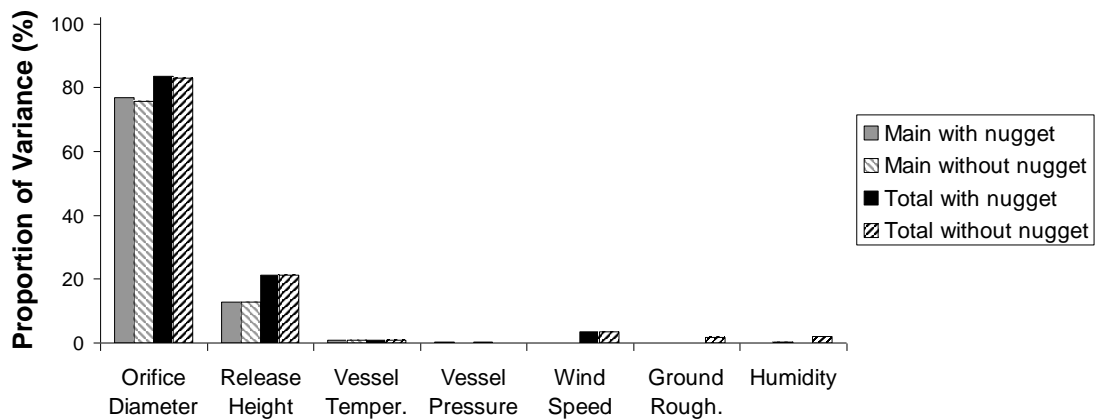**Figure 9** Cross-validation results using the emulator without the nugget

Figure 10 Main and total effects using the emulator with and without the nugget

3.4. Effect of the Sample Size

It is important to check that a sufficient number of model realizations have been performed in order to construct the emulator. The results of cross-validation tests are presented in Figure 11 for emulators constructed using 100 and 400 Phast runs. Each graph shows the emulator's prediction of the dispersion distance against that predicted by Phast. The mean predictions are shown as square symbols and the standard deviation as error bars. For 100 Phast runs, 93 % of the emulator's mean predictions of the dispersion distance are within 4 m of that predicted by Phast. In comparison, with 400 samples, 95 % of the emulator predictions are within ± 4 m of the Phast results.

The reason why the cross-validation results do not show greater improvement is due to the fact that some of the cross-validation emulator runs involve predictions outside of the sample space, i.e. using the emulator to extrapolate rather than interpolate the output. When the

emulator is used in this way, the uncertainty in the model predictions can be significant. Errors are largest for points located close to the vertices of the parameter space. For this reason, even for very large samples sizes, the emulator accuracy in the cross-validation tests may not reach 100 %. The cross-validation tests may therefore be viewed as a severe test of the emulator.

Increasing the sample size from 100 to 400 leads to a reduction in the emulator variance in the main effect predictions (i.e. a reduction in the size of the error bars in Figure 6). The average standard deviation in these results decreases from 1.2 m to 0.4 m when the number of samples is increased from 100 to 400.

A comparison of the main and total effects using 100 and 400 Phast runs shows that increasing the number of samples has a negligible effect on the results of the variance-based sensitivity analysis (Figure 12).

In terms of the computing time required by GEM to construct the emulator and run the sensitivity analysis, using 100 samples took around 1 minute, compared to 29 minutes with 400 samples (excluding the time needed for the additional Phast runs). This rapid increase in computing time with sample size is probably due to the part of the GEM calculation process that involves inverting an n-by-n matrix (where n is the sample size).

Overall, the use of 100 samples in the present study was considered to provide an acceptable degree of accuracy. Increasing the number of samples did not affect the conclusions from the sensitivity analysis.
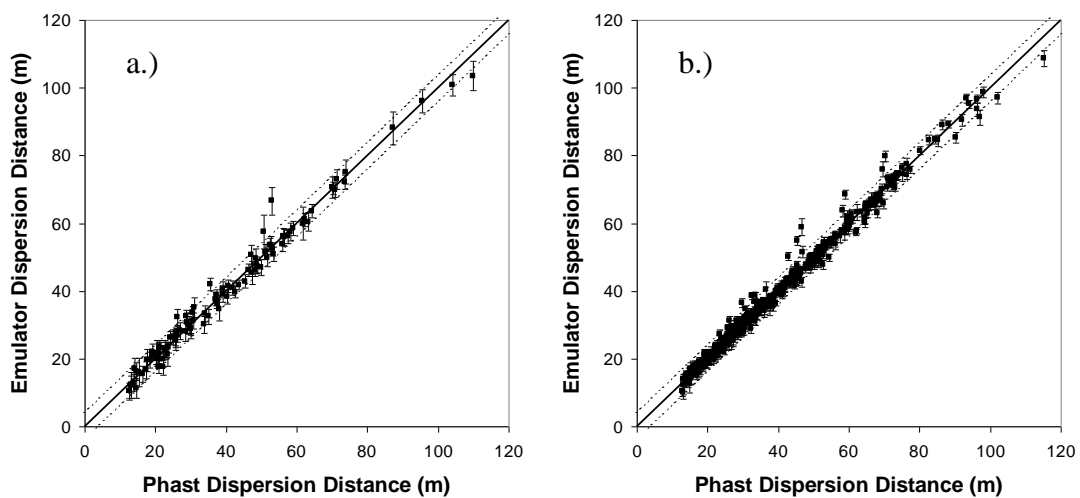

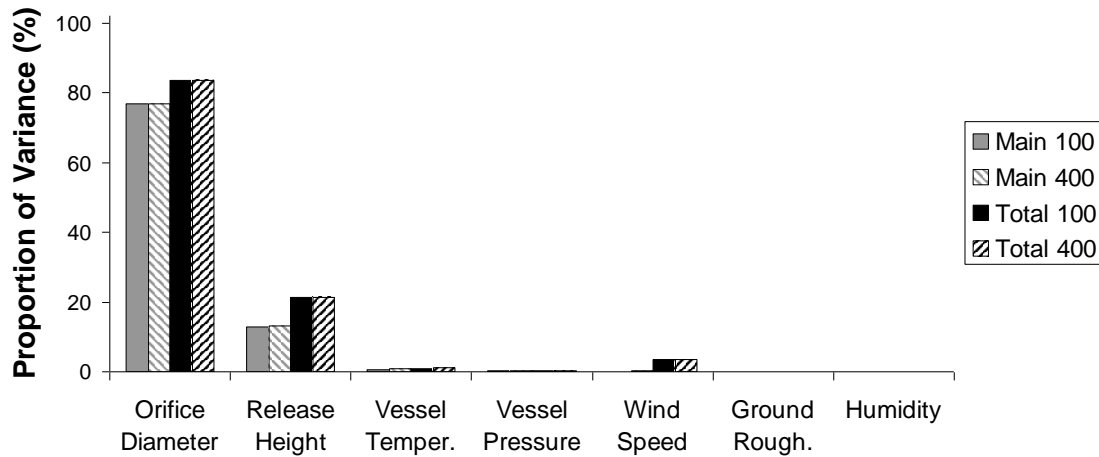**Figure 11** Cross-validation results using: a.) 100 samples; b.) 400 samples

Figure 12 Main and total effects using 100 and 400 samples

3.5.   Effect of Fixing the Orifice Diameter

In risk assessments, one or more generic orifice sizes are typically used to model releases from vessels and pipelines, rather than varying the orifice size continuously over a range of values (HSE, 2012b). For each discrete orifice size, it may therefore be of interest to understand the effect of varying the remaining input parameters, i.e. the vessel temperature and pressure, wind speed, humidity, ground surface roughness and release height. A similar situation may arise in certain design studies or incident investigations, where the orifice dimension may be fixed due to external factors.

Figure 13 shows the results from a global sensitivity analysis in which the orifice diameter has been fixed at 1 inch (25 mm) and the remaining parameters given in Table 1 have been varied over their given ranges. The results have been produced from 100 Phast runs using a nugget, and the dispersion distance has been defined using a concentration of 6.9 % v/v CO_2. As expected, the release height, which had the second-largest main effect in the previous sensitivity analysis (see Figure 5), is now the dominant input parameter. Its main effect is responsible for nearly 80 % of the total output variance. The total output variance with six variable input parameters has reduced significantly from the previous case where seven inputs were varied (a drop from 393 m^2 to only 33 m^2). In absolute terms, the release height is therefore responsible for a similar output variance to before. The three leading input parameters that most affect the dispersion distance are now the release height, vessel temperature and wind speed, ranked in order of importance according to their main effects. Tests were performed using 400 instead of 100 Phast model runs to train the emulator, which produced identical conclusions to those shown in Figure 13.

The main and total effects shown in Figure 13 indicate that there is an interaction between the release height and wind speed input parameters. The joint effects from varying these two inputs simultaneously are shown in Figure 14 in the form of a three-dimensional plot of the predicted dispersion distance. Three trends in model behaviour are highlighted on the graph with arrows marked *A*, *B* and *C*. The path identified by Arrow *A* can be explained by the fact that as the release height is reduced, the plume moves closer to the ground, the entrainment of fresh-air decreases and the dispersion distance therefore increases. For Arrow *B*, at low release heights the plume is grounded and increasing the wind-speed results in a longer, narrower plume and hence a larger dispersion distance. Finally, for Arrow *C,* at large release

heights the jet remains elevated for longer with increasing wind speed, and this leads to increasing air entrainment and smaller dispersion distances.

The balance of all the effects that produces the flow behaviour shown in Figure 14 is complex and not easily understood without an in-depth knowledge of the underlying modelled physics. The results provide a good demonstration of how global sensitivity analysis can provide a useful insight into the underlying model behaviour.
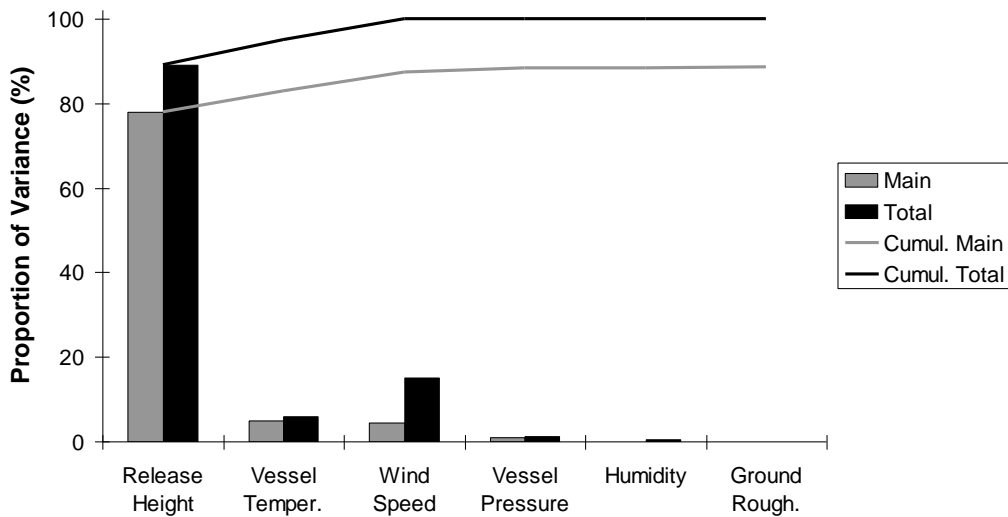

Figure 13 "Lowry" plot showing main and total effects for Phast model input parameters given in Table 1, with a fixed orifice diameter of 1 inch
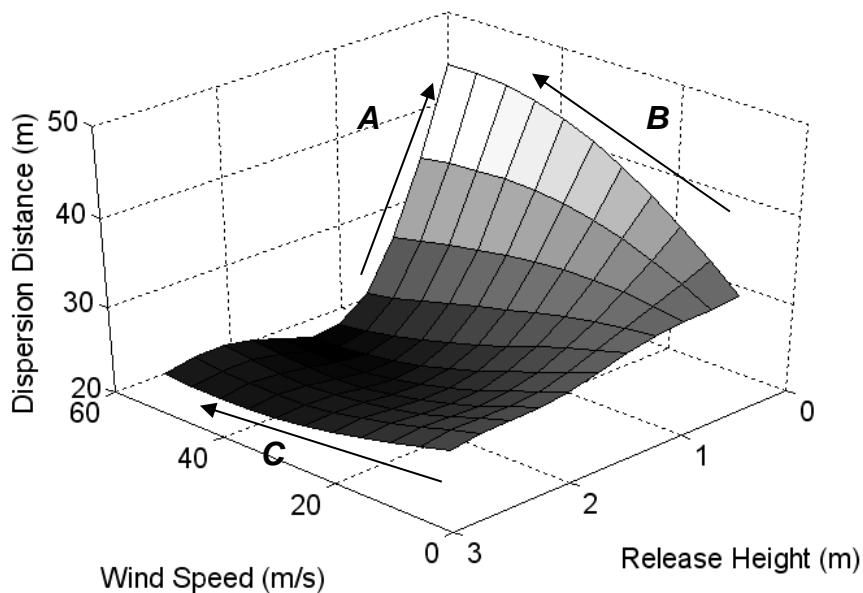

Figure 14 Joint effects resulting from varying the wind speed and release height simultaneously, with a fixed orifice diameter of 1 inch

3.6. Limitations of Gaussian Emulators for Sensitivity Analysis

One of the assumptions implicit in using an emulator is that the output is a homogeneously smooth, continuous function of the input parameters. Consequently, it is not generally recommended to use an emulator with input parameters that are switches or variables that take binary on/off values.

In the present work, the dispersions simulations were all performed with an atmosphere that was neutrally stable (D-class). Further tests were performed with eight input parameters instead of seven, where the eighth parameter represented the atmospheric stability varying between neutral D-class and stable F-class. To do this, the same maximin Latin hypercube algorithm was used to construct a hundred samples, but with an additional eighth input parameter varying from zero to one. If the eighth parameter had a value less than a half then the Phast simulations used D-class stability, and otherwise they used F-class. The results from this emulator showed significant variance in the predicted main effects, demonstrating that the atmospheric stability had a non-negligible effect on the model predictions. However, due to the limitations of the emulator fit, it was not possible to quantify reliably the magnitude of the atmospheric stability effect relative to that of the other input parameters (orifice size, release height etc.).

For this particular issue of atmospheric stability, the limitation of the emulator could in principle be overcome by specifying the atmospheric stability via the Monin-Obukhov length, rather than the Pasquill-Gifford stability class. The Monin-Obukhov length can be varied continuously between neutral and stable conditions, which avoids the sudden jump between D- and F-class states. Unfortunately, Phast version 6.7 does not allow users to specify the Monin-Obukhov length directly, despite the fact that the software internally converts the Pasquill-Gifford stability class selected by the user into a Monin-Obukhov length, although this may be possible in future software releases. Other dispersion models, such as ADMS (CERC, 2012), allow users to specify the Monin-Obukhov length directly. A further complication of undertaking a sensitivity analysis on the atmospheric stability is that the range of observed wind speeds is different for neutral and stable conditions, and it may therefore be simpler to undertake two separate sensitivity analyses for D- and F-class stability classes.

For other model inputs that are switches, such as the choice of whether an indoor or outdoor release is modelled, it is possible to follow a similar approach to that described above to determine whether or not the model input has *any* effect on the output. However, it is not possible to be precise about the main and total effects of this input as compared to other smoothly varying inputs. So-called "treed" Gaussian emulators have been developed which branch and, in effect, fit two different emulators to data that exhibit an abrupt step change (Gramacy & Lee, 2008). Although these emulators provide an improved fit to the data, they are similarly unable to calculate directly the main and total effects values.

Calculating these quantities for input parameters that are switches can be achieved using Monte-Carlo sampling, as demonstrated by Pandya et al. (2012), who analysed the effect of using either constant or power-law atmospheric wind profiles on the dispersion of three toxic chemicals. To decrease the computing time for this analysis, one potential solution would be to fit a treed emulator to the training data and then use Monte-Carlo sampling of the emulator to calculate the main and total effects. Alternatively, a much simpler way to understand the effect of a switch could be to undertake a similar analysis to that presented in Section 3.1 for each value of the switch, and then to plot the average effect of the input parameters on the output (as in Figure 6) with multiple lines for each switch value.

At present, GEM allows the use of input distributions that are either all uniform or all Gaussian. Gaussian distributions could be used to account for the higher likelihood of inputs taking their mean values than their minimum or maximum values. However, GEM does not allow the user to specify a mixture of uniform and Gaussian input distributions, nor is it recommended to try and approximate a log-normal distribution by taking the anti-log of the log-normal input distribution to turn it into a Gaussian distribution that can be used by GEM. Using this approach would produce variances that will differ from variance calculated from a true log-normal distribution. It is also not recommended to use a Gaussian input distribution with an artificially large variance for one or more of the input variables in order to produce a nearly uniform probability distribution between the minimum and maximum values. The methodology presented by Oakley and O'Hagan (2004) has more flexibility in terms of input distributions. However, direct coding of this requires considerable technical expertise.

To overcome the limitation of either all uniform or all Gaussian input distributions, one partial solution would be to create a response surface from GEM, but then rather than calculate the main effects analytically through GEM, to instead calculate these separately by Monte-Carlo sampling the emulator (applying the relevant distributions of the model inputs). This could provide significant savings in computing time compared to sampling the underlying consequence model directly. Advice on running the emulator independently of GEM is provided with the GEM download (Kennedy, 2005).

Finally, GEM has limits on the maximum number of input variables and samples. A maximum of 30 input variables and 400 samples can be used in any one analysis. If it is necessary to consider a larger number of input parameters, a two-stage approach could be taken (Campolongo, Tarantola, & Saltelli, 1999). In the first stage, a screening technique (Morris, 1991) could be used to identify those input parameters of low importance. In the second stage, these unimportant inputs could be fixed at central values, and the more important parameters studied using the techniques described in this paper.

4. Conclusions

A global sensitivity analysis has been performed on Phast using an emulator to identify the important factors affecting the dispersion distance in steady-state horizontal releases of CO_2 over flat terrain. The parameters varied include the CO_2 vessel temperature and pressure, orifice size, wind speed, humidity, surface roughness and height of the release. The output parameter of interest has initially been taken as the distance from the release point to a CO_2 concentration of 6.9 % v/v. The results have shown that for the range of conditions tested, the orifice diameter has a far greater impact than any of the other parameters varied. The second-largest effect was from the release height, with a lower release height producing a plume that extends further, due to the reduction in air entrainment.

When the dispersion distance output was defined differently, using a lower limiting value of the CO_2 concentration, the results showed that the dominant input parameters change and the effect of the ambient wind speed becomes more important. Tests on the sample size used to construct the emulator indicated that a hundred Phast runs were sufficient to provide an acceptable degree of accuracy.

The global sensitivity analysis of Phast typically required less than 30 minutes of computing time on a standard office laptop computer. This includes the time necessary for the hundred Phast runs and the statistical analysis. Due to the speed and ease of implementation, similar analyses could readily be incorporated into industrial design studies, risk assessments and incident investigations at little extra cost. There are significant benefits to be gained from

such analyses in terms of identifying the important physical processes in complex flows, and in narrowing the scope of further simulations or experimental measurements. These methods are therefore likely to become widely used in the process industry in the coming years.

In the present work, uniform probability distributions were applied for each of the input variables. For example, any wind speed was considered equally likely, within the range of conditions modelled. In future work, techniques for uncertainty analysis will be tested which apply realistic probabilities for wind speed, atmospheric stability etc. based on meteorological data. Further extensions to this work may also consider model calibration, using experimental datasets.

Acknowledgements

Disclaimer

References

Andrianakis, I., & Challenor, P. G. (2012). The effect of the nugget on Gaussian process emulators of computer models. *Computational Statistics & Data Analysis*, *56*(12), 4215-4228.

Bastos, L. S., & O'Hagan, A. (2008). Diagnostics for Gaussian process emulators. *Technometrics*, *51*, 425-438.

Brohus, H., Nielsen, P. V., Petersen, A. J., & Sommerlund-Larsen, K. (2007). Sensitivity analysis of Fire Dynamics Simulation. *Int. Conf. on Air Distribution in Rooms (Roomvent)*, Helsinki, Finland, 13-15 June 2007.

Campolongo, F., Tarantola, S., & Saltelli A. (1999). Tackling quantitatively large dimensionality problems. *Computer Physics Communications*, *117*, 75–85.

Carpentieri, M., Salizzoni, P., Robins, A., & Soulhac, L. (2012). Evaluation of a neighbourhood scale, street network dispersion model through comparison with wind tunnel data. *Environmental Modelling & Software*, *37*, 110-124.

CERC (2012). ADMS 4 Dispersion Model, Cambridge Environmental Research Consultants (CERC), Cambridge, UK.

Dixon, C. M., Gant, S. E., Obiorah, C., & Bilio, M. (2012). Validation of dispersion models for high pressure carbon dioxide releases. *IChemE Hazards XXIII Conference*, Southport, UK, 12–15 November 2012.

DNV (2012). Phast Version 6.7 Software, DNV Software, London, UK. Available from http://www.dnv.com/software, accessed 20.08.2012.

Gant, S.E., & Atkinson, G.T. (2011). Dispersion of the vapour cloud in the Buncefield Incident. *Process Safety & Environmental Protection*, *89* (6), 391–403.

Gant, S.E., & Kelsey, A. (2012). Accounting for the effect of concentration fluctuations on toxic load for gaseous releases of carbon dioxide. *Journal of Loss Prevention in the Process Industries*, *25*, 52–59.

Gramacy, R. B., & Lee, H. K. H. (2008). Bayesian treed Gaussian process models with an application to computer modeling. *Journal of the American Statistical Association*, *103*, 119–1130.

Gramacy, R. B., & Lee, H. K. H. (2012). Cases for the nugget in modeling computer experiments. *Statistics and Computing*, *22*(3), 713–722.

HSE (2012a). *Toxicity Levels of Chemicals*, Health and Safety Executive, Bootle, UK. Available from: http://www.hse.gov.uk/chemicals/haztox.htm , accessed 20.08.2012

HSE (2012b). *Failure Rate and Event Data for use within Risk Assessments (28/06/2012)*, Health and Safety Executive, Bootle, UK. Available from: http://www.hse.gov.uk/landuseplanning/failure-rates.pdf, accessed 20.08.2012

Jahn, W., Rein, G., & Torero, J. L. (2008). The effect of model parameters on the simulation of fire dynamics. *Fire Safety Science*, *9*, 1341–1352.

Jansen, M. J. W., Rossing, W. A. H., & Daamen, R. A. (1994). Monte-Carlo estimation of uncertainty contributions from several independent multi-variate sources. In J. Gasman, & G. van Straten (Eds.), *Predictability and Nonlinear Modelling in Natural Sciences and Economics* (pp.334–343). Dordrecht: Kluwer Academic Publishers.

Kennedy M. C. (2005) GEM-SA, Version 1.1. Software: Gaussian Emulation Machine for Sensitivity Analysis. Available from: http://www.tonyohagan.co.uk/academic/GEM/index.html, accessed 20.08.2012.

Khoudja, N. (1988). *Procedures for quantitative sensitivity and performance validation studies of a deterministic fire safety model*. Report NBS-GCR-88-544, U.S. Department of Commerce, National Bureau of Standards, National Engineering Laboratory, Center for Fire Research, Gaithersburg, MD 20899, USA.

MathWorks (2011) MATLAB version 2011b. MathWorks Inc., Natick, Massachusetts, USA.

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33, 161-174.

Oakley, J. E., & O'Hagan, A. (2004) Probabilistic sensitivity analysis of complex models, *Journal of the Royal Statistical Society*, *B 66*, 751–769.

O'Hagan T., 2006, Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering and System Safety, 91*, 1290–1300.

Pandya, N., Gabas, N., & Marsden, E. (2012). Sensitivity analysis of Phast's atmospheric dispersion model for three toxic materials (nitric oxide, ammonia, chlorine). *Journal of Loss Prevention in the Process Industries*, *25*, 20–32.

Ro, K. S., & Hunt, P. G. (2007) Characteristic wind speed distributions and reliability of the logarithmic wind profile. *Journal of Environmental Engineering*, *133*, 313-318.

Saltelli, A., Tarantola, S., & Chan, K. (1999). A quantitative, model independent method for global sensitivity analysis of model output. *Technometrics*, *41*, 39–56.

Saltelli, A., Chan, K., & Scott, E. M. (Eds.) (2000). *Sensitivity Analysis*. Chichester: John Wiley & Sons Ltd.

Sobol, I. M. (2001). Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, *55*, 271–280.

Webber, D. M., Ivings, M. J., & Santon, R. C. (2011) Ventilation theory and dispersion modelling applied to hazardous area classification. *Journal of Loss Prevention in the Process Industries*, *24*, 612–621.

Witlox, H. W. M. (2010). Overview of consequence modelling in the hazard assessment package Phast. *6th AMS Conference on Applications of Air Pollution Meteorology*, Atlanta, USA, 17–21 January 2010.

Witlox, H. W. M., Harper, M., & Oke, A. (2009). Modelling of discharge and atmospheric dispersion for carbon dioxide releases. *Journal of Loss Prevention in the Process Industries*, *22*, 795–802.

Witlox, H. W. M., Harper, M., & Oke, A. (2012). Phast validation of discharge and atmospheric dispersion for carbon dioxide releases. *Proc. 15th Annual Symposium,*

Mary Kay O'Connor Process Safety Centre, Texas A&M University, College Station, Texas, USA, 23–25 October 2012.

Witlox, H. W. M., & Oke, A. (2008). Verification and validation of consequence models for accidental releases of hazardous chemicals to the atmosphere. *Proc. IChemE Hazards XX Symposium & Workshop*, Manchester, UK, 15–17 April 2008.

Witlox, H. W. M., Stene, J., Harper, M., & Nilsen, S. H. (2011). Modelling of discharge and atmospheric dispersion for carbon dioxide releases including sensitivity analysis for wide range of scenarios. *Energy Procedia*, *4*, 2253–2260.