# Compensating for model uncertainty in probabilistic fire simulations.

May 8, 2019

**Abstract**

In the probabilistic simulations, the main aim is to search the output distribution for the given range of input uncertainties. The problem is that the simulated output distribution can be the result of uncertainties other than the input uncertainties. In deterministic analysis, the output distribution is the result of input uncertainty and modelling uncertainty. Whereas in stochastic analysis the output distribution is the result of input uncertainty, modelling uncertainty and sampling uncertainty. How these uncertainty types combine in the context of reliability analyses is not well understood. In this work, using arbitrary examples, we present the connection between various uncertainties involved in the probabilistic simulations. We study how the modelling and sampling uncertainty propagates to the predicted output distributions. We also propose a simple uncertainty model helpful for the quantification and treatment of modelling uncertainties in the probabilistic simulations. The proposed correction method is shown to improve the accuracy of the output distributions. In practical applications, this would lead to more accurate estimates of failure probabilities.

**Keywords:** Uncertainty propagation, Compartment fire, Modeling uncertainty

## 1 Introduction

All numerical models have a certain modeling uncertainty; i.e. the model cannot capture the actual physical phenomenon perfectly. For a particular output, the modeling uncertainty should be quantified in a meaningful way. In fire safety engineering, the most common practice is to express it as a measure of systematic and random deviation from the experimentally observed value. For example, in the validation guide of FDS, the modeling uncertainty is presented for various output quantities. The data obtained from numerous fire experiments are compared with the corresponding model simulations and the model uncertainty is quantified in terms of systematic bias and the second central moment of random errors. These two parameters represent the trending error property of the model, hence can be used to estimate the prediction uncertainty resulting from using the tool [1, 2].

The fire simulation tools have been reported to be used for the probabilistic analysis. For example, Matala [3] used FDS to study the performance of cables in the tunnel fires, Hietaniemi [4] used it to study the performance of load-bearing wood beams in the building fires, Ayala [5] used it for the stochastic simulations of atrium fires, and Anderson [6] used the CFAST zone model to estimate the community-averaged extent of fire damage in homes. The main task in such an

analysis is to calculate the output uncertainty corresponding to the given input uncertainty. The term "output uncertainty" can be used in the case of non-parametric analysis as well but should not be confused with the one used for the parametric analysis. McGrattan [7] presents a method to estimate the output uncertainty based on the available information of the model uncertainty. In his method, the output uncertainty is the interpretation of normally distributed random errors around a single unbiased output. In other words, it is simply the representation of the possible modeling uncertainty resulting from using the tool. In the parametric analysis, the output uncertainty is rather the desired quantity, and should not be dependent on the modeling uncertainty but only the input parameter uncertainty. The problem not addressed in the above-mentioned and similar other studies is that the stochastically inferred output uncertainty is inevitably a combination of both input and modeling uncertainties, being possibly very different from the true output uncertainty [8, 9].

In this study, we present an uncertainty model that can be used to obtain the true output uncertainty from the stochastically simulated one. We use the model to illustrate how the model uncertainty propagates together with parameter uncertainty. We present that the model uncertainty metrics can be used to statistically compensate for their effect in a probability calculation.

# 2 Uncertainty modeling

## 2.1 Parameter Uncertainty

If the inputs of a mathematical model are uncertain then the outputs will be uncertain too. This uncertainty propagation depends upon the characteristics of the model itself. The expression of uncertainty in output, $T = f(\mathbf{X})$, $f$ being continuous and one time differentiable function, can be derived by Taylor expanding $T$ about its mean and utilizing the definition of standard deviation in $T$ [10]. The first order approximation is,

$$\sigma_T^2 = \mathbf{J}^{\mathrm{T}} \mathbf{\Sigma}^{\mathbf{X}} \mathbf{J}, \tag{1}$$

where $\sigma_T^2$ represents variance in $T$, $\mathbf{\Sigma}^{\mathbf{X}}$ is variance-covariance matrix of the input vector, $\mathbf{X}$, and $\mathbf{J} = (J_1, J_2, J_3...)$, $J_i = \partial f / \partial X_i$. If the input variables, $\mathbf{X}$, are independent of each other then the Eq 1 would simply reduce to

$$\sigma_T^2 = \left|\frac{\partial f}{\partial X_1}\right|^2 \sigma_{X_1}^2 + \left|\frac{\partial f}{\partial X_2}\right|^2 \sigma_{X_2}^2 + \left|\frac{\partial f}{\partial X_3}\right|^2 \sigma_{X_3}^2 + \quad ... \tag{2}$$

Figure 1 depicts the uncertainty propagation for a simple model, $T = \pi X$. A normally distributed output, $T \sim \mathcal{N}(\pi 10, \pi^2)$, is obtained for a normally distributed input, $X \sim \mathcal{N}(10, 1)$. For complex and non-linear problems, such derivation is mathematically challenging, therefore, stochastic methods are adopted. Some examples of stochastic methods are Monte-Carlo (MC), Latin Hypercube Sampling (LHS) and Fourier Amplitude Sensitivity Test (FAST) [11–13].

## 2.2 Combining model and parameter uncertainty

The model uncertainty can be decomposed into two components: systematic bias and random error [1]. The systematic bias is assumed to be a measure of the multiplicative factor by which

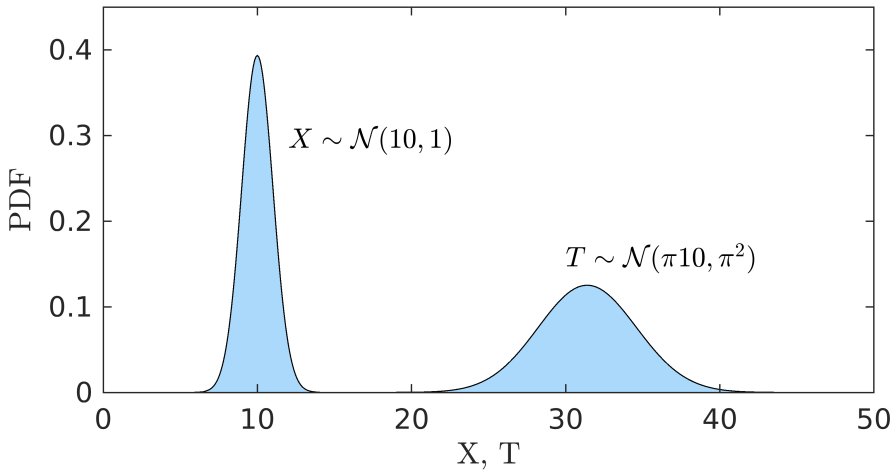


Figure 1: Input and output distribution for $T = \pi X$. The mean and variance of $X$ is 10 and 1 respectively.

the observed output is away from the true value. On average, it is the ratio of observed and true output. The random error is assumed to be an additive error that makes the observed output to fluctuate around the true value. We assume that these parameters can be determined for each output parameter, and are constants for a specific type of fire scenario.

The output for a simulation model, $T = f(X)$, with systematic bias, $\delta$, and random error, $\epsilon$, is

$$\hat{T} = \delta \cdot T + \epsilon, \tag{3}$$

where $\hat{T}$ is the simulated quantity and $T$ is the true quantity. Here, the $T$ and $\epsilon$ are independent and the mean of $\epsilon$ is zero. For such conditions, the mean and variance of the observed quantity can be written as,

$$\mu_{\hat{T}} = \delta \cdot \mu_T \quad \text{and} \quad \sigma^2_{\hat{T}} = \delta^2 \cdot \sigma^2_T + \sigma^2_\epsilon. \tag{4}$$

Where $\mu_T$ and $\sigma^2_T$ are the mean and variance of the true quantity and $\sigma^2_\epsilon$ is the variance of the random error. The derivation for these expressions can be found in the Appendix A.

For a normally distributed output, $T$, Table 1 lists the expressions of distributions in the presence or absence of model uncertainty. Figure 2 shows the histogram plots for specific values of $\delta$ and $\sigma_\epsilon$. The left figure compares the effect of only the bias, the middle one compares the effect of only the random error, and the right one compares the effect of both. Figures show that the bias simply shifts the distribution, while the random error widens it.

Table 1: The output distribution in presence or absence of error.

| S.N. | Distribution of $\hat{T}$ | Condition | Description |
|---|---|---|---|
| 1 | $\mathcal{N}(\mu_T,\, \sigma^2_T)$ | $\delta = 1,\, \sigma_\epsilon = 0$ | No model uncertainty (Blue). |
| 2 | $\mathcal{N}(\delta \cdot \mu_T,\, \delta^2 \cdot \sigma^2_T)$ | $\delta \neq 1,\, \sigma_\epsilon = 0$ | Presence of bias. |
| 3 | $\mathcal{N}(\mu_T,\, \sigma^2_T + \sigma^2_\epsilon)$ | $\delta = 1,\, \sigma_\epsilon \neq 0$ | Presence of random error. |
| 4 | $\mathcal{N}(\delta \cdot \mu_T,\, \delta^2 \cdot \sigma^2_T + \sigma^2_\epsilon)$ | $\delta \neq 1,\, \sigma_\epsilon \neq 0$ | Presence of both. |

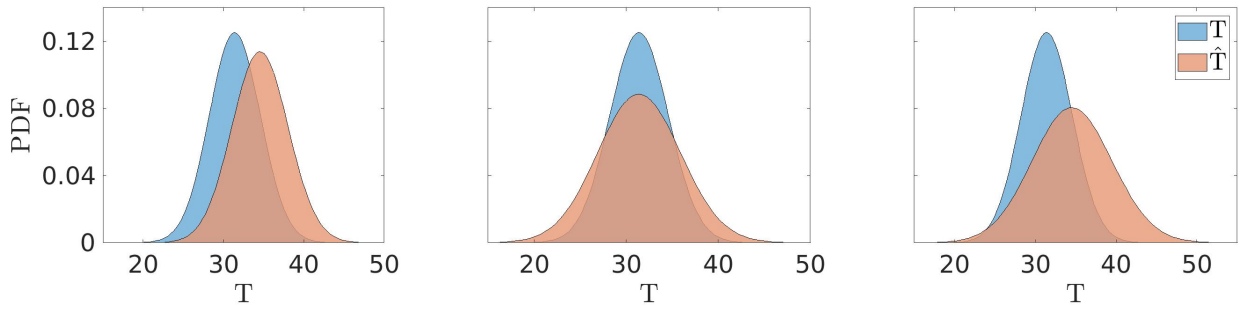


Figure 2: The simulated and true output distribution for, Left: $\delta = 1.1$, $\sigma_\epsilon = 0$, Middle: $\delta = 1$, $\sigma_\epsilon = \pi$ and Right: $\delta = 1.1$, $\sigma_\epsilon = \pi$.

## 2.3 Correction of output distribution

If the prior information of $\delta$ and $\sigma_\epsilon$ is available, one can correct the simulated output towards the true one. The expression of corrected output is

$$T = \frac{1}{\delta}\left[\mu_{\hat{T}} + \left(\hat{T} - \mu_{\hat{T}}\right)\sqrt{1 - \left(\frac{\sigma_\epsilon}{\sigma_{\hat{T}}}\right)^2}\right], \tag{5}$$

where $T$ is the corrected realization corresponding to the observed realization, $\hat{T}$. This expression is derived Eq 4 and the derivation can be found at [14]. We illustrate the correction method using

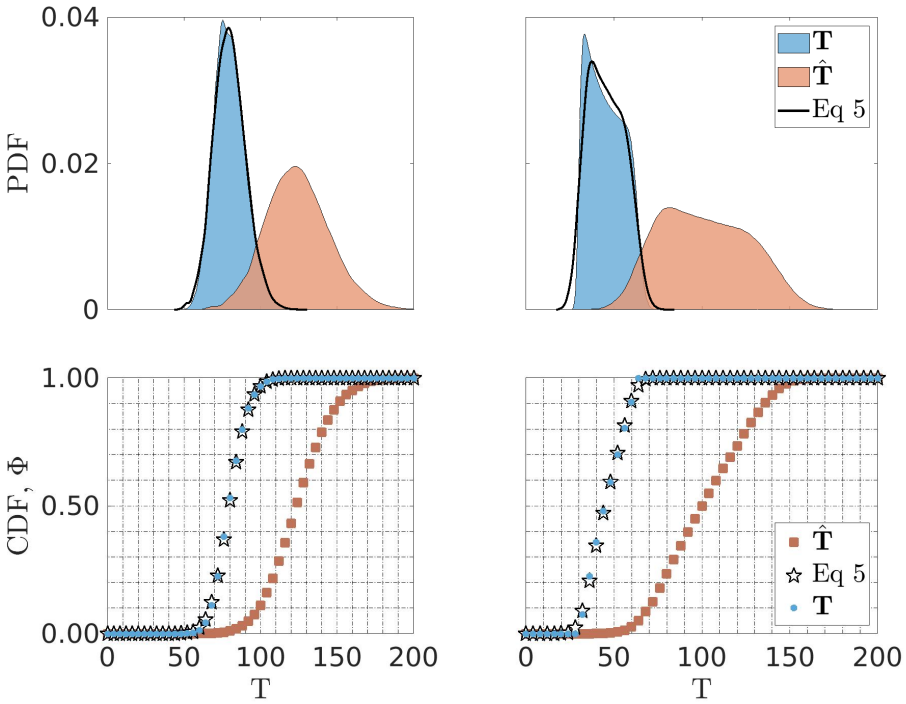


Figure 3: Upper: The true, $T$, simulated, $\hat{T}$, and corrected distributions. Lower: Corresponding cumulative density functions.

two arbitrarily chosen examples [15]. In one of the examples, both the simulated and the true

distribution are Gaussian, while in the remaining one, the distribution shape is irregular. First, we calculate the correction parameters, $\delta$ and $\sigma_\epsilon$, by comparing the simulated and true values,

$$\delta = \frac{\mu_{\hat{T}}}{\mu_T}, \quad \text{and} \quad \sigma_\epsilon = \left[ \frac{1}{N-1} \sum_{i=1}^{N} \left( \hat{T}_i - \delta \cdot T_i \right)^2 \right]^{\frac{1}{2}}, \tag{6}$$

where $\hat{T}_i$ and $T_i$ are the $i^{\text{th}}$ realization of the simulated and the true quantity respectively and $N$ is the sample size. Then, using the correction parameters we estimate the true shape from the simulated one.

Figure 3 shows the true, simulated and corrected distributions along with CDF. In the upper plots, the continuous line represents the distribution generated using Eq 5. Plots indicate that the corrected distribution matches well with the true distribution. The maximum difference between the CDF of true and the corrected distribution is $\sim$ 0.01. The complete trace-backing is not possible because the random error that occurred per realization cannot be known.

## 2.4 Sampling uncertainty

In the stochastic analysis, the inferred moments and the probabilities depend upon the sample size and sampling method. This is known as sampling uncertainty. Figure 4 illustrates such uncertainty using one of the examples presented in the previous section. The simulated distribution $\hat{T}$, corrected distribution, $T$, and the 95 percent fractiles values, $z_{95}$, are presented for sample sizes $N$=100, 1000 and 10000. Higher sample size well represents the distribution and $z_{95}$ values increases with the increase in the sample size.

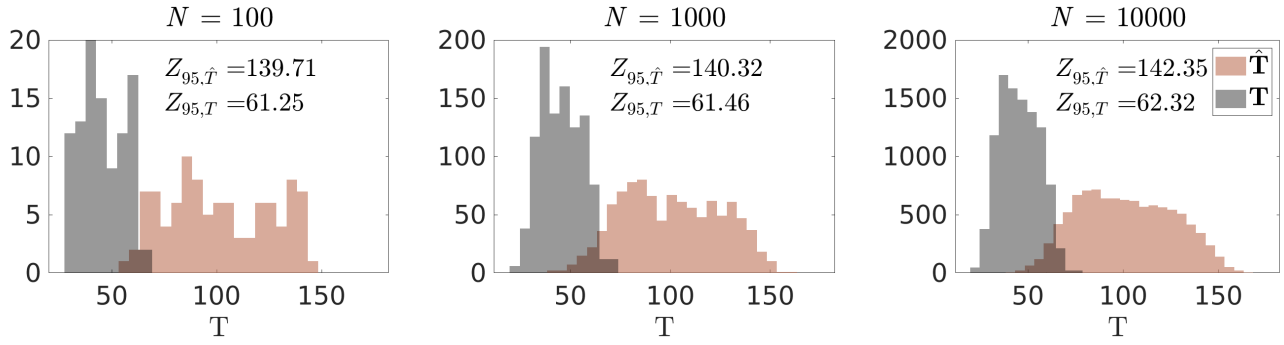


Figure 4: The distributions of simulated values, $\hat{T}$, corrected values, $T$, and 95 percent fractiles for three different sample sizes N=100, 1000 and 10000.

The sampling uncertainty can be presented as $\pm$ bounds from the corrected value. For example, if the probability inferred from the corrected distribution is $p$, then the probability is $p \pm \Delta p$, where $\Delta p$ is the sampling uncertainty. The sampling uncertainty for simple MC simulation having sample size $N$ is $z_a\sqrt{p(1-p)/N}$, where $z_a$ is a multiplier number that determines the level of confidence [16]. For 99% level of confidence $z_a$ is 2.58. For LHS, such analytical expression is not available, and a separate convergence analysis is needed. Figure 5 shows the result of the convergence analysis carried out for the distributions presented in Figure 4. The left plot shows $z_{95}(N)$. The right plot shows their difference with the converged value, $z_{95}(N = 10000)$, and the maximum bound represents the sampling uncertainty. With $N = 1000$, the corrected $z_{95}$ and the sampling uncertainty are 61 and 2 respectively. This means the 95 percent fractiles value is 61$\pm$2.

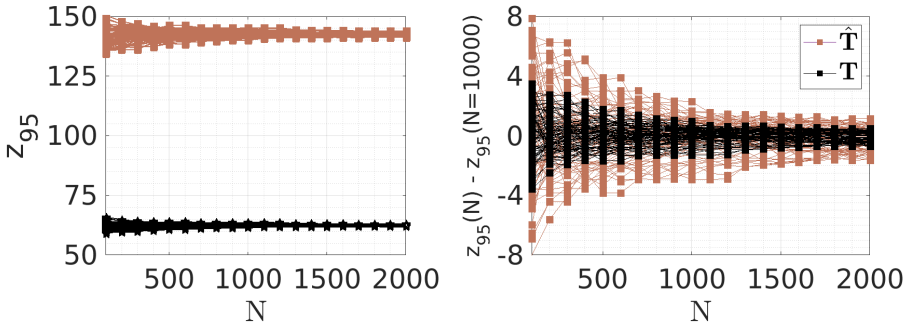


Figure 5: Left: The 95 percent fractiles value, $z_{95}$, of the simulated, $\hat{T}$, and corrected, $T$, distributions for different sample size, $N$. Right: The difference of $z_{95}(N)$ and the converged value, $z_{95}(N = 10000)$.

# 3 Stochastic Analysis

As an example of probabilistic analysis, we simulate fire in an enclosure and compare wall temperatures for a range of inputs listed in Table 2. The enclosure size is $10 \times 7 \times 5\text{m}^3$, and the temperatures were compared on the side wall at 1.3 m height and 4.5 m distance from back wall. The sampling size, N, is 100 and the sampling method is LHS. The selected fire type is t-square

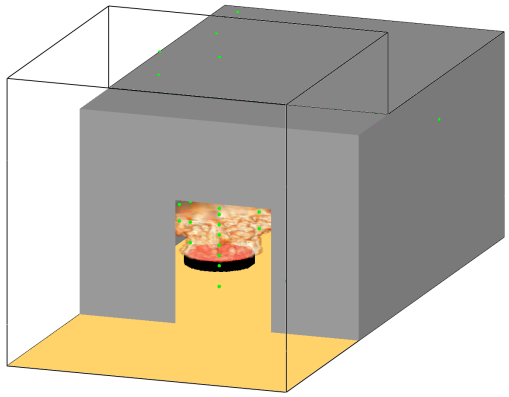

Figure 6: FDS representation of the compartment fire simulation.

fire. For such fire, HRR is calculated using fire growth time, $t_g$, and peak HRR as

$$\text{HRR}(t) = \min\left(1000\left(\frac{t}{t_g}\right)^2, \text{maxHRR}\right)[kW], \tag{7}$$

where $t$ is time in second.

Figure 7 compares the predicted and corrected probability density for wall temperatures. The correction is based on Eq 5 and the model uncertainty values, $\delta = 1.15$, and $\widetilde{\sigma}_\epsilon = 0.16$ obtained from [14]. Figure 8 shows the contour plot for the CDF, $\Phi$, of wall temperatures. The vertical axis shows the temperature range, the horizontal axis shows the time and the embedded text show the $\Phi$ values. The left plot shows the predicted values and right plots show the corrected values. Assuming that the wall fails when it crosses a given temperature threshold, the failure probability would be the fraction of the number of the test cases in which the wall temperature rises above

Table 2: Mean, range and the type of distribution representing the input stochastic.

| Input parameters | Distribution | Mean | Lower value | Upper value | Unit |
|---|---|---|---|---|---|
| Maximum HRR | Uniform | - | 950 | 5400 | [kW] |
| Growth time, $t_g$ | Triangular | 75 | 30 | 150 | [s] |
| Fuel layer thickness | Uniform | - | 20 | 50 | [mm] |
| Pool Diameter | Uniform | - | 0.7 | 1.6 | [m] |
| Pool location, x | Uniform | - | 1.5 | 8.5 | [m] |
| Pool location, y | Uniform | - | 1.5 | 3.5 | [m] |
| Opening door width | Uniform | - | 1.2 | 2.4 | [m] |

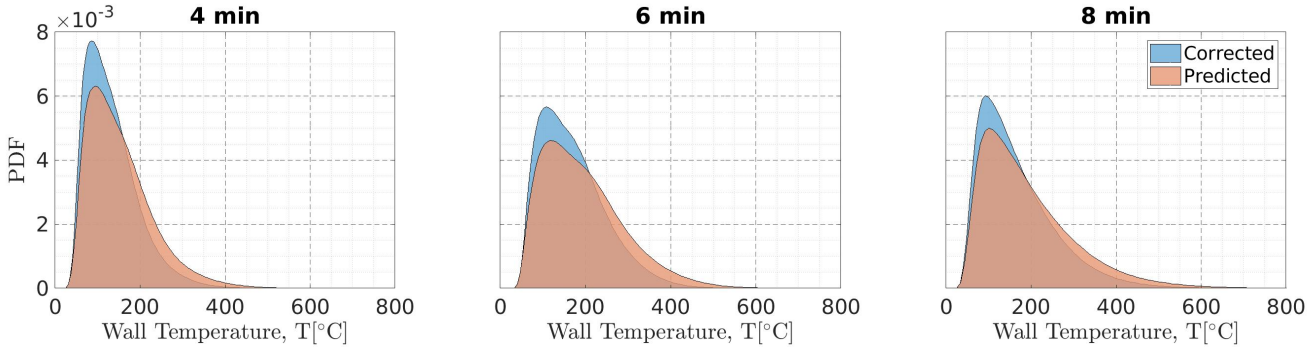


Figure 7: Predicted and corrected probability density of wall temperatures at different times.

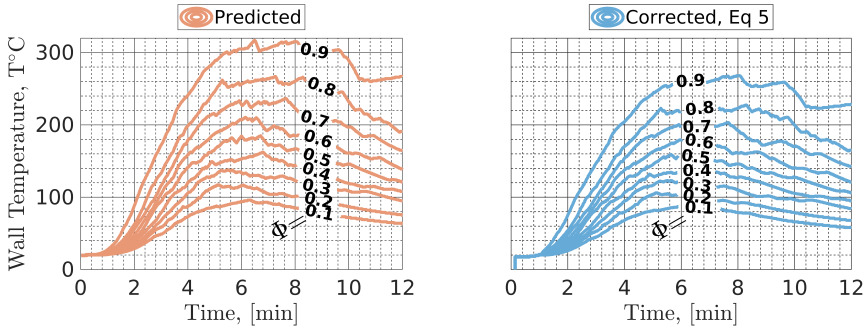


Figure 8: CDF, $\Phi$, of wall temperatures. Left: Predicted. Right: Corrected.

this threshold. From the above CDF plots one can infer the failure probability. For example, the predicted probability that the wall temperature rises above 100 °C before 6 minute is 1 - 0.1 $\sim$ 0.9, where as the corrected probability is 1 - 0.2 $\sim$ 0.8. Similarly the predicted probability for wall to rise above 200 °C before 6 minute is 1 - 0.6 $\sim$ 0.4 and the corrected probability is 1 - 0.7 $\sim$ 0.3. The predicted probabilities are higher than the measured ones. This is due to bias in the temperature prediction.

# 4 Discussion

The study propose correction, Eq 5, for the stochastically simulated output, $\hat{T}$, based on the requirement that the corrected quantity, $T$, and the random error, $\epsilon$, are independent of each other and the mean of $\epsilon$ is zero. In general, the output and the total error are dependent and the mean

of the total error may not be zero. The significance of the uncertainty model presented in this study is that the total error is decomposed into a dependent constant, i.e., the ratio of simulated and corrected mean, $\delta = \mu_{\hat{T}}/\mu_T$, and a random component, $\epsilon = \hat{T} - \delta \cdot T$, which implies that the mean of $\epsilon$ must be zero.

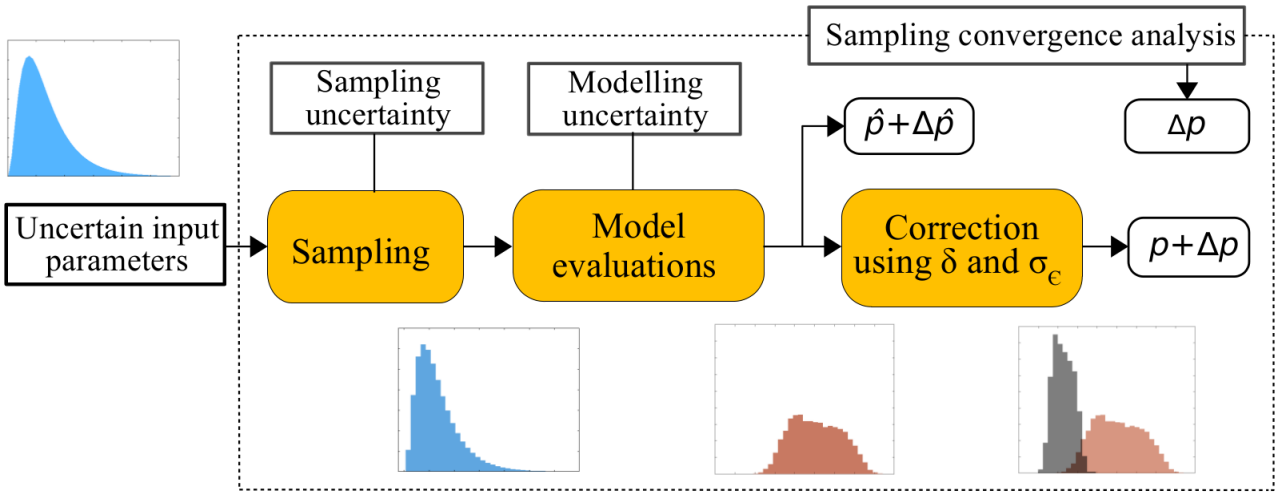


Figure 9: Schematic diagram showing the procedure of uncertainty management in the stochastic simulations.

The proposed correction method handles only one type of different uncertainties appearing in a probabilistic simulation with deterministic models. Other uncertainty types, input uncertainty and sampling uncertainty deserve their own studies when aiming at accurate fire risk analyses. Figure 9 presents an overall procedure for uncertainty management in the stochastic simulation. Estimation of input uncertainty distribution is crucially important for the simulation outcome and can require significant effort if the number of uncertain parameters is high. Luckily, in a nonlinear system, such as fire, the number of dominating input parameters is usually small [17]. For sampling uncertainty, the convergence of the distribution moments can be studied, as explained in Section 2.4. This would be very expensive if a complex numerical method such as CFD is being used. Means to quantify the sampling convergence in LHS could possibly be developed using surrogate models, such as the response surface method.

# 5 Conclusion

In this work, we show that the model uncertainties reported in the context of a model validation can used for correcting the output distributions resulting from parameter (input) uncertainty. The proposed method for the model uncertainty compensation can improve the model predictions significantly provided that the model uncertainty can be generalized. Further work is needed to study the effect of Latin hypercube sampling uncertainty in failure probability calculation.

# Acknowledgements

This work has been funded by the State Nuclear Waste Management Fund of Finland in the scope of the SAFIR-programs, the Finnish Fire Protection Fund (Palosuojelurahasto), Rakennustuotteiden laatu-säätiö and Nordic Nuclear Safety Research (NKS).

# References

[1] McGrattan, K. and Toman, B., Quantifying the predictive uncertainty of complex numerical models, *Metrologia*, 48(3), p. 173, 2011.

[2] McGrattan, K., Hostikka, S., Floyd, J., Baum, H., Rehm, R.G., Mell, W. and McDermott, R., 2010. Fire dynamics simulator technical reference guide volume 3: Validation, *NIST special publication*, 1018(5), 2010.

[3] Matala, A. and Hostikka, S., Probabilistic simulation of cable performance and water-based protection in cable tunnel fires, *Nuclear Engineering and Design*, 241(12), pp. 5263-5274, 2011.

[4] Hietaniemi, J., Probabilistic simulation of fire endurance of a wooden beam, *Structural Safety*, 29(4), pp. 322-336, 2007.

[5] Ayala, P., Cantizano, A., Sánchez-Úbeda, E., and Gutiérrez-Montes, C., The use of fractional factorial design for atrium fires prediction, *Fire technology*, 53(2), pp. 893-916, 2017.

[6] Anderson, A. and Ezekoye, O.A., Quantifying generalized residential fire risk using ensemble fire models with survey and physical data, *Fire Technology*, pp. 1-33, 2018.

[7] McGrattan, K., Peacock, R., and Overholt, K., Validation of fire models applied to nuclear power plant safety, *Fire Technology*, 52(1), pp. 5-24, 2016.

[8] Liang, B. and Mahadevan, S., Error and uncertainty quantification and sensitivity analysis in mechanics computational models, *International Journal for Uncertainty Quantification*, 1(2), 2011.

[9] Olsson, K., Anderson, J. and Lange, D., Uncertainty propagation in FE modeling of a fire resistance test using fractional factorial design based model reduction and deterministic sampling, *Fire safety Journal*, 91(1), pp. 517-523, 2017.

[10] Hamilton, W.C., Statistics in physical science, 1964.

[11] Robert, C. and Casella, G., Monte Carlo statistical methods, Springer Science & Business Media, 2013.

[12] Iman, R., Davenport, J. and Zeigler, D., Latin hypercube sampling (a program users guide): Technical report sand79-1473, *Sandia Laboratories, Albuquerque, NM*, 1980.

[13] Suard, S., Hostikka, S. and Baccou, J., Sensitivity analysis of fire models using a fractional factorial design, *Fire safety journal*, 62(1), pp. 115-124, 2013.

[14] Paudel, D. and Hostikka, S., Propagation of model uncertainty in the stochastic simulations of a compartment fire, *Fire Technology*, https://doi.org/10.1007/s10694-019-00841-9, 2019.

[15] Paudel, D., and Hostikka, S., Propagation of modeling uncertainty in stochastic heat-transfer simulation using a chain of deterministic models, *International Journal of Uncertainty Quantification*, 9(1), pp. 1-14, 2019.

[16] Crow, E.L., Confidence intervals for a proportion, *Biometrika*, 43(3/4), pp.423-435, 1956.

[17] Hostikka, S. and Keski-Rahkonen, O., Probabilistic simulation of fire scenarios, *Nuclear engineering and design*, 224(3), pp.301-311, 2003.