14th International Conference on Structural Safety and Reliability - ICOSSAR'25
1-6 June, 2025 - Los Angeles California, USA
R. Ghanem, E. Johnson, S. Masri and A. Olivier (Eds)

SIGMOID-BASED REGRESSION FOR PHYSICALLY INFORMED TEMPERATURE PREDICTION OF FIRE-EXPOSED PROTECTED STEEL SECTIONS

RAMIN YARMOHAMMDIAN[1], BALSA JOVANOVIC[1] AND RUBEN VAN COILE[1]

[1] Ghent University
9000 Ghent, Belgium
Ramin.yarmohammadian@ugent.be

Key words: surrogate modelling, sigmoid function, physics-informed, steel section, fire

Abstract. Surrogate models offer efficient alternatives to computationally intensive fire simulations. This study evaluates linear and sigmoid-based regression models for predicting the maximum temperature of fire-protected steel elements exposed to physically-based fires. The focus lies on enforcing physical realism through a constrained sigmoid output and a penalty-based loss function targeting conservative predictions. Results demonstrate that while the linear regression model achieves fast and accurate performance with large datasets, the sigmoid model inherently respects physical constraints and offers robust predictions also under limited data.

1 INTRODUCTION

Accurate prediction of steel temperatures during fire exposure is essential for ensuring structural safety and assessing the fire resistance of buildings [1,2]. While detailed numerical models based on physically-based (parametric) fire curves and thermal calculations can provide reliable estimates, their high computational cost makes them impractical for routine design checks, sensitivity analysis, or probabilistic methods.

Surrogate models offer a more efficient alternative by approximating complex simulations with simplified, faster-running models. These techniques have been widely used in other engineering fields, including aerodynamics, heat transfer, and materials science, where they reduce computational effort without compromising predictive capability [3–5]. Methods such as Kriging [6,7], polynomial chaos expansions [8,9], and neural networks [10] are well-established, but their adoption in structural fire engineering has been limited. Moreover, conventional surrogate models often produce unrealistic or non-physical outputs when applied outside the training domain [11].

To overcome these limitations, physics-informed surrogate models (PISMs) incorporate physical laws and engineering constraints into the modelling process. This approach has shown promise in improving accuracy, consistency, and generalization, particularly in scenarios with limited training data or strong physical dependencies [12]. In structural fire applications, embedding thermodynamic principles or code-based rules into the model helps maintain physically consistent predictions.

Effective PISM development requires identifying key input parameters and their relationship to the output, enabling better model transparency and early detection of errors. Incorporating physical constraints—either through tailored model architectures or custom training objectives—further enhances reliability. This study investigates a sigmoid-based regression model that enforces physical bounds on the predicted steel temperature. This method is evaluated in comparison with a baseline linear regression model using data from a parametric fire scenario.

In the next chapter, the detailed numerical model which is used for data development is explained briefly. Then, influencing features are selected to feed into the models.

2 DESCRIPTION OF THE PHYSICAL MODEL

The physical model constitutes a methodology to calculate the maximum temperature attained by protected steel elements exposed to a compartment fire. The compartment fire is modeled through the Eurocode parametric fire curve (EPFC) [13]. The EPFC is an engineering model for physically-based fire exposure, incorporating variables such as fire load, ventilation, and compartment properties [14]. This approach enables modeling of natural fires that include both heating and decay phases, which are essential for reflecting real fire scenarios [15]. The heating phase of the EPFC was derived by Wickström [16], building upon the "Swedish fire curves", i.e., an energy balance model to simulate ventilation-controlled fires [17]. However, because of modifications introduced during the codification [16,18] discontinuities in temperature occur at transitions between fuel- and ventilation-controlled phases, which can lead to unrealistic temperature jumps [14,19].

Once the gas temperature is determined using the EPFC, the lumped mass approach (as defined in EN 1993-1-2:2005 [20]) is applied to calculate the temperature increase of protected steel sections. The lumped mass approach calculates the temperature increase of the steel by accounting for both the insulation's heat capacity and its conductive resistance. This approach is basically a conservation of energy evaluation which takes into account flux due to radiation and convection from the fire curve and the heat capacity of the protected structural element. Using time steps of 0.5 s, the temperature history of the protected steel element is calculated. Since the maximum temperature reached corresponds to the minimum capacity of the member, the maximum is selected to be predicted by surrogate modelling.

3 SELECTION OF FEATURES FOR SURROGATE MODELLING

Selecting appropriate features is crucial for developing a physics-informed surrogate model. Effective feature selection ensures computational efficiency, supports adherence to physical principles, and minimizes the risk of overfitting. A reduced set of significant parameters simplifies training and reduces data requirements.

From the section and insulation properties, four primary features are selected based on prior knowledge: specific area of the steel cross-section, insulation conductivity, volumetric specific heat of insulation, and insulation thickness—a key design parameter for protecting steel members. The specific surface area is essential as heavier steel elements require more energy to heat up, while higher specific areas lead to faster heating. Conductivity and volumetric

specific heat of insulation determine the temperature gradient within the insulation. Since insulation thickness is a critical design parameter, its second-order term was also included for improved accuracy.
Additional parameters were derived from fire characteristics. Parametric fire curves can be described using two key factors: the equivalent opening factor and fuel load for a standard compartment (10×10 m² area, 3 m height, and 1450 thermal resistance of the walls)[21]. Another two independent parameters—maximum fire curve temperature and the time to reach this maximum temperature—were included, as they are strongly linked to the heat transfer physics of radiation and convection. Given the physical significance of these parameters, higher-order terms were also considered, with the fourth power of maximum temperature included to capture the strong radiative heat flux effect.
In total, six primary parameters were considered, with higher-order terms included for some, resulting in nine input features. All features are normalized following Eq. (1).

Table 1. Key parameters (features) with their ranges.

Number	Parameter	Unit	Range
1	Specific surface	[1/m]	90–130
2	Conductivity of insulation	[W/m.K]	0.15–0.25
3	Volumetric specific heat	[J/ m^3.K]	5.6e5-1.68e6
4	Insulation thickness	[m]	0.01–0.05
5	Square of the insulation thickness	[m^2]	
6	Time to maximum temperature of the EPFC	[hour]	
7	Square of time to maximum temperature	[$hour^2$]	
8	Maximum temperature in the compartment	[˚C]	
9	Power 4 of max temperature	[˚C^4]	

$$x_{normalized}^{(n)} = \frac{x_i^{(n)} - x_{\ mean}^{(n)}}{x_{\ st}^{(n)}} \tag{1}$$

where $x_{\ mean}^{(n)}$ and $x_{\ st}^{(n)}$ are mean and standard deviation of the n^{th} selected parameters.

After normalization, surrogate models are trained on the training set and evaluated using cross-validation. Predictions are then visualized using the test dataset to assess model accuracy. The optimal model is selected based on a balance of predictive accuracy, computational efficiency, and adherence to physical principles. Each trained model is evaluated using a dataset of 3,000 instances, including training, cross-validation, and test data.

4 IMPLEMENTING LINEAR REGRESSION

Linear regression is one of the foundational techniques in surrogate modelling, thanks to its simplicity and interpretability [22]. It aims to establish a relationship between a dependent variable y (e.g., maximum steel temperature) and selected features X={x_1,x_2,…,x_n} (i.e., the parameters affecting maximum steel temperature under fire exposure). This relationship is modelled as:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \tag{2}$$

Where y is the predicted outcome (temperature of the protected steel element), β_0 is the intercept term and β_i is the coefficient for each feature x_i.

A key challenge in surrogate modelling is balancing underfitting and overfitting. Overfitting occurs when a model performs well on training data but lacks generalizability to unseen data. Regularization mitigates this issue by penalizing large coefficients, thereby improving generalization. Here ridge regression is used. Ridge regression introduces an additional penalty term to the cost function (Eq.3,4,5):

$$J_R = \frac{1}{2m}\sum_{i=1}^{m}(\beta^i)^2 \tag{3}$$

$$J(\boldsymbol{\beta}) = \frac{1}{2m}\sum_{i=1}^{m}(\tilde{y}^i - y^i)^2 \tag{4}$$

The total cost function with regularization is then:

$$J_{total} = J_R + J(\beta) \tag{5}$$

Using the 9 selected parameters described in Section 3, linear regression with Ridge regularization was implemented. Model optimization was conducted using gradient descent, which iteratively minimized the total cost function by adjusting the model parameters. The training process was governed by a convergence criterion: when the change in weights (measured in L1 norm) and the change in training cost fell below a defined tolerance (10^{-6}), the algorithm terminated.

The evolution of both training and cross-validation cost values across iterations is shown in Figure 1(left). As illustrated, both costs consistently decrease with each iteration, indicating effective learning and regularization.

In addition to monitoring the cost values, Figure 1(right) presents the evolution of the model coefficient for each parameter throughout the optimization process. Since all features were normalized prior to training, the relative magnitudes of these weights provide insight into the influence of each parameter on the model's predictions. Among the features, $t_{\max}$, representing the time at which the compartment reaches its peak temperature, emerged as the most influential. Its substantial positive weight indicates that prolonged fire durations are strongly associated with higher temperatures in protected steel elements—a result that aligns well with physical expectations.

Another key feature is the insulation thickness (d_ins), which consistently exhibits a negative weight across iterations (see Figure 1(right)). Increasing insulation thickness effectively reduces heat transfer to the steel, resulting in lower maximum temperatures which is physically meaningful.

The weight trajectories also offer valuable transparency. They serve as a diagnostic tool for evaluating whether the learned relationships align with engineering judgment and known physical principles. For example, the strong influence of fire duration and the mitigating role of insulation validate the model's internal logic.

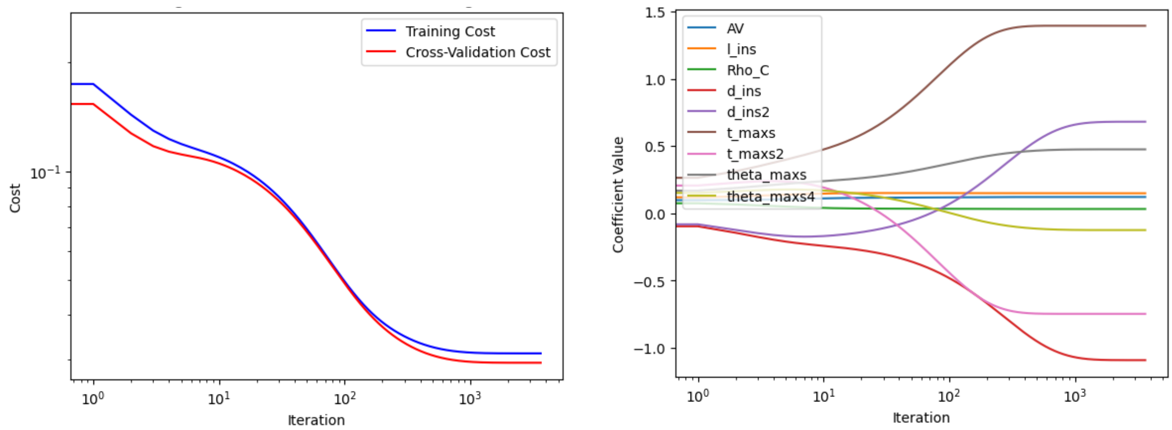

Figure 1. training and cross-validated cost versus iteration (left), coefficient ($\boldsymbol{\beta_i}$) of each parameter in linear regression in each iteration (right)

Figure 2 presents the comparison of maximum steel temperatures predicted by the trained linear regression model against the true values for test data. Notably, just one of the predicted temperatures exhibited errors exceeding 100˚C, demonstrating the model's robustness across the test dataset.

However, while the reference model achieved commendable accuracy overall, certain boundary points—representing very high or very low temperatures—exhibited lower precision. These deviations suggest that the model may have limitations in accurately capturing extreme values, which could be explored further in future refinements.

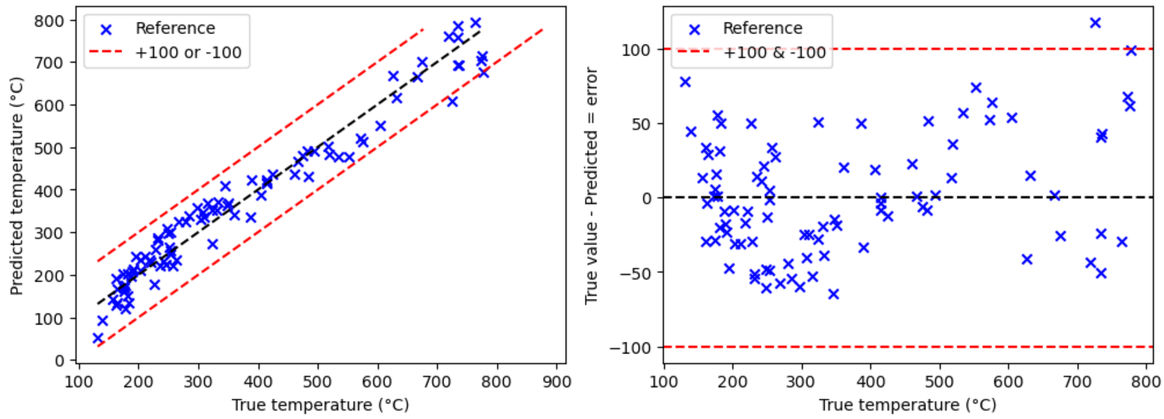

Figure 2. True and predicted values of temperature (left) and residual (right) for protected steel elements in the Reference model.

The R-squared value for predictions in test data is 0.956. Outliers were defined as points with an absolute residual value exceeding 50°C. Within the test dataset of 90 points, 24 outliers were identified. To investigate potential trends not captured by the model, the distribution of outliers is analysed across the range of selected parameters. This visualisation aimed to uncover any relationships between the outliers and the input parameters. For most of 9 parameters, the outliers appear randomly distributed across the selected range. However, for $t_{\max}$, a notable clustering of outliers is observed near the lower boundary of the range. This pattern suggests that the time to reach the maximum temperature, which is indicative of the fire's duration, plays a significant role in the model's performance.

5 SIGMOID REGRESSION

Special functions like the sigmoid function can help surrogate models adhere to physical constraints. A key feature of the sigmoid function is that it outputs values between 0 and 1, making it useful for ensuring positive values or enforcing natural bounds.
In this case, the maximum fire temperature should always be higher than the protected steel element's maximum temperature. To enforce this constraint, the sigmoid function was applied:

$$T_{maxfire} - T_{\max\_protected\_steel} > 0 \tag{6}$$

$$1 > f(x) = \frac{T_{maxfire} - T_{\max\_protected\_steel}}{T_{maxfire}} > 0 \tag{7}$$

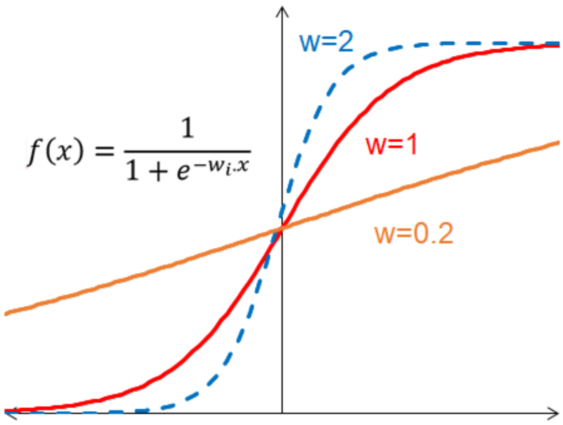

Figure 3. Sigmoid with different weights.

For this model, the same 9 input parameters were used, with each parameter normalized between 0 and 1 to improve training.

$$f(x) = \frac{1}{1+e^{-x}} \tag{8}$$

$$f(x) = \frac{1}{1+e^{-\sum_1^9 xi.wi}} \tag{9}$$

The sigmoid model effectively enforces physics-based constraints, ensuring stable and realistic predictions that align with fire behaviour. The R-squared value for the test data is 0.947, which is comparable to that of the reference model. One significant disadvantage of this method is its relatively slow convergence compared to the reference model. For instance, training this model on 3,000 data points takes approximately 90 seconds, while training the reference model on the same data takes about 1 second. Therefore, there is in this case a trade-off between the extent to which constraints must be adhered to and the speed of surrogate modelling.

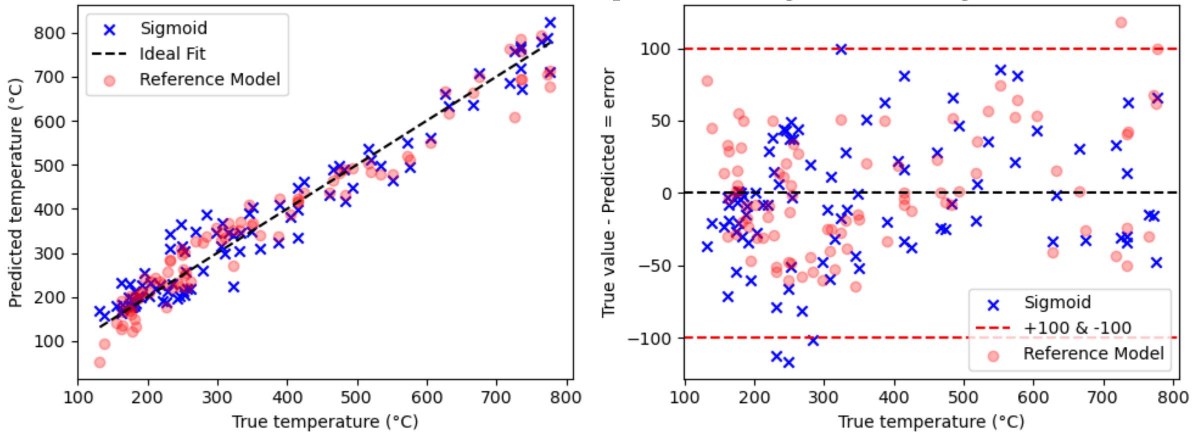

Figure 4. Prediction versus true values (left) and residual values (right) for Sigmoid regression.

6 COMPARISON OF MODELS

To evaluate and compare the performance of the two surrogate models developed, two primary criteria were considered: (i) the number of unphysical predictions, and (ii) the coefficient of determination (R^2). These metrics were analyzed as a function of the number of training data points to assess how model accuracy and physical consistency evolve with data availability.

A total of 100,000 test samples were generated with the physical model using Latin Hypercube Sampling (LHS), which provides a stratified and space-filling distribution across the full range of input parameters. This approach was selected to ensure a comprehensive and unbiased evaluation of model behavior.

Figure 5 (left) illustrates the number of unphysical predictions made by each model as the training set size increases. An unphysical point is defined as a case where the predicted maximum temperature in the protected steel element exceeds the peak temperature within the fire compartment—an outcome that contradicts fundamental thermal principles.

The results indicate that while both models can achieve high predictive accuracy, care must be taken to ensure physical realism. For the linear regression model, unphysical outputs were observed when the training set contained fewer than 1,000 samples. Once the dataset exceeded this threshold, the model produced only physically consistent predictions. This highlights a limitation often encountered in surrogate modelling: although simple models may yield accurate predictions in terms of statistical fit, they may still violate physical constraints when trained on insufficient data. Given the high computational cost of generating detailed simulations, this issue is particularly relevant in practice, where large datasets are often not feasible.

In contrast, the sigmoid regression model, which incorporates a nonlinear mapping and can be designed to enforce physical constraints more naturally, demonstrated a more robust behavior. Even with smaller training datasets, it consistently avoided unphysical outputs. Furthermore, its accuracy for small training sets was comparable to that of linear regression. However, the main drawback of the sigmoid-based model lies in its computational complexity. Training such models is more time-consuming and may present convergence challenges, particularly for large parameter spaces or when tight tolerances are required.

Overall, these results suggest a trade-off between computational efficiency and physical robustness. Linear models are fast and easy to train but may require more data to ensure physical validity, while more complex models like sigmoid regression provide physical guarantees at the cost of increased training time and algorithmic complexity.

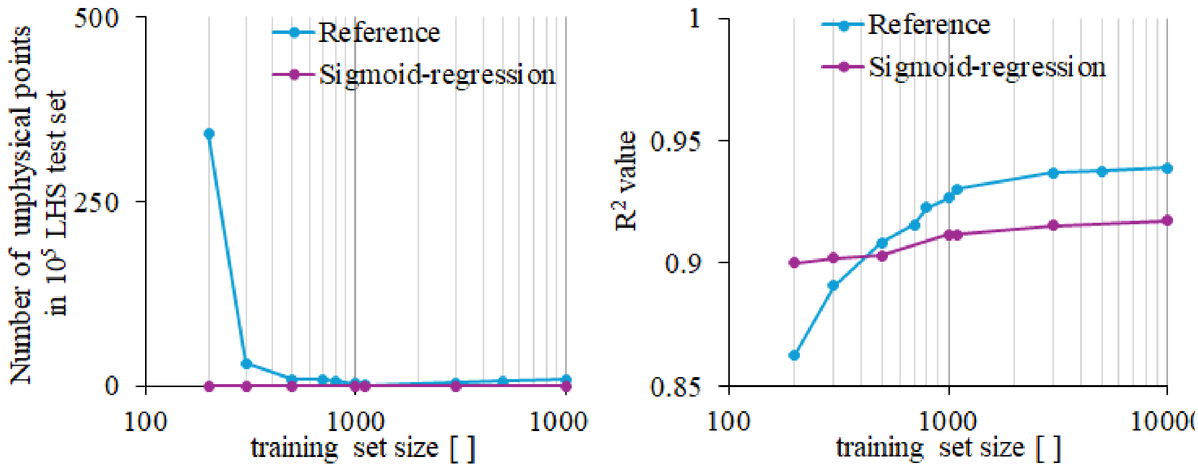

Figure 5. Comparison of models in function of the training set size.

7 Conclusion

This study demonstrated a comprehensive approach to predicting the temperature of protected steel elements under fire exposure by combining physics-based modelling with data-driven surrogate models. The physical model considers an integration of the Eurocode Parametric Fire Curve (EPFC) and lumped mass approach, providing a realistic framework for capturing the thermal behaviour of steel in fire scenarios. Subsequently, surrogate models were trained to allow for physics-based temperature predictions at reduced computational expense. The model adopting traditional Linear Regression was found to be simple and computationally efficient, but less capable of capturing complex, nonlinear relationships, limiting its utility for nuanced fire scenarios. The alternative model adopting Sigmoid Regression consistently ensured physically valid outputs, maintaining temperature predictions within the bounds defined by fire physics. It demonstrated robustness even with limited training data, making it suitable for safety-critical applications. The results demonstrate that physics-informed choices in surrogate modelling architecture can greatly enhance the reliability of the surrogate model, also in situations with a limited training set.

Acknowledgements

Funded by the European Union (ERC, AFireTest, 101075556). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European

Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

REFERENCES

[1] T.T. Lie, K.H. Almand, A Method To Predict the Fire Resistance of Steel Building Columns, Engineering Journal 27 (1990) 158–167. https://doi.org/10.62913/ENGJ.V27I4.560.
[2] A. Abu, R. Shi, M. Jafarian, K. LaMalva, D. Hopkin, Steel and Composite Structures, (2021) 189–233. https://doi.org/10.1007/978-3-030-77123-2_7.
[3] T.D. Robinson, M.S. Eldred, K.E. Willcox, R. Haimes, Surrogate-based optimization using multifidelity models with variable parameterization and corrected space mapping, AIAA Journal 46 (2008) 2814–2822. https://doi.org/10.2514/1.36043.
[4] L. Sun, H. Gao, S. Pan, J.X. Wang, Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data, Comput Methods Appl Mech Eng 361 (2020). https://doi.org/10.1016/j.cma.2019.112732.
[5] M.Z. Naser, Observational Analysis of Fire-Induced Spalling of Concrete through Ensemble Machine Learning and Surrogate Modeling, Journal of Materials in Civil Engineering 33 (2021) 04020428. https://doi.org/10.1061/(asce)mt.1943-5533.0003525.
[6] X. Shang, L. Su, H. Fang, B. Zeng, Z. Zhang, An efficient multi-fidelity Kriging surrogate model-based method for global sensitivity analysis, Reliab Eng Syst Saf 229 (2023) 108858. https://doi.org/https://doi.org/10.1016/j.ress.2022.108858.
[7] Y. Zhou, Z. Lu, An enhanced Kriging surrogate modeling technique for high-dimensional problems, Mech Syst Signal Process 140 (2020) 106687. https://doi.org/https://doi.org/10.1016/j.ymssp.2020.106687.
[8] Z. Liu, Y. Choe, Data-driven sparse polynomial chaos expansion for models with dependent inputs, Journal of Safety Science and Resilience 4 (2023) 358–365. https://doi.org/https://doi.org/10.1016/j.jnlssr.2023.08.003.
[9] J. Yang, B. Faverjon, H. Peters, N. Kessissoglou, Application of Polynomial Chaos Expansion and Model Order Reduction for Dynamic Analysis of Structures with Uncertainties, Procedia IUTAM 13 (2015) 63–70. https://doi.org/https://doi.org/10.1016/j.piutam.2015.01.017.
[10] E.W.M. Lee, Application of Artificial Neural Network to Fire Safety Engineering, in: L.C. Jain, C.P. Lim (Eds.), Handbook on Decision Making: Vol 1: Techniques and Applications, Springer Berlin Heidelberg, Berlin, Heidelberg, 2010: pp. 369–395. https://doi.org/10.1007/978-3-642-13639-9_15.
[11] M.Z. Naser, A Look into How Machine Learning is Reshaping Engineering Models: the Rise of Analysis Paralysis, Optimal yet Infeasible Solutions, and the Inevitable Rashomon Paradox, 2025. https://doi.org/10.48550/arXiv.2501.04894.
[12] M.Z. Naser, Mechanistically Informed Machine Learning and Artificial Intelligence in Fire Engineering and Sciences, Fire Technol 57 (2021) 2741–2784. https://doi.org/10.1007/s10694-020-01069-8.
[13] EN 1991-1-2: Eurocode 1: Actions on structures - Part 1-2: General actions - Actions on structures exposed to fire, 2002.

[14] F. Put, A. Lucherini, B. Merci, R. Van Coile, Model uncertainty in a parametric fire curve approach: A stochastic correction factor for the compartment fire load density, Fire Saf J 144 (2024) 104113. https://doi.org/https://doi.org/10.1016/j.firesaf.2024.104113.
[15] A. Lucherini, B. Jovanovic, R. Van Coile, B. Merci, Background and limitations of the Eurocode parametric fire curves, including the fire decay phase, 2021.
[16] U. Wickström, Temperature Calculation of Insulated Steel Columns Exposed to Natural Fire, Fire Saf J 4 (1981) 219–225. https://doi.org/https://doi.org/10.1016/0379-7112(81)90024-2.
[17] S.E. Magnusson, S. Thelandersson, Temperature - Time Curves of Complete Process of Fire Development, Bulletin of Division of Structural Mechanics and Concrete Construction (1970).
[18] D. Hopkin, R. Van Coile, C. Hopkin, K. LaMalva, M. Spearpoint, C. Wade, Design Fires and Actions BT - International Handbook of Structural Fire Engineering, in: K. LaMalva, D. Hopkin (Eds.), Springer International Publishing, Cham, 2021: pp. 75–114. https://doi.org/10.1007/978-3-030-77123-2_4.
[19] S. Reitgrüber, C. Pérez-Jiménez, C. Blasi, J.-M. Franssen, Some comments on the Parametric Fire Model of Eurocode 1, (2006).
[20] Eurocode 3: Design of steel structures-part 1-2: General rules-structural fire design, London: European Committee for Standardisation (2007).
[21] T. Thienpont, R. Van Coile, R. Caspeele, W. Corte, W. De Corte, Burnout resistance of concrete slabs: Probabilistic assessment and global resistance factor calibration, Fire Saf J 119 (n.d.).
[22] D.C. MONTGOMERY, G.G. VINING, E.A. PECK, Introduction to Linear Regression Analysis, Biometrics 69 (2013) 1087. https://doi.org/10.1111/biom.12129.