Proceedings of the 33rd European Safety and Reliability Conference (ESREL 2023)
Edited by Mário P. Brito, Terje Aven, Piero Baraldi, Marko Čepin and Enrico Zio
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-8071-1_P394-cd

Surrogate modelling of risk measures for use in probabilistic safety analysis applications

Sara Asensio
Department of Chemical and Nuclear Engineering, MEDASEGI research group, Universitat Politècnica de València, Valencia, Spain. https://sn.medasegi.webs.upv.es/ E-mail: sasefal@iqn.upv.es

Isabel Martón
Department of Statistics and Operational Research, MEDASEGI research group, Universitat Politècnica de València, Valencia, Spain. E-mail: ismarllu@ upv.es

Ana I. Sánchez
Department of Statistics and Operational Research, MEDASEGI research group, Universitat Politècnica de València, Valencia, Spain. E-mail: aisanche@eio.upv.es

Sebastián Martorell
Department of Chemical and Nuclear Engineering, MEDASEGI research group, Universitat Politècnica de València, Valencia, Spain. E-mail: smartore@iqn.upv.es

Probabilistic Safety Analysis (PSA) is an efficient tool for assessing, maintaining and improving the Nuclear Power Plant (NPP) safety. In the literature, different PSA applications have been identified such as: PSA to support NPP testing and maintenance planning and optimization, PSA as a tool to monitor level of safety or PSA as a predictive evaluation of risk. In general, these applications require analyzing aging trends, updating reliability parameters and maintenance related to safety equipment. An indispensable tool in PSA is the software such as RiskSpectrum or CAFTA which are widely used in NPP. The main problem with the use of commercial software is its lack of flexibility in modelling. In this context, the use of tools like surrogate models, or metamodeling, emerges as a tool that can improve realism in probabilistic safety analysis. In this approach, the PSA code is substituted by a metamodel in order to obtain the risk measures of interest. In this paper, different metamodels have been considered which have been trained to predict the Core Damage Frequency (CDF). The performance of the different models is evaluated using three quality metrics (Root Mean Square Error, Mean Absolute Error and Mean Absolute Percentage Error) which have been evaluated using k-fold cross-validation technique. The results obtained demonstrate the capacity of the metamodels to provide accurate and computationally efficient estimates of CDF.

Keywords: Probabilistic Safety Analysis (PSA), Core Damage Frequency (CDF), machine learning, metamodel.

1. Introduction

Probabilistic Safety Analysis (PSA) is an efficient tool for assessing, maintaining and improving the Nuclear Power Plant (NPP) safety. In the literature, different PSA applications have be identified such as: PSA to support NPP testing and maintenance planning and optimization, PSA as a tool to monitor level of safety or PSA as a predictive evaluation of risk. In general, these applications require analyzing aging trends, updating reliability parameters and maintenance related to safety equipment (IAEA, 2001).

PSA are typically large models developed using event tree representations together with fault tree models. An indispensable tool in PSA is the software such as RiskSpectrum or CAFTA which are widely used in NPP. The main problem with commercial software such as these is its lack of flexibility in modelling. They do not consider failure rate models explicitly depending on aging and the effectiveness of maintenance and asset management policies, nor the effect of surveillance effectiveness on the availability of safety equipment. In addition, the use of the original PSA model can imply a high computational cost. In this context, the use of advanced computational techniques such as

machine learning emerges as a tool that can improve realism in probabilistic safety analysis.

The applicability of artificial intelligence (AI) and machine learning (ML) in operating nuclear plants is explored at NUREG/CR-7294 (Ma et al., 2022). In this report, the authors review the recent applications of AI&ML in different fields of nuclear industry and identify three main technological application fields: 1) Plant safety and security assessment, 2) Plant degradation modelling, fault and accident diagnosis and prognosis and 3) Plant operation and maintenance efficiency improvement.

This paper focuses on the use of supervised ML algorithms to generate a surrogate model, metamodel or emulator, to substitute the PSA code to obtain the risk measures of interest (e.g. Core Damage Frequency (CDF)).

In literature, different surrogate models have been proposed (James et al., 2017). In this paper, four metamodels are considered (Generalized Additive for location, scale and shape, K-Nearest-Neighbor, Support Vector Regression and Extreme Gradient Boosting). The models have been trained to predict the CDF using 10000 simulations obtained by applying a PSA code with 595 input variables corresponding to basic events which have been modelized according to one distribution probability. To evaluate the performance of the different models, three quality metrics have been considered, Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), which have been evaluated with k-fold cross-validation technique. The results obtained demonstrate the capacity of the surrogate models to provide accurate and computationally efficient estimates of CDF. Finally, the selection of the best surrogate model is made using the Taylor diagram.

The paper is organized as follows. First, a brief introduction of PSA is given in section 2. Section 3 presents the different machine learning algorithms used, the performance metrics and Taylor diagram. In section 4 the case of application and the results obtained are showed. Finally, section 5 presents the concluding remarks.

2. Probabilistic Safety Assessment

The PSA of a nuclear power plant is a risk analysis technique that allows to estimate quantitatively the risk of a nuclear power plant. This risk is traditionally defined as the probability of an accident multiplied by the consequences that result from it. PSA is based on the combined use of Event Tree (ET) and Fault Tree (FT). Event trees characterize the success and failure relationships of safety-critical systems, while fault trees are used to calculate the failure probability of systems based on the probability data of the components. This approach has been proven adequate for analyzing accidents in nuclear reactors, whose safety relies on multiple and redundant safety systems (Keller & Modarres, 2005).

The nuclear industry PSA distinguishes three levels. In this paper, th analysis is focused on Level 1, which estimates the frequency of accidents that cause damage to the core, the sum of which is commonly known as the CDF, and at full power where the initiating events take place in the reactor operating mode.

Various authors have pointed out the usefulness of PSA in evaluating the effects of aging on plant risk (Kančev & Čepin, 2012, Martón et al., 2015) and the importance of including reliability and unavailability models of time-dependent safety components that also incorporate the effect of maintenance and testing programs and their corresponding effectiveness (Martorell et al., 2018).

However, standard PSA does not address adequately these issues. For example, current practice in PSA modelling considers constant failure rates instead of explicit age dependent failure rates of safety related equipment.

In addition, they do not formulate explicitly how a large variety of surveillance, maintenance and inspection programs, which are intended to mitigate or at least keep under control the effects of equipment ageing, impact on failure rates. To consider the above aspects a surrogate model could be useful. Thus, the standard PSA model could be replaced by a surrogate model, which would allow formulate explicitly, for example, the effect of surveillance and maintenance programs in equipment ageing.

3. Methods

As pointed out in the introduction, in this paper, the performance of four metamodels are evaluated in order to predict the CDF. These metamodels are: generalized Additive models of Location, Scale and Shape, k-Nearest Neighbors, support vector regression and extreme gradient boosting.

The above metamodels depend on a set of parameters called hyperparameters. To optimize these parameters the cross validation method has been applied.

The brief descriptions of the highlighted methods are presented in the following subsections. In all models, $\boldsymbol{X}$ is the matrix of inputs (independent variables), Y is the vector of output (dependent variable), n the sample size and p the number of inputs.

3.1. *Generalized additive models of location, scale and shape*

Generalized additive models of Location, Scale and Shape (GAMLSS) (Rigby & Stasinopoulos, 2005) are a type of semi-parametric regression that corrects some of the limitations of generalized linear models and generalized additive models allowing the modelling of dependent variables whose distribution does not belong to the exponential family or which present heterogeneity. The GAMLSS are semi-parametric models since, although they require establishing a parametric distribution model of the response variable, modelling the distribution parameters as a function of the independent variables may involve non-parametric functions, such as smoothing functions.

The GAMLSS models assume that the dependent variable (Y_i) has a density function that can be defined by up to 4 parameters (μ, σ, ν and τ) which determine the location, scale and shape (skewness and kurtosis), respectively. These parameters can vary depending on the values of the independent variables. Thus, for $i = 1,2,..,n$ observations of Y_i the conditional density function is given by $f_Y(y_i|\theta^i)$ being $\theta^i = (\mu_i, \sigma_i, \nu_i, \tau_i)$. Then if D represents the Y distribution:

$$Y_i|\mu_i, \sigma_i, \nu_i, \tau_i \sim D(\mu_i, \sigma_i, \nu_i, \tau_i)$$

(Rigby & Stasinopoulos, 2005) define the original formulation of GAMLSS model as:

$$g_k(\boldsymbol{\theta}_k) = \eta_k = \boldsymbol{X}_k\boldsymbol{\beta}_k + \sum_{j=1}^{J_k} h_{jk}(\boldsymbol{x}_{jk}) \quad (1)$$

for $k = 1,2,3,4$ being $g_k(\boldsymbol{\theta}_k)$ a monotonic link function that relates the distribution parameter $\boldsymbol{\theta}_k$ to the predictor η_k, and $h_{jk}(\ )$ is an unknown function of the independent variables $\boldsymbol{x}_{jk}$.

From Eq. (1) other extensions of the GAMLSS model can be obtained. Thus, a non-linear semi-parametric model is obtained by substituting the linear term for non-linear parametric terms:

$$g_k(\boldsymbol{\theta}_k) = \eta_k = h_k(\boldsymbol{X}_k, \boldsymbol{\beta}_k) + \sum_{j=1}^{J_k} h_{jk}(\boldsymbol{x}_{jk}) \quad (2)$$

The different parameters of GAMLSS models are estimated maximizing the penalized likelihood function (Stasinopoulos & Rigby, 2008).

3.2. *k-Nearest Neighbors*

k-Nearest Neighbors (KNN) is a nonparametric regression method which works on the basis that the similar samples are distributed near to each other in feature space. The simplest approximation of KNN regression is based on the evaluation of the average of the dependent variable values of the K nearest neighbors. Other approaches have been proposed in the literature, such as the use of an inverse distance weighted average of the K nearest neighbors.

The basis of this method is to calculate a similarity measure, for example, using distance functions. The most common distance functions are: Euclidean, Manhattan and Minkowski. The normalized Euclidean metric is generally used which is evaluated as:

$$dist(A,B) = \sqrt{\frac{\sum_{i=1}^{p}(x_{1i}-x_{2i})^2}{p}} \quad (3)$$

where A and B are represented by vectors $A = (x_{11}, x_{12}, \dots, x_{1p})$ and $B = (x_{21}, x_{22}, \dots, x_{2p})$.

3.3. *Extreme Gradient Boosting*

The Extreme Gradient Boosting (XGBoost) is one of the most popular boosting tree algorithms for

gradient boosting machine (GBM) (Friedman, 2001). It generates a weak learner at each step of the algorithm, and accumulates it into the final models. Thus, XGBoost uses many additive functions to predict the dependent variable as:

$$\hat{f}(\boldsymbol{X}i) = \sum_{m=1}^{M} f_m(\boldsymbol{X}_i) \quad (4)$$

where f_m is a regression tree and $f_m(\boldsymbol{X}_i)$ represents the score given by the *m-th* tree to the i-th observation.

Objective function used in XGBoost is the sum of the loss function and the regular term which control the accuracy and complexity of the model, respectively:

$$\sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{m=1}^{M} \Omega(f_m) \quad (5)$$

l is the loss function and Ω is evaluated as:

$$\Omega(f_m) = \gamma K + \frac{1}{2}\lambda \|w\|_2 \quad (6)$$

γ and K are parameters controlling penalization for the numbers of leaves K and magnitude of leaf weights w, respectively.

3.4. *Support Vector Regression*

Support Vector Regression (SVR) is a flexible and powerful machine learning algorithm that can handle both linear and non-linear relationships between input and output variables. Given a data training set, the goal of SVR (Vapnik 1995), is to find a function $f(\boldsymbol{x})$ that has at most ε deviation from the dependent variable values y_i for all the training data:

$$f(x) = w\varphi(x) + b \quad (7)$$

where $\varphi(x)$ represents the non-linearity mapping.

The parameters of the Eq. (7) can be obtained solving the optimization problem:

$$min\left(\frac{1}{2}\|w\|^2 + c\sum_{i=1}^{n}(\xi_i, \xi_i^*)\right) \quad (8)$$

$$subject\ to \begin{cases} \left(w \quad \varphi(x_i) + b\right) - y_i \leq \varepsilon + \xi_i \\ y_i - (w^T\varphi(x_i) + b) - y_i \leq \varepsilon + \xi_i^* \end{cases}$$

being $\xi_i > 0$ and $\xi_i^* > 0$ slack variables (Smola & Schölkopf , 2004).

The above optimization problem can be formulated using the Lagrange multipliers as:

$$max\left[\sum_{i=1}^{n}(\alpha_i - \alpha_i^*)y_i - \varepsilon\sum_{i=1}^{n}(\alpha_i - \alpha_i^*) - \frac{1}{2}\sum_{i=1,j=1}^{n}(\alpha_i - \alpha_i^*)\left(\alpha_j - \alpha_j^*\right)\cdot K(x_i, x_j)\right] \quad (9)$$

$$subject\ to: 0 \leq \alpha_i \leq c$$
$$0 \leq \alpha_i^* \leq c$$
$$\sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0$$
$$K\left(x_i, x_j\right) = \varphi(x_i)^T\varphi(x_j)$$

where $K\left(x_i, x_j\right)$ represents a nonlinear kernel function, such as the radial basis function (RBF) kernel, which maps the input from a lower dimension feature space into a higher dimension.

3.5. *Performance metrics*

The performance metrics used in this study to assess the different metamodels are as follows:

- Root Mean Square Error (RMSE) which measures the error between the estimation of the model and the true value

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)} \quad (10)$$

- Mean Absolute Error (MAE) which is a measures of how accurate the predict values are compared to the observed ones.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (11)$$

- Mean Absolute Percentage Error (MAPE) which is an accuracy measure based on the relative percentage of errors.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (12)$$

To select the best surrogate model a Taylor diagram (Taylor, 2001), which summarizes multiple aspects of model performance in a single diagram, is used. The diagram provides a concise statistical summary in terms of correlation, root-mean-square difference and the ratio of their variances which can be used to compare the performance of different surrogate models.

4. Application case

This application case is focused on obtaining a metamodel that represents the Core Damage Frequency (CDF) of a nuclear power plant as a function of the most important basic events. Simulations of a Level 1 PSA of a Power Water

Reactor (PWR) were used to obtain a representative sampling using a commercial PSA code.

As a preliminary step to sampling, the most important basic events with respect to CDF were studied and selected using a commonly used measure in the importance analysis, the Fussell-Vesely measure (Van der Borst and Schoonakker, 2001).

The set of training observations (10000 observations) were used to obtain the different metamodels. The number the inputs is equal to 595 variables corresponding to the most important basic events which have been modelized according to one distribution probability.

Figure 1 shows the histogram, the probability density function estimated, the mean, and 5th and 95th percentiles obtained using the 10000 simulations.

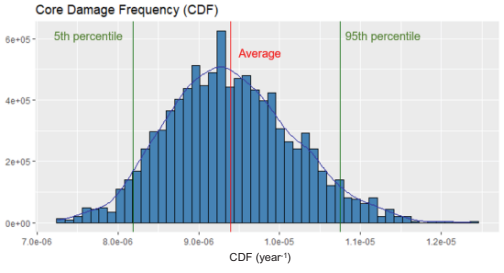

Fig. 1. Histogram and density function of CDF.

The training simulations were randomly partitioned into two groups: 80% of the simulations (8000 runs) were used for training and the 20% (2000 runs) were used for testing. For some metamodels (see section 3) the tuning of some hyper-parameters is required which has been performed using a cross validation strategy.

The performance of the different metamodels on the test set is summarized in Table 1. GAMLSS and SVR have the best results with a very good overall accuracy, as, for these methods, MAPE (%) are approximately equal to 0.5%. The same conclusion is obtained if the values of the RMSE and MAE are analyzed.

Also, the Taylor diagram is used to evaluate each model performance (Figure 2) and select the best. All models, except KNN, have good correlations with observations above 0.95. The green contours indicate the RMS values and it can be seen that in the case of the GAMLSS and SVR the RMS error is less than about 2E-7 hr^{-1}. The standard deviation of the predictions obtained with the surrogate models is proportional to the radial distance from the origin. Again, the two aforementioned models the best behavior since they have similar standard deviations to the observed standard deviation.

Table 1. Performance metrics of the different surrogate models.

Model	RMSE (yr^{-1})	MAE (yr^{-1})	MAPE (%)
GAMLSS	1.732E-07	4.177E-08	4.462E-01
KNN	7.127E-07	5.701E-07	6.102
SVR	5.973E-08	4.771E-08	5.091E-01
XGBoost	1.732E-07	1.379E-07	1.475

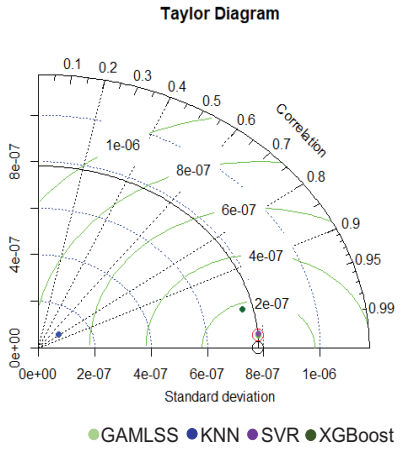

Fig. 2. Taylor diagram displaying a statistical comparison with observations of five surrogate models.

Thus, if the GAMLSS is selected as the surrogate model and 10000 simulations are performed with the model and compared with the observed data a similar behavior is observed. Figure 3 shows the histograms of the observed CDF and predictions using GLM model. Figure 4 shows the scatterplot of the observed and predicted CDF being the Pearson correlation coefficient equal to 0.99, indicating a good model prediction. The same conclusion can be obtained by looking at the box plot shown in Figure 5.

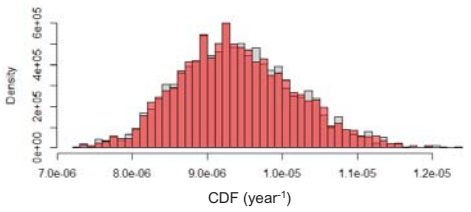

Fig. 3. Histogram of observed vs GAMLSS predictions.

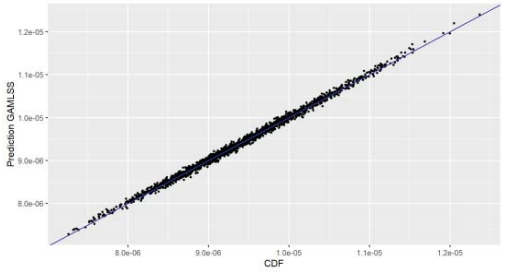

Fig. 4. Observed vs Predicted CDF using GAMLSS model.

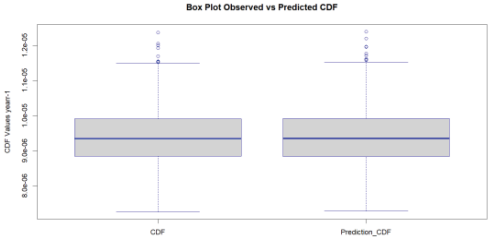

Fig. 5. Box plot of observed and predicted CDF.

10. Concluding remark

This paper presents the capability of the metamodels to obtain a reliable representation of the risk measure of interest, in this application case the core damage frequency (CDF) of a plant, as a function of the most important basic events and to substitute the commercial code PSA by a metamodel in order to obtain the risk measures of interest.

Four metamodels are evaluated in order to predict the CDF : GAMLSS, extreme gradient boosting, k-Nearest Neighbors, and support vector regression. GAMLSS and support vector machine have similar results, obtaining good correlations.

The results obtained in the application case show that surrogate models could be adopted to substitute PSA code which would allow its use in advanced applications of the PSA such as living PSA or Ageing PSA (APSA) providing rapid and usable decision support to decision making.

Acknowledgement

Grant PID2019-110590RB-I00 funded by MCIN/AEI/10.13039/501100011033 “ERDF A way of making Europe”.

References

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.

IAEA-TECDOC-1200. (2001). Applications of probabilistic safety assessment (PSA) for nuclear power plants. IAEA.

James, G. Witten. D. Hastie. T. and Tibshirani. R. (2017). An introduction to statistical learning with applications in R. Springer. New York.

Kančev, D., & Čepin, M. (2012). Uncertainty and sensitivity analyses for age-dependent unavailability model integrating test and maintenance. *Nuclear engineering and design, 246*, 128-135.Author. N. and N. Author (1986).

Keller, W., & Modarres, M. (2005). A historical overview of probabilistic risk assessment development and its use in the nuclear power industry: a tribute to the late Professor Norman Carl Rasmussen. *Reliability Engineering & System Safety, 89(3)*, 271-285.

Ma, Z., Bao, H., Zhang, S., Xian, M. and Mack, A. (2022) Exploring Advanced Computational Tools and Techniques with Artificial Intelligence and Machine Learning in Operating Nuclear Plants. NUREG/CR-7294. (ADAMS Accession No. ML22042A662)

Martón, I., Sánchez, A. I., & Martorell, S. (2015). Ageing PSA incorporating effectiveness of maintenance and testing. *Reliability Engineering & System Safety, 139*, 131-140.

Martorell, P., Martorell, S., Martón, I., Carlos, S., & Sánchez, A. I. (2018). Time-dependent unavailability model integrating on demand-caused and standby-related failures addressing positive and negative effects of testing and maintenance. In *Safety and Reliability–Safe Societies in a Changing World* (pp. 595-599). CRC Press.

Rigby, R. A., and Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical*

Society: Series C (Applied Statistics), 54(3), 507-554.

Smola, A. J., and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing, 14*, 199-222.

Stasinopoulos, D. M., & Rigby, R. A. (2008). Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software, 23*, 1-46.

Van der Borst, M., and Schoonakker, H. (2001). An overview of PSA importance measures. *Reliability Engineering & System Safety, 72(3)*, 241-245.