Energy and AI 17 (2024) 100401

Contents lists available at ScienceDirect

Energy and AI

journal homepage: www.sciencedirect.com/journal/energy-and-ai

Uplifting the complexity of analysis for probabilistic security of electricity supply assessments using artificial neural networks

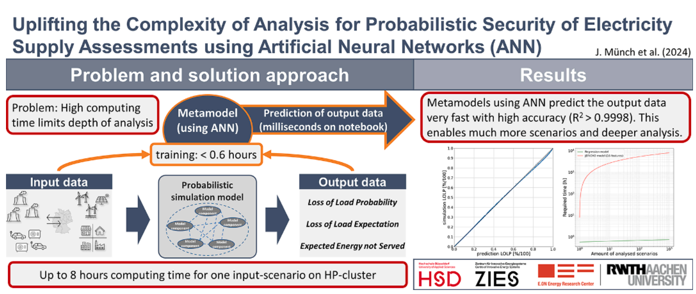
Justin Münch [a,*], Jan Priesmann [b], Marius Reich [a], Marius Tillmanns [b], Aaron Praktiknjo [b], Mario Adam [a,*]

[a] University of Applied Sciences Duesseldorf, Centre of Innovative Energy Systems, Muensterstr.156, 40476 Duesseldorf, Germany
[b] RWTH Aachen University, E.ON Energy Research Center, Chair for Energy System Economics (FCN-ESE), Mathieustrasse 10, 52074 Aachen, Germany

HIGHLIGHTS

- Simulation models for security of electricity supply need high computational time.
- Metamodeling using ANN enables strong time reduction with good prediction quality.
- Metamodeling increases the analysis depth of computationally intensive models.
- Deep learning methods are useful for approximating probabilistic simulation models.
- Statistical methods to sample the design space reduce the required training data.

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Security of electricity supply
Probabilistic simulation
Metamodeling
Artificial neural networks
Regression

ABSTRACT

The energy sector faces rapid decarbonisation and decision-makers demand reliable assessments of the security of electricity supply. For this, detailed simulation models with a high temporal and technological resolution are required. When confronted with increasing weather-dependent renewable energy generation, probabilistic simulation models have proven. The significant computational costs of calculating a scenario, however, limit the complexity of further analysis. Advances in code optimization as well as the use of computing clusters still lead to runtimes of up to eight hours per scenario. However ongoing research highlights that tailor-made approximations are potentially the key factor in further reducing computing time. Consequently, current research aims to provide a method for the rapid prediction of widely varying scenarios. In this work artificial neural networks (ANN) are trained and compared to approximate the system behavior of the probabilistic simulation model. To do so, information needs to be sampled from the probabilistic simulation in an efficient way. Because only a limited space in the whole design space of the 16 independent variables is of interest, a classification is developed. Finally it required only around 35 min to create the regression models, including sampling the design space, simulating the training data and training the ANNs. The resulting ANNs are able to predict all scenarios within the validity range of the regression model with a coefficient of determination of over 0.9998 for independent test data (1.051.200 data points). They need only a few milliseconds to predict one scenario, enabling in-depth analysis in a brief period of time.

* Corresponding authors.
E-mail addresses: justin.muench@hs-duesseldorf.de (J. Münch), mario.adam@hs-duesseldorf.de (M. Adam).

https://doi.org/10.1016/j.egyai.2024.100401

Available online 2 August 2024

1. Introduction

Against the background of the shutdown of conventional power plants and the current energy shortage, the assessment of the security of electricity supply in Germany and other European countries is becoming increasingly important. In order to reduce CO_2 emissions, the German government has created the legal basis for shutting down all hard coal and lignite-fired power plants by the end of 2038 at the latest with the "Act to Reduce and End Coal-fired Power Generation" [1], thus following the energy policy recommendation of the Commission on "Growth, Structural Change and Employment" [2]. Furthermore, as a direct reaction to the nuclear disaster in Fukushima Daiichi, it was decided in 2011 to shut down all existing nuclear power plants by 2023 [3]. In the course of the energy crisis and the resulting switch from pipeline gas to liquefied natural gas from various supply sources, an amendment to the law for a temporary extension of the operating lives of nuclear power plants was passed by the German parliament [4]. This decision was justified by concerns regarding security of electricity. Without a reliable gas supply, this problem will continue to exist in the future. Also, in particular due to the increasing importance of gas supply as the last conventional energie source in electricity generation in the future with a share of 33.3% in 2022 [5] and at the same time as an important component in heat supply with a share of 66.2% in 2021 (Federal Statistical Office, 2021).

Reduced capacities of controllable conventional power plants and growing shares of intermittent feed-in of renewable energies with low storage capacities can lead to a higher probability of situations in which there is a energy supply shortfall. This means that the electricity supply is not sufficient to cover the load, and this can lead to load-shedding measures [6]. Currently, Germany has a high level of security of supply and in recent years there have been only a few interruptions to the electricity supply due to grid instabilities. The annual interruption duration per customer in Germany, the so-called System Average Interruption Duration Index (SAIDI), averaged only 14.8 min between 2006 and 2021 and reached a maximum of 21.5 min [7]. However, this situation could change in the future due to the abovementioned changes in the German/European electricity system. According to the report “Power Supply Security” [8], security of supply in the period from 2025 to 2031 is ensured, even if coal is completely phased out by 2030. However, it also accentuate that a number of developments on the generation and grid side must be implemented as well as the importance of an ongoing assessment of future security of supply. As can be inferred from the foregoing, system analyses that map the uncertainties of future developments are of particular importance for independent scientific advice to decision-makers in politics and industry. This requires high-resolution probabilistic simulation models that are capable of mapping the probability distribution of the available power plants for a large number of scenarios to be investigated. The calculation of these probability distributions is very computationally intensive and is, therefore, associated with long runtimes. The computing times were reduced with advances in code optimization and the use of computing clusters, but still require up to 8 h of computing time per scenario. Due to these long computing times, in-depth analyses, which require the consideration of a large number of scenarios, cannot be carried out in a reasonable time. This often leads either to a reduction in the depth of analysis or to a reduction in the complexity of the simulation model so that results can be achieved within a reasonable period of time. The aim of current research is to overcome this necessary trade-off between depth of analysis, complexity of the simulation model and duration of the analysis using modern methods from the field of machine learning (ML).

The first work in the field of ML dates back to the middle of the 20th century. However, especially in recent years, the use of ML methods in all disciplines has increased significantly due to the improved performance and accessibility of alogrithms, computing hardware and data storage (big data era) [9]. In the review article “Machine Learning: Algorithms, Real-World Applications and Research Directions”, the authors provide a comprehensive overview of ML methods for data analysis and applications in the context of the fourth industrial revolution (4IR or Industry 4.0), digitalization and data such as the Internet of Things (IoT), i.e. cybersecurity data, mobile data, business data, social media data, health data, etc. [10]. The authors discuss how different ML methods can be applied to real-world problems. ML methods are already being successfully applied to many real-world engineering problems. For example, in the paper “A Survey of Machine Learning-Based System Performance Optimization Techniques” [11], the authors review approaches to system performance optimization based on ML methods. The authors come to the conclusion that the use of ML methods is promising and has considerable potential. In the field of energy technology, there are also numerous studies that demonstrate the benefits of ML methods and, in particular, artificial neural networks. In the review article “A comprehensive review of machine learning and IoT solutions for demand side energy management, conservation, and resilient operation” [12], the authors provide an overview of current research efforts to apply ML strategies to energy conservation and management problems, as well as discuss ML approaches and strategies for energy technologies, control methods, conservation and management problems, among other topics. Furthermore, ANNs have already been applied in the field of carbon capture technology using monoethanolamine to model CO_2 capture levels [13]. In another application, an AI-based modelling and optimization system based on ANNs was developed to enhance the performance of coal-fired power plants [14]. A further work is concerned with the improvement of the isentropic efficiency of a high-pressure steam turbine using ANN for modeling [15]. To improve the representation of cross-border exchange capacities defined by the flow-based approach in European resource adequacy assessments, another work proposes a supervised learning-based approach. This improves the mapping between several relevant explanatory variables and the pre-clustered flow-based domains [16]. This is just a very small view into the large research field of machine learning, but all these and other current works demonstrate the high benefit and potential of ML approaches that have been specifically adapted to the problem.

In our previous paper "Can energy system modeling benefit from artificial neural networks? Application of two-stage metamodels to reduce computation of security of supply assessments" [17] we focus on the problem of high computing times of probabilistic security of electricity supply assessments and present a first approach to metamodeling the probabilistic simulation. Here, the benefit of metamodeling was already shown by a significant reduction of the simulation time by 99.7 % with a high prediction quality. However, the probabilistic simulation model could not be completely metamodeled with this approach. This means that the target variables *Lost of Load Probability* (LOLP), *Loss of Load Expectation* (LOLE) and *Expected Energy Not Severed* (EENS) cannot be predicted directly, only the computationally intensive calculation of the hourly resolved probability distributions. These correspond to sigmoid functions, so that the parameters of the sigmoid functions were predicted, and the target variables were then determined from these. In addition, the dimensionality of the problem had to be reduced so that only the conventional power plant park (4 dimensions) could be used as input variables for the metamodeling. This approach is extended in this work so that a direct approximation of the target variables by metamodeling is possible under all relevant input variables, i.e. weather-dependent renewable power plant park, conventional power plant park, planned and unplanned unavailability etc. (16 dimensions). Comprehensive metamodeling of the probabilistic simulation model will significantly accelerate the implementation of analysis methods that require a large number of different scenarios for the input variables.

1.1. Our research questions are as follows

1. Can artificial neural networks predict the key indicators of security of electricity supply assessments by considering all influencing input

variables, which are simulated using a probabilistic simulation model?
2. How can the design space of the probabilistic simulation be scanned as effectively as possible if a target value is limited between values of 0 and 1?
3. What are the runtime gains on the one hand and accuracy losses on the other hand compared to the native simulation?
4. Can the methodology enable in-depth analyses to be performed in a reasonable time?

The rest of the paper is structured as follows: Section 2 presents the simulation model used for the security of supply assessment in more detail. Section 3 describes the methods used to increase the depth of analysis. Methods for generating the data set for training the neural network are presented. Section 4 describes the input data and scenarios for assessing the security of supply. Section 5 shows our results and provides discussions with regards to both, runtime accelerations and accuracy losses. Finally, we provide a conclusion and outlook for possible future research in Section 6.

2. Description of the simulation model

The JERICHO security of supply model is a Python-based simulation model to assess the security of electricity supply. The model has been developed by [18]. A schematic overview of the JERICHO security of supply model is depicted in Fig. 1.

The probabilistic model takes into account stochastic fluctuations of both renewable feed-in and electricity demand, volatile availability of power plant capacities as well as import potentials from neighboring countries. Hourly simulations for 30 different weather years are performed to represent the stochasticity of weather influences. Recursive convolution is used to aggregate the availabilities of all installed conventional power plant units to one distribution curve. Each conventional power plant unit has two possible states: non-availability $p_{non-availability}(t)$, or complete availability $(1 - p_{non-availability}(t))$ and therefore follows a discrete Bernoulli distribution. In this context, the probability of the non-availability of a power plant unit consists of planned and unplanned unavailability. In the event of planned unavailability, the outage start and duration are determined at least four weeks before its occurrence. For unplanned unavailability, the lead time is shorter. The number of conventional power plant units in the German power plant portfolio has a direct influence on the computational effort of the model. For the number n of block units, the possible states grow exponentially with 2^n. Due to the hourly resolution, 8760 h are calculated and aggregated per simulation year. Thus, recursive convolution is a computationally intensive algorithm that requires several hours of computing time even with a parallelized code on computing clusters.

3. Metamodeling to reduce computational effort

In this section, the method for accelerating the generation of results to increase the depth of analysis of the JERICHO security of supply model is explained in more detail. First, the method of metamodeling is presented. Subsequently, the problem statement at hand is explained. Based on this, solutions for effective sampling of the design space and thus for generating data for approximating the underlying simulation are presented.

3.1. Methodology of metamodeling

Metamodeling is a possibility to enable in-depth analysis for complex simulation models when direct use of simulation models is limited due to high computational effort. Metamodels are used to represent the system behavior of the simulation model using regression methods by establishing a relationship between the input and output data. Once a metamodel has been successfully created, it is able to generate the output variables of the simulation model with a strong time reduction through a prognosis. A variety of regression methods can be used for metamodels, achieving different accuracies depending on the present problem. Several classical methods can be used (e.g. polynomial regression). However, in most applications, simulations have a more complex character. This often leads to nonlinear relationships between the input and output data, which cannot be sufficiently approximated by classical methods. In these cases, ML methods can be applied. Fig. 2 shows the methodology schematically for the present problem security of supply assessment with electricity.

Metamodels are generated from real simulation results and are valid for a predefined design space. The design space represents a multidimensional structure spanned by the input data of the simulation model and includes the entire range of all input data combinations. The minima and maxima of these input data define the boundaries of the design space (see section 3.2). Built metamodels are able to predict each response (output data) of each factor combination of the input data within the design space. However, the prediction is limited to the design

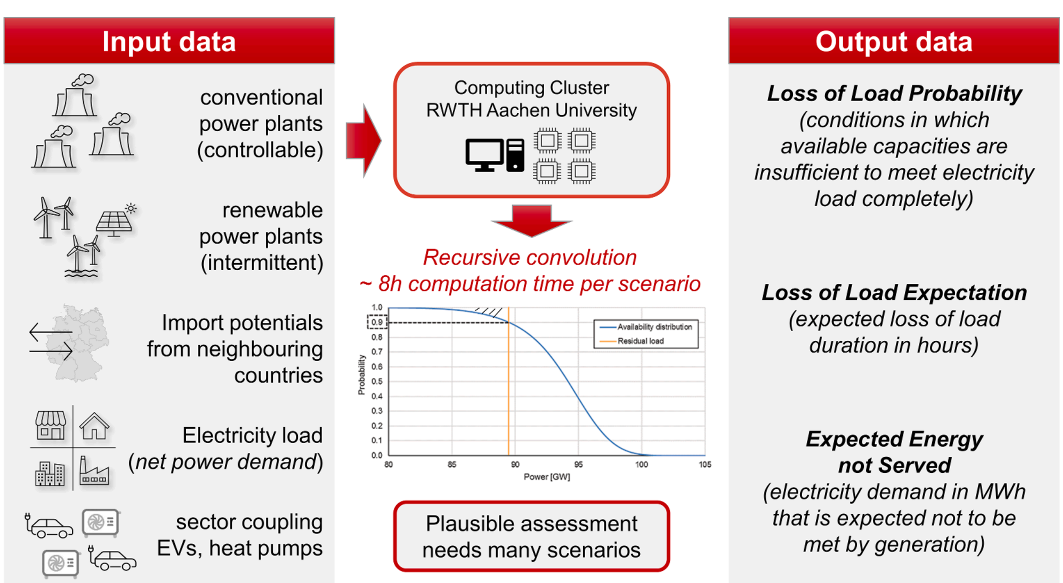


Fig. 1. JERICHO Security of supply model, own representation.

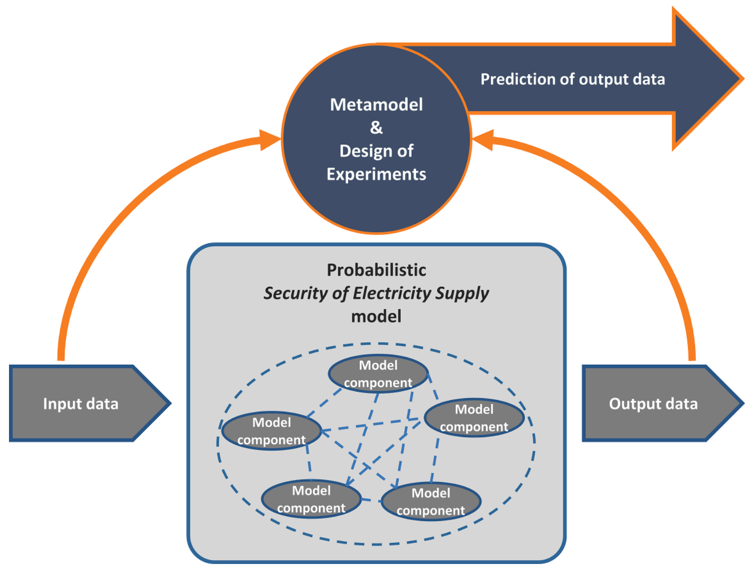


Fig. 2. Schematic representation of the metamodeling method [19].

space in which the metamodel is valid. Fig. 3 demonstrates an example of the design space for a three-dimensional problem.

When sampling the data used to create the metamodel, it is important to ensure that the amount of information obtained is sufficient to represent the system behavior. On the other hand, with the amount of input data combinations rises the sampling respectively simulation time to obtain the related output data. In order to scan the design space as effectively as possible, a variety of statistical methods are available leading to so-called "designs". For sampling the boundaries of the design space, a full factorial design (FFD), which includes all outer corners of the design space, can be applied. To sample the area between the boundary points, for example, Latin Hypercube Designs (LHD) can be used. Another possible principle is to divide the design space into zones. Within each of these zones, a random factor combination is then determined. However, since a uniform and correlation-free coverage of the design space is not automatically guaranteed, other methods such as orthogonal designs or space-filling designs should be used (YE,1998).

In the context of the problem statement at hand, see section 3.2, metamodeling is primarily associated with regression. For this reason, the term regression respectively regression model is used in the following sections.

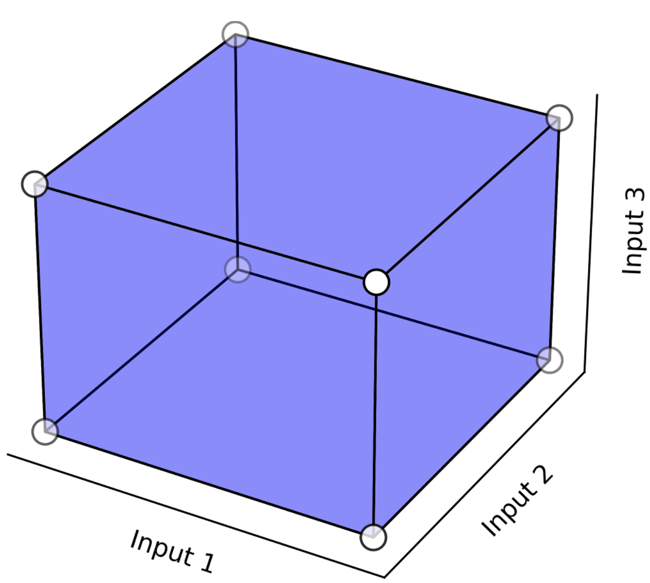


Fig. 3. Demonstration of an exemplary three-dimensional design space.

3.2. Problem statement

The JERICHO security of supply model [18] simulates the target values LOLP, LOLE and EENS on an hourly basis in dependence on the conditions in the power grid, see also Fig. 1. From these, the result values over the 30-year scenarios can be summed up to the LOLE total or the EENS total. The design space is spanned by the 16 time-resolved input variables of the simulation model, which determine the conditions of the power grid at each time step. These include the capacities of the renewable generators (photovoltaics, onshore/ offshore wind power, hydropower, other), the capacities of the conventional power plants (lignite/ hard coal, natural gas and nuclear power plants), the planned and unplanned unavailability of the conventional power plants, the import potential from the neighboring countries and the electricity load. The sizes of the respective minima and maxima of the input values span the 16-dimensional design space of the regression problem. Table 1 shows the chosen boundaries of the design space in which the regression model is valid.

A special characteristic of the regression problem at hand is the target variables, as these have limitations. For example, the LOLP value can only take values between zero and one. In areas of the design space with, for example, a significant load overlap, the LOLP value only takes the value zero. This leads to the fact that there are large areas in the design space in which the target values do not change. A problem arises in the generation of training data for the artificial neural network. A valuable method for selecting a training data set for complex, non-linear simulation models is Latin Hypercube Sampling described above. If this method is applied to the problem at hand, it leads to many selected factor combinations being taken from areas with load overlap. As a result, the training data set has many data points from areas with no further information gain. This means that it has a high number of zero values (approximately 92.0%) for the LOLP value in the training data set. This complicates the generalization of the regression model and leads to worse prediction results. The reason for this lies in an under-representation of areas with high information density. For this reason, sufficient variance in the training data is important. The higher the security of supply in a country, the more space these areas occupy in the design space. As a result, valuable methods for this problem must be adapted for implementation.

Table 1
Boundaries of the design space (power [GW] / unavailabilities [%/100]).

Boundaries	Hard coal	Natural Gas	Lignite	Oil	Nuclear energy	Other
Min.	0.0	0.0	0.0	0.8	0.0	2.0
Max.	25.0	45.0	20.0	1.5	5.0	4.0

Boundaries	Wind Onshore	Wind Offshore	Photovoltaic	Water	Load	Import potential
Min.	0.0	0.0	0.0	1.0	30.0	0.0
Max.	100.0	15.0	100.0	5.0	100.0	50.0

Boundaries	Planned unavailabilities			
	Hard coal	Natural Gas	Lignite	Nuclear energy
Min.	0.0414	0.0109	0.0352	0.0855
Max.	0.1837	0.1597	0.1241	0.1742

3.3. Data selection

As already explained in section 3.2, the regression problem at hand poses a particular challenge when generating factor combinations of the data set for the training/ validation of the regression model. Furthermore, the aim is to use as few factor combinations as possible to reduce the simulation respectively training time while maintaining the highest possible prediction quality simultaneously. To avoid too many factor combinations from areas of low information destiny, the factor combinations of the designs are filtered based on information available before the simulation. The filter criterion used for classification must therefore be able to distinguish the areas with relevant information from those with less relevant information. To avoid additional computational effort, the filter criterion must also be applicable without additional information from the simulation model. Since settings in the design space where the capacities of the power plants are sufficient to cover the electricity demand will always lead to a LOLP of zero, only settings where an energy supply shortfall is present are highly relevance. As thus it is proposed to use the load coverage ratio (LCR) for this purpose.The LCR is calculated on the basis of the input values of the simulation model using the capacities of conventional generators (P_{ce}) considering the planned and unplanned unavailability (p) as well as the electricity production of renewable generators (P_{re}), the import potential of the neighboring countries ($P_{imp.}$) and the current load applied (P_{load}) according to Eq. (1).

$$LCR = \frac{\sum_{i=1}^{n}\left(P_{ce\_i}\cdot(1-p_i)\right) + \sum_{i=1}^{n}\left(P_{re\_i}\right) + P_{Imp.}}{P_{load}} \tag{1}$$

In the following Fig. 4, the filtering principle is exemplified for a three-dimensional problem using an FFD (red) and an LHD (blue). The left schematic illustration of the design space shows the unfiltered case and the right schematic illustration shows the filtered case.

For the most efficient sampling of the relevant regions of the design space, the design is composed of two designs. First, a filtered fractional factorial design (F-FFD) with 2000 factor combinations is created to sample some information from the boundaries of the design space as a random selection of outer corners of a filtered full-factorial design. To ensure that the training data also contains a small proportion of information from areas of the design space with less relevant information, 10 % of the F-FFD is selected without an energy supply shortfall. A complete full-factorial design was not used, because this would increase the simulation time for the design too much even with filtering for a 16-dimensional problem (2^{16} factor combination). This is then combined with a filtered random distribution (FRD) in the second step. For FRD, random factor combinations (= input data sets) are first determined until a desired number of input data sets with energy supply shortfall have been found. Subsequently, the filtered randomly generated distribution of the input data sets in the design space is evaluated. The Pearson correlation *r* is used for this purpose. This factor is a dimensionless measure of covariance that can take values between −1 and 1. The Pearson correlation, unlike the covariance, does not depend on a measurement scale and is therefore universally comparable. If this takes the value zero, then there is no linear relationship between the location of the input data sets in the design space. If, on the other hand, the value approaches an extreme point (±1), then the relationship increases in a positive respectively negative way. This optimization process is repeated until a design is found whose distribution could not be improved 1000 times. An improvement is achieved when the Pearson coefficient is closer to zero. Fig. 5 illustrates this approach.

In the following scatter plot matrix, Fig. 6, the relationships between the individual features are shown graphically using the example of an FRD test plan with 6000 factor combinations. The filtering by the classification works, as can be seen in the point clouds in the figure.

The number of possible factor combinations decreases towards the boundaries of the design space so that the probability of obtaining a factor combination in this area decreases in the case of random generation. As a consequence, fewer factor combinations with a low load coverage ratio are found in the FRD, see Fig. 7. Therefore in addition to the above procedure, an advanced filtered random distribution (AFRD) is considered. In this approach, the factor combinations are uniformly distributed in the energy supply shortfall area. This means that the load coverage ratio is for example divided into 10 zones and the desired

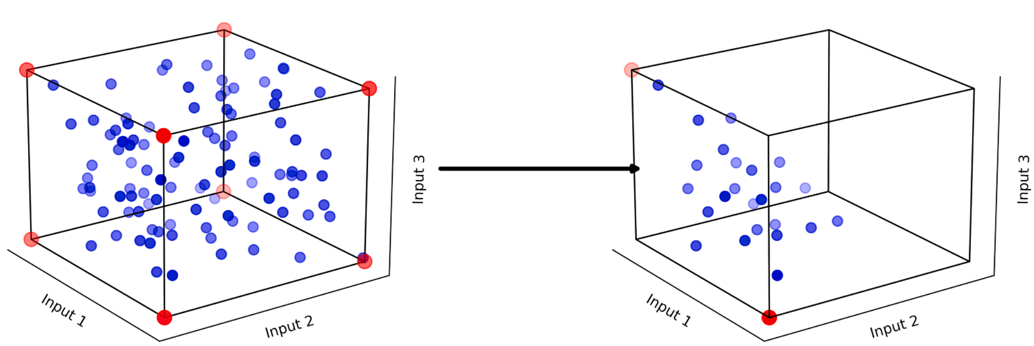


Fig. 4. Schematic illustration of filtering (three-dimensional problem).

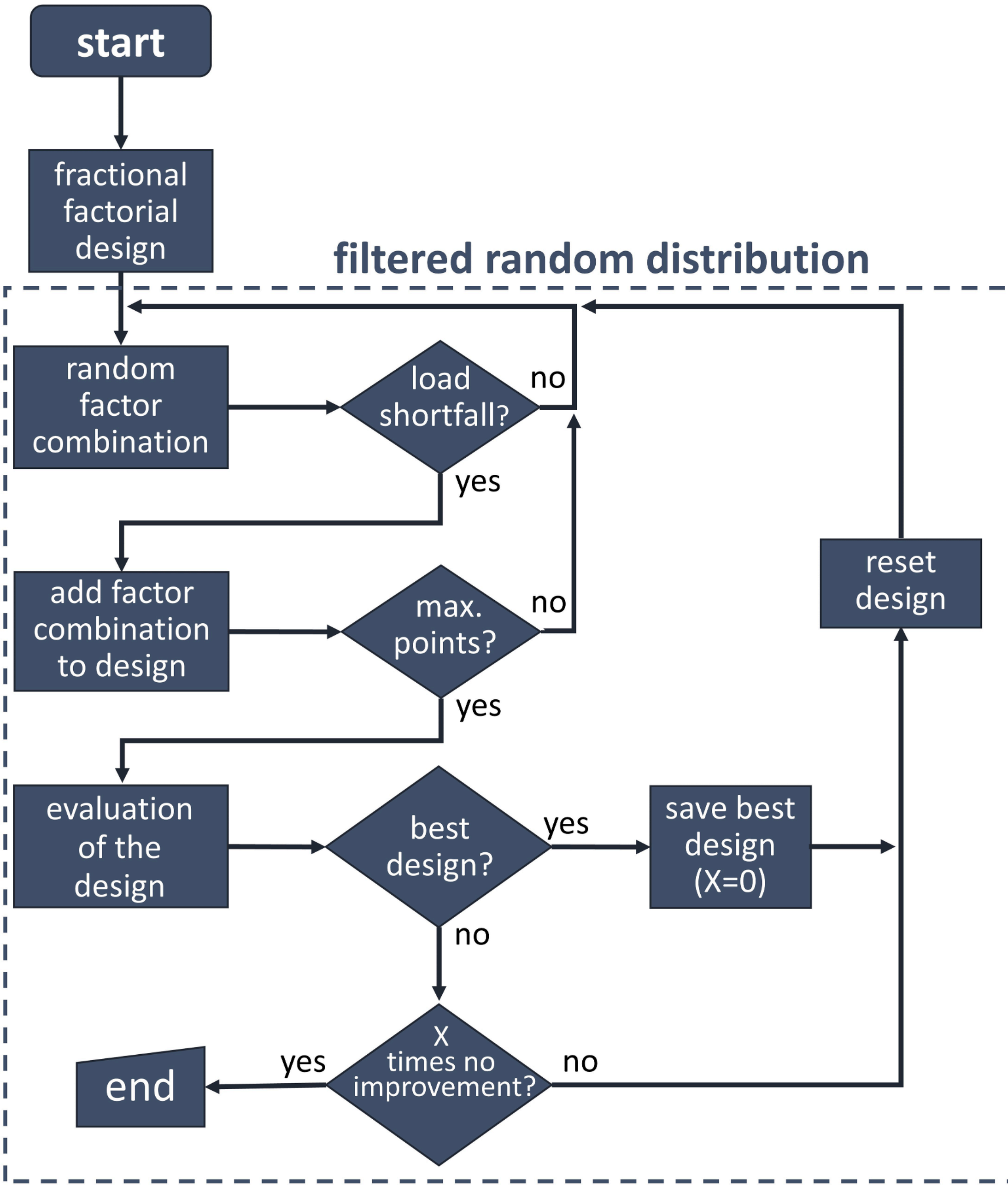


Fig. 5. Schematic representation of the filtered random distribution.

number of factor combinations is distributed equally across these zones. In the case of a design with 4000 factor combinations, 400 factor combinations would be randomly determined in each area. The advantage of this zoning of the filtering is that there is also a uniform distribution over the load coverage ratio. With FRD designs, on the other hand, a large part of the results lie in the area of a low energy supply shortfall. An example of the percentage distribution over the load coverage ratio of an FRD and an AFRD design in the region between 0.6 and 1.0 can be seen in Fig. 7.

3.4. *Training of the artificial neural network*

When complex non-linear relationships are considered, ANN from the field of deep learning, a sub-area of ML, are often used, as they are known to provide powerful and universal approximations [20]. In analogy to the human brain, ANNs are composed of so-called neurons, which are arranged in layers between the inputs and outputs. ANNs consist of different layers of neurons, but at least one input and one output layer. All layers between these two are referred to as hidden layers. Each neuron in a layer is connected to all neurons in the next layer. The strength of this connection can be described by so-called weights. The neurons receive and process the input data by determining weighted sums using activation functions. In an iterative training process, the initially randomly determined unknowns of the neurons are adjusted to minimize a previously defined error function. This requires data with corresponding input and output values (supervised learning). The iterative training process is continued until a criterion is reached. This could be, for example, a previously defined accuracy, the maximum number of iterations (so-called epochs) or a defined difference of the accuracy improvement over a certain number of iterations.

The hyperparameters of the artificial neural network are optimized using a random search algorithm. This optimizes the number of hidden layers and the number of neurons. The optimization is stopped when 20

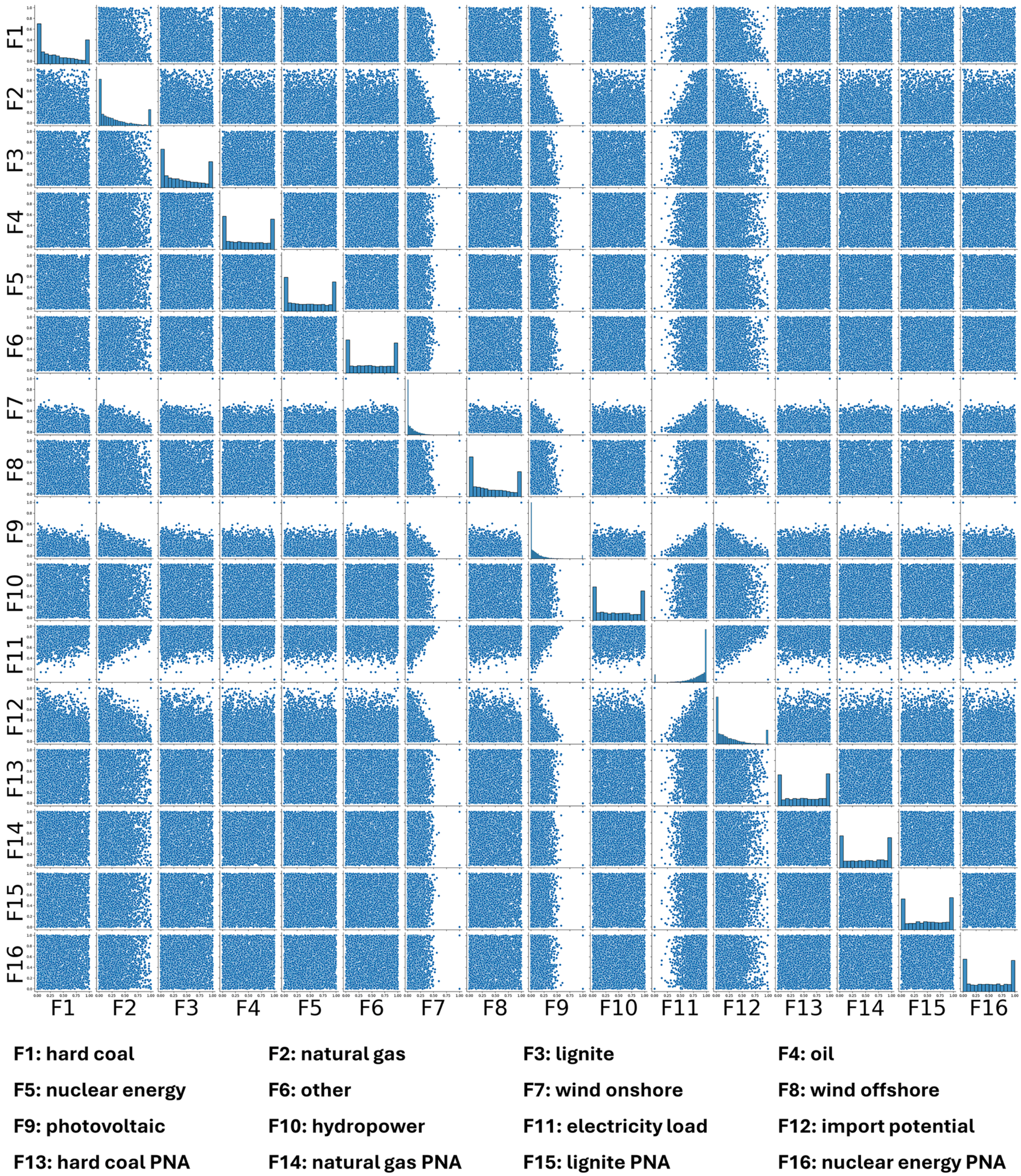


Fig. 6. Scatter plot matrix (FRD with 6000 factor combinations).

subsequent training runs have not led to any improvement. The evaluation is carried out using validation data independent of the training process and the mean square error is used as the evaluation criterion. For this purpose, a random permutation of 10 % is taken from the training data set as validation data. During the training process, the weights and bias values of the ANN are adjusted based on the training data using a limited memory Broyden-Fletcher-Goldfarb-Shanno optimizer and the mean square error is also used as an evaluation criterion. The hyperparameter optimization procedure already detects regression models that have an overfitting, i.e. a decreasing prediction error in the training data and an increasing or high prediction error in the independent validation data. This includes an early stop if the evaluation criterion does not improve over a certain number of epochs. Furthermore, an increase indiscriminately in the number of neurons and hidden layers does not automatically guarantee a higher generalization capability of an ANN and can also encourage overfitting [21]. For this reason, the upper limits of the hyperparameter optimization were set low with a maximum of 5 hidden layers and 50 neurons. The number of neurons in the input and output layer are defined by the number of features or target variables. In this work two target variables (LOLP/EENS value) are relevant and one ANN is trained for each target variable.

4. Description of relevant data and scenarios

We compare scenarios with different situations in terms of security of electricity supply to validate the accuracy of the trained regression model. For this, we take the current power plant park of Germany and manipulate capacities to retrieve hypothetical power plant park

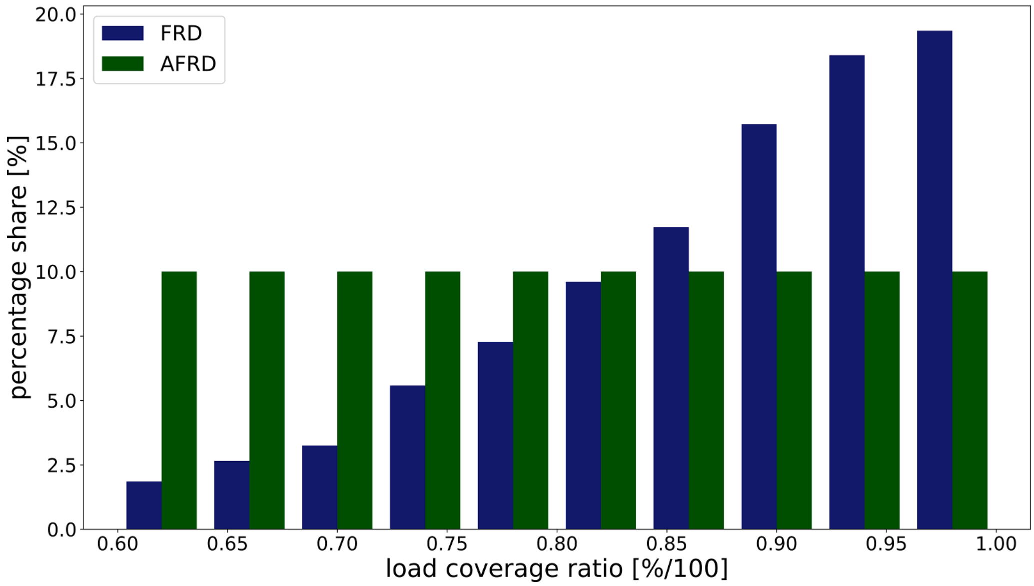


Fig. 7. Percentage of factor combinations in a design over the load coverage ratio.

scenarios. The scenarios used are:

(1) Uncertainty scenario: Consists of an artificially composed power plant park, which can lead to frequent supply uncertainties
(2) Base case: Power plant park as of November 2022 (BNetzA, 2022)
(3) Nuclear phase-out: Phase out of all remaining nuclear power plants in Germany as happened in April 2023
(4) Nuclear phase-out and additional absence of reserve capacities: In addition to case (2), reserve capacities that are not participating in the market but can be activated in case of a supply shortage are no longer available.

In the following, the underlying data is described.

4.1. Power plant park

The power plant capacities of the scenarios are displayed in Fig. 8. Compared to the base scenario (scenario (1)), the total conventional power plant capacities in the other scenarios are ~4 GW lower for scenario (2) and ~12 GW for scenario (3). This affects nuclear power plant capacities as well as lignite, hard coal and natural gas capacities. The installed capacities of the renewable energy plants are constant in all scenarios.

4.2. Time series data

Hourly time series data for 30 weather years is used for the probabilistic assessment of security of electricity supply. This comprises weather-dependent uncertainties influencing electricity load, renewable electricity generation and import potentials for electricity from neighboring countries. Distributions of the values are shown in Fig. 9.

5. Results and discussion

The results of this research are presented and discussed below. Section 5.1 analyzes the required number of factor combinations of the training data and compares the two types of filtering using a sensitivity analysis. Following this, the prediction quality of the regression models is determined in section 5.2 based on independent test data. Finally, section 5.3 evaluates the time reduction compared to the native simulation.

In addition to the prediction quality, the time required to create the regression model is decisive for assessing the usefulness of the regression model. This includes the creation time for the design of the input data sets, the simulation time for the generation of the training data with the original JERICHO security of supply model and the training/prediction time with the regression model. The first step is therefore to investigate how many factor combinations are required for an acceptable prediction in order to keep the simulation time to a minimum. At the same time, it is being investigated whether additional filtering via areas of load coverage ratio (FRD/ AFRD) has a positive influence. Subsequently, the prediction quality is presented in more detail for the individual test scenarios on the basis of the best regression model.

To measure the quality of the prediction for the target values of the probabilistic simulation model (LOLP/ EENS), the coefficient of determination R^2 is initially used as a dimensionless criterion, which represents a well-comparable and problem-independent criterion [22]. It is also very suitable for the limited data available (e.g. 0 values) and offers

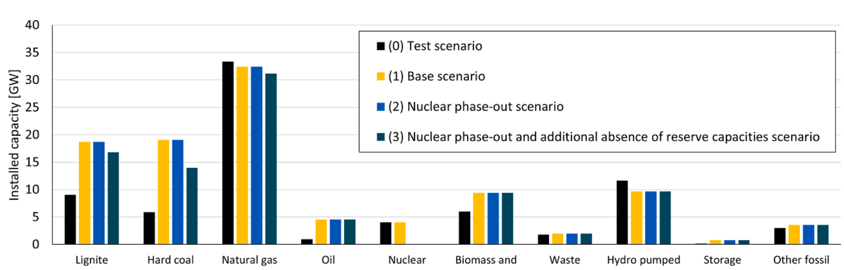


Fig. 8. Installed conventional capacities in Germany according to scenarios.

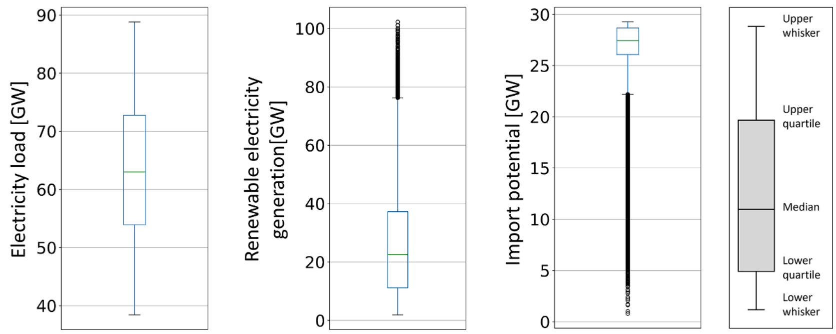


Fig. 9. Boxplots of input data used in the probabilistic simulation model. Displayed are 8760 hourly values for 30 weather years each for electricity load, renewable electricity generation and electricity import potentials.

good comparability between scenarios with different absolute values. Furthermore, the absolute error (AE) and the absolute percentage error (APE) are used for evaluation. In addition, the prediction quality is evaluated on the basis of the total values of the target variables over the respective scenario (30 weather years) or all scenarios together. Therefore, the LOLE/ EENS values are summed up and subsequently referred to as LOLE total/ EENS total. With the exception of Fig. 10, around 262.800 test points (per scenario; = 8.760 h/a x 30 a) to 1.051.200 test points (across all scenarios) are used.

5.1. Time reduction for generating the simulation results

In the following, the validation data are used to evaluate the required number of factor combinations, since these are available in practice without further simulation effort. For both types of filtering (FRD/ AFRD), five designs each with factor combinations between 2500 and 6000 were considered. Due to random processes during training and optimization of ANNs, e.g. initialization of weights, etc., the prediction quality varies in individual training or optimization runs. For this reason, twenty regression models were optimized per design. The coefficient of determination is used as the basis for assessing the quality of the prediction. The results for the LOLP value are shown in a box plot in Fig. 10. For each design, the coefficients of determination, which were determined using the validation data (10 % of the design plan), are shown in the box plot for the 20 optimization runs.

As can be seen in Fig. 10, as the number of factor combinations increases, the prediction quality of the validation data increases. Furthermore, the range between minimum and maximum prediction quality decreases, so that with an increasing number of factor combinations, the result becomes also more robust to the random factors in ANN training. Here, designs with 4000 factor combinations or more can achieve high prediction quality on the validation data with a coefficient of determination above 0.9999. The designs with 6000 factor combinations achieve a coefficient of determination close to 0.9999 even with the statistical outliers. When comparing the two filtering types (FRD/ AFRD), it can be concluded that the median, mean and robustness of optimization (interquartile range) are better in AFRD filtering than in FRD filtering. Only in the results of the design with 3000 factor combinations this cannot be observed.

For a reliable evaluation of the prediction quality, independent test data are required, i.e. data which were not used for the determination of the weights or to optimize the hyperparameters. In the following, the coefficient of determination for the LOLP value of the 20 optimization runs is considered across all test scenarios (1.051.200 data points each). Analogous to the validation data, an increase in the prediction quality and the robustness of the optimization can be observed with an increasing number of factor combinations, so that only the relevant range between 4000 and 6000 factor combinations is shown in Fig. 11.

With this low-resolution sensitivity analysis, the exact minimum number of factor combinations cannot be determined, but it can be stated that well-trained regression models are possible from 4000 factor combinations. A further increase of the number of factor combinations mainly serves the robustness of the training and reduces the number of optimization runs. From 4000 factor combinations, the added value is significantly reduced for a further increase of 1000 factor combinations, so that this added value must be compared to the constant increase in simulation time for 1000 additional factor combinations (see also Section 5.2/5.3).

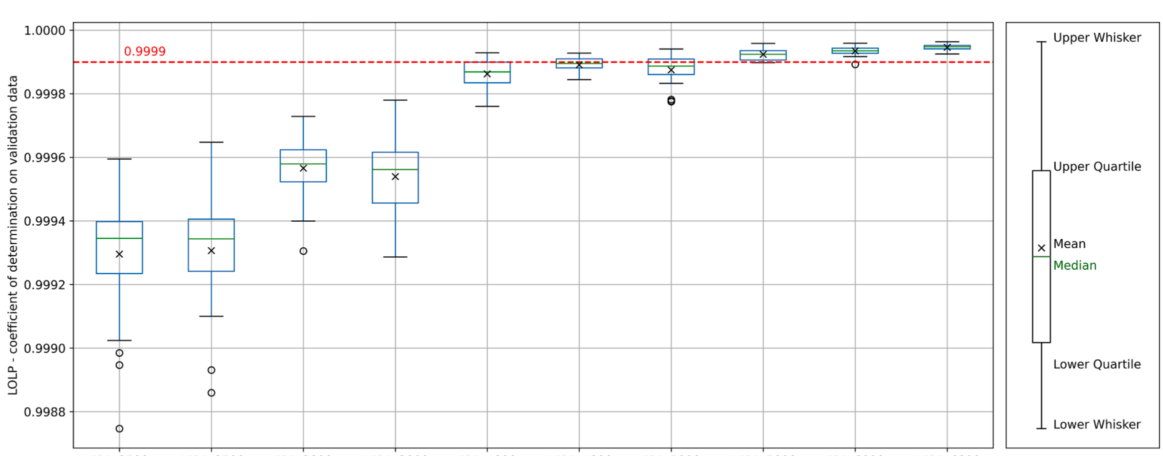


Fig. 10. LOLP - Coefficient of determination for different designs on validation data.

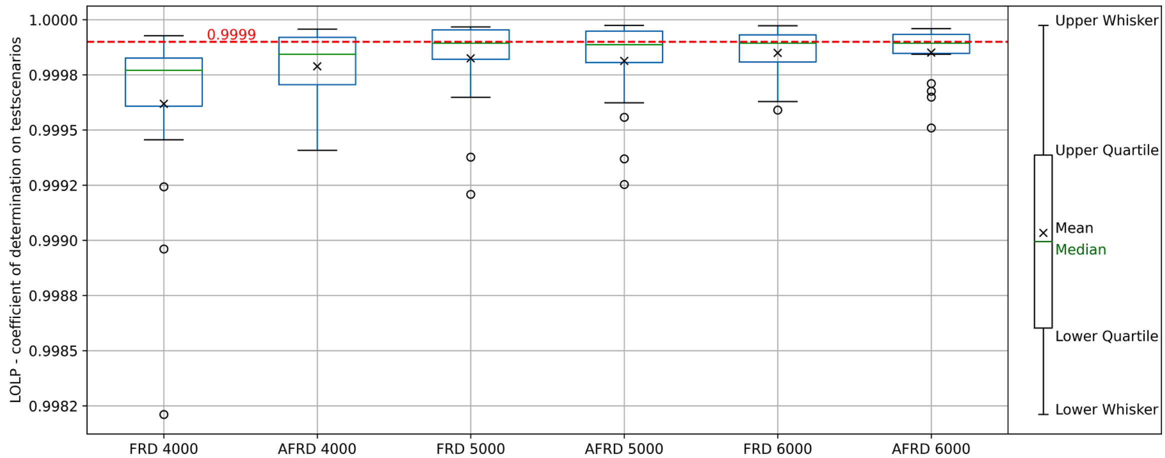


Fig. 11. LOLP - Coefficient of determination for different designs across all scenarios on test data.

5.2. Prediction quality on the scenarios

In this chapter detailed prediction results are presented based on the individual test scenarios on only one selected design. Since the simulation time only increases by about 2 min for another 1000 factor combinations (see Section 5.3), a design with 6000 factor combinations (and AFRD design) is used for robustness of the optimization and prediction quality (see Section 5.1). Fig. 12 shows the coefficient of determination of the LOLP value for the different test scenarios for twenty optimization runs each.

As shown in the figure above, the median coefficient of determination for the test scenarios is between 0.9470 (outlier) and 0.9999. The regression models perform differently in the test scenarios. In the two test scenarios (uncertainty scenario /no nuclear no reserve) with poorer security of supply rating, very high coefficients of determination between 0.9994 and 0.9999 are achieved. In these scenarios, the target values cover a large range of the value range (e.g. LOLP 0–1). For scenarios with a very good security of supply assessment, i.e. with low absolute LOLP values, the prediction quality of the regression models decreases. In these scenarios, for example, only LOLP values close to the lower limit are reached. The median coefficient of determination for these scenarios is 0.9855 and 0.9983 respectively. With maximum values of up to 0.9999, these scenarios also achieve high values for the coefficient of determination. This fluctuation in prediction quality for the base scenario and no nuclear scenario can also explain the outliers in the coefficient of determination across all scenarios, see Fig. 11. The results in Figs. 10 and 11 were generated using the same regression models (AFRD 6000), but they are not identical as the coefficient of determination was determined once over all test scenarios and once for the individual test scenarios each.

To further evaluate the prediction quality, the absolute error (AE) and the absolute percentage error (APE) are now considered. Due to the high number of zero values of the target variable (LOLP), the prediction quality is considered using the sum values of the scenarios. Fig. 13 shows the results of the AE for the total LOLE values for the AFRD design with 6000 factor combinations for the twenty optimization runs of each scenario. The median AE for the scenarios with low absolute LOLP values is around 23 s (base scenario) and around 48 seconds (no nuclear). In contrast, the median AE for scenarios with higher absolute LOLP values is around 9 minutes (no nuclear & no reserve) and around 10.5 hours (uncertainty scenario).

To additionally consider the APE, this is shown in Fig. 14 analogously for the AFRD design with 6000 factor combinations.

This illustrates that the median APE is lower for scenarios with high absolute LOLE/ AE values, e.g. around 0.7% for the uncertainty scenario and 1.2% for the scenario no nuclear & no reserve scenario. The same applies to scenarios with low LOLP/AE values. Here, median APE values of around 21.4% are achieved in the base scenario and around 4.4% in the no nuclear scenario. These increased APE values must be compared with the corresponding AE values of 22.9 resp. 48.1 s over a period of 30 years.

To demonstrate the best possible prediction quality, the best regression model is selected from the 20 optimization runs for the LOLP and EENS. The selection is based on the prediction quality of the validation data, as this information is directly available. The highest coefficient of determination for the LOLP value is 0.99996 and for the EENS value 0.99999 on the validation data. The results of these regression models are shown in Fig. 15 in a prediction/observation plot including

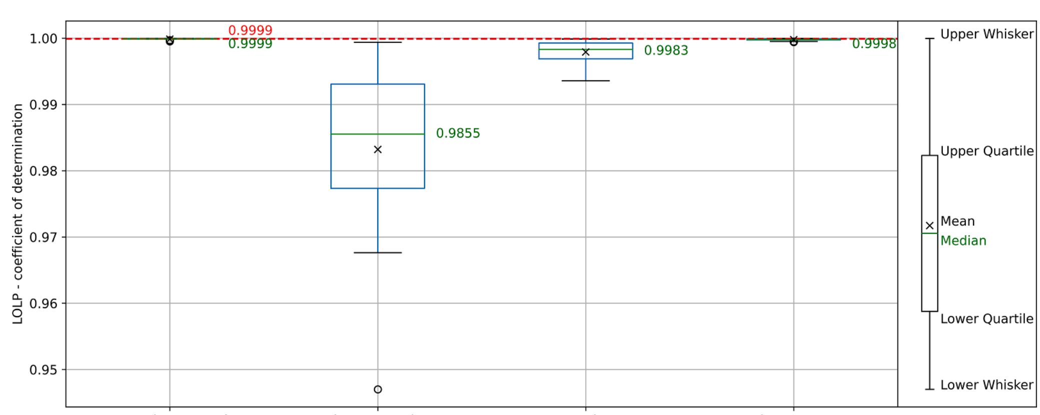


Fig. 12. LOLP - coefficient of determination for the design AFRD 6000 and different scenarios on test data.

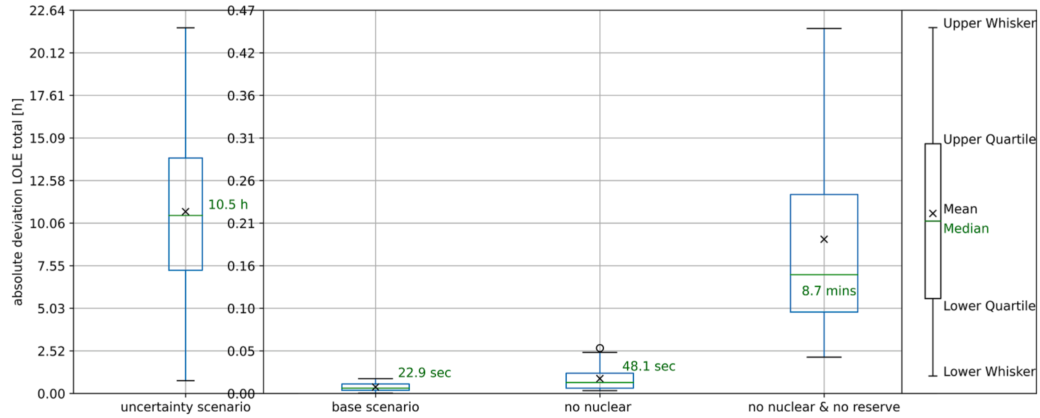


Fig. 13. LOLE - absolute error for the design AFRD 6000 and different scenarios on test data.

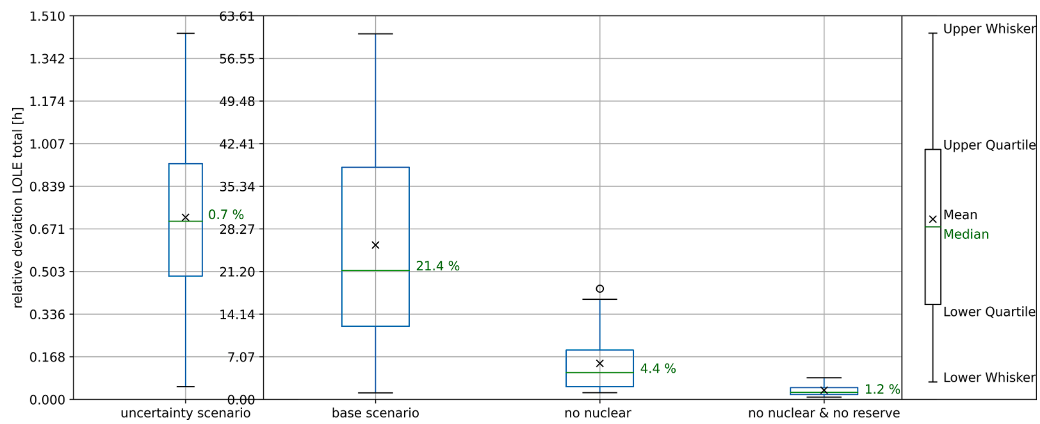


Fig. 14. LOLE - absolute percentage error for the design AFRD 6000 and different scenarios on test data.

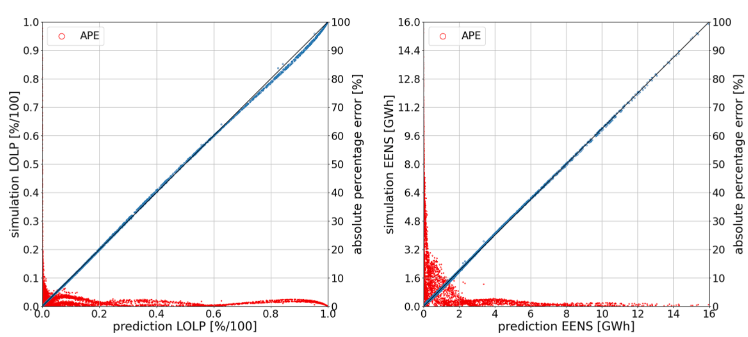


Fig. 15. Prediction / observation plot for LOLP / EENS incl. APE across all scenarios on test data with the best regression model.

APE for all scenarios. In areas close to the zero values (fourth decimal place), APEs greater than 100% also occur; the APE axis was limited for the purpose of visualization.

The best regression models achieve a coefficient of determination of 0.9999 for the LOLP value and 0.9998 for the EENS value on test data across all scenarios. The APE value is 0.5 % for both the LOLE value and the total value of the EENS across all scenarios. This figure also shows that high APE values only occur at low LOLP/ EENS values. The coefficient of determination (R^2) for the LOLP/ EENS value as well as the AE and APE value of the total values (LOLE total /EENS total) of the individual scenarios can be found in Table 2.

As the above results have shown, outliers rarely occur in artificial neural network optimization with robust training data. However, independent test data should also be available in later practice for the identification of outliers. The simulation of whole scenarios is not

Table 2
prediction quality on test data (R^2, AE and APE) of the best regression models each.

scenario	LOLP	LOLE total		EENS	EENS total	
	R^2	AE [h]	APE [%]	R^2	AE [MWh]	APE [%]
uncertainty scenario	0.9999	7.4934	0.5	0.9998	32,277	0.5
base scenario	0.9877	0.0003	1.1	0.9269	30	14.3
no nuclear	0.9993	0.0048	1.6	0.9583	399	17.6
no nuclear & no reserve	0.9998	0.1491	1.2	0.9920	25,510	11.1
all scenarios	0.9999	7.6476	0.5	0.9998	58,216	0.5

necessary here. A selection of factor combinations from the scenarios to be considered later in the energy supply shortfall area is recommended. A number of 1000 factor combinations is sufficient if they have been evaluated using the methodology presented in section 3.3. This means that in an interactive process, random factor combinations from the test scenarios with energy supply shortfall are evaluated based on a good distribution in the design space. After a defined number of iterations without improvement, the search can be stopped and the test data set can be used to identify outliers.

5.3. Evaluation of time-saving through metamodeling

In addition to the prediction quality, the time required to generate the simulation results via the regression modeling method is crucial for evaluating the method for energy research. For the evaluation, the time for the design with the most factor combinations and strongest filtering type (AFRD 6000) is taken as a basis in the following. Here, the time for the generation of the design of the input data sets, the simulation time of the design and the test data (1000 data sets) with the original JERICHO security of supply model, the training/ optimization time for the two regression models (LOLP/ EENS), as well as the prediction time per scenario, has to be considered.

The time required for the individual steps of the regression modeling can be taken from Table 3.

For one scenario with 16 features, the required time is already reduced from 8 h to about 35 min or by 92.7 %. Due to the fact that a regression model is able to predict all scenarios within the validity limits of the regression model, the time to generate the results for each additional scenario increases only by the prediction time, see Fig. 16.

6. Conclusion & outlook

The results of this work have shown that a direct prediction of the output values of probabilistic simulation models for the analysis of the security of supply with electricity is possible by using artificial neural networks and the method of regression modeling. Thereby, a reasonable prediction quality can be achieved with a significant reduction of the needed time for in-depth analyses. This method is capable of performing in-depth analyses with a very large number of scenarios for probabilistic security of electricity supply assessments in a reasonable amount of time. However, this work also shows that regression modeling requires solutions specifically adapted to the problem. If the target variables in the regression modeling problem are limited (here between 0 and 1), it leads to the fact that there are large areas in the design space in which the target variables do not change. An excessively high percentage of data points from areas with low information destiny in the training data makes it difficult to generalize the regression model. For this reason, a method for efficiently scanning the areas of the design space with a high information density via classification is of a high importance. Furthermore, the results also show the importance of test data independent of training for identifying outliers in training. Although these outliers rarely occur in validated designs with sufficient factor combinations, they could not be completely excluded. Thus, independent test data is of great importance not only in research but also in subsequent applications.

The best regression models with the finally selected design achieved an absolute percentage error of 0.5% for the Loss of Load Expectation total and Expected Energy Not Severed total over all independent testscenarios. Larger absolute percentage errors only occur in areas of the target variables with very low absolute values, which only have a minor impact on the overall assessment of the security of electricity supply. This applies to Lost of Load Probability values in the range of a few seconds and the corresponding Expected Energy Not Severed values. In the scenario with the worst prediction quality, the percentage absolute error for the Lost of Load Probability total is 1.6% and the Expected Energy Not Severed total is 17.6%. However, the absolute error in relation to the 30-year period with an annual electricity consumption of around 500 TWh is very low for the Lost of Load Probability total at around 17 s and the Expected Energy Not Severed total at 399 MWh. The simulation time of one scenario with 16 features on a high-performance computing cluster is 8 h. With the presented method of regression modeling, only about 35 min are necessary for the design creation, the simulation of the training data and for the training of the regression model. Following the prediction for the interesting target values of one scenario with a defined power plant constellation and 30 years with different weather conditions requires with the regression model only a few milliseconds. The time benefit increases with each scenario to be considered since regression models can be trained to predict any scenario within the limits applicable to the regression model. In-depth analysis or optimization, which requires results from a large number of scenarios, now becomes possible through regression modeling. The prediction quality can benefit from designs created according to the principle presented with a higher number of factor combinations. Due to the high time reduction, various measures are conceivable to increase the prediction quality while at the same time increasing the time required in a neglectable way. One option is to increase the amount of training data. This can make it easier for the artificial neural networksto recognize patterns in previously underrepresented areas. Therefore it is important to make use of the procedure presented when creating the design. In addition, individual regression models could be created for different dimensions of the regression problem. For example, the regression problem is reduced from 16 to 14 dimensions for some scenarios without nuclear energy. A regression model specially trained for 14 dimensions could achieve a better fit in this area of the design space. Alternatively, it is possible to increase the prediction quality by creating several regression models for different value ranges of the target variables. All this could also benefit the predictive quality of the target variables for example in a small range of values.

Further research can investigate and improve the methodological approach further. A further reduction of the time to create the regression model can be achieved by a stronger reduction of the required factor combinations. Here, the simulation time for generating the training data is reduced as well as the training time, which also reduces the optimization time of the hyperparameters. This could be implemented, for example, by reducing the mapped design space. Regression models could be generated, which only represent a narrow part of the design space around the planned scenarios. Furthermore, the methodology could also benefit from active learning or a reduction of dimensionality (e.g. blending of influencing variables). The foregoing could also make other approximation methods interesting for the method, which in contrast to artificial neural networks have difficulties with high dimensions and high data volumes (e.g. gaussian process regression). In addition, as well as a comparison of different approximation methods using different security of supply assessment models still needs to be carried out to develop a best practice approach for assessing the security of supply with electricity. The probabilistic simulation model used in this work to assess the security of electricity supply is based on the principle of recursive convolution to calculate distribution functions of the available secured feed-in capacities of the power plant park. In future works, the methodological approach should be applied to other approaches for evaluating the security of electricity supply assessment and adapted to the specific requirements. These other approaches

Table 3
Time for the individual steps of regression modeling.

regression modeling steps	Time [min]
design generation (best choice of about 1000 designs)	5.0
simulation time (6.000 + 1.000 = 7.000 factor combinations)	12.8
training time (2 ANN for LOLP and EENS; one optimization)	17.8
prediction time (LOLP, LOLE and EENS for 1 scenario)	<0.01
overall time	35.6

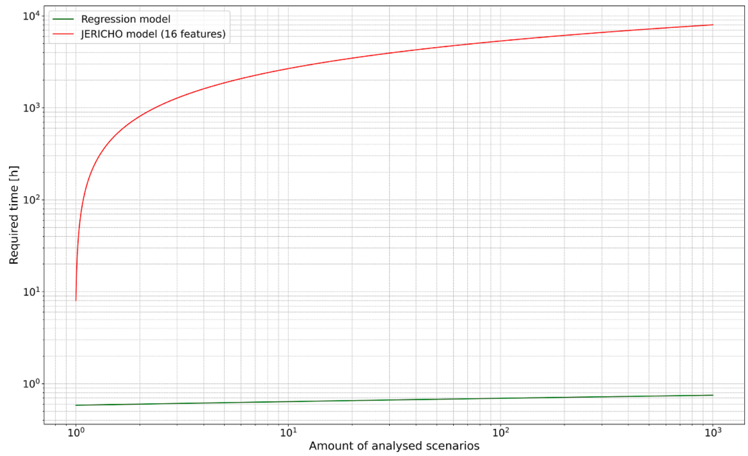


Fig. 16. Time required for the analysis of security of the supply with electricity with the original JERICHO model and with the new ANN metamodel approach.

include, for example, optimization models for power plant deployment planning, which must be coupled with a Monte Carlo simulation. However, such approaches fulfill the current requirements for the European security of supply assessment and are even more computationally intensive. Successful integration of machine learning in these approaches can significantly improve the optimization results through greater depth of analysis. Another research field that arises from the application of regression modeling is a deep analysis method developed specifically for the security of supply assessment with electricity. Due to the possibility to generate the results of a large number of scenarios in a short time, complex optimizations of the power plant park, for example in terms of ecology, economy and security of supply, are now also possible.

Funding

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 532148125 and supported by the central publication fund of Hochschule Düsseldorf University of Applied Sciences.

CRediT authorship contribution statement

Justin Münch: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jan Priesmann:** Writing – review & editing, Writing – original draft, Visualization, Software, Resources, Project administration, Formal analysis, Data curation, Conceptualization. **Marius Reich:** Writing – review & editing, Writing – original draft, Methodology, Funding acquisition, Conceptualization. **Marius Tillmanns:** Writing – review & editing, Visualization, Software, Resources, Project administration, Formal analysis, Data curation, Conceptualization. **Aaron Praktiknjo:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Mario Adam:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

Funding: This work is the result of the research project “Artificial intelligence for assessing the security of electricity supply”, which was funded by the Federal Ministry for Economic Affairs and Climate Action (grant number 03EI1022A).

References

[1] Gesetz zur reduzierung und zur beendigung der kohleverstromung (Act to reduce and end coal-fired power generation). German Government; 2020.
[2] Federal Ministry for Economic Affairs and Energy. Kommission „Wachstum, Strukturwandel und Beschäftigung“ (Commission on Growth, Structural Change and Employment); 2019. [March 18, 2024]; Available from: https://tinyurl.com/528wwevh.
[3] Dreizehntes gesetz zur Änderung des atomgesetzes (Thirteenth law amending the atomic energy act). German Government; 2011.
[4] Neunzehntes gesetz zur Änderung des atomgesetzes (Nineteenth law amending the atomic energy act). German Government; 2022.
[5] Federal Statistical Office. Stromerzeugung 2022: Ein Drittel aus Kohle, ein Viertel aus Windkraft (Power generation in 2022: one third from coal, one quarter from wind power); 2023. [March 18, 2024]; Available from: https://tinyurl.com/yce5dw2s.
[6] Praktiknjo A. The value of lost load for sectoral load shedding measures: the german case with 51 sectors. Energies 2016;9(2):116. https://doi.org/10.3390/en9020116.
[7] BNetzA. Bundesnetzagentur - Kennzahlen der Versorgungsunterbrechungen Strom (Key ratios of supply interruptions electricity); 2023. [March 18, 2024]; Available from: https://tinyurl.com/2s5n4ahr.
[8] BNetzA. Bundesnetzagentur - Stand und Entwicklung der Versorgungssicherheit im Bereich der Versorgung mit Elektrizität (Status and development of security of supply in the area of electricity supply); 2023. [March 18, 2024]; Available from: https://tinyurl.com/54p5f5wd.
[9] Zhou Z-H. Machine learning. ed. Singapore: Springer Singapore; 2021.
[10] Sarker IH. Machine learning: algorithms, real-world applications and research directions. SN Comput Sci 2021;2(3):160. https://doi.org/10.1007/s42979-021-00592-x.
[11] Choi H, Park S. A survey of machine learning-based system performance optimization techniques. Appl Sci 2021;11(7):3235. https://doi.org/10.3390/app11073235.
[12] Elsisi M, Amer M, Dababat A, Su C-L. A comprehensive review of machine learning and IoT solutions for demand side energy management, conservation, and resilient operation. Energy 2023;281:128256. https://doi.org/10.1016/j.energy.2023.128256.
[13] Ashraf WM, Dua V. Machine learning based modelling and optimization of post-combustion carbon capture process using MEA supporting carbon neutrality. Digit Chem Eng 2023;8:100115. https://doi.org/10.1016/j.dche.2023.100115.
[14] Ashraf WM, Dua V. Artificial intelligence driven smart operation of large industrial complexes supporting the net-zero goal: coal power plants. Digit Chem Eng 2023;8:100119. https://doi.org/10.1016/j.dche.2023.100119.
[15] Ashraf WM, Uddin GM, Tariq R, Ahmed A, Farhan M, Nazeer MA, et al. Artificial intelligence modeling-based optimization of an industrial-scale steam turbine for moving toward net-zero in the energy sector. ACS Omega 2023;8(24):21709–25. https://doi.org/10.1021/acsomega.3c01227.
[16] Bakhshideh Zad B, Toubeau J-F, Bruninx K, Vatandoust B, Grève ZD, Vallée F. Supervised learning-assisted modeling of flow-based domains in European resource adequacy assessments. Appl Energy 2022;325:119875. https://doi.org/10.1016/j.apenergy.2022.119875.

[17] Nolting L, Spiegel T, Reich M, Adam M, Praktiknjo A. Can energy system modeling benefit from artificial neural networks? Application of two-stage metamodels to reduce computation of security of supply assessments. Comput Ind Eng 2020;142:106334. https://doi.org/10.1016/j.cie.2020.106334.
[18] Nolting L, Praktiknjo A. Can we phase-out all of them? Probabilistic assessments of security of electricity supply for the German case. Appl Energy 2020;263:114704. https://doi.org/10.1016/j.apenergy.2020.114704.
[19] Priesmann J, Münch J, Tillmanns M, Ridha E, Spiegel T, Reich M, et al. Artificial intelligence and design of experiments for resource adequacy assessment in power systems. Energy Strat Rev 2024;53:101368. https://doi.org/10.1016/j.esr.2024.101368.
[20] Sonnet D. Best Practice: möglichkeiten und Grenzen Neuronaler Netze (Possibilities and limits of neural networks). In: Sonnet D., editor. Neuronale netze kompakt. Wiesbaden: Springer Fachmedien Wiesbaden; 2022, p. 71–99.
[21] Da Silva IN, Hernane Spatti D, Andrade Flauzino R, Liboni LHB, dos Reis Alves SF. Artificial neural networks. Cham: Springer International Publishing; 2017.
[22] Østergård T, Jensen RL, Maagaard SE. A comparison of six metamodeling techniques applied to building performance simulations. Appl Energy 2018;211:89–103. https://doi.org/10.1016/j.apenergy.2017.10.102.