IOSUD – Universitatea Politehnica Timișoara
Școala Doctorală de Studii Inginerești

Bayesian Neural Network based Electrical Load Forecasting

PhD thesis – abstract

for obtaining the scientific PhD degree
Politehnica University Timisoara
in the field of Power Engineering

author Mihaela Daciana Mic (Crăciun)
scientific supervisor Prof. Ștefan Kilyeni

month 10, year 2021

Both for “classical”, regulated, electric power systems (PS) and current, “deregulated” systems, characterized by the wider integration of renewable energy sources, having the knowledge in regards to how electricity consumption will evolve is an important element in decision-making.

Regardless of the temporal horizon for which the forecast is performed, the problem can be approached both through "classical" methods, based on mathematical modeling and known linear or nonlinear optimization techniques [Kilyeni2015], as well as through methods that use modern artificial intelligence techniques, in particular Artificial Neuronal Networks [Kumar 2016], [Singh 2017], [Hsu2018], [Jarndal2020]. In all situations, it is extremely important to know the history of the evolution of consumption and other elements that influence consumption for a period of time significantly longer than the forecast.

In recent years, the use of Bayesian theory, Bayesian networks and Bayesian neural networks [Nabney2002], [Mackay2003], [Bolstad2004], [Koch2007], [Russel2010] for solving applications in the field of electrical power engineering is quite notable. Most applications focus on the reliability of power plants and systems [Sykora2016], [Lorencin2017] and consumption forecast [Sun2019], [Bessani2020], [Sarajcev2020]. Other areas of interest: PS state estimation [Pegoraro2017], PS stability [Chevalier2019], fault diagnosis and location [XuB2019], etc.

In this context, the topic of the PhD thesis is part of the current concerns in the field of transmission management, distribution and consumption of electricity. Generally speaking, two key objectives were considered: forecasting activity (electric energy consumption and load curves) and the use of artificial intelligence techniques (mainly Bayesian artificial neural networks) to obtain forecasts.

Theoretical analyses are completed through original techniques for solving elaborated mathematical models, implemented in their own software tools, which efficiently use the possibilities offered by the various programming environments and current computer systems.

We started from relatively simple cases, in order to validate the proposed methods and "calibrate" the software tools. For the same purpose, comparative studies were performed with the results obtained in other PhD theses [Deacu2015], [Chiş2015]. Next, real situations were analyzed, targeting distribution operators in Romania: Enel, Electrica, Delgaz Grid, etc.

For space related reasons, for the actual applicative part of the thesis, only part of the case studies regarding Enel Distribution Banat were selected. They refer both to the entire distribution company and to the main territorial unit network (TUN) components: Arad, Deva, Reșița and Timișoara. A series of results regarding the 110 kV/m.t transformer stations are also presented from TUN Timișoara. 5 significant 110/20 kV substations were selected (3 from Timișoara, one from an important city of Timiş County and one that supplies an oil exploitation facility): Bucovina, IMT, Musicescu, Deta and Satchinez.

The obtained results and the formulated conclusions are of a special utility both for the distribution operators, in general, and for Enel Distribution Banat, in particular.

The PhD thesis extends on 234 pages, being structured in 8 chapters, a preface, 3 appendices (CD, 82 pages) and an ample bibliography. The thesis also contains 213 figures, schemes and histograms, and 264 tables respectively. The bibliography contains 208 titles, among which some are noted to be significant works in the field, both old and new, of local and foreign origin in the literature.

Chapter 1 has an introductory character. The first part entails the framing and the justification for the PhD thesis’ topic within the context of the current evolution of the field of electric power system engineering, both in the local and wider scale. The second part represents a succinct presentation for each of the following chapters of the thesis. The last part of the chapter highlights both the way of capitalizing the research carried out in the PhD thesis (papers published in specialized journals or in the volumes of international conferences, scientific research or technical assistance contracts, calculation programs) but also the usefulness of the results obtained for electrical energy distribution operators (from Romania, but not only) and for other economic agents (especially those dealing with the implementation of renewable energy sources). Finally, the perspectives opened by this PhD thesis on the possible directions for the continuation and extension of investigations are underlined.

Chapter 2 presents a series of basic theoretical notions related to the Bayesian approach, Bayesian networks (BNs), Bayesian neural networks (BNNs) and related elements. These notions are necessary both for understanding the aspects presented in the following chapters, related to the use of Bayesian concepts and models in the field of electric power engineering, and the mathematical model of energy consumption forecasting, and the related software (Chapter 6). The first part of the chapter reviews a series of elements of probabilities and statistics: conditional probabilities, frequentist vs. subjectivist interpretations of probabilities, Bayes' theory, the laws of probability, etc. The second part refers to machine learning (ML) in the context of probabilistic modelling: Bayesian machine learning (BML), posterior and marginal probability approximation methods, Bayesian inference, etc. The last part is dedicated exclusively to Bayesian ANNs: comparing "classic" ANNs with the Bayesian ones, the principles behind Bayesian ANN training, Bayesian optimization of the control parameters (the notion of evidence framework, including its numerical approach)

Chapter 3 provides an overview of BN applications in the field of PS engineering. A wide range of examples are reviewed, starting with the reliability of electricity transmission and distribution networks, of the PS as a whole, continuing with the estimation of the PS state, with the PS stability analysis, with the diagnosis and location of faults in electrical networks, fault diagnosis related to transformers and generators, with the estimation of the value of the parameters of the network elements (power lines, transformers) etc. The most numerous applications are found in the field of electric energy and power load forecasting. They will be treated separately, in chapter 4, given that the forecasting using BNN is the subject of this PhD thesis.

The increase in the frequency of cybernetic attacks upon SCADA systems has resulted in a decrease of reliability associated to PS, thus [Zhang2014] has taken into consideration six hypothetical cyber-attacks, using a Bayesian model to evaluate the probability of success of such attacks on a SCADA system, which would have the ultimate result the tripping of the system's breakers. A model for the rate of forced interruptions is proposed, taking into consideration the successful attacks on generators and electrical energy transmission lines. Results from the IEEE test system in reliability studies have proven that PSs do become less reliable as the frequency of successful cyber-attacks increase. [Borges2016] uses a specific representation model for statistically dependent time-varying quantities that can be used within the context of mixed methods of reliability evaluation, through non-sequential Monte Carlo simulations. The model has been developed by combining a non-parametric estimation method for the probability-density functions with continuous variables measuring the non-linear statistical dependency and graphically representing the conditioned probability given by the Bayesian Network. The software itself refers to the IEEE test system for reliability studies, applied to 27 fictional wind farms, with varying wind parameters. Simulations have shown that the model can accurately reproduce the historical data and underline the importance of types of correlated factors on the reliability index. In other words, this model is at the same level of precision as the sequential Monte Carlo simulation is. The model can thus be used to represent the correlation between the electrical energy outputted by wind farms and the water flow in a hydroelectric system. [Sykora2016] presents the use of BNs to assess the risks related to the reliability of a generator set within a thermal power plant. In order to implement the statistical approach, more attention is paid to the data related to the failure rates, obtained from the previous history and based on the opinion of the specialists. The information discussed concerns the entire thermo-mechanical and electrical chain: boiler, turbine, generator. The conclusion is that BNs are an effective tool for risk and (in)availability analysis, providing important information on the maintenance process, the necessary repairs and the situations in which it is necessary to replace some equipment. BNs also facilitate the evaluation of the tendency to modify the technical parameters describing the operation of the component installations and equipment. A similar approach is described in [Lorencin2017].

In [Massignan2019] the estimation of the static state of the electric power systems is done by an approach using Bayesian inference. Given the widespread use of the PMU (Phasor Measurement Unit), there is the problem of combining the measurements obtained in this way with those offered by conventional SCADA systems, given the significant differences in sampling rates and accuracy. A two-step approach is proposed. As a first stage, an initial "classical" WLS estimate is made based on the measurements provided by the SCADA systems. In the second stage, the information provided by the PMU is used to achieve a probabilistic interpretation of the solution in the first stage. There is an improvement in the quality of the estimate even in the conditions of a lower number of PMUs, including for nodes that are not monitored by the PMU. [Mestav2019] aims to estimate the state of electricity distribution systems that have observability problems. The proposed method involves in the first phase the distributed learning of stochastic power injections in the nodes of the system. A Monte Carlo method is then used to train an ANN with several hidden layers. Finally, a Bayesian algorithm for detecting and filtering erroneous data is developed. The obtained results highlight the advantages of such an approach compared to the use of pseudo-measurements.

[Augutis2012] presents a method for assessing the high disturbance stability of a PS for various operating regimes using BNs. It is essentially a hybrid analysis technique, which combines the classical method of analysis by modeling in detail the power system with an BN-based approach. The latter uses an estimation model to determine the stability characteristics for certain synchronous generators. In this way all generators can be analyzed by estimation based on the Bayesian approach instead of an analysis using detailed modeling of the system. The major advantage is the substantial reduction of the computational effort, without significantly altering the accuracy of the results. The quality of the results can be improved by grouping the generators based on coherence, assessed on the basis of a correlation coefficient of the swing curves (time variation of the internal angle). Unlike previous works, in [Seppanen2016] and [Ma2013] the analyses refer to the steady state stability of the electric power system (PS), including in the analysis model a BN. In [Chevalier2019] the aim is to locate the source of the oscillations, in the conditions of some uncertainty regarding the parameters of the synchronous generators and the accuracy of the measurements obtained from PMU, elements that justify the use of BN. [Vakili2015] aims to analyze the stability of voltage using a direct Lyapunov method, combined with the Bayesian quickest change-point detection.

In [Li2014] the Bayesian method of fault diagnosis is based on the fault isolation analysis. The model is built from two points of view – breaker isolation and protection isolation, which directly reflects the mode of action of the breaker and the protection at the appearance time of the fault. At the same time, dividing the breakers into three levels, depending on the type of protection and the operating time, reduces the number of suspected switching equipment in the affected area, improving the efficiency of the fault diagnosis.

To deal with a lot of uncertain information existing in the grid when faulted, [WangT 2015] uses a special Bayesian network—Noisy-or, Noisy-and nodes model to build object-oriented Bayesian networks. The connected graph connected without circuits associated to the BN comprises both Noisy-OR and Noisy-AND peaks. If a given peak has a value of "false" when all previous peaks have a value of "false", then it is a Noisy-OR type (similar to the logical OR definition, except that it cannot be state that the peak has the value "true" if one of the previous peaks has the value "true"). If a given peak has a value of "true" when all previous peaks have a value of "true", then it is a Noisy-AND peak (similar to the logical definition of

AND, except that it cannot be stated that the peak has the value "false" if one of the preceding peaks has the value "false"). An error back propagation algorithm is used to train the Bayesian network so as to update the network's parameters. Finally, a case simulation is presented to prove the accuracy and validity of this method for power system fault diagnosis. To minimize the mean square deviation between the calculated and the measured level, the conjugate gradient method is used [Kilyeni2015]. Overall, it is a simple and fast fault location model, using the element (sometimes uncertain) about the action of the protections and the state of the breakers.

In [XuT2010] a method of locating faults in rural distribution networks using Bayesian inference is presented. Based on the information provided by the disturbance calls (mostly unsafe and incomplete) and the experience of the operating personnel (with a relatively low degree of confidence), a probabilistic model of learning, reasoning, based on the Bayesian method was developed. The fault location algorithm was implemented in GIS systems related to rural distribution networks. The fault diagnosis in the case of hydrogenators is addressed in [XuB2019], taking into account both electrical and mechanical or hydraulic causes. The methodology used is based on an expert system and the use of BN. A complete Bayesian model of fault diagnosis has been developed, which is based on in-depth knowledge of the vibrations that occur in various types of faults and the associated fault characteristics. [Zhao2010] presents a fault monitoring and diagnosis system for high power transformers and autotransformers, based on a multi-agent type system, combined with a Bayesian classification algorithm. From a phenomenological point of view, the diagnosis is based on the analysis of the resulting gases in the oil tank in the case of some internal short circuits, which effect is the vaporization of the oil determined by the electric arc. As in the previous case, the [Zheng2010] approach to fault diagnosis in transformers is based on the analysis of the resulting gases in the oil tank in the case of internal short circuits, which have the effect of vaporizing the oil due to the electric arc. This paper compares several types of Bayesian classifiers (NB - Naïve Bayesian classifier, TBA Tree Augmented Naïve Bayesian classifier, GBN - General Bayesian Network classifier), highlighting the advantages and disadvantages. [Zhou2012] presents a model for simulating the lifetime of transformers in electricity distribution networks. Their failure is a fairly rare event, which means a small amount of known initial data. To overcome this aspect, the lifetime of the transformer is treated as a random variable, with a certain probability distribution. Applying this probabilistic model for a group of transformers one can estimate the number of transformers that need to be replaced. The paper proposes that the initial model be made on the basis of information related to the operational safety of a large number of similar transformers. A Bayesian upgrade procedure is then used to incorporate prior (previous) knowledge of actual failures into the original model, resulting in an advanced transformer lifetime model. Finally, a sequential updating of the model is proposed, which leads to a dynamic way of improving the model of the transformer lifetime.

Chapter 4 presents the problems related to the electricity load forecast (peak power, hourly power, energy consumption, load curves, etc.) and the methods used to obtain the forecast. The first part of the chapter includes general aspects related to forecasting, classification of methods used according to various criteria, insisting on the time horizon to which the forecast refers. The actual presentation of the methods follows, depending on the time horizon to which it refers, based on a consistent bibliographic study. Both "classical" methods, based on mathematical modeling, and "modern" ones, more recently, are followed, using techniques of artificial intelligence, fuzzy logic, expert systems, "support vector machine" (SVM), etc.

A separate, consistent subchapter is dedicated to the methods that use BNN, which is the subject of this PhD thesis. The vast majority of examples refer to the electrical load forecast consumed, but there are also situations in which the object of the forecast is the price of electricity, the power generated by wind or photovoltaic power plants, wind speed, solar irradiance, etc. The first mention of the use of the Bayesian approach to power consumption forecasting appears in [Bakirtzis1997], where a mixed Bayesian predictor is proposed, made by combining a predictor based on the use of an ANN with two other predictors specific to linear regression. The application referred to the short-term load forecast in Greece.

Examples of applications for short-term load forecast (STLF):

- In [Ning2010] STLF is performed using a backpropagation ANN with Bayesian training. This type of learning facilitates the obtaining of the most probable values of the hyper-parameters, which lead to an optimal architecture for a backpropagation ANN. The performance testing of the model was performed using real consumption data from Guizhou Province (China), both for network training and for testing the forecasts, the results proving its superiority compared to conventional backpropagation ANN. It was noticed an increase in learning speed, convergence and forecast accuracy.
- A more complex method is presented in [Ghayekhloo2015], which uses an algorithm for preprocessing input data in order to improve the quality of the forecast. A discrete "wavelet" transformation is used to decompose the load components into appropriate resolution levels, based on an entropy criterion, followed by a regression analysis, resulting the best input data set. A correlation analysis with an ANN provides a first estimate of the predicted values associated with the input quantities, followed by a standardization procedure that takes into account the degree of correlation of the output quantities with the set of associated input quantities. In the end, the most appropriate input data for Bayesian ANN are found. A genetic algorithm is used to optimize the weighting coefficients of the various components of the forecast and to minimize the errors of the predicted values. The assessment of the performance and accuracy of the proposed method for STLF is made with the help of a well-known consumption database ("New England load data"), the conclusions being positive. A similarly complex approach appears in [He2019], which presents a hybrid STLF method using recurrent LSTM (Long Short-Term Memory) BNNs.
- [Dagdougui2019] aims to provide a short- and very short-term load forecast in areas of "smart" buildings, comprising several heterogeneous blocks in terms of functionality, using ANN-based models. The paper pursues three objectives: the evaluation of ANN performance considering two backpropagation training techniques – the Bayesian regularization and the Levenberg-Marquardt method; comparative analysis of the model performances for the "hour-ahead" and "day-ahead" forecasts related to the various types of buildings; analysis of the influence of ANN structure (number of hidden layers and neurons in these layers, number of inputs and training data sets) on the accuracy of predicted values. The effectiveness of the proposed method is demonstrated for the case of a neighborhood in downtown Montreal (Canada).
- In [Sarajcev2020] STLF is associated with a technique of clustering consumer data (along with climate data), using an approach based on Bayesian inference. The concrete application refers to the power consumption forecast for the city of Newcastle (Australia).
- [Sun2019a] presents an interesting comparative study of probabilistic STLF methods. The following methods were considered: Bayesian estimation, low-order Bayesian estimation, Ridge type regression (a technique for estimating the coefficients of multiple regression models for situations where independent variables have a high degree of correlation), type estimation LASSO (Least Absolute Shrinkage and Selection Operator – a regression analysis method that performs both variable selection and regularization, in order to improve the quality of the forecast and the degree of interpretation of the statistical model) and ANN with supervised learning. The conclusion of the study highlights that the low-order Bayesian estimation leads to the best results.
- The approach in [Bessani2020] starts from the finding that for residential consumption the uncertainty of the data is much more pronounced than in the case of aggregate consumption at medium or high voltage. In this context, the paper proposes a multivariable model for the very short-term load forecast, based on the use of BN. It takes into account the previous evolution of consumption, climatic factors, socio-economic ones and consumption patterns.

Concrete analysis refers to a total of over 1000 residential consumers in Dublin (Ireland). Comparison with the results obtained by other methods highlights the fact that the model based on the use of BN has superior performance. A similar concern is presented in [Gilanifar 2020], where BNNs are associated with a multitasking learning (training) model. Each task refers to training the STLF model for a specific residential consumer, and "multitasking" means using a combination of individual models for the BNN training process. A more complex "multitasking" model, associated with a Bayesian optimization process, is described in [Yang2020].

Examples of applications for medium and long-term load forecast:

- In [Rivero2015] a long-term load forecast is made. An ANN with Bayesian inference is used, the only input quantity being the previous evolution in time consumption. The ANN output provides the consumption forecast for the coming months. The concrete application refers to the forecast of power consumption for the PS of Argentina. Similarly, [Silva2019] presents a long-term load forecast application for the PS in Brazil, the Bayesian inference being used to estimate the model parameters, which allows the inclusion of the degree of uncertainty of the forecast made with the proposed model.
- The long-term load forecast of electricity is the subject of the study in [Yuan2017], [He2018], [Tang2019] and [Ahmadi2020]. In order to take into account the degree of consumption uncertainty a fuzzy Bayesian approach is proposed. The deterministic results are replaced with the probabilistic ones, being also indicated the confidence interval (the most probable value, accompanied by the minimum and the maximum value).

Examples of power forecast produced by wind and photovoltaic power plants:

- [Yang2013] presents a practical approach to the short-term forecast, in a probabilistic way, of the power generated by a wind farm. The proposed method is based on a sparse Bayesian learning algorithm, which leads to a probabilistic expression of the forecasts results, obtained based on the estimation of the probabilistic density of the weights of the gaussian functions. Due to the "non-stationary" character of the evolution over time of the data, of the information about the generated power, a strategy is proposed based on the decomposition of the time series into components with an increased predictability, using a discrete wavelet transformation. The effective forecast with the Bayesian lacunar algorithm is performed separately for each component, the final result being obtained by summing the partial ones. The concrete application refers to a wind farm in Oklahoma (USA), demonstrating the effectiveness of the proposed method. In [WangY2019] the Bayesian lacunar learning algorithm is used for wind speed forecasting.
- They are on the same line concerns in [Lin2019], noting that a multi-model method is used. In parallel with the Gaussian probability distribution function, the beta (β) type is also used. The testing of the proposed model was performed using the data set GEFC (Global Energy Forecasting Competition) 2014, which includes both the evolution over time of the power generated and the wind speed (at a height of 10 m and 100 m above the ground) for 10 different areas. The data set, from hour to hour, refers to a period of 10 years, being divided into two sub-sets of training and a set of validation of the forecast. The first learning sub-set is used to train the components of the model, while the second is to optimize the weights of the component models.
- Another 20-year period for known data (for IESO - Independent Ontario Electricity System Operator) was also used in [Sahu2019] for the short-term forecast of power generated by wind and photovoltaic sources. Some of those data were used to train ANN, the rest to assess the quality of the forecast. The training was performed by Bayesian regularization, for the time series being taken into account the nonlinear automatic regression, the nonlinear automatic regression with exogenous input data and the input-output model (the second being considered the best).

- In order to assess the risks in the operation of distribution networks in the conditions of large-scale penetration of distributed photovoltaic sources, in [Tao2016] it is proposed to use a dynamic BNN for the probabilistic forecast of power generated by photovoltaic plants. The application refers to the IEEE test system with 53 nodes. A similar approach is presented in [Silva2017], with applications targeting concrete situations in Brazil's PS.
- In [Panamtash2020] a multi-variable Bayesian probabilistic model is proposed for the forecast of the power generated in photovoltaic power plants. In addition to the time series related to the power generated, the dependence on weather conditions (temperature) is also highlighted. For the training of the network, the known data from 2015 are used for a photovoltaic power plant from the USA, the quality of the forecast being verified with the help of the data from 2016 (also known).

Chapter 5 aims to present artificial neural networks (ANN) by providing a theoretical basis for the method chosen to solve the problem for electrical energy consumption and power load forecast (as is presented in chapter 6). The first part of the chapter is a review of all the general aspects regarding ANNs: basic terminology, the structure of an ANN, classification of ANNs. The second part aims to present the structural model of an ANN, starting from the artificial neuron, explaining in more detail the architecture of the ANN. A specific paragraph is dedicated to learning (training) techniques used by ANN – supervised and unsupervised respectively. The last part is dedicated to presenting the perceptron type ANN, with both one and more layers, a special attention being diverted to backpropagation.

The algorithm specific to backpropagation networks has two main steps that have to be followed:

- *A direct pass through the network*, from the inputs to the outputs, in which the ANN is activated and the values of the outputs are determined;
- *A backwards pass through the network*, from the outputs towards the inputs, the determined outputs being compared to the outputs from the examples and a certain error is being estimated; this estimation is then propagated backwards and used to update the weights.

Synthetically, the algorithm of this type of ANN is presented as follows:

❒ *Initialization*

The weights and the biases are randomly initialized with values that are not zero, distributed in a confined interval (for example [–0,1; 0,1] or $\left[-\frac{2,4}{NI};\frac{2,4}{NI}\right]$, where *NI* represents the number of inputs in the ANN).

❒ *Construction of an anterior epoch*

An epoch represents the processing of all examples from the training set. The training of the network entails the passing of multiple training epochs, one single epoch being insufficient.

The weights will be adjusted only after the examples which constitute the training set have been crossed. The weight’s gradients and the current error being initialized with 0.

$$\Delta w_{ij} = 0\,; \quad E = 0 \tag{5.3.6}$$

❒ *Forward propagation*

- On the network input an example from the training set is being applied.
- The outputs for the hidden layer neurons are calculated:

$$y_j(p) = f\left(\sum_{i=1}^{n}\left[x_i(p)\cdot w_{ij}(p) - T_j\right]\right), \quad j = 1, 2, \cdots, b \tag{5.3.7}$$

where: n – the number of inputs of neuron j from the hidden layer; f – represents the sigmoid activation function; p – represents the current learning set.

- The real values of this network are being processed:

$$y_k(p) = f\left(\sum_{j=1}^{b}\left[x_{jk}(p)\cdot w_{jk}(p) - T_k\right]\right), \quad k = 1,2,\cdots,m \tag{5.3.8}$$

  where *m* represents the number of inputs of neuron *k* from the output layer.
- The epoch error is updated:

$$E = E + \frac{\left[e_k(p)\right]^2}{2} \tag{5.3.9}$$

❒ *Backwards error propagation and adjustment of the weights*

- The error gradients are solved for the neurons from the output layer:

$$\delta_k(p) = f'\cdot e_k(p) \tag{5.3.10}$$

  where *f'* is the derivate of the activation function, and the error is:

$$e_k(p) = y_{d,k}(p) - y_k(p) \tag{5.3.11}$$

  where: $y_{d,k}(p)$ – the real value of the output *k*; $y_k(p)$ – the calculated value of the output *k*.

  In the case of the sigmoid function its derivate is the following:

$$f'(p) = \frac{2\cdot a\cdot e^{-a\cdot x}}{(1+e^{-a\cdot x})^2} = \frac{a}{2}\cdot[1-f(x)]\cdot[1+f(x)] \tag{5.3.12}$$

  and the error gradients for the neurons from the output layer thus become:

$$\delta_k(p) = y_k(p)\cdot[1-y_k(p)]\cdot e_k(p) \tag{5.3.13}$$

- The weight gradients from the hidden and output layer are updated:

$$\Delta w_{jk}(p) = \Delta w_{jk}(p) - y_i(p)\cdot\delta_k(p) \tag{5.3.14}$$

- The error gradients for the neurons from the hidden layer are calculated:

$$\delta_i(p) = y_i(p)\cdot[1-y_i(p)]\cdot\sum_{k=1}^{m}\left[\delta_k(p)\cdot w_{jk}(p)\right] \tag{5.3.15}$$

  where *m* is the number of outputs that the network has.
- The weight gradients between the input and the hidden layer are updated:

$$\Delta w_{jk}(p) = \Delta w_{jk}(p) + x_i(p)\cdot\delta_i(p), \quad i = 1,2,\cdots,n, \quad j = 1,2,\cdots,b \tag{5.3.16}$$

❒ *Transition to a new iteration*

If test vectors still exist in this current training epoch, *Forward propagation* and *Backwards error propagation and adjustment of the weights* is executed until all cases are exhausted.

❒ *Verifying the termination condition*

If a training epoch is over, the weights of all connections are updated based on the gradients (η – the learning rate):

$$w_{ij}(p) = w_{ij}(p) + \eta\cdot\Delta w_{ij}(p), \quad i = 1,2,\cdots,n, \quad j = 1,2,\cdots,b \tag{5.3.17}$$

It is tested whether the completion criterion ($E < E_{max}$) has been met or if a maximum number of training epochs has been reached (the maximum number of iterations has been exceeded). If none of the conditions has been met, the next step becomes the creation of a new training epoch.

If one of the termination criteria is met however, the algorithm is considered successful (the network was trained), if the termination criteria is not met, the network is still not trained. In order to find a solution, either the speed of the training is modified, or different initial values for the weights are chosen (or both measures are applied at the same time).

Chapter 6 refers to the mathematical model for electrical energy consumption and power load forecast, as well as to the software tool that was built on that said model. The first part of the chapter goes into details of the Bayesian technique in the context of the forecast application, explaining the principles behind Bayesian inference, ANN priors, how to calculate error functions and the gradient, evidence procedure, the forecast and the error bars. The second part presents the optimization technique used within the application – the scaled conjugated gradient method (SCG). The last part entails details about the software instrument that encompasses all mathematical models presented in this chapter. The software tool was developed in the Matlab environment by effectively utilizing all facilities (interface and portability) specific to modern machines and operating systems. The software application made use of various functions from the Netlab toolbox within Matlab [Nabney2002] which proved to be useful in the simulation of specific ANN algorithms: mlpprior.m, mlp.m, mlpinit.m, netopt.m, scg.m, mlperr.m, errbayes.m, mlpgrad.m, gbayes.m, mlpfwd.m, evidence.m, mlpevfwd.m.

The Scaled Conjugate Gradient Method (SCG) is part of the larger group of conjugate gradient methods, with the classical version being based on the fact that the vector g shows the direction of the largest increase of $f(x)$. Geometrically, the gradient represents the orthogonal vector to the outline of $f(x)$, which is passes through a point x. The most accentuated decrease (searching for the minimum) of $f(x)$'s value is given by $-g$.

Under these conditions, the method's algorithm is the following [Kilyeni2015]:

a) The value of x is initialized with x_0, chosen based on experience;

b) For any iteration j, j=1, 2, … the value of f is solved at the current point

$$f_j = f(x_j) \tag{6.2.1}$$

while the movement direction d_j becomes:

$$d_j = -g_j \tag{6.2.2}$$

where g_j represents the current value of the gradient

c) For the same iteration j, the new point x_{j+1} is solved:

$$x_{j+1} = x_j + \alpha_j \cdot d_j \tag{6.2.3}$$

where the scalar α_j, which is generally determined through the use of parabolic interpolation, shows the size of the movement in the direction of d_j:

$$\alpha_j = \frac{h}{2} \cdot \frac{3 \cdot f^0 - 4 \cdot f^1 + f^2}{f^0 - 2 \cdot f^1 + f^2} \tag{6.2.4}$$

where the scalar h represents the pace of the search, f^0, f^1, f^2 being the values of f in the points $x^0 = x_j$, $x^1 = x_j + h \cdot d_j$, $x^2 = x_j + 2 \cdot h \cdot d_j$;

d) The calculation is considered solved when the absolute value of the gradient becomes 0 (the error threshold being ε **)**:

$$\left\| g_j \right\| < \varepsilon \tag{6.2.5}$$

The main disadvantage of the classical gradient method is related to the orthogonality of the movement directions for two successive iterations, the effect being a crisscross movement towards the minimum value. The convergence towards the minimum that results from this is slow.

The conjugate gradient method eliminates the disadvantage mentioned above [Kilyeni 2015]. The algorithm of the method is similar to that of the simple gradient, the only difference being related to the determination of the direction of movement, which is now of the form:

$$d_j = -g_j + \beta_j \cdot d_{j-1} \tag{6.2.6}$$

where the β_j coefficient takes into consideration the prior "history" ("mixes" in the direction of the current step a weighted correction based on the direction of travel from the previous step), being defined as:

$$\beta_j = \frac{g_j \cdot g_j^T}{g_{j-1} \cdot g_{j-1}^T} \tag{6.2.7}$$

In both versions of the gradient method priorly presented, for each step taken, the value of the scalar α_j needs to be determined, scalar which indicates to the size of the movement in the direction d_j. The Scaled Conjugate Gradient Method [Moller1993] eliminates this disadvantage. It offers a way to choose conjugated search directions without performing the search in the direction d_j and without solving the hessian matrix *H* (matrix of 2nd order derivatives of the function *f*).

Under the hypothesis of using a hessian matrix (in Newton type methods, which make use of the second order derivates), α_j can be calculated as following:

$$\alpha_j = \frac{g_j^T \cdot d_j}{d_j^T \cdot H \cdot d_j} \tag{6.2.8}$$

In order to cut down on the computational time and avoiding a search in the direction d_j, the approximation of $H \cdot d_j$ is proposed, based on a finite differential formula. In this regard we can consider σ_0 a small positive quantity, that can be expressed as following:

$$\sigma = \frac{\sigma_0}{\|d_j\|} \tag{6.2.9}$$

Developed as a Taylor series, withholding terms until the first order derivate (inclusively), the following is obtained:

$$\nabla f(x_j + \sigma d_j) \approx \nabla(x_j) + \sigma \cdot H \cdot d_j \tag{6.2.10}$$

Thus:

$$H \cdot d_j \approx \frac{\nabla f(x_j + \sigma d_j) - \nabla f(x_j)}{\sigma} \tag{6.2.11}$$

Defining:

$$\theta_j = d_j^T \cdot \left[\frac{\nabla f(x_j + \sigma d_j) - \nabla f(x_j)}{\sigma}\right] \approx d_j^T \cdot H \cdot d_j \tag{6.2.12}$$

If the function *f* does not have a quadratic form, then H is possible not to be positively defined and thus the value of *f* is bound to increase (instead of decreasing). This thing can be prevented by adding to the hessian matrix H a multiplicative term θ_j and the unit matrix: $H + \theta_j \cdot I$. Thus, the actualized form is obtained:

$$\alpha_j = \frac{g_j^T \cdot d_j}{d_j^T \cdot H \cdot d_j + \theta_j \cdot \|d_j\|^2} \tag{6.2.13}$$

If θ_j has a large value, then the step α_j is small. This represents approaching the issue via a *trust region model*, because the model (for the function) is to be trusted in a small region around the search point.

In order to obtain a minimum for the function *f*, the hessian matrix must be positively defined (at least $d_j^T \cdot H \cdot d_j > 0$).

The following notation is introduced:

$$\delta_j = d_j^T \cdot H \cdot d_j + \theta_j \cdot \left\| d_j \right\|^2 \tag{6.2.14}$$

If $\delta_j < 0$ then it is a certainty that $d_j^T \cdot H \cdot d_j < 0$ is true, thus θ_j must be increased. Moller's [Moller1993] original algorithm uses:

$$\overline{\theta}_j = 2 \cdot \left(\theta_j - \frac{\delta_j}{\left\| d_j \right\|^2} \right) \tag{6.2.15}$$

Next, the following are set

$$\overline{\delta}_J = \delta_j + (\overline{\theta}_j - \theta_j) \cdot \left\| d_j \right\|^2 = -\delta_j + \theta_j \cdot \left\| d_j \right\|^2 = -d_j^T \cdot H \cdot d_j > 0 \tag{6.2.16}$$

If the condition is also considered a quadratic function, it is solved as such:

$$\Delta_j = \frac{f(x_j) - f(x_j + \alpha_j d_j)}{f(x_j) - f_Q(x_j + \alpha_j d_j)} \tag{6.2.17}$$

where f_Q represents the local quadratic approximation of f on direction d_j:

$$f_Q(x_j + \alpha_j d_j) = f(x_j) + \alpha_j \cdot d_j^T \cdot g_j + \frac{\alpha_j^2}{2} \cdot d_j^T \cdot H \cdot d_j \tag{6.2.18}$$

If $\Delta_j \approx 1$, this approximation is appropriate and θ_j can decrease. If Δ_j has a small value, then θ_j must increase. Simplifying relation (6.2.17) yields:

$$\Delta_j = \frac{2 \cdot \left[f(x_j) - f(x_j + \alpha_j \cdot d_j) \right]}{\alpha_j \cdot d_j^T \cdot d_j} \tag{6.2.19}$$

Equation (6.2.19) can be applied only through the use of a gradient, without calling for a higher order derivate.

Chapter 7 is entirely original, being the main application part of the PhD thesis. The results obtained on the load forecast of power and load curves using artificial intelligence techniques – Bayesian ANN are presented. We started from simple cases and from the test database, in order to validate the proposed methods and "calibrate" the software tools. For the same purpose, a series of comparative studies were performed with the results obtained in [Deacu2015] and [Chiş2015]. Next, real situations were analyzed, referring to distribution operators in Romania: Enel, Electrica, Delgaz Grid, etc. Due to space reasons, only a series of case studies targeting Enel Distribution Banat were selected for the PhD thesis. Some of the results are presented in extenso, the rest in summary, the details being provided in the Appendices (in electronic form). The results obtained for other electricity distribution systems in Romania have been and are used in the contracts developed in recent years by the Research Center for Analysis and Optimization of Electric Power System (EPS) within Politehnica University of Timisoara, the beneficiaries being Enel Distribution Banat, Electrica Muntenia Nord, Delgaz Grid (important operators of electricity distribution in Romania) and economic entities with concerns in the field of implementation of renewable energy resources [UPT 2017], [UPT2018], [UPT2019], [UPT2020a], [UPT 2020b].

The first subchapter has the role of validating the calculation models and software tools developed in the PhD thesis, by making comparisons with the results of [Deacu2015] and [Chiş2015], in order to demonstrate the superior qualities of methods using Bayesian ANN. The first set of comparisons concerns the operator Enel Distribuție Banat and TUN (Territorial Unit Network) components: forecast of load curves for the most significant summer day – June 21, using the values measured from hour to hour (1 o'clock, ..., 24 o'clock) of the active power consumed. The data for 10 years (2001-2010) were used for ANN training, and those for the next 3 years (2011-2013) for verifying the predictions obtained.

Two of the 5 analyzes performed were selected (TUN Resita and the Enel Banat ensemble), resulting in the following conclusions:

- in all situations, Bayesian ANNs lead, in total, to better results than those in [Deacu2015] (with 12-14% for TUN Reșița and 6-10% for Enel Banat) and [Chiş2015] (with 46% for TUN Reșița and 36% for Enel Banat);
- in terms of the breakdown over the 3 years (2011, 2012, 2013), compared to [Deacu 2015] for TUN Reșița, the most pronounced improvement appears for 2011, decreasing constantly in 2012 and 2013, (both for the hourly forecast and for load curve assembly);
- in the same context, for Enel Banat the situation is similar for the hourly forecast, and for the whole load curve it is practically constant every year;
- there are two exceptions to the annual components, when the results obtained with Bayesian ANN are slightly lower, without affecting the overall conclusion from the first point.

The second set of comparisons concerns power substations of 110/20 kV. from TUN Timisoara: forecast of load curves defined by the consumed powers at a certain time of a certain day for each of the 12 months of the year. The first 6 years of the 8 for which consumption data are known (2006-2011) were used for ANN training, and the last 2 years (2012 and 2013) for verifying the obtained forecasts. One of the 5 stations analyzed in [Deacu2015] was chosen: 110/20 kV Victoria, with forecasts made for 9 and 21 o'clock on the last Thursday of each month, the first day of Tuesday and the second day of Wednesday, resulting in the following conclusions:

- in all situations, the methods using Bayesian ANN lead, in total, to better results than in [Deacu2015], the improvement being 13-28% for the overall load curve forecast, respectively 9-39% for the monthly forecast;
- there are 2 cases (for Tuesday) when the improvement is very high (monthly forecast for 9 o'clock (77%) and that of the whole load curve, 9 pm (44%), explainable by the extremely weak correlation of the known load curves);
- also, for Tuesday (9 o'clock) the only situation appears when the result obtained with Bayesian ANN for one of the years for which the forecast was made (2013) is lower (by 3.2%) than in [Deacu2015], without affecting the conclusion from the global comparison.

The second subchapter presents a series of forecast studies carried out both for the entire distribution network within Enel Distribuție Banat and for TUN components: Arad, Deva, Reșița and Timișoara. The load curves are forecast for the most significant summer day – June 21, using the known values of the hourly average power from hour to hour (1 o'clock, 2 o'clock, ..., 23 o'clock, 24 o'clock). The first 10 years (out of the 13 for which consumption data are known), 2006-2015, were used for ANN training, and the last 3 years, 2016-2018, for verifying the forecasts obtained. Approaches based on the use of Bayesian ANNs were applied for the prediction of load curves:

- forecast for the whole load curve - 24 hours (ANN load curve);
- individual forecast for each hour (hourly ANN).

The last part of the subchapter presents a study to predict the quality of forecasts based on known data, using the method from [Deacu2015], based on finite differences. Finally, some comments and conclusions are highlighted, with a more general or particular character, aiming at both the concrete results of the forecasts and the ANN used. Special attention is given to the comparison of the obtained results with the two methods, to highlight their quality, to appreciate the influence of the degree of correlation of the load curves on the quality of the predictions made.

For example, the results for TUN Arad are presented. Figure 7.2.1 shows the data on the load curves for a period of 10 years (2006-2015), related to the most significant summer day. They were used to train ANN. Figure 7.2.2 shows the data on the load curves for the years 2016, 2017 and 2018, used to verify the results of the forecasts for those years.

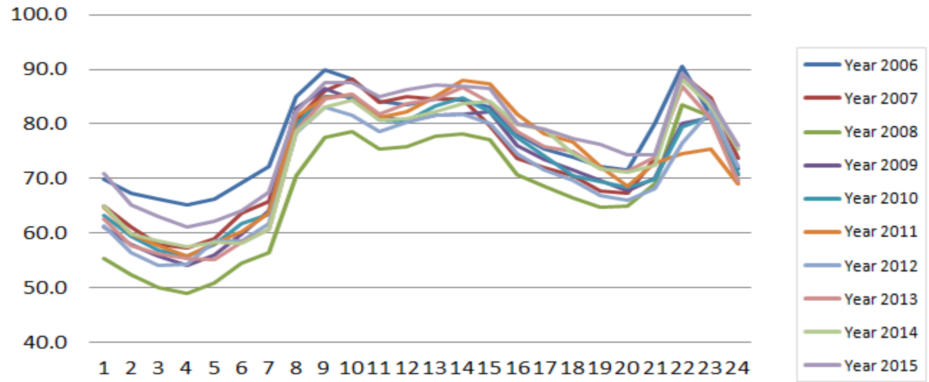

Figure 7.2.1. Load curves for the period 2006-2015 (powers in MW)

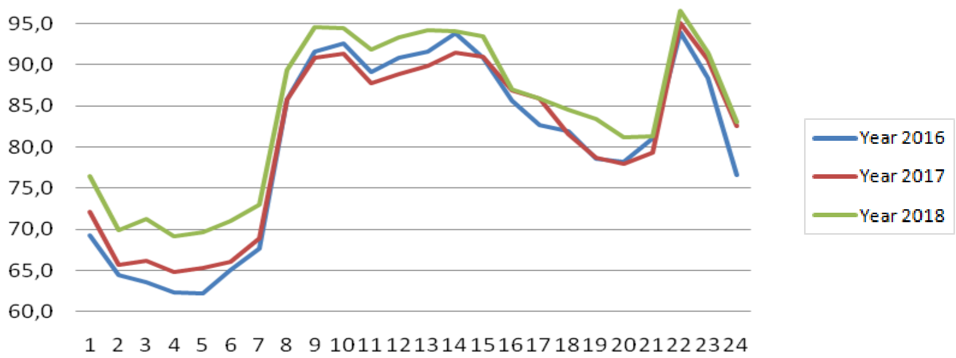

Figure 7.2.2. Load curves for the period 2016-2018 (powers in MW)

The data analysis highlights the following conclusions:

- for the 2006-2015 period the evolution is ambiguous, in other words there is no explicit tendency to increase or decrease the power as a whole;
- the load curves "intersect", which means that their shape differs (especially in certain time zones), signaling "horizontal" correlation problems;
- for the 2016-2018 period, there is a general increasing trend of the average hourly power as a whole (with small "syncopes" for 2017, time zones 9-13 and 18-21), which, correlated with the first observation, can lead to problems regarding the quality of the forecasts that will be obtained;
- the degree of correlation of the load curves is relatively low (evolution over time and shape during a day), which strengthens the prediction from the previous point.

The results obtained - the predicted values, the differences from the real values (in %) and the relative square deviation - are presented in tables 7.2.3 (Bayesian ANN, load curve) and 7.2.4 (Bayesian ANN, hourly). In the last line of the table is given the value of the performance index (s_{total}), defined as the sum of the partial indices (s_{2016}, s_{2017}, s_{2018}) – the sum of the squares of the deviations of the 24-hourly values for the corresponding year. Graphically, these results are presented comparatively (for the two methods used) in figures 7.2.3 (year 2016), 7.2.4 (year 2017) and 7.2.5 (year 2018).

Table 7.2.3. Forecasted load curves (powers in MW) for the period 2016-2018 (Bayesian ANN, load curve)

2016					2017					2018				
Hour	**Kwon value**	**Predicted value**	**Relative deviation [%]**	**Relative square deviation**	**Hour**	**Kwon value**	**Predicted value**	**Relative deviation [%]**	**Relative square deviation**	**Hour**	**Kwon value**	**Predicted value**	**Relative deviation [%]**	**Relative square deviation**
1	69.3	71.1	2.61	6.83	1	72.1	72.3	0.24	0.06	1	76.5	73.4	-3.99	15.91
2	64.4	64.0	-0.68	0.46	2	65.7	65.5	-0.25	0.06	2	69.9	67.1	-3.94	15.56
3	63.5	63.7	0.34	0.11	3	66.2	64.6	-2.41	5.82	3	71.3	67.9	-4.77	22.74
4	62.3	61.6	-1.18	1.40	4	64.8	62.9	-2.96	8.77	4	69.1	65.6	-5.07	25.66
5	62.2	61.0	-1.99	3.94	5	65.3	62.5	-4.26	18.15	5	69.6	65.3	-6.18	38.17
6	65.0	63.6	-2.11	4.45	6	66.1	65.1	-1.49	2.23	6	71.0	66.6	-6.17	38.02
7	67.7	71.2	5.18	26.83	7	68.9	72.1	4.60	21.19	7	73.0	72.7	-0.41	0.17
8	85.7	87.5	2.05	4.22	8	85.7	88.7	3.51	12.32	8	89.3	90.0	0.74	0.55
9	91.6	93.6	2.21	4.89	9	90.8	94.4	4.01	16.11	9	94.5	95.3	0.81	0.66
10	92.6	92.6	0.00	0.00	10	91.3	93.8	2.71	7.35	10	94.4	95.0	0.59	0.35
11	89.1	90.1	1.12	1.26	11	87.8	89.8	2.28	5.19	11	91.8	94.1	2.51	6.28
12	90.8	93.3	2.75	7.58	12	88.8	92.5	4.17	17.36	12	93.3	96.9	3.86	14.89
13	91.6	93.6	2.20	4.86	13	89.9	94.1	4.67	21.83	13	94.2	96.7	2.63	6.91
14	93.8	91.2	-2.75	7.59	14	91.5	92.4	0.95	0.90	14	94.1	93.5	-0.61	0.37
15	90.9	93.0	2.26	5.09	15	91.0	94.4	3.74	13.97	15	93.4	95.9	2.64	6.96
16	85.6	87.4	2.10	4.42	16	86.9	90.3	3.91	15.31	16	87.0	92.4	6.21	38.53
17	82.7	82.5	-0.23	0.05	17	85.9	83.6	-2.71	7.33	17	85.9	84.6	-1.46	2.13
18	81.9	83.7	2.21	4.89	18	81.5	84.7	3.93	15.43	18	84.5	85.7	1.41	2.00
19	78.6	75.7	-3.69	13.64	19	78.7	77.5	-1.50	2.26	19	83.4	79.4	-4.83	23.36
20	78.2	79.7	1.92	3.68	20	78.0	80.8	3.59	12.89	20	81.2	82.9	2.09	4.38
21	81.0	78.6	-2.98	8.88	21	79.3	79.8	0.66	0.44	21	81.3	81.1	-0.28	0.08
22	93.9	93.1	-0.83	0.69	22	95.1	94.2	-0.95	0.91	22	96.5	95.3	-1.28	1.64
23	88.4	90.7	2.60	6.77	23	90.6	92.6	2.21	4.87	23	91.4	93.9	2.74	7.48
24	76.6	79.0	3.11	9.65	24	82.5	80.8	-2.11	4.47	24	83.0	82.6	-0.53	0.28
s_{2016}	132.19				s_{2017}	215.21				s_{2018}	273.05			
s_{total} = 620.45														

Table 7.2.4. Forecasted load curves (powers in MW) for the period 2016-2018 (Bayesian ANN, hourly)

2016					2017					2018				
Hour	**Kwon value**	**Predicted value**	**Relative deviation [%]**	**Relative square deviation**	**Hour**	**Kwon value**	**Predicted value**	**Relative deviation [%]**	**Relative square deviation**	**Hour**	**Kwon value**	**Predicted value**	**Relative deviation [%]**	**Relative square deviation**
1	69.3	71.1	2.54	6.44	1	72.1	73.1	1.45	2.09	1	76.5	75.3	-1.60	2.55
2	64.4	65.9	2.33	5.45	2	65.7	66.8	1.61	2.60	2	69.9	67.6	-3.26	10.63
3	63.5	64.3	1.30	1.68	3	66.2	66.7	0.70	0.49	3	71.3	69.1	-3.12	9.74
4	62.3	63.3	1.55	2.41	4	64.8	65.6	1.22	1.50	4	69.1	68.0	-1.60	2.57
5	62.2	63.7	2.40	5.78	5	65.3	65.7	0.64	0.42	5	69.6	67.8	-2.58	6.66
6	65.0	65.4	0.67	0.45	6	66.1	67.3	1.87	3.48	6	71.0	69.3	-2.43	5.89

2016					2017					2018				
Hour	Kwon value	Predicted value	Relative deviation [%]	Relative square deviation	Hour	Kwon value	Predicted value	Relative deviation [%]	Relative square deviation	Hour	Kwon value	Predicted value	Relative deviation [%]	Relative square deviation
7	67.7	68.2	0.71	0.51	7	68.9	70.1	1.71	2.94	7	73.0	72.0	-1.34	1.78
8	85.7	85.2	-0.64	0.41	8	85.7	86.9	1.35	1.81	8	89.3	88.6	-0.81	0.66
9	91.6	90.6	-1.10	1.22	9	90.8	92.1	1.45	2.10	9	94.5	93.7	-0.89	0.79
10	92.6	91.7	-0.98	0.96	10	91.3	92.7	1.55	2.40	10	94.4	93.7	-0.69	0.48
11	89.1	88.3	-0.88	0.77	11	87.8	89.7	2.16	4.64	11	91.8	91.1	-0.78	0.60
12	90.8	89.4	-1.56	2.44	12	88.8	90.7	2.15	4.63	12	93.3	92.1	-1.33	1.77
13	91.6	90.3	-1.43	2.06	13	89.9	91.7	1.98	3.92	13	94.2	93.1	-1.18	1.39
14	93.8	92.8	-1.09	1.18	14	91.5	92.9	1.53	2.33	14	94.1	93.0	-1.16	1.34
15	90.9	90.6	-0.35	0.12	15	91.0	91.8	0.90	0.81	15	93.4	93.1	-0.36	0.13
16	85.6	85.7	0.16	0.03	16	86.9	86.5	-0.48	0.23	16	87.0	87.2	0.27	0.07
17	82.7	83.4	0.88	0.77	17	85.9	84.9	-1.22	1.48	17	85.9	86.3	0.47	0.22
18	81.9	81.4	-0.64	0.41	18	81.5	82.6	1.38	1.91	18	84.5	83.9	-0.72	0.52
19	78.6	78.0	-0.76	0.57	19	78.7	80.4	2.10	4.42	19	83.4	82.8	-0.77	0.59
20	78.2	77.8	-0.45	0.20	20	78.0	79.3	1.61	2.59	20	81.2	80.7	-0.64	0.41
21	81.0	80.4	-0.71	0.51	21	79.3	80.6	1.66	2.74	21	81.3	80.8	-0.61	0.37
22	93.9	94.2	0.32	0.10	22	95.1	95.3	0.21	0.04	22	96.5	96.3	-0.17	0.03
23	88.4	88.2	-0.23	0.05	23	90.6	89.9	-0.78	0.60	23	91.4	91.6	0.22	0.05
24	76.6	78.3	2.28	5.19	24	82.5	80.9	-1.97	3.88	24	83.0	83.5	0.57	0.32
s_{2016}	39.72				s_{2017}	54.04				s_{2018}	49.57			
s_{total} = 143.33														

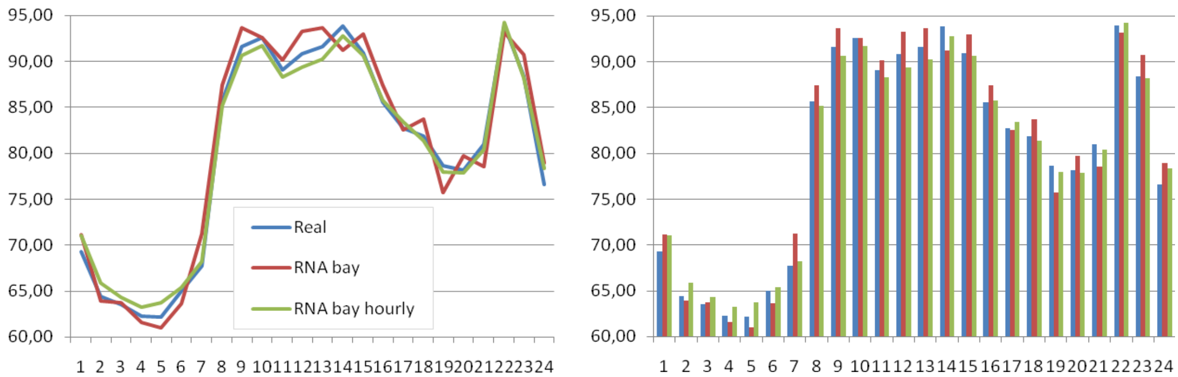

Figure 7.2.3. Comparative analysis of results for year 2016

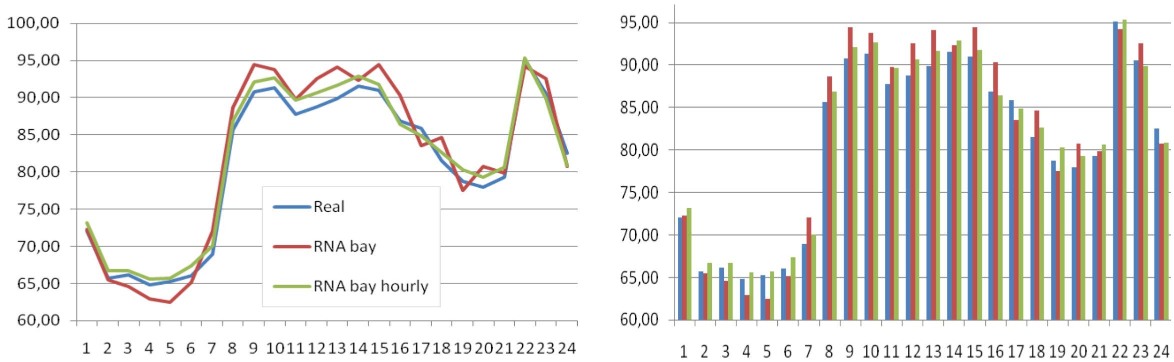

Figure 7.2.4. Comparative analysis of results for year 2017

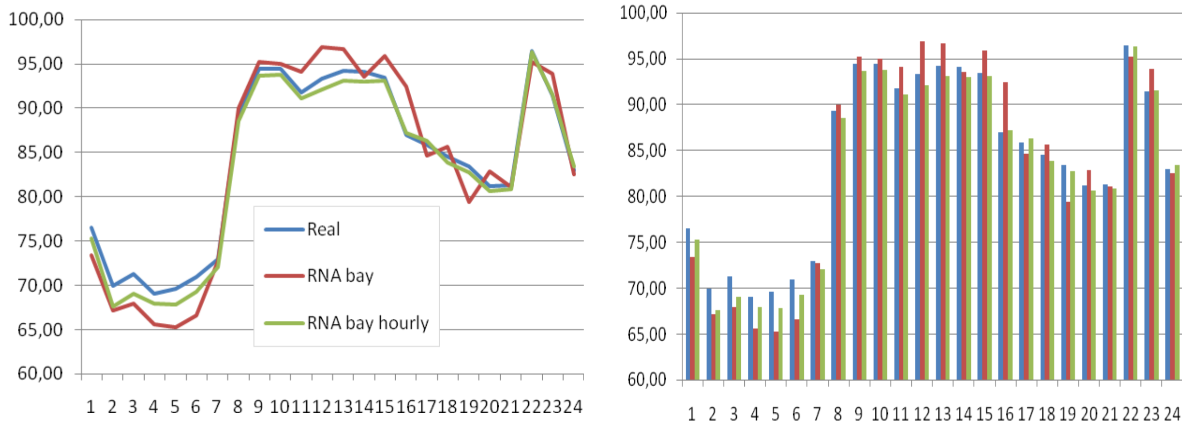

Figure 7.2.5. Comparative analysis of results for year 2018

Table 7.2.5 presents the summarized performance indices for the two forecasting methods.

Table 7.2.5. Comparative value of performance indices

Method / Performance indices	s_{2016}	s_{2017}	s_{2018}	s_{total}
Bayesian ANN, load curve	132.19	215.21	273.05	620.45
Bayesian ANN, hourly	39.72	54.04	49.57	143.33

The comparative analysis of the results obtained with the two forecasting methods leads to the following conclusions:

- the results confirm the observations made in the analysis of the load curves for the 2006-2015 period, and 2016-2018 respectively;
- the hourly forecast offers much better results than those obtained with the forecast of the overall load curve (global performance index 143 compared to 621), a situation that can be explained by the poor correlation of the shape of the load curves;
- however, comparatively, it can be stated that the hourly forecast manages to "catch" better the shape of the load curves for the 2016-2018 period;
- the analysis of the value of the annual performance indices (s_{2016}, s_{2017}, s_{2018}) highlights values of practically the same order of magnitude, however increasing, which means a slight alteration of the quality of the forecasts as we move away from the known area.
- in accordance with the previous conclusions, for the red curves in fig. 7.2.3-7.2.5 (load curve overall forecast) there are areas where the difference is relatively larger than the blue and green ones (real consumption, respectively the one forecasted with ANN hourly), due to the weak correlation of the shape of the load curves.

The third subchapter is dedicated to 110 kV / m.t. substations from TUN Timisoara. For the presentation of the results in detail, 5 significant 110/20 kV substations were selected (3 from Timisoara, one from an important city of Timis County and one that supplies an oil exploitation): Bucovina, IMT, Musicescu, Deta and Satchinez. At the end of the subchapter, some comments and conclusions are highlighted, with a more general or particular character, aiming at both the concrete results of the forecasts and the ANN used. Special attention is given to the comparison of the results, highlighting the quality of the methods used, assessing the influence of the degree of correlation of the load curves on the quality of the predictions made.

For example, the results of the Bucovina substation are presented, at 9 o'clock on the first Wednesday of each month. Figure 7.3.1 shows the data referring the load curves for a period of 8 years (2009-2016), used for ANNs training. Figure 7.3.2 shows the data on the load curves for the years 2017 and 2018, used to verify the results of the forecasts for the respective years.

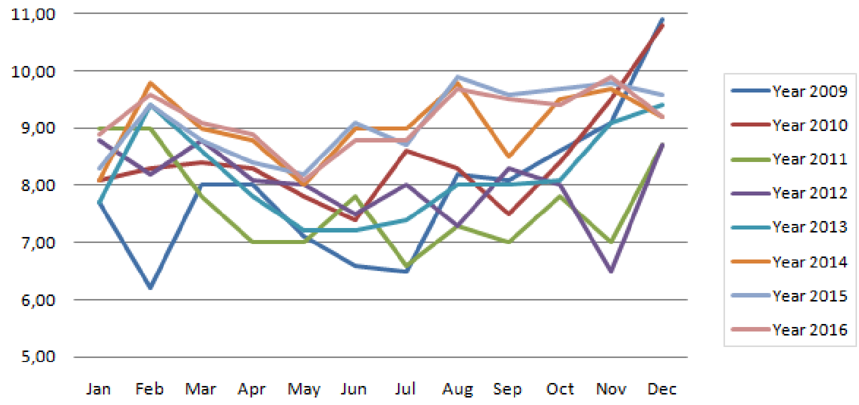

Figure 7.3.1. Load curves for the period 2009-2016 (powers in MW)

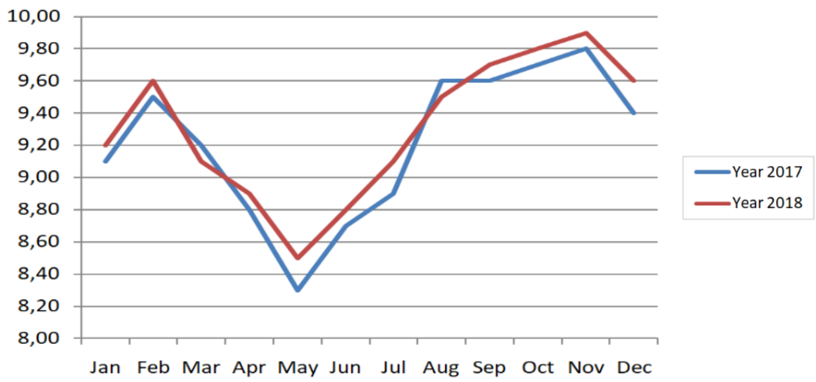

Figure 7.3.2. Load curves for the period 2017-2018 (powers in MW)

The analysis of the results highlights the following conclusions:

- the 2009-2016 period shows a general downward trend in consumption in the first 4 years, followed by an increase in the next 4 years;
- for 2017-2018 the general trend is upward (except March and August);
- the shape of the curves is quite different, with many "intersections";
- consequently, the degree of correlation of the load curves is relatively low, both in terms of evolution over time ("vertical") and shape over a year ("horizontal");
- if we consider all the analyses in this subchapter, the present case is, from this point of view, in the middle area.

The results are presented in graphic form (as a comparison for the two methods used) in figure 7.3.3 (year 2017) and figure 7.3.4 (year 2018).

Table 7.3.5 summarized the performance indices for the two forecasting methods.

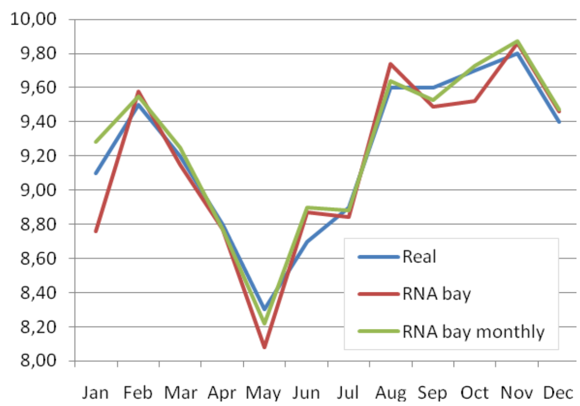

Figure 7.3.3. Comparative analysis of results for year 2017

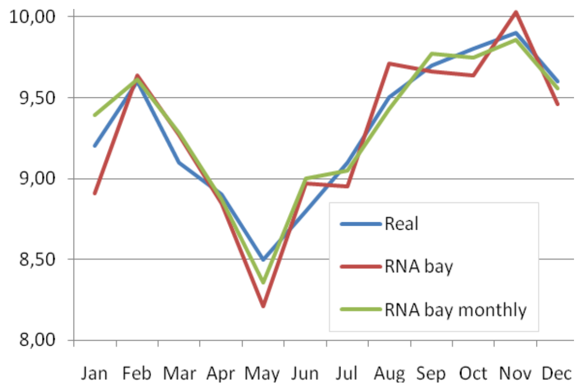

Figure 7.3.4. Comparative analysis of results for year 2018

Table 7.3.5. Comparative value of performance indices

Method / Performance indices	s_{2017}	s_{2018}	s_{total}
Bayesian ANN, load curve	34,04	43,58	77,62
Bayesian ANN, monthly	12,73	18,14	30,87

The comparative analysis of the obtained results highlights the following conclusions:

- the results confirm the observations from the analysis of the load curves;
- the monthly forecast leads to significantly better results than that of the overall load curve (overall performance index 30, compared to 77), which can be explained by the relatively poor correlation of the shape of the load curves;
- however, in comparison, it can be stated that the monthly forecast (green curve) manages to better "catch" the shape of the real load curves (blue curve);
- the annual performance indices (s_{2017}, s_{2018}) have values of the same order of magnitude, slightly higher than those for 2018 (compared to 2017).

Chapter 8 includes the general conclusions of the thesis and the presentation of the original contributions, as well as highlighting the directions and perspectives offered by the PhD thesis for the continuation and extension of research and application of results and experience. The elaborated methodologies and calculation programs are of general applicability, constituting an efficient working tool for the distribution and transmission operators, for the economic agents with preoccupations in the field of electricity consumption and production.

The Electronic Appendices comprise a series of elements and detailed results related to the case studies presented in the PhD thesis.

The obtained results were and will be capitalized through scientific research and technical assistance contracts carried out by the Research Center for Analysis and Optimization of Electric Power System within Politehnica University of Timișoara, the beneficiaries being Enel Distribution Banat and Dobrogea, Electrica Muntenia Nord , Delgaz Grid Iasi (important electricity distribution operators in Romania) and economic entities with concerns in the field of implementation of renewable energy resources [UPT 2017], [UPT2018], [UPT2019], [UPT 2020a], [UPT 2020b].

The activity of preliminary training of the PhD student and the results obtained during the elaboration of the paper were capitalized by 5 ISI indexed papers (2 in journals, 3 in conference volumes) [Bucerzan2010], [Crăciun2013], [Bărbulescu2018], [Crăciun2018a], [Csorba2018], 2 papers indexed in other international databases (BDI) (1 in the journal, 1 in a conference volume, being indexed by ISI) [Crăciun2017], [Bărbulescu2021] and 2 scientific reports.

The theoretical analyses performed in this PhD thesis, as well as the practical results obtained, open a series of clear perspectives for further and in-depth research both in the field of forecasting electrical energy consumption and load curves and in terms of use BN to solve other problems in the field of electric power systems engineering:

- refining solution methods based on the use of BNNs, in order to increase efficiency and improve their performance;
- correlating the history of consumption evolution with a series of other factors (climatic, economic, the degree of implementation of renewable sources), as far as it is possible;
- the use of BNNs in forecasting studies related to renewable sources (power or energy generated, wind speed for wind farms, irradiance for photovoltaics, etc.);
- elaboration of methods to eliminate or correct some obviously erroneous data in the consumption history;
- the use of BNs in studies related to the reliability of electricity transmission and distribution networks, of the PS as a whole, PS state estimation, PS stability, diagnosing and locating faults in electrical networks, transformers, generators, etc.

Bibliography (selective list)

1. [Ahmadi2020] S. Ahmadi, A.H. Fakehi, A. Vakili, M. Haddadi, S.H. Iranmanesh, A hybrid stochastic model based Bayesian approach for long term energy demand managements, Energy Strategy Reviews, vol.28, 2020, pp.1-13
2. [Augutis2012] J. Augutis, I. Zutautaite, V. Radziukynas, R. Krikstolaitis, Application of Bayesian method for electrical power system transient stability assessment, International Journal of Electrical Power & Energy Systems, vol.42, nr.1, 2012, pp.465-472
3. [Bakirtzis1997] A. Bakirtzis, S. Kiartzis, V. Petridis, A. Kehagias, A Bayesian Combination Method for Short Term Load Forecasting, International Journal of Electrical Power & Energy Systems, vol.19, nr.3, 1997, pp.171-177
4. [Bărbulescu2016] C. Bărbulescu, Şt. Kilyeni, A. Deacu, Artificial Neural Network based Monthly Load Curves Forecasting, 11th IEEE International Symposium on Applied Computational Intelligence and Informatics (SACI), 2016, pp.237-242
5. [Bărbulescu2018] C. Bărbulescu, Şt. Kilyeni, A. Deacu, A. Simo, M. Crăciun, Power Consumption Forecast Quality Assessment, 7th International Conference on Computers Communications and Control (ICCCC), 2018, p.103-110
6. [Bărbulescu2021] C. Bărbulescu, Şt. Kilyeni, V. Chiş, M. Crăciun, A. Simo, Daily Load Curve Forecasting. Comparative Analysis: Conventional vs. Unconventional Methods. In: Balas V., Jain L., Balas M., Shahbazova S. (eds.) Soft Computing Applications, SOFA 2018, Advances in Intelligent Systems and Computing, vol. 1221, 2021, Springer, pp.3-18
7. [Bessani2020] M. Bessani, J.A.D. Massignan, T.M.O. Santos, J.B.A. London Jr., C.D. Maciel, Multiple households very short-term load forecasting using bayesian networks, Electric Power Systems Research, vol.189, pp.1-7
8. [Bolstad2004] W.M. Bolstad, Introduction to Bayesian Statistics, John Wiley&Sons, 2004
9. [Borges2016] C.L.T. Borges, J.A.S. Dias, A Model to Represent Correlated Time Series in Reliability Evaluation by Non-SequentialMonte Carlo Simulation, IEEE Transactions on Power Systems, vol.32, nr.2, 2017, pp.1511-1519
10. [Bucerzan2010] D. Bucerzan, M. Crăciun, V. Chis, C. Raţiu, Stream Ciphers Analysis Methods, International Journal of Computers Communications & Control, vol.5, nr.4, 2010, pp.483-489
11. [Chevalier2019] S. Chevalier, P. Vorobev, K. Turitsyn, A Bayesian Approach to Forced Oscillation Source Location Given Uncertain Generator Parameters, IEEE Transactions on Power Systems, vol.34, nr.2, 2019, pp.1641-1649
12. [Chiş2015] V. Chiş, Tehnici de inteligenţă artificială utilizate în studiile de prognoză din domeniul ingineriei energetice, PhD thesis, Politehnica UniversityTimisoara (in Romanian)
13. [Crăciun2013] M. Crăciun, D. Bucerzan, C. Raţiu, A. Manolescu, Actuality of Bankruptcy Prediction Models used in Decision Support System, International Journal of Computers Communications & Control, vol.8, nr.3, 2013, pp.375-383
14. [Crăciun2017] M. Crăciun, Şt. Kilyeni, C. Bărbulescu, Bayesian network applications in power systems engineering. A review, Journal of Sustainable Energy, nr.3, 2017, p. 99-105
15. [Crăciun2018a] M. Crăciun, L.M. Csorba, Application of the Fuzzy-Pay-Off Method in the Valuation of a Financial Instrument. In: Balas V., Jain L., Balas M. (eds.), Soft Computing Applications, SOFA 2016, Advances in Intelligent Systems and Computing, Springer, vol. 634, 2018, pp.235-250
16. [Csorba2018] L.M. Csorba, M. Crăciun, An Application of the Multi Period Decision Trees in the Sustainable Medical Waste Investments. In: Balas V., Jain L., Balas M. (eds.), Soft Computing Applications, SOFA 2016, Advances in Intelligent Systems and Computing, Springer, vol.634, 2018, pp.540-556

17. [Dagdougui2019] H. Dagdougui, F. Bagheri, H. Le, L. Dessaint, Neural network model for short-term and very-short-term load forecasting in district buildings, Energy and buildings, vol.203, pp.1-10
18. [Deacu2015] A. Deacu, Prognoza consumului de energie electrică utilizând reţele neuronale artificiale, PhD thesis, Politehnica UniversityTimisoara (in Romanian)
19. [Ghayekhloo2015] M. Ghayekhloo, M.B. Menhaj, M. Ghofrani, A hybrid short-term load forecasting with a new data preprocessing framework, Electric Power Systems Research, vol.119, 2015, pp.138-148
20. [Gilanifar2020] M. Gilanifar, H. Wang, L.M.K. Sriram, E.E. Ozguven, R. Arghandeh, Multitask Bayesian Spatiotemporal Gaussian Processes for Short-Term Load Forecasting, IEEE Transactions on Industrial Electronics, vol.67, nr.6, 2020, pp.5132-5143
21. [HeY2018] Y. He, B. Lin, Forecasting China's total energy demand and its structure using ADL-MIDAS model, Energy, vol.151, 2018, pp.420-429
22. [He2019] F. He, J. Zhou, Z. Feng, G. Liu, Y. Yang, A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with Bayesian optimization algorithm, Applied Energy, vol. 237, 2019, pp.103-116
23. [Hsu2018] Y.Y. Hsu, T. Tung, H. Yeh, C. Lu, Two-Stage Artificial Neural Network Model for Short-Term Load Forecasting, IFAC PapersOnLine, vol.51, nr.28, 2018, pp.678-683
24. [Jarndal2018] A. Jarndal, S. Hussein, Forecasting of Electric Peak Load Using ANN Cascaded, ANN-NARX and GPR Techniques, International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI), 2020, pp.1-5
25. [Kilyeni2014] S. Kilyeni, Metode numerice. Algoritme, programe de calcul, aplicaţii în energetică, Orizonturi Universitare, 2014 (in Romanian)
26. [Kilyeni2015] S. Kilyeni, Tehnici de optimizare în ingineria energetică, Orizonturi Universitare, 2015 (in Romanian)
27. [Kilyeni2015a] S. Kilyeni, Tehnici numerice de analiză asistată de calculator a regimurilor de funcţionare a sistemelor electroenergetice, Orizonturi Universitare, 2015 (in Romanian)
28. [Koch2007] K.R. Koch, Introduction to Bayesian Statistics, Springer, 2007
29. [Kumar2016] S. Kumar, S. Mishra, S. Gupta, Short Term Load Forecasting Using ANN and Multiple Linear Regression, 2nd International Conference on Computational Intelligence & Communication Technology (CICT), 2016, pp.184-186
30. [Li2014] G. Li, H. Wu, F. Wang, Bayesian network approach based on fault isolation for power system fault diagnosis, International Conference on Power System Technology (Powercon), 2014, pp.601-606
31. [Lin2019] Y. Lin, M. Yang, C. Wan, J. Wang, Y. Song, A Multi-Model Combination Approach for Probabilistic Wind Power Forecasting, IEEE Transactions on Sustainable Energy, vol.10, nr.1, 2019, pp.2226-2237
32. [Lorencin2017] I. Lorencin, M. Pantos, Evaluating Generating Unit Unavailability Using Bayesian Power Priors, IEEE Transactions on Power Systems, vol.32, nr.3, 2017, pp. 2315-2323
33. [Ma2013] H. Ma, H. Li, Analysis of Frequency Dynamics in Power Grid: A Bayesian Structure Learning Approach, 2013, IEEE Transactions on Smart Grid, vol.4, nr.1, 2013, pp.457-466
34. [MacKay2003] D.J.C. MacKay, Information Theory, Inference, and Learning Algorithms, Cambridge University Press 2003
35. [Massignan2019] J.A.D. Massignan, J.B.A. London, C.D. Maciel, M. Bessani, V. Miranda, PMUs and SCADA Measurements in Power System State Estimation through Bayesian Inference, IEEE PowerTech, 2019, pp.1-6

36. [Mestav2019] K.R. Mestav, J.L. Rozas, L. Tong, Bayesian State Estimation for Unobservable Distribution Systems via Deep Learning, IEEE Transactions on Power Systems, vol.34, nr.6, 2019, pp.4910-4920
37. [Moller1993] M.F. Moller, A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning, Neural Networks, vol.6, 1993, pp.525-533
38. [Nabney2002] I.T. Nabley, Algorithms for pattern recognition, Springer, 2002
39. [Ning2010] Y. Ning, Y. Liu, Q. Ji, Bayesian - BP Neural Network Based Short-term Load Forecasting for Power System, 3rd IEEE International Conference on Advanced Computer Theory and Engineeringv, 2010, vol.2, pp.89-93
40. [Panamtash2020] H. Panamtash, Q. Zhou, T. Hong, Z. Qu, K.O. Davis, A copula-based Bayesian method for probabilistic solar power forecasting, Solar Energy, vol.196, 2020, pp.336-345
41. [Pegoraro2017], P.A. Pegoraro, A.Angioni, M. Pau, A. Monti, C. Muscas, F. Ponci, Bayesian Approach for Distribution System State Estimation with Non-Gaussian Uncertainty Models, IEEE Transactions on Instrumentation and Measurement, vol.66, no. 11, 2017, pp.2957-2966
42. [Rivero2015] C.R. Rivero, V. Sauchelli, H.D. Patino, J.A. Pucheta, S. Laboret, Long-term Power Consumption Demand Prediction: a comparison of Energy associated and Bayesian modeling approach, Latin America Congress on Computational Intelligence (LA-CCI), 2015, pp.1-6
43. [Russell2010] S.J. Russell, P. Norvig, Artificial Intelligence. A Modern Approach, Pearson, 2010
44. [Sahu2019] M.K. Sahu, B. Sahoo, M. Khatoi, S. Behera, Short-term Wind and PV Generation Forecasting of time-series using ANN, International Conference on Intelligent Computing and Control Systems (ICICCS), 2019, pp.1328-1333
45. [Sarajcev2020] P. Sarajcev, D. Jakus, J. Vasilj, Ensemble learning with time-series clustering for aggregated short-term load forecasting, 20th IEEE Mediterranean Electrotechnical Conference (MELECON), 2020, pp.553-558
46. [Seppanen2016] J. Seppanen, S.K. Au, J. Turunen, L. Haarla, Bayesian Approach in the Modal Analysis of Electromechanical Oscillations, IEEE Transactions on Power Systems, vol.32, nr.1, 2017, pp.316-325
47. [SilvaT2017] T.V. da Silva, R.V. Monteiro, G. Guimaraes, F.A. Moura, R.M.C. Albertini, M.A. Tamashiro, Performance Analysis of Neural Network Training Algorithms and Support Vector Machine for Power Generation Forecast of Photo-voltaic Panel, IEEE Latin America Transactions, vol.15, nr.6, 2017, pp.1091-1100
48. [Silva2019] FL.C. Silva, F.L.C. Oliveira, R.C. Souza, A bottom-up bayesian extension for long term electricity consumption forecasting, Energy, vol.167, 2019, pp.198-210
49. [Singh2017] S. Singh, S. Hussain, M.A. Bazaz, Short term load forecasting using artificial neural network, 4th International Conference on Image Information Processing (ICIIP), 2017, pp.1-5
50. [Sun2019] M. Sun, C. Feng, J. Zhang, Conditional aggregated probabilistic wind power forecasting based on spatio-temporal correlation, Applied Energy, vol.256, 2019, pp.1-11
51. [Sykora2016] M. Sykora, J. Markova, D. Diamantidis, Bayesian Network Application for the Risk Assessment of Existing Energy Production Units,2nd IEEE International Symposium on Stochastic Models in Reliability Engineering, Life Science and Operations Management, 2016, pp.656-664
52. [Tang2019] L. Tang, X. Wang, X. Wang, C. Shao, S. Liu, S. Tian, Long-term electricity consumption forecasting based on expert prediction and fuzzy Bayesian theory, Energy, vol.167, 2019, pp.1144-1154
53. [Tao2016] L. Tao, J. He, Y. Wang, P. Zhang, H. Zhang, H. Wang, Y. Miao, Operational risk assessment of distribution network with consideration of PV output uncertainties, China International Conference on Electricity Distribution (CICED), 2016, pp.1-6

54. [UPT2017] Contract UPT 21/2017, Elaborare modele de calcul şi aplicarea lor pentru prognoza consumului propriu tehnologic în reţelele electrice de distribuţie, Electrica Muntenia Nord (in Romanian)
55. [UPT2018] Contract UPT 59/2018, Analiza şi optimizarea regimurilor de funcţionare pentru reţeaua electrică de distribuţie, Electrica Muntenia Nord (in Romanian)
56. [UPT2019] Contract UPT 63/2019, Analiza regimurilor de funcţionare în vederea determinării post-calcul a consumului propriu tehnologic, Electrica Muntenia Nord (in Romanian)
57. [UPT2020a] Contract UPT 14/2020, Prognoza consumului propriu tehnologic în reţelele electrice de distribuţie, DelGaz Grid Moldova (in Romanian)
58. [UPT2020b] Contract UPT 75/2020, Evaluarea consumului propriu tehnologic prin metode de tip post-calcul în reţele de distribuţie, Electrica Muntenia Nord (in Romanian)
59. [Vakili2015] S. Vakili, Q. Zhao, L. Tong, Bayesian Quickest Short-term Voltage Instability Detection in Power Systems, 54th IEEE Annual Conference on Decision and Control (CDC), 2015, pp.7214-7219
60. [WangT2015] T. Wang, Y. Zhu, Z. Gao, Fault Diagnosis for Power System Based on a Special Bayesian Network, IEEE Region 10 TTENCON Conference, 2015, pp.58-63
61. [WangY2019] Y. Wang, Q. Hu, D. Srinivasan, Z. Wang, Wind Power Curve Modeling and Wind Power Forecasting with Inconsistent Data, IEEE Transactions on Sustainable Energy, vol.10, nr.1, 2019, pp.16-25
62. [XuB2019] B. Xu, H. Li, W. Pang. D.Chen, Y. Tian, X. Lei, X Gao, C. Wu, Bayesian network approach to fault diagnosis of a hydroelectric generation system, Energy Science & Engineering, vol.7, nr.5, 2019, pp.1669-1677
63. [XuT2010] T. Xu, Y. Zhou, Y. Zhang, Study and Implementation of Rural Distribution Network Fault Location Method Based on Bayesian Reference, 2nd International Conference on Information Science and Engineering, 2010, pp.1-6
64. [Yang2013] M. Yang, S. Fan, J. Lee, Probabilistic Short-Term Wind Power Forecast Using Componential Sparse Bayesian Learning, IEEE Transactions on Industry Applications, vol.49, nr.6, 2013, pp.2783-2792
65. [Yang2020] Y. Yang, W. Li, T.A. Gulliver, S. Li, Bayesian Deep Learning-Based Probabilistic Load Forecasting in Smart Grids, IEEE Transactions on industrial informatics, vol.16, nr.7, 2020, pp.4703-4713
66. [Yuan2017] X.C. Yuan, X. Sun, W. Zhao, Z. Mi, B. Wang, Forecasting China's regional energy demand by 2030: A Bayesian approach, Resources, Conservation & Recycling, vol.127, 2017, pp.85-95
67. [Zhang2014] Y. Zhang, Y. Xiang, L. Wang, Reliability Analysis of Power Grids with Cyber Vulnerability in SCADA System, IEEE PES General Meeting, 2014, pp.1-6
68. [Zhao2010] W.Q. Zhao, S.L. Zhang, D.X. Niu, Multi-Agent and Bayesian Network applied in Transformer Faults Diagnosis, 9th International Conference on Machine Learning and Cybernetics, 2010, pp.43-46
69. [Zheng2010] G. Zheng, Z. Yongli, Research of Transformer Fault Diagnosis Based on Bayesian Network Classifiers, International Conference on Computer Design and Appliations (ICCDA), 2010, vol.3, pp.382-385
70. [Zhou2012] D. Zhou, C. Li, Z. Wang, Power Transformer Lifetime Modeling, Prognostics & System Health Management IEEE Conference (PHM), 2012, pp.1-6