УДК 518.517

Power System Parameters Forecasting Using Hilbert-Huang Transform and Machine Learning [*]

V. G. Kurbatsky, V. A. Spiryaev, N. V. Tomin
Energy Systems Institute, Siberian Branch of Russian Academy of Sciences

P. Leahy
University College Cork

D. N. Sidorov, A. V. Zhukov
Energy Systems Institute, Siberian Branch of Russian Academy of Sciences
Institute of Mathematics, Economics and Informatics, Irkutsk State University

Abstract. A novel hybrid data-driven approach is developed for forecasting power system parameters with the goal of increasing the efficiency of short-term forecasting studies for non-stationary time-series. The proposed approach is based on mode decomposition and a feature analysis of initial retrospective data using the Hilbert-Huang transform and machine learning algorithms. The random forests and gradient boosting trees learning techniques were examined. The decision tree techniques were used to rank the importance of variables employed in the forecasting models. The Mean Decrease Gini index is employed as an impurity function. The resulting hybrid forecasting models employ the radial basis function neural network and support vector regression.

Apart from introduction and references the paper is organized as follows. The second section presents the background and the review of several approaches for short-term forecasting of power system parameters. In the third section a hybrid machine learning-based algorithm using Hilbert-Huang transform is developed for short-term forecasting of power system parameters. Fourth section describes the decision tree learning algorithms used for the issue of variables importance. Finally in section six the experimental results in the following electric power problems are presented: active power flow forecasting, electricity price forecasting and for the wind speed and direction forecasting.

Keywords: time series, forecasting, integral transforms, ANN, SVM, machine learning, boosting, singular integral, feature analysis.

1. Introduction

In recent years power industries worldwide have experienced two major changes: liberalization of the electricity market and the expansion of renewable energy. Liberalization has led to a tremendous increase in the inter-regional wholesale electricity trade between neighboring utilities or regions. The increasing penetration of renewable energy has led to increasing dependency on the volatile nature of renewable energy sources. Because of this, changes in a power system's parameters often feature a sharply variable non-stationary behavior, which limits the efficiency of modern technologies for forecasting time series [23; 29].

In this situation, effective short-term forecasting of state variables is highly important for market participants in wholesale power markets and for the power system as a whole. Over the past decade, many power engineers have focused their efforts on novel prediction problems, especially short-term forecasting of power flow, wind power forecasting and electricity price forecasting. In addition to difficulties with the sharply variable non-stationary behavior of power system parameters, researchers deal with a data requirenïS-ments problem in actual forecasting of such parameters. The required historical data is either not publicly available on power companies' websites or only made available with some delay. Consequently, the correct pattern of possible future variability of power system parameters is difficult to effectively identify, which in turn prevents accurate forecasting of system variables.

In this paper a novel hybrid approach for power system parameters forecasting is suggested, which combines the effective apparatus of analysis of non-stationary time series – the Hilbert-Huang integral transform (HHT) and machine learning technologies.

The paper is organized as follows. Section 2 presents the background and the review of several approaches for short-term forecasting of power system parameters. In section 3 a hybrid machine learning-based algorithm using HHT is developed for short-term forecasting of power system parameters. Section 4 presents the experimental results and discussion.

2. Background

In the electricity grid at any moment balance must be maintained between electricity consumption and generation - otherwise disturbances in power quality or supply may occur. Forecasting power systems parameters may be considered at different time scales, depending on the intended application. From milliseconds up to a few minutes, forecasts can be used, for instance, for turbine active control. This type of forecast is usually

referred to as very short-term forecasts. For the following 48–72 hours, forecasts are needed for power system management or energy trading.

The short-term forecasting of power system parameters can be carried out both with the aid of classical approaches of dynamic estimation, statistical methods of analysis of time series and regressive models, and with the aid of artificial intelligence. Many techniques have been employed for such purposes, including machine learning techniques – artificial neural networks (ANNs) [12; 10], support vector machines (SVMs) [8], random forest models [8; 22] and etc. Moreover, time series models, like ARIMA, GARCH models [13; 11], Kalman filter-based algorithm [9; 30] have also been proven to be effective in the power system parameters forecasting.

It is worth noting that power systems parameters often feature sharply variable non-stationary behaviour, which limits the efficiency of the stated technologies for forecasting of time series. Studies have shown that hybrid approaches, that is a combination of different intelligent techniques, have great potential and are worth pursuing [29]. Previously, several hybrid forecasting approaches have been proposed, including ARIMA-ANN models [1], Fuzzy-ANN-based models [14], Wavelet-ANN-based models [26], Fuzzy-expert system-based models [4] among others [6]. For review and bibliography of the state-ofthe-art medods for dynamical systems identification and forecasting readers may refer to the monograph [27].

It is well known that for effective forecasting of power system parameters a number of additional, correlated characteristics must be taken into account, which can be influenced by a predictable parameter. For instance, everything from salmon migration to forest fires can affect current and future electricity prices. But quite often, information about these characteristics is either unavailable or limited and contradictory. As a result, scientific researchers and power engineers have to work with only one retrospective predictable parameter time series without additional characteristics, which imposes limitations on forecasting performance. We consider that, in such a case, forecasting performance can be improved by using a transition from a time series task to a regression task, when an initial time series is decomposed into some components. An investigation of properties of obtained components can help to improve the performance of a power system parameter forecast.

3.   Proposed Approach

We propose a hybrid machine learning-based algorithm using HHT as a data preprocessing technology [16; 17]. Figure 1 the shows general forecasting scheme. There are five main blocks.

Data collection is performed by the SCADA system, which merely identifies a set of power system parameters. The second block preprocesses input data with HHT. HHT is a two-step algorithm; combining empirical mode

decomposition (EMD) and the Hilbert transform (HT). After the HHT was employed we obtain, through EMD and HT, the sets of intrinsic mode functions (IMFs), instantaneous amplitudes and frequencies. These sets are used as input values to the next block for feature selection and dimensionality reduction. In this block, learning methods are used – Random Forest (RF) and Boosting Trees (BT) to rank a variable's importance [7; 3]. SVM and ANN forecasting models with non-linear optimization algorithms are employed in the next block for optimal selection of parameters. The last block is used for testing and a comparison of obtained forecasting models. As a result, an HHT-ANN-SVM-based forecasting system will be able to predict power system parameters.
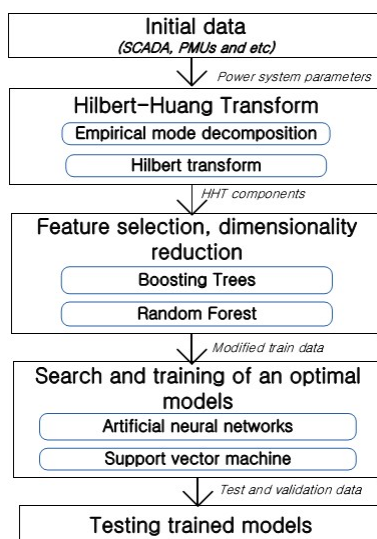

Fig. 1. Block diagram of the proposed hybrid approach scheme

The main goal of the hybrid forecasting approach development was to increase the efficiency of forecasting studies, when we deal with limited retrospective data and sharply variable, non-stationary time series of power system parameters.

4. Decision tree learning algorithms for the issue of variables importance

Forecasting of wind power, power flows, energy price etc. have become a major issue in power systems. Following the needs of today's competitive environment, various techniques are used to forecast wind power, energy price and power demand. Over the past decade, different machine learning

algorithms have been investigated for analyzing power systems parameters' patterns in order to obtain forecasts in power systems with high accuracy.

The importance of a variable is the contribution it makes to the success of the model. For a predictive model, success means good forecasting. Often the forecasting relies mainly on a few variables. A good measure of importance reveals those variables. The better the forecasting, the closer the model represents reality, and the more plausible it is that the important variables represent the true cause of forecasting.

We used the decision tree techniques to rank the importance of variables. These algorithms can output a list of predictor variables that they believe to be important in predicting the outcome. If nothing else, researchers can identify a subset of the data to only include the most "important" variables, and use that with another model.

As it is explained in [19], every binary decision tree is separately represented by a tree structure T, from an input vector $(X_1, ..., X_p)$ taking its values in $(\chi_1 * ... * \chi_p = \chi)$ to a output variable $y \in Y$. Every certain node t represents a subset of the space χ, with the root node being χ itself. Construction of decision trees usually works top-down, by choosing a variable at each step that splits the set of items by the binary test $s_t = (X_m < c)$. The internal node t divides its subset into two subsets[1] corresponding to two children nodes t_L and t_R. For a new instance the predicted value $\hat{Y}$ is the label of the leaf reached by the instance when it is propagated through the tree. Algorithms for constructing decision trees identify at each node t the split $s_t = s^*$ for which the partition of the N_t node samples into t_L and t_R maximizes the decrease

$$\Delta i(s, t) = i(t) - p_L i(t_L) - p_R i(t_R)$$

With regard to variable importance in Random forests, in [3] it was proposed to add up the weighted impurity (loss) decreases $p(t)i(s_t, t)$ for all nodes t where X_m is used, averaged over all N_T trees in the forest:

$$Imp(X_m) = \frac{1}{N_t} \sum_{T} \sum_{t \epsilon T : \nu(s_t) = X_m} p(t)i(s_t, t)$$

and where $p(t)$ is the proportion $\frac{N_t}{N}$ of samples reaching t and $\nu(s_t)$ is a variable which is used in split s_t.

In our paper we have examined two decision tree learning techniques: random forests and gradient boosting trees. We used the decision tree techniques to rank the importance of variables. Also we use Gini index, which is also called Mean Decrease Gini, as an impurity function. It was introduced for decision trees in [2] and later modified for gradient boosting machines [5].

---

[1] More generally, splits are defined by a (not necessarily binary) partition of the range χ_m of possible values of a single variable χ_m

5. Case Study

The novel hybrid approach is realized in STATISTICA 6.0. and Matlab. The efficiency of the developed approach is demonstrated with real time series in the following electric power problems: active power flow forecasting for the Western Baikal-Amur Mainline (Russia) [16; 15], electricity price forecasting for the Australian and Nord Pool Spot (EU) power markets [17], wind speed forecasting for the Valentia region (Ireland) [30].

5.1. Valentia wind speed forecasting (Ireland)

5.1.0.1. *HHT-ANN-SVM model* The initial training data is represented by two per-hour time-series: wind speed and wind direction values for 1 year for the Valentia region. Both time series were decomposed into IMFs by the Huang method, and the Hilbert transform was employed to obtain the amplitudes, A and frequency, F. The latter along with these HHT components were used as input values of the selected HHT-ANN and HHT-SVM models. The final 48 hours of the data were reserved for calculating performance statistics.

Fig. 2 illustrates the feature selection analysis of the modified Valentia wind speed data using regression machine learning algorithms – Boosting Trees and Random Forest. As seen, high-frequency components of the initial wind speed data (IMF1, IMF2, A_1, A_2 and F_1, F_2) have less predictive importance (less than 0.3) by comparison with other components, and therefore they can be excluded from the trained data.
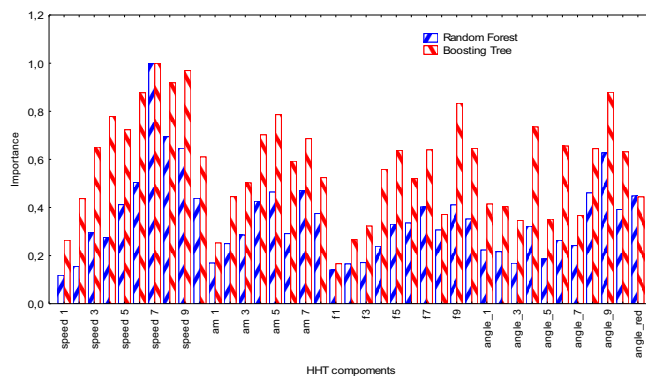

Fig. 2. Predictor importance diagram based on Random Forest and Boosting Trees learning algorithms for the Valentia wind speed data.

As a result, the following hybrid models are constructed, trained and tested: a RBF neural network (HHT-ANN model) and epsilon-SVM regression model (HHT-SVM model). The RBF network is trained by a 2-Phase-RBF scheme (two-step algorithm). After a training-testing process we obtai-

ned the following errors: train error – 0.251%, validation error – 0.263% and test error – 0.257%. The obtained RBF network had the following structure: input neurons 25, hidden neurons 162 and output neuron 1.

The polynomial kernel function is used in the epsilon-SVM regression model. Let us briefly outline the SVM regression model.

Let (x_i, y_i), $i = \overline{1, N}$, be the input data, namely $x_i \in \widetilde{X} \subset X, y_i \in \widetilde{Y} \subset Y$, X is the set of input data, Y is the set of values of the desired objective function. SVM resolves the regression problem by estimation of the unknown function

$$y_i = f(x_i) + \varepsilon_i, \tag{1}$$

where ε_i is error of i-th observation, e.g. in linear case $f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + b$, where $\mathbf{w} \in \mathbb{R}^n$ are weights and $b \in \mathbb{R}$ is the offset parameter. In nonlinear case SVM employs $\phi : \mathbb{R}^n \to H$ and instead of (1) the following function f is under the estimation $f(\mathbf{x}) = \langle \mathbf{w}, \phi(\mathbf{x}) \rangle + b$. Determination of parameters is reduced to a convex optimization problem, i.e. the the following regularized risk functional minimization

$$\min_{\mathbf{w}, b, \xi, \xi^*} \left[\frac{1}{2} ||\mathbf{w}||^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) \right], \tag{2}$$

$$\begin{cases} y_i - \langle \mathbf{w}, \phi(\mathbf{x_i}) \rangle - b \leq \varepsilon + \xi_i, \\ \langle \mathbf{w}, \phi(\mathbf{x_i}) \rangle + b - y_i \leq \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* \geq 0, \, i = \overline{1, N}. \end{cases} \tag{3}$$

Here $C > 0$ is the capacity constant to trade-off between the flatness of f and the amount up to which deviations larger than ε are tolerated, ξ_i are inseparable observations parameters, ε is losses parameter. The second term in the functional (2) penalizes any deviation $f(x_i)$ from y_i based on the constraints (3). This problem is resolved in terms of the dual problem to the problem (2) – (3) as follows $f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(x_i, x) + b$, where $0 \leq \alpha_i, \alpha_i^* \leq C$, and kernel $K(x_i, x)$ satisfies the conditions of Mercer-Hilbert-Schmidlt theorem (symmetry and non-negativity [21]) and can be represented as follows $K(x_i, x) = \langle \phi(x_i), \phi(x) \rangle$, which is the RBF is our case. Using a 10-fold cross-validation method the following optimal training constants and RBF kernel parameters were determined: the capacity constant $C = 1$; $\varepsilon = 1.3$; $d = 3$; $\gamma = 0.03$.

5.1.0.2. *FF-ANN-SA model* In this case example a feed forward ANN is tested for performance comparison with the hybrid HHT approach. The candidate input vectors for the FF-ANN-SA model were: the current value of wind speed at time t; previous values of wind speeds at times $(t - 1)$, $(t - 3)$, $(t - 6)$; wind direction at times t and $(t - 1)$; and time of day. All input vectors and the target vector were normalized to the range $[0, 1]$, and

the data were divided in the ratio 4:1 between training and validation sets, with the final 48 hours reserved for testing as before.

A hybrid simulated annealing/tabu search algorithm for simultaneous optimization of the network structure and weights was applied to the problem of 24h ahead prediction of wind speeds [18; 20]. The structural optimization element of the algorithm uses a network pruning approach, which involves defining an initial, maximal network and then randomly removing intra-neuron connections until a more parsimonious network is arrived at (Fig. 3). The inter-neuron connectivity and weights were optimized using the hybrid algorithm prior to subsequent optimization of weights and biases using Levenberg-Marquadt backpropagation.
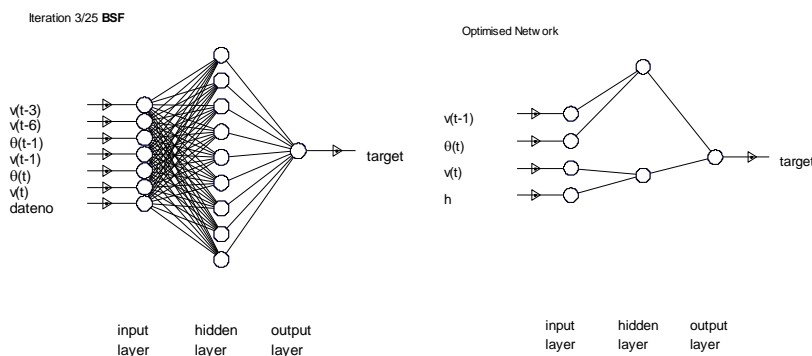

Fig. 3. (a) maximal network; (b) FF-ANN-SA network with optimised structure.

The maximal network was defined as having all seven candidate inputs, and a single hidden layer with nine neurons. Therefore, the maximum number of inter-neuron connections was $N_1 N_2 + N_2 N_3$ where N_1 is the number of inputs, N_2 the number of hidden layer neurons, and N_3 the number of outputs. The cost function specified for the optimization is based on predictive performance and number of inter-neuron connections in the network, with a 3:1 weighting ratio between network complexity and performance. Tan-sigmoidal and linear activation functions were used for the hidden and output layer neurons, respectively. The linear activation function of the output layer was constrained to produce only positive outputs, in order to represent the physical reality of wind speeds which are zero-limited. The algorithm was allowed to run for a maximum of 25 iterations, or until the GL5 early stopping criterion was satisfied [25].

In Table I and Figure 4 the testing results for a "24 hour ahead" wind speed forecasting are presented based on the HHT –ANN, HHT–SVM and FF-ANN-SA models. As shown in Table I, the most accurate forecast was given by the hybrid HHT-SVM and HHT-ANN models (MAE for 0-24h

period – 0.58 ms and 0.55 ms respectively), compared with the conventional approaches: the FF-ANN-SA model – 2.70 ms.
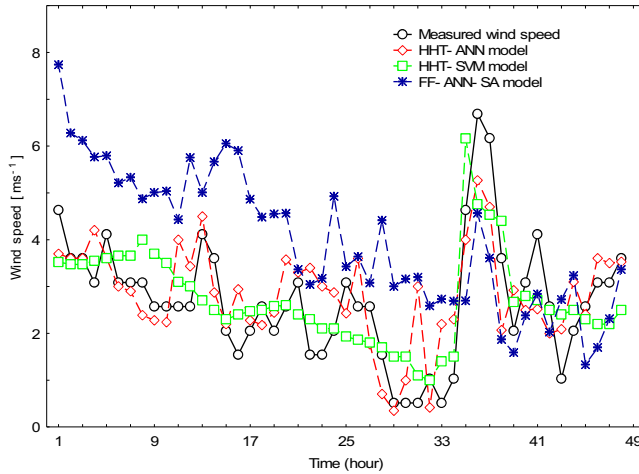

Fig. 4. "24-hour ahead" wind speed forecast based on hybrid HHT-ANN and HHT-SVM models.

Tab. 1

Comparison of "24hour ahead" Wind Speed Forecasts on the Basis of Hybrid and Conventional Neural Approaches for two subsets of the test data.

No.	Test data interval	Error ms	Models		
			HHT-ANN	*HHT-SVM*	*FF-ANN-SA*
1	0-24h	MAE	0.55	0.58	2.70
		RMSE	0.75	0.88	3.30
2	24-48h	MAE	0.73	0.93	1.54
		RMSE	1.05	1.17	1.69

The FF-ANN-SA algorithm did not perform as well as the HHT methods but identified the inputs with the most predictive power as: the current wind speed, the wind speed of the previous hour, the wind direction and the time of day.

5.2. ACTIVE POWER FLOW FORECASTING FOR THE WESTERN BAIKAL-AMUR MAINLINE (RUSSIA)

The BAM external power supply is carried on the 110 kV power network (from "Taishet" to "Lena" substations) and the 220 kV power network (from

" Lena" to "Taksimo" substations). Continuous and random changes in time traction load and the presence of a two-way feed greatly complicates the forecasting of state variables.
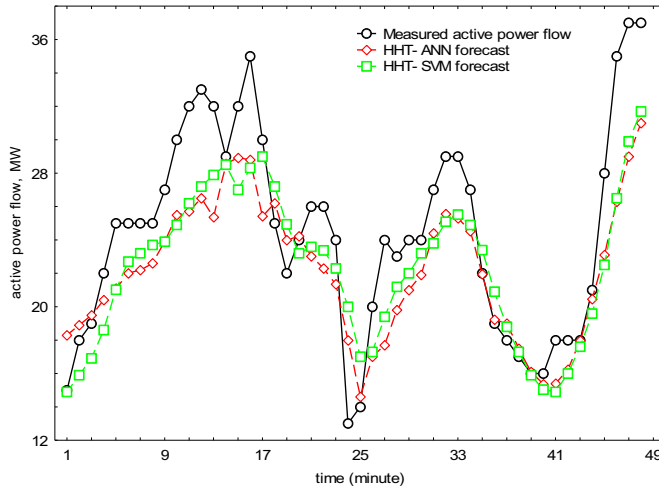

Fig. 5. Result of "1 minute ahead" active power flow forecasting based on hybrid genetic models.

The initial training data is represented by the telemetric data of the per-minute active power flow for 3 days (4320 values). The initial time series were converted using the HHT algorithm. The obtained HHT components are used as input values of the trained machine learning-based forecasting models. The MLP neural network model is trained with the Broyden–Fletcher–Goldfarb–Shanno algorithm [32]. The radial basis function kernel function is used in the epsilon-SVM regression model. Using a 10-fold cross-validation method the following optimal training constants and kernel parameters were determined: the capacity constant, $C = 10$; epsilon, $\varepsilon = 0.4$; degree, $\gamma = 0.001$. To compare the performance of forecasting approaches, the authors also calculated similar forecasts based on conventional approaches: Autoregressive integrated moving average (ARIMA), single ANN and Multiplicative Exponential Smoothing.

Table II and Fig. 5 summaries the numerical results for a "1 minute ahead" active power flow forecasting on based on the HHT-ANN and HHT-SVM models. As Table II shows, the most accurate forecast was given by the hybrid HHT- SVM and HHT-ANN models (MAPE - 11.6% and 12.5% respectively), compared with the conventional approaches: the ANN model – 15,8%, the ARIMA model – 22.4 % and Exponential Smoothing – 32.2%.

Tab. 2

Comparison of a "1 minute ahead" Acive Power Flow Forecasts on the Basis
a Hybrid and Conventional Approaches.

No.	Forecasting models	Error		
		MAPE, %	MAE, MW	RMSE, MW^2
1	HHT-ANN	12,5	3.2	3,9
2	HHT-SVM	11,6	2.9	3,5
3	ANN (MLP)	15,8	6,9	33,6
4	ARIMA	22.4	7.1	7.4
5	Exponential Smoothing	34.2	10.8	10.9

5.3. ELECTRICITY PRICE FORECASTING FOR THE AUSTRALIAN MARKET (AUSTRALIA)

The Australian NEM is a duopoly with a dominant player, resulting in electricity price changes associated with the behavior of dominant players, and is very difficult to predict in the short term. To assess the forecasting performance of different models, each dataset is divided into two samples for training and testing. The training dataset is used exclusively for model development, and then the test sample is used to evaluate the established model. To build the forecasting model for each of the considered weeks, the information available includes hourly price historical data of the four weeks prior to the first day of the week whose prices are to be predicted.

Here we present the comparison of forecast performance with the hybrid ARIMA-ANN model proposed by Phatchakorn Areekul et al. in [1] and our HHT-ANN hybrid using the data of Australian national electricity market, New South Wales. By analogy with [1] for the sake of a fair comparison, the fourth week in one season is selected.

The RF-BT-based importance analysis is shown that in contrast to previous results, in this case study, two first high-frequency components of electricity price data IMF1 (BT – 1.0, RF – 0,79) and IMF2 (BT – 0.95, RF – 0,70) have high prediction importance. Components IMF5 and F2 have less prediction importance (less than 0.20) and therefore they can be excluded from the trained data.

Table III and Figure 6 present the values of RMSE, MAE, and MAPE from the price forecasting performance of the proposed hybrid HHT-ANN-SVM model and hybrid ARIMA-ANN model. The results demonstrate that the proposed HHT-ANN (MAPE – 12.24%) hybrid model could provide a considerable improvement of the electricity price forecasting accuracy comparing to the ARIMA-ANN – 13.03% and the HHT-SVM – 17.16% hybrid models. It is fair to note that the HHT-SVM displays smaller errors

Comparison of "1 hour ahead" Electricity Price
Forecasting on the Basis of Hybrid Approaches.

No.	Forecasting models	Error		
		$MAPE,$ $\%$	$MAE,$ MW	$RMSE,$ MW^2
1	HHT-ANN	12,24	3.86	20,12
2	HHT-SVM	17,16	7.16	23,74
3	ARIMA-ANN	13,03	7.12	28,02

indicated by MAE and RMSE (respectively 7.01% and 23.74%) compared
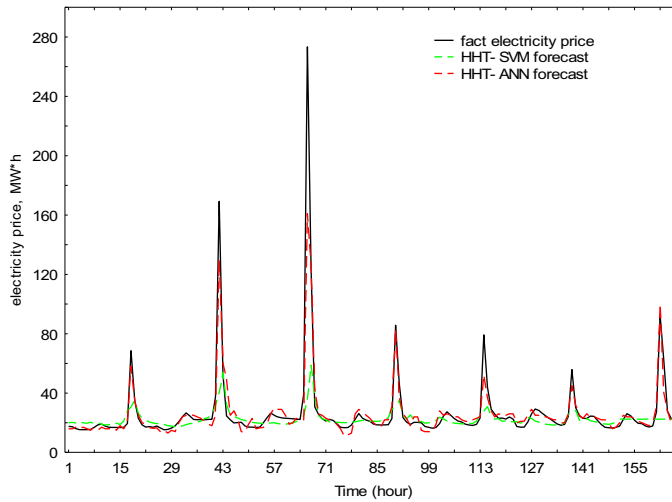to the ARIMA-ANN model (7.12% and 28.02%).

Fig. 6. Forecasting results (in AUD/MWh) of May 21-27, 2006 based on hybrid
HHT-ANN-SVM model.

6. Conclusions

In this paper we studied the problem of short-term forecasting of power
system parameters. The main contribution is the hybrid machine learning-
based algorithm using HHT as a data preprocessing technology. The propo-
sed approach is based on mode decomposition and a feature analysis of
initial retrospective data using HHT and BT-RF-based machine learning

algorithms for a prediction importance evaluation. We employed machine learning technologies such as ANNs and SVM as forecasting models.

Forecast performance comparison studies with other conventional, machine learning and hybrid approaches were also presented, especially with the FF-ANN-SA approach. The hybrid ANN-SA optimization method does not approach the performance of the HHT methods, as indicated by higher RMSE and MAE. Performance of the hybrid optimization ANN method can probably be improved with some further adjustments, but it is not expected that it will approach the performance of the superior HHT methods due to the non-stationarity of the wind speed time series. Also one more circumstance should be noted. As for today there are a lot of results in the HHT time series processing field which are associated with various modifications of the HHT. In [31], a modification of the standard approach to the decomposition of IMF was suggested, and in [28] there is the variation of this proposed modification. Both of these approaches are based on the idea of adding a white noise to the original signal and the construction of an ensemble of different sets of components. Averaging over the components of an ensemble provides a set of IMF, which contains the components reflecting the real components of the investigated process more accurately than the original EMD. These modifications of the proposed hybrid machine learning-based algorithm will be employed in our future works.

References

1. Areekul Ph. Senjyu T., Toyama H., Yona A. A Hybrid ARIMA and Neural Network Model for Short-Term Price Forecasting in Deregulated Market. *IEEE Trans. Power Syst*, 2010, vol. 25, no 1, pp. 524-530.
2. Breiman L., Friedman J., Olshen R., Stone C. Classification and Regression Trees. Belmont California, Wadsworth, 1984.
3. Breiman L. Random Forests. *Machine Learning*, 2001, vol. 45, pp. 5-32.
4. Damousis I.G., Dokopoulos P. A fuzzy model expert system for the forecasting of wind speed and power generation in wind farms. *Proc. of the IEEE International Conference on Power Industry Computer Applications PICA 01*, 2001, pp. 63-69.
5. Friedman J. Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, 2001, vol. 29.
6. Foley A.M. Leahy P.G., Marvuglia A., McKeogh E.J. Current Methods and Advances in Forecasting of Wind Power Generation. *Renewable Energy*, 2012, vol. 37, no 1, pp. 1-8.
7. Friedman J.H. Stochastic gradient boosting. *Computational Statistics and Data Analysis*, 2002, vol. 38, pp. 367-378.
8. Fugon L., Juban J., Kariniotakis G. Data mining for wind power forecasting. *Proc. the European Wind Energy Conference*, Brussels, Belgium, Apr. 2008.
9. Garcia R.C., Contreras J., M. van Akkeren, Batista J., Garcia C. A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices. *IEEE Transactions on Power Systems*, 2005, vol. 20, no 2.

10. Gerikh V., Kolosok I., Kurbatsky V., Tomin N. Application of Neural Network Technologies for Price Forecasting in the Liberalized Electricity Market. *Scien. Journal of Riga Tech. Univer. Ser. Power and Electrical Engineering*, 2009, vol. 5, pp. 91-96.

11. Glazunova A.M., Forecasting Power System State Variables on the Basis of Dynamic State Estimation and Artificial Neural Networks. *Proc. the IEEE Region 8 SIBIRCON-2010*, Listvyanka, Russia, 2010

12. Hippert H. Steinherz, Pedreira C. Eduardo, Souza R. Castro. Neural Networks for Short-Term Load Forecasting: A Review and Evaluation. *IEEE Trans. On Power Systems*, 2001, vol. 16, no 1.

13. Kavasseri R.G., Seetharaman K. Day-ahead wind speed forecasting using f-ARIMA models. *Renewable Energy*, 2009, vol. 34, no. 5, pp. 1388-1393.

14. Khotanzad A.Z., Elragal H. A neuro-fuzzy approach to short-term load forecasting in a price sensitive environment. *IEEE Trans. Power Syst*, 2002, vol. 17, no 4, pp. l273-1282.

15. Kurbatsky V. Sidorov D., Spiryaev V., Tomin N. On the Neural Network Approach for Forecasting of Nonstationary Time Series on the Basis of the Hilbert–Huang Transform. *Automation and Remote Control*, 2011, vol. 72, no. 7, pp. 1405-1414.

16. Kurbatsky V., Tomin N., Sidorov D., Spiryaev V. Hybrid Model for Short-Term Forecasting in Electric Power System. *International Journal of Machine Learning and Computing*, 2011, vol. 1, no 2, pp. 138-147.

17. Kurbatsky V., Sidorov D., Spiryaev V., Tomin N. Forecasting Nonstationary Time Series on the Basis of Hilbert-Huang Transform and Machine Learning. *Automation and Remote Control*, 2014, vol. 75, no 4, pp.12-16.

18. Leahy P. et al. Structural optimisation and input selection of an artificial neural network for river level prediction. *Journal of Hydrology*, 2008, vol. 355, pp. 192-201.

19. Louppe G. et al. Understanding variable importances in forests of randomized trees. *Advances in Neural Information Processing Systems*, 2013, pp. 431-439.

20. Ludermir T.B. et al. An optimization methodology for neural network weights and architectures. *IEEE Trans. on Neural Networks*, 2006, vol. 17, pp. 1452-1459.

21. Mercer J. Functions of positive and negative type, and their connection with the theory of integral equations. *Phil. Trans. of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 1909, vol. 209, pp.415-446.

22. Natenberg E.J. Gagne II D.J., Zack J.W., Manobianco J., Van Knowe G.E., Melino T. Application of a Random Forest Approach to Model Output Statistics for use in Day Ahead Wind Power Forecasts. *Proc. the Symposium on the Role of Statistical Methods in Weather and Climate Prediction*, USA, Austin, 2013.

23. Negnevitsky M., Voropai N., Kurbatsky V., Tomin N., and D. Panasetsky Development of an Intelligent System for Preventing Large-Scale Emergencies in Power Systems. *IEEE/PES General Meeting, Vancouver, BC, Canada*, 21-25 July 2013, pp. 1-6.

24. Neville P.G. Controversy of Variable Importance in Random Forests. *Journal of Unified Statistical Techniques*, 2013, vol. 1, no 1, pp. 15-20

25. Prechelt L. Proben1 - A set of neural network benchmark problems and benchmark rules. Computer Science Faculty, University of Karlsruhe, Germany, Tech.l Rep. 21/94, Sept. 1994.

26. Sinha N., Lai L.L., Kumar Ghosh P., Ma Y. Wavelet-GA-ANN Based Hybrid Model for Accurate Prediction of Short-Term Load Forecast. *Proc. the IEEE Inter. Conf. on ISAP*, 2007, Toki Messe, Niigata, pp. 1-8.

27. Sidorov D. N. Methods of Analysis of Integral Dynamical Models: Theory and Applications (Russian). ISU Publ., Irkutsk, 2013, 293 p.

28. Torres M. E. et al. A complete ensemble empirical mode decomposition with adaptive noise. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2011, pp. 4144-4147.
29. Voropai N.I. Glazunova A.M., Kurbatsky V.G., Sidorov D.N., Spiryaev V.A., Tomin N.V. Operating Conditions Forecasting for Monitoring and Control of Electric Power Systems. *Proc. the IEEE ISGT Europe 2010 Conference*, Gothenburg, Sweden, 2010.
30. Watters C.S., Leahy P. Comparison of linear, Kalman filter and neural network downscaling of wind speeds from numerical weather prediction. *Proc. 2011 10th International Conference on Environment and Electrical Engineering*, Rome, Italia, 2009.
31. Wu Z., Huang N. E. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Advances in adaptive data analysis*, 2009, vol. 1, no. 1, pp. 1-41.
32. Zhu C., Byrd R.H., Lu P., Nocedal J. Algorithm 778: L-BFGS-B, Fortran subroutines for large scale bound constrained optimization. *ACM Transactions on Mathematical Software*, 1997, vol. 23, no 4, pp. 550-560.

Kurbatsky Victor Grigorievich, Doctor of Sciences (Technics), Leading Research Scientist, Energy Systems Institute, Siberian Branch of Russian Academy of Sciences, 130, Lermontov st., Irkutsk, 664033,
tel.: (3952) 500646, ext. 235

Leahy Paul, Ph.D., Science Foundation Ireland lecturer in Wind Energy Engineering, Environmental Research Institute, University College Cork, Cork, Ireland, (e-mail: `paul.leahy@ucc.ie`)

Spiryaev Vadim Aleksandrovich, Junior Research Scientist, Energy Systems Institute, Siberian Branch of Russian Academy of Sciences, 130, Lermontov st., Irkutsk, 664033, tel.: (3952) 500 646, ext. 258

Tomin Nikita Victorovich, Candidate of Sciences (Technics), Senior Research Scientist, Energy Systems Institute, Siberian Branch of Russian Academy of Sciences, 130, Lermontov st., Irkutsk, 664033, tel.: (3952) 500 646, ext. 235

Sidorov Denis Nikolaevich, Candidate of Sciences (Physics and Mathematics), Senior Research Scientist, Energy Systems Institute, Siberian Branch of Russian Academy of Sciences, 130, Lermontov st., Irkutsk, 664033; Associate Professor, Irkutsk State University, 1, K. Marx st., Irkutsk, 664003, (e-mail: `contact.dns@gmail.com`)

Zhukov Aleksei Vitalievich, Postgraduate, Institute of Mathematics, Economics and Computer Science of Irkutsk State University, 1, K. Marx st., Irkutsk, 664003 (e-mail: `zhukovalex13@gmail.com`)

В. Г. Курбацкий, В. А. Спиряев, Н. В. Томин, П. Лихи, Д. Н. Сидоров, А. В. Жуков

Прогнозирование параметров электроэнергетических систем используя преобразование Гильберта-Хуанга и машинное обучение

Аннотация. Разработан гибридный подход прогнозирования нестационарных временных рядов. Подход продемонстрирован на примере задач прогнозирования параметров систем электроэнеретики. Предлагаемый адаптивный подход использует разложение на моды анализ признаков ретроспективных данных на основе преобразования Гильберта-Хуанга и алгоритмов машинного обучения. Используются методы машинного обучения на основе алгоритмов случайный лес (Random forest) и градиентный бустинг над деревьями (Gradient boosting trees). Эти методы, а также индекс Джини, используются для ранжирования значимости признаков в прогнозных моделях. Полученные гибридные прогнозные модели используют нейронную сеть на основе радиальных базисных функций и регресионную модель на основе метода опорных векторов. Помимо введения и списка источников статья организована следующим образом. Второй раздел содержит обзор современных подходов для краткосрочного прогнозирования параметров энергетических систем. В третьем разделе изложен разработанный гибридный подход, основанный на машинном обучении и использующий преобразование Гильберта-Хуанга для краткосрочного прогноза параметров электроэнеретических систем. В четвёртом разделе описаны алгоритмы обучения решающих деревьев для определения значимости переменных. В заключении представлены экспериментальные результаты в следующих электроэнергетических задачах: прогнозирование перетоков активной мощности, прогнозирования цен на электроэнергию и прогнозирование скорости и направления ветра для ветровых генераторов электроэнергии.

Ключевые слова: временной ряд, прогнозирование, интегральное преобразование, ИНС, МОВ, машинное обучение, бустинг, сингулярный интеграл, анализ признаков

Курбацкий Виктор Григорьевич, д.т.н., главный научный сотрудник, Институт систем энергетики СО РАН, ул. Лермонтова, 130, Иркутск, 664033, тел: (3952) 500 646, вн. 235

Спиряев Вадим Александрович, младший научный сотрудник, Институт систем энергетики СО РАН, ул. Лермонтова, 130, Иркутск, 664033, тел: (3952) 500 646, вн. 258

Томин Никита Викторович, к.т.н., старший научный сотрудник, Институт систем энергетики СО РАН, ул. Лермонтова, 130, Иркутск, 664033 тел: (3952) 500 646, вн. 235

Лихи Пол, Ph.D., лектор научного фонда Ирландии по ветровой электроэнергетике, Институт окружающей среды, Университет колледж Корк (Ирландский национальный университет), Ирландия, (e-mail: paul.leahy@ucc.ie)

Сидоров Денис Николаевич, к.ф.-м.н., старший научный сотрудник, Институт систем энергетики СО РАН, ул. Лермонтова, 130, Иркутск, 664033; доцент, Иркутский государственный университет, (e-mail: contact.dns@gmail.com)

Жуков Алексей Витальевич, аспирант, Иркутский государственный университет, (e-mail: zhukovalex13@gmail.com)