

A mathematical model for estimating and forecasting the state of a digital substation based on the wavelet transform method

Roman Oganyan^{1*}, Danil Shaykhutdinov¹, Nuri Narakidze¹, and Dmitriy Shurygin¹

¹SRSPU(NPI), Department of Information and measuring systems and technologies, 346428 Novochoerkassk, Russia

Abstract. The article the possibility of applying the wavelet transform method in combination with a neural-fuzzy approach to solving problems of forecasting the state of digital substations is considered. The optimum level of the wavelet expansion of the time series corresponding to the change in the phase voltage for the day on the basis of the Hurst index is determined. The influence of the sample size and the type of the mother wavelet on the Hurst index is researched. It was revealed, that for wavelet decomposition, the use of the Daubechey wavelet as the mother wavelet is effective, which provides a smoother filtering of noise, compared to the Haar wavelet. Analysis of the original series does not allow to evaluate the optimal level of wavelet expansion if the noise level of the time series under consideration is low (less than 10%), since the Hurst index remains unchanged. However, using the logarithm of changing the time series allows for small fluctuations to be taken into account, which allows to determine the optimal level of the wavelet expansion for their smoothing.

1 Introduction

The technology "Digital substation" is a new level of automation of electrical substations. Based on the fact that this technology is relatively new, there is a need to develop new algorithms for diagnosing complex technical systems. The demand is confirmed by numerous researches and developments in this field [1-9].

As the basis of new algorithm for complex technical systems diagnosing (as example – for digital substation), that earlier proposed in [2], the method of full-scale-model test is used [10]. To diagnostics a computer model based on National Instruments technologies is developed [11]. The model allows to select the monitored parameter (current strength, phase/line voltage, frequency, etc.) and the monitored value (the numerical value of the monitored parameter, exceeding which is dangerous) during program operation. The model also allows to estimate the change of the monitored parameters within the time. Thus, it becomes possible to more accurate diagnose digital substation current state on the basis of an analysis of the parameters values.

However, the diagnosis of the substation current state does not always allow assessing the evolving defects in the equipment. Therefore, it becomes difficult to prevent the emergency operation of the equipment. To solve this problem, it is proposed to use methods of intellectual analysis to predict the substation state. In view of the fact that the data on the digital substation is permanently stored, there is access to a large amount of information provided from the sensors of power equipment. Thus, it becomes possible to use intellectual analysis (modern

methods of signal processing, wavelet analysis, methods of artificial intelligence, machine learning) to diagnose equipment. The use of intellectual analysis allows to assess the current state of equipment, as well as to predict the pre-toxic states on the basis of an analysis of the values of the parameters characterizing the state of the equipment, as well as the relationships of the system parameters [3].

2 The wavelet transform method

To create a predictive model, the time series of parameters describing the state of the substation according to the *C2C power quality monitoring* project are taken. Data about a digital substation can be considered as random time series. The analysis of such time series reflects the behavior of the system in the past and makes it possible to predict its future behavior. The primary task in forecasting complex non-stationary time series is to decompose them as components. The structure of the time series is a set of components, such as the systematic component, the oscillations around the systematic component and the random component.

One of the most effective methods of intelligent data processing are neural networks. However, this method has drawbacks [12, 13]. The disadvantages of this method are partially eliminated by combining neural networks and fuzzy logic – a neuro-fuzzy approach [14]. According to numerous studies [15-22], due to the use of a neuro-fuzzy network, the forecast error is minimized. At the same time, a selection of the architecture of the neural-fuzzy network, taking into account the dynamics of the change in the rapidly changing and slowly

* Corresponding author: roman.work18@gmail.com

changing components of the time series, is a laborious process. In addition, accurate accounting for changes in the rapidly changing time series component leads to significant prediction errors. The solution of the above-mentioned problems, as well as an additional increase in the accuracy of the forecast, can be ensured by applying a wavelet transform. Forecasting is carried out according to the scheme shown in Fig. 1.

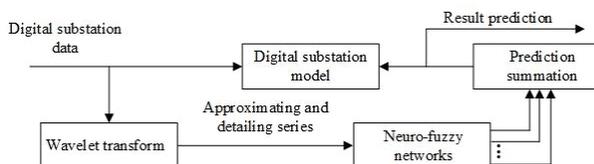


Fig. 1. Prediction diagram.

To solve the problem of predicting discrete time series, an algorithm of excessive discrete wavelet transform is chosen [23,24]. The choice is due to the fact, that other species have disadvantages that do not allow to solve the current task:

- continuous wavelet transformation - a large redundancy, leading to a large time-consuming analysis,
- fast wavelet transformation - presence of decimation operation, i.e. when the time series is shifted by at least one sample, the results of the whole decomposition will change, which will lead to the impossibility of forecasting).

3 The optimal level of detail based on the Hurst index

The prediction process (Fig. 1) includes the decomposition of a time series with a deterministic depth of detail, which is determined empirically from the analysis of the original time series, or on the basis of an analysis of the Hurst exponent [25]. This indicator characterizes the noisiness of the analyzed time series. The best decomposition depth is understood as the obtaining of uncomplicated quasiperiodic sequences at the output of the wavelet filtration with no sharp fluctuations and emissions at the approximating decomposition level (the Hurst index value is different but close to 0.5) [26].

For example, consider the expansion of the time series, which is a change in the phase voltage per day (1440 values, fixing the voltage every minute). The appearance of the change in voltage is shown in Fig. 2.

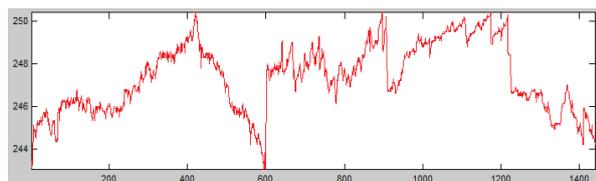


Fig. 2. The one day voltage level variations

To determine the optimal level of detail, an algorithm for calculating the Hurst index based on R / S analysis

was implemented [27,28]. There are two ways of analyzing the time series:

- analysis of the original time series,
- analysis of variation in the original time series.

By variation of the original series N we mean either the difference between two neighboring terms of the series [29]:

$$n = N_{i+1} - N_j, i = 1, 2, 3, \dots, N-1 \quad (1)$$

or the difference between the logarithms of two neighboring terms of the series [28,29]:

$$n = \ln(N_{i+1} / N_j), i = 1, 2, 3, \dots, N-1 \quad (2)$$

According to the E. Neumann research "Calculation of the Hurst index for the purpose of revealing the trend (persistence) of financial markets and macroeconomic indicators," it is recommended to analyze the source series, rather than the original series variation. To calculate the Hurst exponent, used the formula [30] is used:

$$H = \log(R / S) / \log(\pi * N / 2), \quad (3)$$

where H – the Hearst index; S – the standard deviation of a observations number; R – the magnitude of the accumulated deviation; N – the observation periods number.

The results of calculating the Hurst index for the initial series are presented in Table 1.

Table 1. The values of the Hurst index for the original time series.

Partition number	n value (the number of days for the period)	The number (t) of segments	The H value
1	72	20	0,5786
2	120	12	0,6633
3	180	8	0,7553
4	288	5	0,7514
5	480	3	0,7981
6	720	2	0,7876
	1440	1	0,7944

The value of the Hurst index, according to the recommendations [27], should be taken with $t \geq 5$. According [31], at small sample sizes, the Hearst index value is undergoes significant changes, and only starting from a certain sample size, the value of the Hearst index oscillates about a specific value. With respect to the time series under consideration, we take $H = 0.7514$, which indicates the presence of long-term memory in the analyzed series.

To test the possibility of further use of the obtained Hurst index, we select a test developed by Sheinkman and Le Baron [27]. The test consists in mixing the initial data with the subsequent recalculation of the Hurst index. If the Hurst index takes a value close to 0.5, we can conclude that the original series had a long-term memory, and as a result of mixing the original data, their structure was destroyed. In the case under consideration, after randomization of the initial data, the Hurst index

became 0.4741. This result indicates the presence of long-term memory in the original series and, accordingly, indirectly confirms the correctness of the definition of the Hurst index.

To obtain the best decomposition depth, the Hurst index is calculated from the first level of detail, followed by an increase based on the Haar wavelet [32, 33]. The results are shown in Table 2.

Table 2. The Hurst exponent value for the Haar wavelet.

Type of Motherboard Wavelet	Level of detalisation	The H value
Haar wavelet	1	0,7524
	2	0,7533
	3	0,7548
	5	0,7656

It can be seen, that the value of the Hurst index undergoes minor variation with an increase in the level of detail. The next step to get the Hurst index as a function of the level of detail was tested, replacing the mother wavelet with the Daubechies wavelet with a number of zero moments equal to 4 [32,33], which provides a more smooth filtering of noise. The results of the wavelet decomposition are presented in Table 3.

Table 3. The Hurst exponent value for the Daubechy wavelet.

Type of Motherboard Wavelet	Level of detalisation	The H value
Daubechy wavelet	1	0,7521
	2	0,7533
	3	0,7553
	5	0,7623

It can be seen, that the value of the Hurst index also does not undergo significant changes. To study the dependence of the Hurst index on the level of detail, the sample size was increased to 4,096 values (Fig. 3).

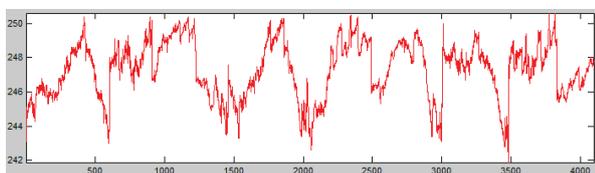


Fig. 3. The voltage level variations for 4096 datapoints

For the original time series, the value of the Hurst index is 0.7855. For the given sample size, the same procedures are repeated as for the previous time series. The results are shown in Table 4.

Table 4. The Hurst index value for the Haar and Daubechies wavelets with a sample size of 4096 values.

Type of Motherboard Wavelet	Level of detalisation	The H value
Haar wavelet	1	0,7859
	2	0,7863
	3	0,7873
	5	0,7914
Daubechy wavelet	1	0,7857
	2	0,7862
	3	0,7870
	5	0,7902

Thus, from the data, it is clear that to form a relationship between the level of detail and the Hurst index, it is necessary to analyze the logarithms of the change in the time series under consideration. This is due to the fact that due to the use of the properties of the logarithm, it will be possible to take into account the detailed changes in the time series under consideration. The results of calculating the Hurst index are presented in Table 5.

Table 5. The Hurst index value in the analysis of changes in the original time series.

Partition number	The n values (the number of days of the period)	The number (t) of segments	The H value
1	72	20	0,3747
2	120	12	0,4104
3	180	8	0,4576
4	288	5	0,4578
5	480	3	0,4448
6	720	2	0,4846
	1440	1	0,4381

According to the E. Neumann research, the value of the obtained Hurst index falls under the classification of a random series, or noisy. We carry out the wavelet decomposition using the mother wavelet Haar and Dobeši. The results are shown in Table 6.

Table 6. The Hurst exponent value for the Haar wavelet and Daubechies for different levels of detail.

Type of Motherboard Wavelet	Level of detalisation	The H value
Haar wavelet	1	0,4886
	2	0,4939
	3	0,4970
	5	0,4788
Daubechy wavelet	1	0,5161
	2	0,5639
	3	0,6068
	5	0,6955

According to the obtained data, it can be concluded that for the time series under consideration, the use of the Daubechy wavelet is more effective. This is due to

the fact that using the Haar wavelet does not allow to smooth out the current noise level, but because of the specifics of this wavelet, it only eliminates the information content of the original signal (Fig. 4).

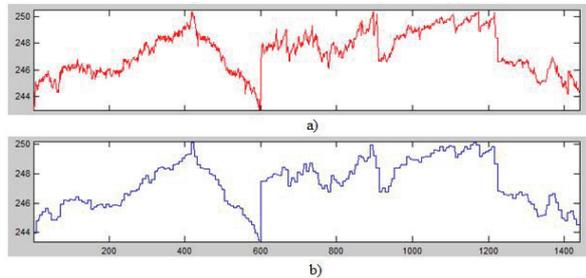


Fig. 4. Wavelet decomposition with a level of detail equal to three for the Haar mother wavelet: a) the original signal; b) an approximating signal

Based on the Daubechey wavelet, the results are similar to the "rough" estimate of the Hurst index without taking into account the logarithms of the change in the original series. It should be noted that with the increase in the level of detail, the Hurst index increases, respectively, indirectly confirms that the Hurst index is a measure of the complexity of the analyzed series.

Thus, the best depth of decomposition obtained as a result of approbation of different levels of detail for 4,096 values is 5. As the level of decomposition increases, a rough change in the configuration of the signal occurs, at a lower level of decomposition the noise component will not be sufficiently eliminated. As a result of the decomposition, we obtain 6 components—the approximating (a_5) and five detailed (d_5, d_4, d_3, d_2, d_1), which together give the original signal (Fig. 5):

$$S \cong a_5 + d_5 + d_4 + d_3 + d_2 + d_1, \quad (4)$$

where S – source signal, a_5 –signal equal to approximation of 5; d_5, d_4, d_3, d_2, d_1 – the signal of detail levels 5, 4, 3, 2, and 1, respectively.

As can be seen from Fig. 5, the original signal is noisy. The use of a neural-fuzzy approach without wavelet transformation entails the appearance of an error, but taking an approximating time series, without taking into account the rapidly changing components, there will also be a significant prediction error [34].

Training is carried out for each level of decomposition separately, accordingly, in the case under consideration, 6 neural-fuzzy networks are formed. After the expansion of the time series, a prediction is carried out, which can be represented in the form of:

$$Y^*(t, l) = \varphi(t, l, Y(t), Y(t-1), \dots, Y(2), Y(1)), \quad (5)$$

where interval $[t, t+l]$ – the lead time interval, l – length of lead time interval, $Y^*(t, l)$ – pretiction, φ – prediction function, depends on the number of observations t , lehgth l and values of $Y(1), Y(2), \dots, Y(t)$. The received predictions $Y^*(t, l) \dots Yn^*(t, l)$, are summarized, to obtain the result prediction value [35].

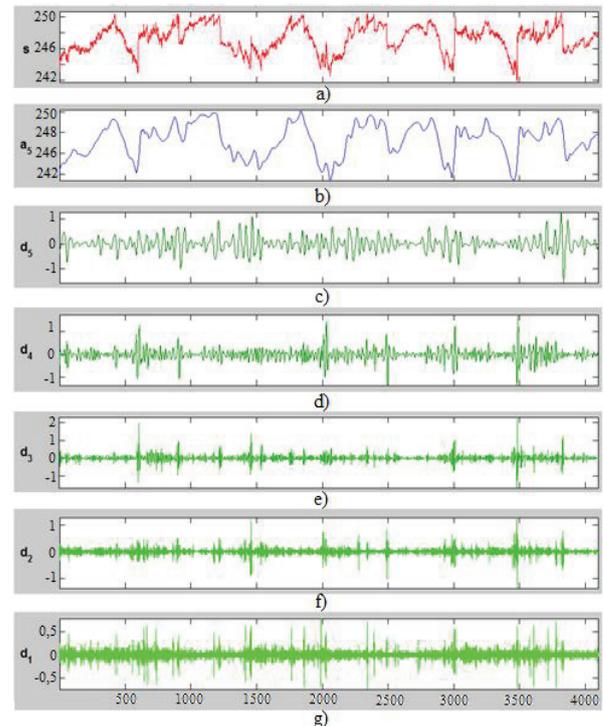


Fig. 5. The decomposition of a signal with a level of detail equal to three: a) the original signal; b) an approximating signal; c), d), e), f), g) the details of 5, 4, 3, 2 and 1 levels.

As a result of the research, it was revealed:

- For wavelet decomposition, the use of the Daubechey wavelet as the mother wavelet is effective, which provides a more smooth filtering of noise, in comparison with the Haar wavelet.
- The increase in the sample size does not have a significant effect on the Hearst index.
- Analysis of the original series does not allow us to estimate the optimal level of the wavelet decomposition if the noise level of the time series under consideration is low (less than 10%), since the Hurst index remains unchanged. However, using the logarithm of changing the time series allows for small fluctuations to be taken into account, which makes it possible to determine the optimal level of the wavelet expansion for smoothing them.

Thus, in order to solve the problems of predicting the state of a digital substation, the application of wavelet transform makes it possible to obtain a signal without noise, which is effective for further application in neural-fuzzy networks. This is due to:

- Eliminating the need to compile a neural-fuzzy network architecture,
- Lack of accurate accounting of the detailed components of the original time series, which leads to an increase in the accuracy of the forecast,
- Forecasting of individual rows after the decomposition of the original signal, which also leads to an increase in forecast accuracy.

This article is based on results of the project No. SP 4108.2015.1, implemented under the program "Scholarship

provided by the President of the Russian Federation for young scientists and graduate students engaged in advanced research and development in priority areas of modernization of the Russian economy” using equipment of shared facility “Diagnosis and energy-efficient electrical equipment” (NPI).

References

1. V.I. Dubrov, R.G. Oganyan, D.V. Shaikhutdinov, E.V. Kirievsky, N.I. Gorbatenko, N.D. Narakidze, *Modern high technology*. **9**, 36-40 (2016)
2. V.I. Dubrov, R.G. Oganyan, D.V. Shaikhutdinov, E.V. Kirievsky, TN. Kruglova, N.D. Narakidze, *Fundamental research*. **9**, 16-20 (2016)
3. L.A. Dar'yan, A.G. Mordkovich, G.M. Tsfasman, *Electro. Electrical engineering, electric power industry, electrotechnical industry*. **5**, 22-26 (2010)
4. Z. Ling, Y. Chaoran, H. Yanxiang, *Mechatronic Science, Electric Engineering and Computer (MEC)*, 1374-1377 (2011)
5. J. Xin, Z. Liao, F. Wen, *Power System Technology (POWERCON), Inter. Conf.* 1-5 (2010)
6. P. Peng, Z. Liao, F. Wen, J. Huang, *SCI* 465, 153-165 (2013)
7. D. La Cascia, V. Di Dio, A. Lipari, *Power Electronics and Motion Control Conf. and Expos (PEMC)*, 1300-1305 (2014)
8. R. G. Oganyan, N. D. Narakidze, D. V. Shaykhutdinov, N. I. Gorbatenko, S. O. Maksuta, *IOP Conf. Ser.: Mater. Sci. Eng.* **177**, 012083 (2017)
9. V.I. Dubrov, R.G. Oganyan, N.D. Narakidze, G.K. Aleksanyan, *Journal of Engineering and Applied Sciences*, **12**, 276-282 (2017)
10. N.I. Gorbatenko, *Full-scale model tests of products made of ferromagnetic materials* (Publishing house SKNC VSh, Rostov-on-Don., 2001)
11. V.I. Dubrov, R.G. Oganyan, N.D. Narakidze, G.K. Alexanian, *Collection of Proceedings of the XV International Conference NIDays-2016*, 300-302 (2016).
12. S.Yu. Galkina, I.P. Kurochkina, *Modern trends in the development of science and technology*. **12**, 28-30 (2016)
13. L.A. Amaeva, *Bulletin of the Kazan Technological University*, **1**, 320-322 (2015)
14. S. Osovsky, *Neural networks for information processing* (Finance and Statistics, Moscow 2004)
15. E.S. Anushina, *Intellectual systems*, **1 (19)**, 84-90 (2009)
16. V.I. Alekseev, *The Bulletin of the Ugra State University*, **3 (30)**, 3-10 (2013)
17. V.Z. Manusov, E.V. Biryukov. *Proceedings of Tomsk Polytechnic University. Engineering of georesources*, **6**, 153-157 (2006)
18. A.A. Dudkin. *System Analysis and Applied Informatics*. **4**, 11-15 (2015)
19. O.P. Soldatova, I.A. Lyozin, I.V. Lyozina, A.V. Kupriyanov, D.V. Kirsch. *Computer Optics*, **5**, 787-794 (2015)
20. A.D. Danilov, Sh.M. Omar, A.L. Rutskov. *Actual directions of scientific research of the XXI century: theory and practice*, **6**, 70-77 (2016)
21. V.L. Burkovsky, V.N. Krysanov, A.L. Rutskov, O.Sh.M. Shukur. *Bulletin of Voronezh State Technical University*. **5**, 41-46 (2015)
22. O.Yu. Kuznetsova, A.A. Solomakh, V.I. Gorbachenko. *Modeling and analysis of data*. **1**, 86-90 (2014)
23. A.V. Pereberin. *Computational methods and programming: new computational technologies*. **3**, 15-40 (2001)
24. I. Dobesi, *Ten lectures on wavelets* (Izhevsk: SRC "Regular and chaotic dynamics", 2001)
25. H.E. Hurst, R.P. Black, Y.M. Simaika, *Long-term storage: An experimental study* (Constable, Leningrad 1965)
26. E.S. Filatova, D.M. Filatov. *International Conference on Soft Computing and Measurement*. **1**, 114 -117 (2014)
27. E. Peters. *Fractal analysis of financial markets: application of the theory of chaos in investment and economy* (Internet-trading, Moscow 2004)
28. E. Peters. *Chaos and order in the capital markets. A new analytical view of cycles, prices and market volatility* (Mir, Moscow 2000)
29. M.S. Chepeleva, S.A. Tkulich, S.A. Chepelev, *Polytechnical network electronic scientific journal of the Kuban State Agrarian University*. **74**, 326-337 (2011)
30. F.F. Pashchenko, E.Yu. Medvedev, *Bulletin of the International Academy of System Studies. Informatics, ecology, economics*. **1**, 74-80 (2013)
31. Yu. A. Kalush, V.M. Loginov. *Siberian Journal of Industrial Mathematics*. **5**, 4(12), 29-37 (2002)
32. O.V. Nagornov, V.G. Nikitaev, V.M. Prostokishin, S.A. Tyuflin, A.N. Pronichev, T.I. Bukharova, K.S. Chistov, R.Z. Kashafutdinov, V.A. Khorkin. *Wavelet analysis in examples: Textbook*. (NNIU MEPhI, Moscow: 2010)
33. A.N. Yakovlev, *Introduction to wavelet transforms: Proc. Allowance* (NSTU Publishing House, Novosibirsk 2003)
34. E.S. Filatova, D.M. Filatov, A.D. Stotskaya. *International Journal of Applied and Fundamental Research*. **10-1**, 46-50 (2015)
35. Filatova, D.M. Filatov, A.D. Stotskaya. *International Journal of Applied and Fundamental Research*. **11**, 26-29 (2014)