

An Approach to the Selection of a Time Series Analysis Method Using a Neural Network

Dmitry Yashin, Irina Moshkina, and Vadim Moshkin(✉)

Ulyanovsk State Technical University,
Severny Venets Street, 32, Ulyanovsk 432027, Russian Federation
taurusrulez@yandex.ru, postforvadim@yandex.ru,
i.timina@ulstu.ru

Abstract. In this paper, a technique for selecting individual methods in a combined time series forecasting model is described. A neural network at the input of which a vector of time series metrics is proposed. The metrics corresponds to significant characteristics of the time series. The values of the metrics are easily computed. The neural network calculates the estimated prediction error for each model from the base set of the combined model. The proposed selection method is most effective for short time series and when the base set contains a lot of complex prediction methods. The developed system was tested on the time series from the CIF 2015-2016 competitions. According to the result of the experiment, the application of the developed system allowed to reduce the average forecast error from 13 to 9 percent.

Keywords: Time series · Forecasting · Combined model · Neural network

1 Introduction

The constant need to improve the accuracy of forecasting time series leads to an increase in the number of publications on this topic. It also contributes to the emergence of new forecasting methods, with the growing number of which, the problem of choosing the most appropriate method is becoming increasingly important.

One of the possible solutions to this problem is the use of the combined forecasting models, which allow to obtain an aggregated forecast based on the results of several methods. However, in this case, the choice of methods from a base set is also a problem since individual inexact forecasts can significantly reduce the overall accuracy.

In this paper, one of the approaches to solving the problem of selecting individual models from a base set is proposed. Used specially developed neural network (Sect. 2) that selects methods according to the values of the metrics corresponding to the characteristics of the time series (Sect. 1). Characteristics relevant to the subject area have been selected for this purpose.

Also in this paper, the set of metrics that is optimal for the solution of the problem is determined, the efficiency of the developed system is estimated (Sect. 3).

The method proposed in this work is most similar to the expert approach and involves the selection of models from the base set by a neural network that uses a set of simply computed time series metrics.

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12249, pp. 692–703, 2020.
https://doi.org/10.1007/978-3-030-58799-4_50

2 Combined Time Series Forecasting Model. Selecting Methods from the Base Set

2.1 Hybrid Systems in Artificial Intelligence

The concept of hybridization arises as a consequence of the development of the principles of the system approach to the study of complex artificial objects in computer science [1]. Professor N.G. Yarushkina highlights the development of hybrid integrated and synergetic systems as one of the leading trends in modern computer science. She defines hybridization as the integration of methods and technologies at a deep level, when the system includes interacting modules that implement various methods for solving problems of artificial intelligence [2]. A.V. Gavrilov, noting the hybridization process as the main trend in the development of artificial intelligence, considers hybridization as a variant of combining the methods of representation and processing of knowledge in hybrid intelligent systems [3].

Hybrid intelligent information systems are an interdisciplinary scientific direction, within the framework of which the applicability of several methods from different classes of the method of formalized representation of systems (analytical, statistical, logical-linguistic, fuzzy, and others) to solving decision-making problems is investigated. While none of the above classes can be considered universal [4]. The interaction of several methods makes it possible to compensate for their drawbacks due to the appearance of an integrative property.

2.2 Combined Forecasting Model

There are many methods for forecasting time series, each of which has its advantages and disadvantages. To take advantage of several methods, combined models are used. According to [5], the combined forecasting model is a prediction model consisting of several individual models, called the base set.

In [6] listed a number of factors that emphasize the effectiveness of the combined model, among which: the inability to choose a single model from experimental data; the presence in each rejected forecast of important for modeling information; the need to choose from a group of models that have similar statistical characteristics when trying to select a single best model.

There are two types of combined models: selective and hybrid. In the selective model, the forecast is calculated by one method selected from the base set at each time point. In the hybrid model, the predicted value is obtained by aggregating the prediction results for several models from the base set (usually, the final forecast is a weighted sum of individual forecasts) [5].

2.3 Selection of Models from the Base Set

According to [7], the main problem of constructing hybrid combined forecasting models is the finding of the optimal weights of individual model predictions that ensure the minimum value of the forecast error of the combined model. However, no less important is the problem of selecting models from the base set.

In this paper, a technique for selecting individual methods in a hybrid combined forecasting model is proposed. This approach can also be used for a selective combined model.

One of the possible ways to solve the sampling problem is to divide the time series into a training and control part with subsequent prediction of the control values by each method from the base set, then it becomes possible to select the models according to the obtained prediction error values. This method of selecting models is ineffective for combined models with a large number of methods in the base set, as well as for forecasting short time series. Another way to solve the problem is an expert evaluation based on the analysis of characteristics and the time series chart by a human expert. A combination of the two methods described above is also used. In this case forecasting the test values of the time series is applied only for models selected by the expert from the base set.

The method proposed in this work is most similar to the expert approach and involves the selection of models from the base set by a neural network that uses a set of simply computed time series metrics.

3 Time Series Metrics for Solving the Prediction Problem

There are many characteristics of a time series, such as the presence of a trend and the seasonal component, stationarity, length, anomaly of values, variance, mean. To predict the time series, the most important are the indicators of its dynamics [8]. According to [9], the two main statistical elements of dynamics are the trend and volatility.

3.1 Trend Metrics

The trend is the characteristic of the process of changing the phenomenon over a long time, freed from random fluctuations [10]. There are many criteria for determining the presence of the trend and for assessing its severity. Below are considered the most popular criteria.

Wallis-Moore Criterion. This criterion assumes the calculation of the differences in the values of the time series ($y_{t+1} - y_t$). Null hypothesis: the signs of these differences form a random sequence. A sequence of identical difference signs is called a phase. The number of phases h (without the first and last phase) is calculated. If the signs form a random sequence, then the actual value of the criterion is calculated by the formula (1) [11].

$$t_{fact} = \frac{\left|h - \frac{2n-7}{3}\right|}{\sqrt{\frac{16n-29}{90}}}, \tag{1}$$

where n is the length of time series, h is the number of phases.

Cumulative T-Criterion. Null hypothesis: no trend in the original time series. The calculated value of the criterion is defined as the ratio of the accumulated sum of

squared deviations of the empirical values of the time series observations from their mean value and the deviations themselves according to the formula (2) [12]

$$T_{\mathrm{p}} = \frac{\sum_n Z_n^2}{\sigma_y^2}, \tag{2}$$

where Z_n is the accumulated result of deviations of empirical values from the mean level of the original time series; σ_y^2 is the total sum of squared deviations, defined by formula (3)

$$\sigma_y^2 = \sum_n y_i^2 - (\bar{y})^2 \cdot n. \tag{3}$$

If a sufficiently long time series is analyzed, then a standardized deviation (4) can be used to calculate the criterion values

$$t_p = \frac{T_p - \left(\frac{n+1}{6}\right)}{\sqrt{(n-2)\frac{2n-1}{90}}}. \tag{4}$$

The calculated values of the cumulative T-criterion and t_p are compared with the critical values for a given significance level α.

Foster-Stewart Criterion. This criterion is used to test the trend of both mean and variance [13]. The statistics of the criterion have the form (5):

$$S = \sum_{i=2}^{n} S_i; \quad d = \sum_{i=2}^{n} d_i, \tag{5}$$

where $d_i = u_i - l_i; S_i = u_i + l_i;$

$$u_i = \begin{cases} 1, & x_i > x_{i-1}, x_{i-2}, \ldots, x_l; \\ 0, & else; \end{cases} \quad l_i = \begin{cases} 1, & x_i < x_{i-1}, x_{i-2}, \ldots, x_l; \\ 0, & else. \end{cases}$$

Statistics S is used to check the trend in variances, statistics d - to detect the trend in the mean. In the absence of a trend, the values of t and $\bar{t}$ (6) have a Student's distribution with $\nu = n$ degrees of freedom:

$$t = \frac{d}{f}; \quad \bar{t} = \frac{S - f^2}{k}, \tag{6}$$

where $k = \sqrt{2 \ln n - 3,4253}; \quad f = \sqrt{2 \ln n - 0,8456}.$

Koks-Stewart Criterion. To test the trend hypothesis, the normalized statistics (7) are applied [14]:

$$S_l^* = \frac{S_l - M(S_l)}{\sqrt{D(S_l)}}, \tag{7}$$

where $M(S_l) = \frac{n^2}{8}; D(S_l) = \frac{n\left(n^2-1\right)}{24}$.

Statistics S_l is calculated by the formula (8):

$$S_l = \sum_{i=1}^{\left[\frac{n}{2}\right]} (n - 2i + 1) h_{i,n-i+1}, \tag{8}$$

where $h_{i,j} = \begin{cases} 1, x_i > x_j \\ 0 \quad else \end{cases}, (i<j)$.

Under (9) the hypothesis of the average trend is rejected.

$$\left|S_l^*\right| < u_{\frac{1+\alpha}{2}}, \tag{9}$$

The Criterion Based on the Mean Comparison Method. To check the presence of a trend, the time series is divided into 2 parts. This criterion is based on the statement: if the time series has a tendency, the averages calculated for each set separately should differ significantly among themselves [15]. Null hypothesis on the absence of a tendency reduces to testing the hypothesis of the equality of the means of two normally distributed populations. The hypothesis is checked using Student's t-test, the calculated value of which is calculated by the formula (10):

$$t_r = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}} * \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}, \tag{10}$$

where $\bar{y}_1$ and $\bar{y}_2$ are the mean values, σ_1^2 and σ_2^2 are the dispersions and n_1 and n_2 are the lengths of the first and second part of the time series.

The calculated value (t_r) of the criterion is compared with its critical (tabular) value (t_{cr}) at a significance level α and the number of degrees of freedom $\nu = n - 2$.

3.2 Seasonal Metrics

There are many methods for estimating the degree of seasonal deviations of the time series. One of these methods is based on the calculation of seasonality indices. The seasonality index reflects the degree of seasonal variability of the phenomenon relative to the mean of the time series. This index is calculated by the formula (11):

$$I_{t,\text{ce3}} = Y_t / \bar{Y}, \tag{11}$$

where Y_t is the average value of the indicator for the month (quarter); $\bar{Y}$ is the overall average value of the indicator.

3.3 Dispersion and Dispersion Trend Metrics

As a metric, can be considered both the value of the variance and the degree of the variance trend. To test the hypothesis of the presence of a trend in dispersion, the Foster Stewart criterion is used. Also used the criterion based on the partitioning of the time series and the comparison of the variances of the two parts. Null hypothesis: the variances of two normally distributed sets are equal $\left(H_0 : \sigma_1^2 = \sigma_2^2\right)$.

The hypothesis is tested on the base of the Fisher-Snedecor F-criterion [16], the calculated value of which is obtained by the formula (12):

$$F_{\mathrm{p}} = \begin{cases} \sigma_1^2/\sigma_2^2, \sigma_1^2 > \sigma_2^2 \\ \sigma_2^2/\sigma_1^2, \sigma_1^2 < \sigma_2^2 \end{cases}, \tag{12}$$

The hypothesis is tested on the base of comparison of the calculated and critical values of the F-test obtained for a given level of significance α and the number of degrees of freedom ν_1 and ν_2 (13).

$$\begin{array}{l} \nu_1 = n_1 - 1, \text{ and } \nu_2 = n_2 - 1, \sigma_1^2 > \sigma_2^2 \\ \nu_1 = n_2 - 1 \text{ and } \nu_2 = n_1 - 1, \sigma_1^2 < \sigma_2^2 \end{array}, \tag{13}$$

3.4 Anomaly Degree Metrics

According to [17], anomalous observations are considered to be separate observations of the time series, the values of which differ significantly from the remaining observations. To assess the degree of anomaly of individual observations of the time series, special statistical criteria are used, for example, the Irwin criterion and the Chauvenet criterion.

Irwin Criterion. For all or only for suspected abnormalities in observations, λ_t is calculated (14). If the calculated value exceeds the table level, then the observation y_t is considered anomalous [18].

$$\lambda_t = \frac{|y_t - y_{t-1}|}{S}, \tag{14}$$

where $S = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \bar{y})^2}{n-1}}$.

Chauvenet Criterion. The sample element y_t is an outlier if the probability of its deviation from the mean is not more than 1/12n (where n is the sample size). For checked observations, the value of K (15) is calculated, which is compared with the tabulated value [9].

$$K = \frac{|y_t - \bar{y}|}{S}, \tag{15}$$

where $S = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \bar{y})^2}{n-1}}$.

3.5 Calculating the Values of the Metrics

To create and train a neural network for selecting forecasting methods, it was decided to establish the requirement that the value of each metric belong to the interval from 0 to 1. Before calculating the metrics, the time series are normalized. It is obvious that the variance of the normalized time series will be within the range from 0 to 1.

To obtain the final values of the metrics that involve computing the calculated value and comparing it with the critical value, we use formula (16).

$$m = \begin{cases} 1 - \frac{t_{crit}}{t_{fact}}, & t_{fact} \geq t_{crit} \\ 0,01, & t_{fact} < t_{crit} \end{cases}, , \tag{16}$$

where t_{fact} - calculated value of the criterion, t_{crit} - critical value of the criterion.

To calculate the metric of the degree of seasonality on the basis of seasonality indexes, the formula (17) is applied.

$$m_{ses} = \sqrt{\frac{\sum_{t=1}^{T}\left(\frac{I_t}{\bar{y}} - 1\right)^2}{T}}, \tag{17}$$

where T is the period of seasonality of the time series, $\bar{y}$ is the mean of the time series, I_t is the value of the average seasonality index for the period t.

To calculate the values of the metrics corresponding to the degree of anomaly of the time series, the ratio of the number of anomalous (according to the criterion) values to the length of the time series is used.

4 Neural Network for Selecting Prediction Methods

A specially designed neural network is used to select the prediction methods from the base set of the hybrid combined model (see Fig. 1). This neural network is created using the R language and the built-in package "neuralnet".

The input values of the neural network are the time series metrics. Each method (m_1, m_2…, $m_{k.}$) from the base set corresponds to one of the output layer's neurons (M_1, M_2…, $M_{t.}$).

The training set of the developed network is a set of TS metrics and the corresponding error values for predicting each TS by each method from the base set.

The neural network is trained to calculate the estimated forecast error values for each model from the base set based on the values of the TS metrics.

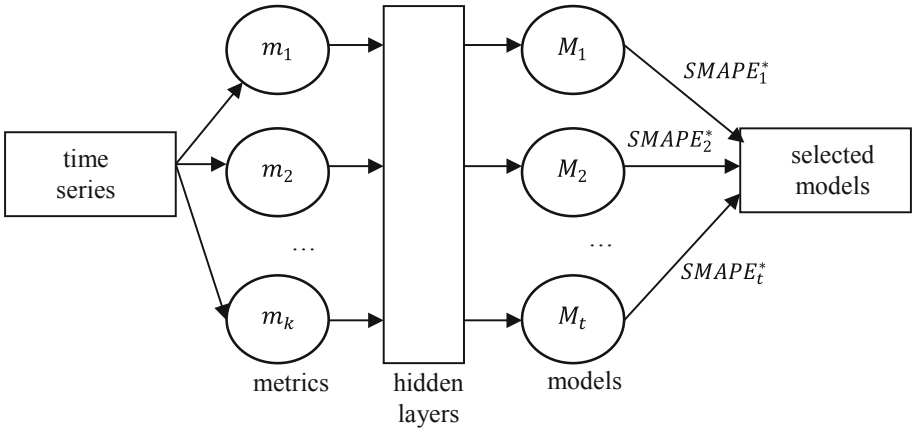

Fig. 1. Neural network circuit

An error back propagation algorithm with a logistic activation function is used for training.

The accuracy of the results of the neural network [19] depends on the similarity of the TS from the training set and the predicted TS and on the completeness and redundancy of the training set. The training set contains metric values and forecast error values for each TS included in it.

The training set consists of a set of time series represented as a set of metrics and the corresponding values of the prediction error (SMAPE) for each forecasting model. In this system, time series are taken from the competition "Computational Intelligence in Forecasting" (CIF) [12] for 2015-2016. The SMAPE values for the training set are obtained as a result of the forecasting of competitive time series using the methods of the combined hybrid forecasting system "Combination of fuzzy and exponential models" (CFEM) [13] developed at the Department of Information Systems of the Ulyanovsk State Technical University. This system was the best in the CIF competition in 2015

The metrics (and the methods for calculating them) for the input layer of the neural network are described in Sect. 3 of this work. The prediction models (corresponding to the output layer of the neural network) are taken from the combined CFEM model. The neural network learns from the metric values to calculate the estimated error values for each model from the base set. Then it is possible to determine which models from the base set are more efficient for predicting the values of the current time series. If the aggregating method involves adjusting the weights of the models, the weights can be set inversely proportional to the predicted error values of each model.

5 The Best Combination of Metrics. Results of Experiments

5.1 The Base Set of Metrics

Initially, 11 metrics were used in the developed neural network of selection of prediction methods: 5 metrics for the mean trend, 2 for the dispersion trend, 1 for the variance, 1 for the degree of seasonal variation, and 2 for the degree of anomaly in the time series.

The results of evaluations for one characteristic according to various criteria differ due to the differences in the regularities on which valuation methods are based. For example, the Wallis-Moore phase-frequency criterion for determining the trend of the mean is based on the comparison of the neighboring observations of the time series. The Foster-Stewart criterion - on the comparison of the current value with all the previous ones. The Cox-Stewart criterion - on the dividing the time series into three parts with the subsequent observation comparison between the third and the first group. The cumulative T-criterion - on the calculating the accumulated total of deviations of the empirical values from the time series mean [9, 11]. Proceeding from the foregoing, it can be argued that metrics that correspond to the same characteristic of a time series, but calculated by different criteria, may not have a high degree of correlation with each other.

In determining the optimal combination of time series metrics, the greatest difficulty is observing a balance between the completeness of the coverage of the regularities of the time series and the low degree of correlation between the characteristics. It was decided to perform an experiment where the metrics were sequentially excluded from the general set and the effect of this exception on the accuracy of the neural network was estimated.

5.2 The Experiment on Finding the Optimal Set of Metrics

Computational experiments were performed for several configurations of a neural network that differ in the number of neurons in the hidden layers. In some configurations, the neural network learning algorithm (based on the back propagation algorithm) did not converge. Such configurations were not taken into account. The experiments are carried out for neural networks with one hidden layer with the number of neurons in it from 1 to 11 and for networks with two hidden layers and the number of neurons in each of them from 1 to 8.

The training set of the neural network was divided 5 times into the training and control parts. After that, for each configuration (a set of metrics and the number of neurons in the hidden layers), the neural network was trained and the deviation of the output values from the test ones was calculated.

The results of the experiment are presented in Table 1. The columns of the table correspond to sets of metrics and show the sequence of excluding metrics from the base set, which allowed increasing the accuracy of the neural network. Each column is divided into two sub-columns containing information about the configuration of the neural network and the corresponding value of the network error. The rows of the table correspond to the best 7 configurations for each set of metrics and the average error for all configurations.

Table 1. The best combination of metrics.

All metrics		All metrics except the dispersion trend by the Foster-Stewart criterion		All metrics except the dispersion trend by the Foster-Stewart criterion and the cumulative T-criterion		All metrics except the cumulative T-test and both variance trend metrics	
Number of hidden neurons	SMAPE	Number of hidden neurons	SMAPE	Number of hidden neurons	SMAPE	Number of hidden neurons	SMAPE
2	9,764	7	8,994	7	9,476	10	9,016
(8,1)	9,794	(4,5)	9,314	(1,3)	9,667	8	9,103
9	9,870	(2,6)	9,314	10	9,720	7	9,392
(4,1)	9,968	(6,2)	9,788	(3,4)	9,777	1	9,553
(2,1)	10,075	(7,3)	9,869	2	9,782	2	9,592
(3,3)	10,099	(1,6)	9,924	(1,4)	9,795	(6,3)	9,777
(8,2)	10,194	4	10,050	6	9,957	(4,3)	9,788
…	…	…	…	…	…	…	…
average SMAPE	10,967	average SMAPE	10,851	average SMAPE	10,656	average SMAPE	10,769

According to Table 1, the greatest average accuracy for all configurations was achieved for a set containing all metrics except the dispersion trend metric by the Foster-Stewart criterion and the metrics based on the cumulative T-criterion. However, the subsequent exclusion of another metric of the dispersion trend allowed to obtain the most accurate results for individual configurations.

Another experiment was performed to evaluate the effectiveness of using the neural network to improve the accuracy of the final forecast. To obtain the aggregated forecast only methods with a low expected forecast error were taken from the base set. It was established that the preliminary selection of methods by a neural network containing all eleven metrics made it possible to reduce the error of the final forecast from 13.131% to 9.764%. The use of the best neural network configuration from the previous experiment (without a cumulative T-criterion and both variance trend metrics, with 10 neurons in the hidden layer) reduced the average final forecast error to 9.016%.

Also as a result of the experiment, it was established that the logistic activation function for the considered neural network allows to obtain significantly more accurate results than the hyperbolic tangent.

6 Conclusions

In this paper, a technique for selecting individual methods in a combined time series forecasting model is described. This approach is based on the use of a specially developed neural network that selects methods according to time series metrics.

We proposed a neural network structure and the algorithms for calculating metric values based on the characteristics of the time series. According to the obtained experimental data, the optimal set of the metrics is determined and the efficiency of the neural network application is estimated: the average forecast error decreased from 13,131% to 9,016%.

The application of the developed neural network is especially effective for short time series and for combined models with a large number of prediction methods (since it uses easily calculated metrics). In these cases, the approach that involves splitting the time series into the training and control parts and then calculating control values by each method is inefficient.

Planned in future work: aggregation of individual forecasts based on error values;

- setting aggregator weights based on the values of prediction errors on the control part of the time series;
- adjustment of aggregator weights based on the values of prediction errors assumed by the neural network;
- development of a neural network model for aggregation of the results of time series forecasting methods.

Acknowledgments. This work was supported by the Russian Foundation for Basic Research. Projects No. 18-47-730035, No. 19-07-00999 and 18-47-732007.

References

1. Tarasov, V.B.: From multi-agent systems to intellectual organizations: philosophy, psychology, informatics. Editorial, Moscow (2002)
2. Yarushkina, N.G.: Fuzzy hybrid systems. Theory and practice. Fizmatlit, Moscow (2007)
3. Gavrilov A.V.: Hybrid Intelligent Systems. NSTU, Novosibirsk (2003)
4. Kolesnikov, A.V., Kirikov, I.A., Listopad, S.V., Rumovskaya, S.B., Domanicky, A.A.: Solution of complex traveling salesman tasks using the methods of functional hybrid intelligent systems. Institute of Informatics Problems of the Russian Academy of Sciences, Moskow (2011)
5. Lukashin, Y.P.: Adaptive methods of short-term forecasting of time series. Finansy i statistika, Moscow (2003)
6. Novák, V.: Mining information from time series in the form of sentences of natural language. Int. J. Approximate Reasoning **78**, 192–209 (2016)
7. Perfilieva, I.: Fuzzy transforms: theory and applications. Fuzzy Sets Syst. **157**(8), 993–1023 (2006)
8. Abayomi-Alli, A., Odusami, M.O., Misra, S., Ibeh, G.F.: Long short-term memory model for time series prediction and forecast of solar radiation and other weather parameters. In: 2019 19th International Conference on Computational Science and Its Applications (ICCSA), pp. 82–92. IEEE (2019)
9. Kobzar, A.I.: Applied Mathematical Statistics. Fizmatlit, Moscow (2006)
10. Afanasyev, V.N., Yuzbashev, M.M.: Time series analysis and forecasting. Finansy i statistika Moscow (2001)

11. Ge, P., Wang, J., Ren, P., Gao, H., Luo, Y.: A new improved forecasting method integrated fuzzy time series with the exponential smoothing method. Int. J. Environ. Pollut. **51**(3–4), 206–221 (2013)
12. CIF. http://irafm.osu.cz/cif/main.php. Accessed 22 June 2020
13. Afanasieva, T., Yarushkina, N., Zavarzin, D., Guskov, G., Romanov, A.: Time series forecasting using combination of exponential models and fuzzy techniques. In: Abraham, A., Kovalev, S., Tarassov, V., Snášel, V. (eds.) Proceedings of the First International Scientific Conference "Intelligent Information Technologies for Industry" (IITI'16). AISC, vol. 450, pp. 41–50. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-33609-1_4
14. Pedrycz, W., Chen, S.: Time series analysis, modeling and applications: a computational intelligence perspective (2013). https://doi.org/10.1007/978-3-642-33439-9
15. Yarushkina, N., Andreev, I., Moshkin, V., Moshkina, I.: Integration of fuzzy OWL ontologies and fuzzy time series in the determination of faulty technical units. In: Misra, S., et al. (eds.) ICCSA 2019. LNCS, vol. 11619, pp. 545–555. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24289-3_40
16. Rybina, G.V.: Modern expert systems: trends towards integration and hybridization, devices and systems. Handling Control Diagnostics **8**, 18–21 (2001)
17. Yarushev S., Averkin A. Time series analysis based on modular architectures of neural networks. In: Procedia Computer Science, 8th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2017 (8th Annual Meeting of the BICA Society), pp. 562–567 (2018)
18. Timina, I., Egov, E., Romanov, A.: Fuzzy models in forecasting time series of project activity metrics. Paper presented at the Journal of Physics: Conference Series, vol. 1096 (2018)
19. Abayomi-Alli, A., Abayomi-Alli, O., Vipperman, J., Odusami, M., Misra, S.: Multi-class classification of impulse and non-impulse sounds using deep convolutional neural network (DCNN). In: Misra, S., et al. (eds.) ICCSA 2019. LNCS, vol. 11623, pp. 359–371. Springer, Cham (2019)