Intelligent time series forecasting system

V V Borisov[1], **P I Komarov**[2] **and V S Luferov**[1]

[1] The Branch of National Research University «Moscow Power Engineering Institute» in Smolensk, Smolensk, Russia

[2] The branch of Financial University of the Russian Federation

Abstract. The developed intellectual system for forecasting time series based on artificial neural networks, adaptive neural-fuzzy models and models based on the decomposition of fuzzy time series is considered. The developed system is designed for modeling time series by various methods in order to select the best model in terms of accuracy for operation in forecasting tasks. The system provides ample opportunities and convenient interface for acquiring knowledge, skills in creating, teaching, comparing the results of applying intelligent models and methods for solving theoretical and practical problems of analyzing and forecasting time series in different subject areas.

1. Introduction

Currently, there is a large number of time-series (TS) forecasting systems [1-2], which, as a rule, are oriented toward the implementation of two classes of methods for forecasting TS: statistical and intellectual. In contrast to statistical methods, intelligent methods for predicting TS are characterized by the ability to take into account mutually influencing (in not only correlating) different-quality factors.

At the same time, for these models there is a problem of determining their type, structure. In addition, IT professionals, rather than domain specialists, who often do not have the required skills in the field of information systems design, carry out the creation of intelligent forecasting systems (IFS). This leads to a more complicated process and a justified increase in the development time of the IFS.

Currently, the most popular language tools for creating intelligent systems are R, Python programming languages using the capabilities of mathematical packages such as MatLab [3]. However, to forecast time series based on intelligent models, domain specialists need to know the relevant software tools.

2. Structure of intelligent time series forecasting system

Based on the above, it is possible to determine the following main characteristics, which must have intelligent time series forecasting system (ITSFS):

- an intuitive interface that allows domain specialists to create, configure, modify and use intelligent models quickly and easily to forecasting time-series;
- ease of specifying interdependencies between input and output parameters, taking into account the required «depth» of autoregression;

- automatic generation and verification of training samples for intelligent models of various types;
- flexible change in the structure and parameters of intellectual models in the process of forecasting TS.

An intelligent time series forecasting system is proposed, the structure of which is shown in Figure 1.

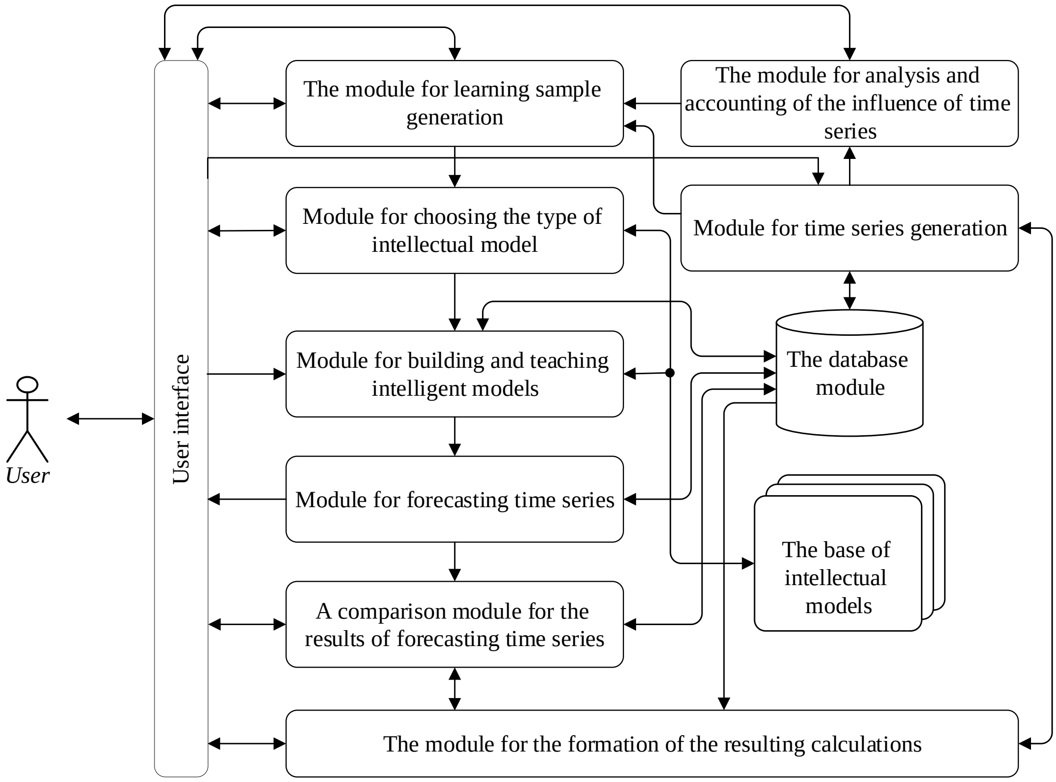

Figure 1. Structural diagram of the proposed ITSFS.

The database module accumulates data from external sources. External data sources can be enterprise database management systems, data tables, specialized data files, expert opinions, etc. The time series generation module converts unstructured data into time series with varying degrees of discretization. Time series can be changed at any stage of the work of the intellectual time series forecasting system If necessary, a time series can be specified by entering data through the user interface.

The module for analyzing and accounting for time series interactions allows the user to perform time series analysis, determine correlation, autocorrelation, etc. The analysis of time series allows specialists to substantiate the correctness of the construction of models, to find out trends and patterns that are subsequently accumulated and used to account for the relationship between input and output variables. When analyzing time series, the user interacts with the module through an intuitive interface.

When the module for forecasting time series predicts, the following steps are performed:

1. construction of linguistic variables based on the clustering algorithm;
2. formation of fuzzy rule base based on clustering algorithm;
3. training model based on retrospective data obtained from the fuzzy partitioning module, and the output of training results
4. forecasting electric load potential forecast horizon.

After the implementation of forecast verification data occurs, if there is reference data, and output the results.

From the set of time series using the module of training sample formation for time series, the relationships between output and input variables are formed and selected with allowance for autoregression:

$$P(t+1) = f\left(P(t),\ P(t-1), U(t+1),\ U(t), ...\right),$$

where $P(t+1)$ – forecasted value of time series; $P(t)$ – the previous value of the forecasted time series; $U(t+1),\ U(t)$ – time series that have a direct or indirect effect on the forecasted time series.

3. Libraries of intelligent time series forecasting system

The subject domain expert selects and configures the models presented in the model database that contains three types of intelligent models from the following libraries:

- FANN (Fast Artificial Neural Network);
- FL (FuzzyLogic);
- FTSD (Fuzzy Time Series Decomposition).

FANN (Fast Artificial Neural Network) is a library for constructing and using artificial neural networks for forecasting time series. This library provides an interface for tuning and subsequent use of artificial neural networks (ANN) such as multilayer perceptron. The library is available on the Internet at: http://leenissen.dk/fann/wp/.

This library has the following features:

- ANN training on back propagation of the error;
- transformation of the ANN structure in the learning process;
- comes with open source;
- library is cross-platform.

FL (FuzzyLogic) is a library for using fuzzy inference models based on the Mamdani and Sugeno algorithms for forecasting time series. The library was developed by one of the authors of this article and is available on the Internet at: https://github.com/Luferov/FuzzyLogic. Based on the Sugeno fuzzy inference algorithm, an adaptive network-based fuzzy inference system output is implemented. The model is trained by back propagation of the error [4-5].

The library has the following features:

- use of *Mamdani* and *Sugeno* fuzzy inference algorithms;
- implementation of the «mountain» clustering algorithm for model training;
- implementation of the algorithm for fuzzy expansion of time series for flexible separation of time series into trend and residual components, depending on the specified «step» of the fuzzy partition [6].

FTSD (Fuzzy Time Series Decomposition) is a library for building forecastion models based on fuzzy time series decomposition. The peculiarity of forecasting time series based on time series decomposition consists in isolating the trend component from the time series at each prediction step and in tracking the dynamics of the process in forecasting.

The module for building and teaching intelligent models initializes the intellectual model chosen by the expert (it`s type and structure). Based on the established relationship, the model is trained, its adequacy is estimated and the reliability of the time series forecasting.

After training, time series forecasting is performed using the module for forecasting time series. In the process of forecasting, the expert can perform structural and parametric adjustment of the model.

The results of the forecast, obtained with the use of different intellectual models, can be compared with each other and with reference values by means of the comparison module for the results of forecasting time series. Depending on the results of the comparison, the decision is made to choose the best or correct the models used, using the module for generating the resulting calculations.

To demonstrate the efficiency of the developed intelligent time series forecasting system, consider the operational and short-term time series forecasts.

As an operational forecast (for 24 hours), let's consider use of ITSFS developed on the example of comparison of the results of the electrical load of the Smolensk region in December 2017 with the use of intelligent models of various types: artificial neural network (from the FANN library), neuro-fuzzy model (from the FL library), models based on fuzzy time series decomposition (from the FTSD library). The graphs of the electric load forecasts obtained using these models are shown in figure 2. The operational projections are presented in table 1 and the comparative evaluation is presented in table 2.

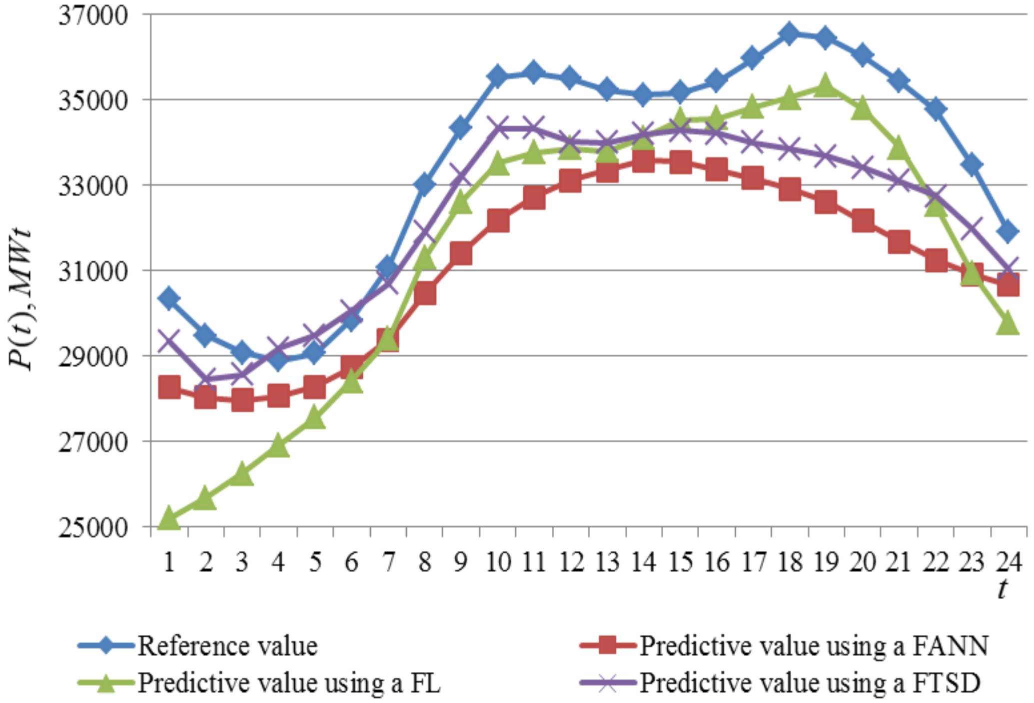

Figure 2. Graphs of operational forecasts based on intellectual models of various types.

Table 1. The result of the operational forecast using various smart models.

hour	Reference value, MWt	Predictive value using a FANN, MWt	Predictive value using a FL, MWt	Predictive value using a FTSD, MWt
1	30341	28267,54	25209,71	29354,52
2	29491	28034,06	25678,35	28447,74
3	29091	27964,68	26246,61	28571,51
4	28896	28073,92	26914,06	29191,89
5	29062	28279,21	27560,84	29490,97
6	29848	28741,88	28422,04	30050,22
7	31091	29385,47	29406,78	30707,12
8	33022	30462,73	31295,97	31900,28
9	34341	31412,77	32614,22	33225,01

10	35537	32178,21	33519,3	34323,55
11	35636	32716,33	33763,95	34335,99
12	35496	33109,36	33867,82	34017,17
13	35225	33342,39	33793,18	33992,74
14	35122	33575,44	34109,01	34192,62
15	35165	33556,03	34530,65	34280,33
16	35425	33365,72	34562,23	34209,4
17	35960	33171,02	34828,68	34005,97
18	36546	32909,38	35047,84	33843,16
19	36449	32612,72	35338,31	33675,62
20	36032	32161,64	34801,14	33419,44
21	35430	31681,77	33882,99	33099,67
22	34772	31244,32	32525,8	32749,12
23	33477	30908,3	30951,73	31976,78
24	31899	30658,4	29766,65	31027,17

Table 2. Comparison of the operational forecasting results.

Intellectual model	MAPE, %
Predictive value using a FANN	6.76
Predictive value using a FL	5.73
Predictive value using a FTSD	4.76

Table 3. Comparison of the short-term forecasting results.

Intellectual model	MAPE, %
Predictive value using a FANN	4.20
Predictive value using a FL	3.85
Predictive value using a FTSD	4.13

As a short-term forecast (for 31 days), let's consider the use of the developed ISP BP by comparing the results of the electrical load of the Smolensk region in December 2017 with the use of intelligent models of various types: artificial neural network (from the Fann library), neuro-fuzzy model (from the FL library), models based on fuzzy time series decomposition (from the ftsd library). The graphs of the electric load forecasts obtained using these models are shown in figure 3. The results of the comparative evaluation are presented in table 3.

The forecasting of time series is carried out on the basis of intellectual models according to the general algorithm. First, the final resultant sample is constructed, depending on the relationship established between the input and output time series by the expert. Then the data is normalized, i.e. all values of the initial data of the training sample are reduced to a single hypercube. This is important for comparing time series to bring them to a single scale. If normalization is not necessary, the training sample is applied to the model in a «raw» form. Then the validation of the intellectual model is carried out, the horizon of possible forecasting is specified and the time series in the training sample is compared. After that, the intellectual model is trained on a trained training sample. As a result, a forecasted series is formed, which can be presented either in the form of a report, or used when compared with the results of forecasting by other models.

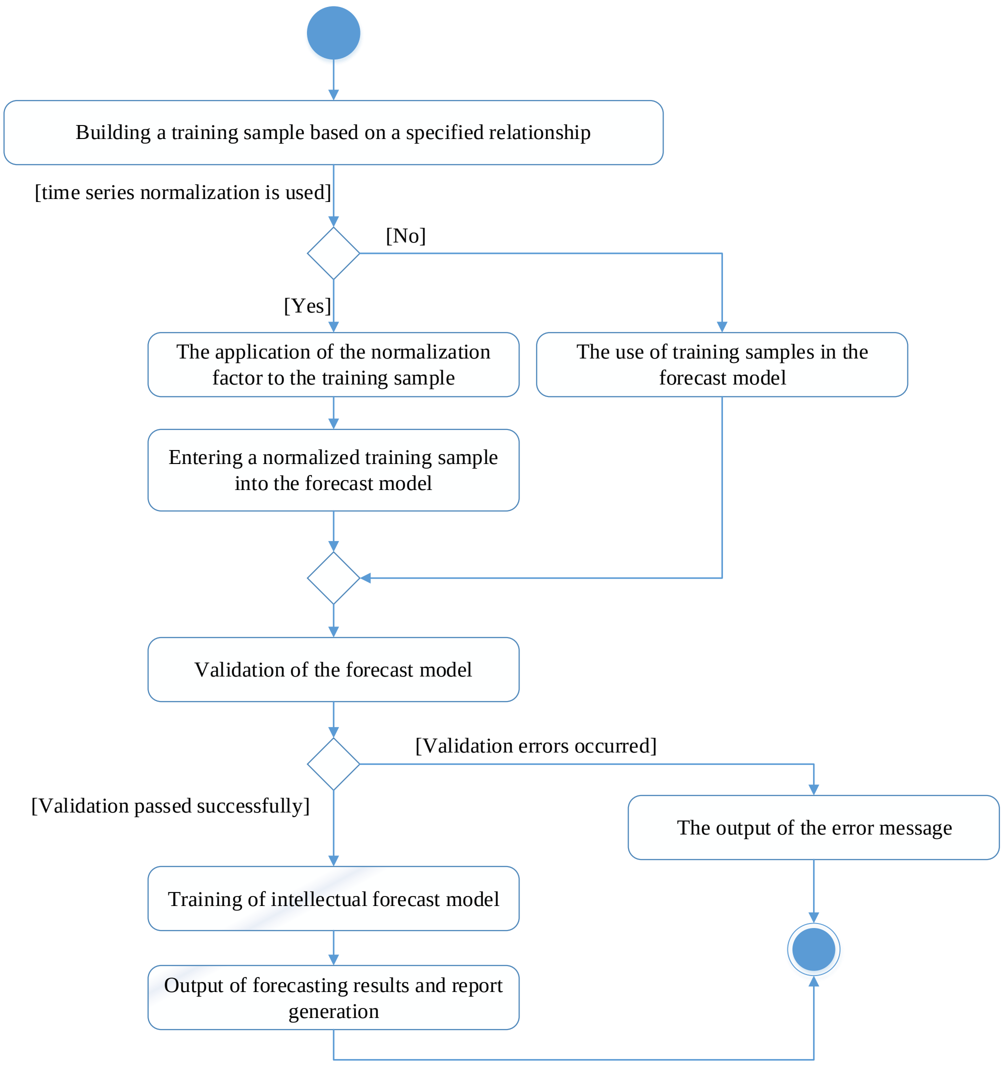

Figure 3. UML diagram of time series forecasting on the basis of mental models.

The developed ITSFS can be recommended as a teaching system for students in the field of «Informatics and Computer Science», and will also be useful for students of other areas and specialists engaged in the analysis and modeling of complex systems and processes under uncertainty, the creation of intelligent information systems and technology.

4. Conclusion

The paper presents the developed intellectual time-series forecasting system that provides ample opportunities and convenient interface for acquiring knowledge, skills and skills in creating, teaching, comparing the results of applying intelligent models and methods for solving theoretical and practical problems of time series analysis and forecasting in various subject areas. The developed system is designed for modeling time series by various methods in order to select the best model in terms of accuracy for operation in forecasting tasks.

5. References

[1] Averkin A N, Kosterev V V *Triangular norms in systems of artificial intelligence* // Izvestiya Akademii Nauk. Theory and control systems, 2000. №5. – PP. 106–109. (in Russian)

[2] Halliwell J, Keppens J, Shen Q *Linguistic Bayesian networks for reasoning with subjective probabilities in forensic statistics*// Proc. of the 5th International Conference on AI and Law, 2003. – PP. 42–50

[3] Leonenkov A V *Fuzzy modeling in MATLAB and fuzzyTECH*. – St.-Petersburg: BHV–Petersburg, 2003. –736 p. (in Russian).

[4] Borisov V V, Kruglov V V, Fedulov A S *Fuzzy models and networks*. 2-nd ed. stereotype. – Moscow: Hot line-Telecom, 2012. –284 p. (in Russian)

[5] Borisov V V, Fedulov A S, Zernov M M *Basics of hybridization of fuzzy models*. Series “Fundamentals of fuzzy mathematics”. Book 9. Textbook for high schools. – Moscow: Hot Line–Telecom, 2017. –100 p.

[6] Afanasyeva T V, Namestnikov A M, Perfilieva I G, Romanov A A and Yarushkina N G *Forecasting time series: fuzzy models; under the scientific*. Ed. N.G. Yarushkina. – Ulyanovsk: UlSTU Publishing House, 2014. – 145 p.

Acknowledgment

The work was supported by the Russian Science Foundation, project No. 16-19-10568.