



DEVELOPMENT OF AN ANFIS NEURO-FUZZY MODEL FOR FORECASTING WEAKLY STRUCTURED PROCESSES

Giyos Pulatov

“TIAME” National Research University, Tashkent, Uzbekistan

giyospulatov1987@gmail.com

Abstract

This article examines the problem of developing an ANFIS neuro-fuzzy model for forecasting weakly structured processes. The relevance of this study lies in the fact that weakly structured processes are typically characterized by uncertainty, nonlinearity, and multivariate dependence, which makes accurate forecasting by means of conventional statistical methods rather difficult. For this reason, an ANFIS model combining the advantages of artificial neural networks and fuzzy logic was proposed. The experimental data were formed as a time-series dataset based on meteorological, geophysical, and target indicators. The input space of the model was constructed using calendar features, a trend component, and lagged variables. The most informative features were selected using the mutual information regression method. The proposed model was built on the basis of a first-order Sugeno-type TSK-ANFIS architecture. Gaussian membership functions were used, rule centers were determined by the K-Means algorithm, and the output parameters were estimated using a ridge-regularized linear solution. According to the experimental results, the model achieved $MAE = 0.388$ and $RMSE = 0.491$ on the test set. The obtained results confirm that the ANFIS model is capable of adequately capturing the general dynamics of weakly structured processes and can be applied in practical forecasting tasks.

Keywords: Weakly structured processes, forecasting, ANFIS, neuro-fuzzy model, Sugeno model, time series, fuzzy rules, Gaussian membership function.



Introduction

A large number of processes observed in modern control systems, medicine, economics, and technogenic environments belong to the class of weakly structured processes. The main feature of such processes is that they are not governed by strictly deterministic regularities; rather, they depend on multivariate, uncertain, and partially stochastic relationships. Therefore, evaluating their dynamics in advance and obtaining reliable forecasts is a complex scientific and practical problem.

Classical regression, smoothing methods, and statistical time-series models can be applied to forecast weakly structured processes. However, such methods are often unable to fully represent nonlinear relationships, the influence of latent factors, and uncertain interactions among variables. For this reason, interest in machine learning and intelligent analytical methods has increased substantially in recent years.

Artificial neural networks are effective in learning complex nonlinear relationships, while fuzzy logic makes it possible to represent expert knowledge through linguistic rules. The integration of these two approaches in the form of the ANFIS model is considered a promising tool for describing uncertain and weakly structured processes. On the one hand, the ANFIS model provides adaptive learning from data; on the other hand, its fuzzy rule base makes the decision-making process more interpretable.

The purpose of this article is to develop an ANFIS neuro-fuzzy model for forecasting weakly structured processes, to define its architecture, and to evaluate its forecasting accuracy on experimental data.

Problem Statement

Within the scope of this study, the problem of forecasting the target indicator at the next time point using multivariate time-ordered data is considered. The input includes meteorological, geophysical, and time-dependent features, while the output is the value of the target indicator at the next time interval.

In formal terms, the problem can be expressed as follows. Let

$$X_t = \{x_1, x_2, \dots, x_n\}$$



be the input feature vector at time t . The forecasting objective is to determine the output value at time $t+1$:

$$\hat{y}_{t+1} = F(X_t, X_{t-1}, X_{t-2}, \dots, X_{t-k}),$$

where F is the nonlinear mapping learned by the ANFIS model, and k is the lag depth.

In this study, the ANFIS model was developed specifically to represent such multivariate and uncertain relationships.

Methods

Data Description. The experiments were conducted using a dataset covering the period from January 1, 2000 to January 18, 2026. The initial number of observations was 9,515. After applying temporal shifting and generating lagged variables, the effective number of observations used in the analysis was 9,486.

The dataset included the following main variables:

- maximum temperature (temp_max_c);
- minimum temperature (temp_min_c);
- mean temperature (temp_mean_c);
- atmospheric pressure (pressure_msl_mean_hpa);
- mean wind speed (wind_speed_mean_kmh);
- maximum wind speed (wind_speed_max_kmh);
- mean humidity (humidity_mean_percent);
- geomagnetic indicators (kp_mean, kp_max);
- target variable (CVD).

The target variable was defined as the value of the indicator at the next time point. To improve model stability, the output variable was transformed into logarithmic form:

$$y_t = \log(1 + CVD_{t+1})$$

Feature Construction. Since the time factor plays an important role in weakly structured processes, additional features were derived from the raw data. In particular, calendar and trend components were introduced:

- day of the week;
- month;



-
- day;
 - time trend ttt.

To account for periodicity, the day-of-week and month features were cyclically encoded using sine and cosine transformations:

$$dow_sin = \sin\left(\frac{2\pi * dow}{7}\right), dow_cos = \cos\left(\frac{2\pi * dow}{7}\right),$$

$$month_sin = \sin\left(\frac{2\pi * month}{7}\right), month_cos = \cos\left(\frac{2\pi * month}{7}\right),$$

In addition, lagged features were generated to capture inertia and delayed effects in the time series. Lag windows of 1, 2, 3, 4, 5, 6, 7, 14, 21, and 28 time units were constructed. As a result, a total of 118 input features were formed for the model.

Data Preparation and Partitioning. The data were divided into three subsets while preserving the temporal order:

- training set — 70% (6,640 observations);
- validation set — 15% (1,423 observations);
- test set — 15% (1,423 observations).

MinMax normalization was applied to the input features in order to bring them into a common range. This made it possible to stabilize the membership functions in the ANFIS model and balance the influence of features measured on different scales.

Feature Selection. The most informative variables were selected using the mutual information regression method. This approach makes it possible to estimate the degree of dependence between each input feature and the target variable. The trend feature ttt was included in the model by default. As a result, eight core features were selected:

- t;
- CVD;
- CVD_lag1;
- month;
- CVD_lag2;
- CVD_lag3;
- CVD_lag4;



- CVD_lag5.

These features were used to form the input space of the ANFIS model.

ANFIS Model Architecture. The proposed model was developed on the basis of a first-order Sugeno-type TSK-ANFIS architecture. This model combines a fuzzy rule base with adaptive parameters to represent the nonlinear relationship between inputs and output.

The general operating principle of the ANFIS model includes the following stages:

1. fuzzification of the input values;
2. computation of the firing strength of each rule;
3. normalization of rule strengths;
4. estimation of the consequent part in the form of a linear model;
5. aggregation of all rule outputs to produce the final prediction.

Gaussian membership functions were used in the antecedent part:

$$\mu(x) = \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right)$$

where c is the center and σ is the spread parameter.

The fuzzy rule centers were initialized using the K-Means clustering algorithm. The consequent parameters were estimated through a ridge-regularized linear solution. The number of rules was tested for 4, 6, and 8 alternatives. The regularization parameter was examined for the values 10^{-4} , 10^{-3} , and 10^{-2} . The best-performing model was selected according to the MAE obtained on the validation set.

Evaluation Metrics. The model performance was evaluated using mean absolute error and root mean squared error:

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$



Experimental Results and Discussion. The proposed ANFIS model was evaluated on the test set. The obtained results demonstrated that the model provides practically acceptable accuracy in forecasting weakly structured processes.

Table 1. Test results of the ANFIS model

Model name	MAE	RMSE
ANFIS	0.388	0.491

As shown in Table 1, the ANFIS model achieved a mean absolute error of 0.388 and a root mean squared error of 0.491. These results indicate that the predicted values are sufficiently close to the actual observations.

Figure 1 presents a comparison between the real and predicted values over time. The graph shows that the model is able to capture the main trend and the general dynamics of the process with a reasonable degree of accuracy. Although some deviations are observed at individual points, the predicted curve generally follows the real series. This suggests that the ANFIS model can also partially capture local variations in the time series.

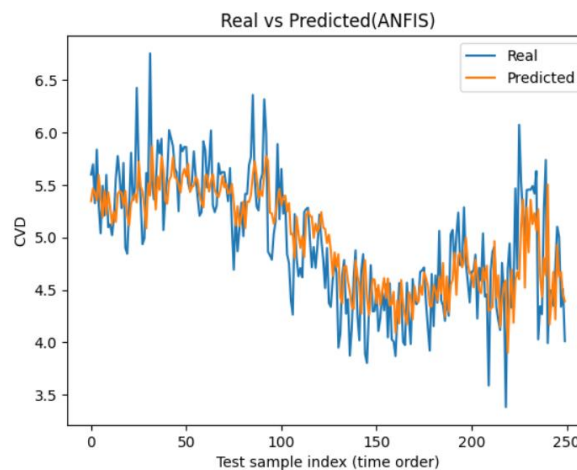


Figure 1. Real vs Predicted (ANFIS)

Figure 2 shows the scatter plot of actual and predicted values for the test set. The concentration of points around the diagonal line indicates a positive agreement between the model predictions and the real values. At the same time, a wider

spread in the regions of small and large values suggests that the prediction error may increase for extreme observations.

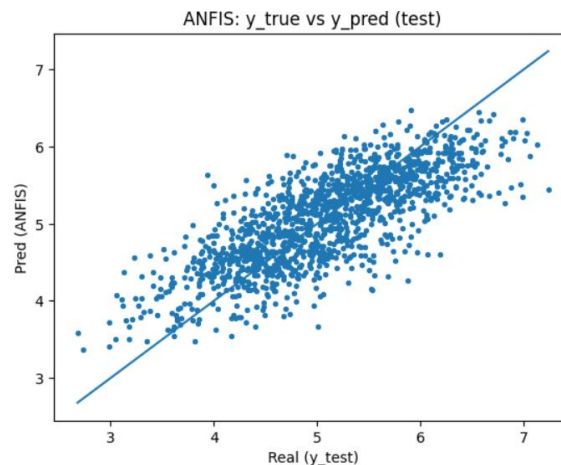


Figure 2. ANFIS: y_true vs y_pred (test)

Figure 3 illustrates the distribution of prediction errors. The histogram shows that most of the errors are concentrated around zero. This indicates the absence of a strong systematic bias in the model. However, the presence of tails on both sides suggests that larger deviations may occur for some individual observations. Such deviations may be explained by the uncertainty and abrupt local fluctuations inherent in weakly structured processes.

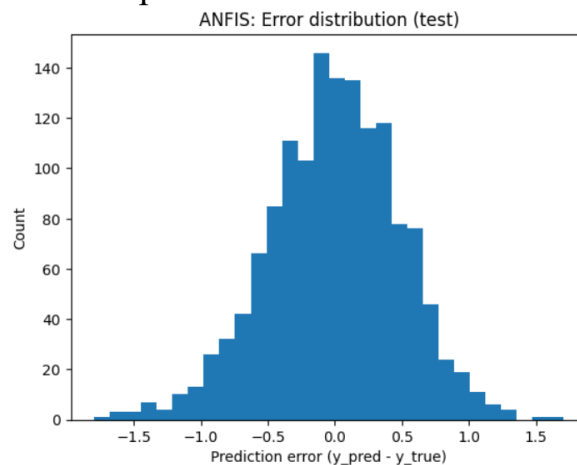


Figure 3. ANFIS: Error distribution (test)

The obtained results indicate that the main advantage of the ANFIS model lies in its ability to flexibly represent uncertain and nonlinear relationships. In particular, the combined use of trend, periodicity, and lag effects contributed to improving forecasting accuracy. The fact that the current and lagged values of the target



variable, together with the temporal component, were identified as the most informative features suggests that historical dynamics and time-related effects play a major role in weakly structured processes.

Discussion

The proposed ANFIS model provides better opportunities for representing complex and uncertain relationships than classical linear approaches. In this study, the combination of fuzzy rules and data-driven adaptive parameters made it possible to capture the internal dynamics of weakly structured processes. In particular, the use of Gaussian membership functions proved convenient for describing smooth transitions within the input space.

At the same time, the model has several limitations. First, increasing the number of rules leads to greater computational complexity. Second, forecasting accuracy may decline for outliers and abrupt changes. Third, model performance depends strongly on the correct selection of input features. Therefore, future research should focus on improving the ANFIS model through hybrid feature selection, adaptive rule generation, and ensemble-based integration with other intelligent algorithms.

Conclusion

This article developed an ANFIS neuro-fuzzy model for forecasting weakly structured processes. The model was built on the basis of a first-order Sugeno-type TSK-ANFIS architecture and used trend, calendar, and lagged features derived from time-series data. Mutual information regression was applied for input feature selection, while the K-Means algorithm was used to initialize rule centers.

The experimental results confirmed the practical effectiveness of the model. On the test set, the ANFIS model achieved $MAE = 0.388$ and $RMSE = 0.491$. The comparison of real and predicted values demonstrated that the model adequately captures the general dynamics of weakly structured processes. The concentration of prediction errors around zero also indicated the absence of a strong systematic bias.



Thus, the proposed ANFIS neuro-fuzzy model can be considered a reasonable approach for forecasting weakly structured processes. Future work may include the incorporation of additional factors, adaptive rule-base formation, and integration with other intelligent algorithms in an ensemble framework.

References

1. Lokman, A.; Wan Ismail, W.Z.; Ab Aziz, N.A.; Ghazali, A.K. Water Quality Index (WQI) Forecasting and Analysis Based on Neuro-Fuzzy and Statistical Methods. *Applied Sciences* 2025, 15(17), 9364. doi:10.3390/app15179364.
2. Lin, S.-S.; Song, J.-H.; Zhu, K.-Y.; Liu, Y.-C.; Chang, H.-C. Applying Adaptive Neuro-Fuzzy Inference System to Improve Typhoon Intensity Forecast in the Northwest Pacific. *Water* 2023, 15(15), 2855. doi:10.3390/w15152855.
3. Zullo Godinho, E.; Dartiere Zulian Fermino, C.; Marques Barreiros, R. Kinetic Modeling and Neuro-Fuzzy Application in Ethanol Production. *Brazilian Journal of Biosystems Engineering* 2025, 19, 1264. doi:10.18011/bioeng.2025.v19.1264.
4. AlRassas, A.M.; Al-qaness, M.A.A.; Ewees, A.A.; Ren, S.; Sun, R.; Pan, L.; Abd Elaziz, M. Advance Artificial Time Series Forecasting Model for Oil Production Using Neuro Fuzzy-Based Slime Mould Algorithm. *Journal of Petroleum Exploration and Production Technology* 2022, 12, 383–395. doi:10.1007/s13202-021-01405-w.
5. Hakimi, A.; Farjami, H. Enhanced Significant Wave Height Prediction in the Southern Ocean Using an ANFIS Model Optimized with Subtractive Clustering. *Scientific Reports* 2025, 15, 35072. doi:10.1038/s41598-025-18140-5.
6. Thejasree, P.; Manikandan, N.; Rajesh, N.; Lokanadham, R.; Krishnamachary, P.C.; Shubham, K.; Olaiya, B.C. Application of Taguchi Grey-Based ANFIS Model for Prediction of Process Parameters in Fused Deposition Modelling of PETG. *Scientific Reports* 2025, 15, 37398. doi:10.1038/s41598-025-21248-3.