ГИДРОЛОГИЯ И КЛИМАТОЛОГИЯ

HYDROLOGY AND CLIMATOLOGY

UDC 519.7:561.58 https://doi.org/10.5281/zenodo.10395740

Arushanov M.L., Umerov Kh.U.

Research Hydrometeorological Institute, Tashkent, Uzbekistan

PREDICTION OF THE TIME SERIES OF TEMPERATURE BY HYBRID MODELS USING WAVELET DECISION AND NEURAL NETWORK

Abstract. *Starting with the works of A. Haar, A. Grosman and J. Morlet [7,8], the wavelet transform, which converts a signal from a time representation to a time-frequency representation, has found wide application in various scientific fields for the purpose of filtering and preprocessing data, analyzing the state and predicting the processes under study. However, in relation to weather forecasting problems based on the use of neural network technologies in combination with wavelet transformation, targeted research has not been carried out.*

This study, in a sense, aims to fill this gap by analyzing the impact of time series preprocessing based on wavelet decomposition combined with neural networks to predict reconstructed temperature series. Approaches using several models are considered and their predictive capabilities in a hybrid combination are analyzed. The wavelet transform is used as a tool for suppressing the noise component of a temperature time series, and the forecast model is built on the basis of neural network technologies and regression methods for forecasting time series.

Key words*: time series, reconstructed temperatures, wavelet decomposition, neural networks, forecasting, hybrid models.*

Арушанов М.Л., Умеров Х.У.

Научно-исследовательский гидрометеорологический институт, Ташкент, Узбекистан

ПРОГНОЗ ВРЕМЕННОГО РЯДА ТЕМПЕРАТУРЫ ПО ГИБРИДНЫМ МОДЕЛЯМ С ИСПОЛЬЗОВАНИЕМ ВЕЙВЛЕТ-РЕШЕНИЯ И НЕЙРОННОЙ СЕТИ

Аннотация. *Начиная с работ А. Хаара, А. Гросмана и Ж. Морле [7,8] вейвлет преобразование, переводящее сигнал из временного представления в частотно-временное, нашло широкое применение в различных научных направлениях в целях фильтрации и предварительной обработки данных, анализа состояния и прогнозирования исследуемых процессов. Однако, применительно к задачам прогноза погоды на основе использования технологий нейронных сетей в сочетании с вейвлет преобразованием целенаправленных исследований не проводилось.*

Данное исследование в определённом смысле направлено на заполнение этого пробела, в котором анализируется влияние предварительной обработки временных рядов на основе вейвлет-разложения в сочетании с нейронными сетями для прогнозирования реконструированных температурных рядов. Рассмотрены подходы с использованием нескольких моделей и проанализированы их прогностические возможности в гибридной комбинации. В качестве инструмента подавления шумовой составляющей температурного временного ряда используется вейвлет-преобразование, а прогностическая модель строится на основе нейросетевых технологий и регрессионных методов прогнозирования временных рядов.

Ключевые слова*: временной ряд, реконструированные температуры, вейвлет-разложение, нейронные сети, прогнозирование, гибридные модели.*

Introduction and problem statement. One of the completely unresolved problems to this day remains weather forecasting, the solution of which determines the development and adaptation of the production and economic sector, which depends on weather conditions.

Despite the huge leap in the development of computing and meteorological technologies in recent decades, it is impossible to make an ideal forecast. This is due, on the one hand, to the limited knowledge of mankind, on the other hand, to the limited resolution of weather observation instruments, as well as the inability to take into account all the factors occurring in the atmosphere. To a greater extent, the above applies to long-term weather forecasts. In the short-term forecast of meteorological quantities, good results have been achieved - correct forecasts are more than 90%. The fact is that the method of short-term forecasts with good reason corresponds to the adiabatic approximation, when sources and sinks of energy can be neglected, i.e. atmospheric processes can be considered with sufficient accuracy as the dynamics of a closed system. Completely different physics is inherent in the long-term forecast methodology, where the main regulators of atmospheric processes are sources and sinks of energy, i.e. the system must necessarily be considered as open thermodynamic. As a result, many factors come into force that cannot be taken into account together. To solve the problem, various mathematical methods are used here that perform the functions of selecting the most informative predictors, various methods of filtering and smoothing the original time series, allowing the most optimal approach to the real physical picture of the dynamics of atmospheric processes.

Time series forecasting is one of the main areas of meteorology. There are many dynamic-stochastic forecasting models, from basic ones (simple moving average, linear regression, etc.) to more advanced models such as autoregressive integrated moving average (ARIMA) and neural networks.

Dynamic-stochastic forecast models are presented in the form of a deterministic component and random noise ε_t

$$y_t+1=A(y_t) + \varepsilon_t, (t=1, 2, \ldots, N)$$

where A is some predictive operator. Thus, the main task of modeling comes down to minimizing the noise component using various expansions of the original function, for example, Fourier and wavelet transforms [3,4,7]. The latter has a significant advantage over the Fourier transform, since the wavelet transform allows one to judge not only the frequency spectrum of the signal, but also how stable in time the period of oscillation of a particular harmonic is at a fixed frequency.

The wavelet transform has become very popular in many fields of science, including physical and mathematical, economic and social [2,6,16,12,13]. However, the wavelet transform and neural networks have not yet found wide application in meteorology for solving forecasting problems.

This article makes an attempt to fill this gap to some extent by considering hybrid models for forecasting average annual temperatures based on the wavelet transform, neural networks [3] and classical methods [15].

Study of the problem. The problem of long-term forecast of meteorological quantities remains to this day, in fact, the main problem of meteorology. This position of long-term forecast methods is, first of all, due to the extremely complex physical mechanism of long-term weather formation and, as a consequence, the impossibility of taking into account all factors. On the other hand, hydrometeorology, as one of the branches of applied physics, is conceptually entirely dependent on the ideology that has developed as a result of theoretical and experimental research in fundamental physics. The reverse influence is almost absent; as a result, geophysics researchers consider their range of problems only as particular applications of long-established (in the case of hydrometeorology - more than a hundred years) physical theories in a specific natural setting. All geophysical specifics come down to the complexity of natural systems, incomplete information about them and, as a rule, to the impossibility of an active experiment, i.e. in replacing experiment, in the physical sense of the word, with passive observation. The

lack of feedback leads to great inertia in meteorology in the perception of a range of new general concepts of fundamental physics. Of course, not every new general physical idea must necessarily have any impact on meteorology. However, it is difficult to give a universal, negative answer without carrying out work in each specific case to verify what some new (usually not generally accepted, and sometimes unsatisfactorily formalized) physical idea gives in specific, geophysical conditions.

Materials and methods. The work used dendrological data of reconstructed temperature on the territory of the Yamal Peninsula [7] for 8 millennia (taken 1220 years - 800-2020). On fig. 1 shows the time course of the reconstructed air temperature on the territory of the Yamal Peninsula and its trend over a 1220-year period.

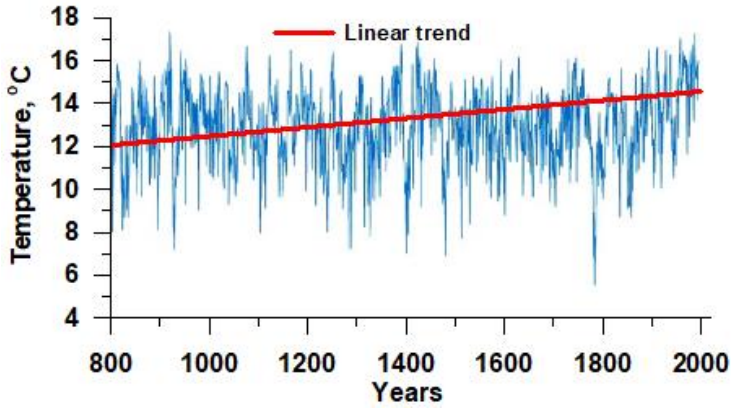


Fig. 1. Time variation of the reconstructed air temperature on the territory of the Yamal Peninsula based on fluctuations in the width of annual tree rings

To expand the original series using the wavelet transform, the following formula is used [1]:

$$W(f(a,b)) = \langle f, \psi_{a.b} \rangle = \frac{1}{|a|^{1/2}} \int_{-\infty}^{\infty} f(t) \cdot \psi\left(\frac{t-b}{a}\right) \cdot dt \qquad (1)$$

where is the mother wavelet function; a is the scaling parameter, b is the shift parameter, is the normalization parameter of the wavelet function, providing. The direct wavelet-transform (1) can be considered as the expansion of the signal over all possible shifts and stretching/compression of the signal f(t). Transformation (1) when specifying a vector a in a certain range (amin ÷ amax) makes it possible to obtain from the original function f(t) a set of functions *W*[*f*(a, b)], varying with different frequencies (from high-frequency to low-frequency). With this signal processing approach f(t), characteristic details that are hidden at one resolution can be detected at another.

Using multi-resolution analysis, both the overall structure and the details of the fine structure of the series can be explored. Denoising with different resolutions using wavelet transform allows you to get a smoother series without changing its basic structure. This turns out to be useful in problems of predictive models [10, 12].

Conceptually, the wavelet coefficients describe the features of the series, and their small values, as a rule, represent the noise component, which is easily eliminated by discarding small coefficients by a given threshold value. This is the advantage of the wavelet transform, which in its content is similar to filtering the original series using expansions in the system of natural orthogonal components [14].

In meteorology, as a rule, data are expressed as discrete sequences, for the decomposition of which the discrete wavelet transform (DWT) is used [3].

The DWT process can be performed by passing the sequence through half-band digital low and high pass filters [10]. In this case, the frequency resolution is doubled, because only half of the original band exists in the filtered sequences. After passing the sequences through

the filters, half of the samples can be eliminated by the Nyquist criterion (i.e., half of the samples are redundant). Therefore, after sub-sampling by two, the temporal resolution is halved. Reconstruction of the original series f(t) from the calculated scaling (approximating) informative coefficients is performed by inverse transformation [1]:

$$f(t)=\frac{1}{\sqrt{M}}\sum_{k}W_{\varphi}(j_0,k)\cdot\varphi_{j_0,k}(t)+\frac{1}{\sqrt{M}}\sum_{j=j_0}^{\infty}\sum_{k}W_{\psi}(j,k)\cdot\psi_{j,k}(t) \quad (2)$$

Let us introduce the notation: a1, a2, …, am – series approximations at the levels (1 2, …, *m*); d_1, d_2, …. d_m – details of the distribution of the series at the corresponding levels. Then the initial series f(t) in formula (2) at the zero level with some accuracy (the accuracy of a given number of levels of detail) can be represented as:

$$f = a_m+d_m+d_{m-1}+\ldots+d_1. \quad (3)$$

In (3), a are low-frequency functions, and the details of d_i with small values of i are very small, like high-frequency functions.

In formula (3), the approximations a_i and details di tend to const as the level of approximation i increases. In fact, a_i and d_i are a time series of varying degrees of smoothness. Note that due to significant distortions at the boundaries of the wavelet transform, boundary values should be excluded from the decomposition [2].

The above procedure can be repeated for further decomposition. Figure 2 illustrates this process for a 3-level decomposition, where f is the original sequence, D1 is the discrete wavelet coefficients (DWC) of the first level, D2 is the DWC of the second level, D3 and A3 are the DWC of the third level [9]. Note that the original sequence f can be synthesized by sequences of low and high frequencies, i.e. $f = a_3+d_3+d_2+d_1$.

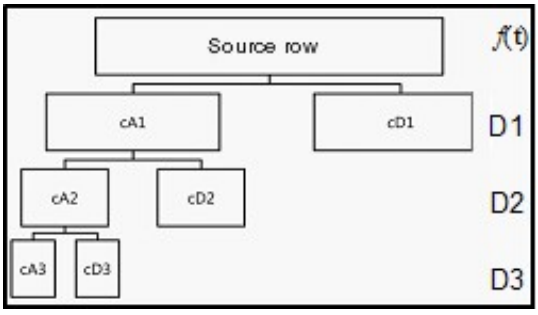


Fig 2. The structure of the wavelet decomposition

1. Apply a continuous wavelet transform to the original series.

2. We use the discrete wavelet decomposition (DWD) of a series of reconstructed average annual July temperatures on the Yamal Peninsula into a series of approximations (approximations) and series of component details: one approximation component (approximator) and five detailed component. Approximation component can be considered as a trend of the original series. Detailed components can represent the noisiness of the original series.

3. Let's choose forecasting methods for the approximator and detailing series. Let us determine the structural parameters of the selected models.

4. Apply forecasting methods for each component.

5. Let's restore the predicted series and compare the results. Subsequently, as expected, it can be concluded whether the detail components represent noise or are an information component.

The most common models in time series forecasting, especially in parametric estimation, are autoregressive integrated moving averages (ARIMA) and generalized

autoregressive conditional heteroscedasticity models (GARCH). GARCH models are based on the idea of inconsistent variance in a common time series [1].

Spline models [18] are widely used in the interpolation of various space-time series, and using the extrapolation strategy can also be used in forecasting problems.

We also note that recently machine learning methods have gained popularity, such as artificial neural networks (ANN), which are used in many predictive studies [2, 9,16], since the versatility of these models allows them to be applied to any geophysical processes, in particular, atmospheric ones.

Based on the foregoing, the following forecast models are considered:

1. GARCH (Generalized AutoRegressive Conditional Heteroscedasticity) – generalized autoregressive conditional heteroscedasticity models [9] (conditional variance model with bias (Gaussian distribution)) – time series $u(t)$ with conditional variance function

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i u_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 ,$$

where p is the order of the GARCH terms of the conditional variance (past conditional deviations); q is the order of ARCH members u^2 (past values of the time series). The choice of ARCH and GARCH lags was carried out experimentally.

2. A spline model based on cubic spline interpolation with a strategy for extrapolating different cubic polynomials. On each segment, the function is approximated by a cubic polynomial

$$f(x) = a(x-x_1)^3 + b(x-x_1)^2 + c(x-x_1) + d$$

для коэффициентов a, b, c, d на интервале $[x_1, x_2]$.

For a regular grid, the nodes of the extrapolation algorithm are represented as a sequence of expansions in point coordinates, and the expansion coefficients are presented analytically. As a result, the ordinate of the predicted point does not depend on the grid spacing, which is essential for estimating the nearest neighbor following in a series of regular observations, when the main thing is not the size of the interval between measurements, but its invariability. The simulation was carried out using the Matlab function interp1 (X_1, Y_1, dTest, 'spline', 'extrap') [5].

3. ANN (Artificial Neural Networks) – artificial neural networks with time delay [13] (Fig. 3).

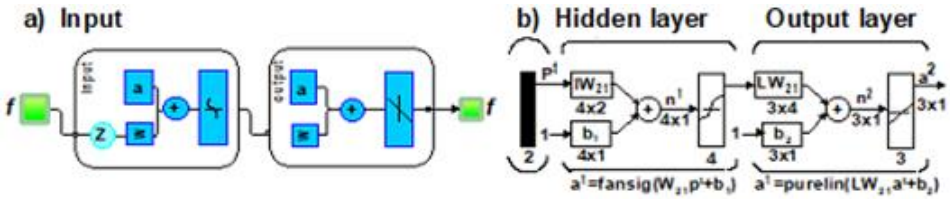


Fig. 3. Feed-forward neural network (ANN): *a*) – graphic and *b*) – block diagram

4. NarNET – Nonlinear autoregressive neural network [13] (Fig. 4.)

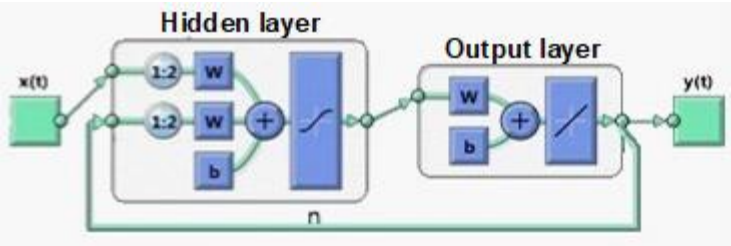


Fig. 4. Structure of a nonlinear autoregressive neural network

Using the closeloop(net) command, the circuit of this network was closed to obtain a recurrent network – generator. The delay step and the number of neurons in the hidden layer were selected experimentally and amounted to 2/14 and 3/12 for ANN and NarNET, respectively.

The forecasting accuracy was assessed by the root mean square error:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \quad , \qquad (4)$$

where $\hat{Y} = \{\hat{Y}_i\}$ – forecast vector, which is generated from a sample of n data points (training sample) for all variables; – vector of observed values of the predicted variable (independent sample).

As part of a preliminary analysis using the Morle wavelet, the amplitude wavelet function (AWF) and the scalogram of the temperature series for the period from 881-2019 were obtained (fig. 5). The AMF displays the details of the frequency image of the temperature series: in the lower part, high-frequency components with a period of 50 years are clearly visible, and in the middle and upper parts, stable oscillations at frequencies corresponding to periods of 480 and 800 years, clearly manifested in the integrated spectrum – scalogram.

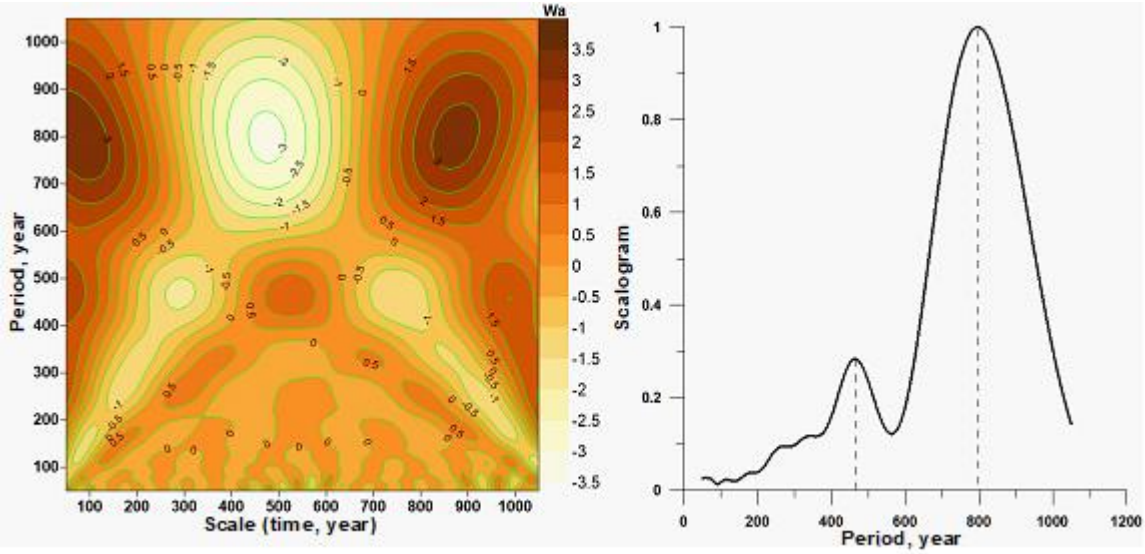


Fig. 5. Wavelet amplitude function Wa and air temperature scalogram

Based on the analysis of the obtained results and taking into account the effect of retraining of neural networks (at the level of 600 years, retraining was clearly manifested), a study period of 320 years was chosen. Of these, a series of 300 years constituted a training sample, a series of 20 years constituted an independent one.

Results and its discussion. *Numerical experiments.* All calculations were performed in the MATLAB environment. To calculate discrete transformations, wavelet functions of the Daubechies family of orders “db4”, “db26” and “db44” are usually used [19]. We have chosen "db44" as the basic wavelet. Figure 6 shows a five-level wavelet decomposition of the series under study for 300 years for the period 1800-1999. using the Daubechies wavelet of the 44th order (filter order N=44).

The simple structure of the approximator A shows the temperature trend, i.e. the big picture, the so-called "forest". When considering "trees", i.e. detailed, at the level of d5, for example, one can see that periods of relative temperature stabilization on a scale of 10-60 years (1710-1760) and 150-220 years (1856-1920) were replaced by significant differences in other periods. If we consider d4, then the time resolution will increase, and we will be able to reflect the periods of stationarity and significant fluctuations in more detail.

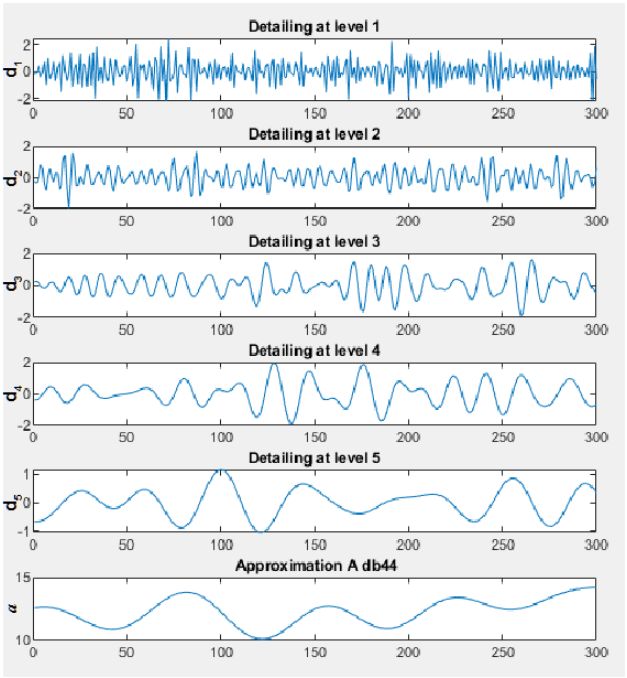


Fig. 6. Wavelet decomposition of the original series into 5 levels

The wavelet reconstruction of the studied series is shown in fig. 7. The root-mean-square error of the reconstruction was $\sigma = 8.21 \cdot 10^{-6}$, which justifies the use of wavelet analysis to solve the problem.

Also in fig. 7 shows the plot of the approximator (low-frequency part of the wavelet decomposition), which clearly describes the main trend in the temperature time series.

The methodology of hybrid models involves training several models on the same time series. Next, an assessment is made of the forecasts obtained by different models, which can be averaged.

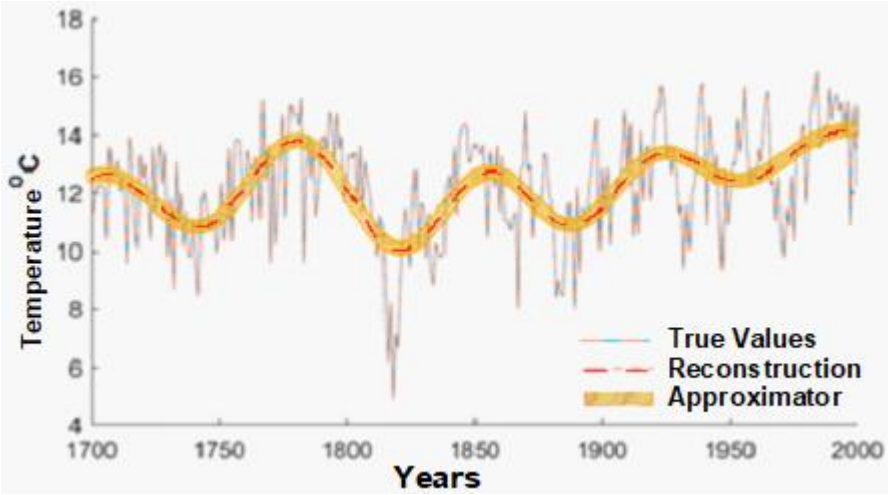


Fig. 7. Wavelet reconstruction of the temperature series

Fig. 8 shows the results of the prediction of the approximator A by various models. The results confirmed the choice of models for predicting the coefficients of the wavelet decomposition: spline and ANN.

The resulting series of coefficients a5, d5, d4, d3, d2 and d1 were predicted for 20 steps (years). To predict a5, an ANN model with two delay steps and 14 steps of a hidden layer neural

network was used. To predict d4, d3, d2 and d1, a spline model with an extrapolation strategy – interp1(X1, Y1, dTest,'spline','extrap') was used. Fig. 9 shows the results of the forecast.

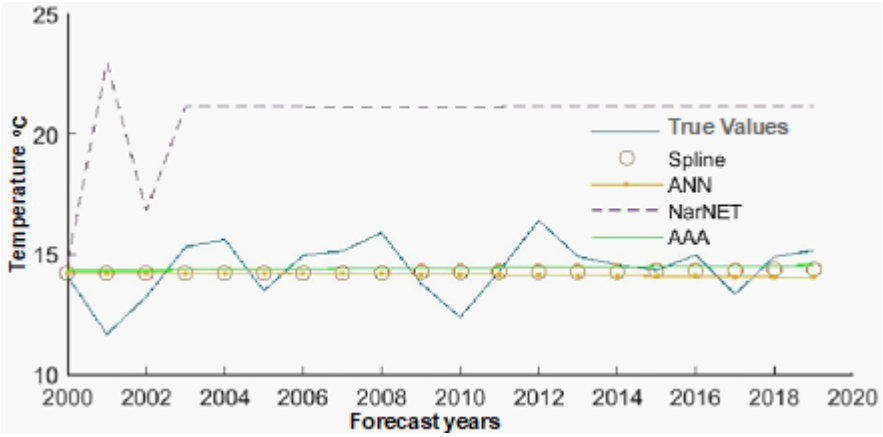


Fig. 8. Approximator prediction by various models

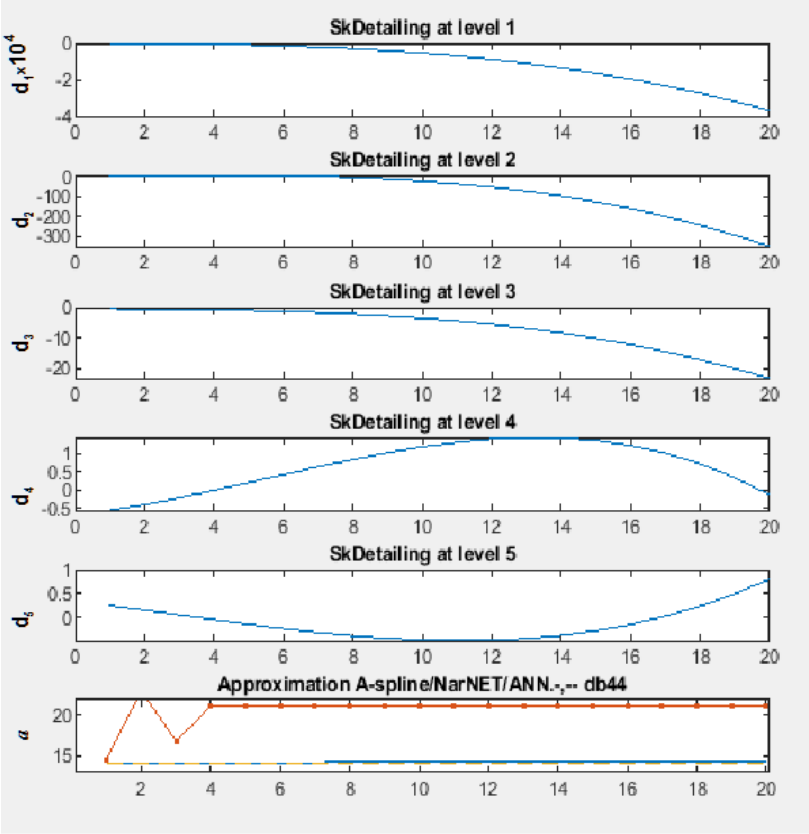


Fig. 9. Forecast of wavelet expansion coefficients for 20 years

Next, the signal is converted back to the time domain by performing an inverse transform (2). One of the disadvantages is that signal decomposition by transformation leads to overlaps at the edges of the corresponding time-frequency series and is enhanced by the choice of decomposition level [18]. When evaluating the obtained detail coefficients at the stage of forecast reconstruction, the first three levels of high-frequency decomposition were not taken into account, because predictive estimates went beyond the acceptable level. As a result of summing the forecast estimates, a reconstruction of the forecast for 20 years (2000-2019) was obtained. Further, the results obtained were compared with the data of an independent sample.

The results of the forecast made according to the model described above are shown in Fig. 10, and in fig. 11 – forecast for various models with preliminary reconstruction.

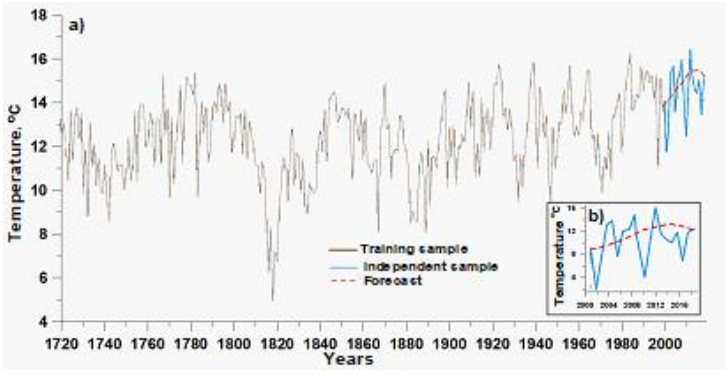


Fig. 10. Prediction results using wavelet decomposition

Note: a) – the time course of a series of training sample temperature + forecast; **b) –** forecast for the period 2000-2019.

We also carried out calculations of the forecast of the series with learning by the approximator by various models and compared the results with the original series (fig. 11). The first row is the initial data, 2-5 are forecasts for various models, 6 is the reconstruction $f=a_5+d_5+d_4$, 7 is the distribution of the approximator a_5.

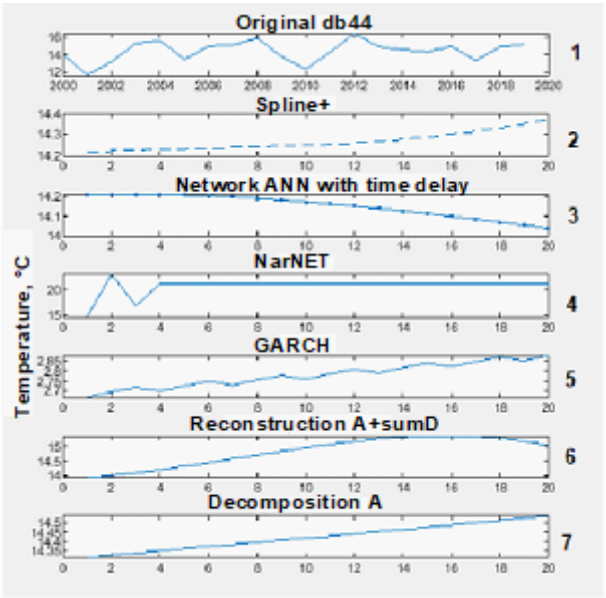


Fig. 11. Forecast results by various models

Tables 1-2 present estimates of the obtained results in the form of the mean square error of forecasts, calculated from the data of an independent sample, and the correlation coefficients of the forecast models and the approximator, depending on the size of the training sample, respectively.

According to the results obtained, it can be seen that a sample of 200 years is inferior to others in almost all indicators (300, 400 years). At the same time, the GARCH model gave a significantly biased forecast, which resulted in a high standard deviation, with a significant correlation coefficient. The NarNET model gave an unsatisfactory result both in terms of the standard deviation and the correlation coefficient. Although on a sample of 400 – the result has improved significantly. It should be noted that with direct prediction, without the use of

wavelets, all models, except ANN, gave unsatisfactory results. As follows from tables 1, 2, the model "A + sumD" showed good results.

According to the chosen method, wavelet decomposition was carried out, training procedures on Neural Network technologies for the period from 1700 to 2019. and the forecast for subsequent years was calculated, starting from 2020 to 2040 (fig. 12).

Table 1

Standard deviation between forecast models and independent sample data

Model	Sample size		
	200	300	400
	Mean square error		
SPLINE	1,344	1,316	1,317
ANN	1,339	1,411	1,351
NarNET	41,847	43,433	2,660
GARCH	132,595	136,820	137,105
A+sumD	208,424	**1,418**	**1,413**
AAA	1,293	**1,274**	**1,274**

Table 2

Correlation of forecast and approximator models

Model	Sample size		
	200	300	400
	Correlation coefficient		
SPLINE	-0,624	0,944	0,944
ANN	-0,613	-0,970	-0,647
NarNET	-0,608	0,393	0,588
GARCH	0,558	0,981	0,981
A+sumD	0,681	0,907	0,907

Conclusions. As a result, it can be noted that the chosen approach has generally justified itself and for long-term forecasting, wavelet decomposition can be used together with neural networks. Particular attention should be paid to the correct exclusion of noise components in the wavelet decomposition.

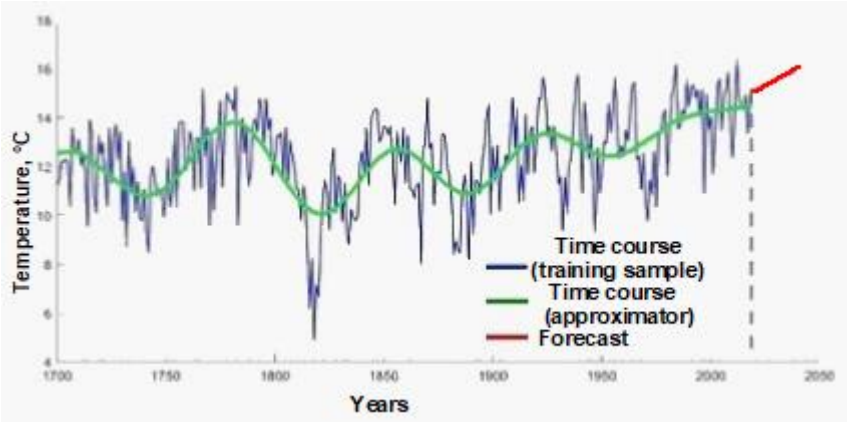


Fig. 12. Results of forecasting according to the developed model

Building the model includes the following steps: 1) decomposition of the series using a given wavelet function and calculation of the wavelet frequency characteristics of the series; 2)

decomposition of the series into approximating components with their subsequent forecasting using Neural Network technologies.

References:

1. Alekseev V.I. (2013), Analysis and forecasting of cyclic time series using wavelets and neural network fuzzy inference rules, *Bulletin of the Yugorsk State University*, No. 3 (30), pp. 3–10. (In Russ.).
2. Alexandridis A.K., Zapranis A.D. (2013), Wavelet neural networks: A practical guide, *Neural Net-works*, No. 42, pp. 3–27.
3. Arushanov M.L. (2015), *Practical issues of using wavelet transform in meteorology*, Tashkent, 51 p. (In Russ.).
4. Budak V.M., Fomin S.V. (1973), *Multiple Integrals, Field theory and Series*, Moscow, 420 p. (In Russ.).
5. Dyakonov V.P. (2012), *Matlab*, Moscow, 768 p. (In Russ.).
6. Hantemirov R.M., Corona C., Guillet S., et. al. (2022), Current Siberian heating is unprecedented during the past seven millennia, *Nature Communications*, No. 3, pp. 3-10.
7. Grossmann A, Morlet J. (1984), Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape, *J. Mathematical Analysis*, No. 15, pp. 723-736.
8. Haar A. (1910), Zur Theorie der orthogonalen Function systeme: Dissertation (Gottingen), *Math. Ann.*, No. 69, pp. 331-371.
9. Jin J, Kim J. (2015), Forecasting Natural Gas Prices Using Wavelets, Time Series, and Artificial Neural Networks, *PLoS ONE*, No. 10 (11), pp. 11-21.
10. Kanasevich E.R. (1985), *Analysis of time sequences in geophysics*, Moscow, 399 p. (In Russ.).
11. Lee G. (1998), Wavelets and wavelet evaluation, *Journal of Economic Theory and Econometrics*, No. 4 (1), pp. 123-157.
12. Malla S. (2005), *Wavelets in signal processing*, Moscow, 671 p. (In Russ.).
13. Medvedev V.S., Potemkin V.G. (2001), *Neural networks. MATLAB 6*, Moscow, 630 p. (In Russ.).
14. Meshcherskaya A.V., Rukhovets L.V., Yudin M.I., Yakovleva N.I. (1970), *Natural components of meteorological fields*, Leningrad, 199 p. (In Russ.).
15. Panovsky H.A., Brier G.W. (1972), *Some applications of statistics to meteorology*, Leningrad, 203 p. (In Russ.).
16. Rioul O. and Vetterli M. (1991), Wavelets and Signal Processing, *IEEE SP Magazine*, No. 10, pp. 14-38.
17. Rumaih M., Mohammad A., Fawzan A. (2001), *Time Series Forecasting Using Wavelet Denoising an Vityazev VP Wavelet analysis of time series*, Saint-Peterburg, 58 p. (In Russ.).
18. Vager B.G., Serkov N.K. (1987), *Splines in solving applied problems of meteorology and hydrology*, Leningrad, 160 p. (In Russ.).
19. Zhmurko D.Yu. (2017), Data analysis of the sugar subcomplex of the agro-industrial complex using wavelet transform (part 2 - discrete wavelet transforms), *Scientific journal of KubSAU*, No. 130 (06), pp. 12-19.

Information about authors:

Arushanov Mikhail – Research Hydrometeorological Institute (Tashkent, Uzbekistan), Doctor of Geographical Sciences, Professor. E-mail: mikl-arushanov@rambler.ru.

Umerov Khikmet – Research Hydrometeorological Institute (Tashkent, Uzbekistan), Candidate of Technical Sciences, Senior Researcher.

Сведения об авторах:

Арушанов Михаил Львович – Научно-исследовательский гидрометеорологический институт (Ташкент, Узбекистан), доктор географических наук, профессор. E-mail: mikl-arushanov@rambler.ru.

Умеров Хикмет Усниевич – Научно-исследовательский гидрометеорологический институт (Ташкент, Узбекистан), кандидат технических наук, старший научный сотрудник.

For citation:

Arushanov M.L., Umerov Kh.U. (2023), Prediction of the time series of temperature by hybrid models using wavelet decision and neural network, *Central Asian Journal of Geographical Researches*, No. 3-4, pp. 49-60.

Для цитирования:

Арушанов М.Л., Умеров Х.У. Прогноз временного ряда температуры по гибридным моделям с использованием вейвлет-решения и нейронной сети // Центральноазиатский журнал географических исследований. 2023. № 3-4. С. 49-60 (На англ. яз.)