Identification of parameters of adaptive time series models

V. D. Polezhaev¹, K. O. Yusupova²
¹Plekhanov Russian University of Economics, Streemanny per, 36, 117997, Moscow, Russia, e-mail: vpolej@gmail.com;
²Bauman Moscow State Technical University, ul. Baumanska2-ya, 5/1, 105005, Moscow, Russia, e-mail: uko2304@mail.ru

Abstract. The work is devoted to the study of adaptive models of time series and the development of methods for their construction. The most problematic stage in the implementation of algorithms for time series forecasting methods is the identification of unknown parameters, on which the adequacy of the prediction depends. On the basis of economic data, a number of adaptive models have been studied: using exponential smoothing, Brown, Holt-Winters and Theil-Wage models, which take into account seasonality. It was found that the main difficulties arise in the selection of constant smoothing and the choice of coefficients, since no universal algorithm for their task exists.

The paper proposes a method for integrating the adaptive method and the method of numerical optimization. As an optimization method, the non-gradient method was chosen - the Particle Swarm Optimization (PSO), the use of which provided less error in the final forecast and convergence in all applications. Algorithms for finding the parameters to be identified, and a predictive model based on real data are presented. It is shown that in some cases the proposed model has an advantage over other models of time series.

Keyword: forecasting, adaptive model, time series, particle swarm optimization, identification, exponential smoothing.

1. Introduction
Prediction of time series values based on its previous values is the basis for planning, managing and optimizing production volumes, etc. The task of forecasting a time series is solved on the basis of creating a forecasting model that adequately describes the process under study. To date, there are many models for analyzing and forecasting time series. The most common because of their simplicity and clarity is the additive model of exponential smoothing [1]. A characteristic feature of adaptive forecasting methods is their ability to continuously take into account the evolution of the dynamic characteristics of the studied processes. Very often there is a need to predict indicators subject to seasonal variations [2]. In these cases, Holt-Winters and Theil-Wage models are also used, which include both a trend and a seasonal component.

Objective: Analysis of the main approaches and the development of methods for improving the methodological tools for short-term forecasting of the series, identifying the possibilities of using adaptive forecasting methods using the methods of identifying model parameters to compile the most accurate forecast.

Modification of a method of exponential smoothing for seasonal ranks are the Holt-Winters and Theil-Wage methods. As a model of a series, its representation is used as a combination of a linear trend with a seasonal component superimposed either multiplicatively (Holt-Winters model) or additively (Theil-Wage model. It is assumed that the trend coefficients and the seasonal component may vary slowly over time. In accordance with this the computational process is arranged as an adaptive procedure, controlled by three adaptation parameters (one parameter is level adaptation, the second is the angle of inclination, and the third is seasonality coefficients). In the course of the calculations, a smoothed series is constructed, which at each moment of time \( t \) is the forecast according to the data up to \((t - 1)\) inclusive.

One of the problems associated with these models is that to start the calculations using the model, you need to somehow set the initial values of the coefficients, but there is no universal principle for specifying these coefficients, and their initial value naturally affects predictive properties of the model. That is, when building a model, it is necessary to select not only the optimal constant smoothing, but
also to choose the principle of specifying the model coefficients, thanks to which one could get the most accurate forecast.

2. Theil-Wage model

The Theil-Wage model (H. Theil, S. Wage) is Holt’s sophisticated model that takes into account seasonality and an additive trend, unlike the Holt-Winters model, it additively includes a linear trend, which is justified in solving some problems.

In economic practice, exponential trends with multiplicatively imposed seasonality are more common. The additive model, which has an independent value in economic research, is also interesting because it allows you to build a model with multiplicative seasonality and an exponential trend. To do this, it is necessary to replace the values of the original time series by their logarithms, which transforms the exponential trend into a linear and simultaneously multiplicative seasonal model into an additive one. Therefore, before using the additive model the members of the analyzed time series are usually replaced by logarithms, transforming the exponential trend into linear, and the multiplicative seasonality into additive. The advantage of the additive model is the relative simplicity of its computational implementation.

The Theil-Wage model is described by an expression that takes into account linear trend and additive seasonal effects:

\[ y_t = \alpha_1 t + \alpha_2 + \alpha_3 t \]

where row level without seasonal fluctuations is \( a_{1,t} = a_{1,t-1} + a_{2,t} \), \( a_{2,t} \) – coefficient of additive growth, \( g_t \) – seasonality coefficient, \( u_t \) – white noise. To fulfill the forecast for \( \tau \) steps forward, it is necessary to determine the estimate of a number of coefficients and substitute it in:

\[ \hat{y}_{t+\tau} = \hat{a}_{1,t} + \hat{a}_{2,t} \tau + \hat{g}_{t+\tau} \]

The coefficients are estimated using the following formulas:

\[ \hat{a}_{1,t} = \alpha_1 (y_t - \hat{g}_{t-\tau}) + (1 - \alpha_1) (\hat{a}_{1,t-1} + \hat{a}_{2,t-1}) \]

\[ \hat{g}_t = \alpha_2 (y_t - \hat{a}_{1,t}) + (1 - \alpha_2) \hat{g}_{t-1} \]

\[ \hat{a}_{2,t} = \alpha_3 (\hat{a}_{1,t} - \hat{a}_{1,t-1}) + (1 - \alpha_3) \hat{a}_{2,t-1} \]

where \( 0 < \alpha_1, \alpha_2, \alpha_3 < 1 \) – three parameters which are subject to definition, \( l \) – number of steps in one seasonal cycle (for monthly observations of \( l = 12 \), for quarter \( l = 4 \)), \( \hat{a}_{1,t}, \hat{a}_{2,t} \) – characteristics of development of the studied process, \( \hat{g}_1, \hat{g}_{t-1}, \ldots, \hat{g}_{t-1+\tau} \) – parameters of additive seasonality.

In order to determine the optimal values of the adaptation parameters model assume searching through various sets of their values and comparing the mean-square errors of the predictions resulting from this.

3. Creation of the forecast by Theil-Wage method

The data on income from radio communications, radio broadcasting, television and satellite communications in the Moscow region were obtained from the Unified Interdepartmental Statistical Information System (UniSIS). It is necessary to make a quarterly forecast for 2018 according to data for the 28 quarters of 2011-2017. Actual data for 2018 are known; an estimate of the prediction error was made using their values.

During the preliminary processing of the model range of 28 points (Fig. 1), it was noted that the time series under study is characterized by the presence of a linear trend and an additive component of seasonality. Therefore, for the compilation of the forecast, the Theil-Wage method was chosen [3]. This model contains three unknown parameters, the selection of which used the Particle Swarm Optimization.

Using the standard tools, we find the equation of the linear trend of the original series (Fig. 1):

\[ \hat{y}(t) = 110345.553t + 1057876.3246, \]
where \( t \) – month number, \( \hat{y}(t) \) – budget volume, thousand rubles (fig. 1). The coefficient of determination this case \( R^2 = 0.432 \).

Let us show a detailed description of the calculations using the Theil-Wage method formulas. For empirical reasons we accept model parameters equal \( \alpha_1 = 0.4, \alpha_2 = 0.8, \alpha_3 = 0.8 \), forecast period \( \tau = 1 \). Let us show the compilation of the forecast for the next quarter 2018 following the processed data. The initial value of the seasonality coefficients is

\[
\hat{g}_t^0 = \frac{1}{2} [(y_t - \hat{y}(t)) + (y_{t+1} - \hat{y}(t+1))].
\]

Model values will then be obtained by the following formulas, \( t = 1 \):

\[
\hat{y}_1 = \hat{a}_{t,0} + \hat{a}_{2,0} + \hat{g}_1^0; \quad \hat{a}_{1,1} = \alpha_1 (y_t - \hat{g}_1^0) + (1 - \alpha_1)(\hat{a}_{t,0} + \hat{a}_{2,0});
\]

\[
\hat{g}_1 = \alpha_2 (y_t - \hat{a}_{1,1}) + (1 - \alpha_2)\hat{g}_1^0; \quad \hat{a}_{2,1} = \alpha_3 (\hat{a}_{1,1} - \hat{a}_{t,0}) + (1 - \alpha_3)\hat{a}_{2,0},
\]

As \( \hat{a}_{1,0} \) and \( \hat{a}_{2,0} \) we take the coefficients of the linear regression equation without taking into account the seasonality: \( \hat{a}_{1,0} = 105787.3246, \hat{a}_{2,0} = 110345.553 \).

For \( t = 2 \) model values are found by the formulas:

\[
\hat{y}_2 = \hat{a}_{t,1} + \hat{a}_{2,1} + \hat{g}_2^0; \quad \hat{a}_{1,2} = \alpha_1 (y_t - \hat{g}_2^0) + (1 - \alpha_1)(\hat{a}_{t,1} + \hat{a}_{2,1});
\]

\[
\hat{g}_2 = \alpha_2 (y_t - \hat{a}_{1,2}) + (1 - \alpha_2)\hat{g}_2^0; \quad \hat{a}_{2,2} = \alpha_3 (\hat{a}_{1,2} - \hat{a}_{t,1}) + (1 - \alpha_3)\hat{a}_{2,1}.
\]

We carry out similar calculations successively up to \( t = 29 \). Thus, we obtain a forecast for one period ahead. For the forecast for 2-4 periods we assume \( \tau \) equal to 2, 3 and 4 respectively. For the
case under consideration, the relative error of the Theil-Wage model is determined by the formula
\[ D_{\text{error}} = \left| \frac{y_t - \hat{y}_t}{y_t} \right| \cdot 100\% \], it is equal to 3.3%.

If we take a different set of parameters of the Theil-Wage model, then the result of the forecast, even with a small change in these parameters, can greatly change. Thus, a change in the two parameters of the model \( \alpha_1 = 0.2, \alpha_2 = 0.7, \alpha_3 = 0.8 \) entailed increase in an error of the forecast up to 24.4%.

We introduce the automatic determination of the optimal set \( \alpha_1, \alpha_2, \alpha_3 \) into the forecast generation algorithm. To do this, we require that the parameters be defined so that the model’s error at all points is minimal over the entire time interval. The prediction algorithm code was implemented using the Matlab application package.

4. Particle Swarm Optimization for identification of model parameters
The Particle Swarm Optimization is a method of numerical optimization, which we will consider in the context of the problems of identifying unknown parameters of the objective function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \). The essence of the method lies in the search for optimal solutions at points that are called particles: the coordinates of particles of the swarm are those parameters, the selection of which will solve the problem [4]. The selection is carried out by the movement of these particles according to the principle of the best positions found in space, which is constantly changing when the particles find more advantageous positions. PSO is a non-gradient method: in its implementation algorithm, the exact gradient of the function being optimized is not required.

PSO has a particle search area, which is a \( n \)-dimensional space. At the beginning of the algorithm (at the first iteration), the particles are scattered randomly throughout the search area, and each particle has its own random initial velocity vector. At each point the particle has visited, the value of the objective function is calculated. Moreover, each particle remembers what and where the optimal value of the objective function it found, and also each particle knows where the point is located, which is the best among all the points that the particles memorized. At each iteration, the particles adjust the module and the direction of their speed in order, on the one hand, to be closer to the best point that the particle found itself, and, at the same time, move closer to the point that is globally better at the moment [5]. After a certain number of iterations, the particles should gather near the most suitable point.

The most interesting thing in the algorithm is the speed correction, the convergence of the algorithm depends on this step. The work uses a modification of the algorithm, which consists in adding a weight coefficient \( \omega \) before the current speed, due to which the speed changes more smoothly:
\[ v_{i,t+1} = \omega v_{i,t} + \varphi_p r_p (p_i - x_{i,t}) + \varphi_g r_g (g_i - x_{i,t}). \]

Here \( v_{i,t} \) – speed component at \( t \) iteration of an algorithm, \( x_{i,t} \) – \( i \) particle coordinate at \( t \) iteration of an algorithm, \( p_i \) – \( i \) coordinate of the best solution found by particle, \( g_i \) – \( i \) coordinate of the best solution found by all particles, \( r_p, r_g \) – random numbers in the range of \((0, 1)\), \( \varphi_p, \varphi_g \) – weighting coefficients which should be selected under a specific objective.

Let’s explain the meaning of these coefficients. \( \omega \) – the inertial coefficient (determining the tendency of a particle to maintain direction), \( \varphi_p \) – the cognitive coefficient (characterizing the tendency of a particle to its own best position), \( \varphi_g \) – the social coefficient (characterizing the tendency of a particle to the best position in the entire swarm). Each of these coefficients should be selected from certain considerations based on the logic of the method.

The correction of the current coordinates of each particle is carried out according to the formula
\[ x_{i,t+1} = x_{i,t} + v_{i,t+1}. \]
After that, the value of the objective function at each new point is calculated, each particle checks whether the new coordinate has become the best among all the points it visited. Then, among all the new points, we check whether we have found a globally better point, and, if we have found, we remember its coordinates and the value of the objective function in it.

When implementing the task, the following parameter values were used:

$$\omega = (0.4 - 0.9) \frac{k}{S} + 0.9, \quad \varphi_p = (0.5 - 1.5) \frac{k}{S} + 1.5, \quad \varphi_g = (1.5 - 0.5) \frac{k}{S} + 0.5,$$

where 0.9 and 0.4 are the largest and smallest values of the parameter, respectively, $S$ is the number of particles in the swarm. At the same time, at the beginning of work, the process tends to a larger value of the parameter, and over time it begins to study the neighborhoods of permissible values $\omega$. Such choice of components allows you to first look for the best position of each particle, and with time the search shifts to determine the best among neighboring particles.

We select the criterion by which the coefficients of the Theil-Wage model will be determined. That is, we make a function, a function, the minimum of which we need to find to determine the parameters $\alpha_1, \alpha_2$. We make the assumption in the problem that the parameter $\alpha_3$ is known. Let such a function be the average approximation error (also when selecting parameters, the condition of minimum forecast error will be required separately):

$$A = \frac{1}{n} \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left| y_i - \hat{y}_i \right|^2}.$$ 

So, the parameter $\alpha_3$ was taken equal to 0.8. The other two parameters were determined programmatically by Particle Swarm Optimization: $\alpha_1 = 0.2367, \alpha_2 = 0.9991$. In Figure 2, a graph of the process model is plotted against the initial time series.

Figure 2. The red graph is the original time series, the blue graph is the forecast time series with a set of parameters $\alpha_1 = 0.2367, \alpha_2 = 0.9991, \alpha_3 = 0.8$. 

5
Using Particle Swarm Optimization an optimal set of parameters was selected. Let's check the quality of the constructed model: the constructed model is considered good if the value of the average approximation error value \( A = \frac{1}{n} \sqrt{\frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{y_i}} \cdot 100\% \) is less than 7%. For a set of parameters obtained using PSO, an indicator \( A = 0.09\% \) that indicates a satisfactory degree of adequacy of the proposed approach.

Table 1 provides for comparison estimates of the forecast error for this and other sets of values for the Theil-Wage model parameters considered in the paper.

**Table 1. Forecast error estimates**

| №  | Set of parameters \( \alpha_1, \alpha_2, \alpha_3 \) | A   | Forecast error |
|----|-----------------------------------------------|-----|---------------|
| 1  | \( 0.2367, 0.9991, 0.8 \)                           | 0.09%   | 0.25%         |
| 2  | \( 0.2, 0.7, 0.8 \)                                    | 1.15%   | 24.4%         |
| 3  | \( 0.4, 0.8, 0.8 \)                                    | 1.17%   | 3.2%          |

5. Conclusions

In this paper, we considered methods for constructing a forecast in the case of source data in the form of time series. As it turned out, most of the methods in this category have one common problem - unknown parameters. The paper describes in detail the essence of the problem and its solution using the example of economic data. It was proposed to search for the parameters using the integration method of two methods: Theil-Wage (chosen for the specifics of the data) and Particle Swarm Optimization. With this approach, the simulation results gave a set of model parameters in which the average approximation error and the forecast error were minimal. Thus, as a result of solving the problem of short-term forecasting of the process with seasonal dynamics, the following results were obtained:

- identified the possibility of using adaptive forecasting methods using the methods of identifying model parameters to compile the most accurate forecast;
- a method of integrating the adaptive method and numerical optimization of the search for the parameters of the model of the exponential smoothing by the gradient-free method of a particle swarm;
- it is shown the possibility to achieve greater accuracy of the forecast, using the search adjusting procedures of the PSO;
- the code of the forecasting algorithm was developed and implemented using the Matlab application software package;
- it is noted that in some cases, the proposed model has an advantage over other time series models.

References

[1] Lukashin Yu.P. Adaptive methods of short-term forecasting of time series / Yu.P. Lukashin. - Moscow: Finance and statistics, 2003. 413 pages. A reference

[2] Polezhayev V. D., Yusupova K.O. Forecasting of rating of television channels on the basis of these media researches//Basic researches. – 2019. – No. 5.

[3] Svetunkov I.S. Methods of social and economic forecasting in 2 t. T. 2 models and methods: the textbook and a practical work for the academic bachelor degree / I.S. Svetunkov, S.G. Svetunkov. Moscow: Yurayt publishing house, 2018. 447 pages.

[4] Kennedy J. and Eberhart R. Particle Swarm Optimization. In Proceedings of the Fourth IEEE International Conference on Neural Networks, Perth, Australia. IEEE Service Center(1995) 1942-1948.
[5] Use of information and communication technologies for teaching physics at the technical university / Polezhaev V.D., Polezhaeva L.N., Kamenev V.V.//AIP Conference Proceedings "Information Technologies in Education of the XXI Century, ITE-XXI 2015: Proceedings of the International Scientific-Practical Conference "Information Technologies in Education of the XXI Century" 2017. Page 030013. DOI: 10.1063/1.4972452