УДК 004.02

Dibrivnyi O. A. *State University of Telecommunications, Kyiv*

COMPARATIVE ANALYSIS OF TIME SERIES FORECASTING BASED ON THE TREND MODEL AND ADAPTIVE BROWN`S MODEL

Article dwells upon statistical methods of analysis of time series, construction of trend and trend-seasonal models of time series and their usage for forecasting of the development of economic processes. A comprehensive comparison of time series forecasting using a trend model and an adaptive Brown model is also performed. The forecasting of the bitcoin rate against the dollar is compared using these two models.

Keywords: *time series, structural-forming components, trend model, trend-seasonal model, forecasting, determination coefficient.*

Дібрівний О. А. *Державний університет телекомунікацій, Київ*

ПОРІВНЯЛЬНИЙ АНАЛІЗ ПРОГНОЗУВАННЯ ЧАСОВИХ РЯДІВ З ВИКОРИСТАННЯМ ТРЕНДОВОЇ ТА АДАПТИВНОЇ МОДЕЛІ БРАУНА

В статті розглянуто статистичні методи аналізу часових рядів. Розглянуто основні етапи алгоритму побудови трендової і тренд-сезонної моделей, який включає в себе: виділення основних структурно-утворюючих компонент часового ряду (тренду, сезонних коливань, циклічної і залишкової компоненти), методи вибору моделі для опису ряду, а також методи перевірки обраної моделі на адекватність та перевірки можливості прогнозування на основі обраної моделі. Основну увагу приділено алгоритмам короткочасного прогнозування часових рядів на основі трендової моделі та за допомогою адаптивної моделі Брауна. Для прогнозування за допомогою трендової моделі описано етапи побудови точкового та інтервального прогнозів, а також критерії вибору найкращої моделі для опису часового ряду серед чотирьох можливих варіантів: лінійної, поліноміальної, логарифмічної та експоненціальної моделей, в залежності від значення коефіцієнту детермінації. При описі адаптивної моделі Брауна розглянуто алгоритм адаптації моделі до результатів прогнозування в залежності від пріоритетності часових моментів, а також описується можливість враховувати зміну тенденцій в ряді та коливань значень, після чого наводиться сам алгоритм побудови адаптивної моделі Брауна. В заключній частині статті проводиться порівняння прогнозу курсу біткоїна до долара, виконаного на основі поліноміальної трендової моделі та прогнозу, зробленого за допомогою адаптивної моделі Брауна. На основі вищевказаного порівняння були зроблені висновки про переваги та недоліки моделей, що розглядалися в статті.

Ключові слова: *часовий ряд, структурно-утворюючі компоненти, трендова модель, тренд-сезонна модель, прогнозування, коефіцієнт детермінації.*

Дибривный О. А. *Государственный университет телекоммуникаций, Киев*

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ПРОГНОЗИРОВАНИЯ ВРЕМЕННЫХ РЯДОВ З ИСПОЛЬЗОВАНИЕМ ТРЕНДОВОЙ И АДАПТИВНОЙ МОДЕЛИ БРАУНА

В статье рассматриваться статистические методы анализа временных рядов, построение трендовой и тренд-сезонной модели, а также ихнее использование для прогнозирования развития экономических процессов. Также проводиться комплексное сравнение прогнозирования временных рядов с использованием трендовой модели и адаптивной модели Брауна. В частности, сравнивается прогнозирование курса биткоина к доллару с использованием этих двух моделей.

Ключевые слова: *временные ряды, структурно-образующие компоненты, трендовая модель, тренд-сезонная модель, прогнозирование, коэффициент детерминации.*

1. Introduction

The need for analysis and forecasting of time series for the development of economic processes occurs quite often. It can be used for: forecasting of stock market indicators, cash flows, changes in inventories, etc. Trend and trend-seasonal models, with all their simplicity, can provide more reliable prediction results than complex economic and mathematical models based on systems of algebraic and differential equations, especially in the short-term and medium-term predictions.

Trend and trend-seasonal models are based on the assumption that the main factors and trends of the previous period will continue for the period of forecast, or that the direction and change of trends can be justified and taken into account, that is, it assumes a large inertia of economic systems. Nowadays, the dynamics of economic phenomena and processes at the level of industries and enterprises are rapidly increasing, as the result collected data of statistical observations becomes useless. In this case, it is convenient to use models based on a small amount of fresh data that can adapt to the process change, so called adaptive models.

2. Preliminary analysis of time series

Dynamic processes occurring in economic systems are usually represented as a series of values of some economic index consistently arranged in chronological order. Changes to this index reflect the course of development of the economic process which is studied. The sequence of observations of a single pointer, ordered according to the successively increasing or decreasing value of another indicator, is called a dynamic row or a series of dynamics. If time is taken as a sign, depending on which ordering takes place, then such a dynamic series is called time series. The elements of the series include: the values observed signs (the levels of the series), moments and intervals of time, which include levels. Time series, in which certain values of the economic index referring to certain points of time, are called momentary ones (e.g. balances on accounts on the first day of each month).

If the levels of the time series are formed by addition, averaging, or some other method of aggregation over selected period, then such rows are called interval intervals (e.g. a number of outputs by month, a number of average monthly wages of workers). The length of the time series can be determined as the time which have passed from the initial to final observation moment, or the number of levels of the series. If in the time series it is possible to allocate a long-term pattern of level change, the trend is considered. Thus, the trend determines the general direction of the development of the economic process. The economic-mathematical model, in which the development of an economic model is reflected through the trend of its main indicators, is called a trend model. To identify the trend of time series, as well as to construct and analyse trend models, the apparatus of probability theory and mathematical statistics is used. But one must consider that this toolkit is intended for the processing of simple statistical aggregates, and therefore the use of methods of probability theory and mathematical statistics requires appropriate corrections. The difference between time series and simple static aggregates is that the levels of the time series depend on each other, while the elements of the statistical aggregate are independent of each other. In addition, the levels of time series are ordered in time and their mixing is inadmissible, and the elements of the statistical aggregate are not ordered. Repositioning these elements does not change the values of statistical indicators (variance, average value, etc.). In the general case, the time series of the index Y, which consisted of n levels $Y_1, Y_2, Y_3, \dots, Y_{n1}$, consist of four structural elements. The main structural element is the trend U_b which leads to a systematic change in the indicator of observation over a specified period of time. Also, in the time series close to usual fluctuations in relation to the main trend can be observed. Fluctuations with a period in a year due to the impact of natural and climatic conditions are called seasonal fluctuations V_t. This effect is most pronounced in agriculture and in energy consumption. According to seasonal conditions, the term of production during the year depends on the influence of natural factors. Seasonality can be observed in marine and annual transportation, fishing and construction.

Alongside fluctuations with a period in a year, in the time series also can be observed fluctuations with a period of several years C_t. This type of fluctuations is called cyclic, and their presence determined by the general recession and rise of the world economy.

Trend, seasonal and cyclic components are called regular or systematic components. The complex time of the time series that remains after the allocation of regular components from time series called the residual component ε_t. It is a compulsory part of any time series of economic indicators, as economic processes always accompany small changes caused by the weak influence of short-acting random factors. If the systematic components of the time series allocated correctly, then the residual component will correspond to the following properties:

- Accidental changes in values;
- The law of distribution corresponding to normal;
- Equality to zero of mathematical expectation;
- Independence of values of levels from each other (absence of autocorrelation).

The verification of the adequacy of trend models is based on checking the conformity of the residual component with these four properties.

If the time series is represented as the sum of the corresponding component, then the resulting model is additive and has the form:

$$Y_t = U_t + V_t + C_t + \varepsilon_t .$$

If the components multiply, then the model is multiplicative:

$$Y_t = U_t V_t C_t \varepsilon_t .$$

Also existed a mixed model and its form is:

$$Y_t = U_t V_t C_t + \varepsilon_t .$$

In most cases, in the analysis of time series, the presence of the cyclic component C_t neglects, and if you need to consider the cyclic component, then special methods based on spectral analysis are used for its selection. In cases where C_t equals zero, such a row is called a trend-seasonal. It`s assumed that U_t – some smooth function whose degree of smoothness is preassigned. The seasonal component U_t has a period T, and the number of levels of a series aliquot to periods of seasonality $n = mT$, where m – is the number of years of observation [1].

Output data of the time series and its components can be represented not only in the form of a series Y_t, but also in the form of the matrix Y_i, where i is the number of the year, j is the month number. The relationship between the indices is as follow:

$$i = \left[\frac{t}{T}\right] + 1; j = t - (i-1)T;\ \ i = \overline{1,m};\ \ j = \overline{1,T}.$$

To select the component of the trend of seasonal time series uses iterative filtration methods, which allow multiple use of the slippery average and simultaneous evaluation of the seasonal wave at each step. Iterative methods differ in simplicity and the appropriate frequency of filtering components of a series. The disadvantage of such methods is the loss of part of the information at the ends of the time series. One such method is the Chetverikov method, with the following algorithm of analysis:

1. The row being studied is adjusted according to the formula of the average chronological with the period of 1 year, $T = 12$. The values from the beginning and the end, which cannot be fitted, are discarded. The result is an estimation of the trend:

$$\widetilde{Y_t} = U'_t ,$$

then the deviation of the initial series from the fitting is calculated:

$$l_{i,j} = Y_{i,j} - U'_{i,j}.$$

2. The mean-square deviation is calculated for each year:

$$\sigma_i = \sqrt{\frac{\sum_{j=1}^{T} e_{i,j}^2 - \left(\sum_{j=1}^{T} e_{i,j}\right)^2 / T}{T-1}},$$

the received deviations are standardized:

$$\tilde{l}_{i,j} = \frac{l_{i,j}}{\sigma_i}.$$

3. The previous average seasonal wave is counted according to standardized deviations:

$$V_j^1 = \frac{\sum_{i=1}^{m} \tilde{l}_{i,j}}{m}.$$

4. The average preliminary seasonal wave is multiplied by the mean square deviation of each row and subtracted from the initial series, thus obtaining the first estimate of the series:

$$U_{i,j}^1 = Y_{i,j} - V_j^1 \sigma_i .$$

5. The resulting trend is smoothed out by a sliding average of five points, getting a new estimate of the trend $U_{i,j}^2$. To avoid loss of points at the beginning and end of the row, they are smoothed out by three points, while for extreme points, special smoothing formulas are used:

$$U_1^2 = \frac{5U_1^1 + 2U_2^1 - U_3^1}{6}; \quad U_n^2 = \frac{5U_n^1 + 2U_{n-1}^1 - U_{n-2}^1}{6}.$$

6. New deviations of the initial series are calculated between trends Y_t and U_t^2:

$$l_t^2 = Y_t - U_t^2 .$$

7. The residual component is calculated:

$$\varepsilon_{i,j} = {l_{i,j}}^2 - V_j^2 \sigma_i ,$$

then the seasonal wave voltage coefficient is determined:

$$k_i = \frac{\sum_{j=1}^{T} l_{i,j}^2 \varepsilon_{i.j}}{\sum_{j=1}^{T} \varepsilon_{i,j}^2}.$$

This is followed by an examination of the adequacy of the model - the compliance of the model of the process being investigated. If during inspections it turns out that the selected model is sufficient, forecasting can be conducted [2].

3. Forecasting based on trend model

During the extrapolation forecasting of economic processes, it is necessary to determine two elements: point and interval forecasts.

The point forecast is the value of the economic index in the future calculated by substituting time into the equation of the chosen growth curve. The match of actual data in the future and the point forecast value is unlikely. Therefore, the point forecast is supplemented by two-sided boundaries - at such an interval in which, with a high degree of probability, the actual value of the predicted pointer is expected. Such a forecast is called interval, it is determined by the confidence interval:

$$Y_\phi(t) = U(t) \pm \Delta Y,$$

where $Y_\phi(t)$ – actual value in the future;

ΔY – confidence interval.

The value of the confidence interval depends on the standard error of the time series approximation by using the growth curve, the time of determination of the forecast, the length of the time series and the level of significance of the forecast.

The standard error of approximation is determined by the expression:

$$S = \sqrt{\frac{\sum_{t=1}^{n}(Y_t - U_t)^2}{n-k}},$$

where k is the number of parameters of the trend model.

For a linear trend, the confidence interval is determined by the formula:

$$Y = t_\alpha S \sqrt{1 + \frac{1}{n} + \frac{3(n+2L-1)}{n^2(n^2-1)}},$$

where L – period of formation (number of steps for which the forecast is divided); t_α – Student\`s criterion for the number of freedom degrees of n-2 of significance level $\alpha = 0{,}2$.

For polynomials of the second and third orders, an expression is used in which the beginning of the time reference is transferred to the middle of the time series of observations:

$$\Delta Y = t_\alpha S \sqrt{1 + \frac{1}{n} + \frac{t_L^2}{\sum t^2} + \frac{\sum t^4 - 2t_L^2 \sum t^2 + n t_L^4}{n\sum t^4 - (\sum t^2)^2}},$$

where t_L is the time of forecast, the addition is performed for all values of the time series:

$$-\frac{n-1}{2} < t < \frac{n-1}{2}.$$

Even though the given formulas allow to determine the forecast for any number of steps, the prognosis attempt for a large period of time leads to a large error. The length of the forecasting period should not exceed 1/3 of the length of the observation series [3].

4. Brown\`s adaptive model

Adaptive models can adapt their structure and parameters to change of the properties of the modulated process. As in trend models, the main factor is time, but observation (series levels) provides different priorities depending on their impact on the current level of the series. This allows you to consider changes in the trend of the series, as well as fluctuations.

Most often, Brown's adaptive model is used for short-term forecasting. It allows you to reflect the development of a linear, or parabolic tendency, as well as a series without tendencies. Accordingly, the models of zero, first and second order are distinguished:

$$Y(t+k) = A_0;$$

$$Y(t+k) = A_0 + A_1 k;$$

$$Y(t+k) = A_0 + A_1 k + A_2 k^2;$$

where t – the current time, k – the time of formation.

The order of the model is determined from the preliminary analysis of the time series and the development laws of the predicted process [4, 5].

The first order model is constructed as follows:

1. The first order model is constructed as follows: using OLS on several first points we find the values A_0, A_1:

$$Y_p(t) = A_0 + A_1 t.$$

2. Using the found parameters, we find the value at the following point:

$$Y_p(t+k) = A_0(t) + A_1(t)k, k = 1.$$

3. Found forecasting error:

$$e(t+k) = Y(t+k) - Y_p(t+k).$$

4. In accordance with the error, we change the values of the parameters of the model:

$$A_0(t+1) = A_0(t) + A_1(t) + (1-\beta)^2 e(t),$$
$$A_1(t+1) = A_1(t) + (1-\beta)^2 e(t),$$

where β – the discount rate of data $0< \beta <1$.

5. Using the model with the corrected parameters, we find the forecast for the next step and return to paragraph 3 if $t<N$ (the training time of the model has not yet expired), at $t \geq N$ we use the obtained value as a prediction without changing the parameters of the model.

6. We supplement the point forecast by interval:

$$\Delta Y = \Delta t_\alpha S_y \sqrt{1 + \frac{1}{n} + \frac{3(n+2k-1)}{n(n^2-1)}},$$

where t_α – value of the Student criterion, S_y – the predicted mean square deviation of the prediction pointer, n – the number of observations of the series.

5. Comparison of the results of forecasting based on the trend model and adaptive brown`s model

This paper continues research presented in [6-8] (table 1).

The forecasting was based on Bitcoin rate to USD, for the period, from January the 15th to December the 1st of 2017.

A polynomial model was chosen to construct a forecasting model for forecasting, since its determination coefficient was higher than the similar coefficients in other types of models.

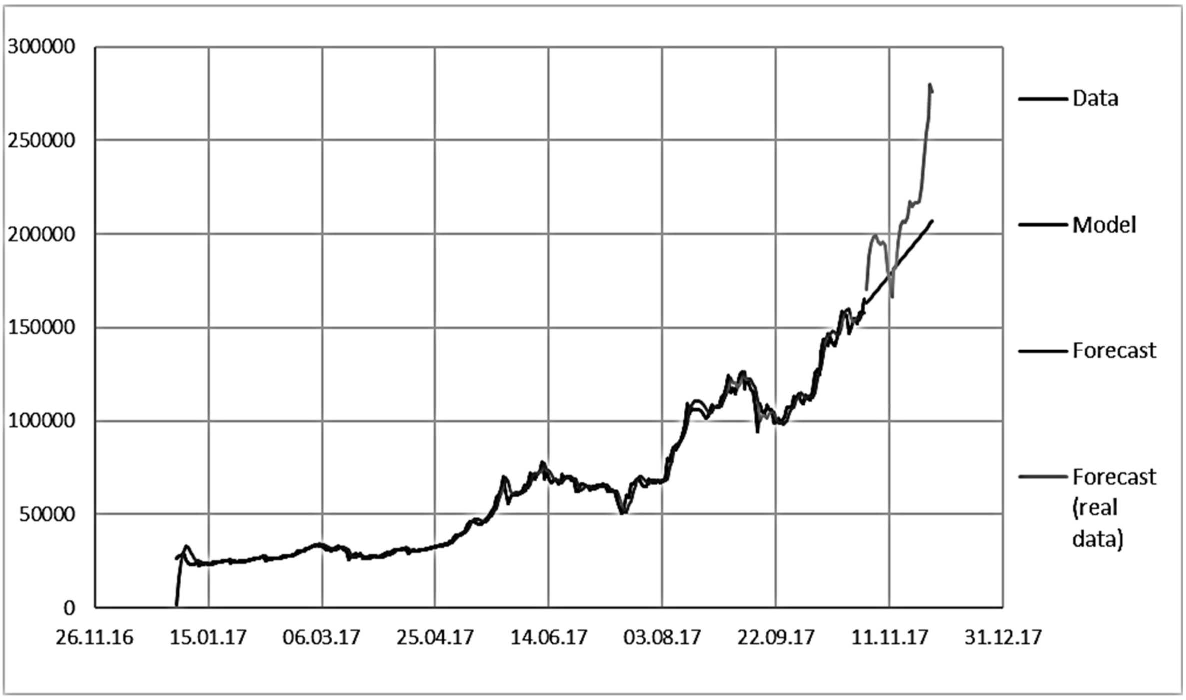
Table 1. Values of the determination coefficient for different types of trend models

Type of the model	**The value of the determination coefficient**
Linear	0.1084
Exponential	0.1143
Logarithmic	0.1085
Polynomial	0.4481

As a result, we have the following forecast for the change of the economic index for 2018 (Pic. 1).

Pic. 1. Forecasting using a trend model of polynomial type

Using the previously described forecasting algorithm using Brown's adaptive model, we build another forecast (Pic. 2)



Pic. 2. Forecasting using a Brown's adaptive model

Comparing the obtained forecasts, one can distinguish the following features:

• For prediction based on the trend model, the biggest error was 4.1% and it was observed at the beginning of forecasting, which is explained by the fact that the trend forecasting model does not consider new economic trends.

• During long-term forecasting, the average relative error was 1.37%, therefore the trend model can be used for long-term forecasting.

• The average relative error for forecasting based on Brown's adaptive model was 2.1%, which is more than in the trend model. This is due to the fact that the adaptive model does not take into account seasonal factors.

• Adaptive model more accurately describes the direction of changes in the indicator, but some time delay is present.

References (MLA)

1. Fedoseev V. V., Garmash A. N., Dajitbegov D. M., Orlova I. V., Polovnikov V. A. "Economic-mathematical methods and applied models." *Moscow: UNITY* (2002): 27-31.

2. Lukashin Yu. P. "Adaptive methods of short-term forecasting of time series." *Moskva: Financy i Statistika* (2003): 15-22

3. Woodward W. A., Gray H. L., Elliott A. C. "Applied time series analysis." *CRC Press* (2012): 141-145.

4. Shumway R. H., Stoffer D. S. "Time series analysis and its applications." *Springer* (2011): 576.

5. Kantrovich G. G. "Time series analysis: lecture and methodological materials." *Economic journal of the higher school of economics* (2002): 21-23.

6. Onyshchenko V. V. "Analysis of the field of communication based on the model of Ferghüst-Pearl." *Economics, management, business* 1(5) (2012): 28 - 32.

7. Matychyn I. I., Onyshchenko V. V. "Sigmoid curves as a model of dynamics of the telecommunication industry." *Economics, management, business* 1(9). (2014): 28-34.

8. Tregub I. V., Tregub A. V. "Application of cointegration analysis for studying the mutual influence of financial time series." *Fundamental research* 8(3). (2015): 620-623.

Література (ДСТУ)

1. Федосеев В. В. Экономико-математические методы и прикладные модели / В. В. Федосеев, А. Н. Гармаш, Д. М. Дайитбегов, И. В. Орлова, В. А. Половников. – Москва: ЮНИТИ, 2002: – 392 с.

2. Лукашин Ю. П. Прикладные методы кратковременного прогнозирования временных серий / Ю. П. Лукашин. – Москва: Финансы и Статистика, 2003. – 416 с.

3. Woodward W. A. Applied Time Series Analysis, CRC Press (2012): P. 141-145Woodward I. Applied Time Series Analysis / W. A. Woodward, H. L. Gray, A. C Elliott. – CRC Press: 2012. – P. 141-145.

4. Shumway R. H., Stoffer (2011), Time Series Analysis and its Applications, Shumway R. H., Stoffer Springer: 2011. – 576 p.

5. Кантрович Г. Г. Анализ временных рядов: лекции и методологические материалы / Г. Г. Кантрович // Экономический журнал высшей школы экономики. – 2002. – №1. – С. 21-23.

6. Онишенко В. В. Аналіз галузі зв'язку на основі модель Фергуста-Перла / В. В. Онищенко // Економіка, менеджмент, бізнес. – 2012. – №1(5). – С 28-32.

7.Матичин І. І. Сігмойдові криві як модель динаміки індустрії телекомунікацій / І. І. Матичин, В. В.Онищенко // Економіка, менеджмент, бізнес. – 2014. – №1(9). – С. 28-34.

8. Трегуб І.В., Трегуб А.В. Застосування коінтегрального аналізу для вивчення взаємного впливу фінансових часових рядів // Фундаментальні дослідження. – 2015. – № 8(3). – С. 620-623.)

Автор статті

Дібрівний Олесь Андрійович – аспірант, асистент кафедри інженерії програмного забезпечення, Державний університет телекомунікацій, Київ, Україна. Тел.: +380 (93) 432 65 43. E-mail: jokeresdeu@gmail.com.

Author of the article

Dibrivnyi Oles Andriiovych – postgraduate student, assistant at software engineering department, State University of Telecommunications, Kyiv, Ukraine. Tel.: +380 (93) 432 65 43. E-mail: jokeresdeu@gmail.com.

Дата надходження в редакцію: 18.01.2018 р.

Рецензент: кандидат технічних наук, доцент А. П. Бондарчук
Державний університет телекомунікацій, Київ