PROBLEM SOLVING OF FORECASTING ON SHORT PERIODS IN THE CASE OF TRANSITIONAL STRUCTURE-CHANGING CHARACTER OF DEVELOPMENT OF ECONOMICS IN THE PRESENCE OF CONSIDERABLE SEASONAL AND STOCHASTIC COMPONENTS IN TIME SERIES.

MODEL OF NONLINEAR-ADDITIVE SEASONAL DEVELOPMENT WITH MULTIPLICATIVE STOCHASTIC CORRECTION BY THE DECLARATIVELY LIMITED AUTOREGRESSIVE ROW OF RELATIVE RESIDUALS.

PART FIRST. GENERAL THEORETIC.

1.1 There are two primary purposes of analysis of time series:
a) determination of nature of time series; b) forecasting (extrapolational calculation of future values of time series based on the present actual today's and past values).

1.2 Unlike methods of interpolations, for which it is enough to choose the approximate interpolating formula such as Taylor, McLoran, Lagrange, Tchebyshev or polynomials series without creation of adequate model, forecasting requires, that the model of series must be identified and, more or less, formally described. As soon as a model is specified, that the extrapolating formula is creates, it becomes possible with its help to interpret analysed information and to estimate its accuracy and adequacy, by the analyzing a difference between calculated by a model and actual values of the forecasted figure. This estimation by the size and statistics of distribution has the name **method of analysis of residuals**.

1.3 In future we will need very important concepts which are used in the analysis of time series, such as: trend of time series, seasonal component of time series. That is why farther we will give short explanation to these elements:

1.4 Formally **a time series** is the series of the supervisions of the analyzed casual figure, wich are done in the successive moments of time. What are the fundamental differences between time series and simple sequence of supervisions that form a random sample? There are two differences: at first, unlike the elements of random sample the members of time series are not statistically independent; secondly, the members of time series are not identically distributed.

1.5 The nonrandom function, formed under action of general or long-term tendencies, which influence on a time series, is named **trend.** For example, the factor of growth of the explored market can come forward as forming tendency. There is not automatic method of trend exposure in a time series. However, if trend is monotonous (proof increase or proof diminish), it is usually not difficultly to analyze such time series.

1.6 The concept **of seasonal component** is used for denotation of nonrandom function, which is formed on the basis of the periodically repeated vibrations of the explored time series in the specified time of year. Often this function is measured in percents, which characterize seasonal deviations from trend. It is typical that every supervision looks very like neighbour one, and also there is the repeated seasonal constituent. It means that every supervision looks like the supervision, which was present in the same month the last year. On the whole, periodic dependence can be formally definite as correlated dependence of order κ between every i-th element of time series and (i-k)th element. It can be calculated by autocorrelation (i.e. correlation between the different members of the same time series); κ usually is named lag (sometimes use equivalent terms: shift, slide, delay). If a measuring error is not too large, seasonality can be defined by sight, examining the behaviour of members of time series through each κ of temporal units.

PART SECOND. STATE-OF-THE-ART REVIEW OF METHODS OF CENSUS.

2.1 Methods of Census, created in Bureau of Census of the USA, carry out additive or multiplicative decomposition of time series X into trend T, cyclic C, seasonal S and casual (irregular, or fluctuation) I components. In a general view:

$$X_t = T_t + C_t + S_t + I_t$$

or

$$X_t = T_t \cdot C_t \cdot S_t \cdot I_t$$

2.2 Components, definite by these methods, are used for a further “component divided” medium-term forecasting by Taylor, McLoran, Lagrange, Tchebyshev, spline or polynomials series, linear extrapolation, exponential smoothing, and other methods, i.e. volume of the forecasted information is calculated no more than on 10% from the volume of supervisions for past periods.

2.3 However, attempt to use this method at least for a forecasting on a year ahead, on the basis of figures of air motion at FIR Ukraine for the last 5 years (minimum number of complete cycles, it is required by the methods of Census) brings to the exposure (in terms of methods of Census) of too large casual (irregular, or fluctuation) I-component already on the stage of decomposition. It considerably worsens the confidence interval of the forecasted figures, i.e. it worsens accuracy of forecasting. For example, at the analysis of total volumes of “units of service” at FIR Ukraine the mean-square value of I-component was about 6% from an average monthly one. Moreover, it was discovered regular, but formally unseasonable, sizeable surges (spikes) with a size about 14% from an average monthly value. The increase of number of the analyzed years yet more worsens accuracy of Census decomposition into the components.

2.4 Such inadequacy of standard method of Census is explained by anomalous (force-majeure) influence on development of air motion at FIR Ukraine of crises events, such as disintegration of USSR and Russian DEFOLT in 1998. The characteristic signs of transitional economy are the rapid changes of figures and qualitative adjustment of its structure. Rapid change of figures, even sometimes change of the directions of development of figures mean that the methods of analysis of stationary processes become less usable ones as a result of maladjustment to the processes with nonlinear and unstationary character. Considerable role in forming of economic processes begin to play unstationary seasonal vibrations. Analyzing character of changes of seasonal vibrations, it is possible to describe in qualitative classes their structure and their progress trends and to build an adequate model for the quantitative analysis and forecasting.

2.5 Also at the analysis was discovered considerable heteroskedasticity of casual I-component, heterogeneity of its dispersion and correlation between I-component and trend-seasonal component.

2.6 Therefore forecasting on the basis of methods of Census is improper and inadvisable.

2.7 Taking it into account, the decision was accepted to develop an own model for medium-term and long-term forecasting of development of air motion at FIR Ukraine on the basis of study of world experience analysis and perfection of existent methods such as method of Census, Fourier analysis, ARIMA, exponential smoothing, Holt’s method and of Robert F. Engle’s conception of autoregressive conditional heteroskedasticity.

PART THIRD. FUNDAMENTAL PRINCIPLES.

3.1 In 2003 the model of nonlinear-additive seasonal development with multiplicative stochastic correction by the declarative limited autoregressive row of relative residuals was created and improved in 2004.

3.2 Unlike the standard methods of Census the stochastic component Δ is entered in this model (actually it consists of steady-state saltatory and gradual interventions). This is not fluctuation noise, but can not be included to trend or seasonal component:

$$X_t = T_t + C_t + S_t + \Delta_t + I_t$$

3.3 The formal and actual monopolist, similar to “Aeroflot”, that would be in a position to enjoy an advantage nonmarket priorities with respect to other users of air space, does not exist in Ukraine unlike in Russia. That is why if temporal, or on new directions, additional commercial interests appear, any airline can make use of them, and vice versa, at the loss of profitability in some period or on some direction, any airline has an opportunity to execute compensating flights on other routes and in other periods. Adopting into account also circumstance that an air transport, as well as any other type of transport, is not industry with direct consumption, and it serves present interests of other industries of economy, the formula of multiplicative correction is used:

$$X_t = TS_t \cdot (1 + \varepsilon_t) + I_t$$

3.4 In this formula TS is the preliminary formed base nonlinear-additive model of seasonal development, ε is stochastic adjuster that is calculated as the declarative limited autoregressive row from relative residuals (errors) between X and TS for the analyzed interval of time.

PART FOURTH. IMPLEMENTATION.

4.1 The base nonlinear-additive model of seasonal development is determined by the formula:

$$TS(m,n) = (1+\alpha_T \cdot (n-n_0))^{\beta_T} \cdot \frac{V_0}{12} + (1+\alpha_S \cdot (n-n_0))^{\beta_S} \cdot \sum_{i=1}^{5} A_i \cdot \cos\left(\frac{2\cdot\pi}{12}\cdot(i\cdot m-\varphi_i)\right)$$

Where following denotations are used:

TS it is base value of the forecast figure in an m-th month of n-th year.
m {1..12} it is number of month
n {>= 2002} it is number of year
n_0 { =2002} it is number "initial year"
α_T {without limitations} it is annual increase to trend of the forecast figure (unknown)
β_T {0..1} it is degree of non-linearity of annual increase to trend of the forecast figure (unknown)
V_0 {without limitations} it is volume of the forecast figure in an "initial year" (unknown)
α_S {without limitations} it is annual increase of seasonal of the forecast figure (unknown)
β_S {0..1} it is degree of non-linearity of annual increase of seasonal component of the forecast figure (unknown)
A_i {without limitations} it is amplitude of *i-th* harmonics of seasonal component of the forecast figure (unknown)
φ_i {0..12} it is phase of *i-th* harmonics of seasonal component of the forecast figure (unknown)

4.2 The simplest variant is used for the modeling of development of seasonal component. In the case of considerable structural changes in an economy it is necessary to examine separately development of every harmonics of seasonal component, and also to regard φ_i not a constant but function from time.

4.3 For compensation of the factors stated in paragraph 2.5 the unknowns are searched by an indirect least-squares method, that as such, that minimize sum of squares of relative residuals (errors) for the analyzed interval of time between the actual values of the forecast figure and values of base model of the forecast figure except for force-majeures periods. The criterion function of minimization is

$$S_{TS} = \sum_{m,n}\left(\gamma(m,n)\cdot\frac{TS(m,n)-X(m,n)}{TS(m,n)}\right)^2$$

Where $\gamma(m,n)$ is sign coefficient, it equals to 0 for force-majeures periods and 1 for all others.

4.4 Unlike the methods of Census and ARIMA, this model determines mutually independent nonlinear character of development both to trend (i.e. it actually interpolates the initial stage of long-period tendency of Kondratieff waves) and seasonal components of development. This improves accuracy of forecasting, and allows to exclude from analytical examination periods of force-majeure, that in other models has the name of surge (spike) neglecting, or ignoring of anomalous supervisions or artificial deintervention. The periods of force majeure are defined as having considerable (more than 10%) deviation between the base value of the forecast figure and the actual value as a result of temporary intervention and for them there is historically fixed motive (reason) without assuredly (probably) foreseen repetition in the future. For example, commencement of operation of the railway rapid lines and tunnel under English Channel greatly influenced on air motion on the corresponding flows, but this influencing will exist during all period of exploitation to the tunnel and to the rapid lines, that is why they can not be considered the force-majeures circumstances. Also like this, building of new runway in Borispol greatly influenced both on development of air motion at FIR Ukraine and on the redistribution of existent flows. In the case of the useful increase of investing in the recreation infrastructure of Crimea, and improving quality of service to the world standards, it is logically to expect the perceptible increase of the air motion with explicitly seasonal character. Unlike these examples, considerable reduction of figures of air motion after events on September 11 had the temporary force-majeure character, and that is why must be excluded from the analysis.

4.5 The difference discovered on this stage between a base model and actual values of the forecast figure for the analyzed interval of time is used in the next stage of multiplicative stochastic correction. With the purpose of simplification of calculations and masking of effect of false autocorrelation the declarative limited autoregressive row is used. It consists of relative residuals (errors) between X and TS for the analyzed interval of time with lags equal -1, -3 and -12 months, that is determined by the formula:

$$F(m,n) = TS(m,n) \cdot \left(1 + K_{-1,0} \cdot \frac{X(m-1,n) - TS(m-1,n)}{TS(m-1,n)} + K_{-3,0} \cdot \frac{X(m-3,n) - TS(m-3,n)}{TS(m-3,n)} + K_{0,-1} \cdot \frac{X(m,n-1) - TS(m,n-1)}{TS(m,n-1)}\right)$$

Where following denotations are used:

F it is the stochastic corrected value of the forecast figure in a m-th month of n-th year

$K_{-1,0}$ {0..1} it is weighing coefficient of influencing of short-range stochastics (unknown)

$K_{-3,0}$ {0..1} it is weighing coefficient of influencing of medium-range stochastics (unknown)

$K_{0,-1}$ {0..1} it is weighing coefficient of influencing of far-range stochastics (unknown)

$$\sum K = 1$$

4.6 For compensation of the factors stated in paragraph 2.5 the unknowns are searched by an indirect least-squares method, that as such, that minimize sum of squares of relative residuals (errors) for the analyzed interval of time between the actual values of the forecast figure and values of base model of the forecast figure except for force-majeures periods. The criterion function of minimization is

$$S_F = \sum_{m,n} \left(\gamma(m,n) \cdot \frac{F(m,n) - X(m,n)}{F(m,n)} \right)^2$$

Where $\gamma(m,n)$ is sign coefficient, it equals to 0 for force-majeures periods and 1 for all others.

4.7 N.B.: Unlike the exponential smoothing and ARIMA method instead of weighted averaging with the values of the forecast figure and their actual values multiplicative correction is realized on the basis of the declarative limited autoregressive row from relative residuals (errors) between the values of the preliminary formed base nonlinear-additive model of seasonal development and actual values, i.e. in the formula of paragraph 4.5 for correction the values of TS are used for past periods not values of F.

Version 1.2

gridl@mail.com