1.4. FINANCIAL TIME SERIES PREDICTION WITH THE TECHNOLOGY OF COMPLEX MARKOV CHAINS

In this research the technology of complex Markov chains, i.e. Markov chains with a memory is applied to forecast financial time-series. The main distinction of complex or high-order Markov Chains [1] and simple first-order ones is the existing of aftereffect or memory. The high-order Markov chains can be simplified to first-order ones by generalizing the states in Markov chains. Considering the «generalized state» as the sequence of states makes a possibility to model high-order Markov chains like first-order ones. The adaptive method of defining the states is proposed, it is concerned with the statistic properties of price returns [2].

According to the fundamental principles of quantum measurement theories, the measurement procedure impacts not only on the result of the measurement, but also on the state of the measured system, and the behavior of this system in the future remains undefined, despite of the precision of the measurement. This statement, in our opinion, is general and is true not only for physical systems, but to any complex systems [3].

Nonlinear systems, in which future states depends on infinite past states are being analyzed. The analysis of above-mentioned systems is possible only in discrete and finite representation, and results of it will be initially and principally approximate, i.e. it contains endogenous uncertainty, which inherits from current system according to the quantum postulates.

With the chosen time discretization, the memory-based model can be described in the following way:

$$x(n+1) = f(x(n); x(n-1); x(n-2)...) \,. \qquad (1)$$

It's necessary to mention, that with the continuing time definition the dynamical behavior of the memory-based model is unable to be represented with some trajectory on the finite-dimensional phase space.

In order to quantify uncertainties in real complex socio-economical systems the probabilistic models are used. However, the usage of probabilistic models is based on the controversial hypotheses, so statistical interpretation of the results is not informative enough, and its results are not corresponding to the real systemic processes. In particular, the 1/*f*-noize problem [4] is widely connected with the existence of long memory in complex systems. From the statistic's point of view it means the absence of the mean value in time series as a limit, when the time window approaches infinity, for any processes in complex systems. So such processes cannot be statistically explained [2].

Modern conceptions in complex system's modeling

The new approaches in complex system's dynamics simulation and prediction are based on the usage of determined chaos and neural-networks technologies [5-7]. The exploration and realization became possible only with appearance of powerful modern computers. The common feature of these technologies is a usage of recurrent computational process:

$$x_{n+1} = f_n\left(f_{n-1}\left(\ldots\left(f_1\left(x_1\right)\ldots\right)\right)\right), n = 1,2,\ldots, \qquad (2)$$

where $f_i(x_i)$ is nonlinear mapping for vector x_i, i is a discrete or real or modeled time. To identify the model (2) means to define parameters of nonlinear function $f_i(x_i)$, the distinctions between determined chaos and neural networks models connected with the type of the function and parameter estimation methods. Convergence of the process (2) in general is not required. In general case a function can take either single-moment vector component's values x_i, or dynamics it's changes in time.

It is possible to convert a particular model (1) to the type of more general model (2) with the help of lag variables addition into the model (1).

Both deterministic (described by integro-differential equations) and stochastic processes (complex Markov chains [1] belongs to it), can be reviewed as particular cases of the determined chaos models of type (2). With time discretization Δt approaches zero, if such a limit exists, the model converges to classical integro-differential equations. With finite Δt it is models with discrete time, which can generate future value's sets in corresponding phase space, also including a lag variables. These sets can be either measurable (discrete or continuous) sets, that accept probabilistic interpretation, or immeasurable sets – fractals [8], for which such an interpretation is in principle unacceptable.

The prominent examples for determined chaos models, acceptable for probabilistic interpretation, are different pseudo random-number generators, which are widely used in simulation modeling. It's necessary to mention, that no exact procedures exist, which can differ «real» random sequence from pseudo-random one. Indeed, any finite «random» sequence definitionally is not random because of its finiteness, and any «non-random» finite sequence may be regarded as one of the possible, but very rare, subsets from real infinite random sequence.

Discrete Markov process $X(t)$ of order $r \geq 1$ with discrete time t, (the Complex Markov Chain of order $r \geq 1$), is defined as conditional probability [1]:

$$p(x_s,t_s/x_{s-r};\ ..,x_{s-1},t_{s-1})=p(x_s,t_s/x_1,t_1;..,\ x_{s-1},\ t_{s-1}). \quad (3)$$

This condition should be fulfilled for any discrete moments of time $t_1<t_2<\ldots<t_r<t_s$. (the tuple (x_i,t_i) is considered as a state $(X(t_i)=x_i)$. Both simple Markov chain (r = 1), and Complex Markov Chain (r > 1) is defined by the distribution of transition probabilities $p(x_s,t_s/x_{s-r};\ ..,x_{s-1},t_{s-1})$ (the conditional probability). This distribution depends on r last states and the distribution of r-th state (unconditional probability):

$$p(x_{s-r}\ ...,x_{s-1},t_{s-1})=P\{(X(t_{s-r})=x_{s-r}),\ \ldots(X(t_{s-1})=x_{s-1})\} \quad (4)$$

(time moments $t_1,t_2,\ldots t_s$ are regarded as discrete integer parameters).

The main distinction of complex high-order Markov chains from simply first order ones is the existence of the aftereffect (memory), because the future state of the system (x_p, t_p) depends not only on the current state (x_q,t_q) (simple Markov chain), but also on sequence of r - 1 past states ($x_{q-r+1},t_{q-r+1};\ldots x_{q-1},t_{q-1}$), $t_{q-r+1}<...<t_{q-1}<t_q<t_p$ in the complex Markov chain. We can simplify the complex Markov chains of order r to simple ones (of order r = 1) by generalizing the state of the system. We consider the «general state» as the sequence of r past states [2].

The technology, which is proposed in the current work, is similar to neural-networks and is based on the following terms:

1. The process has an aftereffect and is generated by some «hidden» model of determined chaos. Classical random and determined processes are regarded as partial cases of more general model.

2. Input data for a model of prediction is only the discrete points of researched value of the system. The time interval of the discretization is constant. This data definitionally is finite and therefore is restricted.

3. We use the quantized discrete relative differential of the input time series. This differentials are counted with certain time steps, that is congruous to the input time-series discretization time interval (The input discretization time interval is considers a unity time interval).

4. The conditional probabilities of the one-step transitions are counted, considering the Markov chain is stationary at the given time-series.

5. We take the difference with the maximum likelihood at the each step as a prediction, and at the next step we consider this probability equals to unity (process is considered as determined one).

6. The optimal choice of hierarchy of time discretizations and the parameters for each discretization time interval (Markov chains order, or memory length, the number and the characteristics of states in the Markov chain) is evaluated with the genetic and learning approach, similarly to neural-networks technologies.

The terms 1-6 should be regarded as conditions, we can prove it only in the set of the experimental researches. Really this postulates a new procedure of indirect measurement, which is based on the current discretized input time-series, the result of this measurement is the prediction as the one of the possible scenarios of the system behavior in the future (the sequence or the vector of predicted time series).

Conceptually this approach may be proved by some analogy with the properties and dynamic and behavior of the quantum-mechanical systems.

The prediction algorithm

The prediction algorithm consists of the following steps.

1. Evaluation the set of time discretization intervals ($t_{\min} \le t \le t_{\max}$), relating to the hierarchy of time steps Δt=1,2,4,8,16,32…$\Delta t_{\max}$, $\Delta t_{\max}= 2^k$, or more complex hierarchies.

2. Chose the number of quantized levels s for the differences (i.e. the number of elementary states for Markov Chains) and coding (discretizing) the differences for every Δt, optimizing the distribution between the states to be uniform.

3. For every discretization time interval Δt and number of quantized states s we estimate the transition probabilities between the states for Markov chains of order r = 1,2,3,4,… and evaluation of transition probability matrices.

4. Doing a prediction for triple (Δt, r, s) and for the last state t_{beg}, $t_{\text{beg}} \le t_{max}$-$\Delta$t using the state with maximum probability at each step.

5. Recurrent conjunction of prediction series of different discretization time intervals Δt in a single time-series.

6. Estimate the optimal parameter values s and r for every Δt.

7. Doing a final prediction using above-mentioned procedures and optimal parameters *s* and *r*, estimated at step 6.

8. Conjunction of resulting time-series with a zero order Markov Chain series. We consider linear trend with sin? as the zero-order series (the function $y = ax + b + \sum_{i=1}^{b} \left(c_i \sin\left(d_i + e_i\right)\right)$). The coefficients of this function are estimated by nonlinear least squares method.

Experimental results and algorithm testing

Based on above-mentioned algorithm, the computer program is created in Matlab environment. Parameters of Complex Markov Chains were automatically estimated through experiments on learning data set. The discretization time step hierarchies $\{\Delta t_i\}$ of two types are used. The simple one is similar to discrete Furier transform and is a set of $\Delta t_i = 2^i$ and the complex one is a natural number's powers productions $\Delta t_i = \prod_{i=1}^{n} p_i^{s_i}$ · and has more wide net of time steps. The time discretizations hierarchy gives a possibility to review long-memory properties of the series without increasing the order of the Markov chains, to make prediction on the different frequencies of the series.

The algorithm was tested on the following time series:

1) on regular dependences of type:

$x_n = a\sin(bn)\exp(-dn) + c$; n=1,2,… (5)

(discrete sin oscillations with different frequencies, time discretizations, with exponential fade out and without it);

2) on time series, generated by discrete model «predator-prey»:

$$\begin{cases} x_{n+1} = x_n\left(1 + \alpha\left(1 - x_n - y_n\right)\right) \\ y_{n+1} = y_n\left(1 - \gamma + \beta x_n\right) \end{cases} \qquad (6)$$

with parameter values α=3,55; β=2,1623; γ=0,8, which causes likely chaotic regime;

3) on real financial time series including EUR/USD Forex course, the World's stock's indices, including Dow Jones, S&P 500, DAX, FTSE, RTS, PFTS and others.

The results of experiments and its analysis give possibilities to make the following conclusions.

1) Replacing the initial time series with its first value and a quantized differences sequence (straight procedure) causes losses in precision because of quantification errors and its cumulating while difference summarizing in inverse transformation procedure. However surplus data representation with time discretization hierarchy and inverse transform procedure can essentially reduce quantification errors. For sine (and all periodical) oscillations the reasons of errors are discretization time step's incoherence with oscillation periods Δt, which causes «pulsation» effects.

2) The prediction quality increases with Markov chain's order r, however while learning set's length is limited, the quality growth is also limited. It is probably caused:

- by reducing the number of chains, for every transition probability and increasing a correlation between them (what is equivalent of it's number reducing because of averaging procedure);

- by chain identification of error number increasing, because of definitely approximate character of state quantification and chain's identification.

3) It's possible to generate two or more possible scenarios, while probability distribution has two similar mode values. The corresponding fork points at the predicted curve can be regarded as possible process bifurcation points.

4) For «predator-prey» models a prediction with Markov chain's order $r = 2$ causes better quality, than a prediction with $r = 3$ (Fig.), what can be explained by model's simplicity and absence of «long» memory (value of x_{n+1} is determined by the values x_n and x_{n-1}). In this case increasing of r does not cause prediction quality to increase, but it can cause the influence of negative factors, described in 2).

Conclusion.

The new prediction technology, similar to neural-network ones is proposed for complex financial system's simulation. The algorithm and its program realization was developed and tested on artificial and real time series. The prediction results for stock indices

S&P 500, DAX, FTSE are reviewed. The results demonstrate the algorithm's ability to predict financial time series and prospect of further researches in the proposed field.

References

1. Tikhonov V.I., Mironov V.A. Markov processes. - Moscow.: Soviet Radio, 1977. - 488 p.

2. Saptsin V., *Experience of application genetically complex Markov chains for neural networks technology prediction,* Visnyk Krivoriz`kogo ekonomichnogo institutu KNEU, Kriviy Rig, **2**(**18**), 56-66 (2009).

3. Saptsin V., Soloviev V. Relativistic quantum econophysics – new paradigms in complex systems modeling // arXiv:physics/0907.1142 [physics.soc-ph],7 Jul 2009.

4. Bookinham M. Noizes in electronic devices and systems. - Moscow., Mir, 1986.

5. Wasserman P. D., *Neural computing: theory and practice* (Van Nostrand Reinhold, New York, 1989).

6. Surovcev I. S., Klyukin V. I. and Pivovarova R. P., *Neural networks* (VGU, Voronezh, 1994) [in Russian].

7. Ezhov A. A. and Shumskiy S. A., *Neurocomputing and his applications in an economy and business* (MIFI, Moscow, 1998) [in Russian].

8. Mandelbrot B., *The fractal geometry of nature* (Freeman, San Francisco, 1982).

9. Rabiner R L (1989), A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proceedings of the IEEE, Vol. 77(2), pp. 257-286.

10. Weigend A. S., Gershenfeld N.A. Time Series Prediction: Forecasting the Future and Understanding the past. Addison-Wesley, 1993.

11. Zhang Y. «Prediction of Financial Time Series with Hidden Markov Models», in School of Computer Science, vol. Master of Applied Science: Simon Fraser University, 2004

12. Soloviev V., Saptsin V. and Chabanenko D. *Prediction of financial time series with the technology of high-order Markov chains,* Working Group on Physics of Socio-economic Systems (AGSOE).-Drezden, 2009, URL http://www.dpg-verhandlungen.de/2009/dresden/agsoe.pdf

Appendix.

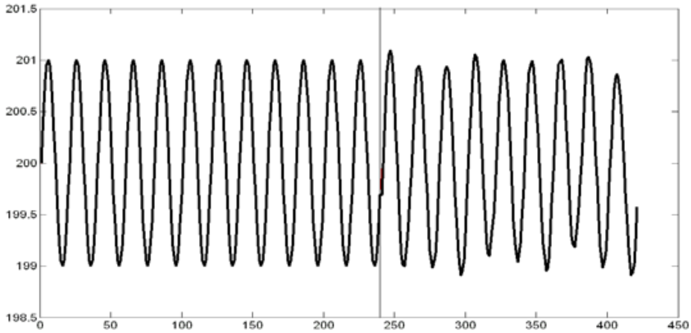

Fig. 1. Prediction of a function y=sin$\left(\frac{2\pi}{20}t\right)$ with parameters k=5, s=9

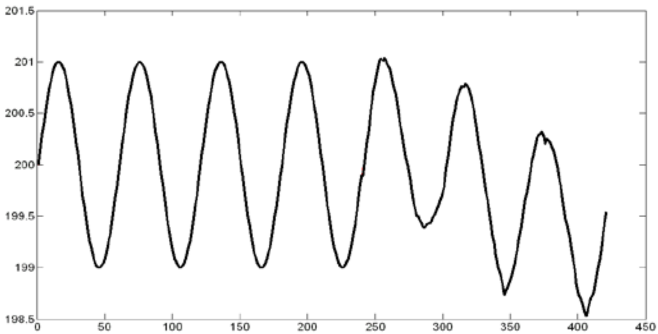

Fig. 2. Prediction of y=sin$\left(\frac{2\pi}{60}t\right)$ with parameters k=2, s=9

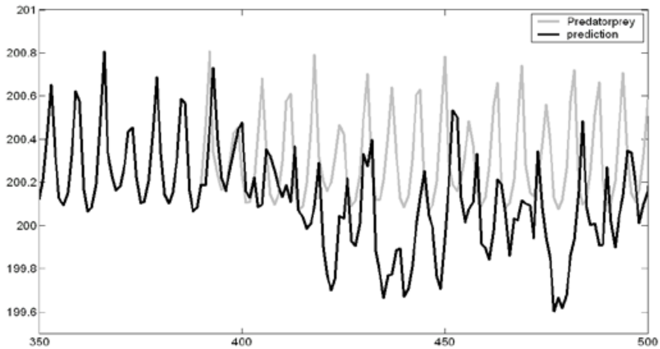

Fig. 3. Prediction of the series from «Predator-Pray» model, *s*=5

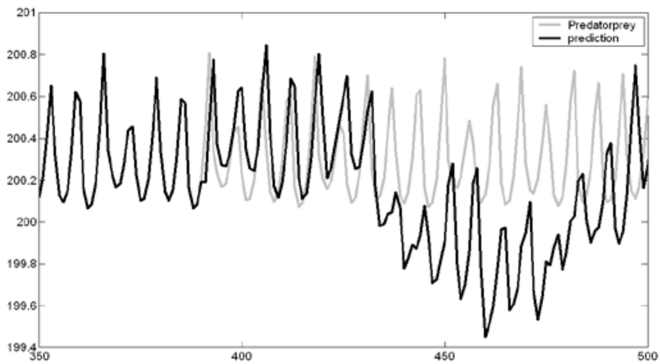

Fig. 4. Prediction of the series from «Predator-Pray» model, $s=\sqrt{\Delta t}$

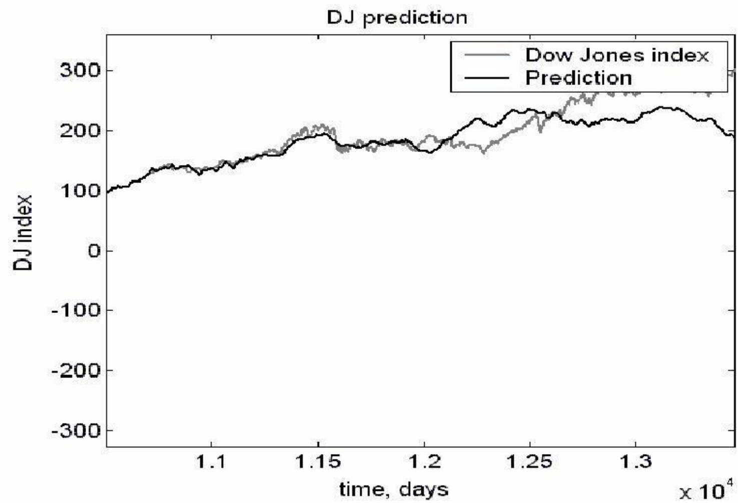

Fig 5. Prediction of Dow Jones index (1940-1953)

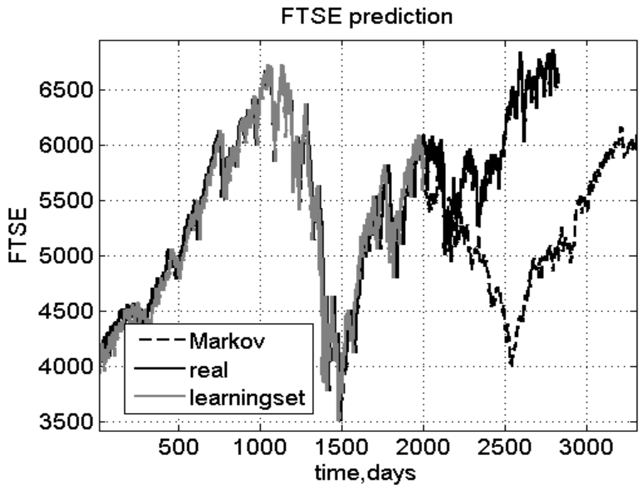

Fig 6. Prediction of index FTSE (Great Britain). March 24, 2011