Vol. 1, No. 2, 2016

ADVANCED MODELING AND FORECASTING OF POLLUTANT CONCENTRATIONS TEMPORAL DYNAMICS IN THE ATMOSPHERE OF AN INDUSTRIAL CITY (GDANSK REGION)

Alexander Glushkov[1*], Timur Safranov[2], Olga Khetselius[1], Georgy Prepelitsa[1],Yuliya Bunyakova[3], Nataliya Bykowszczenko[4]

[1] Department of Applied Mathematics and International Centre of New Mathematical Methods of Modelling Environment, Odessa State Environmental University,
15, Lvivska Str., 65016 Odessa, Ukraine, glushkovav@gmail.com , odeku.intsci@gmail.com
[2] Department of Applied Ecology and Environmental Protection,
Odessa State Environmental University,
15, Lvivska Str., 65016 Odessa, Ukraine, safranov@ukr.net
[3] Department of Economics of Environment, Odessa State Environmental University,
15, Lvivska Str., 65016 Odessa, Ukraine,
[4] Institute of Chemistry and Environmental Protection, Technical University of Szczecin,
Piastów 42 av., 71-065, Szczecin, Poland

Received: 22.04.2016

Abstract. In the paper we present the results of an advanced investigation of dynamics of variations of the atmospheric pollutants (sulphur dioxide) concentrations in the air basins of Polish industrial cities (Gdansk region) by using the improved non-linear prediction and chaos theory methods.

Chaotic behavior of the sulphurous anhydride concentration time series at two sites in the city of Gdansk has been computed. As usually, to reconstruct the corresponding chaotic attractor, it is necessary to determine time delay and embedding dimension. The former is determined by the methods of autocorrelation function and average mutual information, and the latter is calculated by means of the correlation dimension method and algorithm of false nearest neighbours. Further, the Lyapunov exponents' spectrum, the Kaplan-Yorke dimension and the Kolmogorov entropy and other invariants are calculated. An existence of a low-D chaos in the cited system is confirmed and using polynominal algorithm with neural networks block allows making an improved short-term forecast of the atmospheric pollutant fluctuations dynamics.

Key words: air pollution dynamics, studying and forecasting, chaos theory methods.

1. Introduction

Many studies in different fields of science have appeared where the methods of a chaos theory were used to a great number of various dynamical systems [1–35]. Especial interest atytracts its usinf under solving different problems in the Earth and environmental science as athe most of dynamical characteristics of environmental, hydrometeorological and ecological systems manifest typically non-linear chaotic behaviour. Nevertheless the studies concerning this behaviour in the time series of atmospheric constituent concentrations are sparse, and their outcomes are ambiguous [1–3]. The key problem of the Earth and environmental science is carrying out a forecasting model deals with the known problems. In this essense the methods of dynamical system and chaos theories could be especially useful. Let us remind that although a chaos theory puts fundamental limits for a long-rage forecasting [1–8], at the same time it can be used in order to obtain quite effective short-term prediction .

In ref. [5] the NO_2, CO, O_3 concentrations time series analysis was analysed and it was noted that O_3 concentrations in Cincinnati (Ohio) and Istanbul are evidently chaotic, so it is required essentially non-linear modeling the corresponding series [6]. In ref. [5] there is an analysis of the NO_2, CO, O_3 concentrations time series in Gdansk region and evidence of chaos has been definitely received. More over, a short-term forecast of atmospheric pollutants using a non-linear prediction method has been given. These studies have proved that non-linear methods of chaos theory and dynamical systems with satisfactory accuracy can be used for short-term forecasting the temporal dynamics of atmospheric pollutants concentrations, though a prediction model should be maken more exact. It is important to note that the time series are however not always chaotic, and thus chaotic elements should be found for each series.

In the paper we present the results of an advanced investigation of dynamics of variations of the atmospheric pollutants (sulphur dioxide) concentrations in the air basins of Polish industrial cities (Gdansk region) by using the improved non-linear prediction and chaos theory methods. It is presented an advanced analysis and forecasting of chaotic behaviour in the sulphurous anhydride concentration time series at two sites in the city of Gdansk. All calculations are performed with using "Geomath" and "Quantum Chaos" computational codes [3,5,9–39].

2. Testing of chaos in time series

2.1. Data

We have re-analysied the EU monitoring system data for time series concentrations of sulfur dioxide (SO_2) for the city of Gdansk (2003–2004). The multiyear hourly values of the corresponding concentrations (one-year total of 20×8760 data points) were studied. The examples of the time series of SO_2 concentration (in mg/m^3) are listed in Fig. 1.

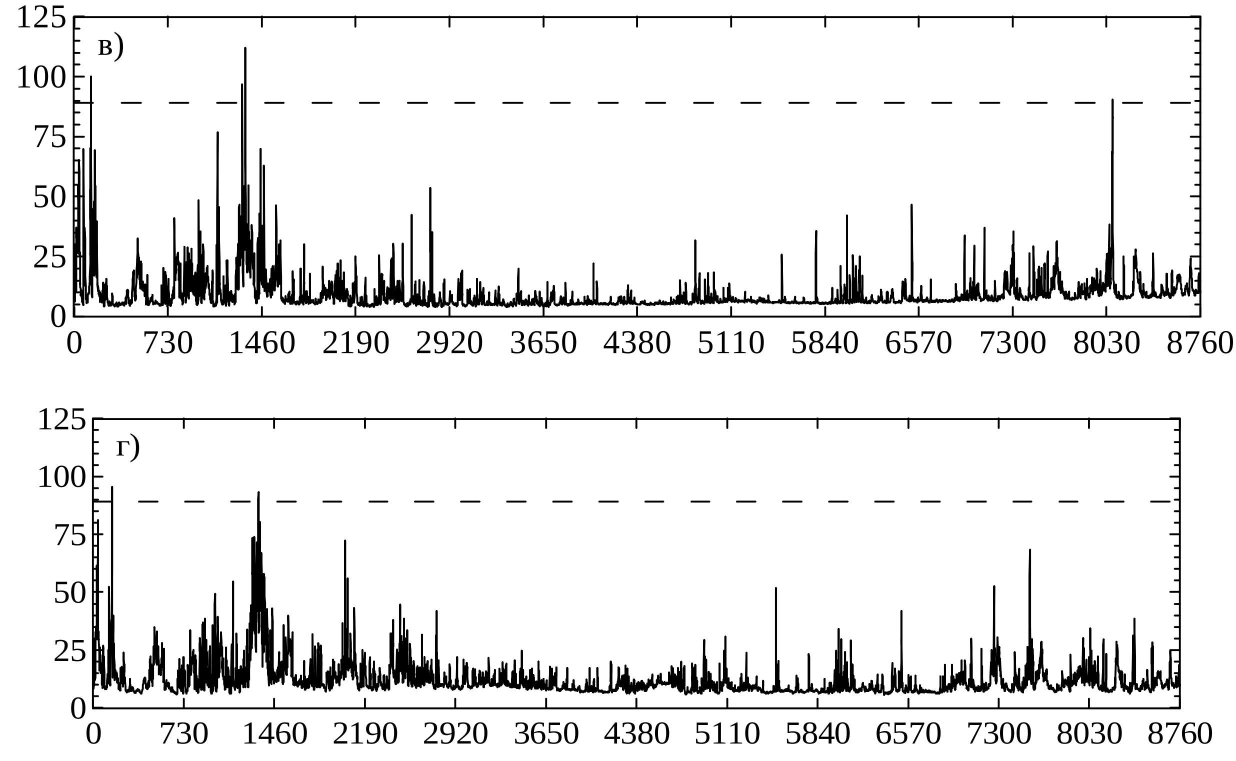

Fig. 1. Time series of SO_2 concentrations (in mg/m^3)

2.2. Testing of chaos in time series

Testing for chaos procedure in time series is described in details in [2, 14–18]. Here we are limited only by the key aspects. AS usually, we consider time series:

$$s(n)=s(t_0+ n\Delta t) = s(n),$$

where t_0 is a start time, Δt is a time step, and n is a number of the measurements. The valuable $s(n)$ in our case means an atmospheric pollutant concentration. The next step is in reconstruction of a phase space using the information contained in $s(n)$. Such reconstruction leads to a set of d-dimensional $\mathbf{y}(n)$-vectors for each scalar measurement of the atmospheric pollutant concentration. The main idea is the direct use of variable lags $s(n+\tau)$, where τ is some integer to be defined, which determines the coordinate system where a structure of orbits in phase space can be restored using a set of time lags to create a vector in d dimensions,

$$\mathbf{y}(n)=[s(n),s(n + \tau),s(n + 2\tau),..,s(n +(d–1)\tau)],$$

the required coordinates are provided. In a nonlinear system, $s(n + j\tau)$ there are some unknown nonlinear combinations of the actual physical variables. The space dimension d is the embedding dimension d_E.

2.3. Time lag

The choice of a proper time lag is important for subsequent reconstruction of phase space. If τ is too small, then the coordinates $s(n + j\tau)$, $s(n + (j+1)\tau)$ are so close to each other in numerical value that they cannot be distinguished from each other. If τ is too large, then $s(n+j\tau)$, $s(n+(j+1)\tau)$ are completely independent of each other in a statistical sense. If τ is too small or too large, then the correlation dimension of attractor can be under-or overestimated. So, the optimal valus has some intermediate position between the above listed cases.

The first well knowjn approach to computing time lag is provided by a linear autocorrelation function $C_L(\delta)$ method. The main idea is to determine such time lag in which $C_L(\delta)$ is the fastest when passing through 0.

Another alternative approach is to use a nonlinear concept of independence, e.g. the average mutual information algorithm. The average mutual information I of two measurements a_i and b_k is symmetric and non-negative, and equals to 0 if only the systems are independent. The average mutual information between any values a_i from system A and b_k from B is averaged over all possible measurements of $I_{AB}(a_i, b_k)$. In ref. [4] it is suggested to choose such value of τ where the first minimum of $I(\tau)$ occurs.

2.4. Embedding dimension

The goal of the embedding dimension determination is to reconstruct Euclidean space R^d large enough sothat the set of points d_A can be unfolded without ambiguity. In other words, we can choose a fortiori large dimension d_E, e.g. 10 or 15, since the previous analysis provides us prospects that the dynamics of our system is probably chaotic. The correlation integral analysis is one of the widely used techniques to investigate chaos in time series. The analysis uses the correlation integral, $C(r)$, to distinguish between chaotic and stochastic systems. If the time series is characterized by an attractor, the correlation integral $C(r)$ is related to the radius r as

$$d = \lim_{\substack{r\to 0 \\ N\to\infty}} \frac{\log C(r)}{\log r},$$

where d is a correlation exponent. If the correlation exponent attains saturation with an increase in the embedding dimension, then the system is generally considered to have a chaotic dynamics. The saturation value of the correlation exponent is defined as the correlation dimension (d_2) of the attractor (see [3,7]).

In Fig. 2 we list the computed dependence of the correlation integral C (r) of radius r for different embedding dimensions d for SO_2 (b) at site 6 of Gdansk in 2003.

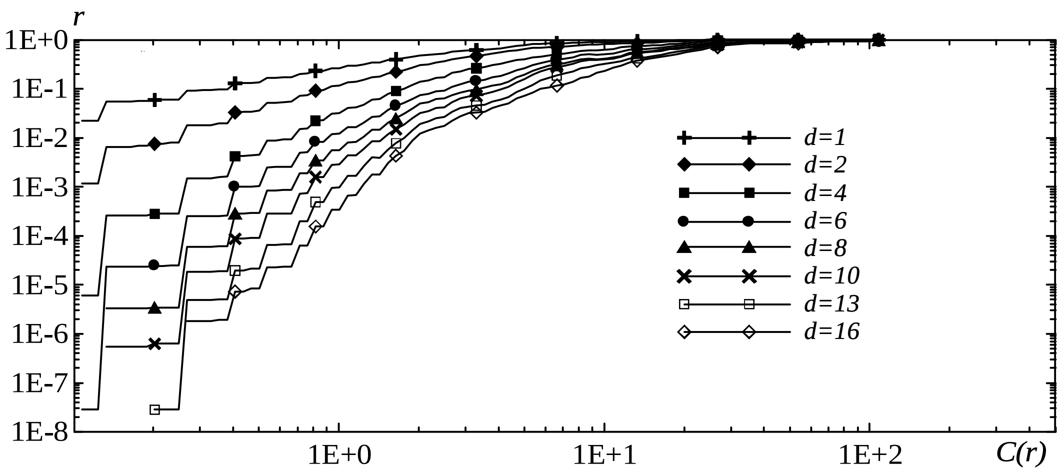

Fig. 2. The dependence of the correlation integral C (r) of radius r for different embedding dimensions d for SO2 (b) at site 6 of Gdansk in 2003–2004

3. Results for atmospheric pollutant time series

3.1. Results

Table 2 summarizes the results for the time lag calculated for first 10^3 values of time series.

The autocorrelation function for all time series remains positive. The values, where the autocorrelation function first crosses 0.1, can be chosen as τ, but in [1] it is shown that an attractor cannot be adequately reconstructed for very large values of τ. So, before making up a final decision we calculate the dimension of attractor for all values in Table 1. The outcomes explain not only inappropriate values of τ, but also shortcomings of the correlation dimension method [7]. If algorithm [1] is used, then a percentage of false nearest neighbours is comparatively large in the case of large τ. If time lag is determined by average mutual information, then algorithm of false nearest neighbours provides $d_E = 6$ for all air pollutants.

Table 1

Time lags (hours) subject to different values of C_L, and first minima of average mutual information, I_{min1}, for the time series of SO_2 at the sites of Gdansk

	$C_L = 0$	$C_L = 0,1$	$C_L = 0,5$	I_{min1}
	Site 6			
SO_2	–	233	13	19
	Site 9			
SO_2	–	148	27	18

3.2. Nonlinear prediction model

First of all, it's important to define how predictable a chaotic system is. The predictability can be estimated by the Kolmogorov entropy, which is proportional to a sum of positive Lyapunov exponents.

The spectrum of the Lyapunov exponents is one of dynamical invariants of non-linear system with chaotic behaviour. The limited predictability of the chaos is quantified by local and global Lyapunov exponents, which can be determined from measurements. The Lyapunov exponents are related to the eigenvalues of the linearized dynamics across the attractor. Negative values show stable behaviour while positive values show local unstable behaviour.

For chaotic systems, that are both stable and unstable, the Lyapunov exponents indicate the complexity of the dynamics. Large positive values determine some average prediction limit. Since the Lyapunov exponents are defined as asymptotic average rates, they are independent of the initial conditions, and hence the choice of trajectory, and they do comprise an invariant measure of the attractor. The estimate of this measure is the sum of positive Lyapunov exponents. The estimate of the attractor dimension is provided by the conjecture d_L and the Lyapunov exponents are taken in descending order. The dimension d_L gives values close to the dimension estimates discussed earlier and is preferable when estimating high dimensions. To compute the Lyapunov exponents, we use a method of linear fitted maps [3], although the maps with higher order polynomials can be used too.

3.3. Nonlinear model

Nonlinear model of chaotic processes is based on the concept of a compact geometric attractor on which observations evolve plus (so-called chaos geometric approach). Since an orbit is continually folded back on itself by dissipative forces and the non-linear part of dynamics, some orbit points $y^r(k)$, $r = 1, 2, .., N_B$ can be found in the neighbourhood of any orbit point $y(k)$, and points $y^r(k)$ arrive in the neighbourhood of $y(k)$ at quite different times k.

One can then choose some interpolation function, which accounts for whole neighbourhood of phase space and how they evolve from near $y(k)$ for the whole set of points near $y(k+1)$. The implementation of this concept is to build parameterized non-linear function $F(x, a)$ which transforms $y(k)$ into

$$y(k+1) = F(y(k), a)$$

and uses various criteria to determine parameters of a. Since one has the notion of local neighbourhoods, one can build up one's model by processing neighbourhood by neighbourhood and, after piecing these local models together, produce a global non-linear model that captures much of the structure in an attractor itself.

3.4. Short-term forecast of atmospheric pollutant time series

Table 2 shows the calculated parameters: correlation dimension (d_2), embedding dimension (d_E), Kaplan-Yorke dimension (d_L), two first Lyapunov exponents, $E(\lambda_1,\lambda_2)$, and the average limit of predictability (Pr_{max}, hours) for time series of the SO_2 at the sites of Gdansk (in 2003-2004).

Firstly, it should be noted that presence of two (from the six) positive λ_i suggests that the system broadens in the line of two axes and converges along the rest four axes in the six-dimensional space. The time series of SO_2 at site 10 have the highest predictability (more than 2 days), and other time series have the predictabilities slightly less than 2 days.

Table 2

The correlation dimension (d_2), embedding dimension (d_E), first two Lyapunov exponents, $E(l_1,l_2)$, the Kaplan-Yorke dimension (d_L), and the average limit of predictability (Pr_{max}, hours) for time series of SO_2 at the sites of Gdansk (in 2003)

	τ	d_2	d_E	λ_1	λ_2	d_L	Pr_{max}	K
	Site 1							
SO_2	19	1,58	6	0,0164	0,0066	5,01	43	0,71
	Site 2							
SO_2	17	3,40	6	0,0150	0,0052	4,60	49	0,73

The concrete example is presented in Figure 3, where the empirical (solid bold line) and theoretical forecasting (solid thin line by the polynomial-type prediction algorithm with neural networks block and dotted line by the standard polynomial-type algorithm [5, 17–19]) concentration lines SO_2 (for the one hundred points) are presented. As one can see, despite the fact that almost all the peaks on the actual curve repeat, as forecasted, the difference between the forecasted and actual data in the case of high concentrations of the ingredients can be quite large. In a whole, the results of this forecast can be considered as very satisfactory.

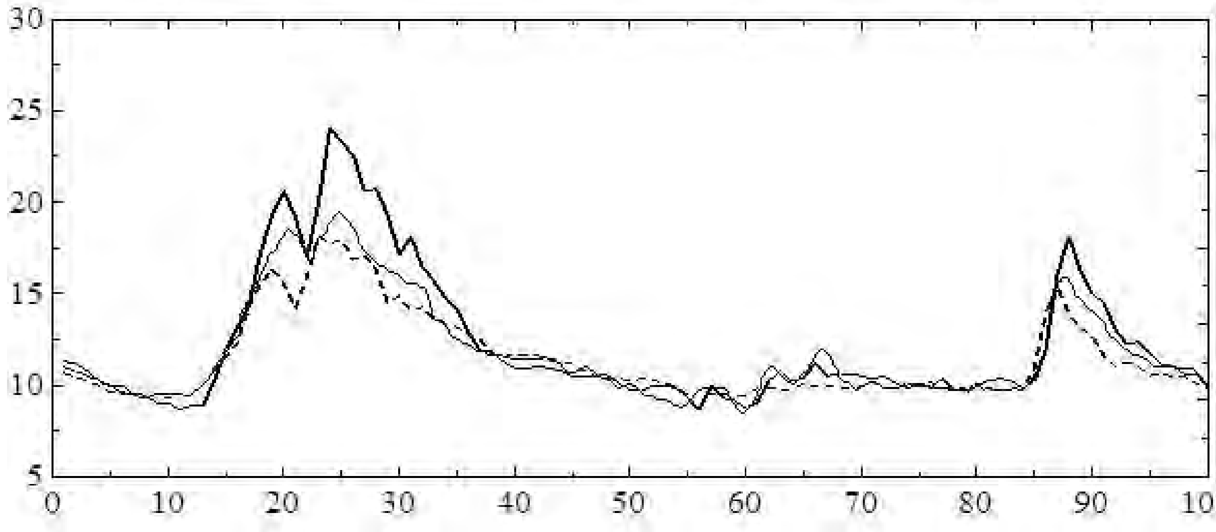

Fig. 3. The empirical (solid bold line) and forecasted (solid thin line and dotted line) curves of SO_2 concentration for the last six hundred members of time series shown in Fig. 1. Axis X – a serial number of the term (see text)

4. Conclusions

In this work, we have studied the dynamics of variations of the atmospheric pollutants (sulphur dioxide) concentrations in the air basins of Polish industrial cities (Gdansk) by using non-linear prediction and chaos theory methods. Chaotic behaviour in the sulphurous anhydride concentration time series at a few sites of Gdansk is numerically investigated. Usually, to reconstruct the corresponding attractor, the time delay and embedding dimension are needed. The former is determined by the methods of autocorrelation function and average mutual information, and the latter is calculated by means of the correlation dimension method and the algorithm of false nearest neighbours. Further, the Lyapunov exponents' spectrum, the Kaplan-Yorke dimension and the Kolmogorov entropy are calculated. The improved results on the short-term forecast of the SO_2 atmospheric pollutant time fluctuations dynamics in Gdansk region are given.

References

[1] Kennel M., Brown R., Abarbanel H.: Phys. Rev. A., 1992, 45, 3403.

[2] Bykowszczenko N., Baranowska-Bosiacka I., Bosiacka B., etal: Pol J Environment. Stud. 2006, 15, 41.

[3] Bunyakova Yu. Ya., Glushkov A. V.: Analysis and forecast of the impact of anthropogenic factors on air basein of an industrial city. Odessa, Ecology, 2010.

[4] Gallager R. G.: Information theory and reliable communication. NY: Wiley, 1968.

[5] Khokhlov V. N., Glushkov A. V., Loboda N. S., Bunyakova Yu. Ya: Atmospheric Environment (Elsevier), 2008, 42, 7284.

[6] Chelani A. B.: Int. J. Environ. Stud, 2005, 62, 181.

[7] Grassberger P., Procaccia I., Physica D, 1983, 9, 189.

[8] Gottwald G. A., Melbourne I.: Proc. Roy. Soc. London. Ser. A. Math. Phys. Sci., 2004, 460, 603.

[9] Glushkov A. V., Khokhlov V. N., Prepelitsa G. P., Tsenenko I. A.: Optics of atmosphere and ocean, 2004, 14, 219.

[10] Glushkov A. V., Khetselius O. Yu., Kuzakon V. M., Prepelitsa G. P., Solyanikova E. P., Svinarenko A. A.: Dynamical Systems - Theory and Applications, 2011, BIF110.

[11] Glushkov A. V., Kuzakon' V. M., Khetselius O. Yu. etal: Proc. of Internat. Geometry Center, 2013, 6(3), 6.

[12] Glushkov A. V., Khetselius O. Yu., Prepelitsa G. P., Svinarenko A. A.: Proc. of Internat. Geometry Center, 2013, 6, No. 1, 43.

[13] Khetselius O. Yu., Glushkov A. V., Gurnitskaya E. P. etal: Spectral Line Shapes, Vol. 15–19th Internat. Conf. on Spectral Line Shapes (AIP), 2008, 15, 231.

[14] Khetselius O. Yu.: Spectral Line Shapes, Volume 15–19th Internat. Conf. on Spectral Line Shapes (AIP), 2008, 15, 363

[15] Rusov V. N., Glushkov A. V., Khetselius O. Yu. et al: Adv. in Space Research (Elsevier), 2008, 42(9), 1614.

[16] Glushkov A. V., Khetselius O. Yu., Loboda A. V., Svinarenko A. A.: Frontiers in Quantum Systems in Chemistry and Physics, Ser.: Progress in Theor. Chem.

and Phys., Eds. S. Wilson, P. J. Grout, J. Maruani, G. Delgado-Barrio, P. Piecuch (Springer), 2008, 18, 541

[17] Bunyakova Yu. Ya., Khetselius O. Yu.: Proc. of the 8th International Carbon Dioxide Conference (Germany), 2009, T2-098.

[18] Glushkov A. V., Khetselius O.Y u., Brusentseva S.V. *et al:* Adv. in Neural Networks, Fuzzy Systems and Artificial Intelligence. Series: Recent Adv. in Computer Engineering (Gdansk), 2014, 21, 69.

[19] Glushkov A. V., Svinarenko A. A., Buyadzhi V. V. et al: Adv.in Neural Networks, Fuzzy Systems and Artificial Intelligence, Series: Recent Adv. in Computer Engineering (Gdansk), 2014, 21, 143

[20] Khetselius O. Yu.: Sensor Electronics and Microsystem Techn., 2008, 3, 28.

[21] Khetselius O. Yu.: Photoelectronics, 2005, 14, 83.

[22] Khetselius O. Yu.: Journ. of Physics: Conf. Series (IOP, UK), 2012, 397, 012012.

[23] Khetselius O. Yu., Florko T. A., Svinarenko A. A., Tkach T. B.: Phys. Scripta (IOP), 2013, T153, 014037.

[24] Rusov V. D., Glushkov A. V., Vaschenko V. N. *et al:* Journ. of Atm. and Solar-Terrestrial Phys., 2010, 72, 498.

[25] Glushkov A. V., Rusov V. D., Ambrosov S., Loboda A.: New projects and new lines of research in nuclear physics, Eds. G.Fazio, F.Hanappe (World Scientific), 2003, 126

[26] Glushkov A. V., Malinovskaya S. V.: New projects and new lines of research in nuclear physics. Eds. G. Fazio and F. Hanappe (World Scientific), 2003, 242.

[27] Glushkov A. V., Malinovskaya S. V., Gurnitskaya E. P., Khetselius O. Yu., Dubrovskaya Yu. V.: Journ.of Physics: Conf. Series (IOP), 2006, 35, 425.

[28] Malinovskaya S. V., Glushkov A. V., Khetselius O. Yu., Lopatkin Yu. M., Loboda A. V., Svinarenko A. A., Nikola L. V., Perelygina T. B.: Int.Journ. Quant.Chem., 2011, 111(2), 288.

[29] Malinovskaya S. V., Glushkov A. V., Khetselius O. Yu., Svinarenko A. A., Mischenko E. V., Florko T. A.: Int.Journ. Quant.Chem., 2009, 109(14), 3325.

[30] Glushkov A. V., Khetselius O. Yu., Malinovskaya S. V.: Molec. Phys. , 2008, 106 (9-10), 1257

[31] Sivakumar B.: Chaos, Solitons& Fractals, 2004, 19, 441.

[32] Packard N. H., Crutchfield J. P., Farmer J. D., Shaw R. S.: Phys. Rev. Lett., 1980, 45, 712.

[33] Glushkov A. V., Khetselius O. Yu., Svinarenko A. A.: Coherence and Ultrashort Pulse Laser Emission, Ed. by F. J. Duarte (InTech), 2010, 159.

[34] Abarbanel H. D. I., Brown R., Sidorowich J. J., Tsimring L. Sh.: Rev. Mod. Phys, 1993, 65, 1331.

[35] Schreiber T.: Phys. Rep., 1999, vol. 308, 1.

[36] Malinovskaya S. V., Glushkov A. V., Dubrovskaya Yu. V., Vitavetskaya L. A., Recent Advances in the Theory of Chemical and Physical Systems (Springer), 2006, 15, 301.

[37] Glushkov A. V., Khetselius O. Yu., Svinarenko A. A.: Advances in the Theory of Quantum Systems in Chem. and Phys., Ser.: Progress in Theor. Chem. and Phys., ed. by P. Hoggan, E. Brandas,J.Maruani, G. Delgado-Barrio, P. Piecuch (Springer), 2012, 22, 51

[38] Glushkov A. V., Khetselius O. Yu., Lovett L.: Advances in the Theory of Atomic and Molecular Systems: Dynamics, Spectroscopy, Clusters, and Nanostructures. Ser.: Progress in Theor. Chem. and Phys., Eds. Piecuch P., Maruani J., Delgado-Barrio G., Wilson S. (Springer), 2009, 20, 125.

[39] Glushkov A. V., Malinovskaya S. V., Sukharev D. E., Khetselius O. Yu., Loboda A. V., Lovett L.: Int. Journ. Quantum Chem.,2009, 109 (8), 1717.