

I.V. Baklan, Y.M. Selin, T.V. Shulkevich, S.V. Vyshemirskaya

## MODELING SUSTAINABLE DEVELOPMENT PROCESSES

**Abstract.** *The paper contains the description of the feasible approach to modeling and forecasting of ecological processes that are the component of the sustainable development paradigm. The suggested mathematical apparatus is based on the statistical methods and comprises hidden Markov models with the linguistic modeling.*

**Keywords:** *hidden Markov models, linguistic model, linguistic modeling.*

### Formulation of problem

We propose the mathematical apparatus that might be used for the modeling and forecasting of the environmental processes of the sustainable development paradigm. The common problem of the statistical methods is the lack of the historical information.

### Analysis of recent research and publication

According to the definition suggested by the Centre “For our common future” (“За наше спільне майбутнє”), that has been working in Geneva since 1988, and the United Nations Conference in Rio de Janeiro (1992), the sustainable development is understood as the kind of development that “satisfies the current needs, but does not compromise the abilities of the future generations to satisfy their own needs”, or that provides “the high quality of environment and the healthy economy for all nations”. Therefore, the problem of the sustainable development is the problem of the mankind salvation from the consequences of its own activity that, by the end of the 20th century, have become critical and manifested themselves in desertification, pollution of the atmosphere, oceans and soil, rapid population growth, poverty, starvation, dangerous diseases, etc. The issues of scientific knowledge and education, industry and innovation technologies, ecological, social and medical problems, international relations and political events and many other phenomena of modern life have intertwined into the single interrelated unit. The consciousness of the situation and the search of the reasonable solutions to the problem have become an urgent matter. It requires focused attention not only on a global, but also on a national, regional and local basis. [1]

Usually, the sustainable development paradigm is presented in the following way (fig.1).

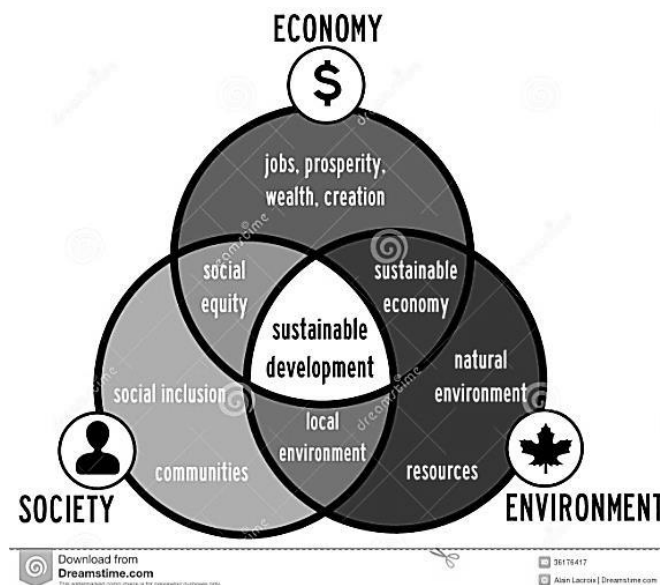


Fig. 1 The triune concept of the sustainable development

The growing rate of exploitation of natural resources, economic decline of the economies in transition raise up the risk of the technogenic disasters, keep the countries from spending considerable efforts and resources necessary for the implementation of measures to reduce the environmental impact.

To accomplish the purpose of the sustainable development the specific mechanisms for its implementation are required. The development and application of the “environmental assessment” system for the assessment of the transition to the sustainable development is of considerable significance in the world practice. The most important are the two of its modes (local and strategic) that nowadays are the well-established mechanisms of revelation and prevention of the environmental consequences of the human induced activity (refer to fig.2) [2].

However, for better understanding of the ecological component of the sustainable development paradigm, it is necessary to analyze some natural phenomena and biospheric processes and their interaction in the “man-environment” system. We will pay careful attention to environmentally hazardous processes.

The task of time series analysis in technical, medical or economical processes where the quality of decision-making to a large extent depends on the prediction of trends in the dynamic process. One of the sub-tasks of such trends is a search for the available templates.. We propose an

approach to solve this problem using a hybrid model of a special kind - the "hidden Markov model - a linguistic model.". The principles of hidden Markov models (HMM) was described Baum and his colleagues in the late 60's [6-8]. In the early 70th Baker, Dzhelinek and their colleagues first used HMM for speech recognition [9, 10]. The linguistic modeling has been for almost the 20th century, especially in the works of Markov A.A. [11,12], Apresian Yu.D.[13], Fu K.S[14]. Andrey Markov produced the first results (1906) for these processes, purely theoretically. In 1913, he applied his findings for the first time to the first 20,000 letters of Pushkin's Eugene Onegin. The main principles of linguistic modeling was described in works [15,16].

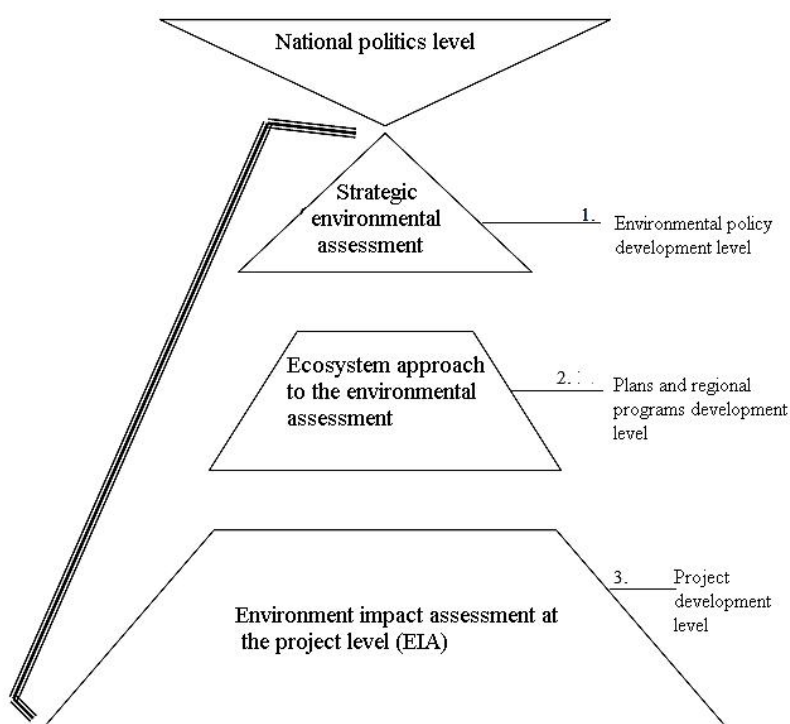


Fig.2. Environmental assessment system (EA)

### Task of investigations

Proposed to build a hybrid model of the time series based on hidden Markov models and linguistic modeling (HMM-LM). The article provides a classification of existing hidden Markov models and ways of constructing a hybrid model HMM-LM.

### Main part (the results of theoretical and experimental research)

Environmentally hazardous processes are generally understood as exogenous and endogenous (anthropogenic), short-term and long-term impacts on the ecosystem as a whole or on its particular element, leading

to the violation or problems of its performance that in its turn disrupts the man-environment balance.

1. Endogenous (anthropogenic) are caused by anthropogenic activity (pollution by harmful organic and non-organic substances of all kinds of the environment: the atmosphere, surface waters, water-saturated underground layers, soil, animate environments; noise and electromagnetic pollution of the atmosphere; light pollution; flooding of soils and slopes; ozone depletion and creating conditions for the greenhouse effect). They may be controlled and, in some cases, reduced or even prevented.

2. Exogenous are the manifestations of the outer space and internal geological processes and changes (solar cycles, hurricanes, tropical cyclones, monsoons, volcanic activity, global warming as a result of the precession of the Earth's axis, and finally, the asteroid hazard). They are purely objective, cannot be controlled, but may be observed and forecasted. Therefore, it is possible to develop a set of measures to minimize the damages or mitigate the consequences.

It becomes clear from the mentioned earlier that the environmentally hazardous processes are of various nature and types. Also, it is evident that “not everything depends on a man”.

In fig.3 there is a schematic classification of the environmentally hazardous processes. [3]

The solar activity (fig.4) is a special type of an exogenous impact on the geological and biochemical processes taking place on Earth. The index for the solar activity characteristics, so called Wolf number, Zurich number or the relative sunspot number is calculated by the following formula:

$$R = k ( f + 10g ) ,$$

where R is the Wolf number; f is the total number of sunspots on the visible sun hemisphere; g is the number of sunspot groups; k is the multiplier (<1) that takes into account the total contribution of observing conditions, the telescope type and that is observed by the standard Zurich numbers.

According to the observation data for the last 304 years, the cycle length in actual practice varies from 8.5 to 14 years between the proximate minimums and from 7.3 to 17 years between the proximate maximums. The connection between the solar cycles and the increased

occurrence of land slides has been proved. The most adverse impact is on that of the innovative technology effects, high-frequency operating devices, such as mobile communications, satellites (and also the related navigational safety problems), etc. Solar radiation is the source of the motion in the atmosphere-ocean system. The force, that the solar radiation generates, produces buoyancy – the source of motion in the atmosphere. It determines the nature of hurricanes, tornadoes, tsunamis and other destructive processes. Also, the solar activity influences the proliferation of some diseases, increases the frequency of traffic accidents, etc.

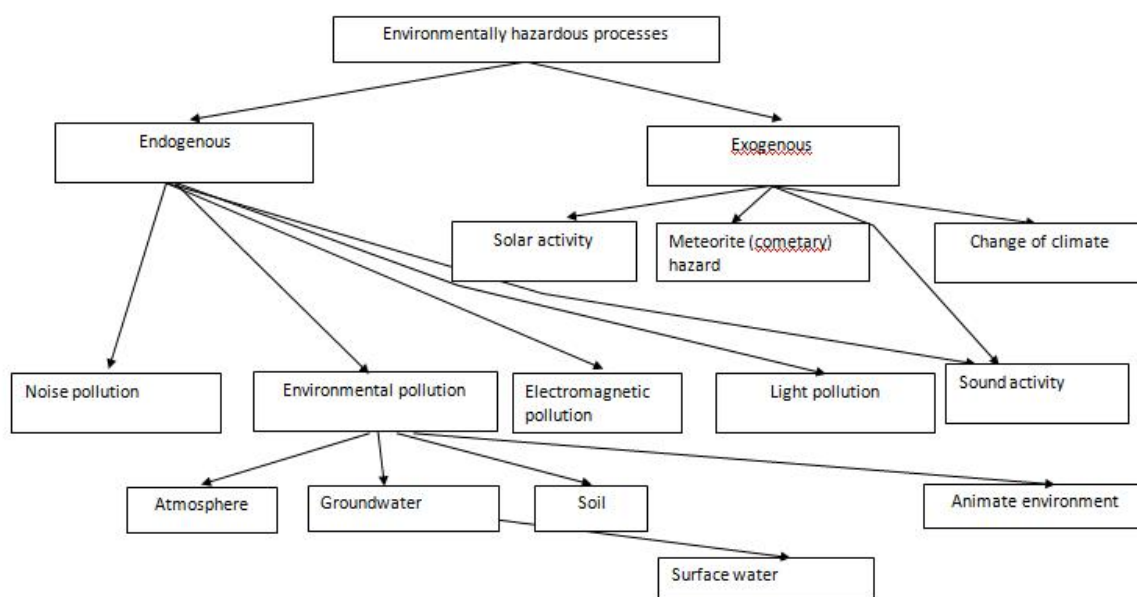


Fig. 3 Environmentally hazardous processes classification

According to the data of US National Centre for Environmental Information

(<http://www.ngdc.noaa.gov>)

The climate change caused by the global warming is not only the result of increased greenhouse gases concentrations (especially, carbon dioxide) in the atmosphere, but also the result of the next cycle of the precession of the Earth's axis, thus it has an exogenous cause and it will be intensified in the near future. Although, according to Kh.Y. Shelnhuber [4], Antarctic and Greenland's ice sheet, Sahara desert, Amazonian rainforest, Asian monsoon system, Gulf Stream and six more "weak spots" of Earth influence the climate, and in the case of at least one of them undergoing changes, the ecological disaster (of the continental level) will be inevitable, the risk may be decreased by reducing the emission rate of gases mentioned earlier.

Considering the literature [3], we can define two principal directions and accordingly, two methodological approaches to the mathematical modeling of the dynamics of the environmentally hazardous processes of various nature. The first approach comprises dynamic-computed approaches based on computational methods for solving various kinds of differential equations that describe the basic laws of physics, and also atmospheric and hydrodynamic processes. They are focused on solving the following basic problems of the most important spatiotemporal patterns of the current natural processes:

- Exposure of the current spatiotemporal interrelations between various atmospheric processes in the observation dynamics;
- Natural processes modeling for the forecasting of their development dynamics.

The second approach, comprising empirical dynamic-statistical approaches based on the use of the long-term field measurement statistics, belongs to the international system for the analysis and forecasting of the ecosystem components. They are focused on the exposure of the basic spatiotemporal patterns typical of the atmospheric processes over decades. The main purpose of these approaches is, in fact, the establishment of deep spatiotemporal correlations between various natural processes based on the long-term statistics. Depending on the purposes of the study it is necessary to perform the development of the mathematical apparatus for the analysis of the dynamics of the environmentally hazardous processes based on either dynamic-computed or dynamic-statistical approaches, taking into account the specific peculiarities and properties of the processes.

Besides, there is the third type of processes that cannot be modeled with the help of dynamic-computed methods, and due to the absence of a peculiar particular periodicity (daily, monthly, annual or another permanent periodicity) they are difficult to describe using empirical-statistical methods. Therefore, the problem of the development of the process analysis and the development of the forecasting methods of such processes for the information support of the environmental situation control and monitoring system is relevant.

To sum up, the conducted analysis allows for the conclusion that the environmentally hazardous processes are characterized by the complex interrelations, interdependencies and interactions of various factors and

causes. They have the following characteristic properties and peculiarities:

- The heterogeneousness and diversity of causes and factors and an activity that causes them;
- The spatial distribution of the triggering events, temporal and spatial uncertainty of the growth dynamics and their impact on the eco-surroundings;
- The nonstationarity of properties and ambiguity of their characteristics.

These properties and peculiarities determine the practical relevancy of studying of all the variety of characteristics, interrelations, interactions, interdependencies of the diverse factors and causes of the environmentally hazardous processes on the basis of the single systematic approach from the perspective of achieving the globally defined objective of the environmental situation management – early prevention and (or) minimization of the adverse effects of their action. However, the analysis shows that nowadays the various types of natural and technogenic environmental processes, their causes, progress, consequences and the sphere of the action are studied separately, without regard to interrelations, interdependencies, and interactions.

Such an approach does not consider some factors of primary importance that influence the active processes, their adverse impact level, the possibility and effectiveness of its prevention.

Considering all of the above-mentioned, we propose the feasible mathematical apparatus that may be used for the forecasting of the environmental processes of all three types.[5]

### **Hidden Markov models**

A hidden Markov model (HMM) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (hidden) states. In a typical Markov model is visible observer, so the probability of transitions - a single option. In HMM, we can monitor only variables that are affected by this condition. Each state has a probability distribution among all possible output values. Therefore, the sequence of characters generated HMM provides information about the sequence of states.

We now turn directly to the classification of HMM. In a broad sense HMM can be classified into the following general classes [12]:

- hidden semi-Markov model;
- classic HMM;
- hybrid Markov model.

*Hidden semi-Markov model (HSMM)* is a generalization of the HMM, which allows a more general distribution of residence time. The basic structure HSMM is shown in Figure 1. Arrows from  $\bar{S}_1$  to  $X_1, \dots, X_{d_1}$  and from  $\bar{S}_2$  to  $X_{d_1+1}, \dots, X_{d_1+d_2}$  indicate conditional dependence of observations on the hidden states.

Hidden semi-Markov model consists of a pair of random processes with discrete time  $\{S_t\}$  and  $\{X_t\}$ . As HMM the observed process  $\{X_t\}$  is associated with a hidden state semi-Markov process  $\{S_t\}$  with conditional distribution.

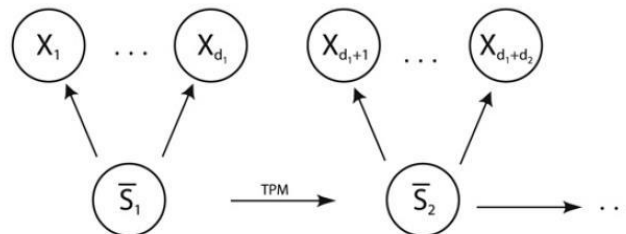


Fig. 4. Basic Structure of HSMM

*Classic HMM* divided in the following classification criteria by:

- number of states (1);
- number of levels (2);
- temporal characteristics (3);
- type of probability distribution that underlie the HMM (10);
- parametric property (11);
- type of topology (4);
- homogeneity (5);
- linearity (6);
- dynamic properties (7);
- structure of matrix transitions (8);
- connectivity graph transitions (9).

The general classification of classic HMM showed in fig.5. Number in parentheses corresponds to the diagram.



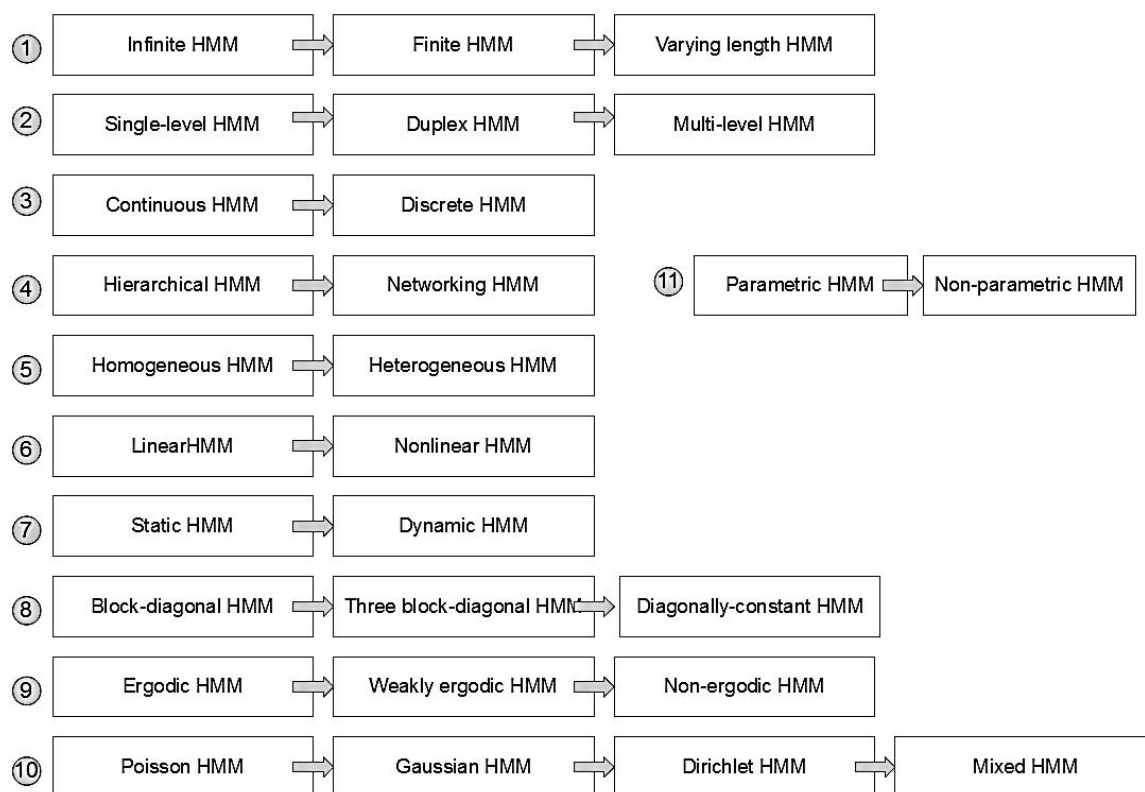


Fig. 5. The general scheme of classification of classical hidden Markov models.

### Conclusions and prospects for further research

Thus, we proposed the mathematical apparatus that might be used for the modeling and forecasting of the environmental processes of the sustainable development paradigm. The common problem of the statistical methods is the lack of the historical information. However, the mankind is engaged in the monitoring of the environment, observation of the weather and natural processes throughout the life. Therefore, the use of the statistical methods in particular is fully justified.

### REFERENCES

1. Програма дій «Порядок денний на XXI століття»: Ухвалена конференцією ООН з навколишнього середовища і розвитку в Ріо-де-Жанейро (Саміт «Планета Земля», 1992 р.): Пер. з англ. — 2-е вид. — К.: Інтелсфера, 2000. — 360 с.
2. Патока І.В., Хлобистов Є.В. Теоретичні засади дослідження екологічної безпеки розвитку і розміщення промислового потенціалу // Економіка промисловості України: Зб. наук. праць. — К.: РВПС України НАН України. — 2001. — С.29-37.

3. Селін Ю.М. Системний аналіз екологічно небезпечних процесів різної природи. Системні дослідження та інформаційні технології, 2007, № 2. – С. 22-32
4. Рамсторф Штефан, Шельнхубер Ханс Йоахим. Глобальное изменение климата: диагноз, прогноз, терапия. ОГИ, М., 2009 г., 272с.
5. Баклан І.В., Селін Ю.М., Шулькевич Т.В. Математичні моделі прогнозування часових рядів різної природи // Вестн. Херсонського національного техн. ун-та. - Херсон: ХНТУ, 2014. - Вып. 3 (50). – С.213-218.
6. Baum L. E. and Petrie T., “Statistical inference for probabilistic functions of finite state Markov chains,” Ann. Math. Statist., vol. 37, pp. 1554–1563, 1966.
7. Baum L. E. and Eagon J. A., “An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology,” Bull. Amer. Math. Soc., vol. 73, pp. 360–363, 1967..
8. Baum L. E. and Sell G. R., “Growth transformations for functions on manifolds,” Pacific J. Math., vol. 27, no. 2, pp. 211–227, 1968.
9. Baker J. K., “The DRAGON system—An overview,” IEEE Trans. Acoust., Speech Signal Processing, vol. ASSP-23, pp. 24–29, Feb. 1975
10. Baker, J. (1979). Trainable grammars for speech recognition. In Speech communication paper presented at th 97th Meeting of the Acoustical Society of America, pp. 547–550 Boston, MA.
11. Markov A.A. "Распространение закона больших чисел на величины, зависящие друг от друга". Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete, 2-ya seriya, tom 15, pp. 135–156, 1906. (in russian)
12. Markov A.A.. "Extension of the limit theorems of probability theory to a sum of variables connected in a chain". reprinted in Appendix B of: R. Howard. Dynamic Probabilistic Systems, volume 1: Markov Chains. John Wiley and Sons, 1971.
13. Apresyan, Yu. D. The ideas and methods of modern structural linguistics (short essay). MA: Education, 1966. (in russian)

14. Fu K. S., Syntactic Pattern Recognition and Stochastic Languages, in Frontiers of Pattern Recognition, ed. S. Watanabe, Academic Press, 1972.
15. Baklan I.V. Linguistic modeling: foundation, methods, some applied aspects // System technologies. N 3(74) – Dnipropetrovsk, 2011. – P.10-19. (in Ukrainian)
16. Baklan I.V. Linguistic modeling for solving of different problems // Intellectual System for Decision Making and Problems of Computational Intelligence: Conference Proceeding. – Kherson: KNTU, 2013. – P.398-400.