

Дослідження спрямоване на оцінювання та прогнозування передачі радіонуклідів із ґрунту у рослини на основі реальних даних. Інструментом моделювання є динамічна мережа Байєса. Фактори впливу проаналізовані разом із можливостями застосування даного підходу до розв'язання такого класу задач. Модель дозволяє будувати довготермінові сценарії для визначення шляхів розвитку сільського господарства на територіях, уражених внаслідок аварії на ЧАЕС

Ключові слова: радіоактивне зараження, динамічна мережа Байєса, імовірнісний висновок

Исследование направлено на оценивание и прогнозирование передачи радионуклидов из почвы в растения на основе реальных данных. Инструментом моделирования является динамическая сеть Байеса. Факторы влияния проанализированы вместе с возможностями использования данного подхода к решению такого класса задач. Модель позволяет строить долгосрочные сценарии для определения путей развития сельского хозяйства на территориях, пострадавших вследствие аварии на ЧАЭС

Ключевые слова: радиоактивное заражение, динамическая сеть Байеса, вероятностный вывод

УДК 519.226

DOI: 10.15587/1729-4061.2014.31613

DYNAMIC BAYESIAN MODELLING FOR RADIONUCLIDE SOIL-TO-PLANT TRANSFER

I. Zagirska

Graduate student*

Email: zahirska@gmail.com

P. Bidyuk

Doctor of technical sciences, Professor*

Email: pbidyuke@gmail.com

D. Levin*

Email: dmtlevin@gmail.com

*Mathematical methods of system analysis

department

National Technical University of Ukraine

"Kyiv Polytechnic Institute"

Peremogy av., 37, Kyiv, Ukraine, 03056

1. Introduction

Chornobyl disaster that happened in 1986, has affected immense fertile soil areas in Ukraine, Russia and Belarus [1]. Thus, the countries have lost vast territories of agricultural land. 28 years after, the problem of soil re-cultivation gains higher and higher importance, because a particular dynamic balance in radionuclide migration processes has been established. This allows researchers to predict that in the following decades this stability will not change at the high grade [2]. The main risk factor lays in the fact that agricultural products, obtained from re-cultivated territories, has to be compliant with the quality standards and to be safe for human life and health.

Contamination process works as follows. The main radionuclides concentrated in soil are cesium and strontium. Plants growing over the contaminated areas absorb radionuclides and accumulate them in the biomass. The reason for that is the similarity of physical and chemical properties of cesium (^{137}Cs) and potassium (K), and strontium (^{90}Sr) and calcium (Ca) [3]. Potassium and calcium are two of the most important elements for human body, indispensable for growth and development, as well as for cell functioning. The other radionuclides (plutonium, europium, uranium etc.) can't substitute any important elements. That's why they are accumulated in plants in small quantities.

The study is of extreme importance, because the tendency shows that getting radionuclides through the nutrition chains has become the major cause of human exposure to radiation, especially inner exposure [4]. Radioactive contamination of human body can be direct and indirect, i.e. by direct consumption of contaminated plants in the first case, and by consumption of meat and dairy products in

the second case. That causes a number of diseases, that's mainly different types of cancer, as well as cardiovascular, neurotic and respiratory system diseases, genetic mutations, only 10 % of which can be seen in the first generation [5]. So, the task of nutrition product safety and the possibility of growing them over certain territories is extremely acute.

To analyze the safety agricultural plant growing over the re-cultivated territories for the human body a mathematical model based on the dynamic Bayesian network was created. Model development specifics and results are to be described further on.

2. Literature overview and task setting

During the time that has elapsed, researchers from Ukraine and the other countries pay much attention to the radiological state of plants and study their characteristics in terms of accumulation of radioactive elements from soil. This is carried out in particular with the aim of agricultural land re-cultivation and of avoiding internal exposure of the human body through the food chain.

Much attention is drawn to the problem of radionuclide transfer evaluation from soil to plants. The authors consider both general principles of radionuclide transfer in theoretical [6] and experimental [7] aspects, and their dependence on various factors. In particular, the most wide spread goal of research is the classification of soil types and the way these types affect the radioactive contamination on plants. The study [8] provides a classification of soils in Ukraine according to this principle, [9] – in Brazil, [10] – in Russia. The study [11] suggests some conclusions about the impact of soil acidity on the rate of plant contamination, and study

[12] – depicts the seasonal changes of this phenomenon using the example of blueberry phytomass contamination. The study [13] provides a detailed analysis of all factors influencing the value of radionuclide transfer to plants. The author presents a classification of internal and external influence factors and the nature of these factors.

Special attention is paid to the definition of plant contamination in the context of influence over the human body. About 25 % of all incomes of radioactive cesium to the body of people living over the contaminated areas is caused by eating contaminated food, taking into account that according to the traditional nutrition principles the level of mushroom and berry consumption is relatively high [14]. The contribution of wild berries and mushrooms to the internal exposure is up to 75–80 % of doses received by the population of the contaminated areas from food.

Currently, the processes of radionuclide migration reached the state of dynamic equilibrium. Thus, it is reasonable to predict that for quite some time, namely several decades, the state will not change or will slightly change [6]. This means that 70–90 % of radionuclides will stay in the upper (5–10 cm) soil layer [12].

As for Bayesian modeling, one should mention that this tool is fairly new, and the possibility of its application is still investigated, in particular in [15]. Currently the main area of its use is economics and finance, insurance, logistics, marketing, medicine, graphic systems. Basic research of mathematical tool was performed by K. Murphy in [16].

Meanwhile, use of Bayesian networks for solving environmental problems is possible, but not common. In addition, in spite of quite a broad representation of radionuclide transfer issues in biological studies, the work on mathematical modeling of this process were not widely carried out, thus it is a part of scientific novelty of this study.

3. Aim and tasks of the study

The aim of this research is to develop a model based on dynamic bayesian network in order to estimate and forecast the level of radionuclide transfer from soil to agricultural plants. This is necessary to define if growing particular plants is safe over the particular part of the affected territory.

The aim can be split into several tasks:

- analysis of the factors having influence over radionuclide transfer into plants;
- development of the model for a particular ecological process using dynamic Bayesian network;
- quality analysis of the model and making conclusions on the possibility of using this mathematical tool for the similar problems.

4. Development of a dynamic Bayesian model for radionuclide soil-to-plant transfer

4. 1. Reference data

The reference plants were comprised of 215 species divided into 9 groups according to the principle of usage and consumption of these plants: beans (soybeans, cat's clover, kidney bean, pea chick, peas etc.), mushrooms (agaric mushroom, boletus edulis, champignon, chanterelle etc.), cereals (barleycorn, wheat, rye, oats, maize), root crops

(potatoes, carrots, beetroot etc.), vegetables (tomatoes, cabbage, cucumbers etc.), oily grasses (flax, hemp, maize, sunflower), grasses (timothy, amaranth, brome etc.), edible and medical herbs (dill, parsley, lettuce, anise etc.), berries (strawberry, raspberry, cowberry, blueberry, blackberry, and cranberry).

The rates of radionuclide concentration in the plants are different depending on the part of the plant that has to be used, and the state of ripeness. Obviously, the parts in use vary for different plants, as well as the state of ripeness when the plant is nutritionally valuable. In some cases even the varieties of the same plant have different radionuclide concentration capacities.

The factors influencing the transfer ratio can be divided into internal (species-specific) and external (ecosystem). These are the model parameters. Internal factors comprise biological peculiarities of plant species; these are:

- systematic position;
- symbiosis with mushrooms;
- need for K^+ , $^{2}Ca^+$ and other cations;
- depth of placement of the root system in the soil;
- contact time (time for growing);
- part of the plant in use.

External factors define the conditions of plant growing. The most significant external factors are:

- type of the soil;
- mineralogical content of the soil;
- wellness (nitrogen level);
- humidity;
- acidity (pH);
- organic content of the soil;
- weather conditions during the vegetation period.

In general, taking into account the internal factors can be eliminated to taking onto account a single factor – systematic position, as this is the factor that univocally defines all the other factors of the same group. External factors can define the possibility of implementation of different scenarios for agricultural activities over the certain territory, so they should be taken into account separately.

4. 2. Model parameter description

The main factors of influence become the model parameters for dynamic Bayesian network. The factors are described in more detail in [13]. 8 factors were considered as the most influential over the radionuclide transfer rate from soil to plants. Further on, we provide some conclusions obtained from analyzing the database using correlations, the data is used as a prior knowledge.

1) Systematic position

This factor has high priority, because it allows estimating the level of radionuclide concentration for certain agricultural varieties based on biological diversity principles. By this factor all samples were divided into 9 groups depending on the usage of the plant (i.e. the influence over human health is not the same in case of plant consumption by animal producing milk and in the case of direct consumption by a human). Systematic position factor includes all the internal factors:

- *Symbiosis with mushrooms*

As far as up to 90 % of the radionuclides are accumulated in the upper layer of the soil (0–30 cm deep), and mushrooms have vast root system, this is an important factor of radionuclide migration. Because of notable presence in the contaminated layer, systematic position and natural properties, mushrooms are immense accumulators of

radionuclides. As for agricultural plants, the closer the symbiosis is the more radionuclides can penetrate from mushroom into root system.

• *Need for K⁺, ²Ca⁺ and other cations*

¹³⁷Cs and ⁹⁰Sr are not the only radionuclides left in soils after Chernobyl disaster, yet being the most dangerous. The major danger is caused by their feature of substituting the vital minerals: ¹³⁷Cs is perceived as K, and ⁹⁰Sr as ²Ca. The plants literally mistake the radionuclides for these cations and get contaminated. In this way, the higher is the need of the plant in K⁺ and ²Ca⁺, the higher is the risk for agricultural products.

• *Depth of placement of the root system in the soil*

As stated above, up to 90 % of the radionuclides are accumulated in the upper layer of the soil. The process of migration into the lower layer is slow, so the radioactive substances persist close to the surface. With depth the concentration of radionuclides decreases exponentially. In this way, generally, the deeper the root system of plant is placed in the soil, the lower is the transfer ratio of the radionuclides. Especially strong is the correlation shown by the difference between varieties of the same species.

• *Contact time (time for growing)*

Ripeness time varies for different species and varieties of plants: some of them ripen faster, thus having less contact with the contaminated soil. During shorter period of time a lower amount of radioactive elements can be transferred into the plants.

• *Part of the plant in use*

As mentioned above, the upper layer contains up to 90 % of the radionuclides present in the soil. For example, in the forest areas the bedding is the major accumulator of the radionuclides (40–80 %), whilst the tree trunks only contain 2–8 % of the dangerous elements. Apparently, for the plants used in nutrition the importance is in using the parts having less contact with contaminated soil. Still depending on the biological features, certain plants accumulate more radionuclides in the parts different from the roots. In this way, the issue should be analyzed, so that the plants having less accumulating parts in use are cultivated.

External factors offer more vast room for possible combinations to change influence over radionuclide transfer.

2) *Humidity*

Normally the plants accumulate more radionuclides with the increase of humidity of the soil. In this case, the peat

soils of Polissya region show the highest rate of radionuclide migration under conditions of high humidity.

3) *Soil type*

The intensity of radionuclide transfer becomes higher with the decrease of fractions with small grain – especially clay. Fertility characteristics also influence the transfer ratio depending on the particular soil type.

4) *Organic content*

Probably the main activity of ¹³⁷Cs is related to the main organic substance of the soil (humus). It is proved by the correlation coefficient of 0.95 between the radionuclide activity and its number in the soil horizons.

5) *Acidity (pH)*

The intensity of transferring the radionuclides to the plants increases on the soils with lower pH level (higher acidity) [5]. Acidity level influences the content of K⁺ and ²Ca⁺, so it is taken into account in the second level of the model.

6) and 7) *Mineralogical content of the soil (K⁺ and ²Ca⁺ included separately)*

K⁺ and ²Ca⁺ are antagonist ions for ¹³⁷Cs and ⁹⁰Sr respectively. The higher is the demand of the plant in K⁺ and ²Ca⁺, the higher is the absorption of ¹³⁷Cs and ⁹⁰Sr. ¹³⁷Cs and ⁹⁰Sr substitute K⁺ and ²Ca⁺ in the plant body. This demand is comprised not only of biological needs regarding the particular plant variety, but also of the mineralogical content of the soil. The more K⁺ and ²Ca⁺ the soil contains, the less elements are mistaken by the plants, meaning less radioactive elements are transferred from the soil.

8) *Granular composition*

Transfer ratio increases for the soils having higher rate of small fractions.

Weather conditions during the vegetation period can be considered as external disturbance, as it can't be predicted and controlled. As far as the feedback from it influences the transfer factor, but the control over it can't be taken, the factor is not considered in the model. Thus it might be a matter for further investigations.

Forecasted value

Transfer factor (TF) is the ration of radionuclide transfer from soil to plants, defined as contamination of plant mass (Bq/kg) by contamination of 1 m² of soil (kBq/m²). The Table 1 below shows correlation between the factors.

Table 1

Correlation values between the influence factors

Influence factors	Systematic position	Humidity	Soil type	Organic content	Acidity	Ca	K2O	Granular composition	TF
Systematic position	1	0,122349661	-0,019587335	0,00081873	-0,056719382	-0,034163079	-0,057787621	-0,105266333	-0,090198428
Humidity	0,122349661	1	0,064127194	-0,039475407	-0,076440668	-0,103962524	-0,167148343	-0,169672769	-0,100636682
Soil type	-0,019587335	0,064127194	1	-0,260950806	-0,040734481	-0,295119383	-0,226791967	-0,112320827	0,018870911
Organic content	0,00081873	-0,039475407	-0,260950806	1	-0,059736446	0,333766273	0,159736252	0,092580391	0,20529436
Acidity	-0,056719382	-0,076440668	-0,040734481	-0,059736446	1	0,190383353	0,34816938	0,126505355	0,063227758
Ca	-0,034163079	-0,103962524	-0,295119383	0,333766273	0,190383353	1	0,345337303	0,256925962	0,001186987
K2O	-0,057787621	-0,167148343	-0,226791967	0,159736252	0,34816938	0,345337303	1	0,114446664	0,021384545
Granular composition	-0,105266333	-0,169672769	-0,112320827	0,092580391	0,126505355	0,256925962	0,114446664	1	0,029888044
TF	-0,090198428	-0,100636682	0,018870911	0,20529436	0,063227758	0,001186987	0,021384545	0,029888044	1

5. Dynamic Bayesian Networks as a Modeling Tool

Dynamic Bayesian networks (DBN) make up a powerful tool for mathematical modeling that can be widely used for a number of problems to solve [15]. Interesting is the fact that usage of DBN allows finding generalizations in the tasks that have nothing in common from the first sight. Learning and inference processes can be facilitated with the help of hierarchical Markov model that shortens time and computational complexity of the operation. There is a number of other benefits of using DBN for modeling. DBN is an extension of simple Bayesian networks, that are static, and used for probability distribution modeling on the sets of random variables, Z_1, Z_2, \dots . Normally the sets are split into the triades $Z_t = (U_t, X_t, Y_t)$, that stand for the set of variables of the incoming, hidden and the outgoing layers of the model [16]. Generally it is assumed that time is discrete, meaning that t increases with each new arrival of observation. Not only the set of variables differs in time, but also the system structure, and this is what the term “dynamic” stands for.

According to the definition, DBN is a couple (B_+, B_-) , where B_+ is a BN that defines prior probability $P(Z_1)$, and B_- is a double-layer BN, that defines $P(Z_t | Z_{t-1})$ using a directed acyclic graph as follows:

$$P(Z_t | Z_{t-1}) = \prod_{i=1}^N P(Z_t^i | Pa(Z_t^i)), \tag{1}$$

where Z_t^i is node number i at the moment of time t , that could be a component of X_t, Y_t or U_t , and $Pa(Z_t^i)$ are the parent nodes of Z_t^i in the graph. Nodes of the first layer of the double-layer BN has no parameters associated with them, but each node in the second layer has a related conditional probability distribution defined as $P(Z_t^i | Pa(Z_t^i))$ for all $t > 1$.

Parents of the node, $Pa(Z_t^i)$, can be in the same or from the previous time layer. It is a common simplification, in general the relations can be far more complicated and there are no mathematic restrictions for that the parent nodes are situated not further than 1 layer away.

Arcs between the layers are oriented from left to right to mark the time flow. If there exists an arc from Z_{t-1}^i to Z_t^i , this node is called constant. Arcs within one layer are conditional, as in

general DBN is a directed acyclic graph. As an exception, within one layer the arcs might be not directed. This stands for marking up the relations, correlations and constraints. In this case DBN becomes a chain graph.

The model is built in concern that the parameters are time invariant, it is only the random variable value that changes in time. If DBN parameters are to be changed, they can be added to the model state-spaces.

The sense of DBN foresees splitting of the network into T time layers. This looks like following:

$$P(Z_{1:T}) = \prod_{t=1}^T \prod_{i=1}^N P(Z_t^i | Pa(Z_t^i)). \tag{2}$$

If the data is not put in interactively, inference can be at easiest performed by splitting the network into layers, making inference statically (on each layer separately) and then performing junction. If the data sets have variable length, one should perform all the operations on the longest one and then cut of the redundant values.

Network development

The model topology was designed as shown at Fig. 1.

Learning set comprised of 5700 samples, while testing set comprised of 570 samples from all 80 time steps. Junction tree inference algorithm was used, as it proved to be accurate with when working with continuous nodes (TF in our case) and expert-created topology. One time step equals to one month.

Graphs below (Fig. 2, a, b) show the errors for learning set – $\log(\text{SSE})$ (Sum of Square Errors) and MAPE (Mean Absolute Percentage Error).

Graphs in Fig. 3 show the deviation between DBN-modeled results and the actual set for one sample of the learning set.

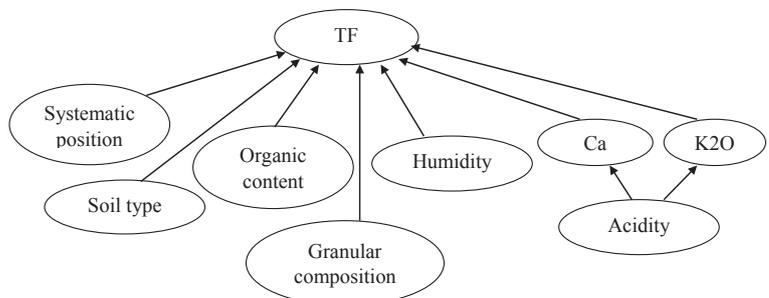


Fig. 1. DBN Topology

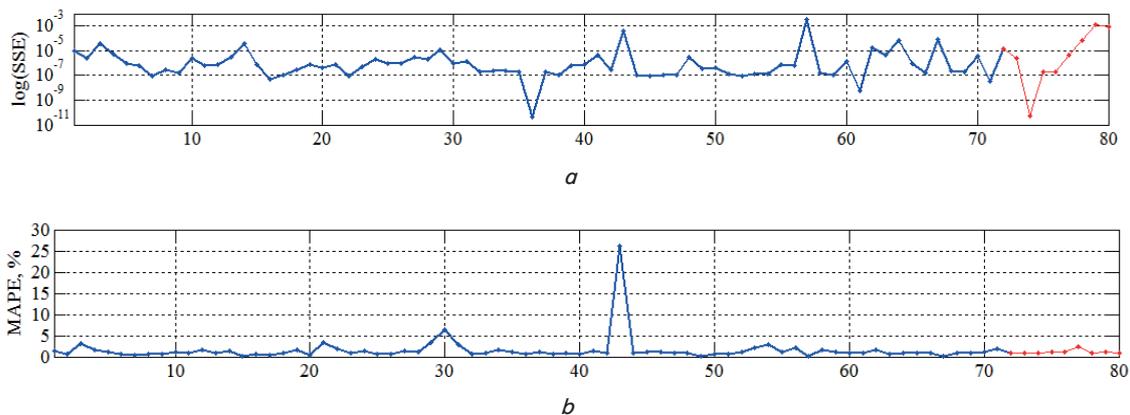


Fig. 2. Illustration of the errors for learning set: a – $\log(\text{SSE})$, b – MAPE

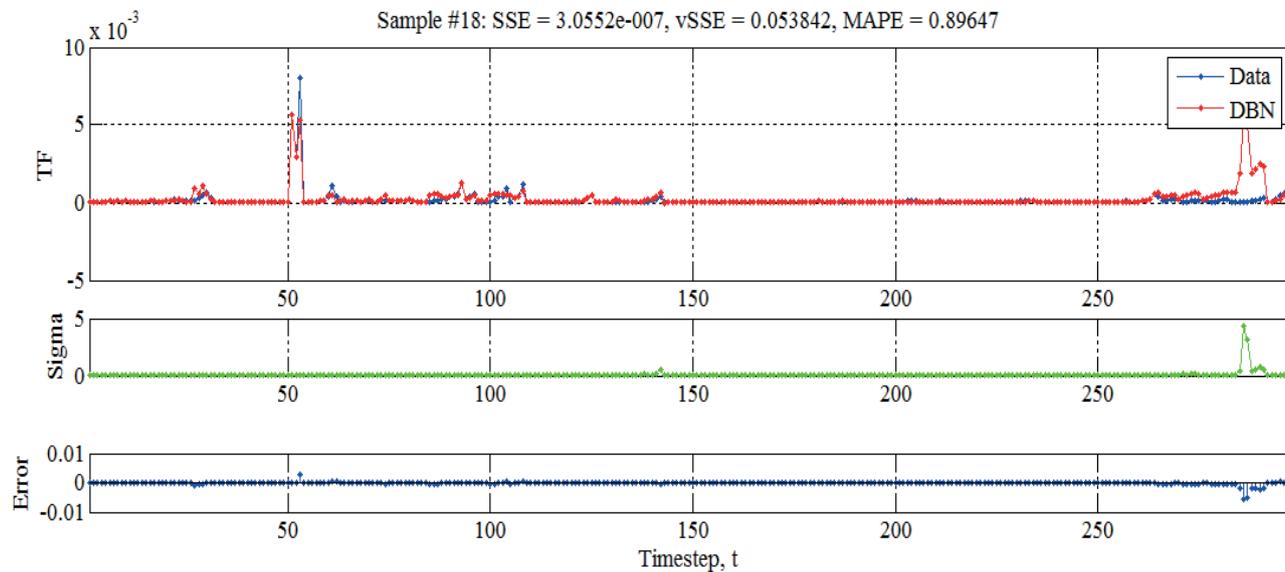


Fig. 3. The deviation between DBN-modeled results and the actual dataset for one sample of the learning set

The results achieved with the model in the form of DBN are highly positive in the sense that the model constructed adequately represents existing reason and consequence interactions between the variables selected. The learning errors are truly small: MAPE does not exceed 5 % except for two measurements, and $\log(\text{SSE})$ is less than 10^{-3} . The deviation error shown in Fig. 3, is close to zero. Thus, the model constructed is suitable for building of respective scenario directed towards forecasting radionuclide contamination.

6. Conclusions

During the research a model based on dynamic bayesian network was developed in order to estimate and forecast the level of radionuclide transfer from soil to agricultural plants. While performing the modelling, a number of problems were solved:

- the factors having influence over radionuclide transfer into plants were analyzed. The following regularities were discovered: the transfer ration decreases with deeper root system, lower humidity, on the fertile soils with low amount

of small fractions, low acidity and high content of K^+ and $^{2}\text{Ca}^+$. Plants meeting those conditions can be grown over the affected territories at lower contamination risk;

- particular ecological process was modelled using dynamic Bayesian network. This approach is reasonable, because the model demonstrates high quality. The learning errors are insignificant: MAPE does not exceed 5 % except for two measurements, and $\log(\text{SSE})$ is less than 10^{-3} ;

- the suitability of the selected mathematical modelling tool for solving problems of this particular class was analyzed.

Apart from that, there are also other benefits of using bayesian approach:

- Bayesian approach provides high level of generalization that allows combining parameters of various natures;

- DBN allows analyzing of the influences between the variables in the time-layer and between the different time-layers, vertically, that is especially efficient for environmental modeling.

All the mentioned factors allow stating that further usage of dynamic Bayesian networks for the tasks related to ecological process modelling is a promising way to solve problems of the similar kind.

References

1. De Cort, M. Atlas on the caesium deposition across Europe after the Chernobyl accident [Text] / M. De Cort, G. Dubois, S. D. Fridman et al. / Office for Official Publications of the European Communities, Brussels, Belgium & Luxembourg, 1998.
2. Hrabovskyy, V. A. Monitoring of radionuclide contamination of plants in the western part of Volyn Polissya (Ukraine) during 1994–2007 [Text] / V.A. Hrabovskyy, O. S. Dzendzelyuk, O. S. Kushnir // Radioprotection. – 2009. – Vol. 44, Issue 5. – P. 639–645. doi: 10.1051/radiopro/20095118
3. Environmental protection in Ukraine 1994–1995 [Text] // Kyiv, Rayevsky publishing, 1997. – 95 p.
4. Chornobyl disaster [Text] / V. G. Baryakhtar (Ed.). – Kyiv, Naukova dumka, 1995. – 560 p.
5. Ciuffo, L. 137Cs and 40K soil-to-plant transfer Process in Semi-natural: Grassland. Assessment of its Impact on Human Food Chain [Text] / L. Ciuffo, M. Belli, H. Velasco, U. Sansone // Journal of Radiation Research. – 2003. – Vol. 44, Issue 3. – P. 277–283. doi: 10.1269/jrr.44.277
6. Priester, B. S. Radionuclide migration in the soils and transfer of them to plants in Chornobyl accident area [Text] / B. S. Priester, N. P. Omelyanenko, L. V. Perepelyatnykova // Soil science. – 1990. – Vol. 10. – P. 51–60.

7. Rahman, M. M. Transfer of radiocaesium from soil to plant by field experiment [Text] / M. M. Rahman, M. Rahman, A. Koddus, G. U. Ahmad // Journal of Biological Sciences. – 2007. – Vol. 7, Issue 4. – P. 673–676. doi: 10.3923/jbs.2007.673.676
8. Prister, B. S. The Classification of Ukrainian soil systems on the basis of transfer factors of radionuclides from soil to reference plants [Text] / B. S. Prister, L. V. Perepelyatnikova, T. N. Ivanova, V. D. Vynogradskaya, L. V. Kalinenko, N. R. Grytsjuk, V. A. Rudenko, G. P. Perepelyatnikov, V. A. Pojarkov // Classification of soil systems on the basis of transfer factors of radionuclides from soil to reference plants. Proceedings of a final research coordination meeting organized by the Joint FAO/IAEA Programme of Nuclear Techniques in Food and Agriculture and held in Chania, Crete, 2003.
9. Handl, J. Accumulation of (137)Cs in Brazilian soils and its transfer to plants under different climatic conditions [Text] / J. Handl, R. Sachse, D. Jakob, R. Michel, H. Evangelista, A. C. Gon alves, A. C. de Freitas // Journal of Environmental Radioactivity. – 2008. – Vol. 99, Issue 2. – P. 271–287. doi: 10.1016/j.jenvrad.2007.07.017
10. Sanzharova, N. The classification of Russian soil systems on the basis of transfer factors of radionuclides from soil to reference plants [Text] / N. Sanzharova, S. Spiridonov, V. Kuznetsov, N. Isamov, S. Fesenko, N. Belova // Classification of soil systems on the basis of transfer factors of radionuclides from soil to reference plants. Proceedings of a final research coordination meeting organized by the Joint FAO/IAEA Programme of Nuclear Techniques in Food and Agriculture and held in Chania, Crete, 2003. – P. 139–144.
11. Djingova, R. Influence of soil acidity on the transfer of caesium 137 and other radionuclides from soil to reference plants [Text] / R. Djingova // Classification of soil systems on the basis of transfer factors of radionuclides from soil to reference plants. Proceedings of a final research coordination meeting organized by the Joint FAO/IAEA Programme of Nuclear Techniques in Food and Agriculture and held in Chania, Crete, 2003. – P. 51–58.
12. Orlov, O. O. Peculiarities of the season dynamics of ¹³⁷Cs accumulation by the phytomass of blackberry (*Vaccinium murtillus* L.) and cowberry (*Vaccinium vitis-idaea* L.) [Text] / O. O. Orlov, O. Z. Korotkova // Naukovi visti, Lviv, UkrDLTU. – 2000. – Vol. 10.2. – P. 34–43.
13. Zarubina, N. Peculiarities of accumulation of gamma-radiation radionuclides by the macromycetes on the territory of the alienation zone and the “southern trace” after Chornobyl disaster [Text] : PhD thesis / N. Zarubina. – Kyiv, KNU, 2002. – 20 p.
14. Krasnov, V. P. Radioecology of Ukrainian Polissya [Text] / V. P. Krasnov. – Zhytomyr. Volyn, 1998. – 112 p.
15. Murphy, K. A Brief Introduction to Graphical Models and Bayesian Networks [Electronic resource] / K. Murphy. – University of British Columbia, Faculty of Science. – Available at: <http://www.cs.ubc.ca/~murphyk/Bayes/bayes.html/> – 07.07.2007.
16. Zgurovskyi, M. Z. System methodics of designing BN [Text] // M. Z. Zgurovskyi, P. I. Bidyuk, O. M. Terentyev / Kyiv, Naukovi visti NTUU “KPI”. – 2007. – № 4. – P. 47–61.