METHODS FOR SELECTING SIGNIFICANT FACTORS IN FORECASTING AND ZONING TERRITORIES

Yakhyaev Khashim
Doctor of agricultural sciences, professor,
Leading Researcher, Research Institute of Quarantine and Plant Protection
Andijan, Uzbekistan
E-mail: egamberdiyevoybek60@gmail.com
Abdullaeva Khuriyatkhon
PhD of Agricultural Science,
Associate Professor, Andijan institute Engineering and Agriculture
Andijan, Uzbekistan
Musaeva Gulbakhor
PhD of agricultural science,
Senior Lecturer, Andijan institute Engineering and Agriculture
Andijan, Uzbekistan
Rahimov Mansurbek
PhD of agricultural science,
Senior Lecturer, Andijan institute Engineering and Agriculture
Andijan, Uzbekistan

ABSTRACT

The ways and principles of choosing essential factors in the development of mathematical models for predicting the development and spread of pests, as well as the issues of zoning territories according to their weather and environmental characteristics are given.

Key words: mathematical models, significant factors, zoning of territories, forecasting, weather and environmental characteristics.

Intoduction. The use of modern information technologies in agriculture of the Republic of Uzbekistan, including the protection of plants from pests and diseases, remains a modern requirement. Currently, farms that cultivate cotton and grain in large areas of the country have been replaced by cultivation based on low volume technologies using farms and clusters. This, in turn, will reduce the cost of growing cotton and grain, reduce the cost of fertilizers, and bring irrigation systems to an alternative level. That is why it is important to create and introduce phytomonitoring bases in cotton and grain cultivation. One of the main tasks of phytomonitoring is to know and analyze the condition of crops, to identify the causes of developmental delays. Timely receipt of such information allows to make clarifications and changes in the technology of cotton and grain cultivation, to determine the impact of certain factors and to create optimal conditions for crop development.

Biological suppression of harmful pests has occupied today a constant predominating position in the concept of integrated suppression of pests. In certain cases, for successful regulation of number of insects - pests any form of biological suppression of their populations is enough; in a number of other cases, suppression is often supplemented with any other method, and at last, in the third, it plays only an auxiliary role. Nevertheless, an effective utilization of biological agents became now a method to be considered and widely applied in all chances. In this connection it is necessary to study the inter-specific ecological communications which play an important role in vital systems of populations.

At the present stage of specialization and intensification of agriculture, due to the need for a general improvement in the strategy and tactics of plant protection, the importance of forecasts for the spread and development of pests of agricultural crops, especially cotton, has sharply increased. The use of mathematical methods based on the specifics of populations, the nature of their reactions to environmental factors, makes it possible to model the dynamics of pest populations and effectively manage them. At present, with very large investments in plant protection, errors in forecasting or choosing a management method can be very costly.

Based on the above, this article discusses the development of methods for selecting significant factors in predicting the dynamics of pest populations and phytosanitary zoning of territories using pattern recognition algorithms.

Results. Methodology for Selecting Significant Factors in Predicting the Dynamics of Pest Populations**.** Effective management of crop production requires the timely collection and processing of extensive agrotechnical, agroeconomic, agrometeorological and agroecological information. The collection, storage and appropriate processing of the above information is necessary for making optimal decisions. The collection of such diverse and extensive information and the solution of prognostic problems require huge labor costs. In this regard, the tasks of developing mechanized data collection systems, improving forecasting methods with a focus on more accessible types of initial information and using modern computer and microprocessor technology for processing, storing information and creating a database (DB) arise.

Moreover, mathematical modeling of the process of pest population dynamics is a complex multifactorial complex, where, along with essential factors, less significant and insignificant ones act. Very often, the mathematical model and the empirical curve do not coincide due to the fact that the presence of insignificant factors obscures the main aspects of the process under study, and some significant factors that give this process a certain character may not be taken into account. In addition, numerous factors contribute to a sharp increase in the amount of work associated with the collection and processing of information. In view of these circumstances, when modeling the dynamics of agricultural pest populations, it is advisable to divide the entire factorial process into areas of significant, less significant and insignificant factors. The solution of this problem will be much easier if we introduce some quantitative measure to assess the significance of a particular factor.

To solve the problem, a class of pattern recognition algorithms is used, which are mainly based on the calculation of estimates [1, 2]. The main problem of pattern recognition can be solved by various methods, for example, correlation (using the theory of statistical decisions), algorithms "Generalized portrait", "Kora-2", "Kora-3", etc. All these methods and algorithms have been applied in solving a number of applied problems and have given good results. Their common distinguishing feature is that when implementing the recognition process, the concepts of importance (informativeness) of the features defining the object are not explicitly used. Therefore, the concept of the measure of importance (information weight) of features is not introduced and it is not calculated [4, 5]. However, in recent years, methods in which the main stage is the calculation and consideration of information weights of features are becoming more widespread. The informational weights of the features are calculated by the frequency of occurrence, the analysis of dead-end tests, and algorithms for calculating estimates.

Let us consider some general principles on which the algorithms for calculating estimates are built.

Let the table Tnm be given. In it, each line Sj (j=1,2,…,m) is an object that can be represented as

$$S_j = (X_{j1}, X_{j2}, \dots, X_{jn}),$$

where elements of a set of numerical or qualitative values that identify an object. Usually these elements are called signs, and their values are obtained, as a rule, as a result of experiments. Пусть задана теперь совокупность множеств q_i , которую будем рассматривать как некоторый алфавит признаков. В качестве алфавита моGут:

Let now a set of sets qi be given, which we will consider as some alphabet of features. As an alphabet can:

a) a set of two elements {0 , 1}, where "1" means the presence of some property, 0 means its absence;

b) a finite set of integers { 1,2,…,P }. The meaningful value of each such attribute reflects the degree of expression of the corresponding property.

An object S_j is called accessible if $X_{ji} \cap q_i$. Note that the concept of a valid object is equivalent to the concept of a valid string.

The set of m admissible objects, each of which is characterized by a set of features, can be reduced to a table Tnm, which we will call an admissible table.

Now suppose that there is a partition of all admissible rows into classes k_1, k_2, ...,k_l, and $K_u \cap K_j \neq 0$ for $u \neq j$. This partition induces a partition of the rows of the table into l non-overlapping classes.

$$K_1, K_2, \ldots, K_l$$

We assume that the last partition is given.

Let us denote the number of elements of class Ke as mu-mu-1 and represent the division of rows by class as

$$K_1: S_1, S_2, \ldots, S_{m1}$$

$$K_2: S_{m1+1}, S_{m1+2}, \ldots, S_{m2}$$

…………………

$$K_e: S_{me-1+1}, S_{me-1+2}, \ldots, S_{me-1}$$

The table T_{nm}, whose rows are divided into L classes, will be denoted by T_{nml}.

The introduced concepts and definitions allow us to proceed to the consideration of the principle caused by the principle of calculating some quantitative estimates obtained by comparing admissible objects or parts and representing the basis of algorithms for calculating estimates.

Let two admissible objects be given:

$$S_1=(x_{11}, x_{12}, \ldots , x_{1n})$$

$$S_1=(x_{21}, x_{22}, \ldots , x_{2n})$$

where x_{jn} $(j=1,2)$ takes the values of one of the above alphabets.

Let further it is required to establish a degree of "proximity". We will understand this term on an intuitive level, comparing it with the concepts of similarity, sameness, etc. The proximity of S_1 and S_2 can be established, for example, based on the analysis of the identifying sets or their parts. This approach is associated with obtaining some quantitative estimate using certain transformations of the sets. Based on this assessment, proximity is established by a certain criterion.

Thus, if the identifying sets of objects are denoted by x_1 and x_2, and the estimate by B, then for the latter we can write

$$B=B(x_1, x_2)$$

The value of B can be intermediate, and, in turn, make it possible to obtain some final estimate characterizing the proximity of the object. This assessment in quantitative terms characterizes the total effect of comparing objects.

Therefore, if the total effect is denoted by G, then we can write $G=G(B)=G(B(x_1, x_2))$

In the theory of recognition algorithms based on the calculation of estimates, the value G is usually called the number of votes given by objects, for example, for the object S_2. It is clear that the number of votes cast by objects S_2 for $S_1(G_{1,2})$ is equal to the number of votes $G_{1,2}$.

Procedures for obtaining the value of G for comparing objects S_1 and S_2 are called voting procedures. Voting procedures can be carried out not only to establish the proximity of two objects, but also a group of objects summarized in a specific table. So, if some table T_{nm} consists of objects, each of which is characterized by a set of n features, then the following voting matrix can be obtained by the voting procedure:

$$\begin{array}{cccc} G_{11}, & G_{12}, & \dots, & G_{1m} \\ & G_{22}, & \dots, & G_{2m} \\ \dots & \dots & \dots & \dots \\ & & & G_{nm} \end{array} \qquad (1)$$

Matrix (1) actually makes it possible to obtain all the necessary values of the votes given by individual objects both for themselves and for other objects of this table. In practical problems, the votes cast for oneself are usually excluded from the matrix (1).

Note that the effectiveness of voting procedures depends on the method of counting votes. The procedure, from the point of view of the speed of implementation, is the more effective, the simpler the method of counting votes.

In the theory of evaluation algorithms, such an apparatus has been obtained, with the help of which the number of votes is calculated relatively simply. The works [1-3] describe the derivation and proofs of the following expressions for counting votes (the value of G), as well as various modifications of the counting formulas, which differ from each other in the specifics of specifying certain stages of the algorithms for calculating estimates. Let us write out expressions for counting votes, which will be used later in this paper.

If the table T_{nml} is given, divided into classes $K_1, K_2, \ldots, K_l$, then the number of votes cast by row S for class K_u is equal to

$$G_u(S)= \sum_{q=m}^{m} C_{n-r}^{h}(S,S_h) \qquad (2)$$

where $S=(x_1, x_2, \ldots, x_n)$, $S_q=(x_{q1}, x_{q2}, \ldots, x_{qn})$ pairs of strings to be compared; h - is the length of the voting set, i.e. the number of columns by which the rows S *u* S_q are compared; $r(S, S_q)$ - Hamming distance between strings S and S_q (if the table is given by binary characters), i.e. the number of columns where rows do not match. When the table is filled with elements that take a value from an arbitrary alphabet, then instead of a value, you should use $r(S, S_q)$ - the distance between the rows S and S_q, equal to the number of fulfilled inequalities where $abs(x_l-x_{qi}) \geq V_i$ - the proximity threshold of the i -th feature. If $h \geq n- r(S, S_q)$, then expression (2) vanishes.

According to [1, 2, 6], the value

$$P_1 = \sum_{u=1}^{e} \frac{1}{m_u-m_{u-1}} \sum_{q=m_{u-1}+1}^{m_u} \left[G_u(S_q^i)\right] \qquad (3)$$

is called the information weight of the i-th feature.

Thus, the above method for determining the information weights of signs was used to select the essential factors necessary in the development of mathematical models for predicting the development of harmful objects of agricultural crops [17].

Tasks and methods of phytosanitary zoning of territories by agricultural pests. The task of optimized zoning of territories described below, in our opinion, can be used to create an ACS for plant protection (APCS). Establishment of the

APMS is unthinkable without an objective preliminary zoning of the territories in relation to the objects of management.

The zoning of agricultural territories according to weather and environmental characteristics are essential links in the scientific knowledge of the continuous environment. Works in this direction are of great practical importance, since zoning is an essential element in many studies, in particular, when predicting the size of the dynamics of the population of agricultural pests, the size of infested areas by pests, the date of appearance of pests, planning pest control measures, and analyzing the causes of outbreaks. breeding pests, etc.

Against the background of the modern development of science and changes in the demands of practice, the traditional group of zoning methods does not always provide satisfactory results.

Zoning in plant protection is a systemic process consisting of a number of interrelated elements, such as the choice of a presentation model, the scale of research, the type of primary objects of observation, methods for locating and describing these objects, as well as subsequent processing of information and interpretation of zoning results. The heterogeneity of these elements creates different possibilities for formalization. The formalization of even individual stages of the zoning process entails a restructuring of the entire system of this process as a whole. Due to the use of formalization methods, an important place is occupied by the linking of all elements of zoning, the development of general rules for making decisions at each stage, the development of general rules for interpreting the results.

At the stage of obtaining the actual source material, the task of primary selection of information is of particular importance. Depending on the specific tasks, it is possible to zoning at the level of farm territories, territories of an administrative district, region, or at the level of observation sites. The objects of zoning include sets of indicators representing the territory being zoned, and they can be conditionally divided into groups of indicators. The first group includes:

- annual and long-term average indicators of the number of harmful objects, the intensity of development of diseases, the percentage of crops populated by pests, the size of infected areas, the percentage of spread of the disease;
- annual and average long-term indicators of reproduction, survival and the timing of the passage of the main phenological phases. These indicators should be collected from different areas of the zoned territory over a number of years according to a single methodology.

The second group of indicators includes parameters that characterize the conditions for the development of harmful organisms and their habitat. These include:

- hydrometeorological factors - the sum of effective temperatures, average ten-day, average monthly, maximum, minimum air temperatures. Hydrothermal coefficient (HTC), the amount of precipitation by decades and months. The timing of the onset of phenological phases and their duration.
- agrotechnical factors - the timing of sowing, harvesting. The timing of the main agrotechnical and protective measures.

Zoning according to the characteristics of reproduction and the timing of the passage of the main phenological phases of harmful organisms makes it possible to identify zones with different intensity of reproduction of pests.

To solve the problems of zoning in plant protection, the same class of pattern recognition algorithms [1,2,6], which was considered above, is used.

In this case, the classification of objects without a standard is used. The problem of spontaneous partitioning of a set of objects is solved using algorithms for calculating estimates using some quantitative measures that characterize the information content of both features and the objects themselves. The informativeness of the object is obtained from expression (3). Computation for all given objects, the information weights are then ordered in descending order.

The division of objects into classes in this case is based on the assumption that when ordering objects by the value of the latter, they will be grouped by rank. The partition obtained in this way is an intermediate stage of spontaneous classification. The final splitting is obtained only after the voting confirms that the object belongs to its class.

The proposed zoning method is based on the rational processing of multidimensional arrays of biological, ecological, agrometeorological and agrotechnical information. Such information is contained in the archival data of the meteorological station network, PPP (plant protection station), strong points and observation points for pests and diseases. The use of recognition algorithms in rational data processing makes it possible to reduce the amount of information required for decision-making by combining points - information meters into groups, and observations - into time classes, according to the principle of similarity of the information contained in them. At the same time, the quality of the initial information is improved by eliminating gross errors while maintaining its information content.

As an example, the table shows the results of zoning the territory of the Andijan region in relation to the cotton bollworm on cotton in terms of abundance, the timing of the passage of the main phenological phases, hydrometeorological (the sum of effective temperatures, HTC, the amount of precipitation) indicators.

Thus, as can be seen from the table, the districts of the Andijan region can be divided into classes (Table 1.):

- Andijan, Asaka, Buz, Shakhrikhan districts - 1st class;
- Balykchi, Bulakbashi, Jalakuduk, Izboskan districts - 2nd class;
- Kurgantepa, Markhamat, Ulugnar districts - 3rd class;
- Oltinkul, Pakhtaabad, Hujaabad districts - 4th class.

Table 1.

Information weights of districts of Andijan region in relation to cotton bollworm on cotton

№	District names	Information weights of districts	District class
1.	Andijan	0.961	1
2.	Asaka	0.915	1
3.	Balykchi	0.890	2

4.	Bulakbashi	0.813	2
5.	Bustan	0.971	1
6.	Jalakuduk	0.872	2
7.	Isboskan	0.810	2
8.	Kurgantepa	0.702	3
9.	Marhamat	0.785	3
10.	Oltinkul	0.621	4
11.	Pakhtaabad	0.544	4
12.	Ulugnar	0.796	3
13.	Shakhrikhan	0,909	1
14.	Hujaabad	0.508	4

For definition of optimum terms of basic cotton pests' occurrence (cotton worms and winter cotton warms), depending on mean annual and real data for different zones of cotton growing in the Republic, the mathematical models and their algorithms are developed. On the basis of these algorithms the program of definition of terms of cotton worms and winter cotton warms' occurrence is made and introduced in practical activities of Plant Protection Centers of Andijan, Kashkadarya, Kharezm and Namangan regions. The research results are shown in Table 2. According to Table 2, it is visible that except the terms of cotton worms' development, the optimum terms of release of useful entomophages such as trichogramma, lacewing and bracon are revealed.

Table 2.

Development of cotton lace wing and optimum terms of carrying out of biological protection (The first cotton generation, 2022)

Areas	Development terms (day, month)				Release terms of entomophages (day, month)		
	Egg-laying	Caterpillars of 2nd age	Caterpillars of 4nd age	Caterpillars of 6nd age	Tricho-gramm	Lace-wing	Bracon
Andijan region							
Andijan	June, 5 th	June, 11 th	June, 17 th	June, 22 th	June, 4-6th	June, 5-11th	June 17-22
Ulugnar	June, 2 nd	June, 8 th	June, 14 th	June, 19 th	June, 1-3st	June, 2-8 th	June 13-18
Kurgantepa	June, 9 th	June, 15 th	June, 20 th	June, 25 th	June, 7-9th	June, 9-15th	June 20-25
Namangan region							
Pap	June,	June,	June,	June,	June,	June,	June

	7 th	13 th	19 th	24 th	6-8th	7-13th	19-24
Narin	June, 9 th	June, 15 th	June, 20 th	June, 25 th	June, 7-9th	June, 9-15th	June 20-25
Kashkadarya region							
Koson	June, 3 rd	June, 9 th	June, 15th	June, 20 th	June, 2-4nd	June, 3-9 th	June 15-20
Nishan	June, 2 nd	June, 8 th	June, 14 th	June, 19 th	June, 1-3st	June, 2-8 th	June 13-18
Khorezm region							
Urgench	June, 3 rd	June, 9 th	June, 14 th	June, 19 th	June, 2,th	June, 4-10th	June 14-19
Shavat	June, 2 nd	June, 8 th	June, 13 th	June, 18 th	June, 2,4th	June, 3-9 th	June 13-18

Conclusions. The pest control system practiced in the past decades, mainly due to the massive use of chemicals, especially when their use was not sufficiently substantiated, in ecological and economic terms, led to serious problems associated with a negative impact on the environment and the emergence of pest resistance to pesticides. plant protection. In part, this even contributed to a direct or indirect increase in the harmfulness of certain types of pests and diseases and the growing dependence of the crop on the effectiveness of measures to combat them. In this regard, in recent years, the biological method of combating agricultural pests has become more widespread, as it most satisfies the principles of environmental protection.

References

1. Журавлев Ю.И., Никифоров В.В. Алгоритмы распознавания, основание на вычислении оценок //Кибернетика 1981. - № 3. С.1-11.
2. Журавлев Ю.И., Камилов М.М., Туляганов Ш.Е. Алгоритмы вычисления оценок и их применение. –Ташкент.Фан. -1984. -119с.
3. Камилов М.М., Ким А.Н., Бузурханов В. Программный распознающий комплекс, основанный на алгоритмах вычисления оценок (ПРАСК) // АлGоритмы и программы. -Ташкент: РИСО АН УзССР. - 1974. - вып.17. -С.3-33.
4. Пискорский Л.Ф. Алгоритмы GП2 и GП3 Глобальной оптимизации многопараметрических функций методом случайного поиска// Вопросы

вычислительной и прикладной математики. –Ташкент: РИСО АН УзССР. –1973. –вып. 20.

5. Пискорский Л.Ф. Программа нахождения глобального экстремума функций многих переменных// Алгоритмы и программы. –Ташкент: РИСО АН УзССР. –1973. – вып. 12.
6. Яхяев Х.К., Холмурадов Э.А. Автоматизация прогнозирования развития и распространения вредителей и болезней сельскохозяйственных культур. -Ташкент, -«ФАКК», -2005, -169 с.
7. Яхяев, Х. К., & Абдуллаева, Х. З. Автоматизированная система мониторинга развития и распространения вредителей сельскохозяйственных культур. Science and World, № 5 (21), том-3. - ВолGoГрад, 2015, с-129-131
8. Wang, R., Li, R., Chen, T., Zhang, J., Xie, C., Qiu, K., ... & Liu, H. (2022). An automatic system for pest recognition and forecasting. *Pest Management Science*, *78*(2), 711-721.
9. Yakhyaev Khashim Kasimovich, Musaeva Gulbakhor Maksudovna, Abdullaeva Khuriyatkhon Zafarbekovna / Modeling the formation of the cotton crop and forecasting the development of the main pests. NeuroQuantology 20 (20), 2022., 2222-2229