Mathematical and information support of the program complex for planning of the harvesting of wild-growing products

Ya Ivanyo, S Petrova and M Polkovskaya

Irkutsk State Agricultural University named after A.A. Ezhevsky, pos. Molodezhny, Irkutsk Region, 664038, Russia

E-mail: iasa_econ@rambler.ru

Abstract. The paper presents the results of the development of information and software for a program complex of modeling the harvesting of wild-growing food resources. Linear programming models with interval parameters are proposed to optimize the harvesting of wild-growing food resources. In addition to models with interval parameters, a model of stochastic programming is developed under the condition of stationary state of the yield series of wild-growing crops. Models may contain dependent and independent coefficients for unknowns. The database includes information on the yield of wild crops, labor resources, product prices, questionnaire data filled in by procurement staff, results of field studies on a preliminary assessment of possible harvest, maps showing the territories of commercial resources for each species. The developed software package can be used to simulate the harvesting of wild food resources in any regions with commercial reserves of wild plants. The development results were tested for some municipal districts of the Irkutsk region and model clusters.

1. Introduction

The task of harvesting and marketing of environmentally friendly forest products is of great importance for the livelihood of a person and his health [1]. In Russia, the Irkutsk Region has significant reserves of wild-growing food resources. Forest land covers about 90% of the region. According to studies of the potential of wild food resources (fruits, berries, pine nuts, mushrooms, seeds, medicinal plants, etc.), the region has significant operational reserves [2 - 4].

The work [5] describes the consumption of wild plants in different countries of the world. It is assumed that in the 21st century, the consumption of edible wild plants will increase both dietary and regular nutrition. This is due to cost savings and human health.

Therefore, the actual task is the planning of harvesting, processing and marketing of wild-growing food products. To solve it, you can use mathematical programming problems with uncertain parameters. The uncertainty of the parameters of extreme tasks is associated with the variability of yields of wild-growing resources, price fluctuations, and the variability of labor resources, which is caused by the instability of weather conditions and the result of human activity.

Since the optimization of wild-product harvesting under uncertainty conditions involves obtaining a large number of optimal solutions, it is necessary to implement algorithms for solving extreme problems using software. Thus, the aim of the work is to describe the developed information and mathematical software for a program complex of modeling wild-growing harvest products.
To achieve the goal, the following tasks are formulated:
- the creation of a database that characterizes the parameters of the models and indicators that affect them;
- building models for optimizing the harvesting of wild-growing products with uncertain parameters;
- creation of a program complex for modeling wild-growing harvesting products.

2. Materials and methods
The materials used are data on the harvesting of wild-growing food products in the Irkutsk region for 1960-1991 [6]. In addition to this, information on the potential of taiga resources is considered according to data from 15 municipal districts of the region [4]. Literary sources on the variability of yields of berries, mushrooms, nuts, medicinal plants and other wild-growing crops, as well as data on labor costs for harvesting are analyzed [2 - 4, 6 - 12]. To assess the yields of wild-growing products, we used a technique that allows us to predict the bioproductivity of various food wild-growing crops, as well as determine their possible values for a long-term perspective [9].

When creating a program complex, appropriate design technologies were used [13 - 18]. For the development of mathematical software, the methods of probability theory and mathematical statistics, mathematical programming in conditions of uncertainty was used. The solution of mathematical programming problems is obtained on the basis of LP Solve 5.5 software using the method of statistical tests for repeated simulation of the values of interval and random parameters. Statistical data processing was carried out with the support of the software SPSS for Windows. The software interface is made in the programming environment Embarcadero Delphi XE10.1 [15]. The system Postgre SQL 9.4 [16, 18] was used to implement the database.

3. Results
With the participation of the article’s authors, a software package for modeling harvesting of wild-growing products was developed. Its components are data sources and applications for the implementation of mathematical support. The sources of information are identified: the Ministry of Agriculture of the Irkutsk region, which provides data on the suppliers of and processing food products. Harvesters of wild-growing products through questionnaires provide data on the volumes of products harvested by them and the work areas. Information from paper and electronic media is transferred to the database. In the software package, methods of simulation, probability theory and mathematical statistics, mathematical programming are implemented. The users of the complex may be the heads of forestry, procurement organizations and specialists of the Ministry of Agriculture of the region.

In this embodiment, the software package was developed for the first time. Its difference from other similar software products is the original mathematical software, based on linear programming problems with uncertain parameters. At the same time, the parameters take into account the features of harvesting wild-growing food resources.

In addition, a database has been designed that allows for the accumulation and use of data on the territories of distribution of wild-growing resources, procurement volumes, labor costs, materials of the survey of procurers, and data on bioproductivity. A part of the database is filled with schematic maps showing locations for the harvesting of wild-growing food products. With the help of the above software package, tasks with a high degree of uncertainty are solved. At the same time, some parameters of the mathematical programming problem are expert assessments, both group and individual. In this sense, the above development can solve a wide range of tasks in the context of covering the situation with inaccurate and insufficient information.

The practical value of the software package is to optimize the production of harvested products of different types, of which there are about 30, for planning procurement within a certain territory, municipal area and region. Estimating income for bad, good, and average conditions of the year is important for risk management [19].
When modeling wild-growing harvesting products, the tasks of mathematical programming with uncertain parameters [2, 20-25] are of key importance. In the conditions of insufficient information on the yields of wild-growing products, extreme tasks with interval parameters can be used.

In this case, the objective function focused on maximizing income will be written as follow

\[ f = \sum_{i,j} \sum_{l,d} \tilde{c}_{ij} x_{ij} \rightarrow \text{max}, \]  

where \( \tilde{c}_{ij} \) is the unit cost of production \( j \), fluctuating in the interval \( [\bar{c}_{ij}, \overline{c}_{ij}] \); \( x_{ij} \) are the required volumes of the obtained products \( j \) in the territory of \( i \).

The limitations of the task are related to the areas where wild-growing products are harvested, the amount of resources received, the labor costs and the costs of procurement:

\[ \sum_{i,d} x_{ij} \leq S_j, \]  

\[ \sum_{i,d} \bar{x}_{ij} \leq \bar{V}_j, \]  

\[ \sum_{i,d} k_j x_{ij} \leq K_j, \]  

\[ \sum_{i,d} \bar{d}_j x_{ij} \leq D, \]  

\[ x_{ij} \geq 0, \]  

where \( \bar{y}_j \) is the biological productivity of wild-growing resources in the territory \( i \) (forestry, municipal district, etc.) of products \( j \) within the limits of the lower and upper estimates \( [\bar{y}_j, \overline{y}_j] \); \( S_j \) is the area of product type \( j \); \( \bar{V}_j \) is the volume of products obtained in the range of lower and upper values \( [\bar{V}_j, \overline{V}_j] \); \( k_{ij} \) are labor costs for obtaining a unit of production \( j \) on site \( i \); \( K_j \) is limitation of labor costs; \( \bar{d}_j \) are reduced costs for ensuring the receipt of products \( j \) on site \( i \); \( D \) are the total allowable costs for the procurement and processing of products.

The problem can be considered in two versions: 1) the interval parameters are the cost indicators of the objective function; 2) as interval estimates, the coefficients for unknown optimality criteria and the right-hand sides of the constraints characterizing the variability of production volumes are accepted. Meanwhile, regression relationships can be observed between some coefficients of the model. In particular, the coefficients \( \tilde{c}_{ij} \) in some cases are related to the yield of wild plants \( \bar{y}_j \). For such a situation, we obtain the dependence \( \tilde{c}_{ij} = z(\bar{y}_j) \), where \( z \) is the deterministic component of the regression equation. Here the assumption is made that the yield in different territories \( i \) is the same.

The article [26] presents the results of modeling of perennial rows of harvesting berries, mushrooms and pine nuts according to the data of the Irkutsk region for 1960-1991. According to this work, sequences are random samples that obey the laws of probability distribution (see table 1).

It is shown that Gaussian distribution is applicable for harvesting berries, logarithmically normal law for mushrooms, and gamma distribution for berries. The statistical parameters of the analytical functions (the average \( \bar{x} \) and standard deviation \( \sigma \), tons) are given. The laws of probability distribution can be used under the condition of stationary state of statistical parameters. Moreover, they characterize the harvesting, and not the yield of wild food resources.

Nevertheless, with a certain approximation, model (1) - (6) can be transformed into a stochastic programming problem by using probabilistic distribution laws to estimate the yield of wild-growing resources and the volume of output. With average product prices, the model will look like:
\[ f = \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \rightarrow \max, \]  
\[ \sum_{i \in I} x_{ij} \leq S_j, \]  
\[ \sum_{i \in I} x_{ij} \leq V_j^p, \]

where \( y_j^p \) is the bioproductivity of wild-growing resources of the form \( j \) on site \( i \), corresponding to the probability \( p \); \( V_j^p \) is the volume of the resulting products of the form \( j \), associated with the probability \( p \). In this task, the restrictions on labor costs and the cost of ensuring the procurement of products correspond to inequalities (4) - (6).

**Table 1.** The laws of probability distribution and their parameters for harvesting berries, mushrooms and pine nuts.

| Product type | Probability distribution law | Parameters of the probability distribution law |
|--------------|-----------------------------|-----------------------------------------------|
| Berries      | Normal \( p(x, \bar{x}, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \) | \( \sigma = 149.6 \) \( \bar{x} = 299.7 \) tons |
| Mushrooms    | Lognormal \( p(x, \bar{x}, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(\ln(x)-\bar{x})^2}{2\sigma^2}} \) | \( \sigma = 149.6 \) \( \bar{x} = 266.7 \) tons |
| Cedar nut    | Gamma \( p(x, b, c) = \left( \frac{x}{b} \right)^{c-1} e^{-\frac{x}{b}} \frac{1}{b^c \Gamma(c)} \) | \( c = \left( \frac{\bar{x}}{\sigma} \right)^2 = 1.08 \) \( b = \sigma^2 / \bar{x} = 1005.6 \) tons |

Since the solution of mathematical programming problems with interval parameters is the upper, lower, and intermediate estimates of the objective function and the corresponding optimal plans, the algorithm involves the use of Monte Carlo method. This method is also applicable in determining the probability distribution of the objective function and the corresponding optimal plans when solving problem (4) - (6) and (4) - (9).

With the help of the program complex, the problem of optimizing income from harvesting wild-growing food products (1) - (6) for 15 municipal districts of the Irkutsk region, taking into account operating stocks, was solved. The fluctuations in the upper and lower bounds for the objective function are 20 - 25% relative to the median value of 0.173 billion US dollars.

Let us consider some additional results of modeling volumes of obtaining wild-growing food products using mathematical models (1) - (6) and (4) - (9). Table 2 shows the optimal solutions for two models of the Ust-Ilimsk municipal district. For the first model with interval estimates of the coefficients for unknowns in the objective function and restrictions, income fluctuations between the lower and upper values are 38.5%. The optimal plans corresponding to the collection of different types of mushrooms undergo strong changes, since the mushroom crop is more dependent on climatic conditions than the harvest of wild berries and medicinal plants [9]. At the same time, harvesting for blueberries, blueberries and fireweed is more stable.

Table 2 shows the simulation results for probabilities of 0.05, 0.5, and 0.95, characterizing the joint events of the yield of wild plants. The probability of joint events was calculated by the formula:

\[ p = \sum_{\alpha=1}^A p_\alpha - \sum_{\alpha=1}^{A-1} \sum_{\beta=\alpha+1}^A p_\alpha p_\beta + \sum_{\alpha=1}^{A-2} \sum_{\beta=\alpha+1}^{A-1} \sum_{\gamma=\beta+1}^A p_\alpha p_\beta p_\gamma - \ldots + (-1)^{A-1} \prod_{\alpha=1}^A p_\alpha, \]  

where \( p_\alpha, p_\beta, p_\gamma \) are the probabilities of the appearance of the species of wild plants.
In this example, six probabilistic estimates are used. The exception was willow-herb.

Table 2. The results of solving the mathematical programming problem with interval and probabilistic parameters on harvesting of wild-growing products in the Ust-Ilimsk district.

| Estimation | Cowberry | Huckleberry | Bog | Whortleberry | Willow-herb | Milk | Yellow boletus | Honey agaric | The objective function, million US dollars |
|------------|----------|-------------|-----|--------------|-------------|------|---------------|-------------|------------------------------------------|
| $p_1 = 0.05$ | 1 427.80 | 220.70 | 448.30 | 138.52 | 0.20 | 0.20 | 39.80 | 9.166 |
| $p_2 = 0.5$  | 1 320.41 | 220.70 | 448.30 | 138.52 | 3.24 | 364.00 | 182.00 | 11.494 |
| $p_3 = 0.95$  | 1 427.80 | 220.70 | 448.30 | 138.52 | 312.00 | 364.00 | 182.00 | 12.994 |

The results shown in table 2 are highlighted from 1,000 solutions to each problem. An increase in the number of iterations above this value leads to a discrepancy between the values of the objective function not exceeding 1%. It should be borne in mind that the assumption of using the laws of distribution of harvest of wild plants to assess their productivity significantly coarsens the model.

We turn to the results of modeling the harvesting of wild food products for clusters of the Irkutsk region. Table 3 shows the optimal solutions for the Ust-Ilimsk cluster, which includes five enterprises, and all clusters in the region. The results were obtained according to the model (1) - (6).

Table 3. The results of solving the mathematical programming problem with interval parameters for harvesting wild-growing products for the Ust-Ilimsk and all clusters of the Irkutsk region according to the plan for 2021.

| Estimation | Walnut cedar | Berries | Mushrooms | Bracken | Wild garlic | Herbal plants | The objective function, million US dollars |
|------------|--------------|---------|-----------|--------|-------------|--------------|------------------------------------------|
| Lower      | -            | 14.0    | 15.0      | 2.0    | 3.0         | 58.0         | 0.598 |
| Median     | -            | 17.0    | 9.0       | 1.0    | 3.0         | 60.0         | 1.028 |
| Upper      | -            | 24.0    | 9.0       | 1.0    | 2.0         | 101.0        | 1.609 |
| Clusters of the Irkutsk region (Angarsk, Irkutsk, Zalari, Ulkan, Ust-Ilimsk) | | | | | | | |
| Lower      | 535.00       | 368.00  | 265.00    | 144.00 | 93.00       | 381.00       | 7.565 |
| Median     | 681.00       | 472.00  | 286.00    | 160.00 | 110.00      | 413.00       | 11.241 |
| Upper      | 763.00       | 474.00  | 314.00    | 197.00 | 98.00       | 597.06       | 16.134 |

The maximum income of the Ust-Ilimsk cluster can reach more than one million US dollars. Medicinal plants make up about 19 - 25% of the total yield of wild plants in all clusters of the Irkutsk region. A comparison of the results given in tables 2 and 3 shows great opportunities for the development of the Ust-Ilimsk cluster.

According to the optimal solutions found, there are large discrepancies between the optimal plans with lower and upper bounds. The number of simulated optimal plans using the Monte Carlo method depends on the required accuracy of the solution to the problem. Repeated experiments show that the number of cycles can be limited to 300. Problem (1) - (6) is solved for the condition that the coefficients for unknowns in the objective function and the constraints are independent. We confirmed this fact according to the results of modeling the synchronism of long-term series of harvesting wild-growing food products [26].
4. Discussion
The result of the work is a software package that allows simulating various optimal plans depending on interval estimates. The complexity of using the stochastic linear programming problem is due to the lack of long-term data on the yield of berries, pine nuts, mushrooms and medicinal plants. In this case, one can use the obtained probability distribution laws for individual territories according to the region. Obviously, the simulation results using this technique are very approximate. Here, to obtain optimal plans that are more close to the real situation, one can use expert estimates taking into account field studies and questionnaires of procurers.

It is also of great importance for planning to solve the problem of forecasting the volume of wild plants in increments of up to 5 years and more. Meanwhile, a certain periodicity in the variability of productivity of different wild crops is not observed. According to the method of mechanical smoothing, the smoothest data for three years is the closest to the empirical values of the harvesting of wild-growing products. However, averaging over three-year periods also leads to large calculation errors. Therefore, further research is needed to supplement the software package with other options for optimization models and prognostic techniques.

5. Summary
A software package for modeling the harvesting of wild-growing food products is proposed, including information, mathematical, algorithmic and software. The developed database includes numerical information, expert estimates and maps of the location of territories of each type of wild-growing resource.

As mathematical support, mathematical programming models with interval and random estimates are proposed. Optimization models with interval estimates are preferable to models with random variables. The second models can be used to assess the risks of harvesting wild plants. At the same time, their low accuracy should be borne in mind.

The developed algorithms for obtaining optimal solutions in the form of upper, lower, and median estimates of the objective function, which correspond to optimal plans, contribute to the selection of effective managerial decisions.

The work of the software package was tested on the example of real objects in the Irkutsk region, in the territory of which about 30 main species of wild plants are harvested.

The proposed mathematical, algorithmic, and informational support can be used to manage harvesting in the territories of those countries where wild resources are harvested. At the same time, planning harvesting of wild plants is strongly related to information support, as well as the specific nature of climatic conditions and the variability of wild-growing food species.

Acknowledgment. The some results study was obtained with funded by RFBR according to the research project no. 19-07-00322

References

[1] International Trade Centre UNCTAD/WTO: Overview of World Production and Marketing of Organic Wild Collected Products (Geneva: ITC) 2007 p 91
[2] Boltvina E and Ivanyo Ya 2016 Models for optimizing the harvesting of wild-growing products with interval parameters (Herald of ISTU) 6 73-81
[3] Method of determining stocks of medicinal plants 1986 (Central Research Institute of State Forestry of the USSR, Moscow) [in Russian]
[4] Report on research work on the topic “Zoning potential reserves of wild-growing resources of the Irkutsk region on harvesting priorities” (final report). State Contract no F.201.30.305501 dated 07.24.2018 [in Russian]
[5] Majundar D, Govil J and Singh V 2002 Edible Wild Plants (In book: Recent Progress in Medicinal Plants, Sci. Tech Publishing LLC: Texas, USA) 8 431-466
[6] Vashukevich Yu and Kruzhkov N 2013 *The history of the Irkutsk trust of cooperative industrial farms* (Irkutsk: Publishing house of the Irkutsk State Agrarian University) [in Russian]

[7] Bonishko S 1923 *Cedar fishery in the forests of the Baikal and Transbaikalia and its future prospects* (Materials for studying the forests of the Far East issue 2 Tipo - lithograph of the Expedition of the Preparation of State Papers, Chita) [in Russian]

[8] *Decree of February 9, 2009 N 23-p on approval of the forest plan of the Irkutsk region* (as amended by decrees of the governor of the Irkutsk region of 26.11.2013, no 445-ug, of 26.03.2015 no 50-ug, of 20.01.2016 no 9) [in Russian]

[9] Ivanyo Ya et al 2017 *Potential reserves of wild-growing resources of the Irkutsk region* (Irkutsk: Publishing house of the Irkutsk State Agrarian University) [in Russian]

[10] Koziakov S 1972 On mapping of mushroom & berry resources *Rastitel Resursy* 8(1) 132-139 [in Russian]

[11] Leontiev D, Novak L, Sopin L, Chudnovskaya G 1991 *Guidelines for the study of wild resources* (IAI, Irkutsk) [in Russian]

[12] Vasilkov B 1968 *Methods of accounting for edible mushrooms in the forests of the USSR* (Science, Leningrad) [in Russian]

[13] Aiken M, Sheng O, Vogel D 1991 *Integrating Expert Systems with Group Decision Support Systems* (ACM Transactions on Information Systems) no 9:1 pp. 75–95.

[14] Egorova A and Kozlova S 2006 Information systems: methods and design tools *Scientific Herald of MSTU GA* 105 84-92 [in Russian]

[15] *Embarcader* (Forums, https://community.idera.com/developer-tools/p/forums, last accessed 2019/02/01)

[16] Lipaev V 2006 *Software engineering* (Methodological foundations (textbook). TEIS, Moscow)

[17] Martin J 1980 *Organization of databases in computing systems* (World, Moscow)

[18] *Russian Postgres Pro and PostgreSQL documentation* (https://postgrespro.com/docs, last accessed 2019/02/01)

[19] Pulwarty R and Sivakumar M 2014 Risk management *Weather and Climate Extremes* 3 14-21

[20] Fidler M, Nedoma Y, Ramik Ya, Ron I and Tsimmermann K 2008 *Problems of linear optimization with inaccurate data* (SRC "Regular and chaotic dynamics"; Institute of Computer Science, Izhevsk)

[21] Ivanyo Ya, Petrova S, Polkovskaya M, Popkova Yu 2018 *Risk management of agrarian production under the influence of heavy rains and droughts* (“Problems and prospects of sustainable development of the agro-industrial complex”: materials of the All-Russian scientific-practical conference with international participation, dedicated to the memory of AA Ezhevsky) pp 125-134 [in Russian]

[22] Shankar S and Quiggin J 2013 Production under uncertainty: a simulation study *Journal of Productivity Analysis* 39 207-215

[23] Yudin D 1979 *Tasks and methods of stochastic programming* (Ow. Radio, Moscow)

[24] Zimmermann K 2006 *Interval linear systems and optimization problems over max-algebras* (In Linear Optimization Problems with Inexact Data, Springer, New York) pp 165–193

[25] Zimmermann K, Korte B, Ritter K 1984 *Some optimization problems with extremal operations* (In Mathematical Programming at Oberwolfach II, eds, vol 22 of Mathematical Programming Studies, Springer, Berlin) pp 237-251

[26] Ivanyo Ya, Luzan A, Petrova S, Polkovskaya M 2017 Probabilistic models for estimating the harvesting of wild-growing products in the Irkutsk region *Actual issues of agrarian science* 25 62-68