Using evolutionary algorithms to determine the environmental projects effectiveness

I B Mamai¹, B I Savelyev¹, S V Pronichkin¹,²,³, and A V Kholstov⁴

¹Federal Center of Theoretical and Applied Sociology of Russian Academy of Sciences, 24/35, Krzhizhanovsky Street, Moscow, 117218, Russia
²Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, 40, Vavilov Street, Moscow, 119333, Russia
³National University of Science and Technology "MISiS", 4, Leninsky Prospect, Moscow, 119049, Russia
⁴N.N. Semenov Federal Research Center for Chemical Physics Russian Academy of Sciences, 4, Kosygina Street, Moscow, 119991, Russia

E-mail: ibmamay@yandex.ru

Abstract. Environmental projects have a high degree of uncertainty and risk of their implementation. In the process of assessing their effectiveness, it is necessary to take into account the interests of all the participants in the environmental project. The environmental projects effectiveness evaluation is formalized in the form of a mathematical problem of nonlinear programming with Boolean variables with constraints such as equality and inequality. To solve it, the interactive genetic algorithm for the environmental projects implementation trajectory formation was developed.

1. Introduction

The development and use of systems for assessing the sustainability of ecosystems is essential to address the depletion of natural resources, as well as the increase in energy consumption and associated greenhouse gas emissions in large cities and suburbs. Composite indicators are an effective means of assessing, comparing and communicating indicators of ecosystem resilience, stimulating active participation and involvement of a wide range of stakeholders in decision-making processes. The efficiency in the use of natural resources and improvement in environmental performance are usually given priority over other dimensions. The peculiarities of tools for ensuring the sustainability of complex socio-economic systems are manifested in innovative projects, which are currently more focused on the environment than on the social and economic dimensions of sustainability. Sustainable development of ecosystems should take into account, first of all, the need to achieve a long-term balance between social and economic characteristics of environmental projects.

Environmental projects have a high degree of uncertainty and the risk of their implementation; a long period passes from idea to implementation. In addition, there are possible cases of revision, which increase both the direct costs of additional research and the time indicators of their implementation. These features directly affect the assessment of the projects economic efficiency, as they increase costs and the payback period of investments.
2. Investment design of environmental projects

Recently, decision support systems (DSS) for investment design of environmental projects are becoming more widespread [1, 2]. Such systems combine data from internal and external sources and generate analytical information on their basis, thanks to which the subjects of management carry out the decision-making process. With the right approach to building a DSS, all the participants in the environmental project should have a powerful and convenient tool to support decision-making at any stage of an investment project. In this case, the system should be built in such a way as to ensure the formation of retrospective, current and forecast reporting. Such a structure will allow solving the following main tasks of investment design:

- evaluating the effectiveness of investment decisions in the environmental retrospective;
- formation of operational reporting on the current state of the environmental project;
- obtaining predictive information of high value and allowing to take preventive measures to manage an environmental project;
- tracking and modeling the dynamics of the ecological project development, taking into account the control actions.

In the process of investment design, many different factors must be considered, such as cash flows in the form of net discounted income, payback period, return on investment and profitability index. At the same time, the internal rate of return, payback period and profitability should not be used as an environmental project effectiveness criterion, but only as a limitation in decision making. And the use of the principle of "maximum net discounted income" seems to be the most correct and theoretically justified.

In the process of effectiveness assessing, it is necessary to take into account the interests of all the participants in the environmental project. The efficiency assessment should be made not in general, but for some specific participant based on a comparison of cash inflows and cash outflows from the project from this participant. Of course, at the first stage of calculations, when specific participants have not yet decided, it is possible and necessary to assess the environmental project effectiveness for one participant. But this participation efficiency in the environmental project is true if the participant implements the entire project at his own expense. Such a calculation of the environmental project effectiveness for one participant can be carried out at the final stages of investment project the implementation in order to advertise it, in order to attract investors, turn them from potential into real ones by providing information about the high efficiency of the project. And then, at the next stage, calculate the efficiency for each participant, for his own capital.

The use of investment decision support systems allows developing special evolutionary strategies for investing funds of participants in the environmental project [3, 4]. In such systems, the optimal trajectories of management decisions for each project participant are used as the initial population.

In existing papers, the sum of the net discounted income (NDI) of the project participants or their average value is used as an objective function characterizing the optimal trajectory of all the project participants [5, 6]. The algorithms for finding the optimal trajectory are based on excluding investments from the initial trajectory that have the lowest NDI until the condition of exceeding the threshold value is met.

Such algorithms have a number of disadvantages, since they do not take into account the limitation on the payback period for each participant and on the level of profitability; the solution may be a trajectory in which the NDIs for each participant are very different.

3. Social effectiveness of environmental projects

The assessment of public effectiveness implies that the only participant in an environmental project is the whole society, which implements the project entirely at its own expense, evaluates inflows and outflows from the point of view of public interests, not at market prices, but at economic prices, reflecting the usefulness of resources and products from a social point of view - the economic interests of society as a whole, and not in the way the market estimates it. These prices differ significantly from the market ones, and the society receives all the incomes from the project calculated on them (including taxes) and bears all the necessary expenses. It is clear that for such a participant, no loans
and other transfer payments (taxes, subsidies, loan repayment, etc.) should be taken into account in the calculation, since from the point of view of the system as a whole society, they represent a zero financial transaction: one element of the system loses a certain amount, and the other gets it, while the system balance is zero.

When assessing the investment project effectiveness, it is also important to take into account the organizational and technical problems that may arise during its implementation, including changes in the nature of relations between managers and line personnel [7, 8]. These tasks include both the organization and management of personnel, and the determination of new environmental products production volume, taking into account the needs and opportunities of the market, the search and attraction of material and technical resources, ensuring the sale of products, conducting timely settlements with suppliers and consumers, determining the competitiveness of new products, etc. In order to effectively implement R&D, the manager must have clearly defined goals. After all, not every enterprise is able to implement science-intensive environmental technology. At the same time, the range of tasks that need to be solved and for which one need to answer is significantly expanding. It is necessary to implement a system of measures to ensure a consistent and predictable for the long term process of mastering the results of scientific and technical activities, taking into account the requirements for the efficiency of the natural resources use by enterprises, the safety of products (services) for the environment and public health and the reduction of energy and material consumption, as well as the determination a system of appropriate rewards and sanctions. The specified system should include such measures as improving the qualifications of management personnel, promoting cooperation between manufacturers, stimulating the formation of manufacturers’ associations, encouraging those who buy and apply new environmental technologies, reducing or abolishing customs duties on the import of modern equipment, personnel training, policy public procurement and the provision of preferences to companies and products that use certain technological solutions [9].

4. Genetic algorithms for the environmental projects implementation the trajectory formation

The following formulation of the problem of forming a trajectory for the environmental projects implementation is proposed - it is required to select one of the many options for management decisions that has the maximum NDI so that NDIs in the group of participants in the investment project do not differ much, and the internal rate of return, payback period and profitability index do not exceed a given value.

The mathematical formulation of the problem is the objective function (1) as a parametric convolution of the mean and standard deviation of NDI $K_d$ for each participant in the environmental project $p_{sd}$. Internal rate of return $IRR_d$, payback period $PP_d$ and profitability $PI_d$ form constraints (2).

$$
g(ps) = (1 - \lambda) \frac{\sum_{d \in D} K_d \cdot p_{sd}}{m} + \lambda \sqrt{\frac{\sum_{d \in D^*} (K_d - K_d \cdot p_{sd})^2}{m - 1}} \rightarrow \min_{ps \in PS^*}.
$$

where $ps = (p_{sd})_{d \in D^*}$ is the group of participants in the environmental project; $PS^* = \{PS_d\}_{d \in D^*}$ is a set of potential participants in the environmental project;

$$ps_d = \begin{cases} 1, & \text{if the } d\text{-th potential participant participates in the project} \\ 0, & \text{otherwise} \end{cases}$$

is the average NDI; $\lambda$ is a model parameter, $\lambda \in [0,1]$. 

where

$$K = \frac{\sum_{d \in D} K_d \cdot p_{sd}}{m}$$

is the
\[
\begin{align*}
\sum_{d \in D} ps_d &= m; \\
\sum_{d \in D} C_{dv} \cdot ps_d &\leq b_v,
\end{align*}
\]

where \( C_{dv} \) is the interest of the participation \( d \)-th potential participant in the environmental project \( v \in \{IRR, PP, PI\} \).

The formulated equation is a nonlinear programming equation with Boolean variables with constraints such as equality and inequality. The upper bound for the complexity of the exact algorithm for solving the equation is exponential, i.e. the task is np-full.

To solve the problem, the algorithm was developed, which is based on the approach of evolutionary computations - genetic algorithms. Genetic algorithms make it possible to obtain a “good” solution relatively quickly (in a finite number of steps) [10]. The proposed algorithm consists of the following basic operations: generation of the initial population; selection of parental couples; the use of crossover operators; calculation of the fitness function; selection; checking the condition of population degeneration, if it is satisfied, apply the inversion operator; check the break condition, if it is not met go to the selection of parent pairs.

1. Set \( T \) – number of iterations, \( m, b, \varepsilon, E \) – population size, \( \lambda = 0 \).

2. \( t = 0 \). Generate (uniform distribution) the initial population \( P^t = \{p_1^t, \ldots, p_E^t\} \), where \( p_e^t = (p_{ed}^t)_{d \in D} \) – \( e \)-th chromosome of the population, \( e = \overline{1, E} \), is satisfying (2).

Calculate \( v(p_e^t) = (v_1(p_e^t), v_2(p_e^t)) = (s^{C_1}_{e}, s^{C_2}_{e}), s^{C_1}_{e} = s^{C_1}_{e} + s^{C_2}_{e}, s^{C_1}_{e}, s^{C_2}_{e} \) – the number of applications of the crossing over operators for \( e \)-th chromosome at the iteration \( t \), for \( t = 0 \),

\[ v(p_e^t) = \left( \frac{1}{2}, \frac{1}{2} \right), s^{C_1}_{e}, s^{C_2}_{e} = 0. \]

3. Calculate \( F^t = \{f_1^t, \ldots, f_E^t\} \), \( f_e^t = g(p_e^t), e = \overline{1, E} \) by (1).

4. Sort \( F^t \), \( F^t = \{f_1^t : f_1^t \leq f_{e+1}^t, e = \overline{1, E-1} \} \). Find quartiles \( f_{\frac{1}{4}}^t : p(f_{1/4}^t < f_e^t) = \frac{1}{4} \) and \( f_{\frac{3}{4}}^t : p(f_{3/4}^t < f_e^t) = \frac{3}{4} \), \( e = \overline{1, E} \). Select randomly (uniform distribution) \( p_{\frac{1}{4}}^t : [f_1^t \leq g(p_{\frac{1}{4}}^t) < f_{\frac{1}{4}}^t], \)

\[ p_{\frac{3}{4}}^t : [f_1^t \leq g(p_{\frac{3}{4}}^t) \leq f_{\frac{3}{4}}^t], \]

\[ p_{\frac{1}{4}}^t : [f_{\frac{1}{4}}^t < g(p_{\frac{1}{4}}^t) \leq f_1^t], \]

\[ p_{\frac{3}{4}}^t : [f_{\frac{3}{4}}^t < g(p_{\frac{3}{4}}^t) \leq f_E^t], \]

\( e = \overline{1, E} \) from \( P^t \). Find

\[ p_{\frac{1}{4}}^t = \arg \min_e D_H(p_{\frac{1}{4}}^t, p_e^t), \]

\[ p_{\frac{3}{4}}^t = \arg \min_e D_H(p_{\frac{3}{4}}^t, p_e^t), \]

\[ p_{\frac{1}{4}}^t = \arg \max_e D_H(p_{\frac{1}{4}}^t, p_e^t), \]

\[ p_{\frac{3}{4}}^t = \arg \max_e D_H(p_{\frac{3}{4}}^t, p_e^t) \), where \( D_H \) is Hamming distance.

5. For \( p_{\frac{1}{4}}^t, p_{\frac{3}{4}}^t, p_{\frac{1}{4}}^t, p_{\frac{3}{4}}^t \) choose randomly with probability \( v_1(p_{\frac{1}{4}}^t), x_1 \in \left\{\epsilon_{\frac{1}{4}}, \epsilon_{\frac{3}{4}}, \epsilon_{\frac{1}{4}}, \epsilon_{\frac{3}{4}}\right\} \)

operator \( C_1 \), or \( C_2 \) with probability \( v_2(p_{\frac{1}{4}}^t), x_1 \in \left\{\epsilon_{\frac{1}{4}}, \epsilon_{\frac{3}{4}}, \epsilon_{\frac{3}{4}}, \epsilon_{\frac{3}{4}}\right\} \). In the case of choosing a one-
point crossing-over operator \( C_i(p'_{x_1}, p'_{x_2}) \), where \( x_1 \in \{ p'_{x_1}, p'_{x_2}, p_{x_1}, p_{x_2} \} \), \( x_2 \in \{ p'_{x_1}, p'_{x_2}, p_{x_1}, p_{x_2} \} \), generate

\( r \) (uniform distribution \((1, |D^v|)\)), calculate \( p'^*_1 = (p'^*_d)_{d=1}^{|p'^*_1|} \), where \( p'^*_d = \begin{cases} p'_{x_1d}, & \text{if } d < r \\ p'_{x_2d}, & \text{if } d \geq r \end{cases} \),

\( s'^{c_1}_{x_1} = s^{c_1} + 1 \), and \( p'^*_2 = (p'^*_d)_{d=1}^{|p'^*_2|} \), where \( p'^*_d = \begin{cases} p'_{x_2d}, & \text{if } d < r \\ p'_{x_1d}, & \text{if } d \geq r \end{cases} \), \( s'^{c_1}_{x_2} = s^{c_1} + 1 \). In the case of choosing a two-point crossing-over operator \( C_2(p'_{x_1}, p'_{x_2}) \), where \( x_1 \in \{ p'_{x_1}, p'_{x_2}, p_{x_1}, p_{x_2} \} \), \( x_2 \in \{ p'_{x_1}, p'_{x_2}, p_{x_1}, p_{x_2} \} \), generate \( r_1 \) and \( r_2 \) (uniform distribution \((1, |D^v|)\)), let \( r_1 > r_2 \), calculate

\[
p'^*_1 = (p'^*_d)_{d=1}^{|p'^*_1|} \text{ with } p'^*_d = \begin{cases} p'_{x_1d}, & \text{if } d < r_1 \\ p'_{x_2d}, & \text{if } r_1 \leq d \leq r_2 \\ s'^{c_2}_{x_1} = s^{c_2} + 1 \end{cases} \text{ and } p'^*_2 = (p'^*_d)_{d=1}^{|p'^*_2|} \text{ with } p'^*_d = \begin{cases} p'_{x_2d}, & \text{if } d < r_2 \\ p'_{x_1d}, & \text{if } r_2 \leq d \end{cases}.
\]

6. Check for \( p'^*_1 \) and \( p'^*_2 \) where \( x_1 \in \{ p'_{x_1}, p'_{x_2}, p_{x_1}, p_{x_2} \} \), \( x_2 \in \{ p'_{x_1}, p'_{x_2}, p_{x_1}, p_{x_2} \} \) the fulfillment of (2). Add in \( P' = \{ p'_1, \ldots, p'_q \} \), \( q \) new chromosomes \( p'^*_1 \) and \( p'^*_2 \) where \( x_1 \in \{ p'_{x_1}, p'_{x_2}, p_{x_1}, p_{x_2} \} \), \( x_2 \in \{ p'_{x_1}, p'_{x_2}, p_{x_1}, p_{x_2} \} \), are satisfying (2). Calculate \( F' = \{ f'_1, \ldots, f'_{q+1} \} \).

7. Conduct sorting \( F' = \{ f'_e : f'_e \leq f'_{e+1}, e = 1, E + q - 1 \} \). Conduct an "elite selection" to take the first \( E \) individuals.

8. Check the fulfillment of the inequations \( |f'_1 - f'^1| \leq \varepsilon \) and \( t + 1 < T \). If the inequations are not met, then go to step 10.

9. Set \( L \). Choose randomly (uniform distribution) \( p'_{e_1}, p'_{e_2}, \ldots, p'_{e_L} \) from \( P' \setminus \{ p'_{e} \} \), where \( f'_1 = g(p'_{e}) \). For \( p'_{x_1}, x_3 = (e_1, e_2, \ldots, e_L) \) apply the "inversion" operator \( R(p'_{x_1}) \), generate \( r_1 \) and \( r_2 \) (uniform distribution \((1, |D^v|)\)), calculate \( p'^*_3 = (p'^*_d)_{d=1}^{|p'^*_3|} \), where \( p'^*_d = \begin{cases} p'_{x_1d}, & \text{if } d \neq r_1, r_2 \\ p'_{x_2d}, & \text{if } d = r_1 \\ p'_{x_3d}, & \text{if } d = r_2 \end{cases} \).

10. \( t = t + 1 \).
11. Check the inequation \( t \leq T \), if the inequation is satisfied go to step 3.

12. Remember \( \lambda \) and \( p'_i \), such that \( g(p'_i) = f'_i \), \( \lambda = \lambda + 0.1 \).

13. Check the inequation \( \lambda \leq 1 \), if the inequation is satisfied go to step 2.

14. If the solution is not found go to step 1.

The main difference between the proposed genetic algorithm and the existing genetic algorithms [11-13] is the step - “choosing a parental pair”. It is proposed to use outbreeding and inbreeding when choosing "parental pairs", and the use of quartiles makes the choice of "parental pairs" robust to the distribution of chromosomes, which is confirmed by computational experiments. In the proposed algorithm, new individuals obtained as a result of the implementation of the crossing-over operator do not replace their parents, but form an intermediate population with them, to which the “elite” selection operator is subsequently applied. Also the inversion operator is used instead of the mutation operator based on the specifics of the problem.

5. Conclusion
The proposed algorithm makes it possible to search more efficiently in local optima, which actually leads to the division of the population into separate local groups around trajectories suspicious for an extremum with a shift towards the global optimum, which is confirmed by testing on experimental data.

At the same time, the proposed approach, together with the parameterization of the "inversion" operation, is aimed at preventing the convergence of the algorithm to the already found local solutions and allows the decision-maker to view new, unexplored combinations in an interactive mode. As the result, the decision maker of the DSS receives the optimal trajectory for the participants of the environmental project, which has the maximum NDI. Moreover, the NDI of the project participants are not very different, and the internal rate of return, payback period and profitability index do not exceed the specified values.

6. Acknowledgments
The article is prepared with the financial support of the Russian Science Foundation, project № 19-78-10035.

References
[1] Zasada I, Piorr A, Novo P, Villanueva A and Valanszki I 2017 Environmental Modelling & Software 98 63-74
[2] Walling E and Vaneckhaute C 2020 Journal of Environmental Management 264 110513
[3] Kamenopoulos S and Agioutantis Z 2020 The Extractive Industries and Society 23 100740
[4] Diez-Rodriguez J, Fischer T and Zio S 2019 Journal of Cleaner Production 220 1239-54
[5] Wu Y, Wu C, Zhou J, He F and Zhang T 2020 Journal of Energy Storage 30 101601
[6] Knoke T, Gosling E and Paul C 2020 Ecological Economics 174 106644
[7] Istrate L, Marian L and Ferencz I 2014 Procedia Economics and Finance 15 1732-9
[8] Pandremmenou H, Sirakoulis K and Blanas N 2013 Procedia - Social and Behavioral Sciences 74 438-47
[9] Ramazani J and Jergeas G 2015 International Journal of Project Management 33 41-52
[10] Costa-Carrapico I, Raslan R and Gonzalez J 2019 Energy and Buildings 210 109690
[11] Wang Z and Sobey A 2019 Composite Structures 233 111739
[12] Lee C 2018 Engineering Applications of Artificial Intelligence 76 1-12
[13] Saracoglu I, Topaloglu S and Keskinturk T 2014 Expert Systems with Applications 41 8189-202