

About a regional development model that takes into account environmental problems with budgeting uncertainty

Artem Novikov*

Sobolev Institute of Mathematics of the Siberian Branch of the Russian Academy of Sciences,
Editorial Department, 630090 Novosibirsk, Russia

Abstract. Raw-materials base (hereinafter RMB) is one of the largest industries for financial investments in Russia. There are various mathematical descriptions for the development of regions with resource-based economy. Earlier in [1] the researchers considered the model based on bilevel integer stochastic programming problems with Boolean variables. This paper proposes a new approach to public-private partnership modelling, including a bilevel linear stochastic programming problem. This model assumes that budget constraints of the state and the investor can vary in a random manner with a specific probability distribution. We put forward two methods to solve this problem: problem's reduction to the deterministic bilevel one and formulation of deterministic problems sequence with help of Monte Carlo methods. In order to solve deterministic problems of integer programming, we suggest two approaches: direct enumeration and heuristic "Game" approach. The numerical experiments for proposed algorithms validation are conducted on the basis of actual data of Zabaykalsky Krai development. Multiple input parameters of the model vary in these experiments. Finally, we present a brief analysis of the obtained solutions to the stochastic linear programming problems with Boolean variables.

1 General statement of planning model

We consider the model containing NP production projects, NI infrastructure projects and NE ecological projects. It is assumed that the notations of all projects for T years (index $t=1, \dots, T$) were developed by the experts.

Production projects:

CFP_i^t - cash flow (difference between cash receipts and cash payments in a given period) of production project i in t year;

EPP_i^t - monetary evaluation of ecological losses during project i implementation in t year;

DBP_i^t - budget revenues from project i implementation in t year;

ZPP_i^t - wages paid during project i implementation, i index possesses the values $i=1, \dots, NP$.

Infrastructure projects:

ZI_j^t - the graph of project j implementation costs in t year;

* Corresponding author: artemka1387@bk.ru

EPI_j^t - monetary evaluation of ecological losses during project j implementation in t year;

VDI_j^t - off-project budget revenues from project j implementation related to general development of region economy;

ZPI_j^t - wages paid during project j implementation, j index possesses the values $j=1, \dots, NI$.

Ecological projects

ZE_k^t - the graph of project k implementation costs in t year ;

EDE_k^t - monetary evaluation of ecological profit during project k implementation in t year;

ZPE_k^t - wages paid during project k implementation, k index possesses the values $k=1, \dots, NE$.

Correlation of projects

μ_{ij} - the indicator of technological interdependence of production and infrastructure projects is equal to 1, if the implementation of infrastructure project j is required for the implementation of production project i and otherwise is equal to 0, $i=1, \dots, NP, j=1, \dots, NI$.

ν_{ik} - interdependence indicator of production and ecological projects is equal to 1, if production project i implementation entails the necessity of ecological project k implementation, and otherwise is equal to 0, $i=1, \dots, NP, k=1, \dots, NE$.

Discounts and budget constraints

DG - discount of the state;

DI - discount of the investor;

DN - discount of the public;

$BudG, BudI$ —are budget constraints of the state and investor set at the beginning of the program.

It is supposed that the state and the investor determine the following values of Boolean variables:

$\xi_i = 1$, if the investor launches the production project i , $\xi_i = 0$ and otherwise;

$\theta_j = 1$, if the state launches the infrastructure project j , $\theta_j = 0$ and otherwise;

$\zeta_k = 1$, if the state launches the ecological project k , $\zeta_k = 0$ and otherwise;

$\omega_l = 1$, if the investor launches the ecological project l , $\omega_l = 0$ and otherwise;

$\lambda_k = 1$, if launching the production project requires compensation for the environmental damage, otherwise $\lambda_k = 0$.

Problem formulation in the previous work [1] did not consider stochastic budget variations. The present paper takes into account state and investor budgets volatilities by using a planning model with random noise. This model can be written as a bilevel integer stochastic programming problem consisting of state problem and investor problem at the top and low levels respectively.

Problem of the state:

To maximise the average discounted cash flow of the region:

$$M(\sum_{t=1}^T (\sum_{i=1}^{NP} (DBP_i^t + ZPP_i^t - EPP_i^t) * \xi_i^* + \sum_{j=1}^{NI} (VDI_j^t + ZPI_j^t - EPI_j^t - ZI_j^t) * \theta_j + \sum_{k=1}^{NE} (EDE_k^t + ZPE_k^t - ZE_k^t) * \zeta_k) / (1 + DG)^t) \Rightarrow \max_{\theta} \quad (1.1)$$

Subject to

$$\theta_j \geq \xi_i^* * \mu_{ij}, i = 1, \dots, NP, j = 1, \dots, NI, \quad (1.2)$$

$$\zeta_k = \lambda_k^* - \omega_k^*, k = 1, \dots, NE, \quad (1.3)$$

$$P(\sum_{j=1}^{NI} ZI_j^t * \theta_j + \sum_{k=1}^{NE} ZE_k^t * \zeta_k \leq BudG_t * \beta 1_t) \geq \alpha_t, t = 1, \dots, T \quad (1.4)$$

the vectors $(\xi_i^*, \lambda_k^*, \omega_k^*)$ —are found from *optimal solution* of investor problem (see below), $\alpha \beta 1_t$ is considered as a random variable with defined density of distribution φ_G .

Here and elsewhere the following designations are used: $\mathbf{M}(a)$ —is a mathematical expectation of the random variable a , and $\mathbf{P}(a \leq b) \geq \alpha_t$ indicates that the probability of the inequality $a \leq b$ execution does not exceed the value α_t .

Problem of the investor:

To maximise the average of net present value (NPV):

$$\mathbf{M}(\sum_{t=1}^T (\sum_{i=1}^{NP} CFP_i^t * \xi_i - \sum_{l=1}^{NE} ZE_l^t * \omega_l) / (1 + DI)^t) \Rightarrow \max_{\xi, \lambda, \omega} \quad (1.5)$$

with the given vector θ at upper management level and the constraints

$$\xi_i * \mu_{ij} \leq \theta_j, i = 1, \dots, NP, j = 1, \dots, NI \quad (1.6)$$

$$\lambda_l \geq \xi_i * v_{il}, i = 1, \dots, NP, l = 1, \dots, NE \quad (1.7)$$

$$\omega_l \leq \lambda_l, l = 1, \dots, NE \quad (1.7')$$

$$\sum_{l=1}^{NE} \omega_l * \sum_{t=1}^T ZE_l^t \geq D * \sum_{l=1}^{NE} \lambda_l * \sum_{t=1}^T ZE_l^t, \quad (1.8)$$

$$\sum_{t=1}^T (\sum_{i=1}^{NP} CFP_i^t * \xi_i - \sum_{k=1}^{NE} ZE_k^t * \omega_k) / (1 + DI)^t \geq 0, \quad (1.9)$$

$$\mathbf{P}(\sum_{l=1}^{NE} ZE_l^t * \omega_l - \sum_{i=1}^{NP} CFP_i^t * \xi_i \leq BudI_t * \beta 2_t) \geq \alpha_t, t = 1, \dots, T \quad (1.10)$$

$$\sum_{t=1}^T (\sum_{i=1}^{NP} (ZPP_i^t - EPP_i^t) * \xi_i + \sum_{j=1}^{NI} (ZPI_j^t - EPI_j^t) * \theta_j + \sum_{k=1}^{NE} (EDE_k^t + ZPE_k^t) * \lambda_k) / (1 + DN)^t \geq 0 \quad (1.11)$$

where $\beta 2_t$ —is a random variable with the density of distribution φ_I

In this statement D parameter shows the minimal share of investor's costs during all ecological projects implementation and forms the commitments of the type (1.8) for the investor. Objective functions (1.1), (1.5) model DM (decision-makers) criteria, that aimed to maximise the average discounted cash flow. Random budgets for the state and the investor are modelled by the constraints of the type (1.3) and (1.10) respectively. Non-negativeness of investor's discounted cash flow (investment project NPV) is verified by the constraints of the type (1.9). Monitoring the efficiency of region development program in terms of benefits and losses that people receive is formulated by the constraints of the type (1.11).

So that we could see that the formulated model is a problem of bi-level integer stochastic programming, and it generates an optimal mechanism for cooperation between the state and the investor. We should also note that in game theory the problem under consideration is interpreted as "Leader-Follower" model [2] (where the leader is the state, and the follower is the investor). Its solution is defined by Stackelberg equilibrium in this model.

2 Solution to the stochastic problem

In order to solve this problem it is proposed to change the stochastic problem to the deterministic one. Below are the two methods of stochastic optimisation problem solving.

2.1 Construction of deterministic optimisation problems sequence

If this approach is used, the stochastic optimisation problem solution is modelled by the solution of the deterministic optimisation problems succession resulted from the substitution of random parameters with their realizations. The arithmetic means of deterministic problems solutions are defined; and the problem which optimal solution is the nearest to the average found is chosen from the sequence. The following computing scheme appears during the realization of the above approach. Let us consider N random variables realizations $\beta 1_t; \beta 2_t$, having the distribution densities φ_G and φ_I respectively. Let us call these realizations as $\beta 1_t^0$,

$\beta 2_t^0$ respectively. We can receive the random variable realization with the defined destiny using the standard Monte Carlo method “Inverse Function Method” [3] and the standard pseudorandom generator RAND (RAND—generates pseudorandom number with the uniform distribution $U(0;1)$). Then we replace the random variables $\beta 1_t, \beta 2_t$ in constraints (1.4), (1.10) with their realizations. $\beta 1_t^0, \beta 2_t^0$. Since we fix the values of the random variables, the functionals (1.1) and (1.5) will be of type (2.1), (2.5) (according to the property of mathematical expectation $M(x)=x$, if $x=const$), and the constraints (1.4), (1.10) will be of type (2.4), (2.10) (because we fixed the value of the random variable).

$$F_1 = \sum_{t=1}^T (\sum_{i=1}^{NP} (DBP_i^t + ZPP_i^t - EPP_i^t) * \xi_i^* + \sum_{j=1}^{NI} (VDI_j^t + ZPI_j^t - EPI_j^t - ZI_j^t) * \theta_j + \sum_{k=1}^{NE} (EDE_k^t + ZPE_k^t - ZE_k^t) * \zeta_k) / (1 + DG)^t \Rightarrow \max_{\theta} \tag{2.1}$$

$$F_2 = \sum_{t=1}^T (\sum_{i=1}^{NP} CFP_i^t * \xi_i - \sum_{l=1}^{NE} ZE_l^t * \omega_l) / (1 + DI)^t \Rightarrow \max_{\xi, \lambda, \omega} \tag{2.2}$$

$$\sum_{j=1}^{NI} ZI_j^t * \theta_j + \sum_{k=1}^{NE} ZE_k^t * \zeta_k \leq BudG * \beta 1_t^0, t = 1, \dots, T \tag{2.3}$$

$$\sum_{l=1}^{NE} ZE_l^t * \omega_l - \sum_{i=1}^{NP} CFP_i^t * \xi_i \leq BudI * \beta 2_t^0, t = 1, \dots, T \tag{2.4}$$

Thus we get the bilevel deterministic problem derived from the stochastic problem (1.1) - (1.11).

Then the cycle process of solving the deterministic problem is launched, where cycle goes down $i = 1 \dots N$. As a result, we have a sequence of values of state's F_{1_i} and investor's F_{2_i} functionals and find their means as $\bar{F}_1 = \frac{1}{N} \sum_{i=1}^N F_{1_i}$ and $\bar{F}_2 = \frac{1}{N} \sum_{i=1}^N F_{2_i}$ respectively. After that, we find the nearest to the means values of the functional $F_{10} = (F_{1_i} | |\bar{F}_1 - F_{1_i}| \rightarrow min)$ and $F_{20} = (F_{2_i} | |\bar{F}_2 - F_{2_i}| \rightarrow min)$. Then we find the optima $\theta^*, (\xi^*, \lambda^*, \omega^*)$ corresponding to the values of F_{10} and F_{20} . These vectors will be taken as the problem solutions (1.1) - (1.11).

Consider next the second approach to the original stochastic problem solving.

2.2 Reduction of the stochastic problem to the deterministic one

Let us consider the distribution densities of random vectors components:

$$\varphi_{I_i}(\beta 2^i) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \varphi_I(\beta 2_1 \dots \beta 2_m) \prod_{i \neq j} d\beta 2_j \tag{2.5}$$

$$\varphi_{G_i}(\beta 1^i) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \varphi_G(\beta 1_1 \dots \beta 1_m) \prod_{i \neq j} d\beta 1_j \tag{2.6}$$

We shall find the roots $\widetilde{\beta 2}_i, \widetilde{\beta 1}_i$ from the integral equations of types (2.5), (2.6) respectively:

$$\int_{\widetilde{\beta 2}_i}^{\infty} \varphi_{G_i}(\beta 1_i) d\beta 1_i = \alpha_i \tag{2.7}$$

$$\int_{\widetilde{\beta 1}_i}^{\infty} \varphi_{I_i}(\beta 2_i) d\beta 2_i = \alpha_i \tag{2.8}$$

If more than one root is found, we shall choose the maximum one.

Thus, for the problem (1.1) - (1.11) solving it is enough to solve the bilevel deterministic problem of Boolean programming, where the functional of the state will be of type:

$$\sum_{t=1}^T (\sum_{i=1}^{NP} (DBP_i^t + ZPP_i^t - EPP_i^t) * \xi_i^* + \sum_{j=1}^{NI} (VDI_j^t + ZPI_j^t - EPI_j^t - ZI_j^t) * \theta_j + \sum_{k=1}^{NE} (EDE_k^t + ZPE_k^t - ZE_k^t) * \zeta_k) / (1 + DG)^t \Rightarrow \max_{\theta} \tag{2.9}$$

The constraints (1.4) in the problem of the state are substituted with the constraints:

$$\sum_{j=1}^{NI} ZI_j^t * \theta_j + \sum_{k=1}^{NE} ZE_k^t * \zeta_k \leq BudG * \widetilde{\beta 1}_t, t = 1, \dots, T \quad (2.10)$$

The functional in the problem of the investor will be of type:

$$\sum_{t=1}^T (\sum_{i=1}^{NP} CFP_i^t * \xi_i - \sum_{l=1}^{NE} ZE_l^t * \omega_l) / (1 + DI)^t \Rightarrow \max_{\xi, \lambda, \omega} \quad (2.11)$$

and the constraints (1.10) are substituted with the constraints:

$$\sum_{l=1}^{NE} ZE_l^t * \omega_l - \sum_{i=1}^{NP} CFP_i^t * \xi_i \leq BudI * \widetilde{\beta 2}_t, t = 1, \dots, T \quad (2.12)$$

Methods of the bilevel deterministic problem solving

Earlier investigations [4] proved that the problems of bilevel linear integer programming are NP hard. Hence, the problems obtained in sec.2 belongs to the class of hard mathematical programming problems. In general case if only $P \neq NP$, there is no algorithm for finding the optimal problem solution in polynomial time. This implies that any exact algorithm of such problem solution will have the exponential complexity at worst.

3.1 Partial enumeration

This approach can be used primarily for small dimension problems. The method is as follows.

All possible variants of the vector components values θ are searched through. The number of variants of enumeration does not exceed 2^{NI} . We shall note that the constraints (2.10) significantly reduce the number of possible variants θ_i . To ensure this, it is enough to check (2.10) for each variant of vector θ , assuming all $\zeta_k = 0$. Then for the permissible set θ_i ($i=1, \dots, NI$) we solve the problem of mathematical programming with Boolean variables $\xi_i, \lambda_k, \omega_l$. It is to be noted that if we solve this problem by enumerating the values of variables, the number of enumeration variants for investor's problem will not exceed 2^{NI+NE} . The constraints (2.12) reduce the number of variants enumerated for the low-level problem, as in the case with the constraints (2.10) for the state problem. When the value $NI+NE$ is small (e.g. less than 50) we can use the existing standard mathematical programming systems (e.g. Gams, MS Excel, PYOMO) for the problem solving.

3.2 "Game" approach

For the first time "Game" approach was described in [5]. This approach can be used primarily for high dimension problems. The low-level problem is supplemented by the constraints (2.10) and a constraint to the functional of the state.

$$\sum_{t=1}^T (\sum_{i=1}^{NP} (DBP_i^t + ZPP_i^t - EPP_i^t) * \xi_i^* + \sum_{j=1}^{NI} (VDI_j^t + ZPI_j^t - EPI_j^t - ZI_j^t) * \theta_j + \sum_{k=1}^{NE} (EDE_k^t + ZPE_k^t - ZE_k^t) * \zeta_k) / (1 + DG)^t \geq N \quad (3.1)$$

Then the iteration process of maximum N search is carried out by the dichotomy method, according to which the Boolean programming problem with linear constraints has an optimal solution of "Leader-Follower" game, where the state plays the role of the leader; in the model proposed it is reasonable for the investor to start the enumeration of variants.

The number $\log_2(N_{max} - N_{min})$ of process iterations is the upper estimate (very rough) for solution complexity, this indicates the enumeration effectiveness.

4 Numerical experiment description

Our numerical experiments were conducted on the basis of the actual data of Zabaykalsky Krai development. This example contained the information about the projects launched for 20 years ($T=20$), and described 5 different infrastructure ($NI=5$), 10 production ($NP=10$), and 10 ecological projects ($NE=10$).

4.1 Construction of the deterministic optimisation problems sequence

While investigating this method of the stochastic programming problem solving, the parameters DG and DI assumed the values 3%, 8%, 13%, 18% and 10%, 15%, 20%, 25% respectively. For this method we modelled the random variables $\beta 1_t$, $\beta 2_t$ with the normal distribution $N(1;0,5)$ different for each T year, then we considered the process from par. 2.1, where $N=40$. The modelling of a random variable was conducted by the typical Monte Carlo method called the “Inverse Distribution Function” method. Since for the experiment we chose the normal distribution, then for the random variables modelling we used Box–Muller formulas [6]

$$\beta 1_t = \frac{1}{2} \sqrt{-2 \ln \alpha_1} \cos \alpha_2 + 1 \quad (4.1)$$

$$\beta 2_t = \frac{1}{2} \sqrt{-2 \ln \alpha_1} \sin \alpha_2 + 1 \quad (4.2)$$

where α_1, α_2 are independent uniformly distributed random variables in the interval $(0,1)$. Then we solved the problems using the “Game” approach.

Through this example we obtain the following dependences of state and investor functionals values on the discounts DG and DI (fig. 1).

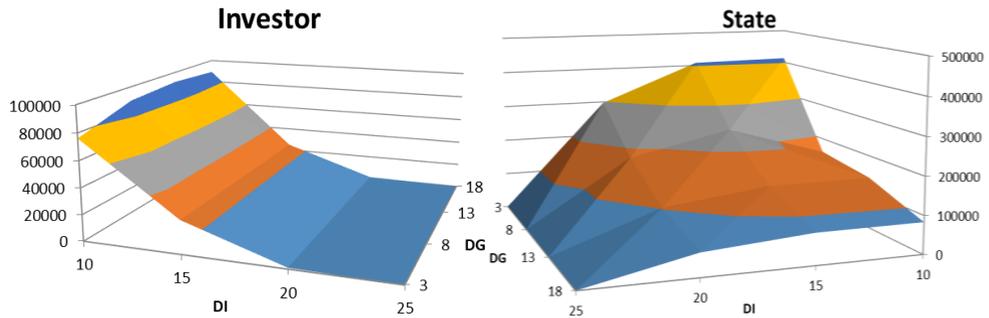


Fig. 1. Investor dynamics (left) and state dynamics (right)

In this approach we observed the instability of the problem due to the lack of the acceptable solutions under small variation of input parameters. We also noted that the value of the investor functional does not depend on the state discount varying the and value of the state functional does not depend on both discounts. Obviously, this fact is true only with the examined coefficients of the bilevel problem.

4.2 Reduction to the deterministic problem

While investigating this method of the stochastic programming problem solution, the parameters DG and DI took the values 8% and 10% respectively. In this method for the values $\beta 1_t$, $\beta 2_t$ we took Simpson’ triangular distribution, equal for each year, in the interval

$(-0,5;0,5)$. Through this example we obtained the following dependences of state and investor functionals values on the confidence level α :

Table 1. Experiment parameters

α	$\widetilde{\beta}1^t$	$\widetilde{\beta}2^t$	State functional	Investor functional
1	0	0	181862.3	89265.65
0.95	0.81	0.81	No solutions	No solutions
0.95	0	0.81	No solutions	No solutions
0.95	0.81	0	181862.3	89265.65
0.9	0.82	0	181862.3	89265.65
0.9	0	0.82	No solutions	No solutions
0.9	0.82	0.82	No solutions	No solutions
0.8	0.86	0.86	No solutions	No solutions
0.8	0.86	0	181862.3	89265.65
0.8	0	0.86	No solutions	No solutions
0.5	1	1	181862.3	89265.65
0.5	1	0	181862.3	89265.65
0.5	0	1	181862.3	89265.65

In this approach we observe the instability of the problem set due to the lack of acceptable solutions under a small variation of the input parameters.

5 Conclusion

To sum up, in this study we have constructed a new public-private partnership model for resource region development program formation using the methods of mathematical modelling, stochastic programming and operations research. We has investigated different approaches to bilevel stochastic programming problem solution and conducted numerical experiments.

The work was carried out within the framework of the state contract of the Sobolev Institute of Mathematics (projects no. 0314-2019-0018) and the Russian Foundation for Basic Research (projects no. 19-010-00910 and no. 20-010-00151)

References

1. S.M. Antzys, S.M. Lavlinsky, I.S. Kalgina, *Transbaikal State University J.*, **11(102)**, 118–125 (2013)
2. Moulin H. *Studies in Game Theory and Mathematical Economics Translated from French*. (Mir, Moscow, 1985)
3. Voitishchik A.V. *Basics of Monte-Carlo Methods: Textbook*. (Novosibirsk State University, Novosibirsk, 2010)
4. O. Ben-Ayed, Ch. E. Blair, *Operation Research*, **38(3)**, 556–560 (1990)
5. Antzys S.M. and others. *Optimization of System Solutions in the Distributed Databases*. (Nauka. Siberian branch, Novosibirsk, 1990)
6. G.E.P. Box, Mervin E. Muller, *The Annals of Mathematical Statistics*, **29(2)**, 610–611 (1958)