CZASOPISMO INŻYNIERII LĄDOWEJ, ŚRODOWISKA I ARCHITEKTURY
JOURNAL OF CIVIL ENGINEERING, ENVIRONMENT AND ARCHITECTURE
JCEEA, t. XXXVI, z. 66 (4/19), październik-grudzień 2019, s. 5-16, DOI:10.7862/rb.2019.19

Leonid DVORKIN[1]
Grzegorz PROKOPSKI[2]
Vitaliy MARCHUK[3]
Andriy HUTS[4]

THE USE OF MATHEMATICAL EXPERIMENTS PLANNING IN BUILDING MATERIALS QUALITY RESEARCH

The issue of ensuring the construction of materials and products, necessary quality characteristics and durability, is associated with the solution of many problems arising both in the construction industry enterprises and in the conditions of a construction site. Most of these (construction-technological) problems are solved due to the rational choice of raw materials, changes in the ratio between the main components, as well as production modes. The method of solving such problems assumes simultaneous consideration of many factors (composition, consumption of components, formation conditions, hardening, etc.) and providing many parameters (workability, strength, frost resistance, water resistance, etc.), which in practice is difficult to implement by traditional methods of experimentation because of high labor intensity of the work and the need for complex analysis. The article considers using methods of system analysis and mathematical planning of experiments in the study of building materials quality. Methods of obtaining mathematical models of various types, their interpretation and analysis, as well as the basic principles of using system analysis for solving the problems of building materials technology are presented. The technique of mathematical planning of experiments is presented. The main types of tasks solved with this method are analyzed, optimization criteria for solving problems of concrete technology are formulated. Methods for constructing linear and nonlinear models, their statistical analysis, and typical plans for carrying out experiments are presented.

Keywords: mathematical planning of experiment, polynomial models, response surface, regression equations, variation levels

[1] Leonid Dvorkin, Institute of Building and Architecture, The National University of Water and Environmental Engineering, 11 Soborna Str, 33-028, Rivne, Ukraine, https://orcid.org/0000-0001-8759-6318, l.i.dvorkin@nuwm.edu.ua
[2] Grzegorz Prokopski, The Faculty of Civil and Environmental Engineering and Architecture, Rzeszow University of Technology, Powstancow Warszawy 6, 35-959, Rzeszow, Poland, https://orcid.org/0000-0002-3601-897X, grzeprok@prz.edu.pl
[3] Vitaliy Marchuk, Institute of Building and Architecture, The National University of Water and Environmental Engineering, 11 Soborna Str, 33-028, Rivne, Ukraine, https://orcid.org/0000-0003-0999-0402, v.v.marchuk@nuwm.edu.ua
[4] Andriy Huts, Corresponding author, The Faculty of Civil and Environmental Engineering and Architecture, Rzeszow University of Technology, Powstancow Warszawy 6, 35-959, Rzeszow, Poland, https://orcid.org/0000-0002-4428-8339, a.huts@prz.edu.pl

1. Introduction

The modern scientific and technological revolution requires a radical increase in the efficiency of production and management processes through the maximum possible consideration of all reserves, introduction of more advanced technology, automation of production and management processes, widespread use of mathematical methods, computer and cybernetic machines.

In these conditions, as well as with a continuous increase in the scale of production and complication of interrelationships between individual factors, processes and phenomena, requirements for the methodology of scientific analysis and problem solving are increasing. In recent years, such a methodology, system analysis is increasingly being used, including a comprehensive feasibility study, comparison of alternatives and the selection of optimal solutions based on wide application of mathematical modeling and computers [1].

The main prerequisite of system analysis is the representation of the aggregate of connections between the phenomena and processes in the form of controlled cybernetic systems [2].

Depending on research and management tasks, the system can be limited to certain scales and divided into subsystems. For example, production of concrete and reinforced concrete structures can be represented as a complex system, in which it is expedient to single out the following subsystems: designing concrete composition, making a concrete mix and reinforcement products, forming and hardening. Each system is in its turn a subsystem in relation to the common system, etc. [3].

The essence of the system approach to the solution of the problem is that each subsystem should function optimally within the framework of the common system and in accordance with the set goal. In this case, the problem of controlling the system consists in:

- goal formulation;
- identifying the features and parameters of the system on which the achievement of the goal depends;
- determination of indicators and performance criteria;
- the construction of a mathematical model and development on the basis of algorithms and programs for determining the optimal values of factors [4].

To solve technological problems, it is possible to use mathematical models, which are the results of physical and mechanical properties research.

2. Research significance

The modern methods at researching of the properties of building materials include mathematical planning of the experiment. This method allows to obtain mathematical models of properties and technological parameters of their manufacture. Using the mathematical planning of experiment, it is possible to analyze the influence of controlled technological factors, their interaction, carry out interpolation calculations, optimize and predict their compositions and properties. For building a mathematical model, it is necessary using the results of experiments that are carried out according to specially developed plans, which allow to achieve a given level of significance with a minimum number of experiments.

3. Mathematical planning of experiment

Mathematical planning of the experiment (MPE) is understood as the formulation of experiments according to a pre-compiled scheme, characterized by optimal properties in terms of experimental work and statistical requirements. The theory of experimental planning is based on probability-statistical methods that make it possible to establish the minimum required number and composition of experiments theoretically justified, as well as the order of their carrying out to obtain quantitative dependencies between the parameter being studied and the factors that influence it [5].

The task of mathematical modeling is reduced to obtaining a certain idea about the surface of the response of factors, which can be analytically represented as a function:

$$M_{\{y\}} = \varphi(X_1, X_2, X_3, ..., X_n) \tag{1}$$

where: y – parameter of optimization, that is the output parameter of the system,
X_i – variable factors of the same system.

The most convenient is the representation of the unknown response function by the polynomial:

$$Y = \beta_0 + \sum_{i=1}^{n} \beta_i X_i + \sum_{i=1}^{n} \beta_{ii} X_i^2 + \sum_{i \neq j} \beta_{ij} X_i X_j + ... \tag{2}$$

The form and exponent of the polynomial degree are chosen either on the basis of a theoretical analysis or are refined statistically. Estimates of the regression coefficients β of polynomial models can be found on the basis of the experiment. According to the modern mathematical theory of experiment, there are statistical and cybernetic approaches and their combinations for the study of complex systems in the methods of mathematical planning of an experiment, which are currently quite well developed for various branches of science and technology, including building materials technology [6-8]. An active multifactorial experiment has several undeniable advantages to single-factor ones. It allows, due to the optimal organization of the study, 2-10 times or more, depending on the number of factors, to reduce the amount of work and to obtain more reliable dependencies considering the interaction of factors. When planning an experiment, mathematical methods play an active role at all stages of the study (setting and conducting experiments, processing results, making decisions) [5].

At the stage of obtaining mathematical models of complex systems it is possible to move from complex causal relationships to simple ones. This is the value of mathematical models that are not only necessary for the operational, automatic control of the system, but also a means of its deep knowledge and the creation of a theory.

The most important purpose of mathematical modeling is to optimize the characteristics of the process. There are two groups of optimization tasks:

a) selection of the most favorable technological mode (composition) - technology optimization;
b) maintaining a predetermined (in particular low-cost) mode.

This tasks can be solved separately and jointly [9].

In general, the problem of finding the optimal state of the system can be formulated as follows, determine positive values of technological factors $x_1 \geq 0$, $x_2 \geq 0 \ldots x_n \geq 0$, that satisfy the system of inequalities (constraints):

$$\left.\begin{array}{l} P_1 = f_1(x_1, x_2, x_3, ..., x_n) \leq \alpha_1 (or > \alpha_1) \\ P_2 = f_2(x_1, x_2, x_3, ..., x_n) \leq \alpha_2 (or > \alpha_2) \\ P_3 = f_3(x_1, x_2, x_3, ..., x_n) \leq \alpha_3 (or > \alpha_3) \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ P_n = f_n(x_1, x_2, x_3, ..., x_n) \leq \alpha_n (or > \alpha_n) \end{array}\right\} \quad (3)$$

which give the minimal (maximal) value of the function:

$$P_\kappa = f_k(x_1, x_2, x_3 ... x_n) \quad (4)$$

herein $P_1, P_2, \ldots, P_n$ are performance indicators, P_κ is efficiency criterion or objective function.

In solving technological problems, mathematical models that differ in the following aspects may be used:

1) the number of equations included in the system of inequalities (constraints);
2) the nature of the functions;
3) a number of factors;
4) the nature of relationships between factors;
5) the degree of intensity and the nature of changes in factors over time [10].

For each of the groups of models, it is necessary to apply specially developed mathematical methods for finding optimal solutions. Finding an extremum is reduced to solving a system of algebraic equations of the form:

$$j = 1,2,3...n, \quad \frac{\partial F(x_1, x_2, x_3...x_n)}{\partial x_j} \tag{5}$$

As the number of variables and the limitations on them is increased, the solution of such a system is associated with growing computational difficulties. In this regard, along with analytical, of practical interest are numerous optimization methods implemented using modern computer technology.

The construction of a mathematical model of the object of the study is the main task of the MPE.

The task of obtaining a mathematical model is to obtain a relationship that characterizes the relationship between the optimization parameter η and independent variables.

In its most general form:

$$\eta = \varphi\ (x_1, x_2,\ldots,x_k) \tag{6}$$

where: $x_1, x_2,\ldots,x_k$ – variables (factors) that can be varied during the experiments.

In the case of using the MPE, the parameter η (response functions) is approximated by a polynomial:

$$\eta = \beta_0 + \sum^{k} \beta_i x_i + \sum_{i<j}^{k} \beta_{ij} x_i x_j + \sum^{k} \beta_{ii} x_i^2 + ... \tag{7}$$

where: $\beta, \beta_i, \beta_{ij}, \beta_{ii}$ – theoretical regression coefficients.

As a result of the experiments, the regression coefficients b_0, b_i, b_{ij}, b_{ii}, which are estimates of theoretical coefficients, are determined. After this Eq. 2 takes the form:

$$\hat{y} = b_0 + \sum^{k} b_i x_i + \sum_{i<j}^{k} b_{ij} x_i x_j + \sum^{k} b_{ii} x_i^2 + ... \tag{8}$$

where: $\hat{y}$ – the calculated value of the optimization parameter.

The magnitude of the regression coefficients can be judged on the effects – the degree of influence of the relevant factors [11]. The importance of the regression coefficients indicates the significance of the respective effects.

For given $\hat{y}$ Eq. 8 can be interpreted as an equation of a certain surface in k-dimensional space (Fig. 1).

Successful application (MPE) depends, first of all, on the correct formulation of the problem. In this case, the experimenter should be able to determine clearly the amount and content of information that must be extracted from experiments, as well as the feasibility and possibility of using the MPE for specific conditions.

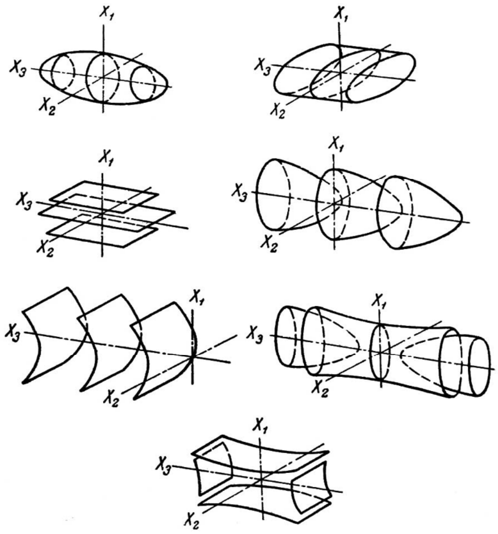

Fig. 1. Some three-dimensional contour surfaces characterizing an almost stationary region, which is described by a second-order equation, with k = 3

In setting up the simplest of tasks, or the first stage of the study, the regression equations of the first degree or incomplete quadratic equations are often designed. The solution of most optimization problems is connected, of course, with the use of second-order polynomials. Third-order polynomial dependences are practically not used in the practice of solving problems of construction materials technology.

The experiment is planned in several stages: first, preliminary research of the object of the study, then the construction of an appropriate mathematical model and its interpretation. At the end, if necessary, technical realization of the results obtained is carried out.

Preliminary research of the object of the study includes: problem statement; collection and processing of information, nomination of a working hypothesis; selection of optimization parameters, independent variables and constraints; preliminary experiment. For optimization tasks, optimization criteria should be clearly defined.

In the study of linear and incomplete quadratic dependences, the full factorial experiment (FFE) and shot replicas are most often used. In using the FFE, planning of experiments is carried out at two levels - the upper (+1) and lower (-1). The plans of experiments that are used, make it possible to realize all the unique variants of experiments, at the indicated levels, for a different number of factors. In this case, the number of experiments N depends on a set of factors k and is 2^k: for example, for two factors - $2^2 = 4$, for three - $2^3 = 8$, for four - $2^4 = 16$, for five - $2^5 = 32$, etc.

4. Choice of factors

An important requirement of the factors in the planning of their experiments is noncorelation. This does not mean that there should be no connection between the factors. It is necessary and sufficient for the connection not to be direct. For example, in the investigation of the concrete mix without introducing plasticizers and changing the quality of raw materials, slump of cone and water consumption cannot simultaneously happen. That is caused by the fact that for each value of the cone slump, water consumption is a constant figure [10].

When choosing factors, the degree of their controllability and the possibility of providing a given level of variation should be taken into account. Planning factors difficult to manage should be carried out using special techniques. It is better to quantify the factors studied. However, factors can be planned for their quality indicators.

For any value of factors, the levels of variation should provide the possibility of conducting an experiment and measuring the output parameter. For example, if the initial parameter is the mobility of a concrete mix, the proportion of sand in the aggregate mix, water content, and the amount of additive should ensure that a concrete mix will have a slump of the cone more than zero. If this is not enough, then it is necessary to narrow the intervals of variation.

The set of all values that can take the factor in the experiment, is called domain of variation. In the planning matrix (tables of planned experiments) factors are given in a coded form. At the same time, the central, so-called zero point is taken as the main level of variation, and denote it X_{io}, the interval of variation is ΔX_i.

The relationship between natural X_i and coded values of factors x_i is determined by the formula:

$$x_i = \frac{X_i - X_{io}}{\Delta X_i} \tag{9}$$

The variation intervals for linear and incomplete quadratic dependences are usually smaller than for full quadratic ones. The conditions for planning an experiment should be put in a special table, which can be constructed as shown at Table 1.

Table 1. FFE matrix of plan 2^2

Points of the plan, u	**Factors**		**Interaction**	**Output parameter**
	x_1	x_2	x_1x_2	y
1	+1	+1	+1	y_1
2	+1	- 1	- 1	y_2
3	- 1	+1	- 1	y_3
4	- 1	- 1	+1	y_4

Two-level plans are used to build linear dependencies, and three-level plans and plans with a large number of levels are used for quadratic dependencies. To avoid systematic errors and to evenly distribute or eliminate undesirable effects on the whole experiment (fluctuations in humidity and air temperature, slight changes in the grain composition of the aggregate, etc.), experiments are not carried out in the order indicated in the matrix, but in a random sequence. The sequence of experiments can be set on the tables of random numbers or depending on specific conditions.

The scheme for obtaining FFE matrices with the number of factors *k* from 2 to 5 is given in Table 2 [10].

The results of the experiments are processed using the methods of mathematical statistics, obtaining dependencies between the initial parameters and the factors that influence them, in the form of linear or incomplete quadratic regression equations.

Table 2. Example completion FFE matrices for k = 2-5

Points of the plan, u	**Factors**				
	x_1	x_2	x_3	x_4	x_5
1	+1	+1	+1	+1	+1
2	+1	- 1	+1	+1	+1
3	- 1	+1	+1	+1	+1
4	- 1	- 1	+1	+1	+1
5	+1	+1	- 1	+1	+1
6	+1	- 1	- 1	+1	+1
7	- 1	+1	- 1	+1	+1
8	- 1	- 1	- 1	+1	+1
9	+1	+1	+1	- 1	+1
10	+1	- 1	+1	- 1	+1
11	- 1	+1	+1	- 1	+1
12	- 1	- 1	+1	- 1	+1
13	+1	+1	- 1	- 1	+1
14	+1	- 1	- 1	- 1	+1
15	- 1	+1	- 1	- 1	+1
16	- 1	- 1	- 1	- 1	+1
17	+1	+1	+1	+1	- 1
18	+1	- 1	+1	+1	- 1
19	- 1	+1	+1	+1	- 1
20	- 1	- 1	+1	+1	- 1
21	+1	+1	- 1	+1	- 1
22	+1	- 1	- 1	+1	- 1
23	- 1	+1	- 1	+1	- 1
24	- 1	- 1	- 1	+1	- 1
25	+1	+1	+1	- 1	- 1
26	+1	- 1	+1	- 1	- 1
27	- 1	+1	+1	- 1	- 1
28	- 1	- 1	+1	- 1	- 1
29	+1	+1	- 1	- 1	- 1
30	+1	- 1	- 1	- 1	- 1
31	- 1	+1	- 1	- 1	- 1
32	- 1	- 1	- 1	- 1	- 1

In general terms, for the k factors:

$$\hat{y}_i = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i \neq j} b_{ij} x_i x_j \tag{10}$$

For example,

- for a two-factor experiment:

$$\hat{y}_i = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 \tag{11}$$

- for a five-factor experiment:

$$\begin{aligned}\hat{y}_i = {} & b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{14}x_1x_4 + \\ & + b_{15}x_1x_5 + b_{23}x_2x_3 + b_{24}x_2x_4 + b_{25}x_2x_5 + b_{34}x_3x_4 + b_{35}x_3x_5 + b_{45}x_4x_5\end{aligned} \tag{12}$$

During the FFE, with increasing amounts of factors, the number of experiments increases sharply. In some cases, for example, on the first stage of the study, that is, when you can slightly lower the preliminary assessment of the degree of influence of factors within the final data, the number of experiments will be greatly reduced. For this, fractional replicas are used (1/2, 1/4, 1/8, etc.), which are obtained by dividing the number of FFE experiments, respectively by 2, 4, 8.

Fractional replica matrices when performing the FFE are obtained by replacing interactions of a higher order (starting with triple - x_1, x_2, x_3, etc.) with new variables. These interactions are usually insignificant. The number of experiments in shot replicas corresponds to 2^{k-p}, where p is the fraction of the replica. For example, FFE of seven factors includes $2^7 = 128$, and 1/2 replicates $2^{7-1} = 64$, 1/4 replicates $2^{7-2} = 32$, 1/8 replicates $2^{7-3} = 16$ experiments, etc. Suppose in it's necessary to study the influence of five factors: x_1, x_2, x_3, x_4, x_5. To make a semi-replica 2^{5-1}, we can take FFE 2^4, and the interaction of factors x_1, x_2, x_3, x_4 can be replaced by a factor x_5. In this case, the number of experiments is reduced by half as compared with FFE 2^5.

5. Calculation of the coefficients of the equations

The free term of the equation b_0 is determined by the formulas:

$$b_0 = \frac{\sum_{1}^{N} y_u}{N} \tag{13}$$

$$b_0 = \frac{\sum_{1}^{N} \overline{y}_u}{N} \tag{14}$$

where: N – number of points of the plan, y_u – experimental value of the output parameter at points u_1 ... u_n of plan, $\bar{y}_u$ – average output parameter value at the point (u), for the case if $y_u = \frac{\sum_1^r y_{ui}}{r}$, i.e. by repeating experiments (r – the number of duplicate experiments on the rows of the matrix).

The coefficients for the linear terms of the equations are determined by the formulas:

$$b_i = \frac{\sum_{i=1}^{N} x_{iu} y_u}{N} \quad (15)$$

$$b_i = \frac{\sum_{i=1}^{N} x_{iu} \bar{y}_u}{N} \quad (16)$$

where: x_{iu} – the value of the i-th factor in the matrix row in the u-th experiment.

The coefficients of paired interactions are determined by the formula:

$$b_{ij} = \frac{\sum_{u=1}^{N} x_{iu} x_{ju} y_u}{N} \quad (17)$$

$$b_{ij} = \frac{\sum_{u=1}^{N} x_{iu} x_{ij} \bar{y}_u}{N} \quad (18)$$

where: x_{ij} – the value of the j-th factor in the u-th experiment.

Construction of the model can be considered complete, and the model itself used to make technological decisions only after an algebraic calculation coefficient estimates will be complemented by statistical (regression) analysis.

At the first stage of the regression analysis, the mean-square errors $S\{b_i\}$ of the estimates of the coefficients of the models are determined. The coefficients are considered significant if the calculated value of t – the Student's criterion turns out to be more than the tabular one, which is set depending on the given level of significance and the number of degrees of freedom. It is determined that the critical value of the coefficient estimate, below which the estimated estimates of b_i are advisable with the risk α considered minor, that is, equal to zero.

At the second stage, the hypothesis of adequacy (matching with experimental data) of a polynomial model with all significant regression coefficients is checked. It uses the minimizing sum of squares, which is called residual in the regression analysis. To verify the adequacy, a null hypothesis is formulated, and if it is considered plausible by the Fisher criterion (F), then the model describes the process adequately for the experiment. From an engineering point of view, this means that the model predicts results $\bar{y}$ on average with an error that is $\sqrt{F}$ times larger than the experimental one.

The tabular value of the F – criterion (FT) is determined depending on the adopted confidence level (significance level) and the number of degrees of freedom. In concrete technology, the confidence level is usually assumed to be 95%. If it turns out that this equation is inadequate, blunders were made in the experiments or the selected polynomial was insufficiently studied. In these cases, it's necessary either to repeat the experiments, or to change the varying intervals, or use a different plan.

6. Conclusions

The experiments to study, the influence of certain factors on the properties of a material, using mathematical planning methods, allows to obtain a result in the form of an equation (mathematical model), which with a certain statistically reasonable probability adequately describes the real process. The presence of the equation allows to solve the problem of providing properties, optimization of the composition and technology both within the framework of the experiment, and outside them. The use of mathematical models, makes it possible to conduct an effective technical and economic analysis of the adopted technological solutions, and the optimization of economic indicators.

The basic principles of using system analysis for solving problems of building materials technology are given. The basic concepts of the systems approach, mathematical modeling of the properties of building materials are explained. The main types of tasks solved using this method are analyzed, optimization criteria are formulated for solving concrete technology problems.

References

[1] Dvorkin L., Dvorkin O.: Basics of Concrete Science, Stroybeton, St.-Petersburg 2006.

[2] Voznesenskiy V., Lyashenko T., Ogarkov B.: Numerical Methods for Solving Problems of Construction Technology Using Computers, Vyshcha Shkola, Kyiv 1989.

[3] Dvorkin O.: Design of Concrete Compositions. Basic Theory and Methodology. NYVGP, Rovno 2003.

[4] Dvorkin L.: Optimal Design of Concrete Compositions, Vyshcha Shkola, Lvov 1981.

[5] Dvorkin L., Dvorkin O., Ribakov Y.: Mathematical Experiments Planning in Concrete Technology, Nova Science Publishers, New York 2011.

[6] Debska B.J., Dobrowolski L., Debska B.: Experiment-design Methods in Innovative Polymer Material Planning, Journal of Applied Polymer science, V.135, Issue 46, December 2018.

[7] Ferrandiz-Mas V., Sarabia L., Ortiz M., Cheeseman C., García-Alcocel E.: Design of Bespoke Lightweight Cement Mortars Containing Waste Expanded Polystyrene by Experimental Statistical Methods, Materials & Design, Vol. 89, 2016, pp. 901–912.

[8] Mohammed Sonebi: Medium Strength Self-compacting Concrete Containing Fly Ash: Modelling Using Factorial Experimental Plans, Cement and Concrete Research, Vol. 34, Issue 7, July 2004, pp. 1199–1208.

[9] Fomina E., Kozhukhova N., Sverguzova S., Fomin A.: Application of Mathematical Model Methods for Optimization Tasks in Construction Materials Technology, Journal of Physic: Conference series 1015 (5), 2018.

[10] Dvorkin L., Dvorkin O., Ribakov Y.: Multi-Parametric Concrete Compositions Design, Nova Science Publishers, New York 2013.

[11] Dolgov N., Zubrets'ka N., Buketov A., Stukhlyak P.: Use of the Method of Mathematical Experiment Planning for Evaluating Adhesive Strength of Protective Coatings Modified by Energy Fields, Strength of Materials, Vol. 44(1), January 2012, pp. 81–86.

Przesłano do redakcji: 15.01.2020 r.