Special test results evaluation features as development of "Innovations management" program - NEFU testing area as case-study

A V Ermakov, A M Bessmertnyy

M.K. Ammosov North-Eastern Federal University, Technological Institute, 58, Belinskogo St., Yakutsk, 677000, Russia

Email: ermakov-it@yandex.ru; bes.am@mail.ru

Abstract. This article gives an overview on the problems of precision in the results evaluation of the tests carried out in the Northern testing areas. One of the significant features of the facilities under study is that they are limited in quantity. In cases when a facility is taken down, that quantity is normally equal to one. The complexity of modern technological equipment and other circumstances require researchers to take into account and evaluate the potential risks. In order to make the sought-for estimations more precise, ways of improving the test result evaluations algorithms are suggested. In particular, one of the productive methods is the Data Mining technology, which presupposes implementing an intellectual analysis of the data with the aim of extracting useful information from the available database which was attained during the tests and other types of activities. Applying the Data Mining technology is becoming more productive when the scenario analysis is carried out, i.e., the analysis of possible alternative solutions. Another perspective trend is an implementation of an interdisciplinary approach. As a result, researchers are able to carry out a complex evaluation of the test results, which will noticeably increase the value of the given results.

1. Introduction

The North-Eastern Federal University (NEFU) testing area was created for testing facilities, which would operate in the cold climatic conditions. A relatively low cost for testing various kinds of facilities in cold climatic conditions in the Sakha Republic (Yakutia) and the possibility to run complex tests using the NEFU intellectual potential allowed beginning constructing a testing area that would help solve at least two existing problems. First, holding such tests would provide NEFU with additional income. Second, taking part in such tests helps the NEFU students and staff members to enhance their professional skills.

What makes the test results evaluation process unique is that, as a rule, the number of tested facilities is limited. The amount of tests is normally equal to one in cases when the facility should be knocked down. In such case scenarios, classic statistic methods cannot be used to process test results data. This article suggests using a procedure of test results evaluation which takes into account the specific NEFU test area operating features.

2. Test results evaluation procedures

The test results evaluation procedure is regarded as one of the trends in the NEFU program devoted to *Innovations Management*. The main goal of such procedures is to use the test results to formulate

theoretical bases which a long-term perspective would allow minimizing the expenses on natural experiments by using mathematical modeling.

The procedure of organizing and carrying our test is the following. Three aspects are usually taken into account in research papers on the topic [1]: preparation, implementation and results analysis. The last of the three components – results analysis – represents the biggest interest. Nevertheless, the preparation and implementation stage remain important for each of the tests at NEFU testing areas. Their content is predominantly defined by the object and the goals of research, whilst a range of other general criteria can also be singled out.

The testing procedure at the NEFU testing area can be represented by joining co-related operations, which are illustrated in Figure 1. They are represented as modules, which are placed in separate rectangles. It is assumed that the main topic of research would always be the object. Processes can also serve as research topics, which do not change the plane of our further discussion in any way.
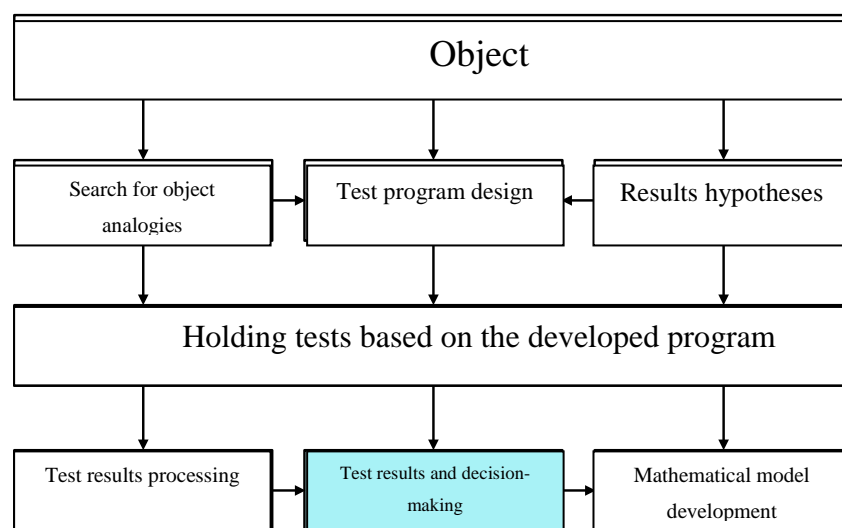

Figure 1. The tests procedure at the NEFU testing area.

The test preparation stage is represented by three operations (modules). The first module includes searching for object analogies. As a rule, the results of such search allow one to increase the effectiveness of the testing process itself. The second module is aimed at developing a testing program with regard to the specific features of the object under study. During the program development, the results of the third module are taken into consideration. The goal of the third module is to form scientifically-based hypotheses on the expected results.

Tests are held further based on the developed program, after which the other three modules are implemented. The first of these modules includes processing the data results. This module is described in detail in the "Intellectual analysis of the test results" subsection of the article.

The second module – test results and decision-making – should be considered as the main one within the framework of the procedure under the description. For this reason, the background of this module is colored blue (see Figure 1). Furthermore, the results of the "Test results and decision-making" operation serve as a source of information for developing a mathematical model of the object under case-study. If the carried out tests do not have any unique value (for instance, they had been held earlier and have not undergone significant changes), the mathematical model should not be redeveloped but simply clarified.

It should be emphasized that the decision-making rules should be represented in the "Test program design" module. Given these conditions, the test results can presumably appear to be obscure. Should this happen, the "Test and program design" module would require some clarifications. In all such cases, the basic theoretical guidelines of decision-making should be taken into close consideration [5].

A number of useful guidelines on how to carry out specific modules included in the test procedure can be found in many other scholar papers, for instance [3, 10].

The test results evaluation procedure should include the evaluation of the number of the possible tests (N) and two other values: ΔC - a portion of the agreement cost, which corresponds to the possible expenses on the tested object, and, finally, the estimated expenses per one test. Given this, N value can be estimated by using the following inequation:

$$N \le 1 + \left\lfloor \frac{\Delta C}{W} \right\rfloor. \qquad (1)$$

The sign $\left\lfloor \dfrac{\Delta C}{W} \right\rfloor$ serves as an indication that an entire part of the final evaluation is taken into account during the estimation. While the N value is estimated in several tens, the data results processing (in addition to the values acquired in accordance with the test program) can be supplemented by rather reliable statistic estimations [2, 11]. Unfortunately, the N value is often estimated in units. Moreover, the N value is equal to 1 for some objects. In such cases, statistic data are of no value in test results analysis.

Whatever may be the case, it would be useful to follow the scheme (Figure 2) in order to clarify the test results. The scheme is based on carrying out tests on two objects. One of them being a physical object, which should be regarded as a "black box" [12]. Such an approach can be justified by the fact that the behavior of some elements of the object cannot be predicted (at least until the very end of the tests). The second test object is a mathematical (imitative) model. It can be regarded as a "white box" [12], since all the model components have been determined.
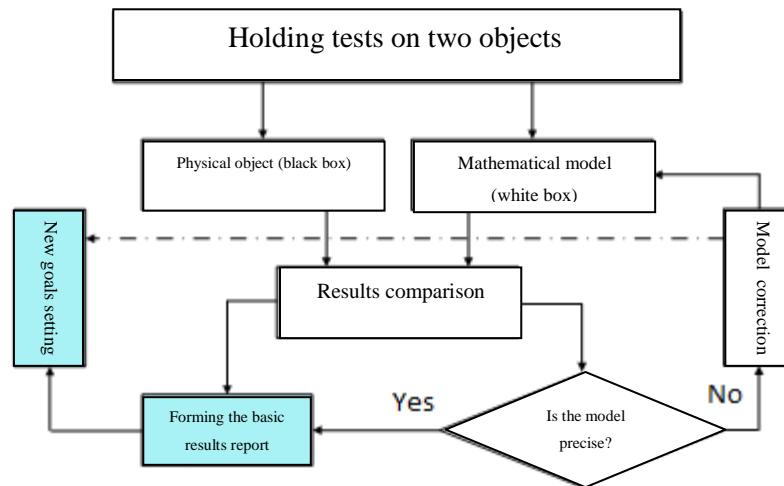

Figure 2. Test results clarification scheme.

Upon finishing the testing of two objects, a comparative analysis of the given results is carried out. The provided information enters the "Forming the basic results report" module. Moreover, it is defined whether the mathematical model was precise or not. If the model was precise enough, this result is reflected in the main report. If the result is contrary, the model is to be refined. Conclusions in the main report and the information from the "Model correction" module allow one, if there is a need, to set new theoretical and practical goals. For precisely this reason the "New goals setting" module background had been labeled the same color as the "Forming the basic results report" module.

Let us assume that the test results are represented by two such sequences [7] as: $\{x_1, x_2,..., x_n\}$ и $\{y_1, y_2,..., y_n\}$. Each of the n-variables of x_i serves as a characteristic of an important feature of the

physical object. Likewise, each of the n-variables of y_i serves as a characteristic feature of the mathematical model. It should be noted that the variable property can be represented by not only a technical parameter expressed by a measured or evaluated value represented in any other form, but also by an expression of a "not supported/not observed" type, or any other similar expressions.

The process of test results comparison can be regarded as two operations. The first operation lies in establishing relations between all the x_i properties and the corresponding norms represented as x_i^{H}. Now let us assume that the norms have been set without losing their common features with the help of the following sequence of inequations:

$$x_1 \leq x_1^{\text{H}}, \ x_2 \leq x_2^{\text{H}}, ..., \ x_n \leq x_n^{\text{H}}. \tag{2}$$

If none of the n-variables of the inequations are violated, the tests are considered to be successful, however, this should happen under a range of conditions which will be listed in the next subsection of the article.

The second operation involves a paired comparison of the x_i and y_i properties. The results of the second operation do not affect the test evaluation. They are used to clarify the mathematic model. Let us suppose that so far as the model retains its general features, the acceptable accuracy of the mathematical model is reached when the difference between the x_i and y_i values is no more than that of the z_i value:

$$\left|x_1 - y_1\right| \leq z_1, \ \left|x_2 - y_2\right| \leq z_2, ..., \ \left|x_n - y_n\right| \leq z_n. \tag{3}$$

Complying with all the n-inequations allows confirming that fact that the mathematic model gives an opportunity to give an adequate assessment of the physical object testing results. This assertion should be considered absolute. It can be considered correct only after an imitational modeling had been held without changing the properties of the physical object and its testing conditions.

In order to carry out an intellectual analysis of the testing results, the Data Mining principles are used [15]. While doing so, it is reasonable to use a scenario approach [13], which assumes finding answers to such questions as "what happens if...". It is worth mentioning that the "Interdisciplinary research" trend [8] played a significant role in the process of developing a NEFU program on "Innovations management". Such approach allows one to give a complex evaluation of the testing results, and in some case to even forecast further changes in the properties of the tested object.

3. Final-decision making

Final decision-making as a result of testing can be minimized down to the interacting pairs of IJD and IMD [5, 9] - individuals justifying the decisions and individuals making the decisions respectively. These are the terms used in the theory of decision-making. Most commonly IJD and IMD are represented by teams of high-qualified experts that have experience working with testing objects. One of the existing peculiarities in assessing results is that it is impossible to hold a large number of tests. For this reason, a huge role is given to the quality researches instead of the statistical results evaluation method [2, 11].

The final decision-making process as a result of the tests can be represented as an algorithm which is illustrated on Figure 3. This algorithm includes five components. If necessary, a certain range of components can be described in more detail by adding additional functional blocks.
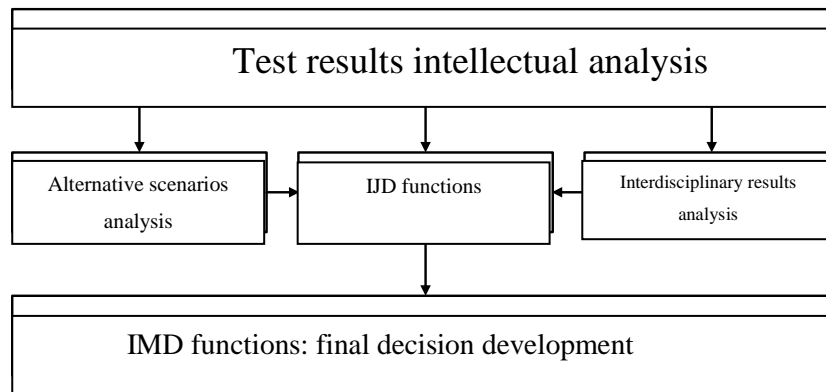
Figure 3. Final decision-making process.

The "Alternative scenarios analysis" block includes a range of operations, the main goal of which is to form judgments that are closely linked to rational (sustainable) decisions [14]. The essence of such judgments lies in stating a fact which is characterized by a maximum level of probability and a minimal chance of an error in the expected changes for the existing functioning conditions of the object under study. For instance, there were three solutions found for using the tested object represented as the values of its efficiency coefficient - η_k and the temperature range - $\left(\alpha_k, \beta_k\right)$ for which the value of η_k remains unchanged:

$$
\begin{aligned}
&1. \quad \eta_1 = 75\%, \quad \alpha_1 = -30C^0, \quad \beta_1 = -50C^0; \\
&2. \quad \eta_1 = 70\%, \quad \alpha_1 = +5C^0, \quad \beta_1 = -70C^0; \\
&3. \quad \eta_1 = 60\%, \quad \alpha_1 = +40C^0, \quad \beta_1 = -80C^0.
\end{aligned} \qquad (4)
$$

Let us now assume that the efficiency coefficient restriction is given as follows: $\eta_k \geq 50\%$, i.e., all three solutions are possible. The alternatives are marked by a temperature range. If the technical requirements for the object indicate an average range of $\left(-30C^0, -50C^0\right)$, then it would be logical to choose the first solution, since it provides a maximum value of the efficiency coefficient. On the other hand, in order to minimize the risks, it is necessary to analyze the other two possible solutions. For example, if the range of $\left(+40C^0, -80C^0\right)$ is seen by the IJD and IMD as too wide, then the third solution should not be regarded at all. In cases when the range of temperature fluctuations of $\left(+5C^0, -70C^0\right)$ is considered nearly impossible, preference should be given to the second possible solution, although $\eta_1 < \eta_2$.

The "Interdisciplinary results analysis" block allows considering factors which are difficult to formulate in the process of developing a testing program. Let us return to the example with different ranges in the temperature fluctuation. Let us now assume that the team of experts selected for carrying out the interdisciplinary results analysis includes a historian. This expert can express personal opinion on the risks caused by aiming the focus on the range of temperature fluctuations $\left(-30C^0, -50C^0\right)$, which is based on one of the crucial errors made by Napoleon during his preparation for the military campaign in 1812. Perfectly knowledgeable in mathematics, he analyzed the temperature fluctuations in the line of the French army's defense within a period of the past twenty years. The statistic results showed that drastic frost weather should not be expected. However, real circumstances did not prove these results to be true. Therefore, IJD (a historian in this case) will receive a more detailed information in order to fulfill the corresponding functions.

The interdisciplinary results analysis may be carried out in various ways. One of the most convenient ones is the Delphi method [16] developed by an American consulting group "RAND

Corporation" over fifty years ago. A collective discussion of the problems is often distorted due to the fact that some experts participating in the discussion subconsciously affect the minds of the listeners in the audience. The Delphi method is based on rejecting the idea of a collective discussion. This results in excluding the possibility of psychological factors influence, i.e., agreeing with the point of view of colleagues with a higher level of authority; or the desire to agree with the major opinion. In fact, rejecting discussions increases the time for getting results. The experts answers are processed, the questions may be refined with new information and sent to the experts once more. All experts opinions in this case are considered equally valuable. Some modifications to the Delphi method presuppose labeling experts with coefficients of professional value that reflect their level of expertise. Several iterations may be held in order to get reliable evaluations.

The "Functions of IJD" block can be considered as somewhat of a sum total of the results attained as a result of performing two blocks (operations) which were described above. The main task of IJD is to systemize the available results and to make a report which should reflect all the advantages and disadvantages of the recommended decision or a list of alternative solutions with a rating for each in accordance with the grading scale applicable to this object. When preparing a IJD report, it is also necessary to consider the factors which were not taken into account by the experts involved in process of solving tasks represented in the "Alternative scenarios analysis" and "Interdisciplinary results analysis" blocks.

The "IMD functions: final decision development" block represents the final stage in the suggested algorithm. Operations carried out by IMD are normally associated with the so called non-structural (non-characterizable) goals [6]. The efficiency of IMD is to a large extent defined by the quality of the report presented by the IJD. Undoubtedly, the qualification of the experts in the IMD group is also an important component in the successful solution to a given problem.

The NEFU testing area holds tests which are, as a rule, connected with the analysis of functioning objects and the low temperature processes. From this perspective, it is very difficult for IMD to select analogies which would allow singling out some general tendencies for the object under study. Perhaps it is easier to some extent to solve an opposing problem: to forecast the values of the object or the characteristics of the process in regular conditions based on the testing results attained in extreme conditions. Nevertheless, finding the analogies method which was described above presents a great practical interest. The possibility to solve such problems can be identified by analyzing the work experience of the NEFU testing area as it is one of the elements of the "Innovative management" system.

4. Conclusion

Thus, based on the information provided in the present article, the test results evaluation held in the NEFU testing areas is a very difficult procedure to carry out. The recommendations given above are aimed at increasing the precision of the evaluations sought-for. However, the complexity of modern technology and other circumstances require taking into consideration the potential risks and their evaluation [4]. For this reason, the algorithms used to process the test results are continuously being updated.

At the initial work stage of the NEFU testing area, the main focus was made on testing physical objects (machine building products and materials). The development of the imitation models complex requires a considerable amount of time and effort on the part of a large team of programmers. Nevertheless, it is the establishment of such a complex that will allow one to effectively solve the problems described in the "Innovations management" program, which was developed by and is used at NEFU.

References

[1]   Blokhin V G, Gludkin O P, Gurov A I and Khanin M A 1997 *Modern Experiment: Preparation, Implementation, Results Analysis* (Moscow: Radio and Svyaz Publishing) p 232
[2]   Ventsel E S 2005 *The Theory of Probability* (Moscow: Akademia Publishing Center) p 616

[3]   Voronin Y A, Chudinsky R M 2003 Correlation of Natural and Model Experiments in Physical Applications *The Teach. Phys. in Univ. Jour* **9(2)**

[4]   Korolyov V Y, Bening V E, Shorgin S Ya 2011 *Mathematical Bases of the Risk Theory* (Moscow: Phismat Publishing) p 591

[5]   Larichev O I 2002 *The Theory and Methods of Decision-making* (Moscow: Logos Publishing) p 382

[6]   Larichev O I, Moshkovich E M 1996 *Qualitative Methods of Decision-making, A Verbal Analysis of Decisions* (Moscow: Phismat Publishing) p 208

[7] 1977-1985 *An Encyclopedia of Mathematics* (Moscow: Sovetskaya entsiklopedia Publishing) p 623

[8]   2002 *New Trends in Synergetics: An Outlook on the Third Millenium/Information Science: Limitless Possibilities and Possible Restrictions* (Moscow: Nauka Publishing) p 480

[9]   Saati T 1993 *Decision-making, A Method of Hierarchy Analysis* (Moscow: Radio and Svyaz Publishing) p 278

[10]   Sidnyaev N I, Vilisova N T 2011 *Introduction to the Theory of Experiment Planning* (Moscow: MSTU Press) p 463

[11]   Shor Ya B 1962 *Statistical Methods of Quality and Reliability Analysis and Control* (Moscow: Radio and Svyaz Publishing) p 553

[12]   Eshbi Yu R 1959 *Introduction to Cybernetics* (Moscow: Foreign Literature Press) p 432

[13]   Yuditskiy S A 2001 *A Scenario Approach to Modeling Business Systems Behavior* (Moscow: Sinteg) p 112

[14] Sokolov A., Sokolov N 2014 Rational solutions for development of telecommunications networks *T-Commun. and Transp.* **8 (6)**

[15]   Zaki M, Meira W 2014 *Data Mining and Analysis: Fundamental Concepts and Algorithms* (Cambridge University Press)

[16]   Linstone H A 2002 *The Delphi Method. Techniques and Application* (New Jersey Institute of Technology) p 618