The new method of COBRA parameters comparison

A A Pavlenko[1]**, V V Tynchenko**[1,2]**, V V Bukhtoyarov**[1,2]**, A A Boyko**[1,2]**, S V Tynchenko**[1,2] **and A A Stupina**[1,2,3]

[1] Reshetnev Siberian State University of Science and Technology, 31, Krasnoyarsky Rabochy Av., Krasnoyarsk, 660037, Russian Federation
[2] Siberian Federal University, 79, Svobodny pr., Krasnoyarsk, 660041, Russian Federation
[3] Krasnoyarsk State Agrarian University, 90, Mira pr., 660049, Krasnoyarsk, Russia

E-mail: ms_f1_pavlenko@mail.ru

Abstract. Comparison between two or more number rows is very difficult and impossible to perform without additional mathematical processing and formulas. We need a tool to determine whether the efficiency of the algorithm has significantly improved, the changes made, or at a certain level of significance, these changes have not made any special improvements to the operation of the optimization algorithm. Analyzed methods of variational series comparison. A new method of variational series comparison has been developed. The methodology was tested when choosing parameters and for comparing the influence of initialization methods on the global optimization genetic algorithm and the collective optimization method based on the Co-Operation of Biology Related Algorithms (COBRA) based on bionic algorithms. Studies have shown that the new method of variational series comparison well fulfills its functions and copes with its task.

1. Introduction

Comparing two numbers with each other is not difficult and cannot be processed by any special method [1]. You can compare variational series with each other by arithmetic averages and dispersion factors. In this case, the comparison of two variational series of numbers is reduced to a simple comparison of two numbers. Such a simple comparison can be carried out for an unlimited number of rows depending on the problem being solved. Such a simple comparison has very weak mathematical power and is not suitable for serious research.

When experimenting with optimization algorithms, we often try to make some changes to the operation of these algorithms: to change the value of the parameters of the algorithms, to replace some modules of the algorithm with others, etc. [2, 3] With such experimentation, we would always like to have a tool - a technique with which one could make the correct and most accurate comparison of changes made to algorithms. To solve this question, I created a method for comparing variational series, i.e. methodology for comparing the effect of changes on the efficiency of optimization algorithms. This methodology was applied to compare the effects of various changes of algorithms on the efficiency of the genetic algorithm of global optimization [4, 5] and the collective optimization method based on common bionic algorithms - Co-Operation of Biology Related Algorithms (COBRA) [6].

Recently, a weak comparison methodology has been used in the work of scientists. Therefore, a new method of variational series comparison was developed and applied.

2. Explanatory part

Figure 1 shows a scheme for identifying factors, factor levels and observations. The factor is the module of the algorithm, the factor level is the version of the algorithm, and the observation is the algorithm implementation.
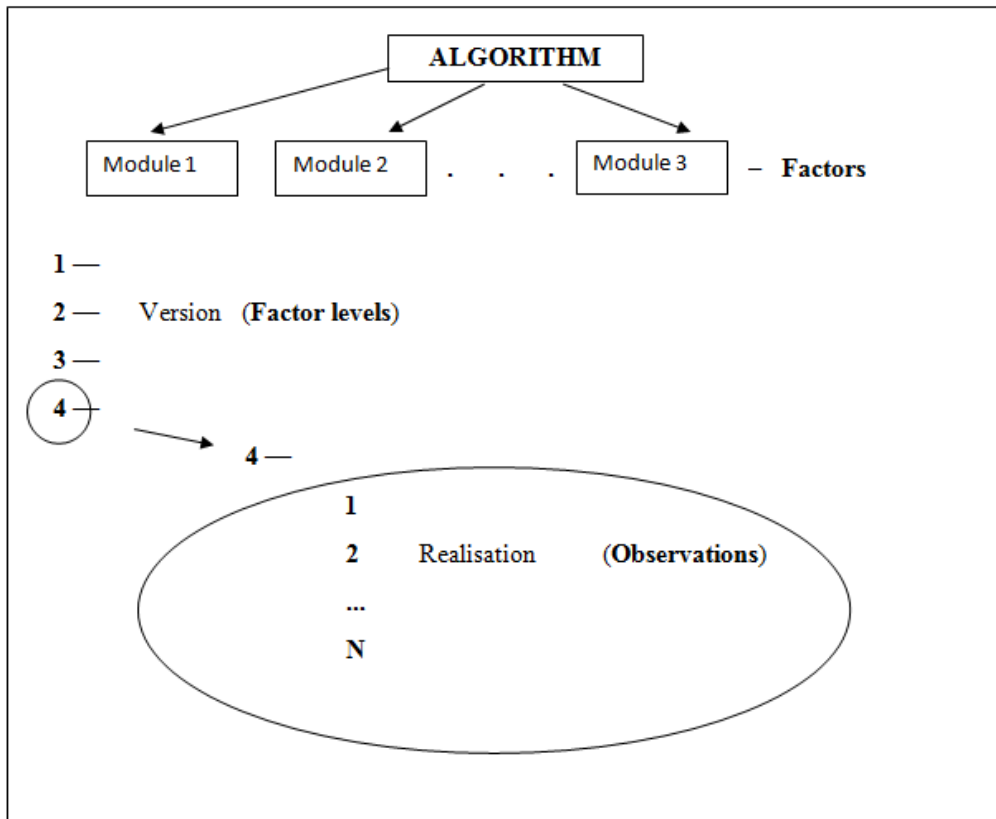

Figure 1. Scheme of factors, factor levels and observers.

The last studies in this area were conducted by A.V. Penenko in 2019 [7]. In his works, a comparison methodology was used that does not allow revealing hidden patterns of binary strings.

3. Experimental part
Figure 2 shows the block diagram of new variation series comparison method.

If you need to compare more than two variational series (samples), for example, when we compare several modifications of the same algorithm module (Figure 1), then we need:

- Get variation series samples of more than three elements.
- Check the samples for statistical stability using the Strangers coefficient [8].
- Check the samples for a normal distribution law [9] by criterion 2 at a 0.05significance level.
- Test the hypothesis about the equality of all samples variances according to the Bartlett criterion at a significance level of 0.05 [10].
- If the variances are equal and the normal law is confirmed, then:
- Test the hypothesis about the equality of the mathematical expectations of all samples for single-factor variance analysis at a significance level of 0.05, i.e. check whether all mathematical expectations are equal to each other (whether the change of module versions (factor levels) affects the efficiency of the algorithm).
- Determine the effect of the module itself on the efficiency of the algorithm by single-factor analysis of variance at a significance level of 0.05, using a sample coefficient of determination.

- If all mathematical expectations are not equal, then in order to identify which mathematical expectations are different from each other and which are not, it is necessary to carry out a multiple comparison of Scheffe.
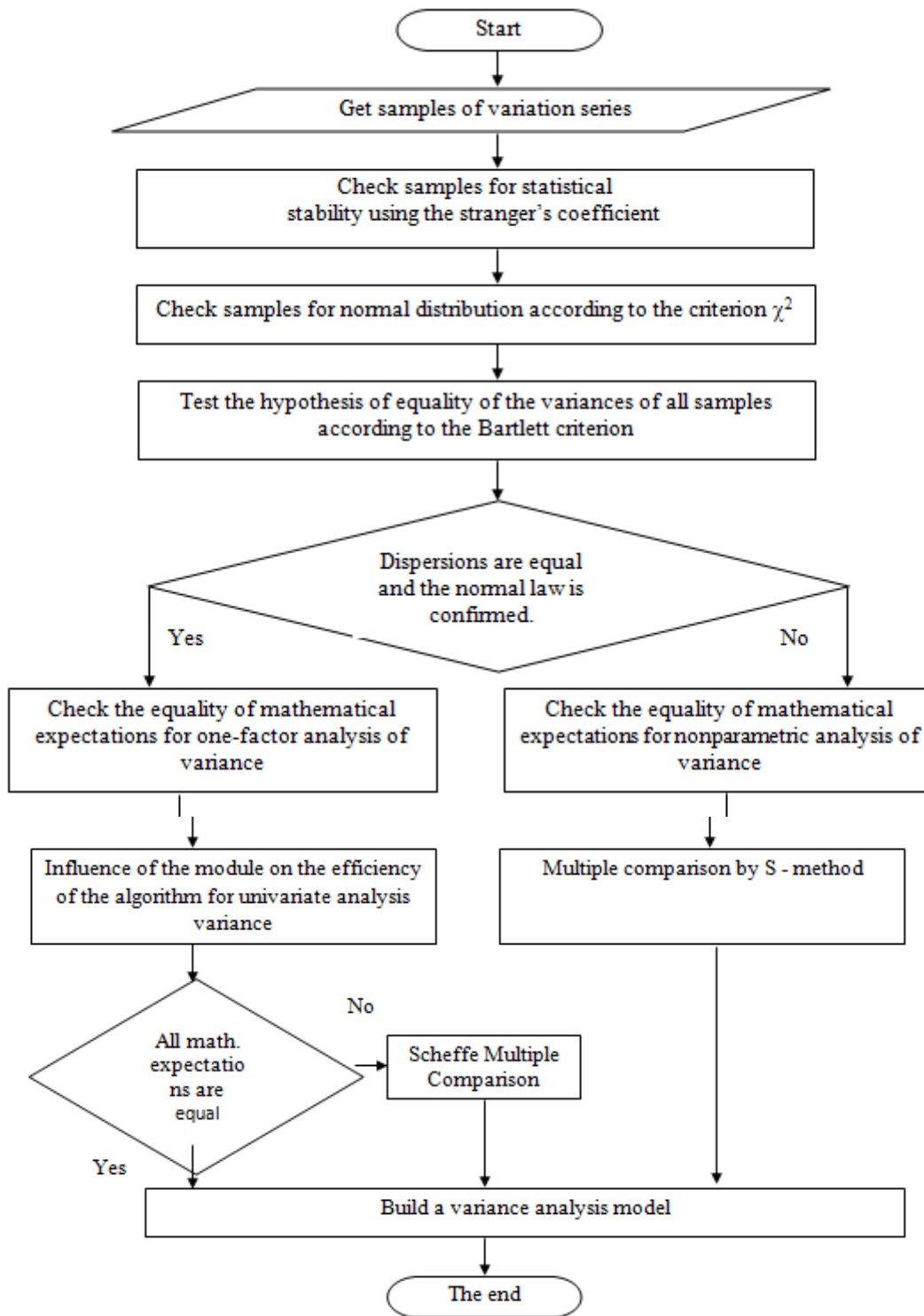

Figure 2. A method flowchart for comparing variation series.

Otherwise:

- Test the hypothesis of mathematical expectations equality for nonparametric variance analysis.
- To conduct multiple comparisons according to the S - method.
- Build a dispersion analysis model to determine the effect of each factor level on the efficiency of the algorithm.

If you want to compare only two variational series, then in sub-Clause 5.1, instead of single-factor analysis of variance, you can test the hypothesis of mathematical expectations equality, at a significance level of 0.05 by Student's statistics, having previously checked the equality of variances at a significance level of 0.05 by Fisher statistics. Sub-clauses 5.2, 5.3, 6.2 and 7 are not implemented in this case. However, in this case we will not be able to determine the effect of the module itself on the efficiency of the algorithm, therefore it is recommended when comparing two variational series also to use the methodology described above.

4. Results

This technique made it possible to compare the values of the genetic optimization algorithm parameters and the collective optimization method based on common bionic algorithms - the Co-Operation of Biology Related Algorithms (COBRA), at a significance level of 0.05. What options are significantly different from each other, and which are nob t.

This technique was applied to compare the effect of initialization methods on the efficiency of the collective optimization method based on the common bionic algorithms - Co-Operation of Biology Related Algorithms (COBRA) at a significance level of 0.05. Testing the hypothesis about the equality of mathematical expectations was carried out by univariate analysis of variance, which led to the subsequent multiple comparison of Scheff.

This technique was also applied to compare the effect of initialization methods on the efficiency of the global optimization genetic algorithm at a significance level of 0.05. Testing the hypothesis about the equality of mathematical expectations was carried out by univariate analysis of variance, which also led to the subsequent multiple comparison of Scheff.

5. Conclusion

Different methods of comparison of variational series were analyzed. A new method of variation series comparison has been developed. A new technique was applied when choosing the parameters of two algorithms and for comparing the influence of initialization methods on the genetic algorithm of global optimization and the collective optimization method based on the common bionic algorithms - Co-Operation of Biology Related Algorithms (COBRA). Studies have shown that the new method of variation series comparison copes well with its task.

References
[1] Zaloga A N, Yakimov I S and Dubinin P S 2018 Multipopulation genetic algorithm for determining crystal structures using powder diffraction data *Journal of surface investigation: X-ray, synchrotron and neutron techniques* **1(12)** 128-34
[2] Tynchenko V S, Tynchenko V V, Bukhtoyarov V V and Agafonov E D 2018 Decision support system for designing an effective configuration of a computing network for distributed complex problem solving *RPC 2018 - Proceedings of the 3rd Russian-Pacific Conference on Computer Technology and Applications* 8482126
[3] Tynchenko V S, Tynchenko V V, Bukhtoyarov V V, Tynchenko S V and Petrovskyi E A 2016 The multi-objective optimization of complex objects neural network models *Indian Journal of Science and Technology* **9(29)** 99467
[4] Zaloga A 2015 Genetic algorithm for automated X-Ray diffraction full-profile analysis of electrolyte composition on aluminium smelters *Informatics in control, automation and robotics 12th International Conference, ICINCO* 79-93

[5]   Tynchenko V S, Petrovsky E A and Tynchenko V V 2016 The parallel genetic algorithm for construction of technological objects neural network models *2nd International Conference on Industrial Engineering, Applications and Manufacturing, ICIEAM 2016* 7911573

[6]   Stanovov V, Akhmedova S and Semenkin E 2018 Selective pressure strategy in differential evolution: exploitation improvement in solving global optimization problems *Swarm and Evolutionary Computation* 183-9

[7]   Penenko A V 2019 Newton–Kantorovich method for solving inverse problems of source identification in product–destruction models with time series data //*Siberian journal of computational mathematics* **1(22)** 57-79

[8]   Karpenko A P 2018 Evolutionary operators of population-based global optimization algorithms. Experience of systematization //*Mathematics and mathematical modelling* **1**

[9]   Akhmedova S, Stanovov V and Semenkin E 2018 Soft island model for population-based optimization algorithms *International Conference on Swarm Intelligence* 68-77

[10]   Yakimov I 2018 Application of evolutionary rietveld method based XRD phase analysis and a self-configuring genetic algorithm to the inspection of electrolyte composition in aluminum electrolysis baths *Crystals* **11(8)** 402