Analysis of the efficiency of obtaining reliable data when operating with number sets in simulation of bridges structural components

Igor Maystrenko, Tagir Zinnurov, Dmitriy Erohin and Tatiana Maystrenko

1Kazan State University of Architecture and Engineering, Kazan 420043, Russia
E-mail: zinnurovta@kgasu.ru

Abstract. The issue of reducing the complexity of a multifactorial simulation experiment to an acceptable level is relevant. The simulation complexity is characterized by the computing capabilities of a PC while maintaining a given level of accuracy and certainty of the problem being solved. Many authors neglect the justification of the sample size, thereby reducing the experiment's reliability. Loads are stochastic, therefore, in the calculations, they are determined by rated and design values. However, the Monte Carlo methods allow the specific realization of loads based on the distribution law chosen. Techniques for optimizing the multifactorial simulation experiment procedures have been proposed on a test example of adding permanent loads due to the bridge pavement layers. They allow excluding the procedure of building multidimensional laws of the final distribution of random parameters, while sequentially transforming them from input to output values. The minimum required simulation experiment size, which should exceed $7.5 \cdot 10^6$ realizations and is determined by the product of the number of tests $N$ by the number of runs $n$, has been experimentally justified for the structural components of bridges.

Keywords: simulation, road bridge, reliability of results, Monte Carlo, distribution function, experiment size.

1 Introduction

Historically, the reliability of engineering systems is determined by the probability of failure and the reliability density function of two random variables, i.e., the bearing capacity $R$ and the load $Q$. Currently, the Monte Carlo simulation method is actively used to evaluate the uncertainties of the $R$ and $Q$ parameters. Despite the Monte Carlo method is widely used in reliability studies, it has a drawback: evaluating highly reliable systems requires a large sample size. This increases the computational efforts and simulation time. To solve this problem, Jebur and Al-Zaidee (2019) have developed an approximate model based on the Monte Carlo method [1]. The authors achieved a significant reduction in the computational effort by formatting the reliability problem into a specific parameter using the dependency of the probability of failure on this parameter.

Li and Cao (2016), R. Rebba, S. Mahadevan (2008) have briefly described the theoretical and numerical elements of set simulation, as well as detailed guidelines for realizing standard algorithms to solve the reliability analysis and structural optimization problems [2, 3]. The paper also considers simple argument checks and post-processing of simulation results. Vidia Sagar and Manjunath (2017) and Pichugin (2018) focused on the stochastic nature of the loads and material properties that allowed them to form the source data realization sets to solve reliability problems. Based on the approach developed, numerical calculations of the reliability have been performed for a wide range of building structures [4, 5]. Some authors, i.e. Shadab Far and Wang (2016) offer a simple algorithm to estimate low failure probability using a small number of samples in combination with the Monte Carlo method [6, 7, 8, 9].
Due to the lack of a unified technique to calculate building structures based on the probability of failure-free operation, the authors have assigned the number set size without any or with insufficient justification in their research. Therefore, the study objective is to scientifically justify the required sample size generated in simulation to obtain reliable results.

2 Materials and methods

The problem of eliminating errors in the course of summation of a large number of parameters when using various number set processing techniques can be solved based on either the corresponding multidimensional distributions or approximating expressions that compensate for the number set deformations. To solve similar problems in other research fields, e.g., in metrology, formulas have been proposed to calculate an error at a given certainty factor for various random variable distribution law combinations [10]. However, the search for an approximating expression has a pronounced specific nature, and the efficiency of eliminating errors in designing structural elements of bridges may prove to be insufficient.

If we presume the possibility of comparing the numerical and realistic experiments, then guided by the requirements for the control reliability tests, in the absence of data on the random variable distribution, the number of tests \( N \) to confirm the reliability level \( P \) at a certainty factor \( \beta \) should be determined by the in equation [11, 12, 13]:

\[
N \geq \frac{\ln(1 - \beta)}{\ln P}.
\]

Usually, the logic of assigning requirements to the minimum number set size is implicit and based on some enlargement of parametric models, e.g., in a typical load-strength system [14, 15]. Solving the problem is not sufficiently detailed, and the statistical characteristics of individual loads are transformed into the total one with new statistical performance with a certain error [1, 2, 16]. These techniques take effect of the first approximation of the problem being solved, however, the results obtained may vary within rather wide ranges. The method requires adequate compensation for this drawback based on additional limitations or hypotheses.

As such compensating techniques, either the hypothesis of a normally distributed random variable is most often used, or the reliability level decreases. Both techniques significantly limit the probabilistic-statistical model capabilities as applied to the highly reliable system calculation. The reliability level should not fall below 0.95 and the reliability of the estimates should preferably be kept at 0.99 or higher.

It should also be noted that the problems of justifying the general engineering facility design solutions are solved under the conditions of personal computers (PCs) operating based on conventional software only. This aspect limits the choice of the initial number set size and algorithms of processing. The Monte Carlo statistical simulation method imposes certain requirements: firstly, adequate mathematical models are needed for modeling; secondly, to obtain reliable analytical results, all system components should be considered; thirdly, a large number of the system, i.e., the simulation model realization copies is simulated.

Researchers treat the efficiency of using the Monte Carlo method to solve multi-parameter problems of calculating complex engineering systems with some skepticism. The classical Monte Carlo method of simulation applies the procedure of complete enumeration of all random realizations with the generation of their combinations [17, 18, 19, 20, 21]. E.g., if six individual loads on a structural component of a road bridge should be summarized while representing each parameter by an array of random realizations of \( 10^6 \) size, then the total load output parameter will be characterized by a number set having a size of \( (10^6)^6 = 10^{36} \) random realizations. To simulate structural components of bridges, a significantly larger number of parameters should be recorded and subsequently processed. Realistic calculation of structural components of complex engineering systems involves more than a hundred of such parameters, therefore, the designer will have to operate with number sets of about \((10^6)^{100} = 10^{600}\) random realizations, which does not meet the PC capabilities. Thus, the hardware and
software progress will not provide the designers with computing systems that allow processing the number sets of such size.

In addition to the above difficulties, the solution of probabilistic problems using the Monte Carlo method requires high precision. This issue arises given the well-known correlation between accuracy and experiment size \((n \cdot N)\) noted, e.g., in the studies of Perelmutter, Melchers, and Beck [11, 14]. The accuracy of the result is proportional to

\[
\frac{D}{\sqrt{n \cdot N}}
\]

(2)

where \(D\) is a constant, \(n\) is the number of runs, and \(N\) is the number of tests in a run; then, to improve the accuracy by an order of magnitude, the counting time should be increased by two orders of magnitude.

Based on the general state and relevance of this issue at the current stage of PC development, effective algorithms should be searched to solve the problems set. Algorithms should minimize possible errors in various algebraic transformations of the large number set arrays in the course of simulating the structural components of bridges.

3 Results

Statistical simulating the structural components of bridges involves the same chain of operations with the deterministic solution of the problem. The difference is in the representation form and the logic of transitions from input parameters to intermediate or output ones [17, 18]. In a deterministic representation, the quantitative estimation of parameters is performed by a single number only. In the statistical simulation, quantitative estimation of parameters is performed statistically using number sets as an interval estimate. Number sets of input parameters are obtained using a random number generator based on the distribution law peculiar to the parameter considered [17, 18].

The statistical simulation technique allows representing number sets of intermediate and output parameters, which seems impossible for statistical transformations based on analytical methods. In the statistical simulation, transitions from input parameters to intermediate and output ones differ from those based on deterministic techniques by the necessity to include all possible combinations of numbers involved in the calculation.

The number set operations are computation-intensive, moreover, operations with sets suppose complete enumeration of numerical values that greatly limits the capabilities of numerical methods. To reduce the initial set size, many authors have proposed various techniques aimed at optimizing the simulation experiment; Law and Kelton have studied techniques based on the sensitivity analysis, i.e. an attempt to determine the dependence between the input parameters and the output estimates in a single run of simulation model has been made [18]. Whitlock and Kalosand in their research have applied a technique based on stepwise comparison of sets; here, the set size effect on convergence to an analytical solution has been analyzed on the example of the structural component reliability problem [19, 20].

Simulation of output and intermediate parameters implies the presence of number sets defining the statistical distribution of input parameters and the determinate functions of transition from input parameters to output or intermediate ones. In engineering systems, intermediate parameters may include total loads, acting load forces, and geometric cross-section parameters. Output parameters usually include the total load stresses, the total displacements of individual points of the structure, and the specific strength of the materials.

Given the variety of approaches, we have analyzed the efficiency of obtaining number sets of output and intermediate parameters by known number sets of input parameters. The technique of working with sets proposed is based on the stepwise use of the set elements according to their position numbers.

As the source data, we have considered design solutions for the steel part of the road bridge trestle. We have performed a simulation of an intermediate parameter - the total permanent load on the trestle
span, which includes seven weight summands: the trestle span steel structure, asphalt concrete pavement, waterproofing, protection fence, utilities and furniture, drainage, and inspection ladders.

Before choosing a number set processing technique, statistical evaluation of input parameters of the permanent loads should be performed. The procedure for obtaining statistical source data is based on general recommendations for assigning the rated load values [11, 12]. The regulatory sources of various countries are not uniform in the assignment of rated values due to the insufficient statistical data. Possible causes of uncertainty in the interpretation of the statistical dispersion of loads due to the own weight are the material density variation, deviations of the geometrical design parameters, uncertainty in the final choice of material, additional dispersion of loads due to joints and fixtures, and possible subsequent structural changes. It is known that the possible load variation is determined by the relation between the rated and design values, however, the current design standards do not provide a clear statistical interpretation of this relation [24, 25, 26]. Most often, the rated values are determined by the project scope.

The common rules for assigning the rated values of permanent loads \( q_{n,k} \) recommend accepting either average statistical values at a variation coefficient of about 0.1 or 0.95 quantiles in the case of a larger variation coefficient at an unfavorable load on the structure. The rated values \( q_{n,k} \) have been taken as the average expected values of the permanent loads \( \mu_{q,k} \), and the ratio between the corresponding standard value \( q_{n,k} = \mu_{q,k} \) of each separate \( k \)th load \( \Delta q_{k} \) and the load reliability factor \( \gamma_{q,k} \) has been used to evaluate the standard deviation of this load [24]:

\[
\Delta q_{k} = \frac{\mu_{q,k} \cdot (\gamma_{q,k} - 1)}{2} \quad (3)
\]

In the Eq. (3) between the average expected value and the standard deviation, the probability of non-exceedance of the design permanent load value is 0.9772. The relation between the variation coefficient \( \nu_{q,k} \) of each permanent load considered and the corresponding average expected value \( \mu_{q,k} \) is determined by the formula:

\[
\nu_{q,k} = \frac{\Delta q_{k}}{\mu_{q,k}} \quad (4)
\]

Herewith, it should be noted that the example considered is a test one. For this reason, the statistical source data of seven loads given in table 1 have been substantially aggregated.

**Table 1.** Statistical evaluation of input parameters of permanent loadson a single main girder of trestle.

| Parameter Index, \( k \) | Permanent Load                     | \( q_{n,k} \) kN/m | \( \gamma_{q,k} \) | \( \mu_{q,k} \) kN/m | \( \Delta q_{k} \) kN/m | \( \nu_{q,k} \) |
|---------------------------|------------------------------------|--------------------|------------------|-------------------|------------------|----------------|
| 1                         | Trestle span steel structure weight| 35.90              | 1.1              | 35.90             | 1.80             | 0.05           |
| 2                         | Asphalt concrete pavement weight   | 19.48              | 1.5              | 19.48             | 4.87             | 0.25           |
| 3                         | Waterproofing weight              | 1.16               | 1.3              | 1.16              | 0.17             | 0.15           |
| 4                         | Protection fence weight           | 0.70               | 1.1              | 0.70              | 0.04             | 0.05           |
| 5                         | Utilities and furniture weight     | 1.50               | 1.1              | 1.50              | 0.08             | 0.05           |
| 6                         | Drainage weight                   | 0.50               | 1.1              | 0.50              | 0.03             | 0.05           |
| 7                         | Inspection ladders weight         | 2.00               | 1.1              | 2.00              | 0.10             | 0.05           |
| **\( \Sigma q \)**        | Total load                        | 61.24              |                  |                   |                  |                |

Figures 1 and 2 show fragments of the total permanent load simulation based on the source data shown in table 1 at an accepted certainty factor of 0.9772 for different simulation experiment sizes \( (n \cdot N) \).
Figure 1. A fragment of the total permanent load simulation at the total simulation experiment size \((n \cdot N) = 10^4\) realizations. Statistical data: sample mean \(\mu_{q,n} = 61.223\) kN/m, variation coefficient \(\nu_{q,n} = 0.086\), confidence interval is \(\sum q \in [61.103; 61.342]\) kN/m.

Figure 2. A fragment of the total permanent load simulation at the total simulation experiment size \((n \cdot N) = 7.5 \times 10^6\) realizations. Statistical data: sample mean \(\mu_{q,n} = 61.242\) kN/m, variation coefficient \(\nu_{q,n} = 0.085\), confidence interval is \(\sum q \in [61.238; 61.247]\) kN/m.

A comparison of the plots in Figures 1 and 2 demonstrated a change in the shape of the distribution density histogram of random total load realization, which increasingly repeated the plot of the normally distributed variable with the increase in the simulation experiment size.

4 Discussion
The output simulation data is stochastic; conclusions and solutions made are based on a single model run may only be considered as a specific parameter realization. This requires multiple runs of the simulation model to choose the statistical analysis techniques appropriate to the output data.

The number of the simulation model runs is a static series of \(n\) terms, in which the random errors of individual runs can be arranged in increasing order \(\Delta_{(1)} \leq \Delta_{(2)} \leq \ldots \Delta_{(n)}\); the static series terms are estimates of the corresponding quantiles that divide the possible probability interval \((0 \text{ to } 1)\) by \((n+1)\) parts of equal probability. The static series of \(n\) terms determines the boundaries of \((n+1)\) intervals, the probability of falling into which is assumed to be equal, then, the number of runs can be estimated with a certainty factor of maximum \((n-1)/(n+1)\), while only discarding the intervals \((-\infty, \Delta_{(1)})\) and \((-\infty, \Delta_{(n)})\).
The technique of working with sets, based on the stepwise use of the set elements according to their position numbers \[ f_k(j)N \text{ type}, \] involves creating a sample of realizations of random variables corresponding to some specific probability distribution with density \( f_k(j) \) and subsequent row-wise transformation of \( j \)th values of the resulting number sets based on deterministic interaction functions.

This processing technique allows creating number sets as random vectors \( k \) of input parameters and ensuring transition from input parameters to output or intermediate ones based on known analytical dependencies, e.g., structural mechanics. In this case, there is no need to generate possible combinations of input parameters, and the required number set size will be determined by trial runs of the simulation model until stable results are achieved. The number of test runs should be determined based on a statistical analysis of output or intermediate parameters, e.g., by studying the confidence interval resistance to deviations from the initial probability distribution of input parameters.

If taking a fixed number set size equal to \( 10^6 \) for each load, then for \( k \) input parameters, the same number of numerical values will be required. For this processing technique, even at 200 runs for each position number, the computational effort will amount to only \( 2 \times 10^8 \) values, which seems to be quite acceptable for designing the bridge structures, including using available random number generators of PC systems.

In the analysis performed, the number of simulation model runs \( n \) is fixed at 50 and 100, which directly correlates with a given reliability level Eq. (5)

\[ \beta \leq \frac{n-1}{n+1} \]

i.e., about 0.96 and 0.98, respectively. Any model with fixed sample size is estimated by the relative entropy \( S_{eq} \) represented by the ratio between the upper fiducial limit values of the parameter considered, obtained based on the Student’s distribution, to the upper fiducial limit values of the same parameter obtained based on the Chebyshev’s inequality [18]. The physical meaning of the relative entropy is to obtain a numerical estimate reflecting the relative dependence of the intermediate parameter uncertainty on the input parameter random value distribution lows chosen in the simulation.

In our example, the total permanent load acts as the input parameter, and the output one is an individual permanent load. The uncertainty level is attributable to an arbitrarily or insufficiently justifiably chosen random value distribution law.

Figure 3a and Figure 3b shows fragments of individual runs of the span total permanent load intermediate parameter simulation model.

![Figure 3a](image-url)  
**Figure 3a.** Dependence of entropy on the number of tests at 50 runs, reliability 0.96.
5 Conclusions
A simulation experiment with processing the number sets using the above technique has shown that the entropy fluctuation intensity decreases with an increase in the number of tests. The main part of the simulation experiment requires test runs that allow identifying the minimum required amount of simulation using, e.g., the criterion of the target parameter fluctuation range.

The span total permanent load simulation results demonstrate that the parameter entropy value asymptotically tends to a certain limit. When increasing the realization number to 7.5 \times 10^6, the total load values flatten out. To experimentally justify the minimum size of the number sets $N$ required in the simulation of structural components of bridges, the required system reliability $n$ can be specified using the equation $N = 7.5 \times 10^6/n$.

Reference
[1] Jebur H and Al-Zaidee S 2019 Non-deterministic Approach for Reliability Evaluation of Steel Portal Frame Civ. Eng. J. 5(8) pp 1684–97
[2] Li H and Cao Z 2016 Matlab codes of Subset Simulation for reliability analysis and structural optimization Struct. Multidiscip. Optim. 54(2) pp 391–410 DOI: 10.1007/s00158-016-1414-5
[3] Rebba R and Mahadevan S 2008 Computational methods for model reliability assessment, Reliability Engineering and System Safety 93(8) pp 1197–1207 DOI: 10.1016/j.ress.2007.08.001
[4] Vidia Sagar J, Manjunath G 2017 Probability of failure of column and beam in steel structure due to plan irregularities Int. Res. J. Eng. Technol. 4(8) pp 1022–28
[5] Pichugin S 2018 Reliability estimation of industrial building structures Magazine of Civil Engineering 7 pp 24–37 DOI: 10.18720/MCE.83.3
[6] Shadab Far M and Wang Y 2016 Approximation of the Monte Carlo Sampling Method for Reliability Analysis of Structures Math. Probl. Eng. 2016 pp 56-78 DOI: 10.1155/2016/5726565
[7] Jirgl M, Bradac Z, Stibor K and Havlikova M 2013 Reliability Analysis of Systems with a Complex Structure Using Monte Carlo Approach IFAC Proceedings 46(28) pp 461–466 DOI: 10.3182/20130925-3-CZ-3023.00031
[8] Buonopane S and Schafer B 2006 Reliability of steel frames designed with advanced analysis Journal of Structural Engineering 132(2) pp 267–276 DOI: 10.1061/(ASCE)0733-9445(2006)132:2(267)
[9] Darmawan M, Refani A, Irmawan M, Bayuaji R and Anugraha R 2013 Time Dependent
Reliability Analysis of Steel I Bridge Girder Designed Based on SNI T-02-2005 and SNI T-3-2005 Subjected to Corrosion Procedia Engineering 54 pp 270–285 DOI: 10.1016/j.proeng.2013.03.025

[10] Novitsky P and Zograf I 1991 Evaluation of errors of measurement results 2nd ed, Revised and add. (Leningrad: Publishing house Energotoatomizdat)

[11] Perelmuter A 2007 Selected problems of reliability and safety of building structures (Moscow: Publishing house of the association of construction universities)

[12] Elishakoff I 2017 Probabilistic Methods in the Theory of Structures (United States of America: World Scientific Publishing Co. Pte. Ltd.)

[13] Kang G 2017 Reliability Analysis Using the Method of Multiplicative Dimensional Reduction, Waterloo, Ontario, Canada p 147

[14] Melchers R and Beck A 2017 Structural Reliability Analysis and Prediction (John Wiley & Sons)

[15] Ghali A, Neville A and Brown T 2009 Structural Analysis 6th ed. (Canada: Taylor & Francis Group)

[16] Morio J and Balesdent M 2016 Introduction to Rare Event Probability Estimation Estimation of Rare Event Probabilities in Complex Aerospace and Other Systems (Woodhead Publishing)

[17] Thomopoulos N 2012 Essentials of Monte Carlo simulation: Statistical methods for building simulation models (Springer Science & Business Media)

[18] Law A and Kelton W 2000 Simulation modeling and analysis (McGraw-Hill)

[19] Kalosand M and Whitlock P 2008 Monte Carlo Methods 2nd ed. (USA: John Wiley&Sons)

[20] Naess A, Leira B and Batsevych O 2009 System Reliability Analysis by Enhanced Monte Carlo Simulation Structural safety 31(5) pp 349–355 DOI: 10.1016/j.strusafe.2009.02.004

[21] Thomopoulos N 2012 Essentials of Monte Carlo simulation: Statistical methods for building simulation models (Springer Science & Business Media)

[22] Maistrenko I and Manapov A 2010 Modeling the process of changing in time the level of reliability of a structural system Izvestiya KGASU 1(13) pp 132–140

[23] Zinnurov T, Kayumov R and Manapov A 2012 On the sensitivity of the results of statistical modeling of constant and wind loads on structures to deviations of the parameters of their distribution laws News of Universities Construction 1 pp 115–121

[24] SP.35.13339.2011. Bridges and pipes. The updated edition of SNiP 2.05.03-84 *. Moscow OJSC TsPP (2011)

[25] ANSI, B. AISC 360-10-Specification for Structural Steel Buildings [J]. Chicago AISC 2010

[26] EN 1991-1-1 Eurocode 1: Actions on structures. Part 1-1: General actions. BSI 2005