Journal of Computers Vol. 33 No. 2, April 2022, pp. 169-176
doi: 10.53106/199115992022043302015

Missing Data Interpolation with Variational Bayesian Inference for Socio-economic Statistics Applications

Yun-Shan Sun[1], Hong-Yan Xu[2*], Yan-Qin Li[1]

[1] School of Information Engineering, Tianjin University of Commerce, Tianjin 300134, P.R. China
sunyunshan@tjcu.edu.cn, liyanqin-2190@126.com
[2] School of Science, Tianjin University of Commerce, Tianjin 300134, P.R. China
2552727224@qq.com

Received: 19 October 2021; Revised: 19 November 2021; Accepted: 24 November 2021

Abstract. The information integrity is needed to solving socio-economic statistical problems. However, the information integrity is destroyed by missing data which is caused by various subjective and objective reasons. So the missing data interpolation is used to supplement missing data. In this paper, missing data interpolation with variational Bayesian inference is proposed. This method is combined with Gaussian model to approximate the posterior distribution to obtain complete data. The experiments include two datasets (artificial dataset and actual dataset) based on three missing ratios separately. The missing data interpolation performance of variational Bayesian method is compared with that which is obtained by mean interpolation and K-nearest neighbor interpolation methods separately in MSE (Mean Square Error) and MAPE (Mean Absolute Percentage Error). The experimental results show that the proposed variational Bayesian method is better in MSE and MAPE.

Keywords: missing data, variational Bayesian, interpolation, posterior distribution, K-nearest neighbor

1 Introduction

The quantitative features and relationships in socio-economic data are mainly studied in socio-economic statistics. Socio-economic data come from national statistics offices and relevant industrial organizations [1]. The accurate, timely and systematic statistical analysis can be offered by the reliable and integrated socio-economic data for policy-making and multi-disciplinary study. However, the inaccurate statistical analysis always appears due to the missing data which leads to incorrect decisions and even huge economic losses [2, 3]. Data missing is caused due to the following reasons, such as insufficient survey, data collection equipment failure [4], the subjective interviewee's will, the investigator's errors, the history reasons, undisclosed sources and so on.

In order to solving data missing problems, deleting missing data directly is the easiest way [5] which is widely used before. However, statistical errors and nonintegrated information are caused. Therefore, in order to ensure the data completeness, interpolation strategy which replaces the missing data with the approximate data is proposed. The aim of interpolation is to obtain approximate data infinitely close to the true value [6].

Interpolation methods are mainly divided into single and multiple. The single interpolation method mainly includes mean interpolation [7], regression interpolation [8], K-means interpolation [9]and Calorie interpolation [10], etc. The single interpolation method is simple and widely used in various fields in recent years. However, the uncertainty of sample distribution is easily ignored in single interpolation. There are also serious deficiencies in complex single data missing problems [11]. Therefore, the overall performance of single interpolation method is not good enough.

The multiple interpolation [12] method usually refers to the construction of multiple complete datasets based on the same group of missing datasets. And then, the optimized interpolation set is obtained based on multiple complete datasets. The candidate data can be got through each interpolation by the multiple interpolation method. The uncertain of single interpolation is complemented effectively by the multiple interpolation method. The main multiple interpolation method is shown as following, regression predictive interpolation [13], preference-based scoring [14] and Markov chain Monte Carlo (MCMC) method [15, 16], etc. The regression predictive interpolation method and the preference-based scoring method are mainly suitable for nonrandom missing pattern while the Markov chain Monte Carlo [17] method is suitable for random missing pattern. The more accurate interpolation results can be obtained by introducing prior information in Markov chain Monte Carlo method. However, this method is complex.

* Corresponding Author

KNN (K-nearest neighbor) and random forest methods are more popular in missing data interpolation problem. The KNN interpolation method performance [18] is obviously better than the traditional method (such as mean, median and mode interpolation). However, high-dimensional data in KNN can be severely degraded because of little difference between the nearest and farthest neighbors. The error can be reduced by random forest interpolation method [19] which is combined with multiple decision trees. However, the random forest interpolation model is easy to over-fit.

In order to solve missing data interpolation problem, the variational Bayesian method is proposed in this paper. The Variational Bayesian is an approximate inference method which combines traditional Bayesian method with machine learning. Not only the original advantages are retained, but also the computing efficiency is improved in the variational Bayesian [20, 21].

This paper is organized as follows. Bayesian correlation theory is introduced in Section 2. The framework of variational Bayesian method is described in Section 3. The experiments are shown in Section 4. Finally, the conclusions are drawn in Section 5.

2 Problem Statement

2.1 Problem Statement

In this paper, $X = (x_1, ..., x_N)$ denotes the complete data and N is the number of complete data. $\hat{X}$ denotes the value of data after interpolation. The data X is generated from prior probability $P(X)$, and a part of data cannot be observed. X can be partitioned as the observed data X_{obs} and the missing data X_{mis}. The relation between X, X_{obs} and X_{mis} is as follows.

$$X = \left(X_1, \cdots, X_N\right) = \left(X_{\text{obs}_1}, X_{\text{mis}_1}, \cdots, X_{\text{obs}_L}, X_{\text{mis}_K}\right) \overset{\text{def}}{=} \left(X_{\text{obs}}, X_{\text{mis}}\right) \cdot \tag{1}$$

where, $X_{\text{obs}} = \left(x_{\text{obs}_1}, \ldots, x_{\text{obs}_L}\right)$ and L is the number of the observed data. $X_{\text{mis}} = \left(x_{\text{mis}_1}, \ldots, x_{\text{obs}_K}\right)$ and K is the number of the missing data.

In some cases, the missing data X_{mis} can't be direct deleted because it holds important information. To solve the missing data interpolation problem, the missing data X_{mis} is estimated and interpolated according to the observed data and prior information in Bayesian method.

2.2 Basic Bayesian Method

The partial interpolation problem of missing data can be solved by Bayesian method, in which the complete data set can be obtained by the deduce posterior distribution with observed data. At present, Metropolis–Hastings and Gibbs sampling method, both of which belong to Markov chain Monte Carlo, are both need to be updated iteratively until they are optimal.

2.2.1 Metropolis–Hastings Method

The Metropolis-Hastings method [22] is an early and relatively general MCMC method. a target distribution π(x) should be sampled, while it is difficult to sample directly. Therefore, an appropriate conditional distribution function is selected to sample π(x). The basic steps are as follows:

The transfer function is defined as $q\left(X, X^{(i-1)}\right)$ and the initial value is defined as $X^{(0)}$. The beginning of the i iteration is defined as $X^{(i-1)}$, then the subsequent iteration process is expressed as follows:

The sample point X' is selected in $q\left(X, X^{(i-1)}\right)$;

Sample from uniform distribution $u \sim U\,[0,1]$;

The acceptance probability is calculated as:

$$\alpha\left(X^{(i-1)}, X'\right) = \min\left\{\frac{\pi\left(X^{(i-1)}; X'\right)}{\pi\left(X^{(i-1)}\right)\mathrm{q}\left(X'; X^{(i-1)}\right)}\right\}. \tag{2}$$

If $\mathbf{u} < \alpha\left(X^{(i-1)}, X'\right)$, then $X^{(i)} = X'$, if $\mathbf{u} \geq \alpha\left(X^{(i-1)}, X'\right)$, then $X^{(i)} = X^{(i-1)}$;

Repeat steps above N times, until the posterior sample $X^{(1)}, X^{(2)}, \cdots, X^{(N)}$ is obtained.

Based on the posterior samples, the posterior distribution moments can be calculated. Finally, the corresponding statistical inference can be obtained.

2.2.2 Gibbs Sampling Method

Gibbs Sampling can be regarded as one of the special cases of Metropolis–Hastings method, which is a statistical simulation method used in Bayesian statistics. The Gibbs sampling method [23] is used to approximate sample sequence from a multivariable probability distribution. The joint distribution and marginal distribution can be approximated by the Gibbs sampling method. And the multivariate distribution of Gibbs sampling method can be obtained by the low-dimensional sampling distribution. Therefore, it is more suitable for higher dimensional problems (two dimensions or more). The basic idea is as follows:

select an arbitrary initial vector $X^{(0)} = \left(X_1^{(0)}, \cdots, X_k^{(0)}\right)$;

extracting sample $X_1^{(1)}$ from $\pi\left(X_1 \mid X_1^{(0)}, \cdots, X_k^{(0)}\right)$;

extracting sample $X_2^{(1)}$ from $\pi\left(X_2 \mid X_1^{(1)}, \cdots, X_k^{(0)}\right)$;

extracting sample $X_j^{(1)}$ from $\pi\left(X_j \mid X_1^{(1)}, \cdots, X_{j-1}^{(1)}, X_{j-1}^{(0)}, \cdots, X_k^{(0)}\right)$;

extracting sample $X_k^{(1)}$ from $\pi\left(X_k \mid X_1^{(1)}, \cdots, X_{k-1}^{(1)}\right)$;

Based on above steps, the complete $X^{(1)}$ is obtained from $X^{(0)}$. After n times iteration, the posterior samples $X^{(1)}, X^{(2)}, \cdots, X^{(N)}$ is obtained. Based on the posterior samples, the posterior distribution moments can be calculated. Finally, the corresponding statistical inference can be obtained.

2.2.3 Variational Bayesian Method

The complex posterior probability of the statistical models in Bayesian need to calculated. To simplify the calculation, an arbitrary distribution $\mathrm{q}(X)$ is used to approximates the posterior probability distribution, where X is missing data and observed data. Therefore, the complex posterior probability distribution is replaced by the approximate distribution $\mathrm{q}(X)$ problem.

Traditional Bayesian formula is defined as

$$P\left(X_{\text{obs}}\right) = \frac{P\left(X_{\text{obs}}, X\right)}{P\left(X \mid X_{\text{obs}}\right)}. \tag{3}$$

where $P\left(X_{\text{obs}}\right)$ denotes the probability distribution of the model, $P\left(X_{\text{obs}}, X\right)$ denotes likelihood function and $P(X_{\mathrm{obs}},\, X)$ denotes the posterior probability distribution. X is hidden variable partitioned as $(X_{\mathrm{obs}},\, X_{\mathrm{mis}})$, where X_{obs} denotes the observed part and X_{mis} represents the missing part.

Introduce arbitrary distribution $\mathrm{q}(X)$, and then take the log function on both side of (2.3). The equation is obtained as,

$$\ln\left(P\left(X_{\mathrm{obs}}\right)\right) = \ln\left(\frac{P\left(X_{\mathrm{obs}},\, X\right)}{\mathrm{q}(X)}\right) - \ln\left(\frac{P\left(X \mid X_{\mathrm{obs}}\right)}{\mathrm{q}(X)}\right) . \tag{4}$$

Take the expectation of distribution $\mathrm{q}(X)$

$$\int \ln\left(P\left(X_{\mathrm{obs}}\right)\right)\mathrm{q}(X)\mathrm{d}X = \int \ln\left(\frac{P\left(X_{\mathrm{obs}},\, X\right)}{\mathrm{q}(X)}\right)\mathrm{q}(X)\mathrm{d}X - \int \ln\left(\frac{P\left(X \mid X_{\mathrm{obs}}\right)}{\mathrm{q}(X)}\right)\mathrm{q}(X)\mathrm{d}X = \ln\left(P\left(X_{\mathrm{obs}}\right)\right) . \tag{5}$$

where

$$F(X) = -\int \ln\left(\frac{P\left(X_{\mathrm{obs}},\, X\right)}{\mathrm{q}(X)}\right)\mathrm{q}(X)\mathrm{d}X . \tag{6}$$

$$\mathrm{D}_{\mathrm{KL}} = -\int \ln\left(\frac{P\left(X \mid X_{\mathrm{obs}}\right)}{\mathrm{q}(X)}\right)\mathrm{q}(X)\mathrm{d}X . \tag{7}$$

equation (5) can be simplified as

$$F(X) = -\ln\left(P(X)\right) + \mathrm{D}_{\mathrm{KL}} . \tag{8}$$

Where free energy $F(X)$ and KL divergence D_{KL} are indexes of evaluate the models.

2.2.4 Bayesian Model Valuation

Before introducing AIC (Akaike information criterion) and BIC (Bayesian Information Criterions), Kullback-Leibler distance should be studied firstly. The difference of models is represented by Kullback-Leibler distance index. The complex KL distance can be obtained by the information loss function and the information content criteria AIC and BIC. The complexity and degree of fitting of statistical models can evaluated by AIC and BIC. The small AIC and BIC means the better model fitting degree. The AIC and BIC are defined separately as:

$$\mathrm{AIC} = 2k + n\ln\left(\frac{RSS}{n}\right) . \tag{9}$$

$$\mathrm{BIC} = k\ln(n) - 2\ln(L) \cdot \tag{10}$$

where k, n and L denote the number of model parameters, the number of observation data sample, and the likelihood function value respectively. RSS is the sum of squares of residuals.

The AIC and BIC both exist penalty terms related to the number of model parameters. The AIC make decision from the perspective of prediction, while the BIC make decision from the perspective of fitting. In addition, because BIC is more punish to model parameters in large data scale, the model with fewer parameters is easy to be selected by BIC.

For equation (5) mentioned above, Let's set $L(\mathrm{q}(X))_{\mathrm{ELOB}} = \int \ln\left(\frac{P(X_{\mathrm{obs}}, X)}{\mathrm{q}(X)}\right)\mathrm{q}(X)\mathrm{d}X = -F(X)$ and $\mathrm{D}_{\mathrm{KL}} = -\int \ln\left(\frac{P(X \mid X_{\mathrm{obs}})}{\mathrm{q}(X)}\right)\mathrm{q}(X)\mathrm{d}X$,

where, $L(\mathrm{q}(X))_{\mathrm{ELOB}}$ denotes the lower bound. Let $F(X) = -L(q(X))_{ELOB}$,where $F(X)$ is Free Energy in this paper. Besides,

$$\begin{aligned}\mathrm{D}_{\mathrm{KL}} &= -\mathrm{E}_{\mathrm{q}}\left[\ln\left(\frac{P(X \mid X_{\mathrm{obs}})}{\mathrm{q}(X)}\right)\right]\\ &= \mathrm{E}_{\mathrm{q}}[\ln \mathrm{q}(X)] - \mathrm{E}_{\mathrm{q}}[\ln P(X \mid X_{\mathrm{obs}})]\\ &= \mathrm{E}_{\mathrm{q}}[\ln \mathrm{q}(X)] - \mathrm{E}_{\mathrm{q}}[\ln P(X, X_{\mathrm{obs}})] + \ln P(X_{\mathrm{obs}})\\ &= -(\mathrm{E}_{\mathrm{q}}[\ln P(X, X_{\mathrm{obs}})] - \mathrm{E}_{\mathrm{q}}[\ln \mathrm{q}(X)]) + \ln P(X_{\mathrm{obs}})\end{aligned} \quad . \qquad \textbf{(11)}$$

Since $\ln P(X_{\mathbf{obs}})$ is independent with q, minimizing the free energy (maximizing the ELBO) is the same as the minimizing D_{LK} divergence. Both ELBO and D_{LK} can be used to measure the approximation degree of $\mathrm{q}(X)$ to the posteriori probability.

3 A VB Framework for the Recovery of Missing Data

The aim of variational Bayesian learning method is to obtain the interpolated data X_{mis} , minimize index $F(X) = -\int \ln\left(\frac{P(X_{\mathrm{obs}}, X)}{\mathrm{q}(X)}\right)\mathrm{q}(X)\mathrm{d}X$. The variational Bayesian learning method is presented as follows.

Step 1: Input parameters or variables (e.g. the observed data X_{obs} , the data missing ratio d and the number of all data N and so on.) and set up GLM (Generalized linear model) with missing data.

Step 2: Initialize parameters and prior information. And store the free energy.

Step 3: Update and optimize parameters by means of iterative repeatedly.

Step 4: Repeat Step 3 until the termination condition is satisfied. The posterior results are showed with charts.

4 Data Analysis and Experimental Design

4.1 Datasets and Experiment Setup

An artificial dataset and a life expectancy at birth dataset (47 African countries) have been selected in these experiments separately. The actual data is from the IMF (International Monetary Fund). The experiments are as follows:

Step 1: For simulated data subject to Gaussian distribution and the life expectancy data, construct missing datasets in different proportion (the missing ratios are 5%, 10% and 20% respectively).

Step 2: Interpolate the missing data by mean interpolation method, KNN interpolation method and variational Bayesian interpolation method successively. For KNN interpolation method, k value is selected as 4 to avoid the model being too simple or over-fitting. The attributes of datasets are listed in Table 1.

Table 1. Datasets attributes

Datasets	Range of data	Size
Simulate data	[0, 1]	10000
Life expectancy	[0,+∞)	987

Step 3: Finally, the interpolation results of each method are compared by means of statistical index MSE and MAPE. They are defined separately as

$$\mathrm{MSE}\left(\hat{X}\right) = \mathrm{E}\left(\hat{X} - X\right)^2 . \tag{12}$$

$$\mathrm{MAPE} = \frac{\left|X - \hat{X}\right|}{nX} . \tag{13}$$

4.2 Comparison Experiments Results

To further explore the performance of the variational Bayesian method on interpolate missing data, we compare the MSE and MAPE with other interpolation (mean interpolation and KNN interpolation) methods' results. The statistical results (the artificial and the actual datasets) are listed in Table 2 and Table 3 separately.

Table 2. Simulation data interpolation comparison results under different missing ratios

Methods	Ratio	MSE	MAPE
KNN	5%	0.0714	15.0655
	10%	0.1276	27.0304
	20%	N/A	N/A
Mean	5%	0.0531	4.9977
	10%	0.0993	9.9738
	20%	0.1956	19.9572
VB	5%	0.0531	4.9989
	10%	0.0992	9.9824
	20%	0.1955	19.9950

Table 3. Actual data (Life expectancy) interpolation comparison results under different missing ratios

Methods	Ratio	MSE	MAPE
KNN	5%	0.6380	0.2432
	10%	0.4315	0.2805
	20%	2.5687	0.7883
Mean	5%	4.3511	0.7712
	10%	5.0685	1.0378
	20%	13.2615	2.3595
VB	5%	0.0070	0.0345
	10%	0.0034	0.0328
	20%	0.0016	0.0314

According to the above results, the MSE and MAPE values of KNN method are empty when the simulated data missing ratio is 20%. That means interpolation is not achieved. For the simulated data, there is no significant difference between the results of mean interpolation method and Variational Bayesian interpolation method, both of which are better than KNN interpolation method. For the missing data interpolation of actual data, the value of variational Bayesian method is significantly lower than that of the other two methods. The numerical results can be drawn as a broken line graph shown in Fig. 1.

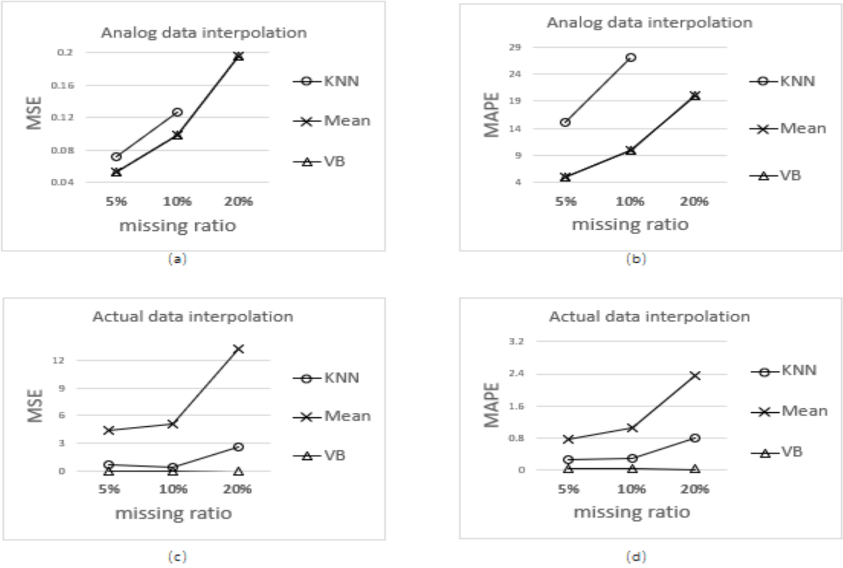

Fig. 1. Comparison diagram of interpolation results

As shown in Fig. 1(a) and Fig. 1(b), the results of mean interpolation method and variational Bayesian interpolation method are too similar, and the trend lines almost coincide. Fig. 1(c) and Fig. 1(d) represent the interpolation results of actual data. There are obvious differences among different methods. In addition, as can be seen from the Table 4, AIC and BIC values in the same dataset with different missing ratios are relatively close without significant difference, which reflects the stability of the model.

Table 4. Gauss model measurement index value

Datasets	Indexes	Ratio		
		5%	10%	20%
Actual data	AIC	3.264e+04	3.173e+04	3.028e+04
	BIC	3.083e+04	2.812e+04	2.307e+04
Simulation data	AIC	-3.900e+03	-4.324e+03	-4.784e+03
	BIC	-4.022e+03	-4.569e+03	-5.269e+03

As we can see from the experimental results, KNN interpolation method cannot effectively interpolate the missing data when the scale of simulated data is large and the missing ratio is high. While the interpolation results of variational Bayesian method are better. At the same time, the mean interpolation method has no advantage under the same condition. Therefore, the interpolation results obtained by the variational Bayesian method have obvious advantage in both two datasets.

5 Conclusion

The information integrity in socio-economic data is achieved based on the variational Bayesian theory in this paper. In addition, the missing data interpolation performance is further optimized. Based on Gaussian model, the Bayesian posterior distribution and the complete data sets are obtained by optimizing relevant parameters. Finally, the missing data interpolation performance of variational Bayesian method is compared with that which is obtained by mean interpolation and KNN interpolation separately. The experiments include simulated and artificial data sets with different missing ratios. The experimental results show that the interpolation results of variational Bayesian method have obvious advantages in MSE and MAPE.

On the other hand, this method also has some limitations. The results obtained by variational Bayesian interpolation and mean interpolation are very close in simulated data interpolation. However, the calculation time of mean interpolation is obviously shorter. In addition, the proposed method is only limited to single variable. In the future work, establishing other missing data statistical models in different fields will be studied to achieve better performance.

Acknowledgments

The authors thank the editor and anonymous reviewers for their helpful comments and valuable suggestions during the revision of this paper. As well as a part of experimental data provided by professor Habtamu Tilaye Wubetie.

References

[1] T. W. Habtamu, Missing data management and statistical measurement of socio-economic status: application of big data, Journal of Big Data 4(2017) 4-47.

[2] J.A. Hussain, I.R. White, M.J. Johnson, M. Bland, D.C. Currow, Performance status and trial site-level factors are associated with missing data in palliative care trials: An individual participant-level data analysis of 10 phase 3 trials, Palliative medicine, 2021.

[3] D. Jeong, C. Park, Y.M. Ko, Missing data imputation using mixture factor analysis for building electric load data, Applied Energy 304(2021) 117655.

[4] L. Song, J.-Z. Wan, Comparative study on missing data interpolation methods, Statistics and decision 36(2020) 10-14.

[5] J.-X. Liu, H.-L. Zhang, Y.-J. Liu, Y.-Z. Liu, Incomplete data Filling algorithm based on missing rate, Statistics and decision 37(2021) 39-41.

[6] G.-L. Huang, Missing data filling method based on linear interpolation and lightgbm, Journal of Physics: Conference Series 1754(2021) 012187.

[7] S. Ujjwol, A. Alsadoon, P.W.C. Prasad, A.A. Sarmad, H.A. Omar, Supervised machine learning for early predicting the sepsis patient: modified mean imputation and modified chi-square feature selection, Multimedia Tools and Applications 80(2021) 20477-20500.

[8] H.-L. Zhu, Y.-L. Tian, Y.-N. Ren, J.-L. Hu, A Hybrid Model for Nonlinear Regression with Missing Data Using Quasi linear Kernel, IEEJ Transactions on Electrical and Electronic Engineering 15(2020) 1791-1800.

[9] S. Han, L. Sun, Y.-Y. Yang, W.-L. Wu, X.-X. Guo, C.-C. Dai, Data cleaning method based on improved k-means clustering and error feedback, Power System and Clean Energy 36(2020) 9-15.

[10]J.-H. Cao, J.-Y. Liu, H. Xu, Z.-X. Lin, K.-F. Zhang, Test efficiency evaluation of high-speed permanent magnet motor based on grey neural network, Computer Measurement and Control 28(2020) 251-256.

[11]C. Song, X. Yang, X. Shi, Y.-C. Bo, J.-F. Wang, Estimating missing values in China's official socioeconomic statistics using progressive spatiotemporal Bayesian hierarchical modeling, Scientific Reports 8(2018) 10055.

[12]Y. Takeuchi, M. Ogawa, Y. Hagiwara, Y. Matsuyama, Non-parametric approach for frequentist multiple imputation in survival analysis with missing covariates, Statistical methods in medical research 30(2021) 1691-1707.

[13]X.-J. Sun, E. Maehle, N. Stoll, WSN blind area predictive regression control model based on interpolation algorithm optimization, Journal of Computational Methods in Sciences and Engineering 19(2019) 1-7.

[14]T. Chris, Add or Multiply A Tutorial on Ranking and Choosing with Multiple Criteria, Informs Transactions on Education 14(2014) 109-119.

[15]A. Reem, A. K. Samer, Use of Bayesian Markov Chain Monte Carlo Methods to Model Kuwait Medical Genetic Center Data: An Application to Down Syndrome and Mental Retardation, Mathematics 9(2021) 248.

[16]T. Masayoshi, Statistical Inference in Missing Data by MCMC and Non-MCMC Multiple Imputation Algorithms: Assessing the Effects of Between-Imputation Iterations, Data Science Journal 16(2017) 37-53.

[17]Z.-H. Ma, G.-Y. Hu, M.-H. Chen, Bayesian hierarchical spatial regression models for spatial data in the presence of missing covariates with applications, Applied Stochastic Models in Business and Industry 37(2020) 342-359.

[18]S. Karshiev, B. Olimov, K. Jaesoo, P. Anand, K. Jeonghong, Missing Data Imputation for Geolocation-based Price Prediction Using KNN–MCF Method, ISPRS International Journal of Geo-Information 9(2020) 227.

[19]R.-H. Feng, D. Grana, N. Balling, Imputation of missing well log data by random forest and its uncertainty analysis, Computers & Geosciences 152(2021) 104763.

[20]P.-H. Ni, J. Li, H. Hao, Q. Han, X.-L. Du, Probabilistic model updating via variational Bayesian inference and adaptive Gaussian process modeling, Computer Methods in Applied Mechanics and Engineering 383(2021) 113915.

[21]X.-Y. Song, G.-H. Zheng, L.-J. Jiang, Variational Bayesian inversion for the reaction coefficient in space-time nonlocal diffusion equations, Advances in Computational Mathematics 31(2021).

[22]C. Sherlock, A. Golightly, D.A. Henderson, Adaptive, Delayed-Acceptance MCMC for Targets With Expensive Likelihoods, Journal of Computational and Graphical Statistics 26(2017) 434-444.

[23]Y. Wang, An Enhanced Markov Chain Monte Carlo-Integrated Cross-Entropy Method with a Partially Collapsed Gibbs Sampler for Probabilistic Spinning Reserve Adequacy Evaluation of Generating Systems, Electric Power Components and Systems 45(2017) 1617-1628.