JP Journal of Biostatistics
© 2025 Pushpa Publishing House, Prayagraj, India
https://pphmjopenaccess.com
https://doi.org/10.17654/0973514325006
Volume 25, Number 1, March 2025, Pages 127-144 **P-ISSN: 0973-5143**

IMPACT OF BOX-COX TRANSFORMATION TECHNIQUE ON THE BAYESIAN MAXIMUM ENTROPY (BME) PREDICTION ACCURACY

Emmanuel Ehnon Gongnet[1,2], Romaric Vihotogbé[3], Codjo Emile Agbangba[1,4], Tranquillin Affossogbé[1], Koye Djondang[2] and Romain Glèlè Kakaï[1]

[1]Laboratoire de Biomathématiques et d'Estimations Forestières (LABEF)
Université d'Abomey-Calavi
Benin

[2]Institut Tchadien de Recherche Agronomique pour le Développement (ITRAD)
BP 5400, N'Djamena, Tchad

[3]Ecole de Foresterie Tropicale
Université National d'Agriculture
BP 43 Kétou, Benin

Received: July 13, 2024; Accepted: August 27, 2024

Keywords and phrases: Box-Cox transformation, skewness, spatial dependence, Bayesian maximum entropy.

Communicated by K. K. Azad

How to cite this article: Emmanuel Ehnon Gongnet, Romaric Vihotogbé, Codjo Emile Agbangba, Tranquillin Affossogbé, Koye Djondang and Romain Glèlè Kakaï, Impact of Box-Cox transformation technique on the Bayesian maximum entropy (BME) prediction accuracy, JP Journal of Biostatistics 25(1) (2025), 127-144. https://doi.org/10.17654/0973514325006

Published Online: December 10, 2024

[4]Laboratory of Research in Applied Biology
Department of Environmental Engineering
University of Abomey-Calavi
Calavi, Benin

Abstract

This study investigated whether increasing the normality of an attribute using Box-Cox transformation improves Bayesian Maximum Entropy (BME) prediction accuracy. Furthermore, we examined if BME accuracy is affected by sample size or spatial dependence. For hard data, the unconditional sequential approach was used to simulate symmetric data (skewness = 0) and data positively skewed (skewness: 1, 3, 6, and 9) with sample size ranging from 100 to 500 at the interval length of 50. Soft data was randomly distributed throughout a square of unit size and a width of 1.5. Data was then transformed using Box-Cox transformation. The prediction accuracy was assessed using the Mean Square Error (MSE) and bias, and transformation methods were compared using the Multivariate Analysis of Variance (MANOVA). The results showed BME accuracy is affected by transformation methods but not the sample size and the spatial dependency. However, in comparing the transformed data with the untransformed data, the MSE and bias of the untransformed data (lambda = 1) were closer to zero than the transformed data (lambda $\neq$ 1). As a result, we concluded that BME is robust to skewness, sample size, and spatial dependency.

1. Introduction

Understanding a system for its improvement based on sound proposals calls for statistical analysis. Depending on the statistical methods to be applied, the data can be summarized using descriptive statistics and examined through statistical inference to test the underlying hypotheses. One of the most frequent hypotheses of statistical tests is the normality of the dataset [1]. Normality is generally characterized by skewness and kurtosis, but the commonly used formal normality tests are: Shapiro-Wilk test and

Kolmogorov Smirnov test [2]. Statistical methods (e.g., multiple regression, ANOVA) are often sensitive to non-normality, heteroscedasticity, and sample size. Data transformations are sometimes required to improve the accuracy of statistical methods, and then the results must be back transformed prior to interpretation.

In geostatistics, the classical method known as kriging is highly sensitive to dataset normality [3]. A transformation on raw dataset is performed to improve its normality and therefore guarantees the prediction accuracy [4]. From this perspective, the logarithmic transformation has been used, regardless of the data nature [5-7]. Since the accuracy of kriging based estimations is frequently criticized due to their failure to handle soft data, the Bayesian Maximum Entropy (BME) [8, 9] was proposed to improve geostatistical estimation [10]. BME approach is built upon a strong mathematical procedure [11, 12].

As a result, BME might be robust regardless of data skewness or sample size [13-15]. However, BME framework hinges on the covariance and variogram models [13], while the variogram, which measures the spatial dependency, might be strongly affected by the data skewness and sample size [16]. Thus, it is demonstrated that variogram calculation for highly skewed data results in an underestimation of spatial dependency [17], whereas an accountable variogram calculation in the spatial analysis process requires a minimum sample size of 100 [18]. The degree of skewness influences the measure of entropy; when the data is highly skewed, the measure of entropy is underestimated [7]. In BME analysis, the choice between raw skewed data (untransformed) prior to BME application, plus back-transformation afterwards, and no application of any transformation, regardless of the nature of the data, occurs [5, 19]. In general, whether a statistical test is considered robust to non-normal data or nonparametric, taking dataset normality into account can improve the accuracy of the results [20]. This study focuses on factors that are countable for better estimation under the BME application. Therefore, relying on the quality of the raw dataset, we provided answers to the following research questions: What are the impacts of using Box-Cox transformation on BME accuracy? Are the effects of managing data normality through Box-Cox transformations influenced by sample size or spatial dependency?

2. Background on the BME

The Bayesian Maximum Entropy (BME) is a robust geostatistical approach designed for spatio-temporal prediction and mapping [8, 9]. BME provides mathematical framework for including different data types into mapping process [12]. The available knowledge is denoted by *K*, and expressed as:

$$K = G \cup S,$$

where *G*, the General Knowledge base, represents the general information (e.g., previous experiences, beliefs, etc.). *S* denotes the site-specific knowledge base. It represents data collected on a natural variable at a specific site. It is divided into two: hard data for exact measurement of the variables and soft data as data with uncertainties [21, 13]. The BME analysis involves three steps, namely: the prior stage, the meta-prior stage and the integration or posterior stage [12, 22, 23]. At the prior stage, only data on general knowledge is collected and processed to build the prior *G* pdf $f_G(x_{\text{map}})$. At the posterior stage, the prior knowledge pdf $f_G(x_{\text{map}})$ is updated with the site-specific knowledge *S* following the total knowledge $K = G \cup S$. For a given Sx_{data}, the Bayes conditionalization rule is applied to $f_G(x_{\text{map}})$ to produce the posterior pdf $f_k(x_k)$ as follows:

$$f_k(x_k) = f_G(f_k/(x_{\text{data}})) = \frac{f_G(f_k, (x_{\text{data}}))}{f_G(x_{\text{data}})}. \tag{1}$$

The posterior pdf expression varies with the soft data. Once the posterior pdf is obtained, the conditional mean, mode, median, etc., can be derived. Moreover, the uncertainty can be assessed using the variance of the posterior pdf, or better still, the BME confidence interval [22, 23].

3. Background on Box-Cox Transformation

In the process of the BME application, it is noticed that the Box-Cox transformation with several lambda (λ) values is widely applied to address the issue of data normality [7, 23-26]. Box-Cox transformation [27] is a transformation method that belongs to the power transformation method family [28]. It includes all traditional methods (e.g., square root, log, inverse, cubic root) and easily leads to an optimal normalizing transformation without the need of exploring several transformation techniques to determine the best option [20]. The article [1] suggested that a method of transformation can be selected based on the relationship between the standard deviation and the mean. Box-Cox transformation offers the possibility of simultaneously correcting normality, linearity, and homoscedasticity in a dataset. The method can be expressed as:

$$g(x, \lambda) = \begin{cases} \dfrac{x^{\lambda} - 1}{\lambda}, & \text{if } \lambda \neq 0, \\ \log(x), & \text{if } \lambda = 0, \end{cases} \tag{2}$$

where y is the Box-Cox transformation of x, λ is the configuration parameter, and x is the value of any attribute in a given dataset. Thus, it increases the variance of homogeneity to improve the accuracy of prediction [28]. According to Osborne [20], λ values can correct non-normality of a variable regardless of whether it is negatively or positively skewed. Most traditional transformation methods are included in Box-Cox transformation: no transformation $(\lambda = 1)$, square root transformation $(\lambda = 0.50)$, cubic root transformation $(\lambda = 0.33)$, fourth root transformation $(\lambda = 0.25)$, natural log transformation $(\lambda = 0)$, reciprocal square root transformation $(\lambda = -0.50)$, reciprocal (inverse) transformation $(\lambda = -1.00)$ and among others. The value of the λ often varies between –5 and 5. In this study, we considered the following lambda values: –3, –2, –1 –0.5, 0, 0.5, 1, 2 and 3.

4. Methodology

The effect of data transformation on BME was evaluated by testing data with different sizes, skewnesses and levels of spatial dependency. Sample sizes considered were 100, 150, 200, 250, 300, 350, 400, 450 and 500. Symmetric data (skewness = 0) and data positively skewed by 1, 3, 6, and 9 were assessed. Data with strong (0.025), moderate (0.25-0.77), and weak (> 0.75) relative nugget effects (NE) [29], which represents the spatial dependency were used. It is denoted by S and expressed as:

$$S = \frac{c_0}{c_0 + c}, \tag{3}$$

where c is the partial sill, S is the spatial dependency level, and c_0 is the nugget. Each spatial dependency level and the ranges were repeated three times as shown in Table 1.

Table 1. Simulation parameters

Spatial dependency	Repetitions	Nugget	Psill	Range	Model
Strong	1	0.9	0.1	6	Spherical
	2	0.8	0.2	5	Spherical
	3	0.7	0.3	4	Spherical
Moderate	1	0.5	0.5	6	Spherical
	2	0.4	0.6	5	Spherical
	3	0.3	0.7	4	Spherical
Weak	1	0.2	0.8	6	Spherical
	2	0.1	0.9	5	Spherical
	3	0.05	0.95	4	Spherical

4.1. Simulation design

A spatial random field was simulated using a spherical covariance model with a sill $c_0 + c = 1$. It ranges over three spatial dependency levels presented in Table 1 [7, 30]. Two types of data were simulated, namely: hard and soft data. Hard data was simulated using the unconditional sequential simulation method (see simulib from the BMELib library, [13]). For each spatial dependency level, 50,000 independent realizations of a standard Gaussian variable were generated, using R 4.0.2 software [31]. A constant (range) was added to ensure that, the minimum value was always positive, and each value was then raised to a power (α). This exponentiation allows to generate a highly positively skewed variable. The data were then standardized to zero mean and unit variance [32].

The simulation function was based on the following arguments: Alpha (the exponentiation power to generate highly skewed variables), skewness (the variable with this skewness is returned out of the 50,000 variables simulated), sample size, nugget effect, range, partial sill and variogram model. Interval-type soft data were assumed to be randomly distributed across a square of unit size, with a width of 1.5. In addition, a function was built to generate location parameters from a shapefile of 1ha extracted from the map of Benin Republic (www.diva-gis.org) by specifying the soft data size and the hard data size.

4.2. Data analysis

Firstly, descriptive statistics (means, skewness, kurtosis, maximum and minimum) were used to summarize the raw data simulated. The datasets were transformed using Box-Cox transformation with lambda values equal –3, –2, –1, –0.5, –0.33, –0.25, 0, 0.25, 0.33, 0.5, 1, 2 and 3, respectively, to check how these methods improve the normality of raw data. Secondly, BME method was applied on each transformed data $(\text{lambda} \neq 1)$ as well as untransformed data (lambda = 1) and the accuracy of the prediction was assessed by computing the Mean Square Error (MSE) and the bias [7]. The best prediction method is expected to have MSE close to 0 [33]. Thus, transformed variables with MSE greater than 1 were excluded. Thirdly, descriptive statistics (means, skewness, kurtosis, etc) were computed on BME predictions. The average MSE and the bias were calculated for different levels of spatial dependency, degree of skewness and the transformation methods (lambda). The MSE and bias were plotted graphically using the ggplot2 package of the R 4.0.2 software [31]. Finally, Multivariate Analysis of Variance (MANOVA) was applied to assess the effect of Box-Cox transformation on BME performance. The response variables were MSE and bias and the explanatory variables were lambda, sample size and spatial dependence.

5. Results

5.1. Data transformation method

Table 2 shows that most lambda values do not improve normality in terms of skewness and kurtosis. However, regardless of the spatial dependency level, lambda values (0 to 0.5) improved the normality of highly skewed data (3 to 9).

Table 2. Skewness and kurtosis values according to lambda values of the Box-Cox transformation

Lambda	Skewness = 1		Skewness = 3		Skewness = 6		Skewness = 9	
	Sk	Kurt	Sk	Kurt	Sk	Kurt	Sk	Kurt
S1								
1	1.01	4.03	3.04	19.12	6.10	51.10	9.19	89.26
–3	–6.86	49.10	–9.85	98.01	–5.55	32.01	–9.06	86.33
–2	–6.44	44.16	–9.84	97.93	–5.48	31.26	–8.17	72.23
–1	–4.81	28.21	–9.28	90.32	–4.80	25.91	–6.10	43.85
–0.5	–2.84	12.86	–6.11	49.53	–3.42	15.90	–3.79	21.39
–0.33	–2.15	8.68	–4.26	28.99	–2.74	11.75	–2.76	14.40
–0.25	–1.86	7.13	–3.44	21.11	–2.40	9.95	–2.24	11.73
0	–1.08	4.00	–1.50	7.30	–1.27	5.87	–0.38	8.13
0.25	–0.47	2.75	–0.31	4.05	0.03	5.76	2.30	18.55
0.33	–0.30	2.61	0.01	4.07	0.53	7.05	3.35	26.12
0.5	0.05	2.59	0.67	5.30	1.78	12.78	5.60	46.80
2	2.54	9.45	7.82	70.95	9.54	93.88	9.84	97.89
3	3.36	13.83	9.44	92.46	9.83	97.72	9.85	98.01

S2								
1	1.02	3.66	3.05	16.34	6.10	50.90	8.78	84.01
–3	–9.85	98.01	–9.85	97.96	–9.82	97.68	–8.22	72.89
–2	–9.85	98.01	–9.78	97.09	–9.62	94.79	–7.31	58.39
–1	–9.84	97.89	–8.31	76.96	–7.54	65.73	–5.88	38.82
–0.5	–9.40	91.97	–4.56	31.57	–4.19	25.74	–3.98	21.38
–0.33	–8.30	77.14	–3.09	17.68	–2.98	15.09	–2.97	14.26
–0.25	–7.29	64.01	–2.49	13.03	–2.48	11.47	–2.47	11.29
0	–3.01	18.12	–1.09	5.32	–1.15	5.24	–0.91	5.65
0.25	–0.78	4.33	–0.09	3.81	0.10	4.92	1.04	9.47
0.33	–0.44	3.39	0.21	4.06	0.56	6.14	1.89	14.18
0.5	0.07	2.75	0.87	5.52	1.73	11.69	4.03	31.04
2	2.23	7.61	6.07	42.29	9.56	94.11	9.83	97.81
3	2.97	11.62	7.16	55.82	9.83	97.72	9.85	98.01
S3								
1	1.02	3.72	3.04	18.05	6.09	51.67	8.44	79.04
–3	–8.69	81.26	–9.24	89.88	–9.85	97.97	–7.14	54.14
–2	–7.40	61.63	–8.13	72.78	–9.78	97.09	–6.60	46.78
–1	–5.22	32.72	–5.61	38.15	–8.23	75.85	–4.80	28.17
–0.5	–3.31	15.78	–3.55	17.79	–4.66	31.17	–2.93	13.08
–0.33	–2.55	10.82	–2.75	12.23	–3.32	17.98	–2.23	9.01
–0.25	–2.20	8.88	–2.37	10.05	–2.78	13.54	–1.90	7.49
0	–1.24	4.71	–1.26	5.39	–1.41	5.97	–0.73	4.99
0.25	–0.49	2.94	–0.26	3.86	–0.16	5.07	1.06	9.41
0.33	–0.29	2.71	0.05	4.00	0.31	6.14	1.85	13.72
0.5	0.10	2.57	0.73	5.28	1.51	11.47	3.84	28.77
2	2.35	8.42	7.37	64.44	9.61	94.87	9.72	95.63
3	3.19	12.93	9.21	90.11	9.84	97.84	9.75	96.00

Legend: Sk: skewness; Kurt: kurtosis; S1: strong; S2: Moderate S3: weak

5.2. Effect of Box-Cox transformation on BME accuracy

Significant differences were observed in the MSE (p-value < 0.001) and bias (p-value < 0.001). This suggests that Box-Cox transformation with various lambdas induces significant changes on BME accuracy.

Table 3. Results of MANOVA illustrating the effect of data transformation on BME

Response	Sources	DF	F-value	Pr (> F)
MSE	L	14	209.041	< 0.001
	Sk	4	105.76	< 0.001
	L:Sk	35	12.10	< 0.001
Bias	L	14	181.54	< 0.001
	Sk	4	0.65	0.6245
	L:Sk	35	4.67	< 0.001

L = Lambda, Sk = skewness

Large variations of the MSE and bias were induced by the transformation methods (Table 3). Data with skewness 1 produced more points below MSE = 0.5 compared to a higher degree of skewness (greater than 3) (Figure 1). This indicates that higher degrees of skewness have a greater impact on BME than lower degrees of skewness. Regardless of the degree of skewness, untransformed data (lambda = 1) produced lower MSE and bias (Figure 1). This demonstrates that despite differences in MSE and bias due to the degree of skewness and data transformation techniques, BME applied to untransformed data is more accurate than BME applied to transformed data. To keep the plot size (Figure 1) reasonable, large values of MSE (greater than 1) were excluded.

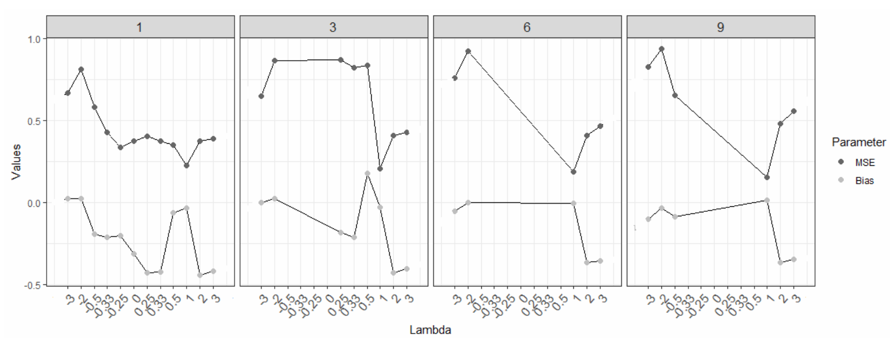

Figure 1. Variations of MSE and bias as impact of data transformation on the BME.

5.3. Effect of sample size and Box-Cox transformation technique on BME accuracy

Data transformation methods significantly affected both MSE and bias ($p < 0.001$). The sample size, however, did not significantly affect either the MSE (p-value = 0.1657) or the bias (p-value = 0.1793) of the prediction. Additionally, the interaction between sample size and transformation techniques showed no significant effect on either MSE (p-value = 0.9145) or bias (p-value = 0.981). These results suggest that BME is robust to changes in sample size.

Table 4. Results of the MANOVA illustrating the effect of sample size and data transformation on BME

Response	Source	Df	*F*-value	Pr (> *F*)
MSE	L	14	158.2379	< 0.001
	Ss	7	1.4909	0.1657
	L:Ss	96	0.8064	0.9145
Bias	L	14	169.9825	< 0.001
	Ss	7	1.4539	0.1793
	L:Ss	96	0.4966	1.0000

Ss: Sample size, L: Lambda

5.4. Effect of spatial dependency and data transformation techniques on BME accuracy

The MSE of transformation technique is not affected by the spatial dependency (p-value = 0.9454) but highly affected by the interaction between transformation technique and spatial dependence (p-value ≤ 0.001). The prediction bias is affected by the spatial dependency level with no interaction between the transformation techniques and spatial dependence.

Table 5. Results of the MANOVA illustrating the effect of spatial dependence data

Response	Source	Df	F-value	Pr (> F)
MSE	L	14	162.0512	< 0.001
	SD	2	0.0561	0.9454
	L:SD	27	2.7955	< 0.001
Bias	L	14	173.1664	< 0.001
	SD	2	3.3335	0.035
	L:SD	27	0.9115	0.5961

Figures 2 and 3 depict the differences in MSE and bias due to spatial dependence. It was observed that the variation of spatial dependence induced changes in the MSE and bias of BME prediction. In the MSE, for moderate and strong spatial dependence, most values were below 0.4. Despite this variation, the best lambda value was 1 (no transformation) irrespective of the spatial dependence.

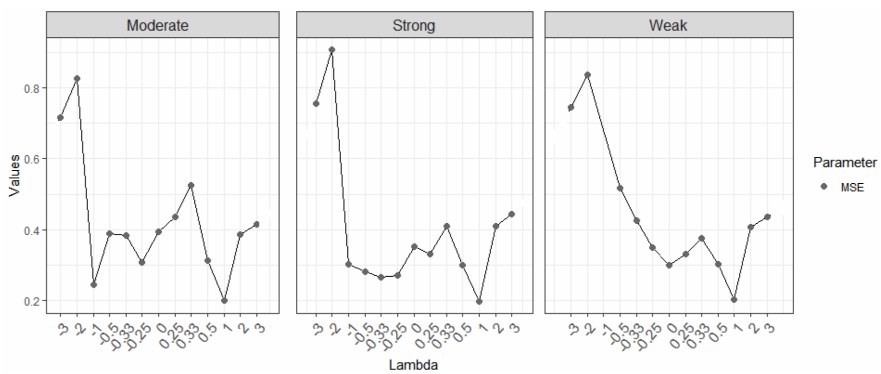

Figure 2. Variation of the prediction's MSE according to sample size and data transformation technique.

For relative bias variation by spatial dependence, the best lambda values are –2 and 1 (Figure 3). Despite the significance difference observed in Table 5, there was a lower variation in bias (–0.22 to –0.25) for spatial dependence, with moderate spatial dependence producing more accurate results.

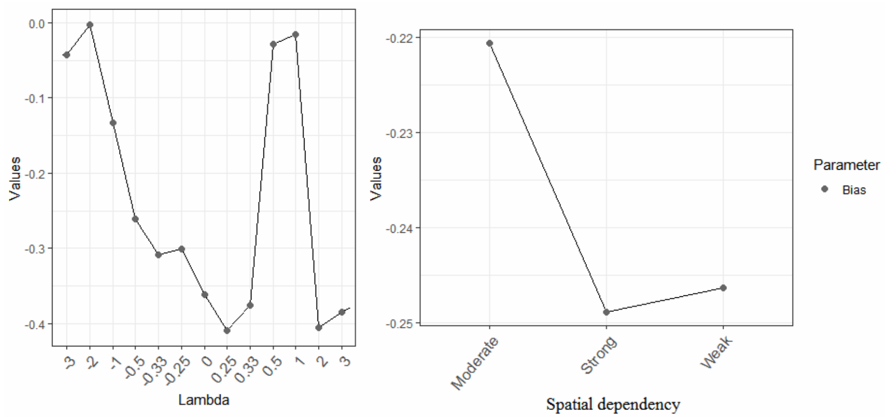

Figure 3. Variation of MSE according to the degree of spatial dependency.

6. Discussion and Conclusion

Bayesian Maximum Entropy (BME) is mostly applied to continuous variables [34-36] and discrete variables [37, 38]. In essence, a transformation is applied when data is skewed [7, 23-25]. Data transformation, as part of the pre-processing phase, plays an important role in ensuring data quality prior to data analysis [39]. In this study, an empirical assessment of the effect of data transformation techniques derived from Box-Cox family was carried out. After applying Box-Cox transformation with different lambda values to skewed data, we noticed that not all Box-Cox transformation methods do correct normality, this includes lambda values such as –3, –0.5, –0.33, –0.25, 0, 1, 2 and 3. However, lambda values such as 0.25, 0.33 and 0.5 improved the normality irrespective of the spatial dependency level. This finding was consistent with the previous works of [40] on trialeurodes vaporariorum populations in which data was transformed.

Firstly, to determine if improving the normality of a dataset can improve the prediction accuracy of BME, our results showed that the interaction between skewness and transformation techniques significantly affected both MSE and bias ($p < 0.001$ for both). This implies that BME accuracy varies with skewness and transformation techniques. However, when data transformation methods were compared by the degree of skewness, the untransformed data (lambda = 1) gave lower MSE regardless of the degree of skewness. This suggests that data transformation can improve data normality but not BME prediction accuracy. As a result, BME can be considered robust to skewness, the transformation might not be necessary when the data is skewed. These results align with Christakos's discussions on BME robustness [8, 13].

Secondly, we investigated if the accuracy of data transformation techniques applied to BME was affected by sample size. We found that the sample size has no significant effect on the MSE and bias of prediction with *p*-values of 0.166 and 0.179, respectively. Furthermore, on the MSE (*p*-value = 0.915) and bias (*p*-value = 0.981), the sample size showed no interaction with data transformation techniques. This finding suggested that the choice of sample size and data transformation had no effect on BME performance.

Finally, we evaluated if the accuracy of data transformation techniques applied to BME is influenced by the spatial dependency of the dataset. Our findings revealed that Box-Cox transformation and spatial dependence had a significant impact on the MSE (*p*-value 0.001) but not the bias of the prediction (*p*-value = 0.596). However, when we compared the performance of transformed data to untransformed data, we discovered that BME applied to untransformed variables (lambda = 1) generated better results than transformed variables regardless of the spatial dependency.

In conclusion, our study revealed that Box-Cox transformation can enhance the dataset normality, which is consistent with prior research [7, 23-25]. However, we found no indication that (1) enhancing normality improves BME prediction accuracy, (2) a given sample size improves BME prediction accuracy after transformation, or (3) any spatial dependence level is more accurate. As a result, BME can be considered robust to sample size and spatial dependency even when the data is skewed.

Acknowledgments

This research received financial support from the German Academic Exchange Service (DAAD). Funding programme FSA/UAC, 2020 (57546598).

We thank Mr. Yassine DOUDOUA, Dr. Nadjiam DJIRABAYE and the staff of Institut Tchadien de Recherche Agronomique pour le Développement (ITRAD) as well as the staff of Laboratoire de Biomathématiques et d'Estimations Forestières (LABEF) for their technical support.

References

[1] S. Manikandan, Data transformation, Journal of Pharmacology and Pharmacotherapeutics 1(2) (2010), 126. https://ncss-wpengine.netdna-ssl.com/wpcontent/themes/ncss/pdf/Procedures/NCSS/Box-Cox_Transformation.pdf.

[2] H. Y. Kim, Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis, Restorative Dentistry and Endodontics 38(1) (2013), 52-54.

[3] H. Arslan, Determination of temporal and spatial variability of groundwater irrigation quality using geostatistical techniques on the coastal aquifer of Çarşamba plain, Turkey, from 1990 to 2012. Environmental Earth Sciences 76(1) (2017), 1-12.

[4] A. Amin, B. Shah, A. M. Khattak, T. Baker and S. Anwar, Just-in-time customer churn prediction: with and without data transformation, 2018 IEEE Congress on Evolutionary Computation (CEC), IEEE, 2018, pp. 1-6.

[5] M. Serre, G. Christakos and S. Lee, Soft data space/time mapping of coarse particulate matter annual arithmetic average over the U.S., geoENV IV - Geostatistics for Environmental Applications, Springer, 2004, pp. 115-126.

[6] B. Rawlins, R. Lark, R. Webster and K. O'Donnell, The use of soil survey data to determine the magnitude and extent of historic metal deposition related to atmospheric smelter emissions across Humberside, UK, Environmental Pollution 143(3) (2006), 416-426.

[7] T. Orton and R. Lark, The Bayesian maximum entropy method for lognormal variables, Stochastic Environmental Research and Risk Assessment 23(3) (2009), 319-328.

[8] G. Christakos, A Bayesian/maximum-entropy view to the spatial estimation problem, Mathematical Geology 22(7) (1990), 763-777.

[9] G. Christakos, Random Field Models in Earth Sciences, Mineola, New York, USA, 2005.

[10] S. Gao, Z. Zhu, S. Liu, R. Jin, G. Yang and L. Tan, Estimating the spatial distribution of soil moisture based on Bayesian maximum entropy method with auxiliary data from remote sensing, International Journal of Applied Earth Observation and Geoinformation 32 (2014), 54-66.

[11] G. Christakos, Spatiotemporal information systems in soil and environmental sciences, Geoderma 85(2-3) (1998), 141-179.

[12] J. He and A. Kolovos, Bayesian maximum entropy approach and its applications: a review, Stochastic Environmental Research and Risk Assessment 32(4) (2018), 859-877.

[13] G. Christakos, P. Bogaert and M. L. Serre, Advanced Functions of Temporal GIS, Springer, 2001.

[14] G. Christakos, On the assimilation of uncertain physical knowledge bases: Bayesian and non-Bayesian techniques, Advances in Water Resources 25(8-12) (2002), 1257-1274.

[15] G. Christakos, Bayesian maximum entropy, M. Kanevski, ed., Advanced Mapping of Environmental Data: Geostatistics, Machine Learning, and Bayesian Maximum Entropy, Wiley, New York, 2008, pp. 247-306.

[16] M. Oliver and R. Webster, A tutorial guide to geostatistics: computing and modelling variograms and kriging, Catena 113 (2014), 56-69.

[17] R. Webster and M. A. Oliver, Geostatistics for Environmental Scientists, John Wiley & Sons, 2007.

[18] R. Webster and M. A. Oliver, Sample adequately to estimate variograms of soil properties, Journal of Soil Science 43(1) (1992), 177-192.

[19] J. Angulo, H.-L. Yu, A. Langousis, A. Kolovos, J. Wang, A. E. Madrid and G. Christakos, Spatiotemporal infectious disease modeling: a BME-sir approach, PloS One 8(9) (2013), e72168.

[20] J. Osborne, Improving your data transformations: applying the Box-Cox transformation, Practical Assessment, Research, and Evaluation 15(1) (2010), 12.

[21] C. L. Wang, S. B. Zhong, G. N. Yao and Q. Y. Huang, BME spatiotemporal estimation of annual precipitation and detection of drought hazard clusters using space-time scan statistics in the Yun-Gui-Guang Region, Mainland China, Journal of Applied Meteorology and Climatology 56(8) (2017), 2301-2316.

[22] M. L. Serre and G. Christakos, Modern geostatistics: computational BME analysis in the light of uncertain physical knowledge - the Equus Beds study, Stochastic Environmental Research and Risk Assessment 13(1) (1999), 1-26.

[23] A. Douaik, M. Van Meirvenne, T. Tóth and M. Serre, Space-time mapping of soil salinity using probabilistic Bayesian maximum entropy, Stochastic Environmental Research and Risk Assessment 18(4) (2004), 219-227.

[24] A. Douaik, M. Van Meirvenne and T. Tóth, Soil salinity mapping using spatiotemporal kriging and Bayesian maximum entropy with interval soft data, Geoderma 128(3-4) (2005), 234-248.

[25] Ö. Baydaroglu and K. Koçak, Spatiotemporal analysis of wind speed via the Bayesian maximum entropy approach, Environmental Earth Sciences 78(1) (2019), 1-21.

[26] C. T. Zhang and Y. Yang, Can the spatial prediction of soil organic matter be improved by incorporating multiple regression confidence intervals as soft data into BME method? Catena 178 (2019), 322-334.

[27] G. E. Box and D. R. Cox, An analysis of transformations, Journal of the Royal Statistical Society: Series B (Methodological) 26(2) (1964), 211-243.

[28] F. Zhang, I. Keivanloo and Y. Zou, Data transformation in cross-project defect prediction, Empirical Software Engineering 22(6) (2017), 3186-3218.

[29] B. B. Trangmar, R. S. Yost and G. Uehara, Application of geostatistics to spatial studies of soil properties, Advances in Agronomy 38 (1986), 45-94.

[30] R. Lark, A comparison of some robust estimators of the variogram for use in soil survey, European Journal of Soil Science 51(1) (2000a), 137-157.

[31] R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2020. URL https://www.R-project.org/.

[32] R. M. Lark, Estimating variograms of soil properties by the method-of-moments and maximum likelihood, European Journal of Soil Science 51(4) (2000b), 717-728.

[33] Z. Jia, S. Zhou, Q. Su, H. Yi and J. Wang, Comparison study on the estimation of the spatial distribution of regional soil metal(loid)s pollution based on kriging interpolation and BP neural network, International Journal of Environmental Research and Public Health 15(1) (2018), 34.

[34] E. Savelieva, V. Demyanov, M. Kanevski, M. Serre and G. Christakos, BME based uncertainty assessment of the Chernobyl fallout, Geoderma 128(3-4) (2005), 312-324.

[35] L. Sanders, A. Daly and K. Regan, Beginning the uncertain journey: Foundation students' expectations and experience, HEA Annual Conference, 2012.

[36] P. Jat and M. L. Serre, Bayesian maximum entropy space/time estimation of surface water chloride in Maryland using river distances, Environmental Pollution 219 (2016), 1148-1155.

[37] L. Han, C. Wang, Q. Liu, G. Wang, T. Yu, X. Gu and Y. Zhang, Soil moisture mapping based on multi-source fusion of optical, near-infrared, thermal infrared, and digital elevation model data via the Bayesian maximum entropy framework, Remote Sensing 12(23) (2020), 3916.

[38] L. Fan, Q. Xiao, J. Wen, Q. Liu, R. Jin, D. You and X. Li, Mapping high resolution soil moisture over heterogeneous cropland using multi-resource remote sensing and ground observations, Remote Sensing 7(10) (2015), 13273-13297.

[39] A. Famili, W.-M. Shen, R. Weber and E. Simoudis, Data preprocessing and intelligent data analysis, Intelligent Data Analysis 1(1) (1997), 3-23.

[40] J. J. Park, Key-Il Shin and K. Cho, Evaluation of data transformations and validation of a spatial model for spatial dependency of Trialeurodes vaporariorum populations in a cherry tomato greenhouse, Journal of Asia-Pacific Entomology 7(3) (2004), 289-295.