



HAL
open science

Interpolation Method of Soil Moisture Data Based on BMA

Wan Shu-Jing, Zhang Cheng-Ming, Liu Ji-Ping, Yu Ting, Ma Jing

► **To cite this version:**

Wan Shu-Jing, Zhang Cheng-Ming, Liu Ji-Ping, Yu Ting, Ma Jing. Interpolation Method of Soil Moisture Data Based on BMA. 8th International Conference on Computer and Computing Technologies in Agriculture (CCTA), Sep 2014, Beijing, China. pp.480-488, 10.1007/978-3-319-19620-6_54 . hal-01420263

HAL Id: hal-01420263

<https://inria.hal.science/hal-01420263>

Submitted on 20 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Interpolation method of soil moisture data based on BMA

WAN Shu-jing¹, ZHANG Cheng-ming^{1,*}, LIU Ji-ping², YU Ting¹, MA Jing¹

1. College of Information Science and Engineering, Shandong Agricultural University, Tai'an, 271018, China

2. Research Center of Government Geographic Information System, Chinese Academy of Surveying & Mapping, Beijing, 100830, China

Abstract: How to consider both the spatial distribution and the time series characters of the soil moisture data, which have an effect on the result of interpolation, is the key to improve the soil moisture data interpolation result. This paper proposes a new interpolation method: the BMA-I (BMA-Interpolation), which contains kriging method, and BMA (Bayes Model Average) method. Spatial forecasting model uses synergetic kriging method. Time series forecasting model uses LS-SVM algorithm. The model average uses BMA method. Taking the project of Shandong Bohai granary as an example, through comparing with kriging method, the proposed approach can improve the problems existing in the kriging method effectively, and can give a reliable forecast uncertainty interval, the simulation results more accurate and reasonable.

Keywords: Soil moisture, the BMA -I method, Interpolation.

INTRODUCTION

Water is the most important resource in the earth, which is the basic element of life. About 97.2 percent water is in the sea, 2.15 percent water is with ice, while the other 0.63 percent water is in the ground, and 0.005 percent water is in the soil all over the world^[1]. Very small portion though the soil moisture occupies in the global water resource, it is a critical factor in the process of energy transition between the earth and the atmosphere^[2]. Water moisture is an important part of land ecosystem water cycle, it has important significance for the study of water stress, drought monitoring and crop yield estimation. Soil moisture data of region or even the world, is an essential parameter in the study of land process model, in improving the regional and global climate, in forecasting the area humidity. How to obtain high quality of the soil moisture data, has always been a focus of the study^[3-5].

The development of sensor technology and communication technology have greatly improved the ability of the ground observation. People can use sensor networks to acquire soil moisture, land surface temperature, wind speed and other related parameters directly and real-time. And the time resolution is very high, even it can up to second order of magnitude. But the observation data on behalf of the space is limited. Therefore, people often use the interpolation method which can transform limited observation data to surface data. Kriging method is a common interpolation method which is often adopted. Kriging method is demanded to satisfy the Second-order stationary hypothesis. Because the soil moisture is a surface variable, which is a kind of high spatial and temporal heterogeneity, have a smooth effect on soil humidity interpolation, even ordinary kriging method, the kriging method, or collaborative kriging method. In other words, the smaller values are often overstated, and larger values are often underestimated. Estimate values can't reflect the real change characteristics of the real space. In order to solve this problem, researchers at home and abroad have made a wide range of research. Journal in Yao^[7] establishing a series of post-processing algorithms called kriging method, which is under the condition of sacrifice local estimation precision, this method processing the smoothing effect problem well. Olea^[7], such as a combination of traditional kriging estimation and the characteristics of the random conditional simulation, puts forward the compensation kriging method, but the results of this method on the global optimality is better than random simulation, on the local accuracy is lower than the traditional kriging estimation. Yamamoto^[8-10] in combination with the idea of Olea, puts forward a set of post-processing on the estimation of kriging method, Yang Yuting^[11] uses his method of soil moisture interpolation and gets post-processing results.

As you can see, in the previous interpolation of soil moisture research, researchers mainly focus on the influence of spatial variation characteristics of interpolation, and the time change characteristics of soil moisture is almost under no consideration. Related studies have shown that besides exclusion of precipitation and irrigation, the change of soil moisture have a relationship with temperature, illumination, wind speed, vegetation and other factors. And these factors can have some regularity changes on the time series. If considering both the space and time effects on the result of the interpolation, we can get a higher accuracy of the interpolated result.

Bayesian model averaging (Bayesian model averaging, BMA) is put forward by Rafer yetc^[11]. Bayesian model averaging uses a statistical processing method which uses multimodal set to make probability predictions. A particular variable probability density distribution function (PDF) predicted by BMA, is average weight which is a probability distribution of every single model forecasting. The PDF is possessed by a deviation correction. The average weight is a posteriori probability of corresponding model and it represents the relative prediction techniques of each model in the stage of model's training. Many researches have been carried out that the method of BMA shows great advantage^[13-16].

This paper chooses the collaborative kriging method as spatial prediction model, and uses the time-series forecasting model as sequential forecasting model, based on the LS - SVM algorithm. This paper also applies the basic principle of the BMA, considering the soil moisture data of space and time changes at the same time. Establishing a kind of the interpolated

method considering time and spatial variational characteristics simultaneously.

1. BMA-I METHODS

Assuming that region Ω to be interpolated can be regarded as composed of M grids, each grid contains up to an observation point, namely $\Omega = \{p_i\}, i = 1, 2, \dots, M$. Observation points set $V = \{v_1, v_2, \dots, v_i, \dots\}, v_i \in \Omega$. Collection of observation points were divided into .and the three V_1, V_2, V_3 mutually collections are disjoint. V_1 is used as the input of kriging interpolation, V_2 is used as the input of forecasting model, V_3 is used as input validation.

The y is on behalf of forecast variables, namely soil moisture value of the grid. $F = \{f_1, f_2\}$ present a set of predicting model, f_1 is spatial interpolated model, f_2 is time-series interpolated model.

According to the principle of the BMA, the probability density function of y is

$$\rho(y|V_2) = \sum_{i=1}^2 w_i g_i \quad (1)$$

For under the condition of given data V_2 , F_i is the optimal probability of the model, g_i is the probability density function of y , where V_2 at a given sample and model forecast variables F_i under the condition. Supposing that smaller regional of underlying surface is nearly the same, so the distribution of soil moisture in the area can be seen as a normal distribution. At this time, for every g_i , the mean or expect meet:

$$y|g_i, V_2 \sim N(\alpha_i g_i + \beta_i, \sigma_i^2) \quad (2)$$

α_i, β_i are deviation corrections

The expectations for (1) :

$$E(y|g_1, g_2) = \sum_{i=1}^2 w_i (\alpha_i g_i + \beta_i) \quad (3)$$

In order to have $w_i, \alpha_i, \beta_i, \sigma_i^2$ the variance is :

$$\text{VAR}(y|V_2) = \sum_{i=1}^n w_i ((\alpha_i v_i + \beta_i) - \sum_{i=1}^n a_i v_i + \beta_i)^2 + \sum_{j=1}^2 w_j \sigma_j^2 \quad (4)$$

Taking the minimum value as constraint conditions, and using EM algorithm to find the solution. According to these formulas, we can obtain the parameters above, and then determine the value of each grid.

2. EXPERIMENT

2.1. Study region

Experimental area is in Shandong "Bohai granary demonstrative project of science and technology" in the project, the area contains Dezhou, Binzhou and Dongying district, 115°45'-119°10' east longitude, 36°24'-38°10' north latitude, a total of 30 counties (city or area), and grain land is 15 million mu, the cotton fields transformed to grain is 1 million mu, the land of sanlinization is 1 million mu. Are shown in figure 1 in the blue box. The layout of ground sensor acquisition node is shown in figure 2.



Fig.1. Experimental area location



Fig.2. Acquisition node distribution

2.2 Study data

(1) The ground observation data

The ground observation data is obtained through the observation network, every node obtain soil moisture, wind speed, temperature and so on. We can obtain a data every 5 minutes. Data is on May 2, 2014 to May 28, 2014.

(2) The historical data of soil moisture

The historical data of soil moisture is being obtained by the writers' method developed by themselves. This method is a kind of inversion method. This method is using passive and positive remote sensing, based on genetic neural network algorithm. First of all, establish a BP neural network, and then use the genetic algorithm of BP network to optimize the node weight value; Lastly, the TM data (TM3, TM4, TM6), different polarized and polarized ratio (VV, VH, VH/VV) ASAR data as the input of neural network. At the same time, the surface soil moisture content is the output of the network, with some measured data training and inversion on the network. The paper of this method have been received by Sensor Letter, details about this paper can be queried if you have interest.

2.3. Experiment Design

(1) observation point data preprocessing

According to the need of the method, the observation point of every area can be divided into V1, V2, V3 three mutually disjoint collections, including V1 used as the input of kriging interpolation, V2 as input of forecast model, the V3 as input validation.

(2) soil moisture history data preprocessing

Soil moisture history data is preprocessed by ENVI software. Main work includes: registration of the vector data and raster data; To determine the grid position of each observation point; Extraction with interpolation grid soil humidity value and arrange in chronological order.

(3) interpolated model and time-series forecasting model

Interpolation model using collaborative kriging method, the method is improved to statistical treatment in the same space domain development also has a statistical correlation and spatial correlation of the ability of multiple variables, is a kind of multivariate statistics of the basic method, can also consider the impact of soil moisture and multiple factors, has the strong ability for data processing.

Time-series model using Wang Yongsheng [15] put forward the based on the LS - SVM algorithm of temporal recursive forecasting method, the method used is equality constraint instead of a classic inequality constraints, the dual problem is a system of linear equations, the solution of the problem into solving linear equations, avoid solving of quadratic programming problem.

(4) the forecast model

Prediction model using EM algorithm and the EM is a kind of iterative algorithm for solving maximum likelihood estimation, the iterative process can be divided into the following two steps: step 1 is called "E", namely according to the result of the previous step to estimate the complete data set expectations of likelihood function; Step 2 is called a "M", namely the estimates made completely data set the parameters of the maximum likelihood function expectations; Then repeat these two steps until the iterative convergence, details refer to relevant literature.

2.4 Experiment and results

In order to compare with the result of the experiment, We have used Ordinary Kriging method (Ordinary Kriging, OK) collaborating Kriging method (Cooperate Kriging, CK) and BMA_Interpose method (BMA - I) to interpolate, and then compared the results. We have chosen the data respectively on 9:00 am at May 3, 9:00 am at May 10, 9:00 am at May 21, 2014 to calculate. To get the interpolation result, 2400 grids were selected as the interpolation points. The results were shown in figure 3 (1) - (9).

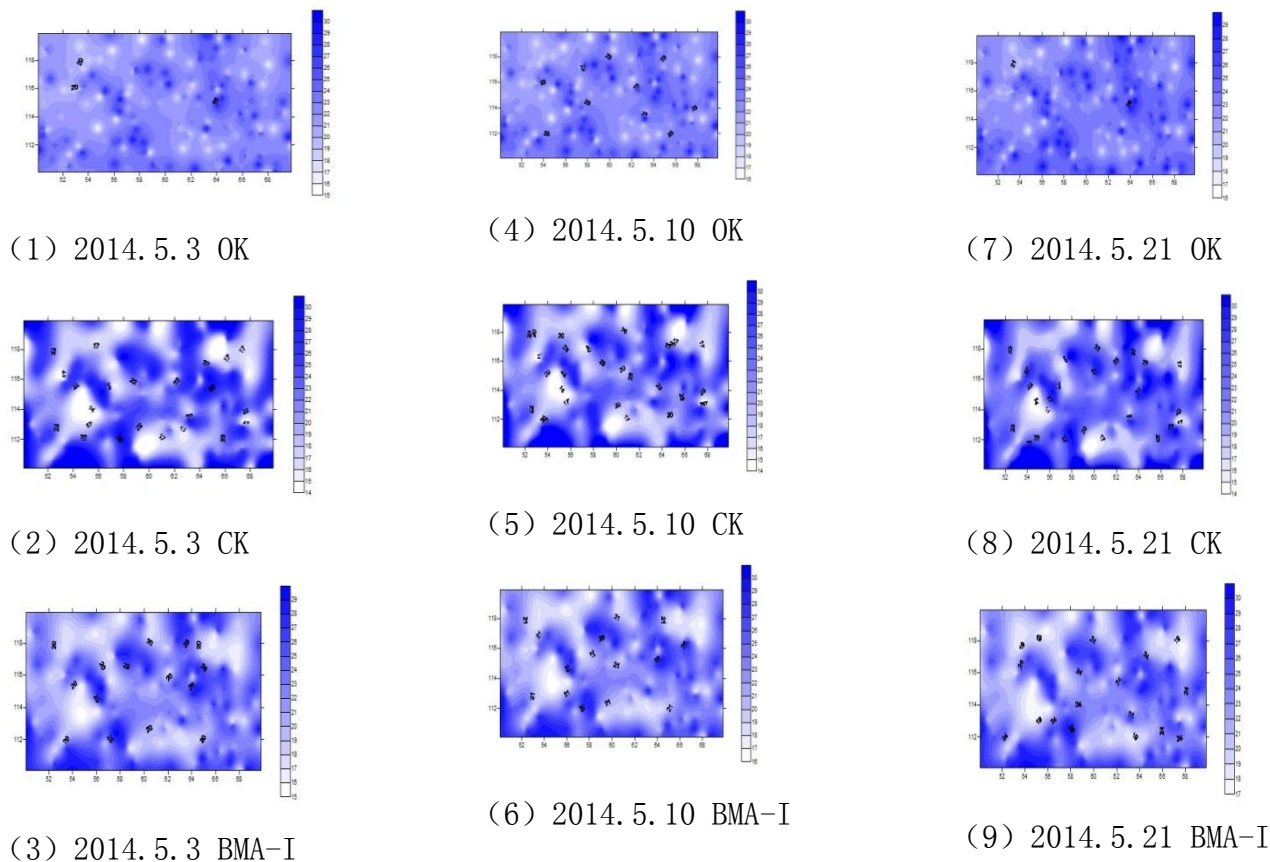


fig.3 .The interpolation result

Table 1 shows the statistical results of the three interpolation methods. By comparing the measured soil moisture and the soil moisture after two methods of kriging interpolation, we can see from the statistical results as follows:

- (1) The measured soil moisture and the mean of two kriging estimation of soil moisture are almost equal, It shows that the kriging estimation conform to the unbiasedness requirements;
- (2) The standard deviation of the measured soil moisture obviously larger than the standard deviation of kriging estimation of soil moisture, this reflects that by using ordinary kriging method for spatial interpolation, produced obvious smooth effect, leading to the estimates of the variation degree is less than actual situation, the decline sharply of range can also explain the existence of the smoothing effect. Compared to the two kinds of kriging method we can found that both standard deviation indicators and The range index , collaborative kriging method is better than the ordinary kriging method. It shows that the result which only use the ordinary kriging method to interpolate can't reflect the real space distribution characteristic of soil moisture, but it will be improved clearly after introduce correlated Variables.
- (3) On the frequency distribution of soil moisture index, the performance of the two methods of kriging has a obvious difference. It shows us that using ordinary kriging method for spatial interpolation of soil moisture can not reflect the real soil moisture spatial distribution features, but after the introduction of relevant variables, it will be improved according to the result of the interpolation.

Table.1.The statistical results of three methods

Name	Grid numbers	means	Standard deviation	range	deviation degree	max	min
Survey sample	178	21.37	2.471	11.39	0.183	29.40	16.60
OK	2400	21.32	1.328	9.41	-0.217	26.91	15.34
CK	2400	21.39	1.711	10.72	0.261	28.12	17.1
BMAI	2400	21.35	2.317	11.17	0.117	29.27	16.93

By comparing the measured soil moisture and statistical features of the BMA-I method it shows as follows:

(1) the soil moisture on the average value between the two ways is very close, indicating that the BMA -i method still keep the unbiased estimation characteristics of the kriging method;

(2) the BMA - I method on the standard deviation index compared with the two kinds of kriging method, has a big improve, and get very closed to the measured soil moisture in standard deviation, the transformation of the range also performs the same tendency.

(3) in the skewness, both of them showed slight positive skewness. and other statistics also approximated.

In addition, using the OK method for interpolation, 107 sample points scatter outside the scope of the measured sample, close to 5%, but CK method and BAMI method don't have this kind of situation. On the scope of data, the data range of CK method is obviously smaller than the measured sample data range, while the BMA -i method in getting data range and numerical range is closer to the sample. In order to illustrate better that the BMA -i method can reflect the soil moisture in the spatial distribution characteristics is improved, we have made the corresponding soil moisture frequency histogram, as shown in figure 4.

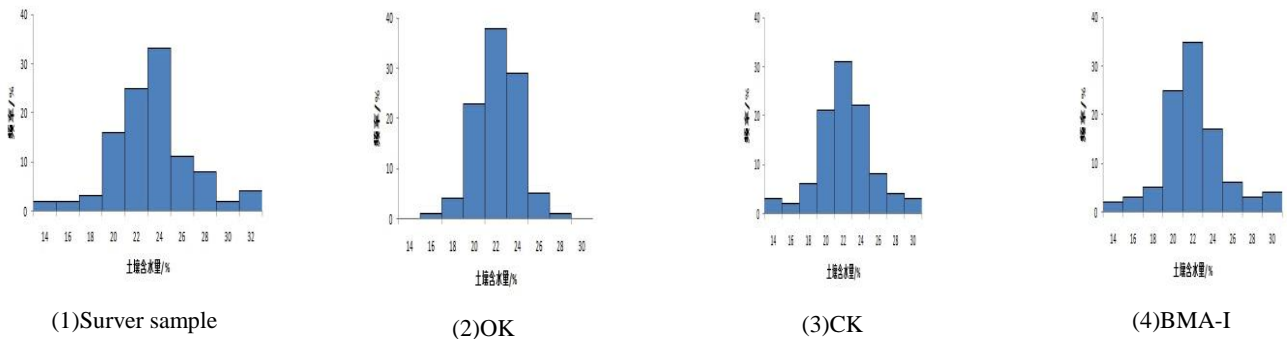


Fig. 4.soil humidity frequency distribution histogram

As can be seen from the figure 4, the ordinary kriging method to estimate the frequency distribution of the histogram is limited in a narrow range, the frequency distribution of the measured values and the histogram obviously has great discrepancy; Collaborative kriging method of frequency histogram has a significant improvement, but still with the measured values have some discrepancy; BMA-I estimate that the results with the measured frequency distribution histogram and are closer.

Through the above analysis, it shows that BMA-I method to obtain the estimate of basic can reflect the actual spatial distribution of the field soil moisture content.

3.CONCLUSION AND DISCUSSIONS

In this paper, we were using the measured soil moisture data, combined with the soil moisture remote sensing inversion historical time-series data, implements a space distribution and time distribution of give attention to two or more things interpolation method. From the point of processing result, due to the introduction of the time-series data, BMA-i method significantly improve the smooth effect brought by the kriging method, spatial variation characteristics of soil moisture also got reappear. Due to the expectations of the BMA -i method is members of the expected average Bayesian method,

which can not only improve the prediction accuracy to a certain extent, and, more importantly, through further processing, this method can also provide reliable predicted uncertain interval.

In temporal changes, considering that irrigation and rainfall has a great influence to the soil moisture variation characteristics, we only use the data weeded out 5 days before and after irrigation and rainfall, the processing methods affect temporal feature extraction. In addition, this article only use small scale data validation and application. To the soil moisture data of different scale, we need to test the effectiveness of the proposed method more.

Acknowledgment: This study has been funded by National Natural Science Foundation of China (41101321), National High Technology Research and Development Program of China (863 Program) (2013AA122003), Foundation for Outstanding Young Scientist in Shandong Province (BS2011DX031), Science and Technology Develop Project in Shandong Province (2012GSF11713), National Special Fund (2012DFA60830) and Funded by Key Laboratory of Geo-informatics of State Bureau of Surveying and Mapping (201319)

References and Notes

- [1] YANG Tao, GONG Huili, LI Xiaojuan, et al. Progress of soil moisture monitoring by remote sensing[J]. *Acta Ecologica Sinica*, 2010, 30(22):6264-6277.
- [2] Han Nianlong. Retrieval of Bare Soil Moisture Using AMSR-E Data[D]. Jilin University, 2007.
- [3] Wang J R, Engman E T, Shi J C, et al. The SIR-B observations of microwave backscatter dependence on soil moisture, surface roughness and vegetation covers [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1996, 24: 510-516.
- [4] Gao Tingting. Study on Soil Moisture Inversion of Bare Random Surface Based on IEM model[D]. Xinjiang University, 2010.
- [5] Li Jinxiang. Soil Moisture Retrieval and Its Spatial Character Analysis in Bare Random Surface[D]. Xinjiang University, 2011
- [6] OLEA R, PAWLOW SKY V. Compensating for estimation smoothing in Kriging [J]. *Mathematical Geology*, 1996, 28 (4) : 407-417.
- [7] GOOVAERTS P. Accounting for estimation optimality criteria in simulated annealing [J]. *Mathematical Geology*, 1998, 30 (5) : 511-534.
- [8] YAMAMOTO J K. An alternative measure of reliability of ordinary kriging estimates [J]. *Mathematical Geology*, 2000, 32 (4) : 489-509.
- [9] YAMAMOTO J K. On unbiased back transform of lognormal kriging estimates [J]. *Computational Geoscience*, 2007, 11: 219-234.
- [10] YANG Yuting, SHANG Songhao, LI Chao. Correcting the smoothing effect of ordinary Kriging estimation soil moisture interpolation. *ADVANCES IN WATER SCIENCE*, 2010, 21(2):208-214.
- [11] WANG Yongsheng, OUYANG Zhonghui, WANG Jianguo, et al. Predict the parameter varying chaotic time series based on LS-SVM. *Computer Engineering and Applications*, 2009, 45(26):114-117.
- [12] Raftery A E, Gneiting T, Balabdaoui F, et al. Using Bayesian model averaging to calibrate forecast ensembles. *Mon Weather Rev*, 2005, 133: 1155-1174
- [13] Ajami NK, Duan Q, Sorooshian S. An integrated hydrologic Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. *Water Resour Res*, 2007, 43: W01403, doi: 10.1029/2005WR004745
- [14] Neuman S P. Maximum likelihood Bayesian averaging of uncertain model predictions. *Stoch Environ Res Risk Assess*, 2003, 17: 291-305, doi: 10.1007/800477-003-0151-7
- [15] Raftery A E, Madigan D, Hoeting J A. Bayesian model averaging for linear regression models. *J Am Stat Assoc*, 1997, 92: 179-191
- [16] Slougher J M, Raftery A E, Gneiting T. Probabilistic quantitative precipitation forecasting using Bayesian model averaging. University of Washington, Department of Statistics, Technical Report 496, 2006