Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2016-401-AC2, 2016

Hydrology and Earth System Sciences Discussions

Open Access

HESSD

Interactive comment

Interactive comment on "An Extended Kriging method to interpolate soil moisture data measured by wireless sensor network" *by* Jialin Zhang et al.

Jialin Zhang et al.

toliuqiang@bnu.edu.cn

Received and published: 29 November 2016

General comments: This interesting study published in HESSD proposes a technique for the spatial interpolation of soil moisture measurements obtained by wireless sensor networks (WSN). Remote sensing data of NDVI and albedo are used as an additional information in the novel approach referred as Extended Kriging. The acronym NDVI is not clearly defined in the text and it is assumed that it refers to the vegetation index abbreviated by VI in line 17 of page 1. The interpolation technique is based on transferring the standard spatial assumptions of Ordinary Kriging to a combination of spatial distance and additional information related assumptions. The results are presented in a good and mostly comprehensible form and the structure of the paper is reasonable. However, the manuscript contains some spelling and grammar errors and the formulations could be more concise at some points. The figure captions could gener-

Printer-friendly version

Discussion paper

ally contain more information. Furthermore, it is not entirely clear how the interpolation techniques were applied. In total, there are several points that need further explanation and require additional work. Due to this, I recommend the paper to be returned for major revision.

ANS:

The anonymous Referee#1 gave us many constructive suggestions and comments. Following his suggestions, the authors have made major revision to the manuscript: besides the changes in the text, a new section (Section 4) is inserted into the manuscript to present the comparison results between Cokriging and the proposed Extended Kriging algorithm, as well as the analysis about the effect of number of observation points to be interpolated. After the revision, the motivation for proposing the new algorithm, instead of using Cokriging or Kriging with External Drift, is explained more clearly. We are very grateful for the referee's help.

We have posted a quick short response to Referee#1 before. But it is not enough. Now we will formally answer Referee#1's questions one by one.

Specific comments:

Specific comments (1): Data pre-processing is often crucial for the interpolation performance of geostatistics. The information given in Sec. 2.2 is not sufficient. It should be explained how the exclusion of abnormal WSN data was performed.

ANS:

Data pre-processing is absolutely necessary. A short description of data pre-processing is added into section 2.2. As we know, the WATERNET is an experimental WSN established for the campaign in summer 2012. Some of the sensor nodes are affected by sensor noise and abnormal conditions during wireless data transfer. Smoothing and noise reduction treatments are necessary. Besides, when the soil moisture is close to saturation, the soil moisture sensors cannot work properly and sometimes give abnormal values. Thus, this part of data also need be removed. We first excluded the abnormal data by assigning the zero value, negative value and abnormally high value (soil moisture content > 50 percent) as invalid data NaN. Then, we averaged the valid data for the whole day, and used the average result as the final soil moisture value for each node.

Specific comments (2): It is not clear why vegetation index in combination with albedo was selected as additional information. This needs to be discussed more. A short correlation analysis for several variables might help to justify this choice.

ANS:

We are sorry that we neglected to explain the acronym NDVI (Normalized Difference Vegetation Index) in the old manuscript. We have added the definition of NDVI in Section 2.2 where it appeared for the first time.

Selecting NDVI and albedo as auxiliary information are out of the following two main considerations:

(a) NDVI and albedo are the spectral indexes which are fairly easy to be obtained from almost all high resolution remote sensing data sources. Although some other spectral indexes, such as TVID or NDWI, may have better ability to indicate soil moisture, they will require remote sensing data from thermal infrared or shortwave infrared channels, which usually are not available for high resolution satellite remote sensors. In order to develop a practical algorithm, it is imperative for us to choose the indexes from the practical data sources.

(b) NDVI and albedo represent most of the remote sensing information in the visible and near infrared spectral range. The raw remote sensing data are expressed as radiance (or reflectance) of multiple spectral channels. Because of the correlation between channels, the raw remote sensing data are redundant. In order to efficiently use the information in remote sensing data, principle component analysis (PCA) can be adopted to extract the most informative component. Usually, the first and second principle component from most images will be the brightness and greenness, which approximately correspond to albedo and NDVI.

Although our reason of selecting NDVI and albedo as spectral variables is not based correlation analysis, it is still necessary to discuss the correlation of NDVI and albedo to soil moisture. So, we have added this correlation analysis into Section 3.2 of the revised manuscript. From the result of correlation analysis (see the revised manuscript), it is concluded that there is no significant LINEAR correlation between soil moisture and NDVI, or between soil moisture and albedo, the correlation coefficients are below 0.5. And this low correlation explains why the Cokriging algorithms cannot achieve satisfactory interpolation of soil moisture.

Specific comments (3): The equations of Ordinary Kriging (Eq. 1, 2, 3 and 4) are not explained entirely. For many readers it might be clear that n refers to the number of adjacent measurements taken into account for the spatial estimation, nevertheless it should be mentioned somewhere in the text. The actual number of points considered for the interpolations using Ordinary Kriging and Extended Kriging should be mentioned as well in the methodology.

ANS:

We have appended the explanation for the equations of ordinary Kriging in the revised manuscript.

As to the actual number of points considered for the interpolations, we plot the number of points for each day in Fig. 1 in this reply. The reason for the fluctuation of actually used numbers is the occasionally malfunctions of WSN nodes. These invalid observations are detected in the pre-processing and excluded in the interpolation. As the fluctuation is not prominent, we think it unnecessary to present this table in the paper. However, the range of the number of points is mentioned in the last paragraph of Section 5.1.

Specific comments (4): It is not entirely clear how the variogram fitting was conducted and whether an automatic or manual approach was used for this. What exactly is shown in Figs. 2, 3, 4 and 5? Are these average experimental semivariograms or the experimental semivariograms of a specific time step? Figure 2 can be omitted.

ANS:

As the number of WSN nodes installed in study area is limited, the node amount is insufficient to gather robust statistics in the process of semivariogram fitting. Thus, we used a soil moisture map of the study area to obtain sampling data. The soil moisture map was derived from the airborne hyperspectral datasets of CASI/TASI, acquired on July 10, 2012.

In the calculation, we sampled 9000 points from the soil moisture map as the sampling data for semivariogram fitting. To reduce the error, the spatial/spectral distance was divided into ten bins. We averaged the semivariance values of each bin as the final semivariance value of each distance. As the amount of the sampling data was insufficient when the spatial/ spectral distance was increased up to a certain extent, its average semivariance value was not stable. Therefore, these invalid data were removed before semivariogram fitting. For example, we only used the data when $h<=4000$ m in the fitting process in Fig. 3. Figure 4 shows the valid average semivariance data after manual exclusion. The color of each grid in Figs. 3 and 4 presents its average semivariance value.

In the process of semivariogram fitting, the nugget value and partial sill value were calculated by automatic approach, and the range distance was set by manual approach. As the sampling data were acquired from the soil moisture map mentioned above, the experimental semivariograms correspond to the time of the soil moisture map. In this manuscript, we assume that the derived semivariograms can be applied to all the dates of interpolation.

Figure 2 in the original manuscript has been removed in the revised manuscript.

Specific comments (5): It is not sufficiently described how the interpolation was performed. The 5 min WSN measurements were aggregated to daily estimates of soil moisture, but only five clearsky satellite images were available for specific days. Did you apply the ordinary kriging and Extended Kriging interpolation only for these five time steps? The performance (Figs. 10 and 11) show more than five sampling points. Theoretically, all days of the investigation period need to be interpolated. Is it possible to apply this method for days without satellite information, for instance by using averaged spectral variables? This might be particularly interesting for the implementation of temperature data.

ANS:

As only five clear-sky satellite images were available in the time range of interpolation, we assumed that the spectral variables derived from satellite images would not change much within a few days. This assumption is generally valid in the visible and near infrared remote sensing. In this way, for days without satellite images, we applied the available satellite image with nearest date to their interpolation. Thus, we could obtain the interpolation results in continuous time series.

In the case of temperature data, however, the problem is different. Because land surface temperature is a quick changing variable, it is not appropriate to the temperature data acquired. So, we think that temperature should not be applied to neighboring days as NDVI and albedo.

Specific comments (6): I suppose that the correlation discussed in section 4.2 refers to the correlation of soil moisture maps interpolated by Extended Kriging with precipitation and irrigation data. Is it possible to show also the correlation using Ordinary Kriging? It might be interesting to see whether the implementation of satellite information can improve it.

ANS:

The original Fig.9 has been revised as Fig.12 in the new manuscript. The interpolation results of ordinary Kriging are added into the figure. As is shown from the curves, the difference of these two methods is not prominent. This can be explained from two aspects: 1) The advantage of the Extended Kriging is to present more details of soil moisture spatial distribution. But averaging the interpolation result to the scale of field blurs the spatial details. 2) The main feature of the Extended Kriging is that it uses remote sensing image to aid spatial interpolation. But Fig.12 demonstrates the temporal variation of soil moisture and the information of temporal variation comes from the WSN data instead of the remote sensing image. So, Fig.12 actually is not suitable to demonstrate the main feature of the Extended Kriging.

The main purpose of the analysis in Sec. 4.2 is to prove that, by inheriting the temporal continuity of WSN observations, the result of Extended Kriging can reflect the changes of soil moisture on time series. So, the soil moisture derived from this approach is continuous both in spatial and temporal dimensions.

Specific comments (7): I strongly recommend comparing the performance of Extended Kriging with the performance of a standard multivariate geostatistical technique, for instance Cokriging or Kriging with External Drift. It is true that Extended Kriging is somewhat simpler. Nevertheless, it would be useful to achieve a better indication of interpolation performance. The second objection stated in the lines 6 to 8 on page 12 is not valid. Multivariate geostatistics is often applied to data without a direct physical relation.

ANS:

One of the reasons to design the Extended Kriging algorithm is that it is simpler to operate than other multivariate geostatistical techniques. So, Extended Kriging can be applied in cases where the pre-conditions of Cokriging or Kriging with External Drift are not satisfied. For example, when we first tried to interpolate the soil moisture measurements from WSN with the Cokriging software package in R language, the interpolation result did not show enough spatial details, possibly because that the correlation of soil moisture with the spectral indexes are weak. Then, we tried the Extended Kriging algorithm, and the result looks better. These comparison results have been added into the revised manuscript as Fig.8 and Table 1 in Section 4.

In this analysis, we use the remote sensing retrieved soil moisture map to extract sampling points and validation points. Firstly, we sampled 9000 points from the soil moisture map to calculate the semivariogram of Cokriging method. Then we used different numbers of points, ranging from 300 points to 30 points, to interpolate soil moisture. The numbers of sampling points and RMSEs of interpolation results are shown in Table 1 of the revised manuscript. All the points of the soil moisture inversion map were used as validation points to calculate RMSE for each interpolation. As the locations of sampling points will influence the interpolation result and accuracy, we sampled randomly and repeated enough times to decrease the disturbing of point locations, and calculate the average RMSE value. From Table 1, we can see that the interpolation uncertainty (indicated both by σ_κ and RMSE) decreases while number of sampling points increases, and the estimator σ_κ can reflect the variation trend of the actual RMSE. We also find out that the RMSE of Extended Kriging and Cokriging are close, except that Cokriging performs a little better when the number of sample points is less than 50.

However, when we made a visual inspection to the interpolated soil moisture map (Fig.8 of the revised manuscript), it looks like that the interpolation of Extended Kriging can present more detailed information for the spatial distribution of soil moisture than that of Cokriging. These test results indicate that in cases where the sophisticated multivariate geostatistical techniques are not applicable, it may be worthy to try the Extended Krigning algorithm proposed in this manuscript.

The method of Kriging with External Drift (KED) is based on the assumption that the target variable (soil moisture) can be predicted from usually a linear function of the covariates. From our analysis in Section 3.2, such a prediction model does not exist, at least on the whole experiment area. If local prediction model should be considered, the KED algorithm could be very complicated; and we are not sure we can fairly present the best result of KED. So, we do not include comparison with KED in the revised manuscript.

Minor technical corrections:

Minor technical corrections (1): What is the reason for the tilted perspective or the masking of the borders in the maps of Figs. 6 and 7? Simple two-dimensional plots might be a better solution. I recommend preparing Fig. 1 in a consistent way, i.e. use the same masking. What is the background colour shading in Fig. 1?

ANS:

The borders in maps of Figs. 6 and 7 of the old manuscript are border of the foci experiment field in the HiWater campaing. The reasons for masking the interpolation results with this border are as follows: (1) WSN nodes in the study area distributed within this border; (2) there are no land surface temperature data outside this border. Thus, we finally used the same masking for all the interpolation results.

The image shown in Fig. 1 was obtained from HJ satellite, combined by band 3(RED), band 4(NIR) and band2(GREEN). It has been masked with the same border in the revised manuscript.

Minor technical corrections (2): The term uncertainty analysis refers usually only to evaluations regarding the kriging standard deviation. I suggest renaming section 4.3 to cross validation and uncertainty analysis.

ANS:

Thanks for pointing out the concept differences. We have renamed this section (5.3 in the revised manuscript) title.

Minor technical corrections (3): Page 3, line 30 and other occurrences: The correct

Printer-friendly version

Discussion paper

spelling is cokriging.

ANS:

We have corrected this spelling mistake in the whole paper as the Cokriging.

Minor technical corrections (4): Page 7, lines 5-6: Why is the spherical model in particular important for structural and spatial interpolation? I recommend removing this clause.

ANS:

We have revised the concerned sentences, and removed the improper clause.

Please also note the supplement to this comment:
http://www.hydrol-earth-syst-sci-discuss.net/hess-2016-401/hess-2016-401-AC2-supplement.pdf

---

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2016-401, 2016.

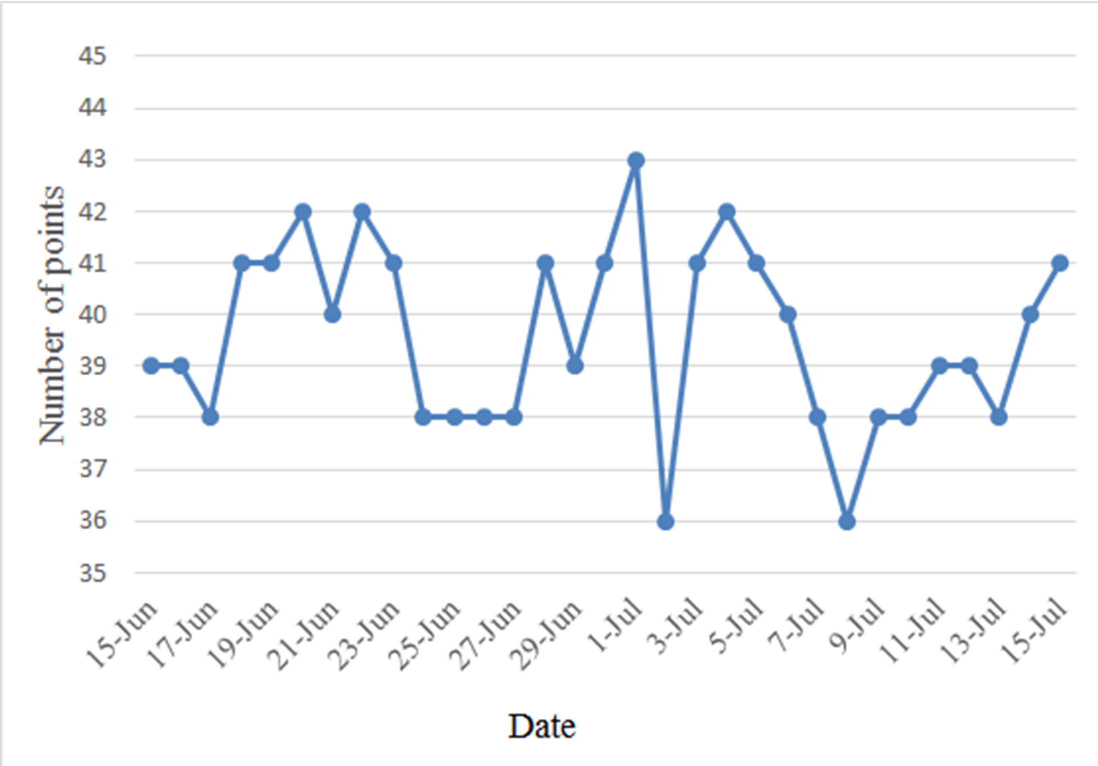

Fig. 1. The actual number of points considered for each interpolation day

Printer-friendly version

Discussion paper