

Article

Optimal Nodes Selectiveness from WSN to Fit Field Scale Albedo Observation and Validation in Long Time Series in the Foci Experiment Areas, Heihe

Xiaodan Wu ^{1,2}, Qing Xiao ^{1,*}, Jianguang Wen ^{1,3,*}, Qiang Liu ⁴, Dongqin You ¹, Baocheng Dou ¹, Yong Tang ¹ and Xiaowen Li ^{4,†}

¹ State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, No. 20A Datun Road, Beijing 100101, China; E-Mails: upcwuxiaodan@163.com (X.W.); youdongqin@gmail.com (D.Y.); dou3516@163.com (B.D.); tangyong@163.com (Y.T.)

² College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China

³ Joint Center for Global Change Studies, Beijing 100875, China

⁴ College of Global Change and Earth System Science, Beijing Normal University, No.19 Xijiekouwai Street, Haidian District, Beijing 100875, China; E-Mails: toliuqiang@bnu.edu.cn (Q.L.); lix@bnu.edu.cn (X.L.)

[†] Deceased.

* Authors to whom correspondence should be addressed; E-Mails: xiaoqing@radi.ac.cn (Q.X.); wenjg@radi.ac.cn (J.W.); Tel.: +86-10-6484-2510 (Q.X.); Fax: +86-10-6480-6255 (Q.X.).

Academic Editors: Xin Li, Yuei-An Liou, Richard Müller and Prasad S. Thenkabail

Received: 31 July 2015 / Accepted: 22 October 2015 / Published: 5 November 2015

Abstract: To evaluate and improve the quality of land surface albedo products, validation with ground measurements of albedo is crucial over the spatially and temporally heterogeneous land surface. One of the essential steps for satellite albedo product validation is coarse scale observation technique development with long time ground-based measurements. In this paper, the optimal nodes were selected from the wireless sensor network (WSN) to perform observation at large scale and in longer time series for validation of albedo products. The relative difference is used to analyze the spatiotemporal representativeness of each node. The random combination method is used to assess the number of required sites (NRS) and then to identify the most representative combination (MRC). On this basis, an upscaling

transform function with different weights for each node in the MRC, which are calculated with the ordinary least squares (OLS) linear regression method, is used to upscale WSN node albedo from point scale to the field scale. This method is illustrated by selecting the optimal nodes and upscaling surface albedo from point observation to the field scale in the Heihe River basin, China. Primary findings are: (a) The method of reducing the number of observations without significant loss of information about surface albedo at field scale is feasible and effective; (b) When only few sensors are available, the most representative locations in time and space should be the first priority; when a number of sensors are available in the heterogeneous land surface, it is preferable to install them in different land surface, rather than the most representative locations; (c) The most representative combination (MRC) combined with the upscaling weight coefficients can give a robust estimate of the field mean surface albedo. These efforts based on ground albedo observations promote the chance to use point information for validation of coarse scale albedo products. Moreover, a preliminary validation of the MODIS (Moderate Resolution Imaging Spectroradiometer) albedo product was performed as the tentative application for upscaling predictions.

Keywords: coarse scale; albedo; heterogeneous; long time series; validation

1. Introduction

Land surface shortwave albedo is a key parameter in climate and biogeochemical models. It directly determines the amount of solar radiation absorbed by the land surface, and is widely used in the study of surface energy budget. Surface albedo can be obtained through land surface models [1], remote sensing retrieval [2–5], or land data assimilation [6], with different levels of accuracy and reliability. However, all these estimates are indirect measurements of albedo that must be validated through accurate ground measurements. Compared to the satellite footprints or coarse model grid cells, ground-based measurements are often considered point measurements because the spatial scale is usually limited to few square meters. For this reason, most of the published studies [7–12] considered the homogeneous sites under the assumption that the ground measurements can represent the same spatial domain as albedo products pixel. In fact, most of the land surface on earth is not homogeneous enough so that the spatial representative of each ground measurement is limited. The accuracy of albedo products over heterogeneous surface should be tested before application in global climate modeling and global environmental change.

The upscaling process is an essential step prior to validation due to the gaps of the footprints of albedo products and the spatial scale of ground measurements over heterogeneous land surface. In recent years, many studies are dealing with methods to scale the ground measurements up to the coarse pixel scale [5,13–16]. One method is to apply a multi-scale validation strategy that uses intermediate scale albedo, e.g., Enhanced Thematic Mapper (ETM) or Huan Jing (HJ) albedo, as an upscaling bridge to reduce the scale discrepancy [5,13]. However, using high-resolution albedo may introduce additional inversion or geometric registration errors [5], so the absolute accuracy of the remote sensing products could suffer from the effects of error propagation. Another method to upscale ground-based measurements is regression kriging (RK) technique combined with remote sensing data, which removes the

heterogeneous component by establishing the trend surface [14,16]. However, these methods relying on high-resolution satellite observations prevent their use in validation in long time series because of the limited time frame of high-resolution satellite observations. Further, the upscaling results may contain errors due to the sophisticated processes, which will introduce much uncertainty to the validation results.

In order to capture the spatio-temporal variation of surface albedo, multiple long-term continuous ground-based measurements are required, which can be upscaled directly to the coarse pixel scale through a transformation function. However, the installation and maintenance of permanent networks of surface albedo measurement sites are expensive and demand a lot of human resources. For this reason, most of the networks are designed for specific experiments of limited duration. For example, in the foci experimental area of Heihe Watershed Allied Telemetry Experimental Research project (HiWATER), 17 wireless sensor network (WSN) nodes were installed to capture the spatial heterogeneity of the surface albedo [17]. However, the maintenance cost of these instruments is so high that the observation time is very short, which began from May 2012 and ended September 2012. Furthermore, all these instruments were removed, except Node 15, after September 2012. For a planned surface albedo network, the balance between the representativeness/reliability and the economic costs should be considered. Vachaud *et al.* (1985) [18] reported a method is required to minimize the number of measuring points without significant loss of information to characterize the behavior of surface albedo at coarse scale.

The number of required sites (NRS) first appeared with an article by Hills *et al.* (1969) [19], and stands for the minimum number of sites required to estimate areal mean value within a predetermined accuracy [18,20–25]. Recently, several methods have been developed to determine the NRS for a specific area, including spatial-temporal variability analysis [18,23,26–30], geostatistical [31–33], stratified sampling [34,35], and bootstrap methods [20,36–39]. The spatial-temporal variability analysis usually uses some statistics, such as relative difference or correlation coefficient, to determine the NRS and the representative sites [18,23,26,27,30,40]. However, because of inadequate samples and the assumption of data independence, the applying condition of this method is strongly restricted [41]. Geostatistics, which consider the spatial autocorrelations of variables, offer an alternative for the optimal samples selectiveness. Nevertheless, this method needs the information of the semivariance. Although stratified sampling method has been used in many research areas, the effectiveness of this method highly depends on the information, which divides the population into internally homogeneous strata [35]. Recently, the bootstrap method has been used to determine the NRS to estimate the population mean. The major advantage of this method is that it does not require any *a priori* assumption on the sampling distribution in the population [20].

In the bootstrapping process, m samples are randomly selected from n available observations ($m = 1, 2, \dots, n$) to calculate the distribution of statistics. An adequate number of bootstrap replicates is needed in each sampling loop [42]. To save computation time, many studies set fixed replicates for the sampling designs [25,28,41]. However, the fixed replicates cannot cover all possible combinations between m and n (for example, when selecting nine samples from 17 available observations, all the combinations are $C_{17}^9 = 24310$), which will result in fluctuations of the confidence level with the sample sizes when determining NRS. To make the samples more representative for the footprint of coarse pixel scale, all possible combinations should be taken into consideration.

In this paper, we proposed a random combination method, which considers all the possible combinations based on the bootstrap method, combined with spatiotemporal representativeness analysis, to determine the NRS and most representative combination (MRC) for the WSN of the foci area in the HiWATER experiment. Furthermore, on this basis, the ground point measurements of the MRC were upscaled to the field scale by applying a transformation function, which is calculated with the ordinary least squares (OLS) linear regression method. This method provides an efficient way to design sparse validation networks and also provides a way to use the existing ground networks for validation.

The specific objective is using the optimal nodes of MRC to obtain the field scale albedo for the validation of coarse scale albedo products in long time series. The mean values of the MCD43B3 pixels corresponding to the study area were assessed as the tentative application for the method. The paper is organized as follows: Firstly, the study area and data are described in Section 2. Secondly, upscaling method based on spatial-temporal representativeness analysis and the random combination method are outlined in Section 3. In Section 4, the results are presented and discussed. Finally, a brief conclusion is given.

2. Study Area and Datasets

The study area (Figure 1) is located in the middle reach of the Heihe River Basin, which was selected as an experimental watershed for conducting the HiWATER because it is a typical inland river basin that has long served as a test region for integrated watershed studies as well as land surface and hydrologic experiments [43,44]. HiWATER is an ongoing, watershed-scale, eco-hydrological experiment designed from an interdisciplinary perspective to address complex problems, such as heterogeneity, scaling, uncertainty, and closing the water cycle at the watershed scale [44].

The core experimental area is approximately 5.5 km by 5.5 km, located between 38°50′–38°54′N and 100°19′–100°24′E. The main crop in the study region is corn, which covers approximately 75% of the total area. The albedo will be different for different regions even if they are all covered with corn because of the different soil moisture levels and irrigation statuses. Other plants, including wheat, vegetable, and orchards, are also present. Additionally, the land cover types include buildings, roads, greenhouses, and irrigation channels. Seventeen ground wireless sensor network (WSN) measurement nodes were selected to capture the spatial heterogeneity of the surface albedo within the core experimental area and each WSN node is a field site. Details concerning specific performances and inter-comparisons can be found in [17]. The spatial distribution of the WSN nodes is shown in Figure 1b. Description of each WSN node can be found in Table 1.

Field sites were instrumented with Kipp and Zonen Radiometers to measure the total downward and upward shortwave radiation (300–2800 nm), and the ground albedo was calculated based on their ratios. When measurements are concerned, the error of the instrument usually deserves attention. The measurement error includes systematic error and random error. The systematic error can be ignored because the radiometers have been strictly calibrated. Although the random error of radiometers still exists with a very small mean relative error of 1.24% [17], it can be reduced by average combination of different nodes. Both radiometers were mounted 3 m above the cropland with a footprint of approximately 26 m in diameter [45]. These radiometers take radiation measurements every 10 min, and the ground-based albedo observations could be precisely synchronized with satellite remote sensing data.

As shown in the Table 1, the measuring times of each WSN node are not consistent. To eliminate the effect of the difference of the measuring time, only the WSN data between 10 June 2012 (Day of year (DOY) 162) and 16 September 2012 (DOY 260), when all WSN nodes worked, were collected from the study area. Node 1 is in a vegetable field, Node 4 is on building ground, Node 17 is in an orchard, and the others are in cornfields. The WSN nodes were divided according to the distribution of, for example, crop structure, windbreak, resident area, soil moisture, and irrigation status [17]. The data quality of Node 16 is poor, and thus the data of this node were removed from the following analysis.

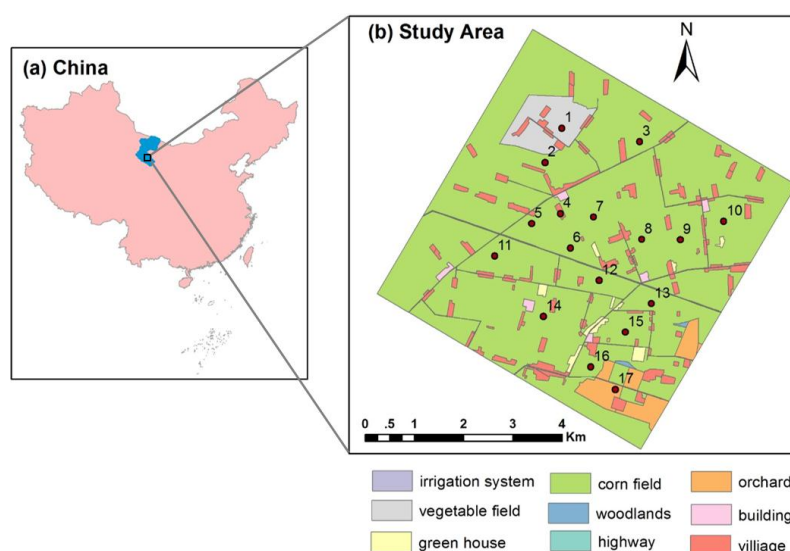


Figure 1. The location of the study area in China (a). The right image (b) is the map (Project: WGS_1984_UTM_Zone_47N) of the study area and the distribution of the wireless sensor network (WSN) nodes.

Table 1. Description of the WSN nodes.

Node	Lat	Lon	Elev (m)	Landcover	Measurement Period
1	100.35813	38.89322	1552.75	vegetable field	10 June 2012 10:30 a.m.–17 September 2012 9:10 a.m.
2	100.35406	38.88695	1559.09	cornfield	3 May 2012 1:50 p.m.–21 September 2012 3:00 p.m.
3	100.37634	38.89053	1543.05	cornfield	3 June 2012 5:10 p.m.–18 September 2012 2:50 p.m.
4	100.35753	38.87752	1561.87	building	10 May 2012 8:10 a.m.–17 September 8:50 a.m.
5	100.35068	38.87574	1567.65	cornfield	4 June 2012 10:30 a.m.–18 September 2012 9:00 a.m.
6	100.3597	38.87116	1562.97	cornfield	9 May 2012 5:50 p.m.–21 September 2012 5:30 p.m.
7	100.36521	38.87676	1556.39	cornfield	28 May 2012 5:10 p.m.–18 September 2012 2:40 p.m.
8	100.37649	38.87254	1550.06	cornfield	14 May 2012 3:20 a.m.–21 September 2012 8:50 a.m.
9	100.38546	38.87239	1543.34	cornfield	4 June 2012 7:10 p.m.–17 September 2012 5:20 p.m.
10	100.39572	38.87567	1534.73	cornfield	1 June 2012 12:00 a.m.–17 September 2012 2:50 p.m.
11	100.34197	38.86991	1575.65	cornfield	2 June 2012 6:20 p.m.–18 September 2012 9:00 a.m.
12	100.36631	38.86515	1559.25	cornfield	10 May 2012 10:40 a.m.–21 September 2012 2:50 p.m.
13	100.37852	38.86074	1550.73	cornfield	6 May 2012 4:30 p.m.–20 September 2012 2:50 p.m.
14	100.3531	38.85867	1570.23	cornfield	6 May 2012 5:40 p.m.–21 September 2012 8:50 a.m.
15	100.37225	38.85557	1559.0	cornfield	10 May 2012 10:40 a.m.–26 September 2012 0:50 p.m.
16	100.36411	38.84931	1564.31	cornfield	1 June 2012 0:00 a.m.–17 September 2012 4:50 p.m.
17	100.36972	38.8451	1559.63	orchard	12 May 2012 9:50 a.m.–17 September 2012 2:30 p.m.

As regard to the validation of albedo products, the albedo from the downward and upward radiation at local noon are derived for each WSN site. The local noon of the study area is 1:10 p.m.–1:20 p.m. The original temporal resolution of the measured radiation is 10 min. For the calculation of albedo at local noon, the mean radiation during local noon (half an hour before and after) was used as the measured value of every single day. Using the hour average of irradiance around local solar noon to construct the ground daily noon surface albedo can smooth the albedo variation due to high frequency process, such as wind blowing and other disturbance.

The 1-km MCD43B3 product was used as a preliminary application for the upscaling results, which was generated by the National Aeronautics and Space Administration (NASA; free download from [46]). It is derived from the integral of the bidirectional reflectance distribution function (BRDF) over hemispheric space. The BRDF is estimated using a three parameter kernel-driven semi-empirical RossThick-LiSparse-Reciprocal (RTLSR) model to characterize the anisotropic reflectivity of the land surface [47,48]. Considering the effects of registration error on the validation results, the mean MCD43B3 albedo within the study area was utilized to represent the average condition of MCD43B3 in an effort to reduce the errors caused by geometric registration.

3. Methods

This paper focuses on the determination of NRS and MRC, and then upscaling ground-based measurements from point scale to field-scale in the long time series. In this paper, a “point” is the location where an observation is made; a “field” is the area where a number of measurement points are made. A “sampling day” is the day when a number of measurements are made. Furthermore, the point scale refers to the footprint of the ground WSN node measurement, which is approximately 26 m in diameter. The field scale refers to the area where the 17 WSN nodes have been worked.

The optimal nodes selectiveness consists of three components (Figure 2). First the spatio-temporal variability analysis was performed to identify the most spatiotemporal representative node. The mean relative difference between the single node values and the field mean is used as the indicator to rank the sensor nodes [18,25–27,30,40,49,50]. It can improve the understanding the OCN and MRC, and eliminate the effects of the order of magnitude.

Second, the random combination method was performed. Compared with the bootstrap technique [20,36–39], the random combinations C_N^k are used in place of the fixed replicates to ensure an unbiased estimation. Besides, since the ground observations of each node and the field mean value are time sequence albedo vectors, the indexes to evaluate vector similarity, including angular cosine (Cosine), Euclidean distance (Euclidean) and correlation coefficient (R) are used as the indicators to determine the NRS of the study area and the OCN for each specific number of selected samples. On the basis of the above results, the MRC of the study area can be determined.

Third, the upscaling transformation function was proposed to estimate the field scale albedo from point observations, which is formulated from the observations of MRC and the average of all WSN nodes observations, based on the OLS linear regression method. Further, the field scale ground albedo from MRC was evaluated with the average of all WSN nodes. In the end, comparison between the field scale ground albedo and MODIS albedo products was performed.

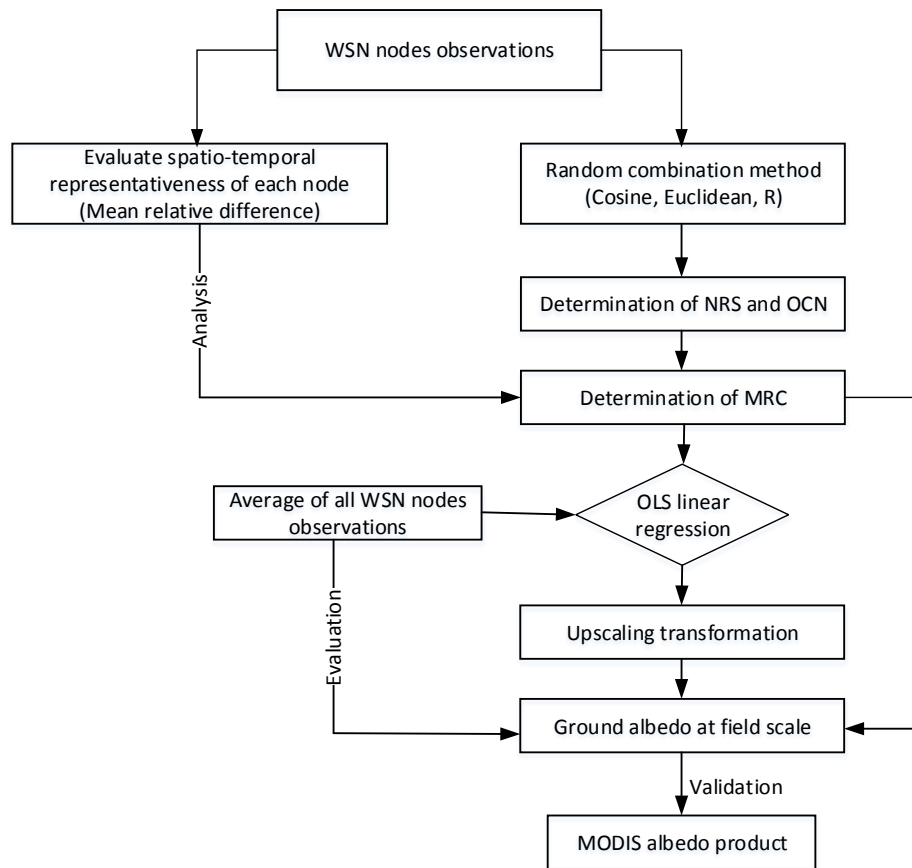


Figure 2. General flowchart of the method.

3.1. Spatiotemporal Characteristics and Assessment of Spatiotemporal Representativeness

To find the optimal nodes from the WSN to perform field scale observation in the long time series, the first step is to analyze the spatiotemporal characteristics of the field mean surface albedo. To demonstrate the spatial variability, the standard deviation (std) and the coefficient of variation (CV) among these WSN nodes are also calculated. The coefficient of variation is used here as statistical descriptor because it allows comparing the variability of different nodes even though characterized by different mean values, and, hence, to analyze the surface albedo variability in field scale.

Let θ_{ij} be the surface albedo observed at node i and day j , then the spatial mean of each sampling day, $\bar{\theta}_j$, is given by:

$$\bar{\theta}_j = \frac{1}{N} \sum_{i=1}^N \theta_{ij} \quad (1)$$

where N is the number of the WSN nodes. The coefficient of the variation of each sampling day, CV_j , is calculated as follows [29,51–53]:

$$CV_j = \frac{\sigma_j}{\bar{\theta}_j} = \frac{\sqrt{\frac{1}{N-1} \sum_{i=1}^N (\theta_{ij} - \bar{\theta}_j)^2}}{\bar{\theta}_j} \quad (2)$$

where σ_j is the standard deviation.

The spatiotemporal representativeness of each WSN node is assessed using the relative difference between the observation of a single node and the field average (denoted as benchmark). In practical, the relative difference (RD) is defined as [18,26–30,49,50]:

$$\delta_{ij} = \frac{\theta_{ij} - \bar{\theta}_j}{\bar{\theta}_j} \quad (3)$$

where $i = 1, 2, 3, \dots, N$ (identification number of nodes), $j = 1, 2, 3, \dots, M$ (number of observations over a measurement period). δ_{ij} indicates the degree of similarity between the node i and the field mean on the sampling day j . The mean relative difference (MRD) $\bar{\delta}_i$ [18,26–30,49,50] and the standard deviation of relative difference (SDRD) $\sigma(\delta_i)$ [27–30] for each site are calculated by:

$$\bar{\delta}_i = \frac{1}{M} \sum_{j=1}^M \delta_{ij} \quad (4)$$

$$\sigma(\delta_i) = \sqrt{\frac{1}{M-1} \sum_{j=1}^M (\delta_{ij} - \bar{\delta}_i)^2} \quad (5)$$

MRD characterizes the mean closeness of a single node to the field average in the entire sampling period. SDRD characterizes the precision of the estimation at a single node, which is calculated as an indicator of temporal stability. A small SDRD indicates that this node presents a similar temporal trend in surface albedo as the field mean surface albedo. Obviously, a stable and representative node of the mean value in time is characterized by a low value of both $|\bar{\delta}_i|$ and $\sigma(\delta_i)$. Especially, the root mean square difference (RMSD) of RD is computed as [25,30,50]:

$$RMSD_i = \sqrt{\bar{\delta}_i^2 + \sigma(\delta_i)^2} \quad (6)$$

RMSD provides a single value combining MRD and SDRD, making it a single metric to identify the most representative node (MRN). The node with the smallest RMSD was identified the most representative in time and space.

3.2. Random Combination Method

In this paper, a random combination method, based on the bootstrap technique [20,36–39], was adopted to determine the NRS for the study area. Cosine, Euclidean and R are used as the indicators to assess the performance of samples and then determine the value of NRS. Cosine and R are used to measure the similarity between two vectors. The closer to 1, the more similar the two vectors are [54]. Euclidean is used to measure the distance (difference) between two vectors. Quantization error is provided by quantifying the deviation degree from the benchmark [54]. Obviously, high values of Cosine and R , and low Euclidean value indicate a good performance of samples.

$$\text{Cosine} = \frac{a \cdot b}{\sqrt{a \cdot a} \times \sqrt{b \cdot b}} \quad (7)$$

$$\text{Euclidean} = \sqrt{(a - b)^2} \quad (8)$$

$$R = \frac{\text{cov}(a, b)}{\sqrt{\text{cov}(a, a) \cdot \text{cov}(b, b)}} \quad (9)$$

where a denotes the average albedo time sequence vector of a number of sampled nodes (subset mean), and b represents the average albedo time sequence vector of all the node observations. In this paper, the population is observations of the 16 site ($N = 16$, after removing Node 16). Specially, the maximum number of subsets for the specific number of measurement points is $C_N^k = \frac{N!}{k!(N-k)!}$ ($k = 1, 2, 3, \dots, k \leq N$). The flowchart of the random combination method is shown in Figure 3. The method consists of the following steps:

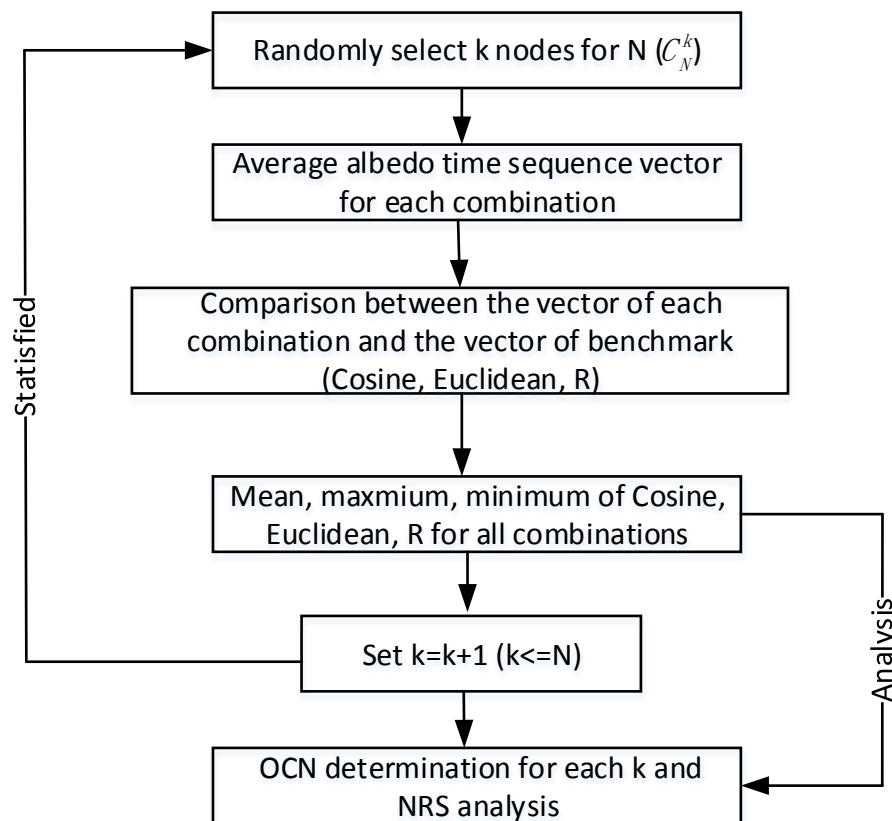


Figure 3. The flowchart of the random combination method.

- (1) Randomly select k node measurements ($k \leq N$) from the available N observations. Thus there are C_N^k possible combinations in all.
- (2) For each possible combination, the time sequence vector of the field mean surface albedo is calculated, thus obtaining C_N^k time sequence vectors in total.
- (3) C_N^k time sequence vectors are compared with the one vector based on all N measurement nodes (denoted as benchmark); for that, the Cosine, Euclidean and R are calculated. In addition, the mean, maximum, minimum value of the Cosine, Euclidean and R are calculated.
- (4) Set $k = k+1$, repeat step 1 to step 4 until $k = N$ to explore all the possible number of sampled nodes.

When this random combination method finished, the mean and maximum of Cosine, Euclidean and R can be considered as a function of k . The NRS can be determined when measurement nodes selected from the observations representing the particular area of the field. During this process, for each specific number of selected nodes, there is an OCN, which is indicated by the maximum of Cosine and R and

minimum of Euclidean. Following this determination of NRS and OCN, the combination that is most representative for the field scale, which is denoted as MRC for concise, is selected to be upscaled to the field scale in the long time series.

3.3. Upscaling Transfer Function

Although the mathematical average was adopted to determine the NRS and the most representative combination [18,20,22,26,27,30,41], in fact, the contribution of each node in the MRC for the field mean surface albedo is different [15]. To estimate field scale albedo from the point data for validation of coarse-scale albedo products, it is helpful to upscale the point data with an existing relationship. Upscaling continuous observations from multiple nodes involved assigning different weight coefficients to each sensor node. Therefore, a regression analysis between the surface albedo values sampled in the MRC and the average value based on all measurement nodes (denoted as benchmark) would allow determining reliable surface albedo on the field scale through few nodes observations.

$$A_{field} = \sum_{i=1}^n \alpha_i w_i \quad (n = 1, 2, 3, 4, \dots, \text{NRS}) \quad (10)$$

where A_{field} represents the field mean surface albedo in the field scale, here, we treated it as the average of all N measurement nodes. α_i is the WSN node observation of MRC, and w_i is the upscaling weight coefficient to be derived from the samples using the ordinary least squares (OLS) method [55].

Once the upscaling weight coefficients were derived, the weighted average of the MRC could be calculated as the upscaling results at field scale. It is important to note that the use of point data for field scale purposes increases the uncertainty of the field scale albedo and may introduce bias. Although the upscaling process can improve precision of the estimation, the uncertainty will increase since the upscaling is never perfect. Prior to utilization, the upscaling results should be evaluated. To assess the accuracy of field scale upscaling results, the average value based on all measurement nodes is used as a reference against which the accuracy and validity of the field scale upscaling values are compared. After this evaluation, the field scale upscaling albedo can be considered to be reliable and rigorous as the reference to validate coarse scale albedo products.

4. Results and Discussion

4.1. Spatiotemporal Characteristics

To demonstrate the spatial variability, the 16 nodes-averaged surface albedo (avg), the standard deviation (std) and coefficient of variation (CV) among these 16 nodes are shown in Figure 4. The spatiotemporal features can be found, as follows:

- (a) The daily local-solar-noon field mean surface albedos reveal obvious day-to-day fluctuations. This phenomenon is strongly correlated with the change of the land surface cover. During the experimental period, the cornfields, as the main landcover, start from seed around DOY 162, and then are in closing with a little gap to the background around DOY 207, and at last mature with withered cornstalks around DOY 245. Besides, crop management (irrigation, fertilization, etc.) is frequent during this period, which has a strong impact on the field mean surface albedo.

- (b) Avg, std and CV show great temporal fluctuations. In addition, the CV shows almost the same trend as the std, while it shows a large difference compared with avg, indicating that the CV is insensitive to average albedo, and the std is the dominant factor to influence the CV. The highly dependence of CV on std indicates the highly spatial heterogeneity of the study area.

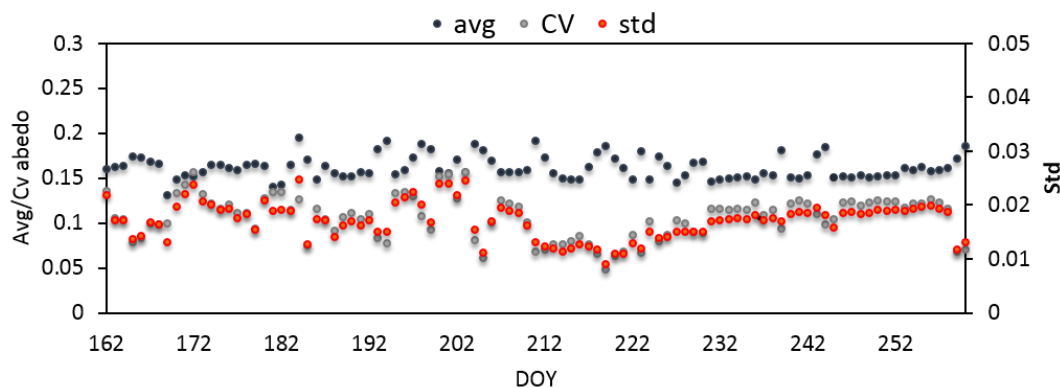


Figure 4. Time series of 16 sites-averaged surface albedo in the study area with standard deviation (std) and coefficient of variation (CV).

4.2. Spatiotemporal Representativeness of the WSN Nodes

The relative difference analysis is conducted independently for the surface albedo observations. Figure 5 shows the rank of nodes ordered by RMSD, meanwhile, the MRD and SDRD of corresponding node are also presented. From this figure, the ranges of the RMSD and MRD are very large, while the range of SDRD is relatively small. The RMSD has the similar magnitude and trend with MRD. This is because MRD in Equation (6) is the dominant variable in calculating RMSD, while the SDRD has approximately the same value for different nodes (except for Node 4, which is on a building surface). The large MRD range is caused by both the spatial heterogeneity and temporal heterogeneity. The latter is caused by the difference of temporal trend in different positions. Particularly, the influence of the spatial heterogeneity may be remarkably, because Node 4 (on building ground) and Node 17 (in orchard) show significantly large values of MRD.

For Node 17, although the MRD value is significantly larger than that of the nodes in the cornfield, the SDRD shows similar value with the nodes in the cornfield. This occurs because this node is in an orchard and thus shows large deviation compared with the field mean albedo (Figure 6), which is dominated by cornfield. However, this deviation does not change significantly over time since the orchard shows a similar temporal trend in surface albedo as the field average albedo as shown in Figure 6. Thus, the RD is relatively stable in time and the SDRD is small. For Node 4, both the MRD and SDRD are the largest among these nodes. This node is on the building ground, which shows the largest albedo in the study area. Thus the deviation of albedo value between this node and the field mean is largest (Figure 6), which contributes to the largest MRD. Further, the difference between building and field mean albedo is changing as shown in Figure 6, thus the standard deviation of RD is large. This is the result of the large SDRD. For Node 1, which is in a vegetable field, the MRD and SDRD is relatively small, compared with the nodes on building and orchard. This is because the deviation between vegetable and field mean

albedo is not as large as that of building and orchard. In addition, the temporal trend in surface albedo of vegetable and field mean is similar as shown in Figure 6.

For nodes installed on the cornfields, these nodes (except for Node 11) show smaller RMSD, compared with the nodes on building, orchard and vegetable (Figure 5). This is because the study area is dominated by cornfields, and the deviations of nodes on cornfields are relatively small. However, the nodes on different cornfields show different MRD and SDRD values. This occurs because different cornfields show different albedo values and temporal trend (Figure 7), due to the different growth, soil moisture and the combined effects of these factors. In particular, the Node 11 (on cornfield) even shows larger MRD and SDRD than the Node 1 on the vegetable field.

According to Figure 5, the MRN identified by rank of RMSD is Node 2. The values of RMSD, MRD and SDRD of this node are all less than 0.03. In fact, the first seven nodes (Node 2, Node 5, Node 14, Node 12, Node 13, Node 9 and Node 15) are very close to each other in RMSD, MRD and SDRD. Furthermore, any of them may be the possible MRN as time goes on. From the Figure 6, it can be seen that even the Node 2 shows a certain bias compared to the field average. Thus, more nodes that can provide the heterogeneous information of the surface albedo within the study area are needed.

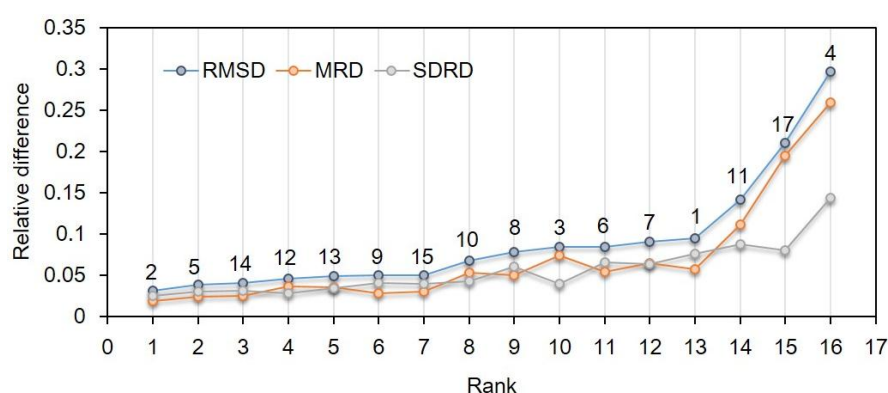


Figure 5. Rank ordered RMSD with its MRD and SDRD for the surface albedo.

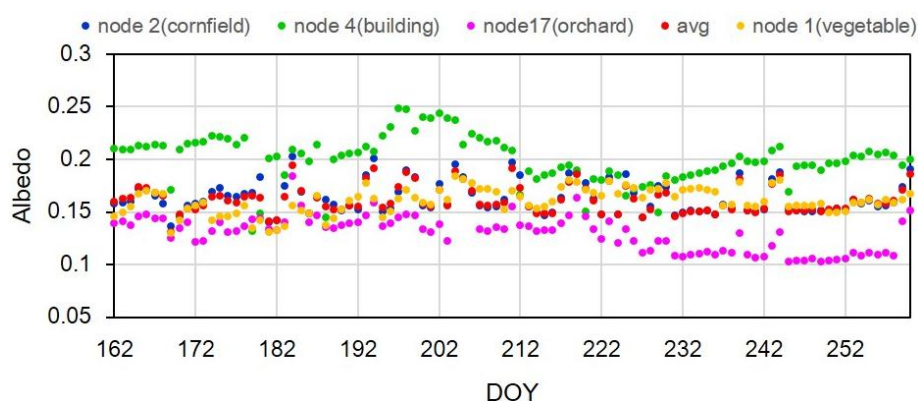


Figure 6. Time series of surface albedo over cornfield (taking Node 2 for example), building, orchard, vegetable and the field mean value over the study area.

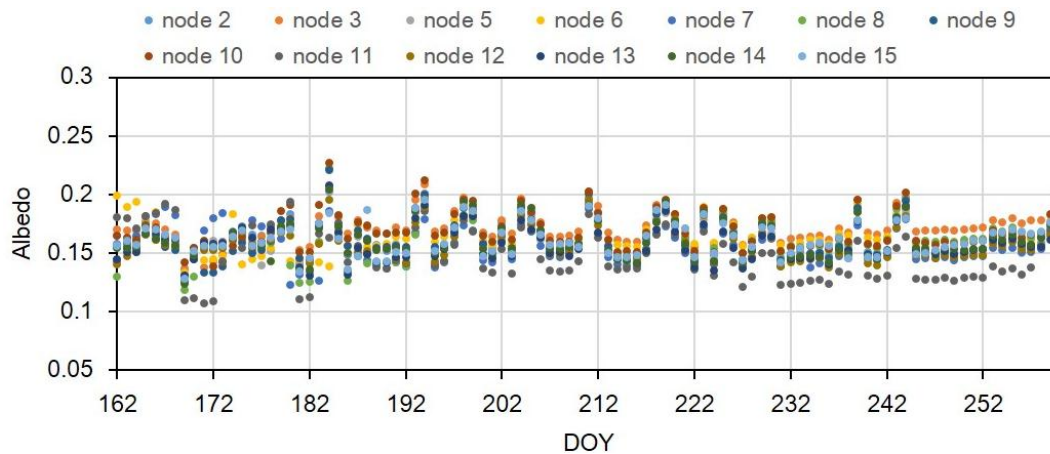


Figure 7. Time series of surface albedo over different cornfields over the study area.

4.3. NRS and MRC Determination

The Cosine, Euclidean and R are computed between the benchmark field-mean surface albedo and average albedo of a number of sampled nodes randomly selected (subset mean). The mean Cosine, Euclidean and R of all possible combinations for the different number of nodes, were plotted against the number of measurements itself (see Figure 8a). As shown in Figure 8a, large Euclidean (or low R) is obtained with small number of sampled nodes (NSN), while its value decreases (or increases) rapidly with increasing NSN, as expected. Then both Euclidean and R change very little and approach relatively stable values with further increased NSN until NSN = 16. The mean Cosine value of all possible combinations for the specific number of nodes is relatively stable, which demonstrates that the mean Cosine value is not sensitive to NSN. The above results are reasonable because larger number of samples may yield more accurate estimate of field mean surface albedo.

A certain value of Euclidean, R and Cosine can be used as criterion to determine NRS. As it can be seen from Figure 8a, an Euclidean of nearly 0.03 and a very high R , equal to 0.99, were obtained with 10 measurement nodes from the study area; By increasing the sample size, the accuracy improves only slightly, indicating a low gain in collecting more than 10 measurements. However, these kinds of indicators, including mean Euclidean, R and Cosine, cannot correctly reflect the NRS of the study area because they do not take into account the most representative node (MRN) in time and space, and the impact of the OCN. In fact, there is an OCN for each specific number of sampled nodes. The determination of NRS is sort of subjective issue and the random sampling process may contain risk [25]. Subjected to different criteria and different conditions, such as spatial scales and total deployed nodes, the estimated NRS values can be various. A more in-depth analysis should be conducted taking OCN and MRN into consideration.

For the newly proposed OCN, all combinations for a specific NSN are marked with an identification number (ID) and ranked by Euclidean value (R and Cosine) from the smallest (largest) to the largest (smallest). Thus the combinations with higher ranks perform relative better than others in estimating field mean surface albedo. Figure 8b–d illustrates the results of the combinations ranking the top (defined as OCN) among all possible combinations for different NSN, characterized by Cosine, R and Euclidean, respectively.

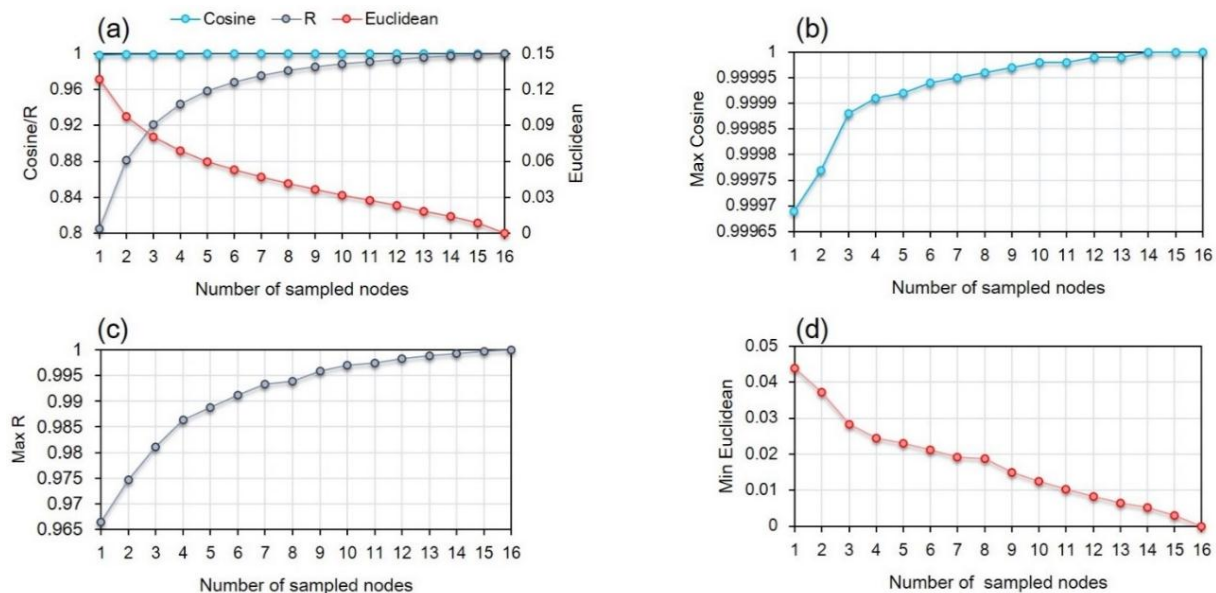


Figure 8. Relationship of angular cosine (Cosine), correlation coefficient (R) and Euclidean distance (Euclidean) for the surface albedo with the number of the sampled sites at daily scales: **(a)** The mean value of Cosine, R and Euclidean of all possible combinations for different number of measurement points. **(b)** The maximum Cosine value of all possible combinations for different number of measurement points. **(c)** The maximum R value of all possible combinations for different number of measurement points. **(d)** The minimum Euclidean value of all possible combinations for different number of measurement points.

As shown in the Figure 8c,d, the ranges of maximum R and minimum Euclidean of all possible combinations with respect to a specific NSN show large difference compared with that of mean R and Euclidean in Figure 8a, indicating that for a specific NSN, the mean values of sampled nodes in different combinations are significantly different. Table 2 summarizes the mean, maximum and minimum value of Cosine, R and Euclidean of all the possible combinations for different number of sampled nodes. It can be seen that with the increase of the NSN, the Cosine values do not have obvious changes, whether for the mean value or the maximum and minimum value, indicating the Cosine values is insensitive to NSN. For the R value, the mean value grows slowly with the increase of NSN, however, the minimum value (Min) changes significantly with the increase of NSN, while the maximum value (Max) shows only slight variations. This result indicates that the R value of the OCN is insensitive to NSN, however, when ignoring the OCN, the NSN plays a vital role in estimating the field mean surface albedo. The Euclidean has similar results with R .

As above analysis, the NRS of the study area is less than 10 when taking OCN into consideration, and the results in Figure 8b–d also prove this conclusion. For Cosine of the OCN, the values are stable when NSN = 6; for R of the OCN, the values grow slowly when NSN reaches seven; while for Euclidean of the OCN, there is no visible break, but the Euclidean values are very small on the whole. Thus, values from 6 to 10 are considered as the potential NRS for further evaluation. In fact, there are so many combinations in total that the top hundreds of combinations perform very close to each other and any of them may be the possible OCN as time goes on.

Because the *in situ* observations are continuous with long time series, the landcover maybe changed as time goes on, and thus the representativeness of node will change accordingly. To test the reliability of potential NRS, the frequency distribution of number of subsets with respect to different levels of R and Euclidean when NSN ranging from 6 to 10 are shown in Figures 9 and 10, respectively. The Cosine was not considered here since its insensitivity to NSN.

Table 2. The mean, max and min value of angular cosine (Cosine), correlation coefficient (R) and Euclidean distance (Euclidean) of all the possible combinations for different number of measurement nodes.

Number of Sampled Nodes	Cosine			R			Euclidean		
	Mean	Max	Min	Mean	Max	Min	Mean	Min	Max
1	0.9981181	0.99969	0.99315	0.8052069	0.96633	0.24541	0.1289426	0.043975	0.44659
2	0.9990732	0.99977	0.9957	0.8812874	0.97472	0.44442	0.0978517	0.037346	0.2704
3	0.9994228	0.99988	0.99786	0.9210191	0.98106	0.64692	0.0806185	0.028328	0.20167
4	0.9995988	0.99991	0.99862	0.944065	0.98641	0.7603	0.0689519	0.024508	0.1629
5	0.9997047	0.99992	0.99906	0.9586796	0.98874	0.83394	0.0600858	0.023139	0.13259
6	0.9997758	0.99994	0.99935	0.9686372	0.99125	0.8903	0.052869	0.021268	0.11133
7	0.9998266	0.99995	0.99957	0.9758096	0.99332	0.92763	0.0467138	0.019214	0.095883
8	0.999865	0.99996	0.99965	0.981203	0.99388	0.95092	0.0412758	0.018718	0.083555
9	0.9998948	0.99997	0.9997	0.9853979	0.99588	0.96487	0.0363307	0.014846	0.074969
10	0.9999188	0.99998	0.99975	0.9887498	0.99699	0.97121	0.0317171	0.012533	0.067144
11	0.9999386	0.99998	0.9998	0.9914875	0.99748	0.97516	0.0273054	0.010248	0.060526
12	0.9999548	0.99999	0.99983	0.9937648	0.99831	0.97937	0.0229754	0.0081909	0.054514
13	0.9999686	0.99999	0.99988	0.9956881	0.9989	0.98462	0.0185927	0.0065248	0.046738
14	0.9999803	1	0.99992	0.9973334	0.99931	0.98969	0.013961	0.0052595	0.038804
15	0.9999919	1	0.99995	0.9987569	0.9998	0.99426	0.0086297	0.0029524	0.029923
16	1	1	1	1	1	1	0	0	0

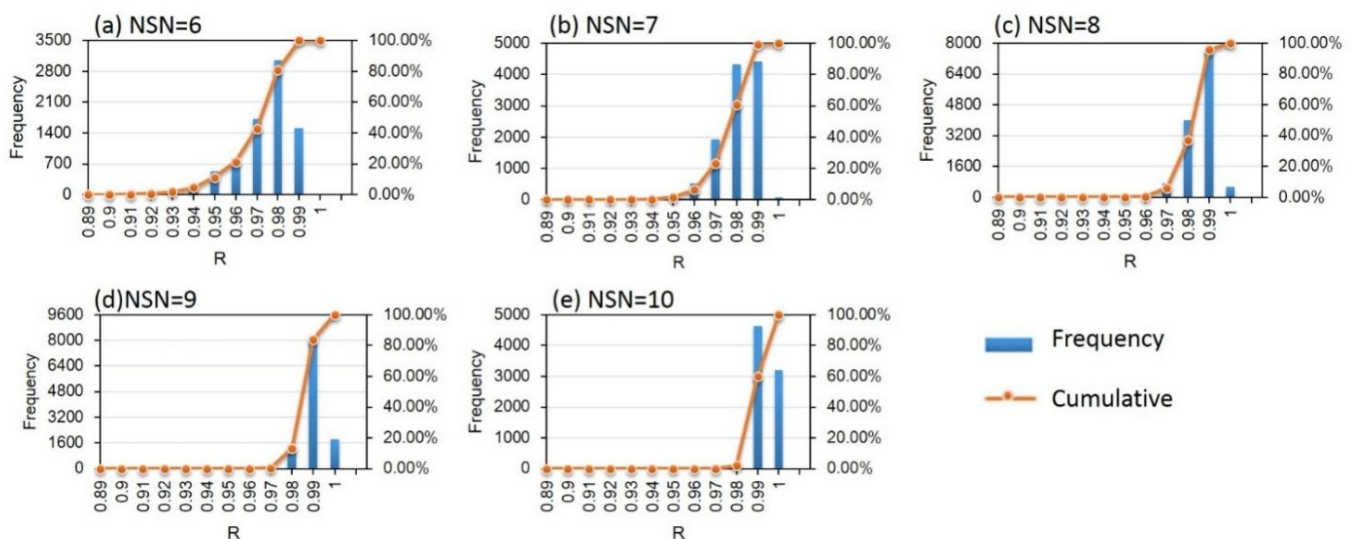


Figure 9. Frequency distribution histogram of number of sampled subsets and the cumulative frequency for R when NSN ranging from 6 to 10. (a–e) are the results when NSN are 6, 7, 8, 9, 10, respectively.

As shown in Figure 9, 86.8% of the subsets in total have R larger than 0.99 when NSN equals nine (Figure 9d), which is much larger than the results for NSN ranging from six to eight (Figure 9a–c). However, the improvement of R is not remarkable when the NSN increases to 9. Similar results can be found for Euclidean (Figure 10), the percentage of the Euclidean within 0.03 increases very quickly with the increase of NSN when NSN is less than nine. However, the improvement is slightly when NSN reaches nine. With increasing NSN, the cost of the fieldwork will increase as well. Thus, conclusion can be drawn that the NRS of the study area is nine. Table 3 lists the OCN characterized by Cosine, R and Euclidean for different NSN. Nodes 1–17 are identified as 1–17, respectively.

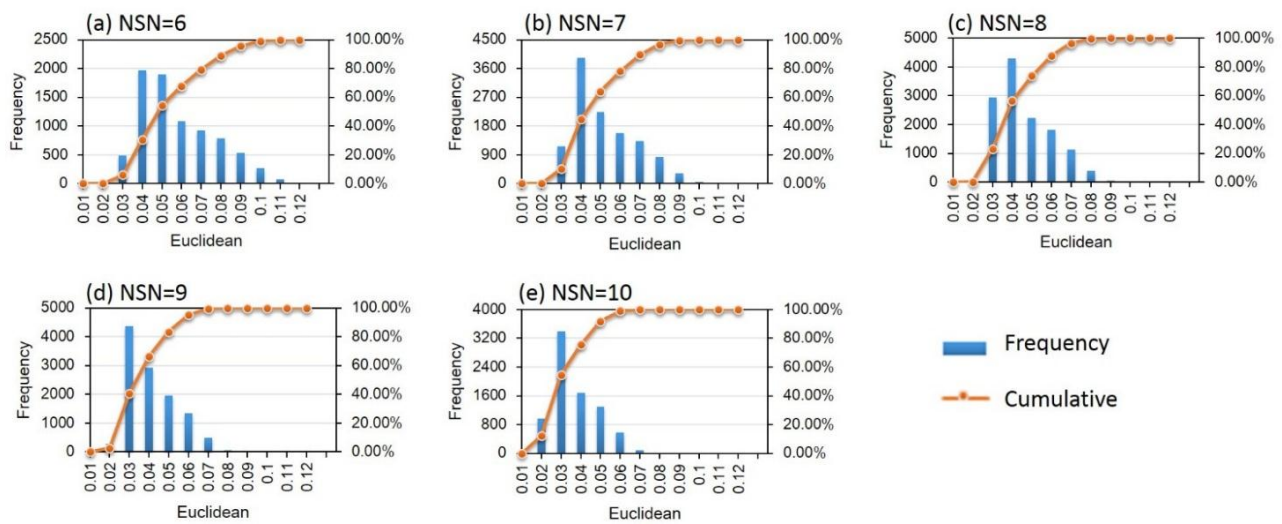


Figure 10. Frequency distribution histogram of number of sampled subsets and the cumulative frequency for Euclidean when NSN ranging from 6 to 10. (a–e) are the results when NSN are 6, 7, 8, 9, 10, respectively.

As shown in Table 3, the case for NSN = 1 is similar to the study of MRN in the Figure 5. Node 2 is the most representative in time and space. When only few nodes (less than four) are available, the locations of nodes on cornfields should be the top-priority. When more than four nodes are available, more potential OCN may appear. The OCN in this case includes the nodes with different ranks, even the nodes that rank last according to RMSD in the Figure 5. In other words, researchers may have more choices in maintaining surface albedo network in the wild when a number of nodes are available. Since the NRS in the study area has been identified, we just focus on the OCN for NSN = 9. Therefore, combinations with identification numbers of “1,2,4,5,8,9,10,11,12”, and “4,6,8,10,11,12,14,15,17” are considered as the potential MRC for further evaluation. As shown in Figure 1b, the study area is not only covered by vegetation with a relatively low albedo values, but also buildings, green house and villages with high albedo values, which is the result of the heterogeneity of the study area. Physically, since most of the nodes are installed on the vegetation cover, it is reasonable to have Node 4 (on building) with the largest albedo values in the MRC. In addition, the node on the orchard (Node 17) has the lowest albedo values as shown in the Figure 6, and thus should be included in the MRC. The deviation of observed surface albedo at these nodes contains important information on spatial heterogeneity. The OCN “4,6,8,10,11,12,14,15,17” integrates the nodes on the building (Node 4) and orchard (Node 17), and thus is identified as the MRC of the study area to provide more robust estimate of field mean surface albedo.

Table 3. The OCN identified by the maximum angular cosine (Cosine), maximum correlation coefficient (R) and minimum Euclidean distance (Euclidean) for different number of measurement nodes.

Number of Sampled Nodes	Cosine (Max)	R (Max)	Euclidean (Min)
1	2	2	2
2	2,3	5,12	7,10
3	3,5,7	7,9,14	3,5,7
4	2,3,5,7	2,3,5,7	3,5,7,12
5	2,6,8,13,17	2,6,8,13,17	1,2,5,7,10
6	3,6,7,13,14,17	3,6,7,13,14,17	1,3,5,7,9,13
7	2,3,6,7,13,14,17	2,3,6,7,13,14,17	1,2,3,5,7,9,13
8	2,3,6,7,13,14,15,17	2,3,6,7,13,14,15,17	1,2,3,5,7,10,11,14
9	1,2,4,5,8,9,10,11,12	1,2,4,5,8,9,10,11,12	4,6,8,10,11,12,14,15,17
10	1,2,4,5,8,9,10,11,12,15	1,2,4,5,8,9,10,11,12,15	3,4,6,8,11,12,13,14,15,17
11	1,3,4,5,7,9,10,11,12,14,15	1,3,4,5,7,9,10,11,12,14,15	2,3,4,6,8,11,12,13,14,15,17
12	1,4,6,8,9,10,11,12,13,14,15,17	1,4,6,8,9,10,11,12,13,14,15,17	1,2,4,6,8,9,10,11,13,14,15,17
13	1,2,4,6,8,9,10,11,12,13,14,15,17	1,2,4,6,8,9,10,11,12,13,14,15,17	1,2,4,6,8,9,10,11,12,13,14,15,17
14	1,2,4,5,6,7,8,9,10,11,13,14,15,17	1,2,3,4,6,7,8,9,10,11,13,14,15,17	1,2,3,4,5,6,8,9,11,12,13,14,15,17
15	1,3,4,5,6,7,8,9,10,11,12,13,14,15,17	1,3,4,5,6,7,8,9,10,11,12,13,14,15,17	1,3,4,5,6,7,8,9,10,11,12,13,14,15,17

In order to see whether the effect of the random measurement error can (or should) be neglected on the determination of NRS and MRC, the relative difference between the average albedo of a number of sampled nodes randomly selected (subset mean) and the benchmark field-mean surface albedo was calculated. The mean relative difference of all possible combinations for the different number of sampled nodes, were plotted against the number of sampled nodes in Figure 11. It can be seen that when NSN is small (less than three), the mean relative difference is several times larger than the random measurement error (relative error = 1.24%). Although the mean relative difference is decreasing with the increasing of NSN, meanwhile, the random measurement error of different nodes cancel out each other by being averaged. Thus, the random measurement error would not affect the determination of NRS and MRC since the above analysis are based on the average of a number of sampled nodes randomly selected (subset mean).

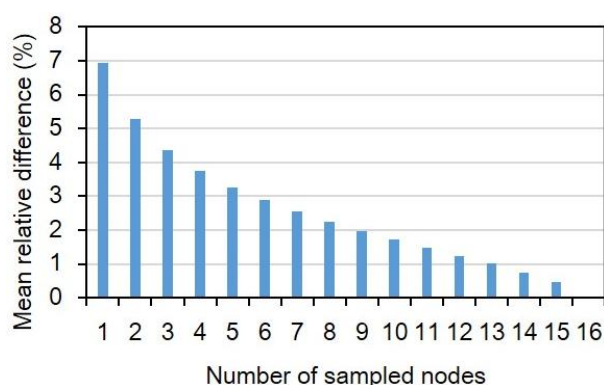


Figure 11. The mean relative difference of all possible combinations when different numbers of nodes are chosen.

4.4. Upscaling Results and Evaluation

Table 4 lists the upscaling weights for each node in the MRC, which were established statistically with the OLS method. In addition, the rank orders of spatial-temporal representativeness for each node in the MRC are also presented. To assess the accuracy of the upscaling results at the field scale, the average values based on all measurement nodes are used as the benchmark and compared with the field scale albedo, which is calculated from the MRC combined with the upscaling weights. Figure 12 shows a time series of upscaling results at the field scale *versus* the benchmark. A scatterplot between them is shown in Figure 13.

Table 4. The upscaling weights and the rank order of representativeness for each node in the MRC.

Node	Weight	Rank
node4	0.090953	16
node6	0.10031	11
node8	0.14274	9
node10	0.14665	8
node11	0.07954	14
node12	0.19748	4
node14	0.073957	3
node15	0.086302	7
node17	0.079639	15

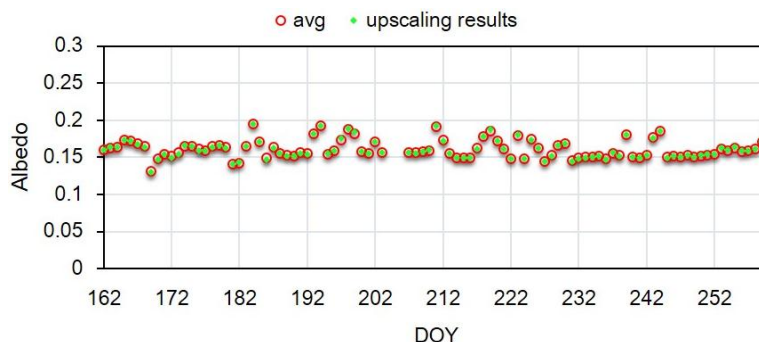


Figure 12. Comparison between daily time series of the average value based on all measurement nodes and the upscaling results in the field scale.

As shown in Figure 12, the upscaling results of the MRC at field scale are consistent with the benchmark over the entire experimental period. Both types of albedo have reasonable day-to-day fluctuations. The upscaling results of the MRC at field scale and the benchmark are correlated with a very high determine coefficient (R^2) value of 0.996 as shown in Figure 13, and the RMSE is as small as 0.0008. The mean bias between the upscaling results of the MRC and the benchmark is close to 0. Furthermore, the maximum difference between these two kinds of value is 0.002. These statistics (Figure 13) demonstrate the accuracy of the upscaling results at field scale that was upscaled from point measurements at multiple nodes from the MRC with the upscaling coefficients. Thus, the MRC and the upscaling coefficients can accurately represent the mean albedo of the entire study area. They can guide the design of sparse validation networks and the upscaling transformation from point scale to field scale.

for direct validation of coarse scale albedo products. To show a preliminary application, the upscaling results during the experimental period were used as the most valid (unbiased) reference to directly evaluate the mean MCD43B3 albedo within the study area in the time series.

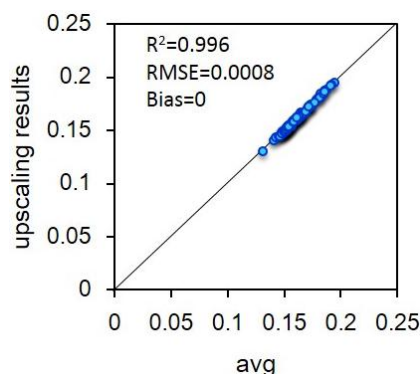


Figure 13. Scatterplots between daily time series of the average value based on all measurement nodes and the upscaling results in the field scale.

4.5. Preliminary Application for Assessment of MCD43B3 Product

MCD43B3 is a 16-day product, while the temporal resolution of upscaling value at field scale is one day. Thus, the 16-day span albedo is computed by averaging the daily upscaling value under clear sky conditions corresponding to the MODIS 16-day product. Figure 14 shows preliminary validation results for the mean MCD43B3 albedo within the study area. The geometric displacement uncertainties can be reduced to some extent by averaging operation. The statistical results of the RMSE, bias and R^2 calculated in the time series are also displayed. The MCD43B3 albedo data are always higher than the upscaling values at field scale with a bias of 0.008. The RMSE is 0.009 for MCD43B3 albedos, and the R^2 between the MODIS albedo and field scale albedo is as high as 0.67. The low RMSE and high R^2 demonstrate the good accuracy of the MODIS albedo.

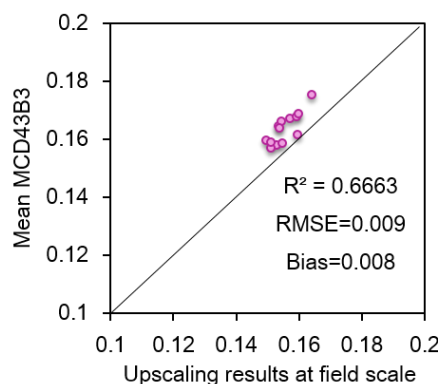


Figure 14. Scatterplots between the mean MCD43B3 and upscaling results in the field scale.

The discrepancy between the MCD43B3 and the 16-day upscaling values are mainly caused by two factors. First, the MODIS BRDF/albedo algorithm assumes that the physical parameters of the land surface remain constant within 16 days; the algorithm was set to screen out obvious changes in the measured albedo. However, the surface albedo changes a lot over a 16-day period as shown in Figure 4.

Furthermore, the variation of the surface albedo will introduce large errors in the MODIS-derived albedo. Second, the MCD43B3 is a 16-day composite product, while the upscaling results is based on *in situ* observations every day and averaged as composite albedo during this period, which will also contribute to the discrepancy between these two kinds of values.

5. Conclusions

Land surface albedo networks are necessary to validate albedo products estimated through land surface model, satellite remote sensing, and land data assimilation at a coarse scale. A fundamental requirement of successful validation is the quality of ground-based albedo measurements at the corresponding pixel scale. This paper describes a new approach that uses less WSN observations to provide the field scale ground albedo over heterogeneous land surfaces, including spatiotemporal representativeness assessment, the NRS and MRC determination based on random combination method to fit field scale, and a method for upscaling multi-node measurements in the MRC from point scale to field scale.

In this paper, the relative difference analysis is introduced to investigate temporal-spatial stability of each node albedo and then to identify the MRN for determination of MRC. Even the observation of the MRN cannot accurately represent the field mean albedo due to the heterogeneity of surface albedo. Joint observation with other nodes, which contain the heterogeneous information of surface albedo, is required. However, with the increase of the number of nodes, the cost of the fieldwork will increase as well. The balance between the accuracy and the cost can be established through a random combination method.

The random combination method is used to assess the NRS and the MRC approach is proposed to estimate field scale surface albedo. NRS is much less than the number of WSN nodes established before, and finally it has proved the existence and efficiency of MRC in estimating field scale surface albedo. However, this does not mean that the NRS is mandatory required. In fact, other number of nodes may show reasonable results, too. Subjected to spatial scales of ground measurements, accuracy levels, and the total deployed nodes, the estimated NRS values can be various.

For a continuously operating albedo monitoring network, we also investigated if it is preferable to install a number of albedo sensors in the most representative nodes. From the results of the OCN for each specific number of sampled nodes, it is clear that if only few nodes are available in the study area, it is preferable to install the sensor in the most representative sites. If more than four nodes are available, more optimal combinations will appear, including the nodes in different landcover. This occurs because the study area is not homogeneous and includes not only high albedo value surfaces (buildings, villages and greenhouses) but also low albedo value surfaces (cornfields after irrigation). In this case, it is always preferable to install them in different land surfaces, even if a “representative” node was detected. This result is in agreements with other findings [28].

Another important issue in this paper is the upscaling process of ground measurements from point scale to the field scale. Weight coefficients with the continuous combined observations of multiple node sensors in the MRC were created using the OLS method. The upscaling results calculated from the MRC and weight coefficients were evaluated and prove the effectiveness of MRC and upscaling transformation on spatial and temporal scales. Thus the MRC and upscaling coefficients proposed in this paper can guide the design of sparse validation networks in future and the upscaling transformation from point scale to field scale for direct validation of coarse scale albedo products.

Although the upscaling results show reasonable results in the time series, there are still many deficiencies in this paper. The shortcoming of the study is using the average of all measurement nodes as the benchmark to perform the spatiotemporal representativeness analysis, the determination of NRS and MRC, and the calculation of upscaling weight coefficients. This is reasonable only when all nodes can capture the heterogeneity of surface albedo in field scale and thus the average of all measurement nodes can accurately present the mean albedo of the study area. In addition, due to the limitations in length of time series, the analyses are only conducted within a short period. Thus, the results in this study may be biased when applying to longer time series. However, the method of determination of NRS and MRC provides important guidance for the establishment of future surface albedo network. Furthermore, this paper provides a beneficial method for upscaling ground point measurements to field scale for the validation of coarse pixel albedo products in the long time series over heterogeneous surfaces.

Acknowledgments

This research is jointly supported by the Natural Science Foundation of China under Grant 91125003 and in part by the National Basic Research Program of China under Grant 2013CB733401. The authors would like to thank all of the scientists, engineers, and students who participated in the HiWATER campaigns, which provided WSN data products for this work.

Author Contributions

Xiaodan Wu was responsible for the data analysis and writing the manuscript. Jianguang Wen and Qing Xiao contributed to the main research ideas, data collection and manuscript organization. Qiang Liu contributed to the research ideas and data collection. Dongqing You, Baocheng Dou, and Yong Tang preprocessed the remote sensing data. All of the authors thoroughly reviewed and edited this paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Zhang, X.; Liang, S.; Wang, K.; Li, L.; Gui, S. Analysis of global land surface shortwave broadband albedo from multiple data sources. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2010**, *3*, 296–305.
2. Sellers, P.J. Remote Sensing of Land Surface Climatology Change. In *NASA/GSFC International Satellite Land Surface Climatology Project Report*; NASA Goddard Space Flight Cent.: Greenbelt, MD, USA, 1993.
3. Liang, S. Narrowband to broadband conversions of land surface albedo. I: Algorithms. *Remote Sens. Environ.* **2001**, *76*, 213–238.
4. Wen, J.G.; Zhao, X.J.; Liu, Q.; Tang, Y.; Dou, B.C. An improved land-surface albedo algorithm with DEM in rugged terrain. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 883–887.
5. Peng, J.J.; Liu, Q.; Wen, J.G.; Liu, Q.H.; Tang, Y.; Wang, L.Z.; Dou, B.C.; You, D.Q.; Sun, C.K.; Zhao, X.J.; *et al.* Multi-scale validation strategy for satellite albedo products and its uncertainty analysis. *Sci. China: Earth Sci.* **2015**, *58*, 573–588.

6. Liang, X.; Xu, M.; Gao, W.; Kunkel, K.; Slusser, J.; Dai, Y.; Min, Q.; Houser, P.R.; Rodell, M.; Schaaf, C.B.; *et al.* Development of land surface albedo parameterization based on Moderate Resolution Imaging Spectroradiometer (MODIS) data. *J. Geophys. Res. Atmos.* **2005**, *110*, 1341–1355.
7. Jin, Y.; Schaaf, C.B.; Woodcock, C.E.; Gao, F.; Li, X.; Strahler, A.H.; Lucht, W.; Liang, S. Consistency of MODIS surface bidirectional reflectance distribution function and albedo retrievals: 2. Validation. *J. Geophys. Res.* **2003**, *108*, 4159–4159.
8. Román, M.O.; Schaaf, C.B.; Woodcock, C.E.; Strahler, A.H.; Yang, X.; Braswell, R.H.; Wofsy, S.C. The MODIS (Collection V005) BRDF/albedo product: Assessment of spatial representativeness over forested landscapes. *Remote Sens. Environ.* **2009**, *113*, 2476–2498.
9. Román, M.O.; Gatebe, C.K.; Shuai, Y.M.; Wang, Z.S.; Gao, F.; Masek, J.G.; He, T.; Liang, S.L.; Schaaf, C.B. Use of *in situ* and airborne multiangle data to assess MODIS- and Landsat-Based estimates of directional reflectance and albedo. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1393–1404.
10. Wang, Z.S.; Schaaf, C.B.; Chopping, M.J.; Strahler, A.H.; Wang, J.D.; Román, M.O.; Rocha, A.V.; Woodcock, C.E.; Shuai, Y.M. Evaluation of moderate-resolution imaging spectroradiometer (MODIS) snow albedo product (MCD43A) over tundra. *Remote Sens. Environ.* **2012**, *117*, 264–280.
11. Wang, Z.S.; Schaaf, C.B.; Chopping, M.J.; Román, M.O.; Shuai, Y.; Fitzjarrald, D.R. Evaluation of MODIS albedo product (MCD43A) over grassland, agriculture and forest surface types during dormant and snow-covered periods. *Remote Sens. Environ.* **2014**, *140*, 60–77.
12. Cescatti, A.; Marcolla, B.; Santhana Vannan, S.K.; Pan, J.Y.; Román, M.O.; Yang, X.; Schaaf, C.B. Intercomparison of MODIS albedo retrievals and *in situ* measurements across the global FLUXNET network. *Remote Sens. Environ.* **2012**, *121*, 323–334.
13. Liang, S.; Fang, H.; Chen, M.; Shuey, C.J.; Walthall, C.; Daughtry, C.; Morisette, J.; Schaaf, C.; Strahler, A. Validating MODIS land surface reflectance and albedo products: Methods and preliminary results. *Remote Sens. Environ.* **2002**, *83*, 149–162.
14. Wu, X.D.; Xiao, Q.; Wen, J.G.; Liu, Q.; You, D.Q.; Dou, B.C.; Lei, F.; Tang, Y.; Li, X.W. Upscaling albedo over heterogeneous surfaces using ground measurements and high-resolution imagery. *Remote Sens.* **2015**, submitted.
15. Wu, X.D.; Wen, J.G.; Xiao, Q.; Liu, Q.; Peng, J.J.; Dou, B.C.; Li, X.H.; You, D.Q.; Tang, Y.; Liu Q.H.; *et al.* Coarse scale wireless sensor network-based albedo observations over heterogeneous land surfaces and preliminary results of validating Modis remote sensing products over Northern China. *Remote Sens. Environ.* **2015**, submitted.
16. Kang, J.; Jin, R.; Li, X. Regression kriging-based upscaling of soil moisture measurements from a wireless sensor network and multiresource remote sensing information over heterogeneous cropland. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 92–96.
17. Xu, Z.; Liu, S.; Li, X.; Shi, S.; Wang, J.; Zhu, Z.; Ma, M. Intercomparison of surface energy flux measurement systems used during the HiWATER-MUSOEXE. *J. Geophys. Res.: Atmos.* **2013**, *118*, 130–140.
18. Vachaud, G.; Passerat de Silans, A.; Balabanis, P.; Vauclin, M. Temporal stability of spatially measured soil water probability density function. *Soil Sci. Soc. Am. J.* **1985**, *49*, 822–828.
19. Hills, R.C.; Reynolds, S.G. Illustrations of soil moisture variability in selected areas and plots of different sizes. *J. Hydrol.* **1969**, *8*, 27–47.

20. Dane, J.H.; Reed, R.B.; Hopmans, J.W. Estimating soil parameters and sample-size by bootstrapping. *Soil Soc. Am. J.* **1986**, *50*, 283–287.
21. Famiglietti, J.S.; Devereaux, A.; Laymon, C.A.; Tsegaye, T.; Houser, P.R.; Jackson, T.J.; Graham, S.T.; Rodell, M.; Oevelen, P.J. Ground-based investigation of soil moisture variability within remote sensing footprints during the Southern Great Plains 1997 (SGP97) hydrology experiment. *Water Resour. Res.* **1999**, *35*, 1839–1851.
22. Famiglietti, J.S.; Ryu, D.; Berg, A.A.; Rodell, M.; Jackson, T.J. Field observations of soil moisture variability across scales. *Water Resour. Res.* **2008**, *44*, W01423.
23. Hupet, F.; Vanclooster, M. Intraseasonal dynamics of soil moisture variability within a small agricultural maize cropped field. *J. Hydrol.* **2002**, *261*, 86–101.
24. Brocca, L.; Morbidelli, R.; Melone, F.; Moramarco, T. Soil moisture spatial variability in experimental areas of central Italy. *J. Hydrol.* **2007**, *333*, 356–373.
25. Zhao, L.; Yang, K.; Qin, J.; Chen, Y.Y.; Tang, W.J.; Montzka, C.; Wu, H.; Lin, C.G.; Han, M.L.; Vereecken, H. Spatiotemporal analysis of soil moisture observations within a Tibetan mesoscale area and its implication to regional soil moisture measurements. *J. Hydrol.* **2013**, *482*, 92–104.
26. Cosh, M.H.; Jackson, T.J.; Bindlish, R.; Prueger, J.H. Watershed scale temporal and spatial stability of soil moisture and its role in validating satellite estimates. *Remote Sens. Environ.* **2004**, *92*, 427–435.
27. Martínez-Fernández, J.; Ceballos, A. Mean soil moisture estimation using temporal stability analysis. *J. Hydrol.* **2005**, *312*, 28–38.
28. Brocca, L.; Melone, F.; Moramarco, T.; Morbidelli, R. Spatial-temporal variability of soil moisture and its estimation across scales. *Water Resour. Res.* **2010**, *46*, W02516.
29. Brocca, L.; Tullo, T.; Melone, F.; Moramarco, T.; Morbidelli, R. Catchment scale soil moisture spatial-temporal variability. *J. Hydrol.* **2012**, *422–423*, 63–75.
30. Vanderlinden, K.; Vereecken, H.; Hardelauf, H.; Herbst, M.; Martínez, G.; Cosh, M.H.; Pachepsky, Y.A. Temporal stability of soil water contents: A review of data and analyses. *Vadose Zone J.* **2012**, *11*, 280–288.
31. Mcbratney, A.B.; Webster, R. How many observations are needed for regional estimation of soil properties? *Soil Sci.* **1983**, *135*, 177–183.
32. Di, H.J.; Trangmar, B.B.; Kemp, R.A. Use of geostatistics in designing sampling strategies for soil survey. *Soil Sci. Soc. Am. J.* **1989**, *53*, 1163–1167.
33. Chung, C.K.; Chong, S.; Varsa, E.C. Sampling strategies for fertility on a stony silt loam soil. *Commun. Soil Sci. Plant Anal.* **1995**, *26*, 741–763.
34. Cochran, W.G. *Sampling Techniques*, 3ed.; John Wiley & Sons, Inc.: New York, NY, USA, 1977; pp. 89–110.
35. Miyamoto, S.; Chacon, A.; Hossain, M.; Martinez, I. Soil salinity of urban turf areas irrigated with saline water: I. Spatial variability. *Landsc. Urban Plan.* **2005**, *71*, 233–241.
36. Efron, B. Computers and the theory of statistics: Thinking the unthinkable. *Soc. Ind. Appl. Math.* **1979**, *21*, 460–480.
37. Efron, B. The Jackknife, the Bootstrap and Other Resampling Plans. *Soc. Ind. Appl. Math.* **1982**, *38*, doi:10.1137/1.9781611970319.
38. Efron, B.; Gong, G. A leisurely look at the bootstrap, the jackknife, and cross-validation. *Am. Stat.* **1983**, *37*, 36–48.

39. Diaconis, P.; Efron, B. Computer intensive methods in statistics. *Soc. Ind. Appl. Math.* **1983**, *248*, 116–130.
40. Brocca, L.; Melone, F.; Moramarco, T.; Morbidelli, R. Soil moisture temporal stability over experimental areas in central Italy. *Geoderma* **2009**, *148*, 364–374.
41. Wang, C.M.; Zuo, Q.; Zhang, R.D. Estimating the necessary sampling size of surface soil moisture at different scales using a random combination method. *J. Hydrol.* **2008**, *352*, 309–321.
42. Manly, B.F.J. *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 2nd ed.; Chapman & Hall Press: London, UK, 1997; pp. 34–35.
43. Cheng, G.D. *Integrated Management of the Water-Ecology-Economy System in the Heihe River Basin*; Science Press: Beijing, China, 2009; p. 581.
44. Li, X.; Cheng, G.; Liu, S.; Xiao, Q.; Ma, M.; Jin, R.; Xu, Z. Heihe watershed allied telemetry experimental research (HiWATER): Scientific objectives and experimental design. *Bull. Am. Meteorol. Soc.* **2013**, *94*, 1145–1160.
45. Sailor, D.J.; Resh, K.; Segura, D. Field measurement of albedo for limited extent test surfaces. *Sol. Energy* **2006**, *80*, 589–599.
46. NASA's Earth Observing System/ Data and Information System. Available online: <http://reverb.echo.nasa.gov/reverb/> (accessed on 26 October 2015)
47. Lucht, W.; Schaaf, C.B.; Strahler, A.H. An algorithm for the retrieval of albedo from space using semiempirical BRDF models. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 977–998.
48. Wanner, W.; Strahler, A.H.; Hu, B.; Lewis, P.; Muller, J.P.; Li, X.; Barker Schaaf, C.L.; Barnsley, M.J. Global retrieval of bidirectional reflectance and albedo over land from EOS MODIS and MISR data: Theory and algorithm. *J. Geophys. Res.: Atmos.* **1997**, *102*, 17143–17161.
49. Grayson, R.B.; Western, A.W. Towards areal estimation of soil water content from point measurements: Time and space stability of mean response. *J. Hydrol.* **1998**, *207*, 68–82.
50. Jacobs, J.M.; Mohanty, B.P.; Hsu, E.-C.; Miller, D. SMEX02: Field scale variability, time stability and similarity of soil moisture. *Remote Sens. Environ.* **2004**, *92*, 436–446.
51. Bell, K.R.; Blanchard, B.J.; Schmugge, T.J.; Wiczak M.W. Analysis of surface moisture variations within large field sites. *Water Resour. Res.* **1980**, *16*, 796–810.
52. Owe, M.; Jones, E.B.; Schmugge, T.J. Soil moisture variation patterns observed in hand county, South Dakota. *JAWRA J. Am. Water Res. Assoc.* **1982**, *18*, 949–954.
53. Gilbert, R.O. *Statistical Methods for Environmental Pollution Monitoring*; Van Nostrand Reinhold Company, Inc.: New York, NY, USA, 1987; pp. 26–54.
54. Zhang, Y.; Liu, Y.D.; Zhao, J.I. Vector similarity measurement method. *Tech. Acoust.* **2009**, *28*, 532–536.
55. Neter, J.; Kutner, M.H.; Nachtsheim, C.J.; Wasserman, W. *Applied Linear Statistical Models*, 4th ed.; McGraw-Hill: Boston, MA, USA, 1996; p. 1391.