

Research Article

Time Domain Similarity of Lightweight Parameters Based Soil Respiration Sensor Network Deployment

Guoying Wang,^{1,2} Lufeng Mo,² Jizhong Zhao,¹ and Chaoqi Zhang³

¹School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China

²School of Information Engineering, Zhejiang Agriculture and Forestry University, Lin'an 311300, China

³Xi'an Thermal Power Research Institute Co., Ltd., Xi'an 710032, China

Correspondence should be addressed to Jizhong Zhao; zjz@xjtu.edu.cn

Received 2 November 2014; Accepted 20 March 2015

Academic Editor: Shaojie Tang

Copyright © 2015 Guoying Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Deployment is a basic and key issue for all kinds of wireless sensor networks applications. Most of existing researches on deployment problem of wireless sensor network are on generic network level and not for specific application scenarios and also do not utilize any domain knowledge of practical applications. Focusing on the problem of limited number of sensing nodes in soil respiration sensor networks, using domain knowledge such as some lightweight parameters (temperature, humidity, etc.) influencing soil respiration and soil respiration having a day-periodic trend, we proposed a deployment method, *TimSim*, for soil respiration sensor network based on time domain similarity of lightweight parameters. Lightweight parameters data from positions in the region to be monitored are collected before the deployment of soil respiration sensing nodes, and then time domain similarities of lightweight data among different positions are analyzed, according to which these positions are divided into some groups. A representative position in each group is chosen to deploy a soil respiration sensing node. The experimental results show that *TimSim* method can place nodes to proper positions so as to monitor regional soil respiration carbon flux effectively with a smaller estimation error than uniform and random deployment methods.

1. Introduction

Deployment strategy is very important for the application of wireless sensor network technology. On the one hand, it affects the connectivity and coverage performance of the entire network; on the other hand, it determines the number of nodes that the network needs. Good deployment strategy can maintain network performance and the monitoring accuracy using fewer nodes in the network, which can effectively reduce the cost of whole network when the nodes are expensive.

Soil respiration is one of the main ways of carbon exchange between the terrestrial ecosystem and atmosphere system and plays an important role in the global carbon cycle and carbon balance. CO₂ discharged to the atmosphere through soil respiration is about 68 PgC/yr, accounting for about 10% of the total atmospheric CO₂ [1]. Therefore, objectively evaluating the contribution of soil respiration

carbon flux to the global carbon budget and analyzing the influence of soil respiration on global climate change are especially important and have become core issues of the global carbon cycle.

As to the measurement of soil respiration flux of carbon, accurate soil carbon flux data of a single position can be achieved using some current equipment, for example, LI-8100 series of soil respiration measuring systems from LI-COR company [2]. However, due to the influence of a variety of factors such as soil pore size and type, soil temperature, soil water content, wind speed, and CO₂ concentration gradient, spatial and temporal heterogeneities of soil respiration carbon flux are very obvious [3], which makes existing methods hardly achieve accurate measurement of regional soil respiration carbon flux.

Wireless sensor networks have characteristics of wide monitoring range, sustainable monitoring, collaborative work, and transferring information in time and can meet

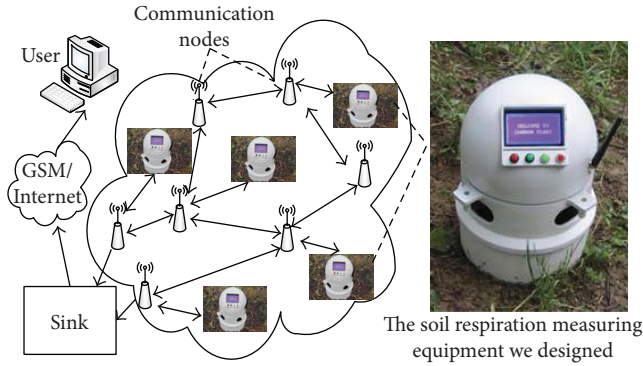


FIGURE 1: Wireless sensor network for soil respiration monitoring.

the requirements of regional soil respiration monitoring such as area coverage, persistent monitoring, and synchronous sampling. In order to apply wireless sensor networks to soil respiration monitoring, we have designed an instrument for soil respiration carbon flux measurement with communication ability using wireless sensor networks protocol, as is shown in Figure 1, which can be used as soil respiration sensing nodes [4].

But we encountered a problem during the practical construction process of soil respiration monitoring sensor networks; that is, the number of soil respiration sensing nodes is limited due to the high cost, which makes a well-designed deployment strategy become an urgent need to use fewer nodes monitoring a target region.

Although there have been a lot of research works aimed at nodes deployment issues in wireless sensor networks, most of them are to study the coverage of monitoring area and network connectivity issues. Furthermore, most of these studies are generic methods to solve the deployment problem and not for a specific application scenario. However, there is usually some specific domain knowledge in the relevant fields of each practical application, and using such domain knowledge may lead to noneligible promoting effect on the solution to the deployment problem when wireless sensor network technology is applied in practical fields.

Because soil respiration in a position has an approximately day-periodic variation trend [5], the analyzing result of the former sensing data can play a guiding role for the later measurement. The diel variation trend of a position is affected by the environment parameters, so positions with similar parameters may have similar time domain features. It is very possible that some positions in a region with the same vegetation have similar parameters. As a result, if the initial measurement data of different positions have similar time domain features, the same similarity will be maintained in the process of subsequent measurement.

Therefore, we designed a deployment strategy of soil respiration sensor networks based on time domain feature similarity of lightweight parameters, *TimSim*. At first, according to the domain knowledge of soil respiration, the soil respiration is described by some lightweight parameters. Then positions are divided into groups according to time domain similarity of lightweight parameters, and positions

in each group have similar trends of soil respiration. Then a representative position is chosen from each group, in which a soil respiration sensing node will be deployed. Because the time domain similarity exists, the soil respiration data of other positions in the same group can be obtained by a certain conversion of the soil respiration data from the deployed position. As a result, soil respiration data of the positions in the monitored region can be obtained by a small number of deployed nodes.

What should be emphasized is that although this method is put forward for a specific application of soil respiration monitoring, in fact the idea has universality and can be used for other applications with lightweight parameters, in which the monitoring data is periodic and sensing nodes costs are high. Therefore, the presented method can provide reference for solving problems during the application of wireless sensor network technology to other fields and promote wider application of wireless sensor networks.

The remaining parts of this paper are organized as follows. Section 2 introduces some previous work related to the subject this paper. Section 3 describes the nodes deployment problem in soil respiration monitoring sensor network to solve in this paper. Section 4 introduces in detail the proposed deployment strategy of soil respiration measurement nodes based on time domain similarity of lightweight parameters, *TimSim*. Section 5 verifies the performance of *TimSim* by experiments and compared with other two deployment methods. Finally, Section 6 carries on the summary and conclusions of the full paper and discusses the directions for further research.

2. Related Work

2.1. Deployment of Wireless Sensor Networks. Wireless sensor network technology can be used for many applications such as industry automation, localization and track tracing, health monitoring, and environment monitoring. Deployment is an important and basic issue for each of such applications. The concerns of the deployment problem include coverage, connectivity, and energy-efficiency.

The coverage problem of a wireless sensor network for different applications includes field coverage and barrier coverage, and the field may be two-dimensional area or three-dimensional space or surface. The basic deployment method for 2D area coverage is cell-based, and the shape of each cell may be square, triangle, hexagon, and so on. Ammari and Das [6] studied the coverage and connectivity of wireless sensor networks in 3D space using integrated concentric sphere model and continuum percolation theory, so as to determine the critical density of sensor nodes for the deployment and topology control. Liu and Ma [7] analyzed the expected coverage ratios on both regular and irregular 3D surface using stochastic sensors deployment method. For barrier coverage, there are line-based model [8] and curve-based model [9] for the deployment of wireless sensor networks. Furthermore, the coverage requirement may be k -coverage for redundancy in some applications. Ammari and Das [10] studied the node space density required for full k -coverage in an area so as to guide the deployment of sensor nodes. Li et al. [11] proposed

a field k -coverage autonomous deployment method when the count of nodes is not abundant with the consideration of load-balance, using the localization and mobility of nodes. Mahboubi et al. [12–14] studied the distributed deployment of mobile sensor networks using weighted Voronoi diagram model, in which nodes can move to sparse or empty area so as to fulfill better coverage.

The connectivity and energy-efficiency are also considered by many deployment methods besides coverage. If there are obstacles in the field, a serpentine movement policy can be used to handle obstacle [15]. The positions of relay nodes are important for the connectivity of a network. Lee and Younis [16] studied the selection and deployment problem of relay nodes, when some nodes failed and the network is separated into segments. The deployment of relay nodes can also be guided by the communication load of the network, so as to place more relay nodes in heavy load region [17]. Liu et al. [18] studied the deployment of sensor nodes using minimum cost when the required lifetime of network is given. Energy can also be saved through optimizing the distance of movement of sensor nodes in mobile sensor networks [19].

Compared with the previous work, the main novelty of *TimSim* lies in its different angle of view to solve the deployment problem of sensor nodes. These previous researches about the deployment of wireless sensor networks focusing on the coverage, connectivity, and energy-efficiency are on generic level and without specific application field. What we are considering in this paper is the utilization of domain knowledge of soil respiration, which is necessary and important.

2.2. Wireless Sensor Networks for Environment Monitoring.

There are many systems using wireless sensor network technology in the recent years. Barrenetxea et al. [20] established a wireless sensor network system for wild environment monitoring. GreenOrbs is a wireless sensor network system in forest for canopy closure sustainable estimation [21]. Besides, wireless sensor network technology is also used for volcano monitoring [22], wildlife monitoring [23], marine environment monitoring [24], soil property monitoring [25], and so on. In these systems, the focus is the application rather than deployment, so they do not have particular design for the optimization of the count and positions of sensor nodes.

Some environment monitoring applications use graph theory to optimize the positions of sensor nodes [26, 27]. However, these graph-based methods are generic and do not utilize any knowledge of applications. Du et al. [28] introduced a wireless sensor network system for wind distribution monitoring around an urban reservoir, in which sparse deployment is accomplished by adjusting distance of communication. However, it did not introduce how to determine the positions of sensor nodes.

Du et al. [29] segmented a year into different monsoon seasons according to the analysis of historical wind distribution data of the field and used computational fluid dynamics to learn the spatial correlation of wind in the presence of surrounding buildings. The most informative locations are the outputs of the deployment method. This method uses knowledge of application, which is similar to our method.

However it cannot be used in soil respiration monitoring sensor network, because we cannot have historical soil respiration data of a monitored field and computational fluid dynamics cannot be used for soil respiration.

As to the previous works, the deployment methods of them are all different from the method based on similarities of lightweight parameters, which is to be studied in this paper.

3. Problem Descriptions

3.1. Soil Respiration Sensor Network. In order to monitor soil respiration carbon flux data in a region, traditional methods are to carry out measurements in one or a few sampling positions using devices such as LI-8100 and then estimate soil respiration carbon fluxes across the region. But the results of such methods using a small number of measuring positions are very inaccurate. On the one hand, for the spatial heterogeneity of soil respiration, using a small amount of measuring positions to represent other positions in the region as a whole is not appropriate [30]; on the other hand, soil respiration varies over time and the result measured at a certain moment cannot represent the measurements of other moments [31]. So the traditional measuring methods are not representative in both time and space aspects, and regional soil respiration carbon flux cannot be accurately estimated using traditional methods.

Therefore accurate measurement and estimation of regional soil respiration should meet the following three requirements. (1) Measurements in multiple positions: due to the spatial heterogeneity of soil respiration, only measurement results of multiposition can describe the entire regional soil respiration accurately. (2) Measurements at the same time in multiple positions: because of the time heterogeneity of soil respiration, measurements in the same position at different time may be different, so only measurements at the same time in multiple positions can describe the regional soil respiration at the measuring time. (3) Continuous measurements: for the temporal heterogeneity of soil respiration, few measurements cannot reflect the details of the soil respiration varying with time, and only continuous measurements can reflect the varying details of soil respiration along with time.

Wireless sensor network technology can well meet the three aspects of requirements of measurements: multiple positions, synchronization, and continuation. So we adopt this technology to carry out the measurement and estimation of regional soil respiration carbon flux. During the construction of a soil respiration sensor network, because the soil respiration measuring equipment is expensive which leads to limited quantity of the equipment, we add some routing nodes which bear only communication tasks and are not responsible for the measurement task to guarantee the connectivity of the network, as shown in Figure 1. In order to save energy of these nodes, soil respiration measuring nodes act as leaf nodes in the data collecting tree, and measured data will be transferred to the sink node by routing nodes. In addition, we have studied segmental dynamic sampling strategy to reduce energy consumption of soil respiration measuring nodes and prolong the continuous working time of the whole system in our previous study [32].

3.2. Deployment of Soil Respiration Sensing Nodes. As mentioned above, a soil respiration sensor network contains some routing nodes to guarantee the connectivity of the network, so the deployment positions of soil respiration nodes mainly consider the requirements of application layer rather than network layer. At the application layer, in order to effectively measure regional soil respiration using a limited number of soil respiration sensing nodes, each node should be deployed in a more reasonable position.

Deploying a soil respiration sensor network in the region Z , in addition to the sink node, there are N soil respiration sensing nodes $S = \{s_1, s_2, \dots, s_N\}$ and some routing nodes R . Soil respiration nodes S are deployed in $P_s = \{p_{s_1}, p_{s_2}, \dots, p_{s_N}\}$, and $p_{s_i} = (x_i, y_i)$, ($1 \leq i \leq N$) represents the physical position where the soil respiration measuring node s_i deploys. The number and positions of routing nodes R depend on the quantity and location of soil respiration nodes and are determined according to the connectivity and robustness requirements, which is not in the scope of this paper.

The goal of this paper is to determine small N and appropriate P_s , according to which soil respiration sensing data $SR = \{sr_1, sr_2, \dots, sr_N\}$ are obtained by sensing nodes $S = \{s_1, s_2, \dots, s_N\}$ in the positions P_s . The errors between spatial interpolation results $SR' = \{sr_1, sr_2, \dots, sr_N, sr_{N+1}, sr_{N+2}, \dots, sr_{N'}\}$ and the real soil respiration data in N' positions, $E = \{e_1, e_2, \dots, e_N\}$, are computed and small mean error and small error variance are preferred.

4. Our Solution

4.1. Main Idea. In order to solve the problem of reasonable layout of sampling positions, a relatively easy way is to arrange more sampling positions in subregions with more drastic changes of soil respiration so as to get more details of the changes and to choose fewer sampling positions in subregions with relatively flat changes of soil respiration. This approach is based on the spatial correlation of soil respiration, which means each measurement in a position represents the data in surrounding area, and the correlation can be exploited by spatial interpolation. As a result, the distribution state of sensing data in the entire region can be obtained by interpolation using a small number of measurement points.

The method above only pays attention to the local spatial correlation of soil respiration, the main consideration of which is the relationship of soil respiration of a position and that of positions nearby it. But the approach does not take into account the correlation between soil respiration of positions which are not adjacent. In addition to the correlation among adjacent positions which are naturally used in the spatial interpolation process, the correlation of soil respiration data measured in nonadjacent positions may play a significant role in promoting the rationality of the layout of soil respiration sensor network, which is the focus of this paper.

In this paper we designed a soil respiration nodes deployment method exploiting the correlation of soil respiration data in nonadjacent positions. According to the similarity of time domain features of soil respiration data in different positions of monitored area, these positions

are divided into groups regardless of whether positions in each group are adjacent or not in physical space, and thus soil respiration data monitored in one position from each group can represent the remaining positions of the same group for the time domain similarity. During the monitoring stage, once soil respiration data are measured in the chosen sampling position of each group, soil respiration data of other positions in the same group can be obtained through certain transformation relations which are found during the grouping process. As a result, we can have soil respiration data of positions much more than sensing nodes actually deployed before final interpolation to simulate the real state of regional soil respiration, which makes the interpolation result closer to the real world.

However, obtaining the time domain similarities of soil respiration data of all positions in the monitored area is very difficult, since it needs fine-grained soil respiration data of each position in the monitored area for a period of time, which is the ultimate goal of soil respiration monitoring sensor network, and can only be realized through dense and regular deployment of soil respiration measuring nodes. In fact, it is not realistic to densely deploy soil respiration measuring nodes in the monitored area, because we cannot have enough soil respiration measuring nodes usually due to expensive cost, which is just the root cause of what to be studied in this paper.

To solve the problem above, we introduced the temporal similarity of lightweight parameters into the solution. According to the domain knowledge of soil respiration, soil respiration sensing nodes use open chamber method to measure the soil respiratory carbon flux, and the calculation is shown by [4]

$$C_{\text{flux}} = \frac{100VP_0(1 - W_0/1000)}{rs(T_0 + 273.15)} \frac{\partial C'}{\partial t}. \quad (1)$$

In the equation, C_{flux} means soil respiration carbon flux, V means chamber volume (cm^3), P_0 means initial pressure in the chamber (kPa), W_0 means initial water vapor mole fraction ($\text{mmol} \cdot \text{mol}^{-1}$), r means ideal gas constant $8.314472 \text{ m}^3 \cdot \text{Pa} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$, s means soil surface area (cm^2), T_0 means initial temperature ($^\circ\text{C}$), and $\partial C'/\partial t$ means initial changing rate of CO_2 ($\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$).

From (1) we can see that soil respiration is affected by temperature, humidity, pressure, and other factors. Because price and energy consumption of sensors for these parameters are relatively low, we refer to these factors such as temperature, humidity, and pressure as “lightweight parameters” of soil respiration in this paper.

In order to solve the problem that the similarities of time domain features of fine-grained positional soil respiration data in the monitored area cannot be obtained directly, we converted the similarities into the similarities of time domain features of lightweight parameters in different positions. Because of the low cost and low energy consumption of lightweight parameter sensing nodes, lightweight parameters can be measured densely and regularly in the monitored area and the similarities of time domain features of lightweight parameters in different positions can be obtained.

Therefore, we proposed a strategy for the deployment of soil respiration monitoring sensor network based on similarity of time domain features of lightweight parameters, which is called *TimSim* in this paper. Firstly densely and regularly deploy lightweight parameters sensing nodes in the monitored area and measure for a period of time, and then analyze the measurement results to observe the similarities of time domain features of lightweight parameters, group the positions according to the similarities, and finally choose a representative position for each group to deploy a soil respiration measurement node.

What should be emphasized is that because deploying a soil respiration measuring node requires some complicated special processes such as smoothing the ground surface and embedding the device bottom into the soil, once the layout scheme of sensor nodes is determined, it is usually assumed that soil respiration sensing nodes in the monitoring period will not be moved anymore after deployment. Therefore, these complex preparatory works including the measurement, analysis, and computation of lightweight parameters are very necessary for the optimal layout scheme of sensor nodes.

4.2. Temporal Similarity of Data in Different Positions. In order to group similar positions in the monitored area using the time domain features of the data, the similarity of time domain features of data between different positions need be defined.

Let $D_p = \{d_{p_i} \mid 1 \leq i \leq l\}$ and $D_q = \{d_{q_i} \mid 1 \leq i \leq l\}$ denote two time series of data measured in positions p and q , respectively, in which l means the length of the data series. In order to describe the similarity between D_p and D_q , they are regarded as the two l dimensional vectors, and then the cosine value of angle θ between these two l dimensional vectors can be used as the similarity metrics. So we define the temporal similarity of data in the two positions of the p and q , $\text{Sim}(p, q)$, as

$$\begin{aligned} \text{Sim}(p, q) &= \cos(\theta) = \frac{D_p \cdot D_q}{\|D_p\| \|D_q\|} \\ &= \frac{\sum_{i=1}^l (d_{p_i} \times d_{q_i})}{\sqrt{\sum_{i=1}^l d_{p_i}^2} \times \sqrt{\sum_{i=1}^l d_{q_i}^2}}. \end{aligned} \quad (2)$$

Because all the data in this application, either soil respiration data or lightweight parameters data, are all positive, the components of D_p and D_q are not negative, and hence the value of $\text{Sim}(p, q)$ is in the zone $[0, 1]$. Obviously, if the directions of D_p and D_q are identical, that is, there exists the same linear relationship between D_p and D_q components, the similarity between D_p and D_q is the best, and $\text{Sim}(p, q)$ equals 1. On the contrary, if the two vectors, D_p and D_q , are orthogonal to each other, the similarity is minimum and $\text{Sim}(p, q)$ equals 0.

4.3. Grouping Positions according to Temporal Similarity. After the temporal similarity of data between two positions is defined, we can divide multiple positions into different

groups according to the similarities between each other, and the grouping result is that positions in each group have high similarities of time domain features of data while there are less similarities among positions from different groups. There are two cases about the grouping problem in practice, one of which is that the number of groups, k , is determined and the other is that the lower bound of similarity among positions in the same group, LB_{Sim} , is given.

When the number of groups k is identified, which means the count of soil respiration sensing nodes is limited, we can use clustering algorithms based on partition, for example, k -means, to divide n positions $P = \{p_i \mid 1 \leq i \leq n\}$ into k groups $G = \{g_i \mid 1 \leq i \leq k\}$. The stop condition of clustering procedure is that the differences among average similarities of positions in every group g_i ($1 \leq i \leq k$) are below a preset threshold. Obviously, the bigger the k is, the more the groups are divided, the fewer the positions are in a group, the smaller the θ_i is, and the higher the similarities among positions in a group are.

When the lower bound of group similarity LB_{Sim} is given, which means the count of soil respiration sensing nodes is relatively abundant and what we are pursuing is the accuracy of results, we can use clustering algorithm based on hierarchy, such as *ROCK* [33], to divide n positions $P = \{p_i \mid 1 \leq i \leq n\}$ into k' groups $G' = \{g_i \mid 1 \leq i \leq k'\}$. Minimum similarity in each group g_i ($1 \leq i \leq k'$), Sim_i , is bigger than the preset similarity threshold LB_{Sim} . Obviously, the bigger the LB_{Sim} is, the higher the similarities of positions in each group are, the fewer the positions are in a group, the more the group number k' is, and the more the sensing nodes are needed.

After the grouping scheme of positions in the monitored area is determined and the grouping result is achieved, we choose the position where the data vector has the highest similarity with the center vector of all data vectors in the same group as the position to deploy a soil respiration sensing node. When the deploying positions in all groups are chosen, the positions of routing nodes R can be determined according to the connectivity requirement of soil respiration sensing nodes.

One position in each group is selected as a representative for the soil respiration sampling. Therefore, if the number of soil respiration measuring nodes to deploy is certain, we use the case of determined k , the number of groups. Otherwise, the situation is corresponding to the case of similarity lower bound LB_{Sim} which is identified. Higher LB_{Sim} leads to not only higher monitored accuracy of regional soil respiration, but also bigger group count k' which needs more soil respiration sensing nodes, that is, higher network cost, so an appropriate LB_{Sim} should keep a compromise between accuracy of results and cost of network.

4.4. Reconstruction of Data for Positions of a Group. A following problem is the reconstruction of data for the rest of positions using the data of the sampling position in a group. Data reconstruction should be done after the deployment and during the working stage of soil respiration sensor network system, and it seems that this issue is not a part of the deployment of a soil respiration sensor network. But in fact, as for the proposed deployment strategy based on time domain

similarity of data, its main idea is taking advantage of the correlation of data from nonadjacent positions, so we should clarify that the method is able to rebuild the same correlation in data before putting the deployment strategy into practice. If the data correlation among nonadjacent positions used by the deployment strategy cannot be reconstructed during the network works, this deployment method will be not able to be applied into any applications.

The reconstruction problem of data of each position in a group is described as follows. Assume that a position group g contains m positions $P_g = \{p_i \mid 1 \leq i \leq m\}$ and the sampling position in the group is gc , ($gc \in P_g$), whose data vector is the closest to the centroid of vectors of all group member positions in the grouping result. Before the grouping process, the temporal data sequence measured in position p_i is $D_{p_i} = \{d_{p_i,j} \mid 1 \leq i \leq m, 1 \leq j \leq l\}$ and data sequence measured in position gc is $D_{gc} = \{d_{gc,i} \mid 1 \leq i \leq l\}$. After the deployment of whole sensor network (i.e., only in position gc a soil respiration sensing node is deployed in this group), we obtain data d'_{gc} measured in position gc . The data reconstruction problem we need solve is how to build data for the rest of $m - 1$ positions in group g using the data d'_{gc} .

To fulfill this task, we need compute and save the conversion relations between the data vectors of the rest of $m - 1$ positions in the group and that of gc during the grouping procedure. These positions are grouped into a group because the similarities of temporal features of data vectors in these positions are high; that is, the directions of the data vectors are closer and can be considered as the same direction approximately, which means that the ratio relationships among data of these positions at the same time remain still. The norm of data vector of position p_i is $\|D_{p_i}\|$ and that of gc is $\|D_{gc}\|$. If we calculate and save the ratio between these two norms $\text{ratio}_i = \|D_{p_i}\|/\|D_{gc}\|$, the data of position p_i at the same time to time d'_{gc} was measured in position gc , d'_{p_i} , can be estimated by $d'_{p_i} = d'_{gc} \times \text{ratio}_i$.

During the analysis of premeasured data for grouping, record the ratios Ratio between data vector of every nodes p_i in group P_g and that of sampling position gc :

$$\text{Ratio} = \{\text{ratio}_i\} = \left\{ \frac{\|D_{p_i}\|}{\|D_{gc}\|} \right\}, \quad 1 \leq i \leq m. \quad (3)$$

Then data of all positions in the group can be reconstructed using Ratio after data d'_{gc} are measured in position gc after the soil respiration measuring node is deployed in position gc and during the persistent practical monitoring procedure:

$$D'_{p_g} = \{d'_{p_i}\} = \text{Ratio} \cdot \|d'_{gc}\|, \quad 1 \leq i \leq m. \quad (4)$$

But in the approach we designed in this paper, the pre-measured data is lightweight parameters data rather than soil respiration data, and the ratio vector Ratio is the relation of lightweight parameters data of each position within the group rather than that of soil respiration data. So this ratio vector Ratio can only be used to reconstruct lightweight parameters data and cannot be directly used for the reconstruction of soil respiration data.

Because the soil respiration carbon flux is calculated according to (1), the lightweight parameters such as temperature and humidity measured in position pc by the soil respiration sensing device can be used to calculate the soil respiration carbon flux on the one hand; on the other hand, these lightweight parameters data in position pc can be used to get lightweight parameters data in other positions of the same group using (4), and then the soil respiration data in these positions can be calculated according to (1).

4.5. TimSim Algorithm. Now we formally describe the deployment strategy for soil respiration monitoring sensor network based on time domain similarity of lightweight parameters, *TimSim*, proposed in this paper.

Algorithm 1 requires the following: P : positions uniformly and densely distributed in the monitored area; D_p : lightweight parameters data serials in these positions P ; k or LB_{Sim} : limiting conditions required for grouping in two cases, k for count limit of groups and LB_{Sim} for lower bound of similarity in groups. In the process of the algorithm, firstly, outliers in data are rectified and other preprocessing operations are done, and then similarities of temporal features of every two data sequences in different positions are calculated. In the grouping process, positions in P are divided into k groups using k -means algorithm if parameter k is provided or divided into k' groups using *ROCK* algorithm if parameter LB_{Sim} is provided. For each group of positions, the centroid vector of lightweight parameters data vectors is calculated, and choose the position where lightweight parameters' data is the closest to the centroid vector as the representative position of the group, gc_i , and then the ratios between data vectors in every other position and data vector in the representative position, ratio_i , are calculated and recorded for the data reconstruction of subsequent monitoring process later.

As for the soil respiration sensing network, the length of time lightweight parameter data sequence D_p covered is usually 24 hours, considering the soil respiration has day-periodic trends. Furthermore, to avoid the occasionality of data sequence, several data sequences of different days should be used as the basis of grouping operation after computing the average data sequence with time-correspondence.

5. Experiments and Analysis

5.1. Experimental Data. In order to verify the proposed deployment strategy of soil respiration monitoring sensor network based on similarity of time domain features of lightweight parameters, *TimSim*, the ideal experiment procedure should be as follows: densely and regularly deploy lightweight parameters sensing nodes in the monitored area Z and carry out a few days of continuous measurement, and then divide the positions into groups based on the lightweight parameters data sequences measured, and find the positions closest to the centroid of light weight data sequences to deploy soil respiration sensing nodes. In order to examine the performance of this method, we need collect soil respiration data in positions where soil respiration sensing nodes deployed and then do spatial interpolation with them

Require: $P = \{p_i \mid 1 \leq i \leq n\}$ // n positions regularly and densely distributed in the monitored area.
 $D_P = \{D_{p_i} \mid 1 \leq i \leq n\} = \{\{d_{p_{ij}} \mid 1 \leq j \leq l\} \mid 1 \leq i \leq n\}$ // lightweight data sequence from each position
 k // limited number of soil respiration sensing nodes
 LB_{Sim} // limited lower bound of similarity in each group

Ensure: $G = \{g_i \mid \cup g_i = P, \cap g_i = \emptyset, 1 \leq i \leq k \text{ (or } k')\}$ // result of positions group
 $GC = \{gc_i \mid gc_i \in g_i, 1 \leq i \leq k \text{ (or } k')\}$ // gc_i is the position to place a node in group g_i
 $Ratio = \{gratio_i = \{ratio_j \mid p_j \in g_i\} \mid 1 \leq i \leq k \text{ (or } k')\}$ // data ratio between positions to node position

- (1) preprocess(D_P); // rectify the outlier in data
- (2) **for** $i = 1$ to n **do**
- (3) **for** $j = 1$ to n **do**
- (4) $Sim(p_i, p_j) = (D_{p_i} \cdot D_{p_j}) / (\|D_{p_i}\| \|D_{p_j}\|)$;
- (5) **end for**
- (6) **end for**
- (7) $G = \{g_i \mid 1 \leq i \leq k\} = k\text{-means}(P, Sim, k)$; // for the case of k is given
or $G = \{g_i \mid 1 \leq i \leq k'\} = ROCK(P, Sim, LB_{Sim})$; // for the case of LB_{Sim} is given
- (8) **for** $i = 1$ to k (or k') **do**
- (9) $Dc_i = \text{centroid}(\{D_{p_j} \mid p_j \in g_i, \cup \{p_j\} = g_i\})$; // centroid of data vectors from all positions in g_i
- (10) $Dist_i = \{\text{dist}_j = 1/Sim(D_{p_j}, Dc_i) \mid p_j \in g_i, \cup \{p_j\} = g_i\}$; // distance to centroid data vector
- (11) $gc_i = \{p_j \mid \text{dist}_j == \text{MIN}(\{Dist_i\}), p_j \in g_i\}$; // position with nearest data vector to centroid
- (12) $gratio_i = \{\|D_{p_j}\| / \|D_{gc_i}\| \mid p_j \in g_i, \cup \{p_j\} = g_i\}$;
- (13) **end for**
- (14) **return** $G, GC = \{gc_i\}, Ratio = \{gratio_i\}$;

ALGORITHM 1: Time domain similarity of lightweight parameters based deployment algorithm (*TimSim*).

to estimate regional soil respiration of the monitored area, SR_{TimSim} . The estimated results should be compared with the real soil respiration data SR_{real} of the monitored area at the same time to analyze the measuring error of the method. But the real soil respiration data of the area, SR_{real} , cannot be obtained, because this requires to deploy many soil respiration measuring equipments very densely throughout the area, while we cannot have so many soil respiration measuring equipments for the high cost.

We adopt a substitutional method to simulate real soil respiration data of many dense positions in the monitored area using temporal and spatial shift. We deploy x soil respiration sensing nodes in x different positions of the experimental area and let them keep working for d days to obtain data of soil respiration and lightweight parameters. Thus d serials of daily soil respiration data can be collected by the equipment in each position in the d days. We shift the d daily serials of data of a position to d different positions for one day. So in order to simulate real data of one day in n positions of the experimental area, SR_{real} , such x equipment need keep working for n/x days.

Taking into account that there are several factors affecting soil respiration such as temperature, humidity, and air pressure, too many lightweight parameters may complicate the experiment. We carry out the experiment process in smooth lawn with uniform illumination, in which the temperature and pressure of different positions are almost the same; thus humidity is the only main factor affecting soil respiration in this area. So the lightweight parameter in our experiment only remains humidity data.

In the experiment, we used 20 soil respiration measuring equipments, deployed them in different positions of

the monitored area, let them keep working for 5 days, and obtained 100 position-day serials of data. Then we changed the positions of such equipment and repeated the data collection process for 2 times. As a result, we obtained 3 data sets, each of which includes 100 position-day serials of data, and we distributed each data set regularly with 10×10 grids pattern in the monitored area as a set of experimental data, SR_{real} .

5.2. Experimental Scheme. Suppose that we have measured regional soil respiration real data SR_{real} ; then we can evaluate a specific deployment strategy using the deviation of $SR_{strategy}$, the measured results of soil respiration data using the strategy, from SR_{real} at each position p_i :

$$E = \{e_i\} = \left\{ \left| SR_{strategy_{p_i}} - SR_{real_{p_i}} \right| \mid 1 \leq i \leq n \right\}. \quad (5)$$

Once we get the error between $SR_{strategy}$ and SR_{real} , e_i , in position p_i , we can use mean error and error variance to analyze the performance of a deployment strategy. The mean error indicates an overall performance of a deployment scheme, and the error variance is helpful to analyze whether the errors of the different positions are close and to analyze the spatial distribution of error combining e_i and the location of p_i . Of course we expect that our method, *TimSim*, may have small mean error and small error variance.

In order to evaluate the performance of *TimSim*, we use the scheme in Algorithm 2 to carry out the experiment. As shown in Algorithm 2, we use multiple sets of the experimental data to calculate the error of regional soil respiration carbon flux measured using *TimSim*, *Uniform*, and *Random* deployment strategies.

Require: $P = \{p_i \mid 1 \leq i \leq n\}$ // n positions regularly and densely distributed in the monitored area.
 $D_P[] = \{\{D_{p_i} = \{d_{p_{ij}} \mid 1 \leq j \leq l\} \mid 1 \leq i \leq n\}_t \mid 1 \leq t \leq \text{times}\}$ // times sets of lightweight data sequences of all positions
 $\text{SR}_{\text{real}}[] = \{\{\text{sr}_{\text{real}_i} \mid 1 \leq i \leq n\}_t \mid 1 \leq t \leq \text{times}\}$ // times sets of data for experiment

Ensure: $\text{Err}[]$ // error for each strategy.

- (1) $\text{times} = 3, n = 100, N_s = \{6, 9, 12, 15\}$; // Initialization
- (2) $\text{STRATEGY} = \text{TimSim}, \text{Uniform}, \text{Random}$; // using three strategies respectively
- (3) **for** each strategy $\in \text{STRATEGY}$ **do**
- (4) **for** each $n_s \in N_s$ **do**
- (5) **for** $t = 1$ to times **do**
- (6) $P_{\text{sample}} = \text{strategy}(P, D_P[t]) = \{p_i \mid 1 \leq i \leq n_s\}$; // determining n_s sample positions using strategy
- (7) $\text{SR}_{\text{sample}} = \text{sample}(\text{SR}_{\text{real}}[t], P_{\text{sample}}) = \{\text{sr}_{p_i} \mid p_i \in P_{\text{sample}}, 1 \leq i \leq n_s\}$; // sampling n_s data from $\text{SR}_{\text{real}}[t]$ according to P_{sample}
- (8) $\text{SR}_{\text{strategy}} = \{\text{sr}_{p_i} \mid 1 \leq i \leq n\} = \text{Kriging}(\text{SR}_{\text{sample}}, P_{\text{sample}}, P)$ // interpolation using Kriging method
- (9) $E[\text{strategy}, n_s, t] = \{e_i \mid 1 \leq i \leq n\} = |\text{SR}_{\text{strategy}} - \text{SR}_{\text{real}}[t]|$
- (10) **end for**
- (11) **end for**
- (12) $\text{Err}[\text{strategy}, n_s] = (1/\text{times}) \sum_{t=1}^{\text{times}} E[\text{strategy}, n_s, t]$;
- (13) **end for**
- (14) **return** Err ;

ALGORITHM 2: Experiment procedure.

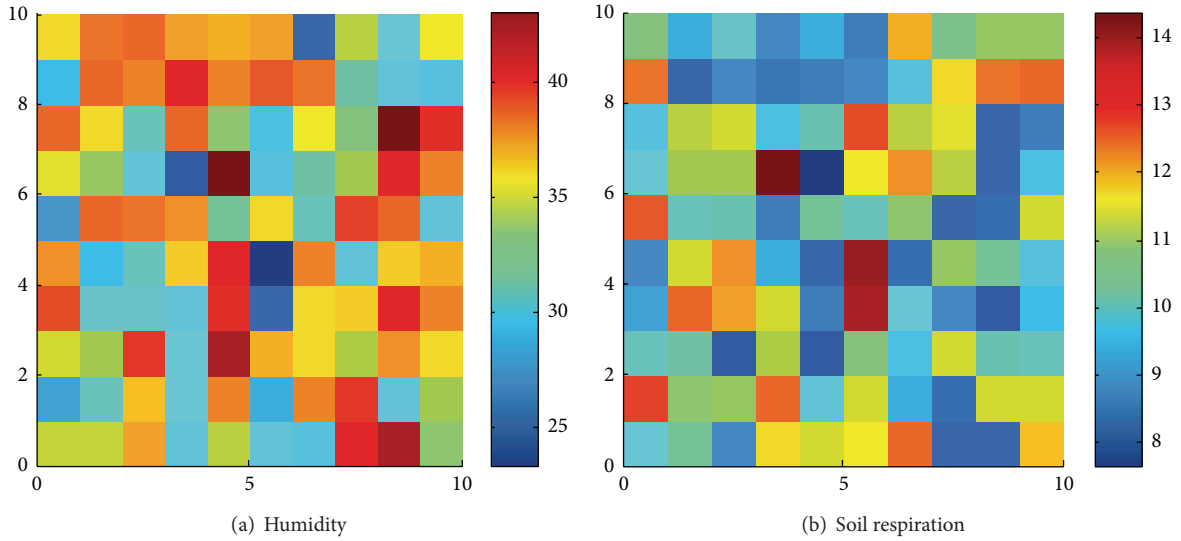


FIGURE 2: Distribution of soil humidity and soil respiration carbon flux in the experimental area.

5.3. Experimental Results and Analysis

5.3.1. Deployment Results and Performance Analysis of TimSim. In order to demonstrate the procedure and result of *TimSim* deployment method, we choose a set of experimental data, including lightweight parameter data, humidity, and soil respiration carbon flux data of 100 positions, and the data distribution of these positions at a moment is shown in Figure 2.

Then the similarities of time domain features of lightweight parameter data are calculated using the data set, according to which the 100 positions are divided into groups using k -means algorithm. Figure 3 shows the grouping results of k -means algorithm for the case when the count k of groups

is 6, 9, 12, and 15, respectively, in which positions marked with the same number are in the same group. Furthermore, the position to deploy a soil respiration sensing node in each group is marked with a small square around the number, whose data vector of lightweight parameter is the closest to the centroid of data vectors of all positions in the same group.

As we can see from Figure 3, there is not any obvious regular pattern in the spatial distribution of positions in each group. Positions in each group are not all adjacent in the space and their distribution is very disperse and irregular. When comparing the lightweight parameter data of positions in the same group, their values are not very close to each other at the same time. This is because the grouping method is based on the similarities of temporal pattern of data rather than the distribution of data at a certain moment.

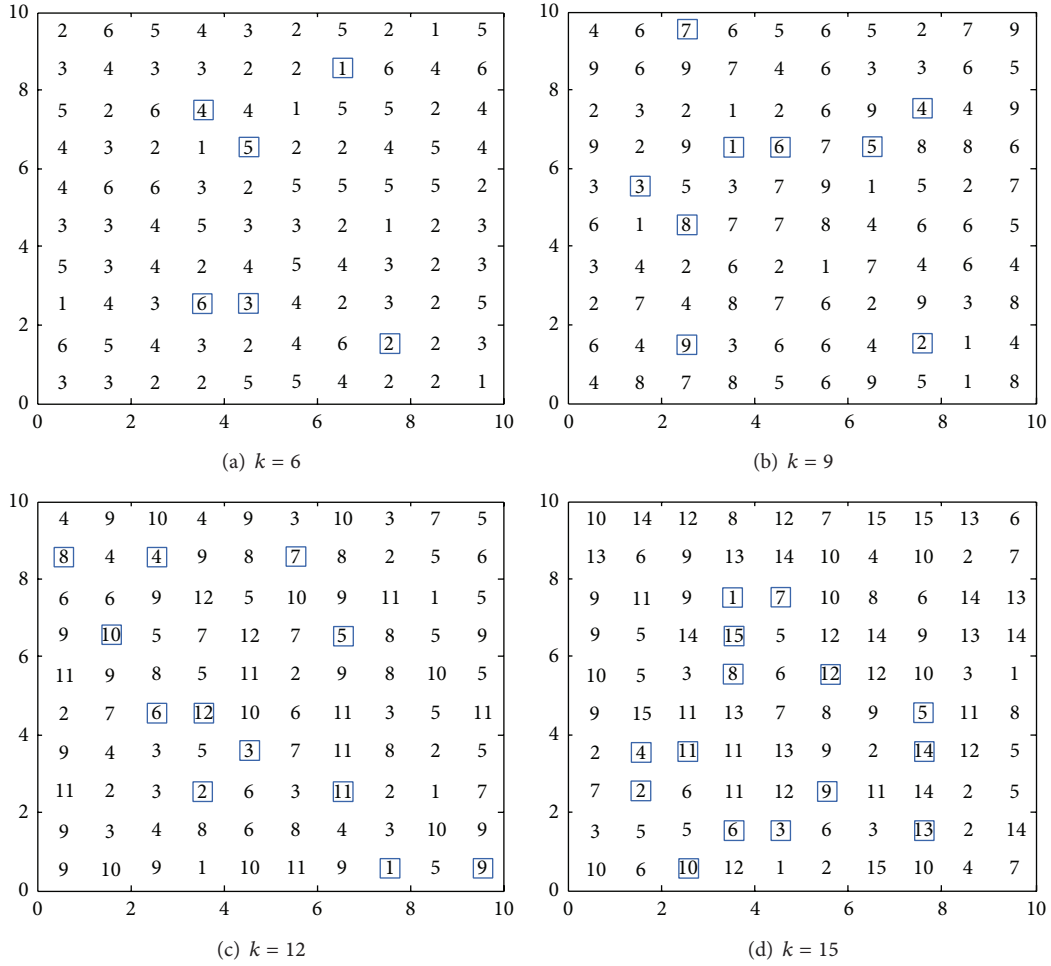


FIGURE 3: Position grouping results for different group numbers.

As we can also see from Figure 3, the selected deployment positions of soil respiration sensing nodes are not in the geographical central positions of each group, because in our *TimSim* method the deployment position is determined by the centroid direction of data sequences in positions of a group, and the selected position to deploy soil respiration sensing node is the position whose direction of data vector is the closest to the centroid direction among all directions of data vectors in positions of the same group. So the selected deployment position is independent of the geographical distribution of positions of the same group.

After the deployment positions are determined, the distribution of error between SR_{TimSim} and SR_{real} can be obtained, which are shown in Figure 4. Figure 4 shows the distribution of average value of errors in each position using SR_{TimSim} method for 3 times, as described in Algorithm 2, when group number k is 6, 9, 12, and 15, respectively. And the mean errors and error variances of all these cases are shown in Figure 5.

From Figure 4 we can see that errors vary within a wide range when the count of sampling positions is small and there are large errors in some positions. For example, the biggest error among all positions reaches 4.8 in Figure 4(a). The reason is that the group count is low, and there are

weak similarities between data vectors of some positions and that of the deployment positions in the same group, which brings big errors during the reconstruction process of data in these positions with weak similarities according to centroid vector. For the same reason, the mean and variance of errors of different positions are also relatively big when the group count is low, as shown in Figure 5.

With the increase of the number of sampling positions, the mean and variance of the errors decrease, as shown in Figure 5. From Figure 4 we can see that errors of most positions decrease with increasing the number of sample points with also some exceptional positions, which is caused by the alternation of directions of centroid vectors with the change of count of groups. Grouping results alter with the growth of count of groups, which makes the data vectors of some positions farther from the centroid vector directions than the former case leading to bigger error in these positions. While with the growth of count of groups, quantity of positions within a group is reduced at large, the similarities in groups become better and there are no positions with very large error. As shown in Figure 4(d), the maximum error among all positions is only 0.9.

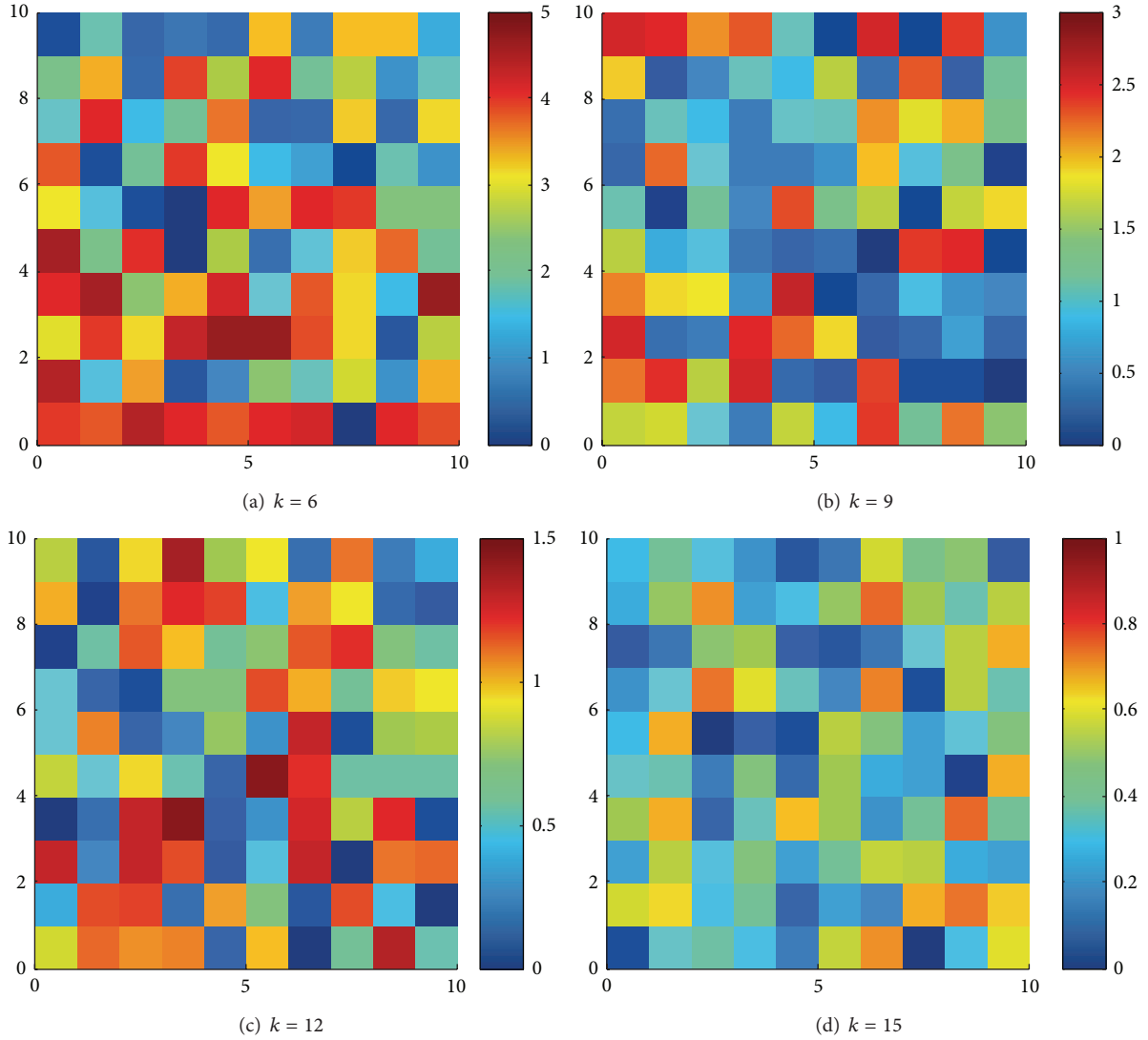


FIGURE 4: Error distribution of *TimSim* for different group numbers.

5.3.2. Comparison of *TimSim* and Other Methods. In order to show the advantage of the *TimSim* method, we also evaluate other two deployment strategies, uniform deployment (*Uniform*) and random deployment (*Random*), using the same experimental data in our experiment as described in Algorithm 2, as shown in Figure 5. The experimental results show that the performance of *TimSim* is better than those of the other two strategies.

As can be seen from Figure 5, when the number of sampling positions is low, the mean errors of *Uniform* deployment and *TimSim* deployment are close to each other and are 3.2 and 3, respectively, and both of them have significantly better performance than random deployment, whose mean error is 3.9. The reason is that both *Uniform* deployment and *TimSim* can make the sampling positions cover the entire area well through spatial interpolation, while the coverage of *Random* deployment strategy is seriously inadequate due to the small number of nodes, resulting in large errors in the positions of uncovered area.

In addition, the error variance of *TimSim* has obvious advantages compared to those of both *Uniform* and *Random* deployment strategies in the case of few sampling nodes and the reason is that *TimSim* strategy takes the correlation among nonadjacent positions into account, which makes no positions have too large error. The other two strategies use the spatial correlations of soil respiration during the interpolation operations while the correlations may be not the actual case, so there may be large local deviations in some positions where nodes are sparse, which makes the errors relatively dispersed and leads to large error variances.

When the count of sampling positions is high, *TimSim* deployment strategy has obvious advantage in the aspect of either mean error or error variance compared with *Uniform* and *Random* deployment. When the number of sampling positions is up to 15, the coverage of *Random* deployment increases and the mean error was about to catch up with the *Uniform* deployment while the error variance is relatively high due to the inhomogeneity of deployment positions.

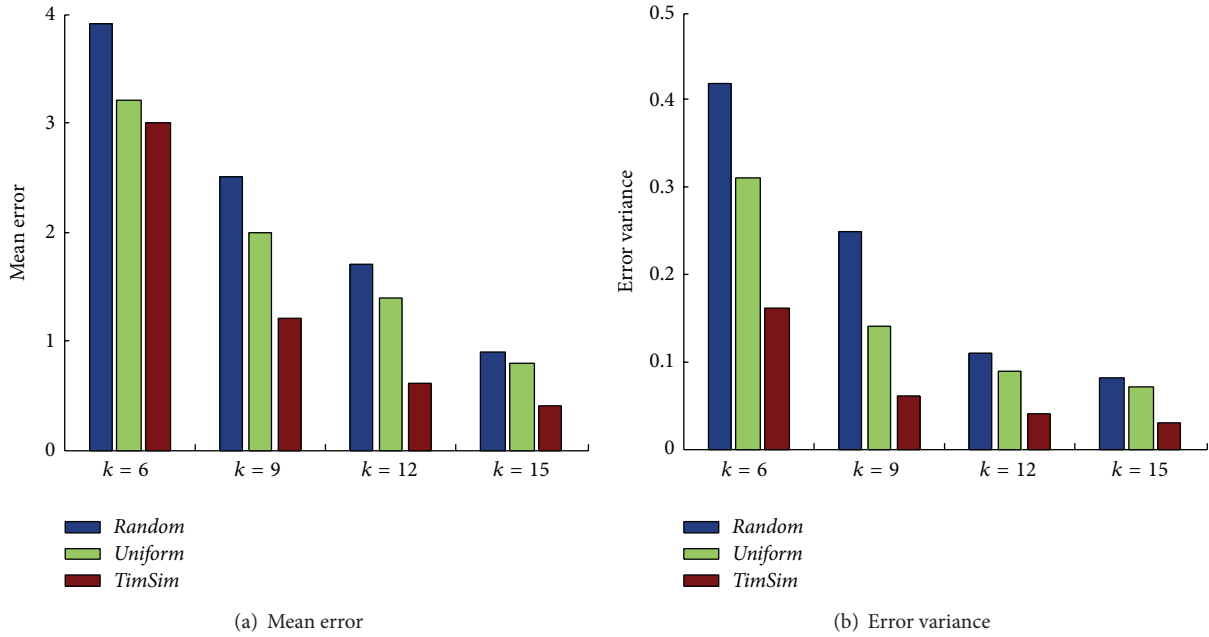


FIGURE 5: Error comparison of three deployment strategies.

With the increase of deployment positions, number of groups increases in the *TimSim* deployment method and number of positions included in each group is reduced, which makes the similarities of temporal features of lightweight parameters in positions of the same group become high. As a result, the data in positions can be reconstructed with small error using data in the deployment position, which means that real data measured in 15 positions can be extended to approximate real data far more than 15 positions, so both mean error and error variance of *TimSim* deployment method are smaller than the other two deployment methods when the number of sampling positions is large to a certain extent.

6. Conclusions

Focusing on the problem of limited number of nodes for its expensive cost in soil respiration sensor networks, we proposed a new deployment strategy based on the time domain similarity of lightweight parameters, *TimSim*. This method is inspired by the knowledge that soil respiration is mainly affected by some lightweight parameters such as temperature and humidity. Using this method, the count of sampling positions can be reduced while the accuracy of results remains by proper arrangement of the node positions. The performance of *TimSim* is examined by experiments, and the results show that *TimSim* can measure regional soil respiration carbon flux effectively and can achieve smaller deviation than uniform or random deployment when using the same amount of measuring equipment.

In the proposed method, lightweight parameters data need to be measured densely and regularly in advance, which is infeasible when the area to be monitored is very large. In this case lightweight parameters can be relatively sparsely

measured and then be interpolated to increase the density. Although this can reduce the accuracy of the final result, it is a feasible way anyway.

The method proposed in this paper is more convenient to be used for single factor than multiple factors of lightweight parameters, and the experiment in the paper also emphasized the guidance of a single lightweight parameter, humidity, to the deployment of soil respiration sensing nodes. When there are multiple factors, the problem is that the grouping results may be different using different lightweight parameters. A possible solution is to fuse multiple lightweight parameters as a complex lightweight parameter and analyze its temporal similarities among different positions using *TimSim* deployment strategy and the detailed solution to this problem is a direction we will study in future.

Finally, although the proposed deployment method based on lightweight parameters in this paper is described as a solution to soil respiration sensor networks, this idea for deployment is applicable and can provide a reference and guidance to other applications of wireless sensor networks.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This study is supported by the NSF China under Grant nos. 61190114 and 61303236, the State Bureau of Forestry 948 Project under Grant no. 2013-4-71, and Zhejiang Provincial Science Technology Plan Projects Key Science Technology Specific Project under Grant no. 2012C13011-1.

References

- [1] J. W. Raich and W. H. Schlesinger, "The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate," *Tellus Series B*, vol. 44, no. 2, pp. 81–99, 1992.
- [2] LI-COR, "LI-8100A Automated Soil CO₂ Flux System," <http://www.licor.com/env/products/soil.flux/>.
- [3] R. Juszcak, E. Humphreys, M. Acosta, M. Michalak-Galczevska, D. Kayzer, and J. Olejnik, "Ecosystem respiration in a heterogeneous temperate peatland and its sensitivity to peat temperature and water table depth," *Plant and Soil*, vol. 366, no. 1–2, pp. 505–520, 2013.
- [4] G. Wang, Y. Jiang, X. Wu, L. Mo, and L. Xu, "Soil respiration measurement system based on wireless sensor networks," *Chinese Journal of Sensors and Actuators*, vol. 26, no. 6, pp. 901–906, 2013.
- [5] Q. Liu, N. T. Edwards, W. M. Post, L. Gu, J. Ledford, and S. Lenhart, "Temperature-independent diel variation in soil respiration observed from a temperate deciduous forest," *Global Change Biology*, vol. 12, no. 11, pp. 2136–2145, 2006.
- [6] H. M. Ammari and S. K. Das, "Critical density for coverage and connectivity in three-dimensional wireless sensor networks using continuum percolation," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, no. 6, pp. 872–885, 2009.
- [7] L. Liu and H. Ma, "On coverage of wireless sensor networks for rolling terrains," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 1, pp. 118–125, 2012.
- [8] A. Saipulla, C. Westphal, B. Liu, and J. Wang, "Barrier coverage of line-based deployed wireless sensor networks," in *Proceedings of the 28th Conference on Computer Communications (INFOCOM '09)*, pp. 127–135, IEEE, Rio de Janeiro, Brazil, April 2009.
- [9] S. He, X. Gong, J. Zhang, J. Chen, and Y. Sun, "Curve-based deployment for barrier coverage in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 724–735, 2014.
- [10] H. M. Ammari and S. K. Das, "Centralized and clustered k -coverage protocols for wireless sensor networks," *IEEE Transactions on Computers*, vol. 61, no. 1, pp. 118–133, 2012.
- [11] F. Li, J. Luo, W. Wang, and Y. He, "Autonomous deployment for load balancing k -surface coverage in sensor networks," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 279–293, 2014.
- [12] H. Mahboubi, J. Habibi, A. G. Aghdam, and K. Sayrafian-Pour, "Distributed deployment strategies for improved coverage in a network of mobile sensors with prioritized sensing field," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 451–461, 2013.
- [13] H. Mahboubi, K. Moezzi, A. G. Aghdam, K. Sayrafian-Pour, and V. Marbukh, "Distributed deployment algorithms for improved coverage in a network of wireless mobile sensors," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 1, pp. 163–174, 2014.
- [14] H. Mahboubi, K. Moezzi, A. G. Aghdam, and K. Sayrafian-Pour, "Distributed deployment algorithms for efficient coverage in a network of mobile sensors with nonidentical sensing capabilities," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 8, pp. 3998–4016, 2014.
- [15] C.-Y. Chang, C.-T. Chang, Y.-C. Chen, and H.-R. Chang, "Obstacle-resistant deployment algorithms for wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 6, pp. 2925–2941, 2009.
- [16] S. Lee and M. Younis, "Optimized relay placement to federate segments in wireless sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 5, pp. 742–752, 2010.
- [17] F. Wang, D. Wang, and J. Liu, "Traffic-aware relay node deployment: maximizing lifetime for data collection wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 8, pp. 1415–1423, 2011.
- [18] H. Liu, X. Chu, Y.-W. Leung, and R. Du, "Minimum-cost sensor placement for required lifetime in wireless sensor-target surveillance networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 9, pp. 1783–1796, 2013.
- [19] Y. Song, B. Wang, Z. Shi, K. R. Pattipati, and S. Gupta, "Distributed algorithms for energy-efficient even self-deployment in mobile sensor networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 5, pp. 1035–1047, 2014.
- [20] G. Barrenetxea, F. Ingelrest, G. Schaefer, M. Vetterli, O. Couach, and M. Parlange, "SensorScope: out-of-the-box environmental monitoring," in *Proceedings of the 7th International Conference on Information Processing in Sensor Networks (IPSN '08)*, pp. 332–343, St. Louis, Mo, USA, April 2008.
- [21] L. Mo, Y. He, Y. Liu et al., "Canopy closure estimates with GreenOrbs: sustainable sensing in the forest," in *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems (SenSys '09)*, pp. 99–112, November 2009.
- [22] W.-Z. Song, R. Huang, M. Xu, B. Shirazi, and R. Lahusen, "Design and deployment of sensor network for real-time high-fidelity volcano monitoring," *IEEE Transactions on Parallel and Distributed Systems*, vol. 21, no. 11, pp. 1658–1674, 2010.
- [23] V. Dyo, S. A. Ellwood, D. W. Macdonald et al., "WILDSSENSING: design and deployment of a sustainable sensor network for wildlife monitoring," *ACM Transactions on Sensor Networks*, vol. 8, no. 4, article 29, 2012.
- [24] C. Xu, Z. Zheng, and C. Qi, "Practical deployments of SEMAT on wireless sensor networks in the marine environment," in *Proceedings of the 9th IEEE International Conference on Mobile Ad-Hoc and Sensor Networks (MSN '13)*, pp. 442–448, Dalian, China, December 2013.
- [25] Z. Li, N. Wang, A. Franzen et al., "Practical deployment of an in-field soil property wireless sensor network," *Computer Standards and Interfaces*, vol. 36, no. 2, pp. 278–287, 2014.
- [26] S. Xiong, L. Yu, H. Shen, C. Wang, and W. Lu, "Efficient algorithms for sensor deployment and routing in sensor networks for network-structured environment monitoring," in *Proceedings of the IEEE Conference on Computer Communications (INFOCOM '12)*, pp. 1008–1016, March 2012.
- [27] F. M. Al-Turjman, H. S. Hassanein, and M. A. Ibnkahla, "Efficient deployment of wireless sensor networks targeting environment monitoring applications," *Computer Communications*, vol. 36, no. 2, pp. 135–148, 2013.
- [28] W. Du, Z. Li, J. C. Liando, and M. Li, "From rateless to distanceless: enabling sparse sensor network deployment in large areas," in *Proceedings of the 12th ACM Conference on Embedded Networked Sensor Systems (SenSys '14)*, pp. 134–147, Memphis, Tenn, USA, November 2014.
- [29] W. Du, Z. Xing, M. Li, B. He, L. H. Chua, and H. Miao, "Optimal sensor placement and measurement of wind for water quality studies in urban reservoirs," in *Proceedings of the 13th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN '14)*, pp. 167–178, Berlin, Germany, April 2014.

- [30] Q.-H. Song, Z.-H. Tan, Y.-P. Zhang et al., "Spatial heterogeneity of soil respiration in a seasonal rainforest with complex terrain," *iForest—Biogeosciences and Forestry*, vol. 6, no. 2, pp. 65–72, 2013.
- [31] G. Miao, A. Noormets, J.-C. Domec et al., "The effect of water table fluctuation on soil respiration in a lower coastal plain forested wetland in the southeastern U.S.," *Journal of Geophysical Research G: Biogeosciences*, vol. 118, no. 4, pp. 1748–1762, 2013.
- [32] L. Mo, Y. Jiang, G. Wang, and J. Zhao, "Segmental dynamic duty cycle control for sampling scheduling in wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2013, Article ID 267935, 10 pages, 2013.
- [33] S. Guha, R. Rastogi, and K. Shim, "Rock: a robust clustering algorithm for categorical attributes," *Information Systems*, vol. 25, no. 5, pp. 345–366, 2000.

