Hindawi Publishing Corporation
International Journal of Distributed Sensor Networks
Volume 2013, Article ID 267935, 10 pages
http://dx.doi.org/10.1155/2013/267935

Research Article

Segmental Dynamic Duty Cycle Control for Sampling Scheduling in Wireless Sensor Networks

Lufeng Mo,[1,2] Yujia Jiang,[2] Guoying Wang,[1,2] and Jizhong Zhao[1]

[1] *School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China*
[2] *School of Information Engineering, Zhejiang Agricultural and Forestry University, Lin'an 311300, China*

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

Received 5 July 2013; Accepted 13 September 2013

Academic Editor: Yuan He

Wireless sensor networks for environment monitoring are usually deployed in the fields where electric or manual intervention cannot be accessed easily. Therefore, we hope to minimize the times of sampling to reduce energy consuming. Energy-efficient sampling scheduling can be realized using compressive sensing theory on the basis of temporal correlation of the physical process. However, the degree of correlation of neighboring data varies over time, which may lead to different reconstructive quality for different parts of data if constant duty cycle is used. We proposed SDDC, a segmental dynamic duty cycle control method, for sampling scheduling in wireless sensor networks based on compressive sensing. Using a priori knowledge obtained by means of analysis on earlier sensing data, dynamic duty cycle is determined according to the linear degree of data in each segment. The experimental results using data from soil respiration monitoring sensor networks show that the proposed SDDC method can lead to better reconstructive quality compared to constant duty cycle of the same average sampling rate. That is to say, the SDDC method needs smaller sampling rate if the reconstructive error threshold is given and consequently saves more energy.

1. Introduction

Wireless sensor networks for environment monitoring are usually deployed in the fields where electric or manual intervention cannot be accessed directly. Therefore, the entire system must be energy efficient, so that the sensor networks could run unattended as long time as possible. Processing, sensing, and radio are main operations that consume energy in wireless sensor networks [1]. In this paper, we focus on the second operation: sensing. To collect detailed information of physical process which changes with time, the ideal sampling scheduling strategy is sampling at a very high frequency. However, some measurement operations are time, and energy-exhaustive processes, such as soil respiration speed measurement operation in the soil respiration monitoring sensor networks. Therefore, the main objective of this paper is to design appropriate sampling scheduling policy for the environment monitoring sensor network nodes with energy exhaustive measurement process, so as to reduce the duty cycle of sensor nodes and save energy.

To achieve required reconstructive quality of soil respiration process using as less duty cycle as possible, we can usually use methods like interpolation or fitting. In this paper, we achieve sparse sampling using compressive sensing theory: real values to the physical world can be sparsified on the basis of temporal correlation of soil respiration carbon flux, and soil respiration carbon flux of time series data can be reconstructed using sparse sample data with accuracy requirement.

When compressive sensing theory is used for sparse sampling and data reconstruction, it is needed to determine two matrices: the representation basis matrix Ψ, used to the sparse of true value of soil respiration, and the measurement matrix ϕ, used to indicate the sampling scheduling policy, which is usually a random or uniform sampling with certain duty cycle.

Duty cycle in the measurement matrix determines the number of measurement on soil respiration for the nodes, namely, determines the energy-saving effect compared with dense measurement, and also affects the accuracy of the data reconstruction. Intuitively, the lower the sampling rate is, the better energy-saving effect of the sampling scheduling policy will be, but this may lead to larger reconstructive error. So there is a contradiction between sampling rate and the accuracy of reconstructed data. In designing of sampling scheduling policy, we should find the balance which is based on the demanded accuracy of data reconstruction. On the other hand, the degree of correlation of neighbouring data varies over time, which may lead to different reconstructive quality for different parts of data if constant duty cycle is used.

In this paper, we proposed SDDC, a segmental dynamic duty cycle control method based on compressive sensing. According to SDDC, dynamic sampling rates are adopted when constructing measurement matrix on the basis of data changes over time: higher sampling rates for drastically changing physical stages but lower sampling rates for slightly changing stages. This method can lead to a dynamic trade-off between energy-saving effect and accuracy of data reconstruction. Because we cannot know the true data in advance, the earlier measurement data are analyzed to find a priori knowledge, according to which segmental dynamic sampling rates are obtained. And we analyzed SDDC method with real data from wireless sensor networks for soil respiration monitoring.

The remaining contents of this paper are arranged as following. The temporal sampling scheduling problem is analyzed and modeled based on compressive sensing in Section 2. Section 3 presents the SDDC sampling scheduling method we proposed. Section 4 introduces the experimental data and the design of experimental process, evaluates the SDDC method using measured data from soil respiration monitoring sensor networks, and analyzes the measurement performance of SDDC through the comparison of experimental results. The last section is a summary of this paper and also analyzes future directions so that we may continue our study.

2. Sampling Scheduling Based on Compressive Sensing

2.1. Compressive Sensing. We can obtain a large amount of data through dense, periodic sampling strategy. However, is this the best way to recognize the real physical process? The increase in data volumes does not really mean the increase of the amount of information. On the contrary, too much redundant noisy data may cover up the valid data which contains main structure (the principal components), and at the meantime it increases the difficulty of sampling and sample price.

Compressive sensing mainly relies on data sparseness characteristic and low rank characteristic of original data. Under the condition of less than the Nyquist sampling rate, we get a small amount of discrete samples and then reconstruct the signal and algorithm through nonlinear method [2, 3]. This theory has been applied to data compression [4], channel coding [5], analog signal perception [6], routing [7], data collection [8], and other aspects.

For the discrete signal which is represented by a vector x ($\|x\|_0 \ll N$) whose size is N, the measurement of x can be represented as a matrix Φ with a size of $M \times N$, which is called as measurement matrix, and then we can get vector y with a size of M:

$$y = \Phi x. \tag{1}$$

So the question is: how many times of measurement are needed at least to reconstruct the signal x? According to linear algebra, to have existent and unique solutions of (1), $M \geq N$ is necessary, which means at least N times of measurement are needed. However, if x is sparse ($\|x\|_0 \ll N$), there is probability to reduce the observation volume M, theoretically.

In practice, x may not be sparse, while it is likely to have sparse expression in another domain. Specifically, using a matrix Ψ with the size of $N \times N$, x can be written as

$$x = \Psi s. \tag{2}$$

Here, s is a $N \times 1$ sparse vector from Ψ, $\|s\|_0 = KK \ll N$. Matrix Ψ is also referred to as the representation basis. So, sampling vector can be written as

$$y = \Phi \Psi s. \tag{3}$$

Consequently, there are three main problems in the research and application of compressive sensing: (1) before sampling, designing a representation basis matrix Ψ which is required for the sparsification of x according to the characteristics of x. (2) When sampling, design a measurement matrix Φ in the size of $M \times N$, where M is as small as possible. (3) when reconstructing a signal, using the given y and known matrix Φ and Ψ, we determine the s according to reconstructive m A (the reconstructive algorithms like LP, MP, OMP, ROMP, and so on). Then the original signal can be reconstructed with $x = \Psi s$.

For the first problem, the most important thing is to choose the representation basis matrix Ψ which would transform x into a sparse matrix. Usually using the wavelet as basis matrix can achieve approximate sparse for the smooth data, most absolute value of expansion coefficients is small.

For the second problem, the measurement matrix Φ is used in the third task, so it should be chosen seriously, and it is necessary to meet the restricted isometry principle (RIP) [5]. Currently, the measurement matrix usually adopts Gaussian random measurement matrix or Fourier matrix, such as Bernoulli matrix.

For the third problem, (3) is a nondetermined linear system because $M \ll N$; the solution of this system is widely studied in recent years. The first way is to find s, who has the minimum Paradigm l_0:

$$\begin{aligned} &\min_{s \in \mathbb{R}^N} \|s\|_0 \\ &\text{s.t.} \quad y = \Phi \Psi s. \end{aligned} \tag{4}$$

It is very difficult to solve directly [5, 9]. If N is large, there is no solution. But there are fast methods to smooth paradigm l_0, for example, SL0 [10]. The second way is to use the minimization of paradigm l_1 instead of paradigm l_0 which can reduce the complexity of the algorithm; it is called basis pursuit (BP) [11]:

$$\begin{aligned} &\min_{s\in\mathbb{R}^N} \|s\|_1 \\ &\text{s.t.} \quad y = \Phi\Psi s. \end{aligned} \tag{5}$$

It can be solved using the linear programming (LP) method. There exists the polynomial time algorithm to solve these problems, including interior-point method, as well as some fast algorithms for the large-scale systems [12, 13]. In addition to the linear programming method, the commonly used algorithms include matching pursuit (MP), the OMP [12], and ROMP [13]. They are thought to be faster than the LP method, but they are worse in quality, especially when the signal is not sparse enough.

If M meets the following equation:

$$M \geq C\mu^2 (\Phi, \Psi) K \log N \tag{6}$$

there is very high possibility to reconstruct the K sparse signals from M measurements using any of the above reconstructive algorithm A. C is a positive constant, N is the size of signal, and $\mu^2(\phi,\psi)$ is the relevance between Φ and Ψ. Given a heap of orthogonal basis Φ and Ψ which rely on RN, coherence can be defined as

$$\mu (\Phi, \Psi) = \sqrt{N} \max_{1\leq i,j\leq N} \left|\left\langle \phi_i, \psi_j \right\rangle\right| \in \left[1, \sqrt{N}\right], \tag{7}$$

and are column vectors of Φ and Ψ, respectively. When Φ and x are certain, we should select the Ψ carefully. X must be expressed sparsely in the domain of Ψ, and at the meantime, $\mu(\Phi,\Psi)$ must be as small as possible.

2.2. Modeling the Sampling Scheduling Problem in Time-Domain. In the real physical world, carbon flux of soil respiration in the sample point is continuous in the time. It can be treated as discrete while the time unit is small enough compared with the time scale of the changes in soil respiration. In reality, no matter how high sampling frequency is, the operations of the measuring equipment are discrete, and the carbon flux data on soil respiration is obviously discrete.

We use the discrete time model in modeling the sampling schedule of soil respiration monitoring sensor network [14].

(a) Time-series data of soil respiration carbon flux over a period of time in the sampling location can be expressed as $X = \{x_t\}$, $(t = 1, 2, \ldots, N)$, t is for time, with a total of N times.

(b) The sampling scheduling policy π is expressed as $T^\pi = \{t_1, t_2, \ldots, t_n\}$, $(t_i \in \{1, 2, \ldots, N\}$, $1 \leq i \leq n)$, that is a subset of the real time series, and we will sample at those moments.

(c) Assuming that there is no measurement error and noise, after several times of measurement which relies on the sampling scheduling policy π, we will get the sample data sequence $X^\pi = \{x_{t_1}, x_{t_2}, \ldots, x_{t_n}\}$ that is part of the real physical world time-series data.

(d) In order to understand the real process of soil respiration, it is necessary to measure several times according to the sampling scheduling policy π and then reconstruct the time-series data of whole process of soil respiration carbon flux with the sampling data. That is to say, generate estimation of the original sequence $\widehat{X}^\lambda = \{\widehat{x}_t\}$, $(t = 1, 2, \ldots, N)$ according to the estimation function λ and the sample data sequence X^π. If $t \in T^\pi$, then $\widehat{x}_t$ equals x_t; otherwise, $\widehat{x}_t$ equals the value of the estimation function, $\widehat{x}_t^\lambda(x^\pi)$.

(e) Basing on the above description, the goal of sampling scheduling policy is to select the best sampling strategy π and estimate function λ, so as to minimize the evaluated error between the reconstructed soil respiration data sequence $\widehat{X}^\lambda$ and the original real physical world soil respiration data sequence X, and the average sampling rate should be in a certain range. Namely,

$$\begin{aligned} &\min_{\pi,\lambda} \ \mathrm{Err}\left(X, \widehat{X}^\lambda \left(X^\pi\right)\right), \\ &\text{s.t.} \quad \frac{n}{N} \leq \alpha. \end{aligned} \tag{8}$$

Err is a specific error metrics, and α is the maximum sampling rate threshold value.

2.3. Model of Sampling Scheduling Based on Compressive Sensing. In the application of sensor network for soil respiration monitoring, we design the sampling scheduling policy using the compressive sensing theory, namely, to design according to the original data sequence X, the sampling scheduling policy π, sampling data X^π, the reconstructed estimate sequences $\widehat{X}^\pi$, and the error metrics Err which are presented in Section 2.2 with compressive sensing theory.

We model sampling scheduling based on compressive sensing as follows.

(a) The raw data sequence X: we express it with a vector $X = \{x_t,\ t = 1, 2, \ldots, N\}$ whose size is $N \times 1$, which is the x in the compressive sensing equation (1).

(b) The sampling scheduling policy π: we express the sampling scheduling policy measurement matrix $\Phi_{M\times N}$ from the equation of compressive sensing. M is the row number of the matrix, and it is the times of sampling. The columns of the matrix mean the sampling time, and each row contains a design of a sampling time. If the value of the matrix ϕ in row m and column n is 1, that means the mth measurement takes place at time n.

(c) Sample data sequence X^π: according to the Sampling scheduling policy $\pi = \Phi_{M\times N}$, if the value $\Phi(m,n)$ is 1, then the measured data is x_n, the whole sample data sequence is ΦX, it can be recorded as y in (1).

(d) Reconstructed data sequence $\widehat{X}^{\pi}$ according to the given y, Φ, and Ψ, we can get s through a certain method based on the compressive sensing, namely, Ψs.

(e) Reconstructed error metrics Err: we can evaluate the quality of reconstruction through the average error $\|\Psi s - x\|_1/N$ or the mean square error $\|\Psi s - x\|_2/N$ between Ψs and x.

3. The Segmental Dynamic Duty Cycle Control Method

To estimate one soil respiration carbon flux data, it is necessary to measure the soil temperature, humidity, air pressure in the closed chamber, and CO_2 concentration using the soil respiration measurement instrument. The measurement of temperature, humidity, and air pressure is of low energy consumption, while the measurement of changes in CO_2 concentration is a complicated and energy-consuming process.

A soil respiration carbon flux value measuring cycle is three minutes, and in this period it measures the CO_2 concentration every three seconds and the chamber keeps closed. Then the chamber opens automatically for ventilation with the outside world for one minute. This procedure ensures the following measurement to reflect the real process of soil respiration. In the measurement period, there are 60 CO_2 concentration data. Firstly, use the linear fitting method on these data and then calculate the slope, and thus get the change rate of CO_2 concentration during the measurement period. Then combining with parameters such as soil temperature, humidity, air pressure in the closed chamber, soil respiration flux data are obtained through carbon flux calculation formula. Soil respiration measurement is large in energy consumption; reducing the sampling frequency through compression perception theory can effectively extend the life span of the equipments.

Using compressive sensing theory to carry on the sampling schedule of the sensor network for soil respiration monitoring, we should confront several problems of the compressive sensing research and application which are described in Section 2.1. We focus on the second one, namely, the design of the measurement matrix.

As described in Section 2.3, the rows count M in the measurement matrix $\phi_{M\times N}$ that we designed represents the number of samples; every sampling time is represented by the nonzero elements of the column number n in the same row. Because measuring soil respiration can only measure data of one point in time in each sampling, every row in measurement matrix ϕ has a nonzero element, and every column has no more than one nonzero element. Therefore, the design of measurement matrix comprises two parts: (1) row number M, that is, sampling frequency; (2) column number n of the nonzero element in every row, that is, every sampling point in time.

This paper mainly focuses on the former one in the design of measurement matrix, which is the determination of sampling times M. According to the original data sequence X in real physical world, determining the number of samples equals determining the sampling rate M/N. If the data in the sequence X changes linearly with time, we can get the sparse matrix through the appropriate linear transformation of basis matrix Ψ; thus we can get a good reconstruction result through a lower sampling frequency. But data sequence X does not change linearly in the real physical world, and it is necessary to increase the sampling rate in order to get abundant data change information, so that the reconstructed data sequence can meet needs of overall accuracy.

Although X does not change linearly, it is possible to get approximate linear change in some parts of the X through a further decomposition of X. In the whole measuring process, if we use the fixed sampling rate to design the measurement matrix, we may get better reconstruction results in the approximate linear change part, but in the nonlinear change part it is bad. If we increase the sampling rate in order to improve the reconstruction results in nonlinear part, there will be a certain redundancy in a linear change part of the sampled data. This paper studied SDDC, a segmental dynamic sampling scheduling policy, in which dynamic sampling rate is used to construct the measurement matrix in different time interval according to the trends of X. Under the condition of meeting the required accuracy, we lower the sampling rate of the linear change part, and increase the sampling rate of the nonlinear change part so that we can reduce the sampling rate as far as possible. Soil respiration measurement is an energy-consuming and time-consuming process, and the reduction of sampling rate can reduce the energy consumption of the whole monitoring system and thus extend the field working time of the system.

Studies show that soil respiration relates to the change of time. Soil respires slowly in the day but respires relatively quickly in the night. This is because one of the causes for soil respiration is the respiring effect of plant root system. Therefore, the changes of soil respiration are influenced by plant physiological processes [15]. Plant conducts photosynthesis to sequestrate carbon in the noon when it is the best time. Carbon is transported to the root several hours later and is released through root respiration in the night [15, 16]. And root respiration lags behind photosynthesis for 7–12 hours [17]. In addition, temperature is higher than that of soil at noon and the gas pressure is also stronger, which restrains the spread and release of soil CO_2, so value of soil respiration is relatively low in this period [18]. The intensity for the respiring effect of plant root system relates to the location and season. During the summer when plant grows vigorously, rate of soil respiration attains peak value at night when rate of soil respiration is higher than that in the day. But during the winter when plant grows slowly, rate of soil respiration at night is a little higher than that in the day without apparent peaks and valleys [19].

There is regularity and similarity in the soil respiration, changes over time by day cycle. And there is little difference among neighboring days on the temperature in a day which caused by the sun as well as the difference of plant growth caused by the season. The trend on the change of soil respiration can be estimated by the soil respiration data which is measured a few days before. So we proposed a SDDC method based on a priori knowledge. Everyday sampling time is divided according to the observation and analyses of the experimental data or the reconstructed data. Then sampling rate of each segment is differed according to the historical data curve. In the time period of which the data sequence is highly nonlinear, the sampling rate is increased; on the contrary, the sampling rate is reduced.

In order to get the sampling time fragmented, piecewise function can be fitted and subdivided completely according to the changing trend of historical data. It is aimed for the nonlinear degree of data in each piecewise function so as to reduce the sampling rate at the extreme in the context of reconstructing quality. However, considering the situation of soil respiration monitoring sensor network, except for measurement, each sensor node can communicate. And this requires that each sensor node should coordinate mutually when communicating with other nodes. It will result in the difference of different fragmented length in a node if fragmented by data changing trend. Moreover, due to the spatial heterogeneity of soil, the temporal segmental results by sensor nodes at different sampling locations are likely to be different, which makes it difficult for the cooperation between measurement and communication of sensor nodes.

This paper will employ the fixed segmentation method which segments the sampling time evenly. On the one hand, with the same segmentation, the original physical world data sequence X can be divided into the subsequence X_i with the same number of data, which is described in Section 2.3. Assume that each subsequence X_i has an element number of N, we can use the same basis matrix $\Psi_{N\times N}$. On the other hand, the same time segmentation method is good for the synchronous communication among different nodes. Each node uses the same communication scheduling method, for example, transfer the measurement data of the former period at the beginning of subsection sampling period.

On the basis of fixed segmentation, this paper presents a segmental dynamic duty cycle method based on a priori knowledge, as shown in Algorithm 1. When we get the segmental dynamic measurement matrix Φ_i, soil respiration measurement instrument will measure with different measurement matrix according to different sampling segment.

In Algorithm 1, we choose R^2 as the evaluation index of the linear degree of each data sequence. Determination coefficient R^2 is often used to evaluate the fitting degree between fitting results and the corresponding real data; the numerical values range between 0 and 1. When R^2 equals 1 or close to 1, there is high correlation between those data, on the contrary, the correlation is low.

The β in line 10 is incremental adjustment factor, the higher it is, the larger the difference of sample rate between different segments with different nonlinear degree will be. According to the expression in line 10, the larger in ${R^2}_i$ for a certain segment means a higher linear degree of changes in soil respiration of that segment; thus, there will be smaller sampling rate, and conversely, there will be larger sampling rate.

Algorithm 1 uses the data derived from the soil respiration data sequence X_{P_i} $(1 \le P_i \le P_d)$ just P_d days before, and this leads to certain timeliness for the calculation results. Therefore, in the long-term process of soil respiration monitoring, soil respiration monitor needs to (the period can usually be set in three days to a week) repeat the above algorithm periodically so as to update the segmental sampling rate. As there is no a priori data at the beginning of the measurement, we adopt a static scheduling policy with the same sampling rate for each segment.

4. Simulation Experiment

As is mentioned in Section 2.1, the reconstruction quality of compressive sensing is influenced by three factors, measurement matrix Φ, basis matrix Ψ, and the reconstructive algorithm A. The following paper, respectively, introduced the design and choice on these three factors in the experiment, as well as the experimental data and solution.

4.1. Experimental Data. We have made dense measurement outside for 10 days with the self-designed soil respiration measurement instrument and got the original data sequence in the real physical world.

As mentioned above, there is one soil respiration carbon flux data every 4 minutes. So there are 3600 data in the dataset which we used in this experiment. The dataset is divided into several subsequences evenly, and then we sample and reconstruct on each sequence which is treated as an experimental data.

Evenly segment the sampling time as is mentioned in Section 3, then the data sequence X is divided into several subsequence, and sample and reconstruct on each subsequence. According to the model in Section 2.3, the element number in each sequence is N, so the number of subsequence is $d = 3600/N$, and then record the ith subsequence as X_i $(1 \le i \le d)$.

4.2. The Construction of Measurement Matrix. As is described in Section 3, the design of measurement matrix includes aspects: the determination of the row number of M and the column numbers of n where the value is nonzero. Section 3 put forward SDDC method which solves the first problem completely. The policy determines the sampling frequency SR_i of the subsequence X_i in the ith time segment, by the time interval of the sampling frequency is $M = (\mathrm{SR}_i \times N)$, so the measurement matrix is the matrix of $(\mathrm{SR}_i \times N)$ rows and N columns.

For the second aspect, determine the column number N of non-zero element in all rows, namely, sampling time of concrete measure of M times; this problem does not belong to this research. In the experiment, two simple but RIP constraint solution schemes are chosen: the periodic sampling (PS) and pseudorandom sampling (RS). Periodic sampling means that the nodes are measured M times according to the cycle of N/M, this measurement matrix noted as Φ_P. Pseudo random sampling means that the sampling time segment is distributed according to the uniform random probability, this measurement matrix noted as Φ_R.

Require: $X_{P_i}, N, \mathrm{SR}_{\mathrm{base}}$
(1) X_{P_i} = soil respiration data reconstructed by the sampling data of last P_d days, everyday data consists of N_d elements $((1 \le P_i \le P_d))$.
(2) X = the mean data sequence calculated according to the corresponding time for X_{P_i};
(3) **for** $i = 1$ to $d = \lceil N_d/N \rceil$ **do**
(4) X_i = subsequences with a length of N segmented from X;
(5) ${R^2}_i$ = determination coefficient of linear fitting of X_i;
(6) **end for**
(7) ${R^2}_{\mathrm{mean}} = \sum_{i=1}^{d} {R^2}_i/d$;
(8) $\mathrm{SR}_{\mathrm{base}}$ = a given sampling rate base, which is the sampling rate while the determination coefficient of the linear fitting equals to ${R^2}_{\mathrm{mean}}$;
(9) **for** $i = 1$ to d **do**
(10) $\mathrm{SR}_i = \mathrm{SR}_{\mathrm{base}} + ({R^2}_{\mathrm{mean}} - {R^2}_i)\,\beta$;
(11) Φ_i = the dynamic measurement matrix constructed according to SR_i for the present segment
(12) **end for**
(13) **return** $\Phi_i, (1 \le i \le d)$;

Algorithm 1: Segmental dynamic duty cycle method based on a priori knowledge (SDDC).

4.3. The Selection of Basis Matrix Ψ and the Reconstructive Algorithm A. The change of soil respiration in the physical world is smooth, so the raw data sequence X_i can be sparse according to the correlation between the adjacent sampling data. This paper adopts two schemes to express the basis matrix Ψ [14].

(1) The difference matrix

$$M_D = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & -\gamma \end{pmatrix}. \quad (9)$$

The last element γ should be $0 < \gamma < 1$, so as to make sure that M_D is reversible. In this experiment γ is 0.001. Project X_i on M_D, then $s = M_D X_i$ is a vector that contains a number of 0 elements and small absolute value elements. If so, the original signal x can be sparsely expressed as $x = M_D^{-1} s$. Therefore, in this paper we use M_D^{-1} as a basis matrix, which is recorded as Ψ_D. (2) The Haar wavelet transform M_H can be used to sparsification of smooth data, so we use M_H^{-1} as the basis matrix, which is recorded as Ψ_H.

Algorithm SL0 and BP (the LP) mentioned in Section 2.1 are used as the reconstructive algorithm. Based on the study of Wu and Liu [14], in this paper, when the basis matrix is Ψ_D, reconstructive algorithm SL0 is adopted, and when the basis matrix is Ψ_H, reconstructive algorithm LP is adopted. Code of algorithm LP is acquired from SparseLab [20], SL0 is from [21].

4.4. Experimental Scheme. Based on the measurement matrix Φ (Φ_P and Φ_R) which is obtained by periodic sampling and pseudo random sampling, when the basis matrix Ψ (Ψ_D and Ψ_H) and reconstructive algorithm A (SL0 and LP) are confirmed, we can analyze the dynamic changes of sampling rate and the reconstruction results on soil respiration measurement data using the SDDC method described in Section 3. The experimental scheme is shown in Algorithm 2.

As mentioned above, the equipment can collect 15 data per hour. The value range is 30, 60, 90, 120, 180, and 360, which is the length of the subsequence N; they correspond to the segmental cycle of time as 2, 4, 6, 8, 12, and 24 hours, respectively. We adopted the average error to evaluate the reconstructed results. Its calculation method is shown in Algorithm 2, line 16. Where R is the times of random experiment, it is set as 20. There is scheduling mechanism for the random factors in the experiment.

4.5. Experimental Results and Analysis. Firstly, we analyzed the dynamic sampling rate calculated by the dynamic sampling strategy. In SDDC method, use experimental data of the first five days and get its mean value according to the corresponding relation with time, using Algorithm 1; the results are shown in Figure 1. Secondly, segmental linear fitting on these data, get the determination coefficient R^2 of each segment; results are shown in Figure 2. Finally, calculate the dynamic sampling rate for each segment according to Algorithm 1, line 10. The results are shown in Figure 3. The referencing sampling rate $\mathrm{SR}_{\mathrm{base}}$ is 10% and incremental adjustment factor β is 0.2.

Observing the fitting coefficient R^2 of each segment in Figure 2, the relatively high R^2 happens in the fifth segment (time is 8 to 10) where $N = 30$, the ninth segment (time is 16 to 18) and the tenth segment (time is 18 to 20) where $N = 30$, and the fifth segment (time is 16 to 20) where $N = 60$. This means that the sequential variation on soil respiration carbon flux data of these segments has relatively high linear degree. Moreover, the dynamic sampling rates of the four segment are all below 5%, lower than half of the referencing sampling rate, observed from Figure 3. According to Figure 3, the highest sampling rate of the segments is the 12th segment

$\Phi\_method = \{\Phi_P, \Phi_R\}$; $\langle \Psi\_method, A\_method \rangle = \{\langle \Psi_D, \mathrm{SL0} \rangle, \langle \Psi_H, \mathrm{LP} \rangle\}$;
$N = \{30, 60, 90, 120, 180, 360\}$;
for each $\langle \psi\_method, a\_method \rangle \in \langle \Psi\_\mathrm{method}, A\_\mathrm{method} \rangle$ **do**
  for each $\phi\_method \in \Phi\_method$ **do**
    for each $n \in N$ **do**
      divide dataset X into subsequence X_i $(1 \le i \le d = \|X\|_0/n)$ with the size of n;
      ψ = the $n \times n$ basis matrix constructed according to method ψ_method;
      for $i = 1$ to d **do**
        SR_i = the sampling rate of X_i using SDDC;
        ϕ = the $\lceil \mathrm{SR}_i * n \rceil \times n$ measurement matrix constructed according to method ϕ_method;
        $y = \phi X_i$;
        $s = a\_method(y, \phi, \psi)$;
        $\widehat{X_i} = \Psi s$;
        $Err_{(\phi\_method, \psi\_method, i)} = \left\| \widehat{X_i} - X_i \right\|_1$
      end for
      $ErrAvg_{(\psi\_method, \phi\_method, n)} = (1/R) \sum_{r=1}^{R} (1/d) \sum_{i=1}^{d} Err_{(\phi\_method, \psi\_method, i)}/n$
    end for
  end for
end for
return $ErrAvg$;

Algorithm 2: Experimental procedure experiment.

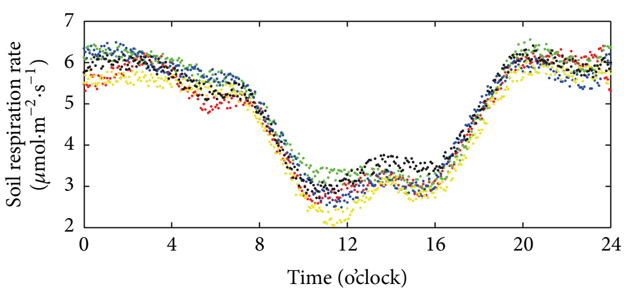


Figure 1: Soil respiration carbon flux data of five days.

where $N = 30$ (time is 22 to 24); the R^2 of the corresponding segment in Figure 2 is about 0.16. In Figure 2, when $N = 360$ (time is 0 to 24), the whole segments have the lowest R^2. But these segments had adopted the referencing sampling rate (Figure 3) rather than a high sampling rate. As there is only one segment one day and there is no other good segment which has a better linear degree to balance it, so it is equal to the condition which has nonsegmental fixed sampling rate.

Figure 4 has shown the average error for the whole reconstruction between SDDC method (SDDC) and constant duty cycle (CDC) with a nonsegmental and fixed sampling rate. We selected 10% (10CDC, 10SDDC in figure), 20% (20CDC, 20SDDC in figure) and 30% (30CDC, 30SDDC in figure), as the fixed sampling of the constant sampling strategy and the referencing sampling rate of dynamic sampling scheduling policy. According to the expression of SR_i in Algorithm 1, line 10, when using the SDDC method, the mean value of the dynamic sampling rate SR_i for each segment is equal to the referencing sampling rate $\mathrm{SR}_{\mathrm{base}}$. Therefore, as is shown in Figure 4, the mean sampling rates under the condition of 10CDC and 10SDDC are 10%. Namely, the energy consumed in sampling is the same.

According to Figure 4, when measurement matrix Φ and Ψ choose different construction methods, there will be smaller reconstruction error from the SDDC method than from the constant sampling strategy under the same mean sampling rate. Furthermore, as can be seen from the figure, no matter which method (Φ_P or Φ_R) we choose to construct the measurement matrix, the reconstruction error is always smaller while using the basis matrix Ψ_D than using Ψ_H. That means Ψ_D is more suitable for the processing of soil respiration carbon flux data. Especially when the element number of the subsequence is small (e.g., 30), it will have larger reconstruction error if we use Ψ_H. In general, the reconstruction error derived by using Ψ_H is almost three or more times larger than using Ψ_D. When the basis matrix Ψ is determined, there is little difference between the selecting of measurement matrix Φ_P and Φ_R. This also indicates that to optimize the sampling quantity scheduling strategy (the construction of measurement matrix), the most effective way is to lower the sampling rate. This is one of the reasons why we focus on the study on dynamic sampling scheduling policy but not on the construction of the measurement matrix.

Figure 4 also showed the difference on global reconstruction effect between dynamic sampling schedule and the constant sampling schedule. In Figure 4(a), under condition that the size of segmental subsequence is 30 (namely, 2 hours) and the constant sampling rate and the referencing sampling rate are 10%, the mean reconstruction errors of these two methods are about 0.182 and 0.101, respectively.

In order to analyze the reason that causes the difference, we calculated the mean reconstruction error of the subsequence of each segment; the results are shown in Figure 5.

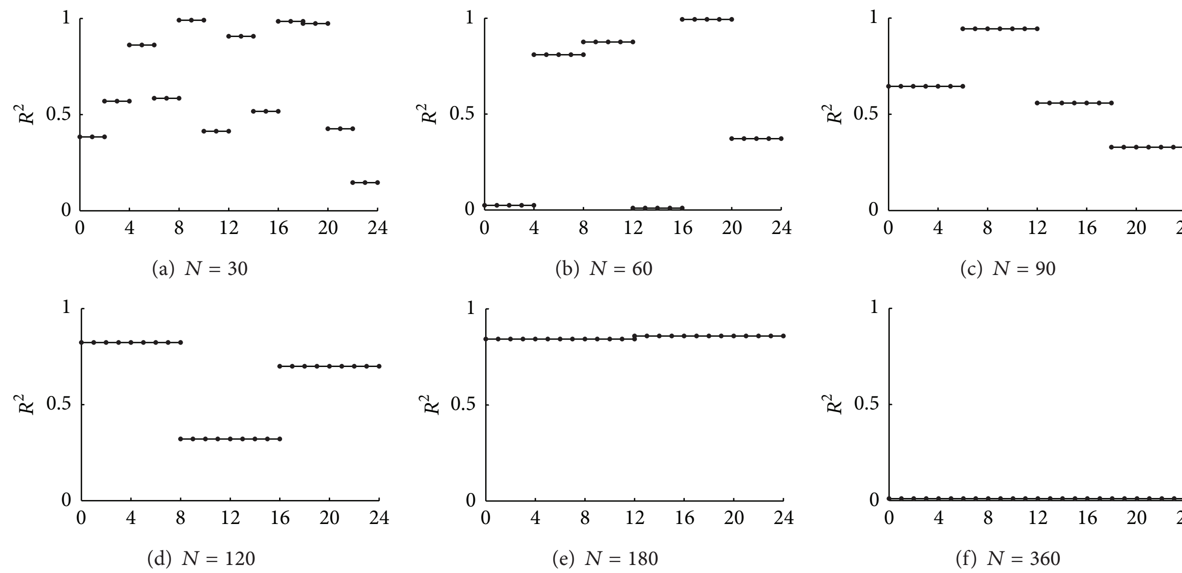


Figure 2: R^2 of segmental linear fitting.

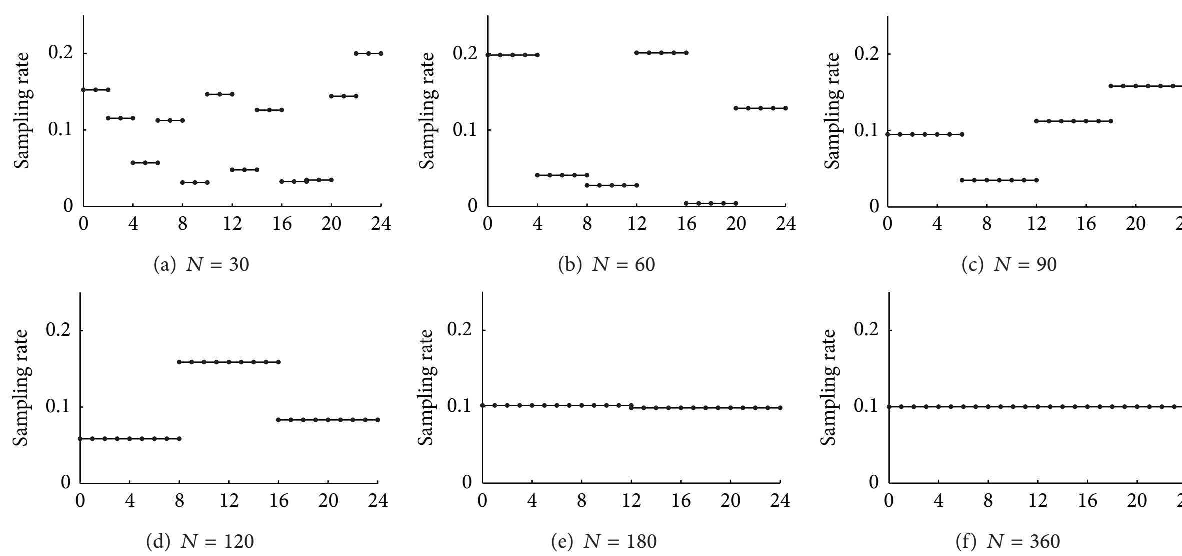
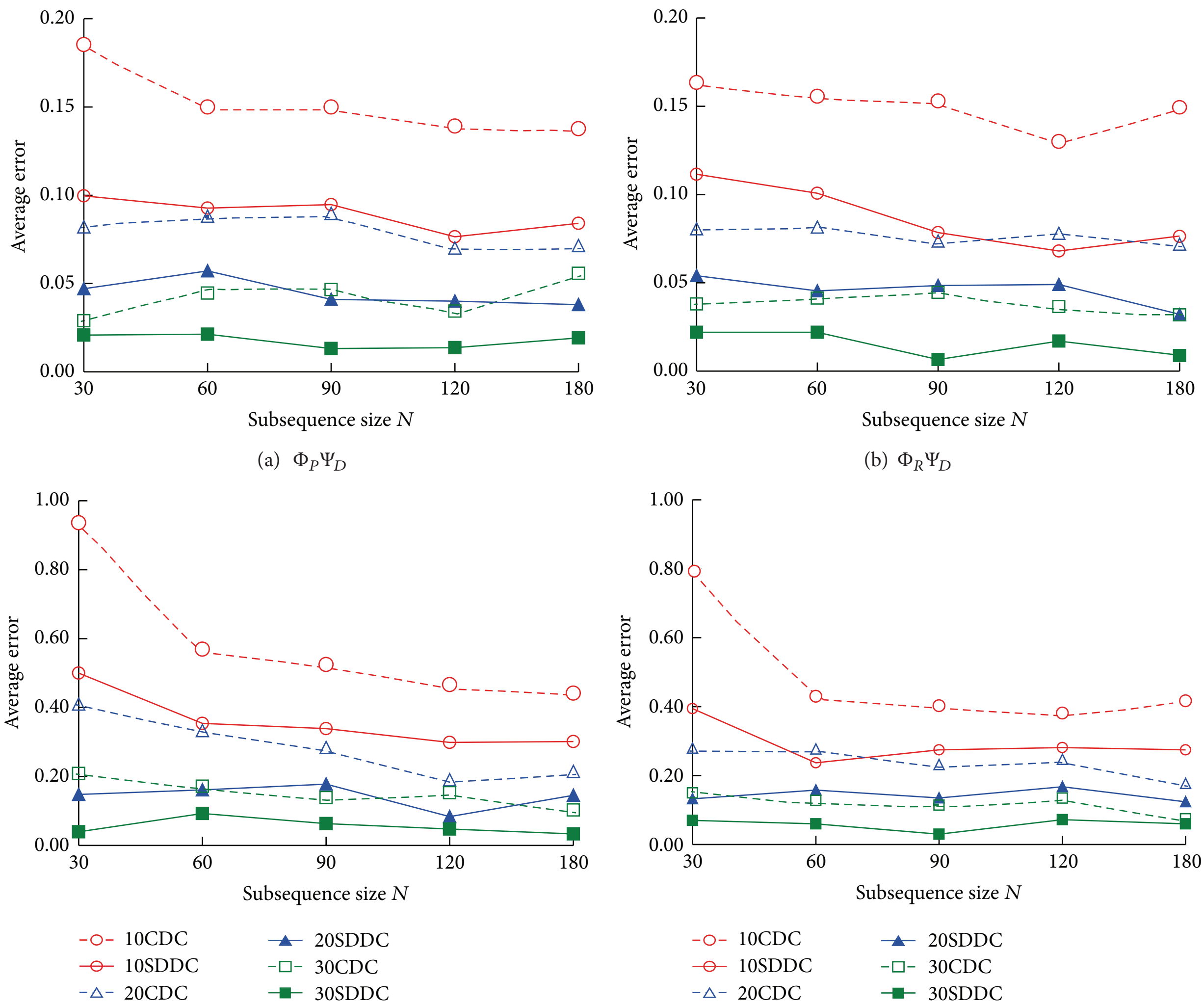


Figure 3: Segmental dynamic sampling rate.

As is seen from Figure 3, when segment with the $N = 30$ (namely, 2 hours), in periods 1, 4, 6, 8, 11, and 12, the dynamic sampling rate conducted by Algorithm 1 is higher than the referencing sampling rate. That is to say the corresponding subsequence of soil respiration carbon flux data is with high nonlinear degree. We should increase the sampling rate to get a better reconstruction effect. According to Figure 5, in the periods corresponding to segment 1, 4, 6, 8, 11, and 12, the segmental reconstruction error calculated by the fixed sampling rate is far large than the dynamic sampling rate. In the other time segment 2, 3, 5, 7, 9, and 10, the dynamic sampling rate is lower than the referencing sampling rate, which means the linear degree is high in these data sequences. We can get better reconstruction effect without a high sampling rate. Although the high sampling rate of constant sampling policy leads to a higher quality of the reconstruction than the dynamic sampling rate, the increase is not large. The reconstruction error acquired by the fixed sampling rate is a little smaller than by the dynamic sampling rate during the time segment 2, 3, 5, 7, 9, and 10 in Figure 5.

According to the segmented error which corresponds to the two kinds of sampling rate, the difference on the reconstruction error, which is calculated by the constant sampling strategy, is relatively large. As the linear degree of each time segment is different, there is bigger difference in the reconstruction results when calculated with the indiscriminate method. The reconstruction errors calculated by dynamic sampling and scheduling policy in all segments were similar. This is because we use different sampling rate in different segments, so that we can balance the difference in



(a) $\Phi_P\Psi_D$ (b) $\Phi_R\Psi_D$

(c) $\Phi_P\Psi_H$ (d) $\Phi_R\Psi_H$

FIGURE 4: Reconstructive performance comparison between SDDC and CDC.

reconstruction accuracy which is carried by the change of linear degree.

5. Conclusion

On the energy-saving sampling issues of soil respiration monitoring, we proposed a segmental dynamic sampling scheduling policy based on compressive sensing (SDDC). We found that SDDC method can adapt to the dynamic changing of monitoring objects better so as to reduce the sampling rate and save energy and achieve the effect of relatively uniform segmental sampling error and better overall reconstructive quality. Though SDDC needs the soil respiration instrument to carry out extra sampling rate updating algorithm and produce same energy, the energy saving of reducing sampling times can far outnumber the energy consuming of updating the sampling rate because the measurement of soil respiration is a relatively energy-consuming and time-consuming process. Although the SDDC sampling scheduling method in this paper is proposed based on sensor networks for soil respiration monitoring and the related performance analysis is carried out using these measured data, SDDC can be used commonly, and it is widely applicable for other sparse sampling application scene with a priori regular pattern.

Soil respiration includes the root respiration, soil microbial respiration, and heterotrophic respiration of soil animal. These respirations are affected by soil temperature and humidity. Main environmental factors which affect soil respiration rate are soil moisture and temperature, in both spatial gradient and time level [22]. Soil respiration measurement requires a dynamic-open box method and other methods, including some time-consuming and energy-consuming process like movement of box, measurement after the pumping of air. It is relatively easy on the measurement of soil temperature and humidity. We plan to find the relevance among temperature and humidity and these sensing data of soil respiration, so as to optimize and adjust the dynamic sampling policy for soil carbon flux, and to further reduce the energy consumption.

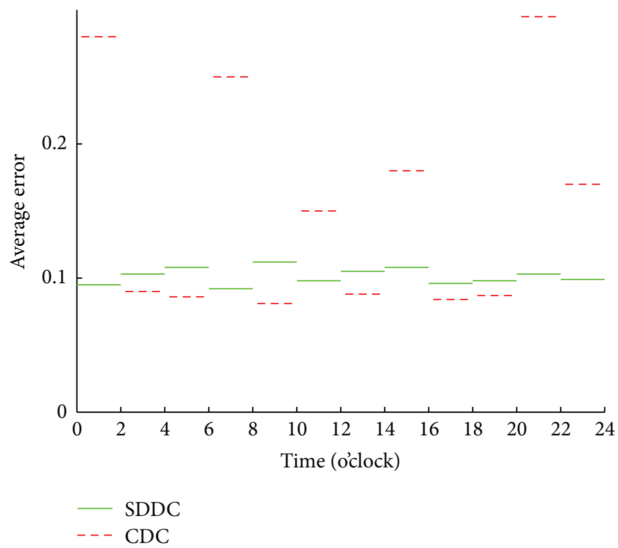


Figure 5: Segmental reconstruction performance comparison between SDDC and CDC.

Acknowledgments

This study is supported by the NSF China under Grant no. 61190114 and 61303236, Zhejiang Provincial Natural Science Foundation of China under Grant no. LY12F02016, and Zhejiang Provincial Science Technology Plan Projects Key Science Technology Specific Project under Grant no. 2012C13011-1.

References

[1] M. N. Halgamuge, M. Zukerman, K. Ramamohanarao, and H. L. Vu, "An estimation of sensor energy consumption," *Progress in Electromagnetics Research B*, no. 12, pp. 259–295, 2009.

[2] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] E. J. Candes and M. B. Wakin, "An introduction to compressive sampling: a sensing/sampling paradigm that goes against the common knowledge in data acquisition," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[4] D. Baron, M. Wakin, M. Duarte, S. Sarvotham, and R. Baraniuk, "Distributed compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[5] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[6] Z. Tian and G. B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '07)*, pp. IV1357–IV1360, April 2007.

[7] G. Quer, R. Masiero, D. Munaretto, M. Rossi, J. Widmer, and M. Zorzi, "On the interplay between routing and signal representation for Compressive Sensing in wireless sensor networks," in *Proceedings of the Information Theory and Applications Workshop (ITA '09)*, pp. 206–215, February 2009.

[8] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Compressive data gathering for large-scale wireless sensor networks," in *Proceedings of the 15th Annual ACM International Conference on Mobile Computing and Networking (MobiCom '09)*, pp. 145–156, September 2009.

[9] D. L. Donoho, M. Elad, and V. N. Temlyakov, "Stable recovery of sparse overcomplete representations in the presence of noise," *IEEE Transactions on Information Theory*, vol. 52, no. 1, pp. 6–18, 2006.

[10] H. Mohimani, M. Babaie-Zadeh, and C. Jutten, "A fast approach for overcomplete sparse decomposition based on smoothed $\ell 0$ norm," *IEEE Transactions on Signal Processing*, vol. 57, no. 1, pp. 289–301, 2009.

[11] D. Donoho, "For most large underdetermined systems of linear equations the minimal l1-norm solution is also the sparsest solution," Technology Report, 2004, http://stats.stanford.edu/~donoho/Reports/2004/l1l0EquivCorrected.pdf.

[12] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

[13] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *IEEE Journal on Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 310–316, 2010.

[14] X. Wu and M. Liu, "In-situ soil moisture sensing: measurement scheduling and estimation using compressive sensing," in *Proceedings of the 11th ACM/IEEE Conference on Information Processing in Sensing Networks (IPSN '12)*, pp. 1–12, April 2012.

[15] 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.

[16] N. B. Dilkes, D. L. Jones, and J. Farrar, "Temporal dynamics of carbon partitioning and rhizodeposition in wheat," *Plant Physiology*, vol. 134, no. 2, pp. 706–715, 2004.

[17] J. Tang, D. D. Baldocchi, and L. Xu, "Tree photosynthesis modulates soil respiration on a diurnal time scale," *Global Change Biology*, vol. 11, no. 8, pp. 1298–1304, 2005.

[18] J. Cao, L. Song, G. Jiang et al., "Diurnal dynamics of soil respiration and carbon stable isotope in Lunan stone forest," *Yunnan Province, Carsologica Sinica*, vol. 24, no. 1, pp. 23–27, 2005.

[19] W. Feng, X. Zou, L. Sha et al., "Comparisons between seasonal and diurnal patterns of soil respiration in a montane evergreen broad-leaved forest of Ailao mountains, China," *Journal of Plant Ecology*, vol. 32, no. 1, pp. 31–39, 2008.

[20] SparseLab, http://sparselab.stanford.edu/.

[21] SL0, http://ee.sharif.ir/~SLzero/.

[22] J. W. Raich and A. Wtufekcioglu, "Vegetation and soil respiration: correlations and controls," *Biogeochemistry*, vol. 48, no. 1, pp. 71–90, 2000.