

Lawrence Berkeley National Laboratory

Lawrence Berkeley National Laboratory

Title

Time-windows-based filtering method for near-surface detection of leakage from geologic carbon sequestration sites

Permalink

<https://escholarship.org/uc/item/2hv8m72g>

Author

Pan, L.

Publication Date

2010-04-01

Peer reviewed

Time-Windows-Based Filtering Method for Near-Surface Detection of Leakage from Geologic Carbon Sequestration Sites

Lehua Pan, Jennifer L. Lewicki, Curtis M. Oldenburg, and Marc L. Fischer

Earth Sciences Division, Lawrence Berkeley National Laboratory
Mail-Stop 90-1116, One Cyclotron Road, Berkeley, CA 94720, USA

Corresponding author:

Lehua Pan

Email: lpn@lbl.gov

Telephone: (510) 495-2360

Fax: (510) 486-5686

Abstract

We use process-based modeling techniques to characterize the temporal features of natural biologically controlled surface CO₂ fluxes and the relationships between the assimilation and respiration fluxes. Based on these analyses, we develop a signal-enhancing technique that combines a novel time-window splitting scheme, a simple median filtering, and an appropriate scaling method to detect potential signals of leakage of CO₂ from geologic carbon sequestration sites from within datasets of net near-surface CO₂ flux measurements. The technique can be directly applied to measured data and does not require subjective gap filling or data-smoothing preprocessing. Preliminary application of the new method to flux measurements from a CO₂ shallow-release experiment appears promising for detecting a leakage signal relative to background variability. The leakage index of ± 2 was found to span the range of biological variability for various ecosystems as determined by observing CO₂ flux data at various control sites for a number of years.

Keywords: Geologic carbon sequestration; monitoring, eddy covariance; CO₂ flux; Leakage detection; Signal filtering;

1. Introduction

Geologic carbon sequestration (GCS) is being widely investigated as one part of a portfolio of approaches needed to reduce greenhouse gas emissions from fossil-fuel energy production (e.g., Pacala and Socolow, 2004). Among the areas of active research is monitoring to ensure the overall effectiveness and safety of GCS (Benson, 2006). Monitoring begins with the injection process, but extends to include monitoring carbon dioxide (CO₂) migration within the reservoir, and may include monitoring of unexpected leakage of CO₂ up wells and conductive faults, and possibly into the near-surface environment. The challenge and expense of subsurface geophysical monitoring motivate development of near-surface monitoring approaches to assure the public and regulators that CO₂ is not seeping out of the ground (e.g., Oldenburg et al., 2003). The challenge in near-surface detection of CO₂ leakage from GCS sites is to discern a leakage signal from within natural background CO₂ variability, especially when the signal is of very small magnitude and/or spatial extent (e.g., Lewicki et al., 2005). Figure 1 shows an example of a raw net CO₂ flux time series measured using the eddy covariance (EC) method over a time frame during which two controlled shallow releases of CO₂ were carried out in the shallow subsurface. Surface CO₂ flux signals from these releases are difficult to detect due to the high variability of background ecosystem fluxes.

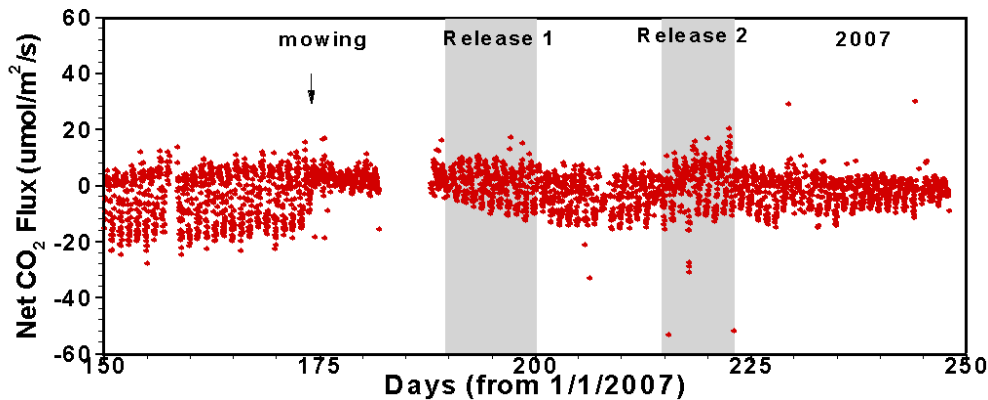


Figure 1. Time series of raw net CO₂ fluxes measured by the EC method during summer of 2007. Gray areas indicate the periods over which CO₂ was released from a shallow subsurface horizontal well (experiment details are given in Lewicki et al. (2007; 2009)).

The objective of this study is to develop a signal-enhancing technique to detect a CO₂ leakage signal from net near-surface CO₂ flux measurements, such as those made using the eddy covariance (EC) approach. Based on an analysis of the diurnal variations in the biological flux, we developed a new technique for detecting leakage events from the time series of measured net CO₂ flux data. The new approach consists of a novel time-window splitting scheme, a simple median filtering, and an appropriate scaling method.

2. Methods

2.1 Release experiments and observation data

Shallow CO₂ release experiments were conducted by the ZERT (Zero Emissions Research and Technology) project team at Montana State University (MSU), Montana Agricultural Experiment Research Center, Bozeman, MT (45°39'N, 111°04'W). The MSU study site was an approximately 0.12 km², nearly flat field, with vegetation composed primarily of prairie grasses, alfalfa, and Canadian thistle (Lewicki et al., 2007; 2009). In July and August 2007, two controlled releases of CO₂ were carried out at different rates from a shallow (~2 m deep) horizontal well of length 70 m, with six packed-off injection zones to simulate CO₂ leakage from the subsurface. Release 1 lasted for 10 days at 0.1 t CO₂ d⁻¹ whereas Release 2 was for 7 days at 0.3 t CO₂ d⁻¹. An EC station was deployed in the field from 8 June 2006 to 4 September 2006, and from 28 May 2007 to 4 September 2007, to measure the average half-hour net CO₂ flux across the atmosphere-canopy interface. Additional information about the site, the release experiments, modeling, and flux measurements is available in (Lewicki et al., 2007; 2009; Oldenburg et al., 2009) and will not be repeated here. Note that the raw measured fluxes were fed to the filter developed in this study without any pre-filtering or gap filling. This is because the objective of this study is to detect the leakage signals, which are basically “abnormal events” or outliers, whereas the most existing gap filling or filtering techniques are usually based on some type of assumption of “normality” (e.g., reflecting average behaviors of the ecosystem) and tend to remove the outliers. The filtering methods based on friction velocity and wind direction as conducted by Lewicki et al. (2009) were not used in this study either, because (1) the time-windows and scaling approaches used in the filter developed in this study tend to cancel the effects of the underestimation of flux caused by low turbulent conditions during nighttime and (2) the leakage location relative to EC station is usually unknown priori.

To test the proposed CO₂ leakage detection technique, we also used several data sets of natural CO₂ flux from other sites as control sites (Table 1). Gap percentage is the fraction (in percent) of data missing from the complete time series and occurs randomly within the “Duration” (period of the time series).

Table 1. Data sets of the observed CO₂ flux (EC flux)

Data set	Location	EcoType	Duration	Gap (%)	Measurement Interval (hr)	Source
MSU2007	45.650N/111.067 W	Grassland	5/28/2007-9/4/2007	14.89	0.5	Lewicki et al. 2007
MSU2006	45.650N/111.067 W	Grassland	6/8/2006-9/4/2006	17.03	0.5	Lewicki et al. 2007
Fermi-Agri	41.859N/88.223W	Corn/soybean rotation	7/11/2005-9/30/2007	15.42	0.5	public.ornl.gov/ameriflux/
Santarem (km67)	2.857S/54.959W	Tropical forest	1/3/2002-1/25/2006	25.50	1.0	Saleska et al 2003
Sgp2002	36.605N/97.489W	Crop (dry year)	1/1/2002-6/6/2002	11.88	0.5	Fischer et al. 2007
Sgp2003	36.605N/97.489W	Crop (wet year)	1/1/2003-12/31/2003	11.99	0.5	Fischer et al. 2007

2.2 Theory

In this section, we will begin with an analysis of biological-flux temporal patterns to find an effective time-domain filtering method by which to discern leakage signals from the time series of measured net CO₂ flux data.

The biological CO₂ flux, $b(t)$, is the net effect of two processes: (1) the uptake caused by plant photosynthesis and (2) the emission caused by the respiration of plant and soil, which can be mathematically described as

$$b(t) = r_p(t) + r_s(t) - p(t) \quad (1)$$

where, r_p , r_s , and p are plant respiration, soil respiration, and plant photosynthesis, respectively. Conceptually, both the respiration and the photosynthesis can be seen as the response of the given ecosystem to the input of meteorological driving forces. Therefore, their temporal variations, although very complex, are basically controlled by two kinds of variables: (1) external driving forces (e.g., solar radiation, precipitation, air temperature, irrigation, and fertilization) and (2) properties of the ecosystem (e.g., biomass, soil temperature and wetness, plant leaf index). Among these controlling variables, solar radiation is the most fundamental one, especially for temporal variation patterns in the biological CO₂ flux. As a result, the most important temporal variation patterns are the diurnal (due to Earth's rotation) and seasonal (due to Earth's orbiting) variations. The quantitative relationships between CO₂ flux and the controlling variables are very complicated, and often nonlinear and interrelated. As a result, accurately predicting the net CO₂ flux of a particular ecosystem is not an easy task, and often requires intensive site characterization and model calibration. However, for our purpose of finding an effective CO₂ surface leakage detection method, understanding the general patterns of the CO₂ flux under natural conditions (i.e., no leakage) using numerical models of the ecosystem is still feasible. Figure 2 shows the simulated photosynthesis, respiration, and net CO₂ flux from a grassland over a one-year period using CLMT2 (Pan et al., 2008; Pruess et al., 1999) with observed meteorological driving-force data as input. Both respiration and photosynthesis fluxes seasonally increase in the summer and during the day (see the smaller inserted figure). The net CO₂ flux shows strong diurnal oscillation besides the seasonal variations, while photosynthesis flux becomes zero during the night.

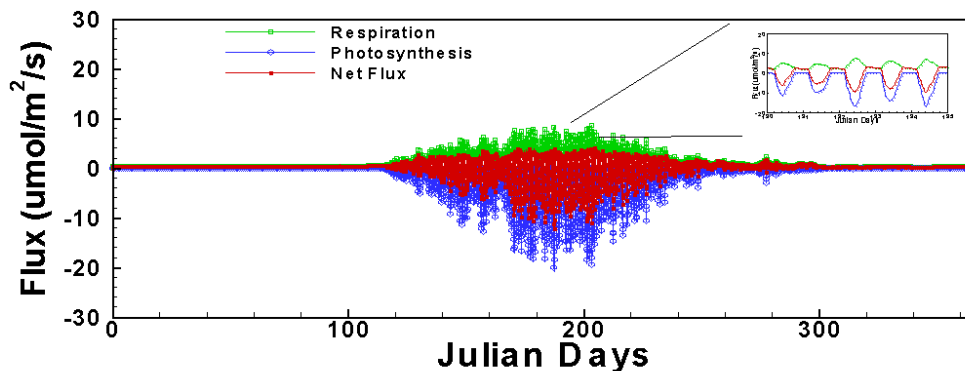


Figure 2. Simulated natural CO₂ fluxes of a grassland

Because the diurnal oscillation is so strong, the first-order approach for describing the temporal features of the net CO₂ flux is to calculate the daily mean flux and standard deviation (STD) of the net flux. Figure 3 shows the mean and STD in the net flux data of the same simulation shown in Figure 2. Both the mean and the STD show strong seasonal variations (Figure 3). The daily mean flux reaches the smallest value in summer because of strong photosynthesis, whereas the daily STD reaches the maximum at the same time. Furthermore, many peaks of both the mean and STD occur during the growing season, possibly because of a transient weather or climate changes (e.g., a cold front, a storm, a drought) or some other events. In a managed ecosystem, the situation would be more complicated, because of management activity such as tillage, irrigation, fertilizer application, harvest, etc.

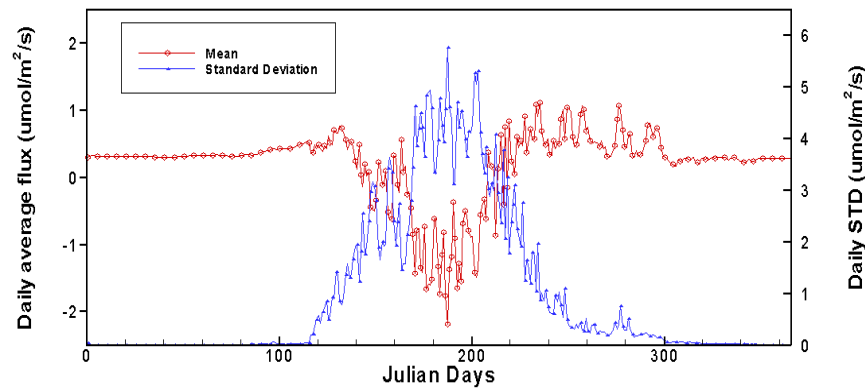


Figure 3. Daily mean and STD of the flux data in Figure 2

The challenge in distinguishing a CO₂ leakage signal from the natural signals of the biological fluxes is that the surface leakage from GCS sites could also be subject to temporal variations, because of the dynamics of subsurface processes, which could be difficult to discern from the peaks of the mean biological flux (as shown in Figure 3). However, no matter how leakage occurs—as prolonged smooth flux or individual eruption events—the factors controlling the subsurface processes are fundamentally different from those factors controlling the biological CO₂ fluxes. Specifically, the biological flux is sensitive to factors that affect plant growth and microbial activity (e.g., sunlight, irrigation, fertilization, and harvest), whereas the leakage flux is a result of subsurface flow processes and is therefore insensitive to those variables. In particular, the leakage flux is insensitive to solar radiation, so that temporal variations in solar radiation (e.g., the strong diurnal and seasonal cycling due to visual movement of sun) will not influence leakage as they do the biological flux. As a result, the leakage flux, unlike the biological flux, is not in general subject to such strong diurnal and seasonal oscillations. Furthermore, an important feature of the biological flux is that a decrease in the mean flux is almost always accompanied by an increase in the STD (Figure 3). This is mainly because the amplitude of the diurnal oscillation is dominated by photosynthesis, the CO₂ uptake process. Although both respiration and photosynthesis contribute to the strong diurnal oscillation of CO₂ flux, their detailed diurnal-variation structures are quite different. Figure 4 shows the simulated photosynthesis flux and respiration flux, the net

flux, and the corresponding measured solar radiation (input) for every 30 minutes during several typical summer days.

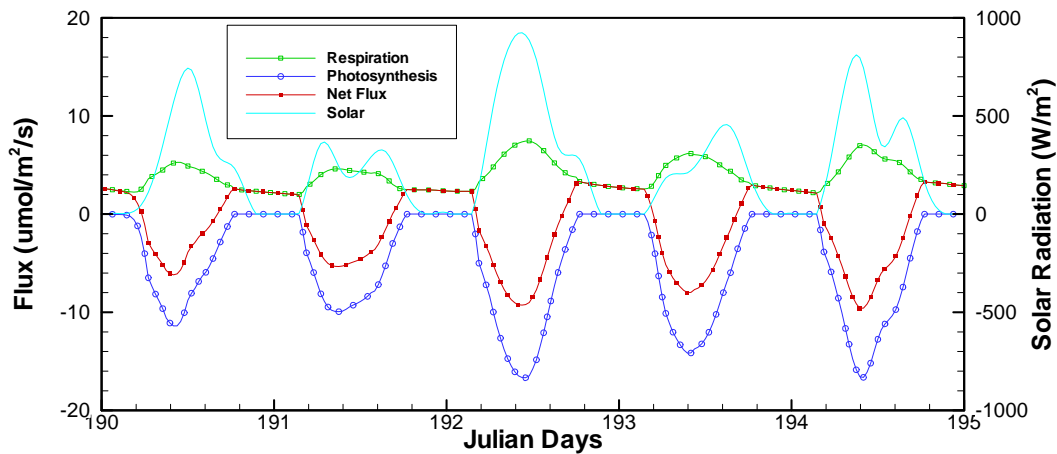


Figure 4. Simulated photosynthesis, respiration, and net CO₂ flux, as well as the observed solar radiation (axis-Y2) during typical summer days using CLMT2.

Clearly, photosynthesis occurs only in the daytime, whereas respiration takes place all the time. Even though respiration is generally higher during daytime than during nighttime, photosynthesis often dominates over respiration during the growing season and is able to pull the net flux down into negative territory during most of the daytime. As a result, the larger the photosynthesis flux, the smaller the daily mean flux and the larger the daily STD (Figure 3). This holds for seasonal variations, too (Figure 3). However, no matter how the mean and the STD vary over time, the photosynthesis flux and the respiration flux during the growing season usually reach a balance (i.e., the net flux reaches zero) at some point within two time windows: (1) the early morning and (2) the late afternoon (Figure 4). Because both photosynthesis and respiration are physiological processes and functions of the ecosystem, they have similar responses to the most-controlling variables (e.g., temperature, soil moisture, soil nutrition, and biomass). In other words, a higher photosynthesis flux is often accompanied by a higher respiration flux. As a result, the net flux within those two time windows (the early morning and the late afternoon) would be much less variable than either the nighttime flux (dominated by respiration) or the noontime flux (dominated by photosynthesis). On the other hand, leakage could lift the daily mean flux, but would not alter the daily STD (at least not decreasing the STD when higher CO₂ concentration might favor photosynthesis).

In terms of the probability distribution for the net CO₂ flux data within a given time window, a leakage event could shift its center, but would not alter its shape. To make the point clear, we define four time windows for each day (Figure 5). While the first three windows are straightforward, the fourth window, W₃, consists of two subwindows (early morning and late afternoon), each one-quarter of a day in length. The time of solar noon is used to label a given day so that each day always starts with a nighttime period. As shown in Figure 5, the net flux typically reaches zero within the W₃ window (twice each day) as a result of balance between respiration and photosynthesis. Our main hypothesis

is that a leakage event would lift the W_3 flux, but not reduce the magnitude of the diurnal variations. This is the basis of the proposed new technique.

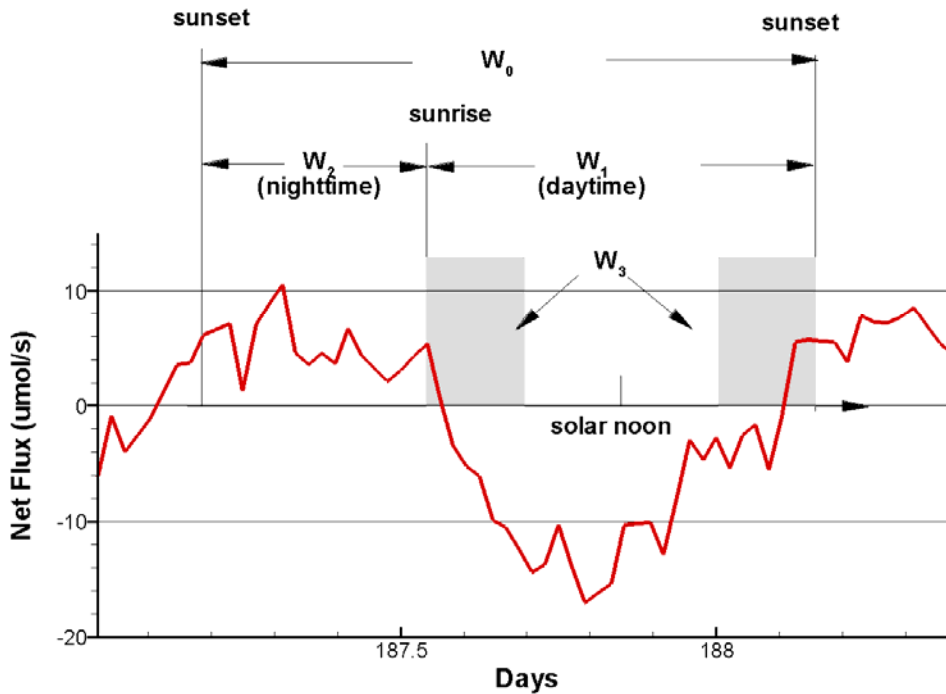


Figure 5. Sketch of the four time windows used for data filtering and their relationships with the typical net flux

2.3 Implementation

The time series of the measured net CO_2 flux data often contain significant noise. Although a number of standard noise-filtering methods are available in the literature (Press et al., 1992), missing data is a common obstacle to filtering. Numerous data-filling methods have been used to fill the data gaps (Falge et al., 2001). However, all filling methods inevitably introduce some assumptions regarding the time series, either its temporal features (e.g., filled with the mean diurnal variations) or its dependence upon other environmental variables (e.g., look-up table method). Therefore, gap-filling is not possible for our purpose, for at least two reasons: First, our objective is to detect possible leakage events from the time series of the measurements, i.e., to detect the “outliers.” Therefore, any estimated data from the “average” behaviors are useless or even misleading. The second reason is that it is impossible to include unknown leakage events in a look-up table that often describes “normal” relationships between the biological flux and the external driving forces and/or the status of the ecosystem. To solve the problems of noise and missing data together, we use the median and the mean absolute deviation as a robust filtering method to extract the information (i.e., the center of the probability distribution) from the measured net CO_2 flux data, which often have gaps.

The median of a distribution, x_{med} , can be calculated from x_j for $j = 1, \dots, N$ depending on whether N is even or odd (Press et al., 1992):

$$x_{med} = \begin{cases} x_{(N+1)/2} & N \text{ odd} \\ \frac{1}{2}(x_{N/2} + x_{N/2+1}) & N \text{ even} \end{cases} \quad (2)$$

The four median values of CO₂ flux, m_0 , m_1 , m_2 , and m_3 , are calculated for the data within the four windows, W_0 , W_1 , W_2 , and W_3 , respectively (Figure 5), following Equation (2) where N is the number of available flux data values (x) (i.e., excluding the missing data points). If 60% or more of the data points within either W_1 (daytime) or W_2 (nighttime) are missing, the value of m_3 is always set to be zero, so that no unreliable leakage indications will occur.

To estimate the magnitude of the diurnal variation, we calculate the mean absolute deviation, D_0 , defined by

$$D_0 = \frac{1}{N} \sum_{j=1}^N |x_j - m_0| \quad (3)$$

Note that D_0 is needed for each day to calculate the growth factor (discussed later), even though the data gap is too big to estimate m_3 for that day. Therefore, two rules are used to estimate the value of D_0 , depending on the availability of data points. The first is set D_0 to be the value of the previous day if either daytime or nighttime has some data present but more than 60% of the data are missing. The second, for the case in which no data exist at all for that day, is for D_0 to be linearly interpolated between the last D_0 before the gap and the first D_0 after the gap. In both cases, m_3 is also set to zero, because the available data are considered to be insufficient to have a reliable estimate for that day, as mentioned above.

To filter out possible isolated peaks in the given data series due to noise, we define a three-point median filter, $Med^i(x)$, as follows:

$$Med^i(x) = Median(x^{i-1}, x^i, x^{i+1}) \quad (4)$$

where the superscript indicates the position of a data in the series.

The leakage index of the i th day, LI^i , is defined as a scaled W_3 flux as follows:

$$LI^i = \max(0, f^i) \alpha^i \beta \quad (5)$$

where, α and β are the growth and ecosystem factors, respectively, to be described below. The W_3 flux of the i th day, f^i , is defined as follows:

$$f^i = \begin{cases} Med^i(m_3) & \text{if } \alpha^i > 0 \\ Med^i(0.5 m_2) & \text{if } \alpha^i \leq 0 \end{cases} \quad (6)$$

The growth factor of the i th day, α^i , is defined as follows:

$$\alpha^i = \max \left[\frac{0.5(m_2^i + m_2^{i+1}) - m_1^i}{\frac{1}{14} \sum_{j=1}^{14} D_0^{i-j}}, -1 \right] \quad (7)$$

Note in Equation (7) that the i th and the next nighttime medians envelop the i th daytime median. Therefore, the numerator actually measures how deep the net flux can be drawn down during daytime by photosynthesis. If α^i is less than or equal to zero, the photosynthesis is extremely weak and the flux at daytime is higher than that at nighttime (note that the denominator would never be negative), because the dominant respiration processes usually emit more CO₂ responding to the higher temperature during daytime. If no or very weak photosynthesis were the case, a higher median within window W_3 would not necessarily be related to any possible leakage flux, because of the lack of balance between photosynthesis and respiration. Therefore, we propose to use half of the nighttime median ($0.5 m_2$) as an estimate of m_3 (as if there were photosynthesis). Such situations could occur when all the plants are dead, because of either natural reasons or human activities. Furthermore, the denominator of Equation (7), the average diurnal variations of the previous two weeks, can be seen as an estimate of the potential respiration capacity on the current day. As a result, the growth factor is actually the seasonally adjusted strength of the photosynthesis. The larger the growth factor, the higher the probability that a larger W_3 flux will indicate a leakage event. Therefore, we use it as a scaling factor to represent the adjustments caused by changes in plant growth status including seasonal or non-seasonal changes (e.g., harvest). In the case of no photosynthesis (i.e., negative α), α no longer reflects the strength of photosynthesis and the W_3 flux is no longer the median flux within the window W_3 . Therefore, the factor loses its original meaning, and we limit its magnitude to less than 1. This ad hoc approach to deal with the situations of no photosynthesis, although possibly not be the best solution and subject to further investigation, will be workable for most cases because the temporal variability of CO₂ flux at sites without vegetation or when not in the growing season is always small. Therefore, detection of a leakage event for those situations is relatively easy.

As discussed above, the W_3 flux actually reflects the relative strength of the respiration flux over the photosynthesis flux under natural conditions. Obviously, it varies for different ecosystems. For example, it would be very different for a mature tropical forest than it would be for a growing temperate cropland, because of very different temperature-moisture patterns and growing stages. To make the leakage index comparable between different ecosystems, we introduce the ecosystem factor, β , analogous to the soil respiration rate (Qi et al., 2002):

$$\beta = \frac{C}{D_{\max} \exp(0.069 T_m)} \quad (8)$$

where C is a constant that can be seen as an indicator of the respiration strength of the reference ecosystem (we use $C = 10Q_{10} = 20$, equivalent to a site with $D_{max} = 10$ and $T_m = 10$), and D_{max} is the maximum D_0 (diurnal variation) of a given ecosystem over a period that includes a growing season. The coefficient (0.069) is obtained as $\ln(Q_{10})/10$ for $Q_{10} = 2$ (Fisher et al., 2007). T_m is the mean annual air temperature of the given ecosystem. Here, D_{max} is used to measure the maximum diurnal variability of the given ecosystem (which will not be sensitive to any leakage events, but will reflect the overall biological capability of the ecosystem). The term representing mean annual air temperature (T_m) is used to represent an estimate of the strength of the natural (base) respiration of a given ecosystem. The procedure for processing the data is presented graphically in Figure 6.

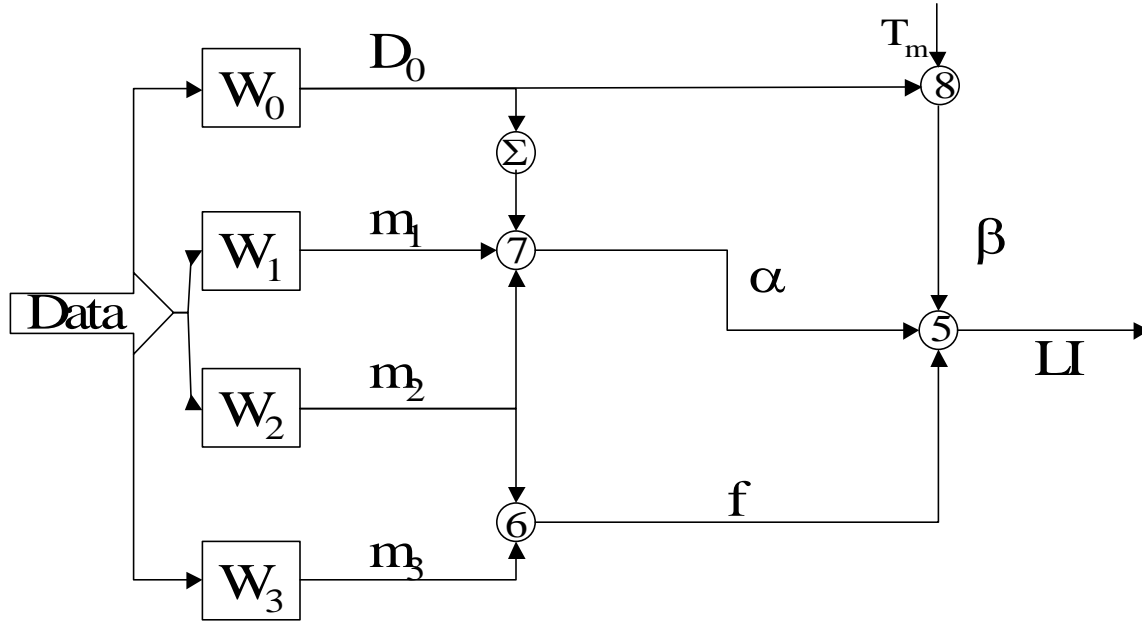


Figure 6. Flow chart of the data processing through four time windows: W_0 (entire day), W_1 (daytime), W_2 (nighttime), and W_3 (the first and the last quarter of the daytime), to the leakage index (LI). The numbers (6, 7, 8, and 9) indicate the corresponding equation numbers.

As can be seen, we must determine the sunrise and sunset times for each day. Although the measured photosynthetically active radiation (PAR) data are available at many sites, the noise in those data will make it difficult to accurately determine the times of sunrise and sunset when PAR values should be close to zero. To avoid the trouble caused by noise on the measured PAR data, the times of sunrise and sunset for each day at a particular site (given the longitude and latitude) are calculated as the roots of the equation of zero solar zenith angle using the bisection method:

$$\cos \mu = \sin \phi \sin \delta + \cos \phi \cos \delta \cosh = 0 \quad (9)$$

where h is the solar hour angle, δ is the solar declination angle, and ϕ is the latitude of the location. The calculations of the solar zenith angles are based on Meeus' algorithms (1998). A half hour is added to the day length of each day (evenly distributed) to reflect the atmospheric refraction effects. The calculated sunrise and sunset times were shifted by a small constant based on visual comparison with the locally measured PAR curves.

3. Results and Discussion

The net CO₂ flux from the area of the ZERT shallow-release facility (consisting of prairie grasses, alfalfa, and Canadian thistle) was measured every 30 minutes during the summers of 2006 and 2007, respectively, using the EC method (Lewicki et al., 2009). Approximately 17% and 15% of the data points were missing in the raw CO₂ flux data sets of 2006 and 2007 (excluding the gap between Days 182–186), respectively (Figure 7).

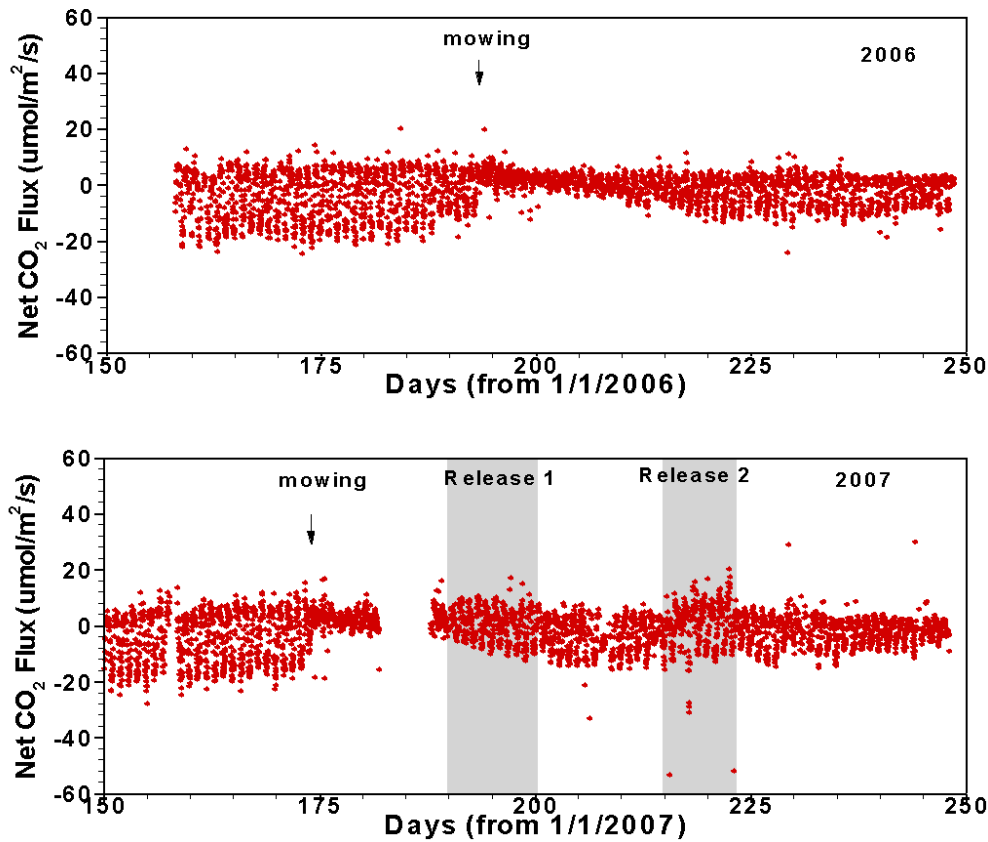


Figure 7. Measured net CO₂ fluxes during summer seasons of 2006 (a) and 2007 (b). Gray areas indicate the periods when CO₂ was released from the subsurface shallow horizontal well.

As shown in Figure 7, besides natural variations, some special events happened during the observation periods. Specifically, the field was mowed in both 2006 (at Day 193) and 2007 (Day 173), and two shallow subsurface CO₂ release experiments were carried out during Days 190–199 and 215–222 in 2007. The variations in flux are quite large, making it difficult to discern the leakage signals from the other signals in the plots. This is

especially the case for the first release, during which the mean flux decreased with time, mainly because of the increase in photosynthesis as a result of the regrowth of the mowed grasses.

For comparison purposes, we plot the (EC-) measured net CO₂ flux data from three different ecosystems (Figure 8). At the Femi-Agri site, the crop is soybean (2005 or 2007) or corn (2006) growing in summer. Besides diurnal variations, the seasonal variation in the net flux is obvious. During summer, both nighttime respiration flux and daytime photosynthesis flux reach their peak values (Figure 8a). For winter crops at the SGP site, the peak photosynthesis occurs during April–May. The year 2002 is relative dry, whereas 2003 is wet in terms of precipitation (Fischer et al., 2007). Note that 2002 only has about 150 days of EC data available because of the monitoring schedule. In contrast to the temperate farmlands, the mature tropical forest at Santarem shows unique behavior for the temporal patterns of net CO₂ flux. First, its nighttime positive flux and daytime negative flux are very similar. Second, the seasonal and interannual variations are much less than those of the temperate farmlands. The latter are often subject to both human management (e.g., irrigation, fertilization, or even rotation with different crops) and the stronger seasonal changes of climate. This diversity of ecosystems and conditions provides a good control case for the proposed detection method.

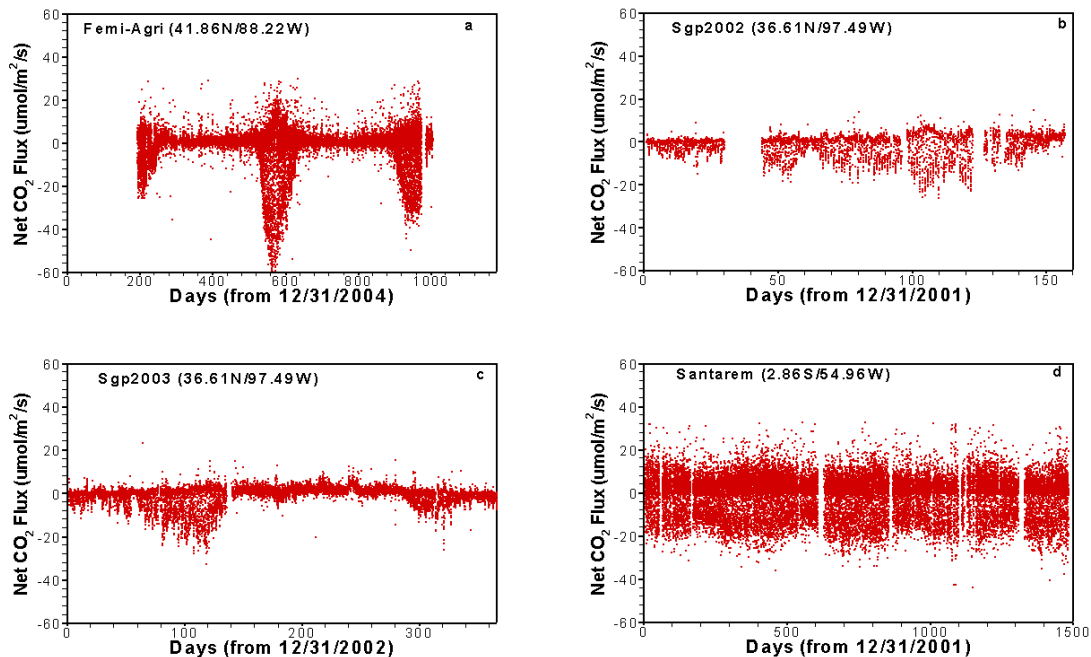


Figure 8. Measured net CO₂ flux data at different sites over multiple years: (a) the Femi-Agri site is a managed farmland with Soybean/corn rotation; (b) and (c) SGP (2002-dry year and 2003-wet year, respectively) is also a managed farmland but with winter crop only; (d) Santarem is a mature tropical forest.

In Figure 9 we present the leakage index curves calculated by the method outlined in Figure 6 for the MSU site for 2006 and 2007 observations, respectively. During most of the growing seasons, the leakage index is very close to zero, because the median CO₂

flux within the W_3 window is close to or below zero. However, two peaks corresponding to the two shallow release events in 2007 very clearly stand out above the other small peaks. Among those small peaks, two are known to be caused by the mowing events (Day 173 in 2007 and Day 193 in 2006). The negative peak during 2006 reflects the period during which daytime flux is larger than nighttime flux, because of the extremely low photosynthesis activity after mowing. Some other peaks also exist, possibly corresponding to short-term meteorological variations or unknown farm-management activities. However, those non-leakage peaks are usually small or isolated (i.e., a single day peak). Their magnitudes are usually less than two. If we smooth the leakage index with a three-point moving average (i.e., a 3-day moving average), all non-leakage peaks are within the range of ± 2 of the leakage index, whereas the leakage peaks rise to 8. Therefore, $|LI| > 2$ can be used as a criterion to indicate that a leakage event has occurred.

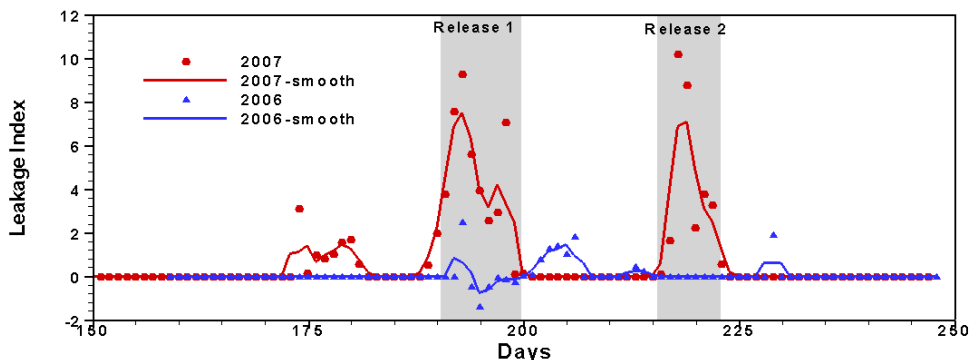


Figure 9. The Leakage Index for 2006 (blue, no shallow release) and 2007 (red, with two shallow-release events marked by gray areas). The lines are the three-point moving average curves.

To test the proposed filtering approach on the data observed over multiple seasons under various ecological and geographical conditions, but without known leakage, we plot the calculated leakage index and its three-point moving averages on Figure 10. At temperate sites, the leakage index is higher in summer no matter whether there is significant crop growth (e.g., Femi-Agri, Figure 10a) or not (e.g., SGP site, Figure 10b and c). This may be because the respiration of the ecosystem is favored at higher summer temperature, so as to elevate the median CO_2 flux within the W_3 window (early morning and late afternoon). Although the crop yield at the SGP site was significantly affected by the amount of the precipitation in the growing season (Fischer et al., 2007), the leakage index seems not to be sensitive to such interannual variations (Figure 10b and 9c; the crop growth ended by about the end of May). The peak after Day 150 most likely results from the effects of the harvesting of the crop, very similar to the situation from mowing that we see at the MSU site (Figure 9). Because there are no data during most of the summer in 2002 at the SGP site, we cannot clearly see the effect of such yearly precipitation variation on the leakage index behavior in the summer season. However, the results at the Femi-Agri site show that the leakage index is quite stable between neighboring years, even though the corn year (see the middle of Figures 9a and 7a) has larger diurnal variations and usually accumulate more biomass than the soybean year. This is key to a leakage index effectively detecting any possible leakage without false alarm, due to a change in the ecosystem owing to natural or management events. At the mature tropical

forest site, respiration and photosynthesis are more balanced than in the farmland. As a result, the W_3 flux is almost always positive year round, and so is the leakage index (Figure 10d). As shown in Figure 10d, the interannual variation in leakage index at Santarem is at least no less than the seasonal variation. In all cases, the leakage index is within the range of ± 2 , except for a few isolated points (Figure 10).

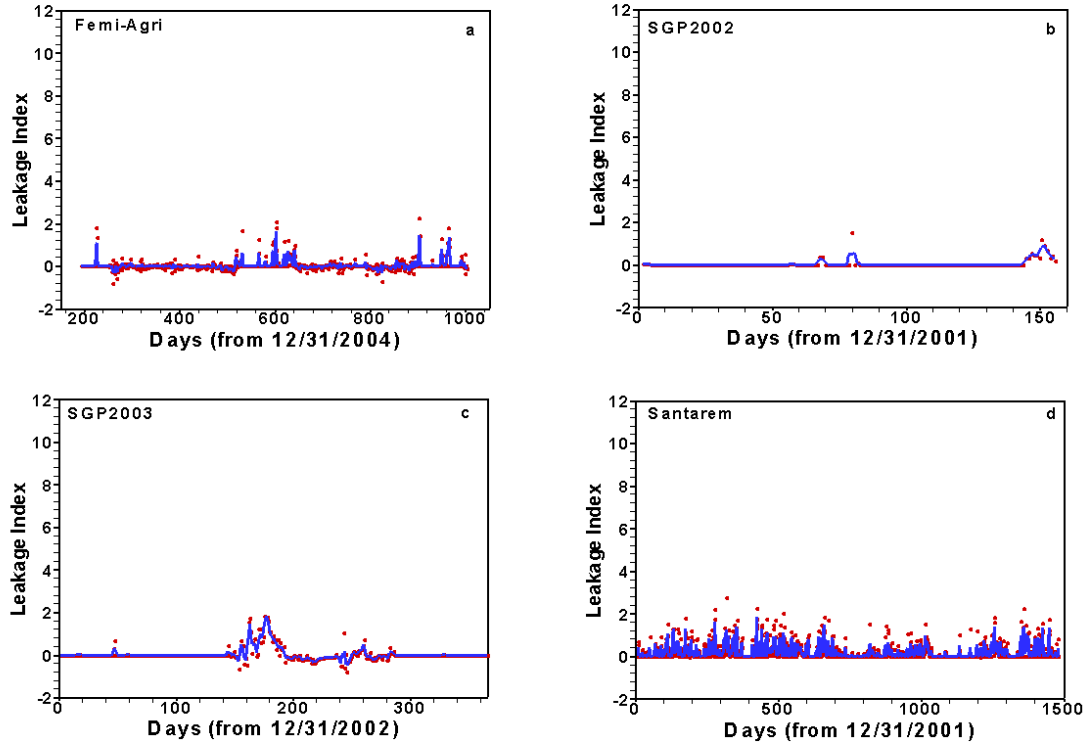


Figure 10. The leakage index (dots) and its three-point moving average (line) at four (non-leaking) sites: (a) Femi-Agri (7/11/2005-9/30/2007); (b) SGP2002 (1/1/2002-6/6/2002); (c) SGP2003 (1/1/2003-12/31/2003); and (d) Santarem (1/3/2002-1/26/2006).

A three-point moving average can remove all of these isolated peaks and present a clear picture of leakage events in terms of the leakage index (Figure 11). Typically, the leakage index peaks resulting from natural driving-force variations or land management activities at farmland sites (i.e., MSU, Femi-Agri, and SGP) usually occur from June to September (Days 150–270), whereas the peaks at tropical forest sites (Santarem) occur year around, probably because of much fewer seasonal variations in climate. The fact that all the nonleakage peaks are within the range of ± 2 implies that β is an effective scaling factor, one that makes possible the comparison of data from different geographic and ecological systems. The maximum diurnal variation and the mean annual temperature, both of which are insensitive to leakage events, are the best choices for characterizing varied ecosystems in terms of detecting abnormal CO_2 flux due to leakage events.

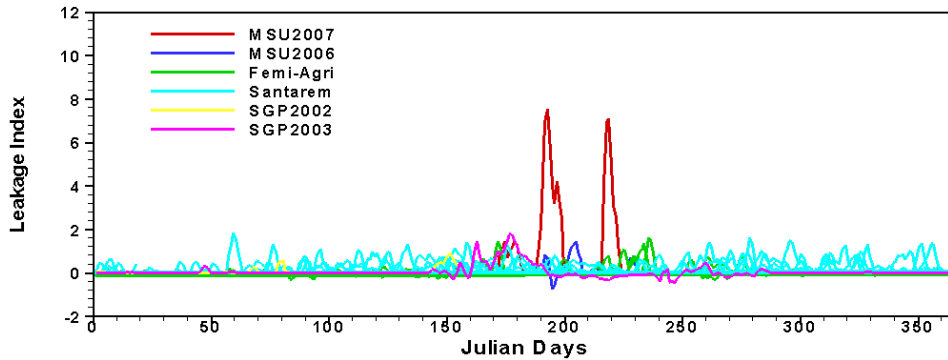


Figure 11. Comparison of the leakage index (the three-point moving average) with Julian days (the first day of each year is Day 1) among all sites. The two peaks at MSU2007 (red line) corresponding to two release events are very clear.

Although comparison between different ecosystems is valuable for understanding the patterns of natural CO₂ flux, the interannual comparison of the leakage index at the same site is probably more critical to the goal of discerning leakage signal from the background biological signals in the time series of the net CO₂ flux. At the MSU site, two leakage peaks are much higher than the other peaks (e.g., those caused by mowing), in the summers of 2006 and 2007 (Figure 9), which make it very easy to tell when the leakage **starts** and ends. As shown in Figure 10a (Femi-Agri site), the crop type does not significantly affect the leakage index peaks (corn in 2006 summer and soybean in 2007 summer), even though the photosynthesis of corn is much stronger than that of soybeans in terms of CO₂ assimilation (Figure 9a). Furthermore, the growth of winter crop at the SGP site in both 2002 (a dry year) and 2003 (a wet year) seems not to contribute to any leakage index peaks until the crop was harvested in early summer. This could result from the relatively low temperature in winter and spring causing low respiration flux compared to the photosynthesis flux, so that the balance between photosynthesis and respiration leans toward photosynthesis within the W_3 window. As a result, the corresponding W_3 flux and (in turn) the leakage index are often close to zero. These results confirm that the proposed filtering method is quite stable with respect to seasonal, ecological, and geographical variations.

4. Conclusions

A new method for detecting potential surface leakage signals of CO₂ from GCS sites from naturally variable flux data in the near-surface environment has been developed. The new method utilizes the unique temporal features in the biological CO₂ flux caused by the competition between photosynthesis and respiration processes. Specifically, the median flux within the W_3 time window reflects the natural balance and is found to be an effective and stable indicator of any abnormal flux of the given ecosystem. The different effects on the magnitude of diurnal variations between a leakage event and the biological events (e.g., harvest) offer an excellent opportunity for identifying leakage events from other abnormal CO₂ flux events. A scaling factor consisting of the maximum diurnal variation and the mean annual temperature of a given site, analogous to soil respiration rate, makes it possible to compare the results between different ecological and

geographical sites. Exceeding the leakage index range of ± 2 can be used as a criterion for determining whether or not a leakage event occurs.

The new method consists of novel time-window splitting, a simple median filtering, and an appropriate scaling with the ecosystem and plant-growth factors. It can be directly applied to the measured data and does not need any subjective gap filling or data-smoothing preprocessing.

Application of the new method to the EC measured data from the ZERT shallow-release tests of 2006 (no CO₂ release events) and 2007 (two CO₂ release events) shows that the proposed method can effectively detect leakage signals, even when the daily mean flux was decreasing (e.g., during the first release experiment). The results from applications to the data from three very different control sites (from managed farmlands to natural tropical forest) show that the new method accommodates natural peaks without producing false positive leakage peaks. The leakage index is quite stable with respect to handling seasonal, interannual, and vegetation variations.

Acknowledgments

This work was funded by the Assistant Secretary for Fossil Energy, Office of Sequestration, Hydrogen, and Clean Coal Fuels, NETL, of the U.S. Department of Energy under Contract DE-AC02-05CH11231. We thank the U.S. Department of Energy as part of the Atmospheric Radiation Measurement Program Climate Research Facility for use of the ARM-SGP flux data and thanks also to Dr. Roser Matamala and Dr. Steven Wofsy as well as their coworkers for the use of data from their sites.

5. References

- Benson, S.M., September 24–27, 2006, Monitoring carbon dioxide sequestration in deep geologic formations for inventory verification and carbon credits, Society of Petroleum Engineers, SPE102833, SPE Annual Technical Conference and Exhibition, San Antonio, TX.
- Bonan, G.B., K.W. Oleson, M. Vertenstein, S. Levis, X.B. Zeng, Y.J. Dai, R.E. Dickinson, and Z.L. Yang, 2002, The land surface climatology of the community land model coupled to the NCAR community climate model, *Journal of Climate*, 15, 3123-3149.
- Falge, E. et al., 2001, Gap filling strategies for long term energy flux data sets. *Agricultural and Forest Meteorology*, 107, 71–77,
- Fischer, M. L., D. P. Billesbach, J. A. Berry, W. J. Riley, and M. S. Torn, 2007, Spatiotemporal variations in growing season exchanges of CO₂, H₂O, and sensible heat in agricultural fields of the Southern Great Plains. *Earth Interactions*, 11, 1–21.
- Lewicki, J.L., G.E. Hilley, and C.M. Oldenburg, 2005, An improved strategy to detect CO₂ leakage for verification of geologic carbon sequestration, *Geophys. Res. Lett.*, 32, L19403, LBNL-57414.
- Lewicki, J. L., C. M. Oldenburg, L. Dobeck, and L. Spangler, 2007, Surface CO₂ leakage during two shallow subsurface CO₂ releases. *Geophys. Res. Lett.*, 34:L24402.

- Lewicki, J.L., G.E. Hilley, M.L. Fischer, L. Pan, C.M. Oldenburg, L. Dobeck, L. Spangler, 2009, Eddy covariance observations of leakage during shallow subsurface CO₂ releases, *J. Geophys. Res.*, 114: D12302, doi:10.1029/2008JD011297.
- Meeus, 1998. *Astronomical algorithms* (2nd Edition), Willmann-Bell.
- Oldenburg, C.M., J.L. Lewicki, and R.P. Hepple, October 2003, Near-surface monitoring strategies for geologic carbon dioxide storage verification, Lawrence Berkeley National Laboratory Report *LBL-54089*.
- Oldenburg, C.M., J.L. Lewicki, L. Dobeck, and L. Spangler, Modeling Gas Transport in the Shallow Subsurface During the ZERT CO₂ Release Test, *Transport in Porous Media*, *LBL-1529E*, in press.
- Pacala, S., and R. Socolow, August 2004, Stabilization wedges: Solving the climate problem for the next 50 years with current technologies, *Science* 305(5686), 968-972, 13.
- Pan, L., J. Jin, N. Miller, Y.-S. Wu, and G. Bodvarsson, 2008, Modeling hydraulic responses to meteorological forcing: From canopy to aquifer. *Vadose Zone Journal*. 7:325-331.
- Press et al., 1992, *Numerical Recipes* (2nd Edition), Cambridge University Press.
- Pruess, K., 2008, On CO₂ fluid flow and heat transfer behavior in the subsurface following leakage from a geologic storage reservoir, *Environmental Geology*, 54, 1677-1686.
- Pruess, K., C.M. Oldenburg and G.J. Moridis, November 1999, TOUGH2 User's Guide Version 2. E. O. Lawrence Berkeley National Laboratory Report *LBL-43134*.
- Qi, Y., M. Xu, and J. Wu, 2002, Temperature sensitivity of soil respiration and its effects on ecosystem carbon budget: nonlinearity begets surprises. *Ecological Modeling* 153:131-142.
- Saleska, S. R. and coauthors, 2003, Carbon in Amazon Forests: Unexpected seasonal fluxes and disturbance-induced losses. *Science* 302:1554-1557.