Ecological Informatics 5 (2010) 492–502

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Uncertainty analysis of eddy flux measurements in typical ecosystems of ChinaFLUX

Honglin He [a], Min Liu [a,b], Xiaomin Sun [a], Li Zhang [a], Yiqi Luo [c], Huiming Wang [a], Shijie Han [d], Xinquan Zhao [e], Peili Shi [a], Yanfen Wang [b], Zhu Ouyang [a], Guirui Yu [a,*]

[a] Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China
[b] Graduate University of Chinese Academy of Sciences, Beijing 100049, China
[c] Department of Botany and Microbiology, University of Oklahoma, 770 Van Vleet Oval, Norman, OK 73019, USA
[d] Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang, Liaoning 110016, China
[e] Northwest Plateau Institute of Biology, Chinese Academy of Sciences, Xining, Qinghai 810001, China

ARTICLE INFO

Article history:
Received 9 January 2010
Received in revised form 6 July 2010
Accepted 9 July 2010

Keywords:
Eddy covariance
Measurement error
Random error
Flux
Uncertainty
ChinaFLUX

ABSTRACT

Fluxes of CO_2 (FCO_2) and energy (latent heat, LE; sensible heat, H) exchange between ecosystems and atmosphere, as measured by the eddy covariance technique, represent a fundamental data source for global-change research. However, little is known about the uncertainties of flux measurements at an ecosystem level in China. Here, we use data from six eddy covariance tower sites in ChinaFLUX, including two forested sites, three grassland sites, and one agricultural site, to conduct a cross-site analysis of random flux errors (RFEs) of FCO_2, LE, and H. By using the daily-differencing approach, paired observations are obtained to characterize the random error in these measurements. Our results show that: (1) The RFEs of FCO_2, LE, and H in different ecosystems of ChinaFLUX closely follow a double-exponential (Laplace) distribution, presumably due to a superposition of Gaussian distribution for high flux magnitude. (2) The RFEs of FCO_2, LE, and H are not homogeneous and appear to be a linear function of flux magnitude. (3) Except for H, the RFEs of FCO_2 and LE exhibit a distinct seasonal pattern. For FCO_2, the dependence of RFEs on wind speed varies somewhat according to vegetation type, whereas for LE and H, there is no such dependence. The effect of temperature on RFEs is not statistically significant ($P<0.05$). Both the distribution and the relationship of RFEs with flux magnitude in ChinaFLUX are essentially in accord with those in AmeriFlux and CarboEurope.

1. Introduction

The long-term and continuous eddy covariance (EC) measurements of ecosystem fluxes such as CO_2, water, and energy between biosphere and atmosphere at tower sites around the world (e.g., Baldocchi et al., 2001; Yu et al., 2006; Mizoguchi et al., 2009) offer various opportunities to improve our understanding about the fundamental processes of ecosystem functions in both time and space (Baldocchi, 2003; Friend et al., 2007). However, there is a growing recognition within the eddy-flux community that more attention needs to be paid to the uncertainties inherent in these EC measurements (Hollinger and Richardson, 2005; Richardson et al., 2006; Lasslop et al., 2008). With the development of a model-data fusion method in terrestrial ecosystem research, data uncertainties are as important as data themselves and play a major role in determining the outcome (Raupach et al., 2005). Therefore, how to quantify the uncertainty of flux data and acquire the probability density function (PDF) and its statistical characteristics have become a frontier issue in global flux research.

Flux data actually are not deterministic; rather, they can be expressed as the "correct" value plus or minus measurement error, which is called uncertainty. Specifically, a flux measurement (x) represents a sum of the "true" flux (F) and the potential measurement errors, which can be further divided into systematic errors (ε) and random measurement errors (δ), namely $x = \mathrm{F} + \varepsilon + \delta$ (Richardson et al., 2006). The systematic errors and random measurement errors are usually evaluated separately. The energy imbalance and incomplete nocturnal data may cause the systematic errors, which are difficult to identify. However, the systematic errors can be eliminated by calculating the bias. Identifying the source of systematic error and how to reduce this error represent an active research area in flux study (Goulden et al., 1996; Moncrieff et al., 1996; Mahrt, 1998; Twine et al., 2000; Massman and Lee, 2002; Morgenstern et al., 2004). In contrast to systematic error, random error is related to the observational systems (e.g., gas analyzers, ultrasonic apparatus, data acquisition system, and the calculation method), turbulent transport, and the heterogeneity in flux footprint (Moncrieff et al., 1996). In most cases, random measurement errors cannot be eliminated, but their numerical value can be obtained by statistic analysis. Here, as regards the uncertainty of flux observation data, we mainly focus on RFEs.

Extensive studies on the random errors of EC data have been conducted by a repeated sampling method in a single tower or twin towers (Hollinger and Richardson, 2005; Richardson et al., 2006; Rannik et al., 2006) or statistical analysis of model residuals (Hagen et al., 2006; Chevallier et al., 2006; Lasslop et al., 2008). Hollinger and Richardson

* Corresponding author. Tel.: +86 010 6488 9268; fax: +86 010 6488 9399.
E-mail address: yugr@igsnrr.ac.cn (G. Yu).

1574-9541/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecoinf.2010.07.004

(2005) studied flux data uncertainty by using repeated sampling method in two nearby towers in Howland Forest, pointing out that random measurement error follows a double-exponential (Laplace) distribution rather than a normal (Gaussian) distribution. Meanwhile, a daily-differencing approach was proposed to quantify the RFEs in a single tower (Hollinger and Richardson, 2005). Rannik et al. (2006) discussed the uncertainty of flux observation data through use of the repeated-sampling method at the same time in two nearby towers (within a distance of 30 m) in Hyytiala of Finland. Richardson et al. (2006) conducted a cross-site study on flux measurement errors in AmeriFlux, including forest, grassland, and farmland ecosystems. They demonstrated that flux measurement errors of different ecosystems follow a double-exponential distribution as well. The relationships between measurement errors and environment variables and flux magnitudes are also examined. Based on model residuals, Richardson et al. (2008) conducted a systematic analysis of the statistical characteristics of CO_2 flux random errors in several forest ecosystems in Europe, and they suggested that the random error analysis method based on model residuals is a supplement to the daily-differencing approach based on single (double)-tower data.

Based on these investigations, a number of studies on parameter optimization and model-data fusion were conducted (Richardson and Hollinger, 2005; Hagen et al., 2006; Lasslop et al., 2008). Regarding analysis of random measurement error, Richardson and Hollinger (2005) compared the effects of two different error distributions (Gaussian distribution versus double-exponential distribution) on model parameter selections. They also explored the relationship between random error and environment factors using the maximum likelihood method for parameter optimization. Hagen et al. (2006) conducted uncertainty analysis of gross ecosystem exchange (GEE) derived from 7 y of continuous eddy covariance measurements in Howland Forest. In China, Liu et al. (2009) analyzed the random error of CO_2 flux measurements, and they employed the bootstrapping method to evaluate different models and optimization methods in influencing the estimate of key parameters and CO_2 flux components. Zhang et al. (2008) explored the effect of error distribution of CO_2 flux on key parameters in ecosystem carbon-cycle models. However, the statistical properties of random measurement errors remain currently under debate and need to be tested and verified at more flux towers around the world (Richardson et al., 2006; Lasslop et al., 2008; Williams et al., 2009).

This paper seeks to obtain the statistical characteristics of RFEs for FCO_2, LE, and H; quantify the uncertainty of flux data; determine its influencing factors; and compare the differences of the RFEs among ChinaFLUX, AmeriFlux, and CarboEurope. Data used in our analyses are from six sites in ChinaFLUX, including two forested sites, three grassland sites, and one agricultural site. We first focus on evaluating the statistical characteristics and distribution of RFEs in CO_2 and energy (latent energy, LE and sensible heat, H). Then we examine the relationship between RFEs and flux magnitude as well as wind speed. The seasonality of RFEs and the influence of temperature on the random error in all three fluxes are also discussed. Finally we compare our results with similar studies for the AmeriFlux and CarboEurope. All these works tend to provide technical support for quantifying flux observation uncertainty and properly evaluating flux observations, which in turn will be helpful for model-data fusion research and model evaluation.

2. Data and method

2.1. Data

Data used in our analyses are obtained from six eddy covariance sites within the ChinaFLUX network, representing a diverse range of ecosystems: subtropical evergreen coniferous plantation (QYZ), northern warm temperate deciduous broad-leaved forest (CBS), Qinghai-Tibet Alpine meadow (HBGC), Qinghai-Tibet alpine grassland meadow (DX), Inner Mongolia typical grassland (NM), and Huang-Huai-Hai farmland (YC). The flux and routine meteorological measurements are operated with the same set of instruments and program at the six forest sites (Yu et al., 2006). For most sites, at least 3 or 4 y of continuous measurements are available. An overview of these sites is given in Table 1. Extensive data and site information are available online at the ChinaFLUX Web site (http://www.chinaflux.org/).

The datasets are processed by using the flux data processing system at ChinaFLUX (Li et al., 2008). The processing includes: (1) coordinate rotation for 30-min flux data (Wilczak et al., 2001), (2) Webb–Pearman–Leuning (WPL) correction (Webb et al., 1980; Leuning, 2004), (3) storage calculation for forested sites (Hollinger et al., 1994), (4) outlier rejection (Papale et al., 2006), and (5) nighttime filtering with u^* threshold obtained by evaluating the relationship between temperature and CO_2 flux (Reichstein et al., 2005).

2.2. Method

2.2.1. Determination of flux uncertainty

Uncertainty associated with the measured eddy covariance flux can be defined as the variance of high-frequency data in average time (e.g., 30 min), which can be detected by taking multiple measurements when the data are relatively independent and the condition is stable and then using the variability of these measurements to estimate the standard deviation. However, flux is usually not stable, because of the influence of phenologic and climate conditions. Therefore, simultaneous measurements from two towers located nearby can be used to meet the assumption of the repeated-sampling method (Hollinger and Richardson, 2005; Rannik et al., 2006). Given the fact that there are very few sites where two adjacent towers can simultaneously measure fluxes for the same ecosystem in ChinaFLUX, we use the daily-differencing approach as described by Hollinger and Richardson (2005) to quantify the random measurement errors. Specifically, a measurement pair (x_1, x_2) is considered valid only if both measurements are made under "equivalent" environmental conditions (PPFD within 75 μ mol m^{-2} s^{-1}, air temperature within 3 °C, and wind speed within 1 m/s) in the same successive two days. These criteria are chosen as a trade off for two conflicting requirements: (1) environmental conditions sufficiently similar that the difference between the measured fluxes can be attributed to random error instead of the differences in forcing variables; and (2) a large enough set of measurement pairs to accurately characterize the probability distribution function (PDF) of the random error (Richardson et al., 2006). Regarding the limitation of sampling length, the sample should be obtained for more than 1 y. We use $(x_1\text{-}x_2)/\sqrt{2}$ to express the measurement errors, δ. The standard deviation of random errors is used to characterize flux measurement uncertainty. Finally, we can estimate the RFEs by calculating the standard deviation of the differences, which is expressed as:

$$\sigma(\delta) = \frac{\sigma(x_1 - x_2)}{\sqrt{2}} \quad (1)$$

2.2.2. Analysis of the RFEs

According to the traditional micrometeorologic method based on turbulence theory (Lenschow et al., 1994; Mann and Lenschow, 1994), the relationship between random measurement error and environmental variables can be described as:

$$\sigma \propto |\overline{F}| \sqrt{\frac{h_\tau}{\bar{u}T}} \quad (2)$$

where $|\overline{F}|$ is the absolute value of mean flux, h_τ is the appropriate height measure for the integral timescale, $\bar{u}$ is the mean wind speed at the measurement height, and T is the sampling period (e.g., $T = 1800$ s for our study). From Eq. (2), we can expect that the flux magnitude and wind speed may be the main factors influencing random measurement error.

In this study, we focus on the scaling of RFEs with $|\overline{F}|$ and $\bar{u}$. The inferred random errors are divided into many bins on the basis of flux