Agricultural and Forest Meteorology 206 (2015) 163–171

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

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

Statistical uncertainty of eddy covariance CO_2 fluxes inferred using a residual bootstrap approach

Huei-Jin Wang*, William J. Riley, William D. Collins

Earth Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

ARTICLE INFO

Article history:
Received 6 April 2014
Received in revised form 12 March 2015
Accepted 17 March 2015
Available online 30 March 2015

Keywords:
Model-data fusion
Multi-model ensembles
Gap-filling comparison
Long-term measurements
Monte Carlo
Net ecosystem CO_2 exchange

ABSTRACT

High-frequency eddy-covariance measurements of net ecosystem CO_2 exchange (NEE) with the atmosphere are valuable resources for model parameterization, calibration, and validation. However, uncertainties in measured data, i.e., data gaps and inherent random errors, create problems for researchers attempting to quantify uncertainties in model projections of terrestrial ecosystem carbon cycling. Here, we demonstrate that a model-data fusion method (residual bootstrap) produces defensible annual NEE sums, through mimicking the behavior of random errors, filling missing values, and simulating gap-filling biases. This study estimated annual NEE sums for 53 site-years based on nine eddy-covariance tower sites in the USA, and found that our annual estimates were, in most cases, comparable in magnitude with those obtained from AmeriFlux gap-filled data. Additionally, compared to the AmeriFlux standardized gap-filling, our approach provides better NEE estimates for moderate to longer, and more frequent, data gaps. Annual accumulated uncertainties in NEE at the 95% confidence level were $\pm$30 gC m^{-2} year^{-1} for evergreen needleleaf forests; $\pm$60 gC m^{-2} year^{-1} for deciduous broadleaf forests; and $\pm$80 gC m^{-2} year^{-1} for croplands. The residual bootstrap performed worst when gap length was greater than one month or data exclusion greater than 90% during the growing season, common to other gap-filling techniques. However, this study produced robust results for most site years when monthly data coverage during the growing season is not extremely low. We therefore suggest that the inclusion of NEE uncertainty estimates and better estimation for moderate to longer, and more frequent, data gaps as provided by the residual bootstrap approach can be beneficial for ecosystem model evaluation.

1. Introduction

Despite progress in developing terrestrial ecosystem models over the past several decades, there is still very limited knowledge of the performance skills of these process models (Schwalm et al., 2010; Keenan et al., 2012; Luo et al., 2012). In order to evaluate and improve performance of terrestrial ecosystem models, more attention needs to be placed on validation against observations. Eddy covariance observations of ecosystem-atmosphere CO_2 exchanges are essential for evaluating dynamics of model predicted fluxes because these net ecosystem exchange (NEE) measurements are on a continuous basis of typical 30-min averaging intervals (Falge et al., 2001; Baldocchi, 2003). However, recent studies have revealed that data uncertainty is a systematic cause of the low agreement between model predictions and observations (Schwalm et al., 2010; Dietze et al., 2011; Keenan et al., 2012). Many of the observational flux datasets used to develop and test ecosystem models are subject to systematic and random measurement errors, which weaken the quality of data and complicate model evaluation (Baldocchi, 2003; Hollinger and Richardson, 2005). Therefore, knowing the makeup of uncertainty in observed data is a prerequisite to quantifying the performance of ecosystem process models.

In the eddy covariance technique, uncertainties of flux measurements can be roughly categorized into systematic and random errors. Systematic errors often occur under stable, low-wind conditions at night due to insufficient turbulence mixing and are notoriously difficult to quantify (Lee, 1998; Loescher et al., 2006). The most common solution to these types of systematic errors is data filtering and data filling. Friction velocity (u^*) filtering has been developed to reject suspicious NEE measurements when u^* falls below a critical threshold (Gu et al., 2005; Barr et al., 2013b), and then data gaps created by u^* thresholds are filled using various gap-filling methods (Falge et al., 2001; Moffat et al., 2007). In addition, instrument failures and data quality controls (Foken and Wichura,

* Corresponding author. Tel.: +1 510 486 6397; fax: +1 510 486 7897.
E-mail address: hjwang@lbl.gov (H.-J. Wang).

http://dx.doi.org/10.1016/j.agrformet.2015.03.011
0168-1923/

1996; Mahrt, 1998) result in further gaps in the data record. In general, data coverage over the course of a year is only ~65% (Falge et al., 2001). Consequently, these extensive, non-random data gaps are a major source of bias in estimating the magnitude of NEE integrals at various timescales, ranging from hours and years. Further, data gaps pose a challenge to quantitatively assess how well terrestrial ecosystem models simulate the processes governing NEE, i.e., gross primary production (GPP) and ecosystem respiration (RE). Because GPP and RE estimates rely only on a small amount of reliable nocturnal NEE measurements and are likely to be biased, they in turn complicate the ecosystem model validation (Reichstein et al., 2005; Desai et al., 2008).

Apart from data gaps, random errors are inherent in flux measurements at non-gap time points. Random errors are stochastic and include turbulence sampling errors, statistical errors associated with time-varying flux footprints, and errors relevant to the measurement equipment, among others (Moncrieff et al., 1996). To characterize this type of data uncertainty, Hollinger and Richardson (2005) compared two adjacent tower measurement series in an evergreen needleleaf forest and found that random measurement errors were double-exponentially distributed with zero means and heteroscedastic variances. This heteroscedasticity depended on the flux magnitude, which varied in time, i.e., flux uncertainty was greater during the growing season than dormant season and greater in the daytime than nighttime. Therefore, these findings suggest that when not account for, this heteroscedastic random uncertainty has the potential to undermine model-measurement intercomparisons. Although Dietze et al. (2011) added artificial double-exponential errors to ecosystem synthetic data for the purpose of assessing model-measurement mismatch, it is unclear from the study of Hollinger and Richardson (2005) to what extent the application of the distribution parameter estimates is appropriate at other sites. In subsequent work, Richardson et al. (2008) used model residuals (mismatches between observed and modeled fluxes) directly to quantify the uncertainty distribution characteristics of a number of CarboEurope sites. However, their residuals did not reflect the nature of flux random errors (as could be inferred from Monte Carlo simulations) and were closely tied to an underlying model structure.

Due to a lack of two adjacent tower measurement series for most sites, Monte Carlo simulations, in conjunction with model residuals, have been used to resolve the problem of estimating uncertainty due to the random nature of any individual NEE observation. Also, when model residuals are resampled and added back to the model output, gap-free flux datasets can be constructed so that uncertainty in sums of flux estimates can be quantified at various timescales (Hagen et al., 2006; Stauch et al., 2008). Conceptually, this method requires a good model to give reasonable residuals, so that the resampled residuals reflect the behavior of the true measurement random errors even though residuals do not have mean zero (Hardle and Bowman, 1988). In this context, Hagen et al. (2006) and Stauch et al. (2008) used empirical models under the Monte Carlo framework. Although these empirical models are closely tuned to the data, their model parameters are tied to "non-gap-point" data and they in turn exert less capacity for extrapolation at gap points. The resampled residuals may therefore not reflect the behavior of random errors at gap points.

In this paper, rather than using empirical models, we used process models to separate residuals from NEE observations, for several reasons. First, process models contain useful prior functional constrains about ecosystem NEE fluxes and maintain mechanistic consistency in gap and non-gap predictions. Second, although process models exhibit persistent bias at certain times of year, they generally can adequately capture the diurnal cycle (Schwalm et al., 2010; Dietze et al., 2011; Stoy et al., 2013). Because our approach does not require mean zero residuals, resampled residuals have the potential to mirror the behavior of measurement random errors. Third, the gulf between process-based and empirical approaches to predicting NEE fluxes may be bridged by the use of process model-data fusion. Because little agreement on model performance metrics exists to separate "good" and "bad" process models (Gleckler et al., 2008; Reichler and Kim, 2008; Luo et al., 2012), using multi-model ensemble means has been advocated because ensemble means generally provide more reliable information than any single model by alleviating individual model bias (Cantelaube and Terres, 2005; Thomson et al., 2006; Schwalm et al., 2010).

The goal of this study is to quantify data uncertainty, in association with random measurement errors and gap-filling errors, from eddy covariance measurements at nine sites spanning three vegetation types. We applied a Monte Carlo approach (residual bootstrap) to simulate multiple runs of gap-free NEE time series, and hence, estimates the mean NEE response at each point in time (pseudo data). To evaluate the degree to which process model errors confound random measurement errors, we differenced posterior residuals from eddy covariance observations and pseudo data in line with non-gap points. Having evaluated the confounded effect, we assessed the performance of residual bootstrap simulations at timescales longer than the measurement time intervals to ensure consistent error propagation. Finally, we inferred the annual NEE sum with uncertainty limits, for the purpose of assessing the consequence of random errors and gap-filling errors in long-term measurements.

2. Materials and methods

2.1. Observed and modeled NEE data

All eddy covariance data used were obtained from the AmeriFlux network (http://public.ornl.gov/ameriflux/). The obtained 30-min NEE values had been processed using a standardized protocol, including storage correction, spike removal, u^* filtering (Gu et al., 2005), and gap-filling using marginal distribution sampling (MDS; Reichstein et al., 2005) or artificial neural network (ANN; Papale and Valentini, 2003). The valid NEE observations (non-gap data) had data coverage ranging between 30% and 70% over the course of a year (Table 1).

Mean model ensemble (mean simulated value across all models) data were analyzed from 15 ecosystem models (Table 1): 13 models obtained through the NACP (North American Carbon Program) interim site synthesis model output (Barr et al., 2013a; Ricciuto et al., 2013), and two versions of the Community Land Model (Lawrence et al., 2011; Koven et al., 2013; Tang and Riley, 2013). Modeled NEE fluxes were model-specific runs using standardized meteorological data, soil types, and management history. Meteorological data, such as air temperature, precipitation, solar radiation, and humidity, were gap-filled using National Oceanic and Atmospheric Administration (NOAA) meteorological station data and Daymet reanalysis products following Ricciuto et al. (2009). Locally observed values of soil texture and management history by model simulations were given by the AmeriFlux BADM templates (Law et al., 2008). All models were simulated at a 30- or 60-min step using the standardized meteorological data as driving variables (http://nacp.ornl.gov/mast-dc/docs/Site_Synthesis_Protocol_v7.pdf).

Concerning the interannual variation in NEE provided by Monte Carlo simulations, we selected sites in the AmeriFlux network across the U.S. with at least five years of data collected between 2000 and 2007 and at least nine model outputs, with plant functional types that were represented by at least three sites. This resulted in nine eddy covariance sites spanning 53 site-years. Of these sites, three were characterized as evergreen needleleaf forest (US-Ho1, US-Me2, and US-NR1), three as deciduous broadleaf forest