

## Statistical uncertainty of eddy flux–based estimates of gross ecosystem carbon exchange at Howland Forest, Maine

S. C. Hagen,<sup>1</sup> B. H. Braswell,<sup>1</sup> E. Linder,<sup>2</sup> S. Frolking,<sup>1</sup> A. D. Richardson,<sup>1</sup>  
and D. Y. Hollinger<sup>3</sup>

Received 29 April 2005; revised 29 September 2005; accepted 20 October 2005; published 11 March 2006.

[1] We present an uncertainty analysis of gross ecosystem carbon exchange (GEE) estimates derived from 7 years of continuous eddy covariance measurements of forest-atmosphere CO<sub>2</sub> fluxes at Howland Forest, Maine, USA. These data, which have high temporal resolution, can be used to validate process modeling analyses, remote sensing assessments, and field surveys. However, separation of tower-based net ecosystem exchange (NEE) into its components (respiration losses and photosynthetic uptake) requires at least one application of a model, which is usually a regression model fitted to nighttime data and extrapolated for all daytime intervals. In addition, the existence of a significant amount of missing data in eddy flux time series requires a model for daytime NEE as well. Statistical approaches for analytically specifying prediction intervals associated with a regression require, among other things, constant variance of the data, normally distributed residuals, and linearizable regression models. Because the NEE data do not conform to these criteria, we used a Monte Carlo approach (bootstrapping) to quantify the statistical uncertainty of GEE estimates and present this uncertainty in the form of 90% prediction limits. We explore two examples of regression models for modeling respiration and daytime NEE: (1) a simple, physiologically based model from the literature and (2) a nonlinear regression model based on an artificial neural network. We find that uncertainty at the half-hourly timescale is generally on the order of the observations themselves (i.e., ~100%) but is much less at annual timescales (~10%). On the other hand, this small absolute uncertainty is commensurate with the interannual variability in estimated GEE. The largest uncertainty is associated with choice of model type, which raises basic questions about the relative roles of models and data.

**Citation:** Hagen, S. C., B. H. Braswell, E. Linder, S. Frolking, A. D. Richardson, and D. Y. Hollinger (2006), Statistical uncertainty of eddy flux–based estimates of gross ecosystem carbon exchange at Howland Forest, Maine, *J. Geophys. Res.*, *111*, D08S03, doi:10.1029/2005JD006154.

### 1. Introduction

[2] Efforts to accurately predict patterns of carbon dioxide exchange between terrestrial ecosystems and the atmosphere are currently limited by our ability to represent the relevant biogeochemical processes in unifying models, which typically parameterize fluxes as a function of environmental variables. Models of the global carbon cycle need to accurately capture the dynamics of terrestrial biosphere-atmosphere exchange at a range of timescales, because forcings and responses occur across a broad temporal spectrum, from seconds (e.g., light capture by leaves) to years (e.g., community dynamics). Field

biometric studies have historically been used to validate model predictions at long timescales, and evaluation of the rapid ecophysiological mechanisms has been limited to important, but temporally sparse, leaf and soil chamber measurements.

[3] In the past decade, at several hundred locations around the world, eddy flux tower measurement programs have been established to quantify ecosystem-atmosphere CO<sub>2</sub> exchange with high-frequency, near-continuous, multi-year measurements. These net ecosystem exchange (NEE) measurements provide another data source for ecosystem model evaluation. One primary advantage of using eddy flux data for process studies and model evaluation is the continuity of the measurements, with time intervals typically 0.5–1 hour. Many time series are now between 5 and 15 years in duration (e.g., Harvard Forest [Wofsy *et al.*, 1993], Walker Branch Watershed [Balldocchi and Vogel, 1996], and Howland Forest [Hollinger *et al.*, 2004]). Another advantage is that the measurements are associated with a growing and coordinated effort (e.g., AmeriFlux) to establish networks of

<sup>1</sup>Complex Systems Research Center, University of New Hampshire, Durham, New Hampshire, USA.

<sup>2</sup>Department of Mathematics and Statistics, University of New Hampshire, Durham, New Hampshire, USA.

<sup>3</sup>U.S. Department of Agriculture Forest Service NE Research Station, Durham, New Hampshire, USA.

towers that span a range of ecosystem types and environmental conditions. Also, eddy flux sites tend to be foci for a suite of other measurements including meteorological variables, biometry, and other types of flux measurements. The primary disadvantage, with respect to understanding terrestrial biogeochemistry, is that measurements of eddy flux do not themselves directly quantify specific ecosystem processes but rather the net result of several processes. Of secondary concern are occasional instrument failures and other normal data collection gaps and errors.

[4] Net ecosystem exchange observations record the typically small imbalances between the gross component fluxes of ecosystem respiration and photosynthesis [Wofsy *et al.*, 1993], and while NEE data can be compared to model predictions, it is often more desirable to validate modeled component fluxes independently. The gross fluxes individually reflect distinct sets of processes whose mechanisms might influence one another but are largely separable. The net flux does not constrain the overall dynamics as well as the component fluxes because the net flux could be mistakenly modeled by gross fluxes having large compensating errors. Furthermore, some models, for example those driven by remote sensing observations, focus on uptake by photosynthesis, also known as gross ecosystem exchange (GEE), with little or no attempt to predict respiration [e.g., Prince and Goward, 1995; Xiao *et al.*, 2004]. Models such as these require independent GEE estimates for validation, and eddy flux observations of NEE can be useful in estimating these independent GEE data sets.

[5] In principle, the eddy flux data, along with associated meteorological drivers (e.g., temperature, solar radiation, humidity) contain enough information that will allow separation of the net flux into its gross components [Goulden *et al.*, 1996a; Braswell *et al.*, 2005], though there is currently no agreed upon approach for doing so, and the underlying uncertainties are not well quantified. The basis for this disaggregation is the fact that nighttime NEE reflects respiration processes only, and to the extent that respiration can be predicted during the day on the basis of relationships with predictor variables at night, daytime GEE can be estimated essentially as the difference between NEE and modeled respiration. Thus GEE estimates rely heavily on model predictions for large contiguous intervals (i.e., all daylight hours). Like any statistical inference, this process carries with it some prediction uncertainty that should be quantified in order to compare tower-based GEE with independent observations or model predictions.

[6] An additional factor that must be considered in utilizing eddy flux data is the existence of missing data resulting from inevitable instrumental lapses. Also, periods of low atmospheric turbulence result in CO<sub>2</sub> flux measurements that are not representative of the actual ecosystem-atmosphere exchange, and these data typically are removed prior to analysis [Goulden *et al.*, 1996b]. Altogether, the resulting gaps can be extensive and nonrandomly distributed in time. The implication for estimating GEE is that an additional model to fill daytime NEE gaps must be defined and parameterized, which adds some amount of quantifiable prediction uncertainty.

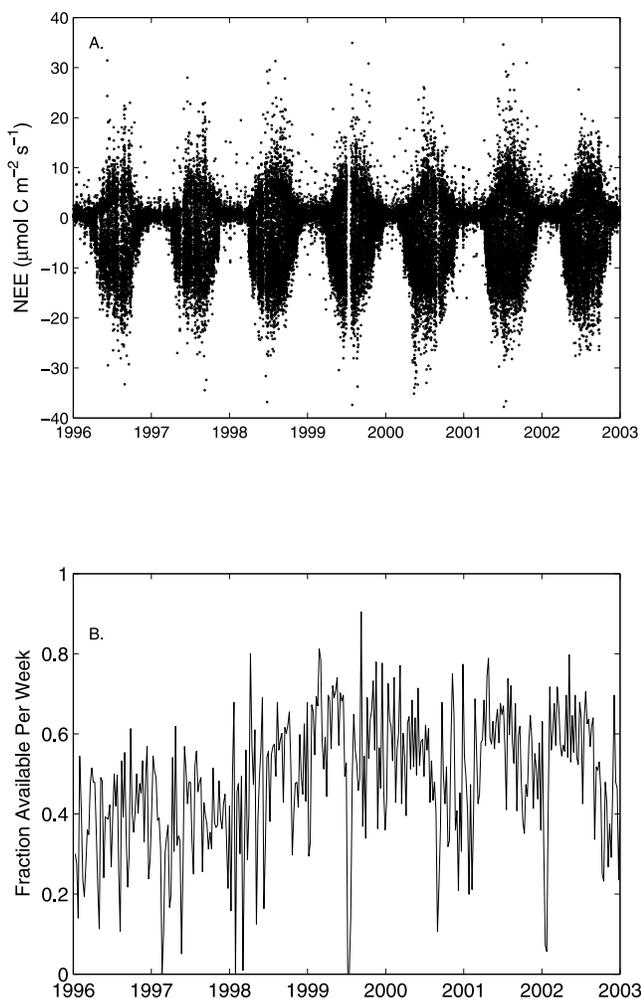
[7] One possible framework for constructing a time series of ecosystem uptake (GEE), given the data and a choice of models, is

$$G = \left\{ \begin{array}{ll} 0 & \text{Night} \\ \hat{R} - F & \text{Day, No Gap} \\ \hat{R} - \hat{F} & \text{Day, Gap} \end{array} \right\}, \quad (1)$$

where  $G$  is GEE,  $F$  is the observed net flux (NEE), and  $\hat{R}$  and  $\hat{F}$  are the modeled respiration and daytime NEE, respectively. Several previous studies have focused separately on issues related to “gap filling” [e.g., Falge *et al.*, 2001], i.e., defining and evaluating the model  $\hat{F}$ , as well as the general problems of disaggregating NEE into component fluxes, which has focused principally on choosing an appropriate regression model for  $\hat{R}$  [e.g., Goulden *et al.*, 1996a]. More recently, however, data assimilation techniques have been used to both fill gaps in flux records and disaggregate NEE into component fluxes [Jarvis *et al.*, 2004; Gove and Hollinger, 2006].

[8] To most appropriately use eddy flux derived GEE for comparison with process models, satellite data, or other field observations, the statistical uncertainties associated with the inference of daytime respiration and NEE during gaps should be quantified so that error bars can be applied at any given choice of timescale. Commonly used statistical approaches for providing error bounds using analytical formulas, such as the formula used to estimate the prediction interval for least squares regression predictions, are not applicable to these data because the underlying assumptions of these approaches do not hold [Hollinger and Richardson, 2005]. For example, eddy flux CO<sub>2</sub> data and the predictions obtained from regressions using these data have (1) non-constant variance, (2) nonindependence of residuals, (3) non-Gaussian noise, and (4) potential sampling bias due to the nonrandom distribution of data gaps. Hollinger and Richardson [2005] conclude that the first three properties listed above result from a combination of the stochastic nature of turbulence, occasional large instrument errors, and the nonuniform occurrences of environmental driving conditions (e.g. over 24 hours, there are far more instances of zero solar radiation than higher values).

[9] Monte Carlo based statistical techniques such as resampling with replacement (“bootstrapping”) [Robert and Casella, 1999] provide a computational solution to the problem of estimating statistical uncertainty in nonlinear model predictions and data with complicating features such as severe heteroscedasticity. Previous studies have utilized ad hoc approaches inspired by bootstrapping to estimate uncertainties of net CO<sub>2</sub> exchange. Often, the technique is used to estimate uncertainty in a sum of flux estimates over time. The most common application includes the random simulation and filling of additional data gaps [Falge *et al.*, 2001; Griffis *et al.*, 2003]. Another Monte-Carlo technique applied to net flux data involves modeling and repeatedly resampling residuals to estimate uncertainty [Saleska *et al.*, 2003]. Uncertainty due to gaps has also been estimated by creating seasonal populations of daily carbon balance that are randomly sampled for comparison with actual fluxes [Goulden *et al.*, 1996b]. Quantification of the measurement uncertainty in flux observations has recently been addressed (this



**Figure 1.** (a) Time series of valid observations of NEE at Howland Forest, Maine, and (b) the fraction available data per week. In this study we used half-hourly data for intervals in which  $u^* > 0.25$ . Overall, the remaining observations amounted to 49% of the total time intervals.

includes defining a suitable probability density function and some measure of the variance) [e.g., *Hollinger and Richardson, 2005*]. Following model parameter optimization using maximum likelihood techniques, random noise with the same statistical characteristics as the measurement uncertainty of the original data can be added back to the model output [*Press et al., 1993*]. By using repeated simulation, as in a Monte Carlo approach, uncertainty limits can be estimated for model parameters, gap-filled values, or annual sums [e.g., *Richardson and Hollinger, 2005*].

[10] In this paper, we present an example of statistical uncertainty estimation and error analysis for a GEE time series, based on eddy flux data from the Howland Forest in Howland, Maine, USA. Our analysis differs from previous work in several ways. First, we are focusing on gross ecosystem exchange, a component flux that reflects a distinct set of ecosystem processes, as opposed to ecosystem respiration or net flux. Second, we account for uncertainty due to model parameterization as well as the

uncertainty associated with the random nature of the flux observations (earlier studies have focused on one or the other). We recognize that uncertainty in ecosystem flux arises from sources other than the statistical modeling, including different choices of friction velocity thresholds for filtering, variability in tower footprint, and changes in the system (i.e., insect infestations, large tree blow downs, etc.). In this analysis, we estimate patterns of uncertainty that are related only to statistical inference. Third, our method does not require the generation of additional gaps and therefore allows us to estimate statistical uncertainty at any timescale, from half hour to multiyear. Last, we perform a sensitivity analysis of the uncertainty of half-hourly to annual GEE estimates using different modeling approaches and different statistical assumptions, in an attempt to understand the effect of model choice on the estimates. We examine and quantify the 90% prediction intervals for one site, but our discussion of the general implications of our results for the role of data and models in understanding ecosystem processes is not site specific.

## 2. Data

[11] Howland Forest is an AmeriFlux research site located at  $45.20^\circ\text{N}$  and  $68.74^\circ\text{W}$ , about 35 miles north of Bangor, ME. The site is dominated by red spruce and eastern hemlock. The vegetation, soils, and climate of this site have been thoroughly described elsewhere [*Hollinger et al., 1999*]. The main eddy flux research tower has been operational since 1995.

[12] We examined 7 years of  $\text{CO}_2$  flux data (NEE) measured half-hourly from 1996 through 2002 (Figure 1a). We screened out flux data with low friction velocity ( $u^* \leq 0.25 \text{ m s}^{-1}$  [*Hollinger et al., 2004*]). The friction velocity screening, primarily, and the occasional instrument failure, secondarily, combine to reduce the amount of available data. There are also other periods when data do not meet quality standards and are rejected. The resulting time series of NEE data contain available observations for 49% of all half-hour intervals (Figure 1b). To compute GEE for each of the 61,362 daytime half hours in 1996–2002, we need to model all 61,362 (100%) respiration values and 24,295 (40%) missing daytime NEE values. The NEE time series is missing 39,382 (64%) nighttime observations. While the nighttime measurements are not used directly in the GEE estimates because we assume no photosynthesis occurs in the dark, the valid nighttime NEE observations are used fit the respiration model.

[13] Half-hourly meteorological data (including air temperature, soil temperature, solar PPFD, and vapor pressure deficit) from the Howland tower were used as driving variables for the GEE modeling.

## 3. Methods

[14] To estimate GEE and the associated uncertainty range given an observed NEE time series, the following components are needed: (1) a statistical regression model, (2) an expression for the likelihood of the data given a model (which implicitly provides a cost function), and (3) a strategy for calculating distributions that represent the probability that a missing flux observation would have

taken a certain value. From these distributions, attributes such as the mean and variance (i.e., uncertainty) of the GEE estimates can be derived for any desired timescale.

[15] Our goal is to present a general analysis framework to bracket GEE estimates, rather than to present a comprehensive exploration of all possible model formulations that could be used in this context. Therefore we chose two previously employed models for respiration and daytime NEE, one physiologically based [Hollinger *et al.*, 2004], and the other a fully empirical, nonlinear regression model [e.g., Papale and Valentini, 2003]. Our priority is to evaluate the magnitude and uncertainties associated with each approach, but not to compare the relative usefulness of the two approaches, primarily because they utilize different amounts of information from independent variables. We also evaluate two assumptions about the underlying error distribution of the modeled flux (i.e., the likelihood of the data given the model). One is a Gaussian error distribution, giving rise to least squares estimates; the other is a two-sided exponential error distribution, giving rise to minimization of absolute differences.

[16] We disaggregated the valid half-hourly CO<sub>2</sub> flux measurements into nighttime (PAR < 5 μmol m<sup>-2</sup> s<sup>-1</sup>) and daytime (PAR ≥ 5 μmol m<sup>-2</sup> s<sup>-1</sup>) periods. To model daytime respiration, we fit both a typical physiological ecosystem respiration model and an artificial neural network to the observed nighttime flux data. These models, which relate ecosystem respiration to observed biophysical variables (e.g., nighttime soil temperature), are then used to estimate daytime ecosystem respiration on the basis of daytime observations of the same variables. To fill gaps in daytime NEE data, we again fit the same two types of models to the observed daytime flux data, on the basis of environmental drivers (e.g., daytime air temperature and PAR), and then used the model to estimate daytime NEE on the basis of the available data. We then estimated GEE using equation (1), and calculated the uncertainty associated with the modeling using a bootstrapping approach, which produces empirical distribution functions for the modeled missing data.

[17] We examined the influence of three factors on GEE estimates, resulting in eight sets of model results, parameters, and posterior distributions. We used two different models (physiological and neural network), assumed two different error models (Gaussian and two-sided exponential), and applied the method to the two flux data sets (respiration and daytime NEE) (equation (1)). In the following sections we discuss the details of these cases, and of the bootstrap algorithm.

### 3.1. Physiologically Based (PB) Model

#### 3.1.1. Respiration Component

[18] For respiration modeling, we used available nighttime respiration data to train a simple physiological model of respiration: a three-parameter exponential function of soil temperature at 5 cm depth,  $T_{soil}$  [Lloyd and Taylor, 1994; Hollinger *et al.*, 2004], with one set of parameters, regardless of season:

$$\hat{R} = Ae^{\frac{-E_0}{T_{soil}-T_0}} \quad (2)$$

where  $A$  is a scaling factor,  $E_0$  is the soil temperature-adjusted activation energy (in degrees Kelvin), and  $T_0$  is a reference soil temperature between 0°C and  $T_{soil}$ . Because  $A$  and  $E_0$  are highly correlated parameters [Richardson and Hollinger, 2005], we fixed the value of  $E_0$  at 113.4 K [Hollinger *et al.*, 2004] and optimized the two remaining independent parameters, using a constrained minimization algorithm.

#### 3.1.2. Daytime Net Ecosystem Exchange Component

[19] The physiological model we used to fill gaps in daytime NEE combines the respiration component above with a rectangular hyperbolic equation that relates photosynthesis to PAR, regulated by an optimum air temperature. This Michaelis-Menten type functional relationship requires fitting three additional parameters, for a total of five independent parameters:

$$\hat{F} = \begin{cases} Ae^{\frac{-E_0}{T_{soil}-T_0}} - \frac{P_m I_{PAR}}{I_{PAR} + K_m} \left( \frac{T_{air}^2}{a^2} - \frac{2T_{air}}{a} \right) & T_{air} > 0 \text{ and } T_{soil} > 0 \\ Ae^{\frac{-E_0}{T_{soil}-T_0}} & T_{air} \leq 0 \text{ or } T_{soil} \leq 0 \end{cases} \quad (3)$$

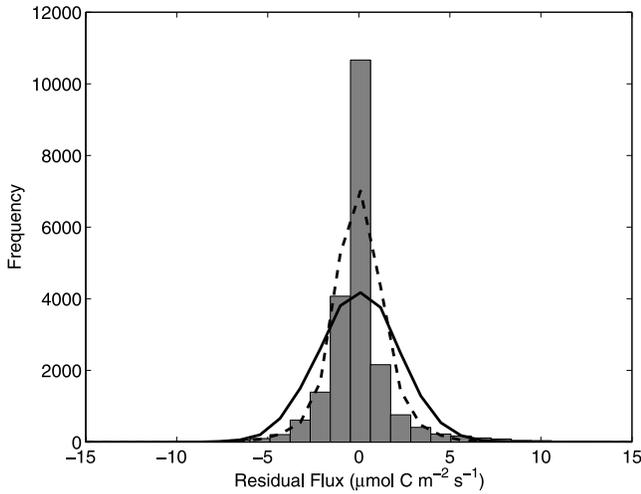
where  $I_{PAR}$  is the incident horizontal photosynthetically active radiation and  $T_{air}$  is the air temperature. The parameters are  $P_m$ , the maximum rate of photosynthesis,  $a$ , the normalized parabolic air temperature response with an intercept of zero, and  $K_m$ , the photosynthetic half-saturation constant. We used the previously optimized nighttime values for  $A$  and  $T_0$  (section 3.1.1). When  $T_{air}$  or  $T_{soil}$  is less than 0°C we assume that  $GEE = 0$  and  $\hat{F} = \hat{R}$ .

[20] The PB model was chosen for its simple representation of the system (i.e., five parameters) and its relatively wide use in the forest ecosystem community. For additional simplicity, the parameters are assumed constant across the years. Other analyses with Howland data suggest that fitted parameters of similar models change seasonally and between years [e.g., Hollinger *et al.*, 2004; Gove and Hollinger, 2006].

### 3.2. Artificial Neural Network (ANN) Model

[21] The physiological models used here represent a family of functions whose characteristic shapes are constrained by prior knowledge of, or assumptions about, the relationships between a set of independent variables and the response (e.g., the soil temperature control of respiration). In contrast, the ANN approach focuses solely on characterizing the relationship between the valid NEE measurements and the climate measurements, making no assumptions about physiological processes, so the functional dependence of daytime NEE and respiration on biophysical predictor variables is not prescribed. Other studies have used this modeling approach for the purpose of gap filling flux data [e.g., Papale and Valentini, 2003].

[22] We apply essentially the same ANN architecture separately to valid nighttime NEE data for modeling ecosystem respiration, and to valid daytime NEE data to model NEE where it is unavailable. The respiration model is driven by soil temperature, air temperature, surface soil moisture, and a seasonal indicator in the form of sine and cosine functions of the day of the year. The daytime NEE model adds photosynthetically active radiation (PAR), vapor pres-



**Figure 2.** Residuals of a model fit to nighttime NEE, or ecosystem respiration (shaded bars), distributed with a kurtotic peak around zero. This distribution resembles a two-sided exponential distribution (dashed line) more than a normal distribution (solid line).

sure deficit (VPD), and sine and cosine functions of the hour of the day as additional input drivers.

[23] An artificial neural network model is a multistage nonlinear regression function where the intermediate values are called hidden nodes. For example, with two stages  $y = f(g(x))$ , or more specifically:

$$y_k = f\left(\sum_{j=0}^M w_{kj}^{(2)} g\left(\sum_{i=0}^D w_{ji}^{(1)} x_i\right)\right), \quad (4)$$

where  $x$  represents the collection of independent variables in the regression (in our case the biophysical drivers). The outer function  $f(\cdot)$  is usually linear and the inner function  $g(\cdot)$  is a nonlinear, typically sigmoidal function, such as the hyperbolic tangent. The free parameters in this regression are weights  $w_{ji}$  and  $w_{kj}$ , which represent the strength of the connection between the  $i$ th input and the  $j$ th intermediate value (represented by the  $M$  evaluations of  $g$ ) and also between the  $j$ th intermediate value and the  $k$ th output value  $y$  (in our case NEE or respiration). This ANN has  $D$  inputs and  $M$  hidden nodes.

[24] This regression approach is referred to as a network because all inputs can influence all outputs, depending upon the values of the weights. For estimating the parameters, we use the standard backpropagation algorithm [Bishop, 1995], which updates the weights for each pair of  $\{y_k, x_i\}$  data vectors in order to minimize the error. We also incorporate a Bayesian modification of artificial neural networks [MacKay, 1994] that limits the complexity of the model to that which is supported by the data, avoiding the common neural network problem of overfitting. In our study, we independently verified that the models do not overfit (as part of the K-fold validation exercise below) and, therefore, that the results are not dependent on the choice of the number of hidden nodes  $M$ .

### 3.3. Error Distribution

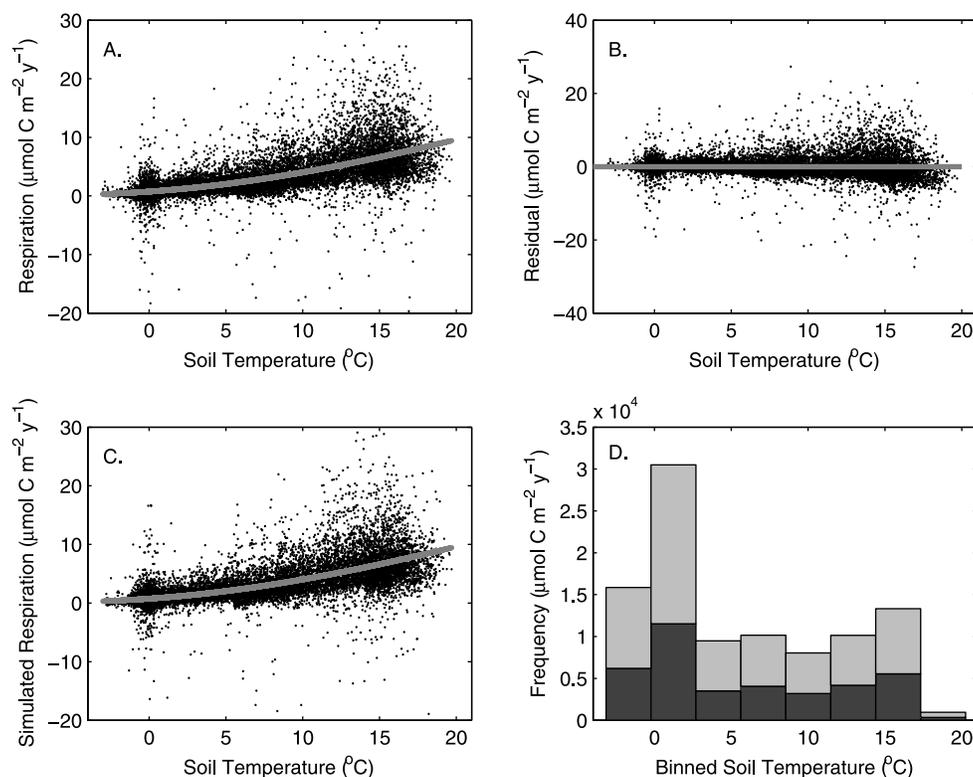
[25] There is evidence that errors associated with eddy flux observation are better represented by a two-sided exponential distribution than a Gaussian distribution, i.e., they are leptokurtic with outliers (Figure 2) [Hollinger and Richardson, 2005; Richardson and Hollinger, 2005]. We performed a multipart analysis with the two types of regression models, considering in each case both an underlying Gaussian and an underlying two-sided exponential distribution. We evaluated the assumptions of underlying error distribution by posterior analysis of the model residuals.

[26] We alter our assumption of how the error is distributed by specifying the form of the cost function that is minimized in the optimization routine. When assuming a Gaussian error distribution, we minimized the usual least squares error function. In the case of the two-sided exponential distribution assumption, we minimized the weighted absolute value of the residuals. We used weights based on the recommendation of Richardson and Hollinger [2005] that the intrinsic observational uncertainty is well represented by an exponential function of soil temperature. More specifically, as can be seen in the data, the uncertainty in flux observations scales with the magnitude of the flux (i.e., absolute error is larger when the absolute flux is larger), and to obtain an independent estimate of that uncertainty, we express the uncertainty as a function of soil temperature.

### 3.4. Uncertainty Analysis

[27] Ecosystem carbon flux is an aggregate property of a system containing many physical, chemical, and biological interactions. For example, nighttime NEE generally increases exponentially with increasing soil temperature, and a simple physiological model captures the basic relationship (Figure 3a). However, substantial noise (i.e., model residuals) remains after this simple relationship has been accounted for (Figure 3b). This residual noise is due to both measurement uncertainty and model uncertainty (i.e., noisy data and an imperfect model), with model uncertainty potentially due to both parameterization and choice of functional form. In addition, the variance of these residuals can be heteroscedastic (i.e., not constant with respect to one or more of the independent variables); in this case, the residual variance varies with soil temperature (Figure 3b).

[28] Many approaches to uncertainty estimation (e.g., least squares regression) assume that the data have constant variance and Gaussian noise, and that the regression model has independent identically distributed residuals. Eddy flux observations and associated models generally do not conform to these assumptions, but computational solutions exist. The bootstrapping approach (resampling with replacement) to uncertainty assessment is one of several techniques more appropriate than conventional analytic methods for data with heteroscedastic and nonnormally distributed errors. This method assumes that the observed data represent only one possible realization out of many, and reconstructs a large number of alternate realizations based on random resampling of residuals. Bootstrapping brackets the range of unobserved values conditioned on the assumption of the model and its associated likelihood function [Efron and Tibshirani, 1993].



**Figure 3.** (a) Nighttime flux (respiration) fit using an Arrhenius function [Lloyd and Taylor, 1994] of soil temperature (shaded line; see equation (2) in text). (b) Variance of the residuals from this model, which is heteroscedastic, with variance increasing at higher temperatures. (c) One example of the 1000 artificial data sets, constructed by randomly adding residuals (bootstrapping) to the simple fitted function in Figure 3a. (d) A histogram of soil temperature for the entire time period. Each bar is split into fraction of half-hours having a valid respiration observation (dark shading) and fraction needing modeled respiration (light shading). Bin locations and sizes from this histogram (Figure 3d) were used to construct the artificial data sets (Figure 3c).

[29] Previous studies have used Monte Carlo analyses for estimating modeling uncertainty in NEE and GEE, but most provide a measure of variability centered on the mean response of a model prediction at a point in time, and do not consider the additional uncertainty due to the random deviations from the mean response of any individual eddy flux observation [e.g., Griffis *et al.*, 2003; Richardson and Hollinger, 2005]. Other studies have accounted for the random processes associated with NEE, but have not considered the uncertainty in the mean response [e.g., Saleska *et al.*, 2003]. Uncertainty about the mean response is known as the confidence interval, while this same uncertainty plus the additional uncertainty due to inherent variations in the data is called the prediction interval. In this study, we present uncertainty as a 90% prediction interval, which brackets uncertainty about an estimate based on new data (i.e., gap filling), which is an appropriate statistical measure of our knowledge (or lack of knowledge) about predicted values. Below, we outline our implementation of the nonparametric resampling approach (bootstrapping), which is based on the statistical theory of Efron and Tibshirani [1993] and recent algorithms described by others [e.g., Robert and Casella, 1999].

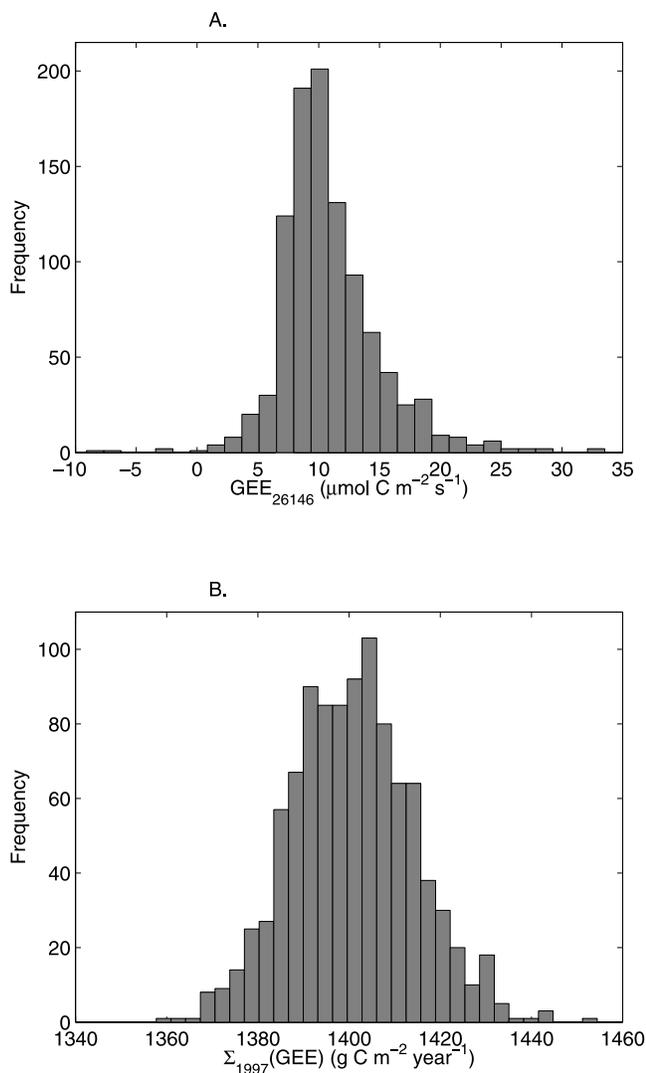
[30] The bootstrap is a simulation based calculation of the properties of an arbitrary estimator, typically the bias or the

standard error, and can also be used to calculate confidence and prediction intervals. Since in the bootstrap algorithm the data are resampled, there is no underlying assumption about the statistical distribution. In regression models, where the statistical assumptions pertain to the model errors, the residuals are resampled and added back to the fitted values to create bootstrap replicates of the data. The regression model is then refit to each replicate, and the resulting empirical distribution of the recalculated estimators provides the desired properties. In our case we evaluate the statistical properties (90% prediction intervals) of the estimators for the response where the original data were missing. This procedure makes no assumptions of the statistical distribution of the residuals. To account for heteroscedasticity as a function of a covariate variable we propose a simple residual binning (step 3 below). The text below outlines the bootstrap algorithm:

[31] In step 1, the regression model (either PB or ANN) is fit to the valid observations (e.g., Figure 3a).

[32] In step 2, the residuals from this fit are calculated (e.g., Figure 3b), and the variance of the residuals is examined for a significant dependence on the driving variables (e.g., soil temperature).

[33] In step 3, if there is significant heteroscedasticity, the main driver of the nonconstant variance is identified. The



**Figure 4.** (a) Bootstrapping algorithm produces empirical probability distributions for each daytime half hour. Most half-hourly distributions of simulated GEE are leptokurtic and skewed, like the example displayed here (1630–1700 LT on 28 June 1997). (b) Aggregating (by summing) the half-hour GEE simulations to the annual scale, for each bootstrapped data set produces an annual empirical distribution. These predictions are generally approximately normally distributed.

range of this driving variable is divided into several intervals, and the residuals are binned on the basis of the value of the driving variable at the time of measurement (e.g., Figure 3d). In the analysis, both daytime NEE and respiration residuals, for both the PB and ANN models, were divided into eight bins on the basis of soil temperature.

[34] In step 4, an artificial data set (e.g., Figure 3c) is created by adding the “model fit” predicted values (the line in Figure 3a) to random residuals drawn with replacement from the correct bin (Figure 3b).

[35] In step 5, a revised PB or ANN model is fit to the bootstrapped data set (e.g., Figure 3c).

[36] In step 6, this bootstrap model is used to predict flux values for the gap points (e.g., Figure 3d).

[37] In step 7, a residual (from step 2) is added to the predicted value (from step 6) in the same manner as described in step 4, to simulate the effect of random noise on any predicted or gap filled point. This step ensures that we capture the statistical prediction error, not just the uncertainty due to model parameterization.

[38] In step 8, repeat steps 4–7 above  $N$  times (we used  $N = 1000$ ).

[39] In step 9, predicted values and prediction intervals are calculated using the empirical distributions of the results (e.g., Figure 4a). Every gap point in the time series will have  $N$  estimated values from  $N$  realizations of the resampled and refit time series. Calculation of the quantiles of these values yields many metrics, including the median and 90% prediction limits.

[40] In step 10,  $N$  complete component flux time series are generated by using the measured value at every point in the time series where there is an observation and by using a bootstrap-predicted value for those time steps with no measurement. Expected values and prediction limits for sums of fluxes are estimated from these  $N$  synthetic time series (Figure 4b).

### 3.5. Validation

[41] We used two measures of performance to evaluate both the PB and the ANN models both for filling unavailable daytime flux and for estimating daytime respiration. First, we conducted a standard  $K$ -fold cross validation of the nighttime respiration models and daytime NEE models [Hastie *et al.*, 2001], which allowed us to quantify out-of-sample model error. We split all the valid data into  $K$  randomly distributed groups. Initially, group 1 is set aside for testing, while the models are parameterized on the basis of groups 2 through  $K$ . The fitted models are then used to predict the group 1 observations. Next, group 2 is set aside for testing, while groups 1 and 3 through  $K$  are used for training. This pattern proceeds until all  $K$  groups have been withheld for testing. We then computed the root mean squared error (RMSE), the weighted absolute value of the error (WAD), the correlation coefficient ( $R^2$ ), and mean bias as measures of model performance.

[42] A second evaluation of model performance allows us to investigate the accumulation of uncertainty as model predictions are aggregated (by summing) into longer temporal intervals. There are few long periods without missing observations (Figure 1b), but we identified in the 7-year Howland NEE time series 13 days having zero gaps and 73 days having only one gap. We compared the observed 48 half-hour total NEE to the 1000 model predicted NEE values for these 86 complete and near-complete days. While the models used in this analysis were generated without the data from the 86 days of interest, the uncertainty estimates were taken from the bootstrapping analysis described in section 3.5.

## 4. Results and Discussion

### 4.1. Half-Hourly Time Step

#### 4.1.1. Parameter Optimization

[43] The physiological parameter values that minimize the cost functions applied to the observed data are similar to the parameter values fit by Hollinger *et al.* [2004] in their

**Table 1.** Optimal Parameter Values for the Physiological Models<sup>a</sup>

	Gaussian	Exponential
<i>Respiration</i>		
A	149.1	149.9
E <sub>0</sub>	113.4	113.4
T <sub>0</sub>	251.8	252.8
<i>DayNEE</i>		
P <sub>m</sub>	22.3	18.8
K <sub>m</sub>	344.8	300.1
a	22.4	24.4

<sup>a</sup>E<sub>0</sub> is fixed in this exercise.

analysis (Table 1), though they used a different subset of the data (1996 only). The artificial neural network used four hidden nodes ( $M = 4$ ; equation (4)) for both the respiration model and the daytime NEE model. The optimized neural network parameters (i.e., weights) are not physiologically meaningful and therefore their values cannot be compared with other studies.

[44] The residuals generated from the respiration model fit resembled a two-sided exponential error distribution more than a Gaussian distribution (Figure 2), which is in agreement with the observation that flux measurement uncertainty follows a Laplace rather than a Gaussian distribution [Hollinger and Richardson, 2005]. This was also true for residuals from other models' fits (not shown). By changing the assumption of how the error is distributed, one changes the optimal parameters. There are many combinations of parameter values that fit the data nearly equally well. The flatness of the cost function near the optimum has been described thoroughly elsewhere [Radtke et al., 2002; Hollinger et al., 2004; Hollinger and Richardson, 2005].

#### 4.1.2. Model Validation

[45] The K-fold cross validation results show that both modeling approaches (ANN and PB) reproduce observed daytime NEE and nighttime respiration reasonably well at the half-hourly timescale, with all coefficient of determination values ( $R^2$ ) greater than or equal to 0.49 (Table 2). For respiration, the ANN and PB models fit the data approximately equally well, probably because both models are based primarily on soil temperature (with the addition of the time variables in the ANN approach). However, there is a larger discrepancy between the ANN and PB model fits to the daytime NEE observations.

[46] The ANN modeling approach has a lower mean error than the PB approach in every case, expressed either as root mean squared error (RMSE) or weighted absolute deviation (WAD). This is expected because ANN provides more flexible choices for the functional dependence than the physiological model and a larger set of input variables. The daytime NEE models are less accurate (i.e., they have higher RMSE or WAD) than the respiration models, likely because daytime NEE observations have higher variance than nighttime respiration observations.

[47] Changing the assumption of error distribution has a small effect on the cross-validation results of the respiration model, increasing the error (e.g.,  $\text{RMSE}_{\text{gauss}} - \text{RMSE}_{\text{exp}}$  or  $\text{WAD}_{\text{exp}} - \text{WAD}_{\text{gauss}}$ ) by at most 5%. This change has a slightly larger effect on the daytime NEE models, increasing the error by up to 10%. The magnitude of change in this K-

fold error statistic is an indication of the model's sensitivity to assumptions about the error distribution and the daytime NEE models are more sensitive to this assumption.

[48] The cross-validation results indicate that all models assuming a Gaussian error distribution have no statistically significant model bias (Table 2). The models using weighted observations and a two-sided exponential distribution in the cost function, however, all show a significant bias. This bias is an expected by-product of the model assumptions, particularly the weighting of observations. The weighting scheme assumes that the observations taken during high soil temperatures are less reliable and, therefore, the influence of residuals taken at high soil temperatures is reduced. These assumptions reflect a belief about how best to accommodate heteroscedastic data and occasional large instrumentation errors [Richardson and Hollinger, 2005].

#### 4.1.3. GEE Estimates

[49] Each modeling approach (PB/ANN and Gaussian/Exponential) produces one time series of daytime NEE and a second of daytime respiration, both at half-hour intervals. The daytime NEE time series contains observed fluxes where data are available, and modeled fluxes where they are not. The daytime respiration time series has only modeled fluxes. By applying the bootstrapping algorithm, we generate one thousand time series, each representing a simulated potential time series that includes uncertainty in the model parameters as well as uncertainty due to the random nature of the flux observation. One thousand GEE time series are estimated by subtracting the 1000 daytime NEE time series from the 1000 respiration time series (equation (1)). Thus each daytime half hour has 1000 simulated GEE estimates that approximate the distribution of values that could have been observed given the data and the modeling assumptions. The simulated GEE estimates for any half hour can be displayed as a histogram (Figure 4a). From this histogram, we can extract several statistics of interest, including the mean, median, upper 90% value, and lower 90% value.

[50] At the half-hour timescale, the GEE estimates generated from the bootstrapping algorithm are often skewed (Figure 4a). This skewness reflects a skewness in the model residuals and, ultimately, in the flux observations themselves. The nighttime flux (i.e., respiration) record contains

**Table 2.** K-Fold Validation Results for All of the Model Filling Approaches

	Gaussian Error		
	R <sup>2</sup>	Bias, $\mu\text{mols m}^{-2} \text{s}^{-1}$	RMSE
<i>Respiration</i>			
Artificial neural net	0.53 ± 0.01	0.00 ± 0.02	2.21 ± 0.04
Physiological	0.51 ± 0.01	-0.03 ± 0.02	2.28 ± 0.03
<i>DayNEE</i>			
Artificial neural net	0.75 ± 0.01	-0.00 ± 0.01	3.12 ± 0.04
Physiological	0.50 ± 0.01	-0.01 ± 0.03	4.56 ± 0.04
	Two-Sided Exponential Error		
	R <sup>2</sup>	Bias, $\mu\text{mols m}^{-2} \text{s}^{-1}$	WAD
<i>Respiration</i>			
Artificial neural net	0.52 ± 0.01	0.27 ± 0.02	0.67 ± 0.01
Physiological	0.50 ± 0.01	0.28 ± 0.01	0.71 ± 0.01
<i>DayNEE</i>			
Artificial neural net	0.70 ± 0.01	0.16 ± 0.03	1.31 ± 0.02
Physiological	0.49 ± 0.00	-1.04 ± 0.03	2.86 ± 0.02

**Table 3.** Complete Day Validation Results for the NEE Gap Filling Approaches, Based on 86 Days With Fewer Than Two Missing Half-Hour Intervals

	$R^2$	RMSE, $\text{g C m}^{-2} \text{ day}^{-1}$	Daily Mean Bias, $\text{g C m}^{-2} \text{ day}^{-1}$
Gaussian error			
Artificial Neural Net	0.75	0.74	-0.23
Physiological	0.53	1.15	-0.50
Two-sided exponential error			
Artificial Neural Net	0.72	0.74	-0.05
Physiological	0.48	1.16	-0.52

more unusually high flux measurements (i.e., positive; flux out of the canopy and into the atmosphere) than unusually low (i.e., negative) flux measurements, while the daytime flux record is skewed in the opposite direction. To estimate GEE, we subtract daytime NEE flux from respiration, which magnifies the skewness in the GEE estimates.

[51] At the half-hour scale, the GEE estimates generated from the four approaches are never significantly different at the 90% prediction limit level. While the median bootstrapped estimates predicted from any approach at any half hour are different, the statistical uncertainty reflected by the 90% prediction limits is large relative to this difference. The ANN models generally predict slightly higher GEE during half hours with high  $I_{PAR}$  than the PB models. During low  $I_{PAR}$  levels, the PB models predict higher GEE than the ANN models.

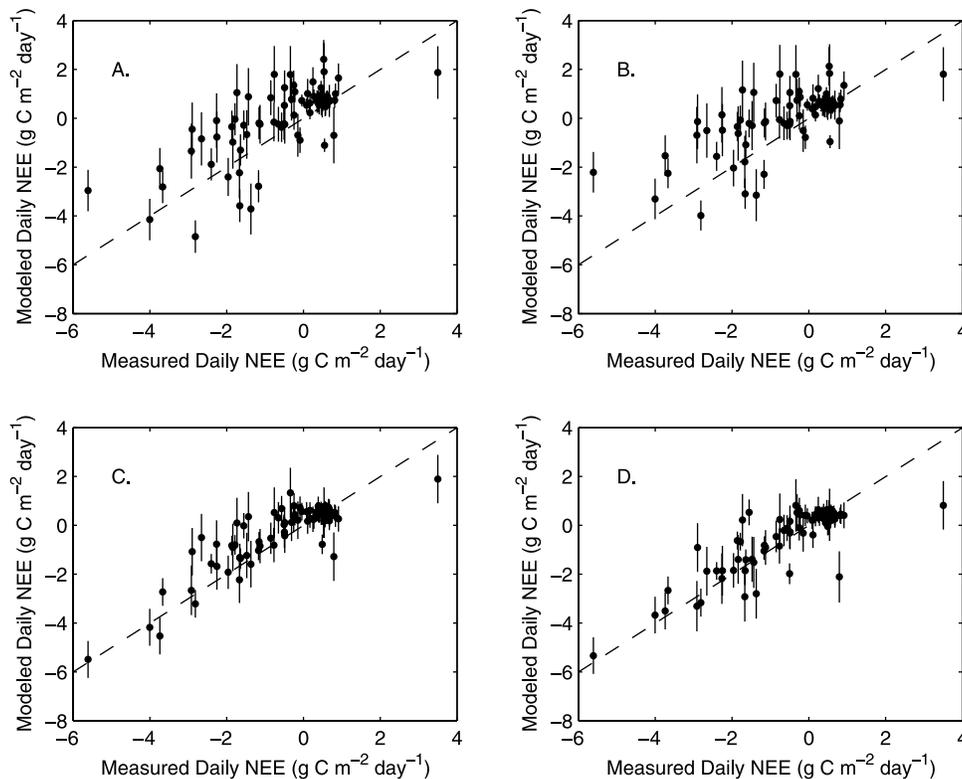
#### 4.2. Daily Time Step: Validation of Complete-Day NEE

[52] Both model approaches validate reasonably well using the 86 complete-day data points (all  $R^2 > 0.48$ ),

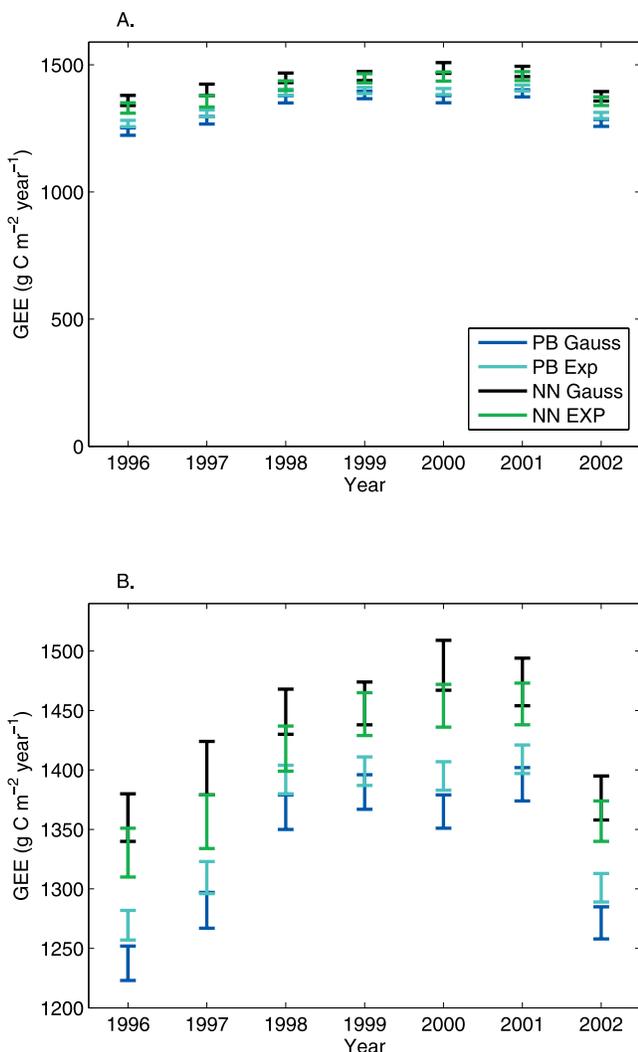
though the ANN has a higher correlation and a lower RMSE and mean bias (Table 3 and Figure 5). In the context of this analysis, changing the assumption of normally distributed residuals to an assumption of two-sided exponentially distributed residuals does not improve the accuracy of the predictions. The 90% prediction limits around each daily prediction in this small sample are apparently underestimates of the actual uncertainty, as only about 70% of the prediction limits touch the 1:1 line.

#### 4.3. Annual Time Step: GEE Estimates and 90% Prediction Limits

[53] Annual GEE estimates for each modeling approach are generated by aggregating each of the 1000 individual GEE time series to the annual scale. At this scale, annual GEE estimates are approximately normally distributed (Figure 4b). They are no longer significantly skewed or kurtotic, so that the mean estimates and the median estimates are effectively equal.



**Figure 5.** Modeled versus measured daily NEE for 86 complete or nearly complete days in the Howland Forest time series, using four model and error distribution combinations: (a) PB Gaussian, (b) PB exponential, (c) ANN Gaussian, and (d) ANN exponential. Error bars represent 90% bootstrap intervals.



**Figure 6.** Time series of annual GEE. (a and b) Same data at different scales. In Figure 6a, the difference in estimates of annual total GEE from the four modeling approaches is small relative to the magnitude of GEE, as is the statistical uncertainty. In Figure 6b, the annual GEE estimates do exhibit dependence on the method chosen for gap filling daytime NEE and respiration modeling. The statistical uncertainty due to model fitting and the random variability of the observations is comparable to the uncertainty due to model selection. Interannual variability in GEE is partially masked by statistical uncertainty and nearly completely masked by model selection uncertainty, but the overall patterns are almost identical (i.e., the rank correlation is very high).

[54] The annual GEE sums estimated in this analysis (Figure 6) are generally consistent with previous estimates for the Howland site using the same data [Hollinger *et al.*, 2004], and with those based on mechanistic model predictions (e.g., PnET model [Aber *et al.*, 1996]). This similarity includes the overall absolute values of the magnitude of the flux as well as the rank order of annual values. However, focusing especially on interannual patterns, there is a consistent offset between the modeling approaches.

[55] The bootstrapped estimates of the annual 90% prediction intervals average  $40 \text{ g C m}^{-2} \text{ year}^{-1}$  for the ANN approach and  $30 \text{ g C m}^{-2} \text{ year}^{-1}$  for the PB approach. The year-to-year variability in GEE is smaller than the magnitude of uncertainty at the annual timescales in at least three of the six pairs of adjacent years (i.e., three of six pairs in the PB and four of six pairs in the ANN). Changing the cost function to reflect the assumption of exponentially distributed error slightly reduces our estimates of statistical uncertainty (Figure 6). All methods agree in predicting higher GEE at Howland over the 1998–2001 period than before or after this time.

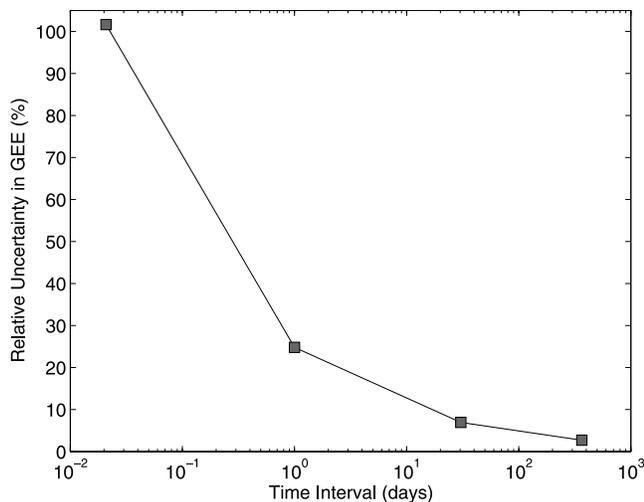
#### 4.4. Statistical Uncertainty in GEE Estimates Across Time

[56] The 90% annual prediction intervals from the different methods are generally offset from one another and in many cases do not overlap. This may be due to the fact that our analysis accounts only for uncertainties associated with statistical modeling, and is consistent with the likely influence of other external factors. The offset of prediction intervals within a year also shows that uncertainty related to model selection contributes considerably to the overall range of possible GEE estimates. This overall range is difficult to quantify comprehensively because the total number of models that can be used is not finite. However, the two models used here represent two extremes, both in terms of the number of variables and the way in which the variables are used.

[57] Though the nonlinear regressions are quantitatively more accurate than the physiologically based regression, there is no objective basis for choosing one approach over the other. A process-oriented model (e.g., equations (2) and (3)) may contain useful prior functional constraints about ecosystem carbon fluxes. Alternatively, a regression model that synthesizes the data record most accurately (e.g., equation (4)) may be the best choice if we desire estimates that mimic the behavior of the data rather than provide insights about the processes or capacity for extrapolation.

## 5. Conclusions

[58] Tower-based estimates of GEE represent a potentially important source of ecosystem information that is derived by a combination of data and models. As such, they require more analytical processing than most data sets that are considered “observations,” but they also are likely to be used as data to a greater extent than most quantities that are considered “model outputs.” The objective of this analysis was to provide a framework for estimation of uncertainty in tower-based GEE time series. Specifically, we are interested in quantifying the prediction intervals associated with regression models that are needed to (1) extrapolate respiration into the day and (2) fill missing NEE values in the day. These prediction intervals correspond to the range of values we would likely observe, given the valid data and the model assumptions. We have used a computational technique that is intended to bracket the range of likely observations, but it is not guaranteed to bracket the unknown “true” values of GEE flux. We did not explore a large number of different regression models, but instead illustrated the issue by using two different modeling



**Figure 7.** Relative uncertainty, expressed as the magnitude of the mean 90% prediction interval divided by the mean prediction value, of the ANN-modeled Gaussian GEE as a function of time step on a log scale. This relative uncertainty, as with the other approaches (not shown), drops dramatically as GEE is aggregated over time (half-hourly, daily, monthly, and annually). This result is attributable to the fact that the statistical uncertainty adds approximately in quadrature and reflects the law of large numbers in estimating mean quantities (i.e., standard errors shrink with increasing sample size).

approaches. Valid arguments could be made for the use of either approach, and we do not recommend one over the other.

[59] The statistical uncertainty in annual GEE estimates at Howland Forest associated with each model type, is about 30–40 g C m<sup>-2</sup> year<sup>-1</sup> (90% prediction limit). Our results indicate that the uncertainty due to model assumptions is greater than the statistical uncertainty associated with any particular model. The combined uncertainty due to modeling in the GEE estimates is nearly the same magnitude as the interannual variability. These estimates are similar in magnitude to the uncertainty in NEE arising from systematic errors associated with choice of nocturnal  $u^*$  threshold [Hollinger *et al.*, 2004].

[60] While our analysis indicates a relatively small amount of uncertainty in the absolute value of GEE at the annual scale, this relative uncertainty is much larger at shorter timescales and is a dominant feature when considering half-hourly to daily fluxes (Figure 7). Furthermore, the interannual variability of the GEE flux, which is a key focus point for research into process controls linking environment and ecosystems, is often masked by the uncertainty from one year to the next. The implications of this result are potentially significant, and should be investigated independently at other sites and with other methods. On the other hand, the consistent patterns of the variability between model types indicate that some insight can still be gained about larger trends without considering explicitly the absolute magnitude of the GEE flux (Figure 6b).

[61] The complexity of this data set and the nature of the GEE calculation make error estimation sensitive to statisti-

cal assumptions. The impact of our choice of underlying error distribution assumption was significant, but less so than the differences associated with the selection of a model. While future work is needed to further integrate sources of uncertainty, evaluate alternate modeling techniques, and generalize results across multiple sites, this paper represents an initial step in the characterization of uncertainty in gross ecosystem fluxes from the bottom up (e.g., in situ observations) and is useful in conjunction with top-down estimates (e.g., satellite observations, model inversions).

[62] **Acknowledgments.** This work was supported by the NASA Terrestrial Ecology Program (contract NAG5-12876) and the Office of Science (BER), U.S. Department of Energy, through the Northeast Regional Center of the National Institute for Global Environmental Change under contract 123246-06. The Howland flux research was supported by the USDA Forest Service Northern Global Change Program, the Office of Science (BER), U.S. Department of Energy, through the Northeast Regional Center of the National Institute for Global Environmental Change under Cooperative Agreement DE-FC03-90ER61010 and by the Office of Science (BER), U.S. Department of Energy, Interagency Agreement DE-AI02-00ER63028. We are grateful to Julian Jenkins, Scott Ollinger, Dave Schimel, and Bill Sacks for helpful discussions before and during this analysis.

## References

- Aber, J. D., P. B. Reich, and M. L. Goulden (1996), Extrapolating leaf CO<sub>2</sub> exchange to the canopy: A generalized model of forest photosynthesis validated by eddy correlation, *Oecologia*, *106*, 257–265.
- Baldocchi, D. D., and C. A. Vogel (1996), Energy and carbon dioxide flux densities above and below a temperate broad-leaved forest and a boreal pine forest, *Tree Physiol.*, *16*, 5–16.
- Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Oxford Univ. Press, New York.
- Braswell, B. H., B. Sacks, E. Linder, and D. S. Schimel (2005), Estimating ecosystem process parameters by assimilation of eddy flux observations of NEE, *Global Change Biol.*, *11*, 335–355.
- Efron, B., and R. J. Tibshirani (1993), *An Introduction to the Bootstrap*, CRC Press, Boca Raton, Fla.
- Falge, E., *et al.* (2001), Gap filling strategies for defensible annual sums of net ecosystem exchange, *Agric. For. Meteorol.*, *107*, 43–69.
- Goulden, M. L., J. W. Munger, S. Fan, B. C. Daube, and S. C. Wofsy (1996a), Exchange of carbon dioxide by a deciduous forest: Response to interannual climate variability, *Science*, *271*, 1576–1578.
- Goulden, M. L., J. W. Munger, S. Fan, B. C. Daube, and S. C. Wofsy (1996b), Measurements of carbon sequestration by long-term eddy covariance: Methods and a critical evaluation of accuracy, *Global Change Biol.*, *2*, 169–182.
- Gove, J. H., and D. Y. Hollinger (2006), Application of a dual unscented Kalman filter for simultaneous state and parameter estimation in problems of surface-atmosphere exchange, *J. Geophys. Res.*, doi:10.1029/2005JD006021, in press.
- Griffis, T. J., T. A. Black, K. Morgenstern, A. G. Barr, Z. Nescic, G. B. Drewitt, D. Gaumont-Guay, and J. H. McCaughey (2003), Ecophysiological controls on the carbon balances of three southern boreal forests, *Agric. For. Meteorol.*, *117*, 53–71.
- Hastie, T., R. Tibshirani, and J. Friedman (2001), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, New York.
- Hollinger, D. Y., and A. D. Richardson (2005), Uncertainty in eddy covariance measurements and its application to physiological models, *Tree Physiol.*, *25*, 873–885.
- Hollinger, D. Y., S. M. Goltz, E. A. Davidson, J. T. Lee, K. Tu, and H. T. Valentine (1999), Seasonal patterns and environmental control of carbon dioxide and water vapour exchange in an ecotonal boreal forest, *Global Change Biol.*, *5*, 891–902.
- Hollinger, D. Y., *et al.* (2004), Spatial and temporal variability in forest-atmosphere CO<sub>2</sub> exchange, *Global Change Biol.*, *10*, 1689–1706.
- Jarvis, A. J., V. J. Stauch, K. Schulz, and P. C. Young (2004), The seasonal temperature dependency of photosynthesis and respiration in two deciduous forests, *Global Change Biol.*, *10*, 939–950.
- Lloyd, J., and J. A. Taylor (1994), On the temperature dependence of soil respiration, *Functional Ecol.*, *8*, 315–323.
- MacKay, D. (1994), Bayesian methods for backpropagation network, in *Models of Neural Networks*, vol. III, edited by E. Domany, J. van

- Hemmen, and K. Schulten, chap. 6, pp. 211–254, Springer, New York.
- Papale, D., and R. Valentini (2003), A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization, *Global Change Biol.*, *9*, 525–535.
- Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (1993), *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, 992 pp., Cambridge Univ. Press, New York.
- Prince, S. D., and S. N. Goward (1995), Global primary production: A remote sensing approach, *J. Biogeogr.*, *22*, 815–835.
- Radtke, P., T. E. Burk, and P. V. Bolstad (2002), Bayesian melding of a forest ecosystem model with correlated inputs, *For. Sci.*, *48*, 701–711.
- Richardson, A. D., and D. Y. Hollinger (2005), Statistical modeling of ecosystem respiration using eddy covariance data: Maximum likelihood parameter estimation and Monte Carlo simulation of model and parameter uncertainty applied to three different models, *Agric. For. Meteorol.*, *131*, 191–208.
- Robert, C. P., and G. Casella (1999), *Monte Carlo Statistical Methods*, Springer, New York.
- Saleska, S. R., et al. (2003), Carbon in Amazon forests: Unexpected seasonal fluxes and disturbance-induced losses, *Science*, *302*, 1554–1557.
- Wofsy, S. C., M. L. Goulden, J. W. Munger, S. M. Fan, P. S. Bakwin, B. C. Daube, S. L. Bassow, and F. A. Bazzaz (1993), Net exchange of CO<sub>2</sub> in a midlatitude forest, *Science*, *260*, 1314–1317.
- Xiao, X., D. Y. Hollinger, J. Aber, M. Goltz, E. A. Davidson, Q. Zhang, and B. Moore (2004), Satellite-based modeling of gross primary production in an evergreen needleleaf forest, *Remote Sens. Environ.*, *89*, 519–534.
- 
- B. H. Braswell, S. Frolking, S. C. Hagen, and A. D. Richardson, Complex Systems Research Center, Morse Hall, University of New Hampshire, Durham, NH 03824, USA. (steve.hagen@unh.edu)
- D. Y. Hollinger, U.S. Department of Agriculture Forest Service NE Research Station, 271 Mast Road, Durham, NH 03824, USA.
- E. Linder, Department of Mathematics and Statistics, University of New Hampshire, Durham, NH 03824, USA.