**Preserving the variance in imputed eddy-covariance measurements: alternative
methods for defensible gap filling**

Sujit Kunwor¹, Gregory Starr¹, Henry W. Loescher^{2, 3}, Christina L. Staudhammer^{1, §}

¹University of Alabama, Department of Biological Sciences, Tuscaloosa, AL 35487,
USA

²Battelle - National Ecological Observatory Network Inc., Boulder, CO 80301, USA

³Institute of Alpine and Arctic Research (INSTAAR), University of Colorado, Boulder,
CO 80301, USA

[§]Corresponding author: Christina Staudhammer

Phone:

Email: cstaudhammer@ua.edu

Address: P.O. Box 870344, Tuscaloosa AL 35487

Abstract

The high utility of eddy covariance (EC) data has made it the cornerstone of carbon dynamics research for more than two decades. However, a substantial number of measurements from EC data can be missing for various reasons. Robust gap-filling methods are required to estimate carbon budgets from net ecosystem exchange measurements of CO₂ (NEE) with high precision and accuracy. While the gap-filled methods used have provided unbiased estimates of annual NEE, little research has been done on preserving the variance structures associated with gap-filled flux data. In this project, we used EC data from a longleaf pine ecosystem located in the southeast US to investigate variance preservation in gap-filling methods.

We used three non-linear regression approaches to impute artificially created gaps of different sizes via light and temperature response curves: 1) “traditional” fixed monthly window, 2) moving window, and 3) moving window with parameter prediction using physiological drivers. The results of gap-filling simulations showed that the variability of NEE estimates made with moving window and parameter prediction methods were closer to that of observed NEE, whereas the traditional method had overall lower variability. The average root mean square errors (RMSE) of predictions was lower for moving window and parameter prediction (3.38 and 3.22, respectively), versus that of the traditional method (3.42) over one year, including both daytime and nighttime data across all gap sizes. The variances associated with moving window and parameter prediction methods were 52% and 57%, respectively, of the observed variance, versus that of the traditional (51%), while the average of first-order autocorrelation coefficients was 0.76 for each method compared to 0.58 for observed. The results showed that the moving window approaches provided better estimates (lower RMSEs and more similar variance) at annual scales, yet underestimated the observed variance. These results contribute toward the development of a framework of standardized gap-filling approaches which maintain variation inherent in EC data. Moreover, these results call for further research on potential environmental drivers and their interactions for inclusion in gap-filling models, as well as exploration of

42 sampling size of estimation windows and averaging time (half hour) of flux data to promote
43 variance maintenance and decrease the autocorrelation of predictions.

44

45 *Key words: Eddy Covariance, net ecosystem exchange, gap-filling, variance*

46 *preservation, parameter prediction, orthonormal wavelet transformation*

1. Introduction

The study of terrestrial carbon dynamics plays a crucial role in understanding the biological processes that contribute towards atmospheric CO₂ concentrations (Baldocchi et al., 1996; Goulden et al., 1996; Papale et al., 2006). In the present context of global climate change, the eddy covariance (EC) method has improved our understanding of temporally and spatially integrated net ecosystem exchange of CO₂ (NEE) (Baldocchi et al., 1996; Schimel, 2001; Schimel et al., 2001; Stoy et al., 2009). The data produced by the EC method is not only useful to estimate ecosystem-level annual carbon budgets, *i.e.*, temporal integrals of NEE, but also helps to establish functional relationships between NEE and micrometeorological variables, and to devolve below-canopy ecological processes (Loescher et al., 2003; Sierra et al., 2011).

Despite the mathematical and theoretical foundations of the EC method, data cannot be collected under all climatic conditions (Burba, 2013). Datasets routinely have missing data, which arise during the data collection period or from filtering out data that do not meet micrometeorological assumptions (Goulden et al., 1996; Gu et al., 2005; Hollinger and Richardson, 2005; Papale et al., 2006). The EC datasets often have missing periods as a result of: i) rain and condensation in the sampling path, ii) incomplete datasets at the desired averaging times (*i.e.*, 30-min) during system calibration or maintenance, iii) poor coupling of the canopy with the above atmospheric conditions, as defined by the friction velocity, u^* , or iv) excessive variation from the half-hourly mean based on an analysis of standard deviations for wind and CO₂ statistics (Whelan et al., 2013). The series of missing data may extend over hours or days, and are non-randomly distributed (Hagen et al., 2006). For example, at an EC site on the Coastal Plain in the southeastern US, missing data accounted for 63%, 52% and 66% of nighttime values for mesic, intermediate and xeric longleaf pine savannah ecosystems, respectively (Whelan et al., 2013; Starr et al., 2015). EC data gaps are commonly filled utilizing various empirical or semi-empirical methods (Falge et al., 2001; Moffat et al., 2007).

Communities of long-term carbon measurement such as FLUXNET (Baldocchi et al., 2001), AmeriFlux, and the Integrated Carbon Observation System (ICOS) have stressed standardized and statistically defensible gap-filling approaches in order to make synthesis activities feasible, primarily for annual carbon budgets. Despite this effort, a number of gap-filling approaches are used and include non-linear regression (NLR), look-up tables (LUT), mean diurnal variation methods (MDV) (*e.g.*, Falge et al., 2001), artificial neural networks (ANN) (*e.g.*, Papale and Valentini, 2003; Braswell et al., 2005; Moffat et al., 2007), Marginal Distribution Sampling (MDS) (*e.g.*, Reichstein et al., 2005), data assimilation (DAM) or Bayesian model approaches (Gove and Hollinger, 2006). However, their performances vary over time and scale (Hollinger et al., 2004; Moffat et al., 2007), and may not always be statistically defensible. The non-random nature of missing data also introduces bias into predictions, and the autocorrelation among observations artificially deflates their standard errors (Anderson 1954).

A study of a few of these gap-filling techniques (Falge et al., 2001) concluded that MDV, NLR and LUT have similar performance in terms of the systematic bias that is introduced in annual sums of NEE. Furthermore, this bias was found to be proportional to the percentage of gaps filled. While their methods included light and temperature relationships to fill NEE, they recommended adding additional predictors, such as vapor pressure deficit and soil water content, to reduce this systematic bias.

Moffat et al. (2007) followed on this work with a comparison of common gap-filling methods applied to a wider range of datasets that included artificially created gaps of varying lengths in time. This study also showed less bias of gap-filling approaches using additional new techniques such as ANN. However, studies thus far have not adequately addressed preserving the variance in gap-filled data as that observed from the measured NEE.

The variance in data from environmental drivers (*e.g.*, air temperature) used in gap-filling methods has been shown to amplify or dampen flux estimates depending on the degree of

convexity in the flux-temperature relationship (Sierra et al., 2011). Models that rely on average temperature underestimate flux values compared to those that preserve observed variance as a consequence of Jensen's Inequality, and therefore long-term estimates of NEE will be further biased (Moffat et al., 2007). This logic would extend to any case using flux data to estimate any non-linear behavior, as would be expected with changing chronic disturbance (Smith et al., 2009). Hence, additional development of techniques is needed that advances our ability to (i) preserve the variance structures inherent in EC data, and (ii) reduce the systematic bias as well as the error due to gap-filling techniques.

End-to-end systematic and random error propagation from EC data collection and processing steps has not been ensembled into a defensible uncertainty budget, *i.e.*, classic metrology (JCGM 2008, ISO 1995). Instead, the state-of-the-science has been assessing individual sources of both systematic and random errors. For example, observations before gap-filling have systematic and random errors (Goulden et al., 1996; Loescher et al., 2006a), which can be further propagated during gap-filling practices. This issue has been examined by a few studies. Hollinger and Richardson (2005) compared flux measurements from two adjacent tower fluxes to estimate random errors and showed double exponential distribution and heteroscedastic variances. Uncertainty budgets in calibrations and random sources of error were estimated from the AmeriFlux network (Ocheltree and Loescher, 2006), and overall uncertainties in EC techniques can be found in Loescher et al. (2006a). Moffat et al. (2007) compared the error contribution by different gap-filling methods. Wang et al. (2015) used a residual bootstrap approach to quantify uncertainty associated with random measurement and gap-filling errors. The limitations currently found in gap filling techniques establish a need to further investigate gap filling techniques that are more robust, statistically defensible, and work toward preserving the native variance. In this study, we develop statistically defensible methods to re-introduce the variance structure in gap-filled NEE data, and compare and contrast the results of three model approaches. It is not the intention of this study to develop uncertainty budgets.

We evaluated the ability of gap-filling methods to preserve variance inherent in the data structure as a means to provide more robust estimates of annual sums of NEE. Three different gap-filling approaches based on non-linear light and temperature response models were compared. First, a traditional method utilizing static monthly data was used to parameterize these models. In the second method, the same models were used with a monthly moving window rather than the static window. The third method used models with a monthly moving window and the incorporation of additional micrometeorological measurements as additional explanatory variables to better account for the controls on NEE variability. We hypothesized that the gap-filling method with moving window (method 2) would better capture variance in NEE by the use of varying parameters estimated from moving the sampling window. We also hypothesized that using additional environmental variables in the model (method 3) would improve the variance preservation of NEE from 30-min to annual time scales.

2. Methods

After EC data were collected from the field and processed into 30-min estimates (rf. Whelan 2015, Starr et al. 2015, 2016), our first step was to choose a suitable base dataset for analyses. From this base dataset, four case study datasets were created for testing different gap-filling methods. While two of our methods could be directly tested with the case study datasets, implementation of our third method required additional preliminary analyses using the base dataset (Fig. 1).

2.1 Data collection and processing

Carbon fluxes and micrometeorological data were obtained from three established sites along an edaphic gradient (mesic, intermediate and xeric) at the Joseph Jones Ecological Research Center (JJERC) in southwestern Georgia, USA (31.2201°N, 84.4792°W). NEE has been measured continuously at all three sites starting in October 2008 with open path EC techniques, which includes an infrared gas analyzers (IRGA, LI-7500, LICOR Inc., Lincoln, NE), and 3-D sonic anemometers (CSAT-3, Campbell Scientific Inc., Logan, UT). The

instruments were installed in towers above the forest canopy at heights: 34.4, 37.5 and 34.9 m for mesic, intermediate and xeric sites, respectively (~4.0 m above mean canopy height). Micrometeorological data were collected and stored in a datalogger (CR5000, Campbell Scientific Inc., Logan, UT).

Fig. 1. Steps for developing test datasets and models. The base dataset was selected from 12 site years. The PP method required additional preliminary steps to develop parameter prediction (PP) models using environmental variables. Case study datasets were developed from the base dataset, and contained simulated gap scenarios of different lengths. Model runs were made with each of the three gap-filling methods to test for variance preservation capacities.

Raw EC data were processed following methods of Starr et al. (2016) and Whelan et al. (2013) using EdiRe (v.1.4.3.1184; Clement 1999), which uses a 2-d coordinate rotation of the horizontal wind velocities to obtain turbulence statistics perpendicular to the local streamline. The covariance between turbulence and scalar concentrations was maximized through cross correlation of the time series at 0.1 s intervals on both sides of a fixed lagtime (in this case, ± 0.3 s) following Loescher et al. (2006b). Fluxes were calculated for half-hour intervals and then corrected for the mass transfer resulting from changes in density not accounted for by the IRGA (Massman, 2004), and differences in the frequency response between the CSAT3 and the Li-7500 (Massman, 2000). All fluxes were standardized to 1 ATM.

Flux data were filtered to remove systematic errors such as: (1) rain and condensation in the sampling path as indicated by an automatic gain control value >62 (Li-7500 LICOR Manual), (2) incomplete 30-min datasets during system calibration or maintenance, (3) absolute values of the standard deviation for u, v, and w winds > 3.0 , (4) CO₂ concentration outside a plausible range, and (5) poor coupling of the canopy as defined by the friction velocity, $u^*, < 0.20 \text{ m s}^{-1}$ (Whelan et al., 2013). Quality assurance of the flux data was also maintained by examining stationarity criteria, and integral turbulent statistics (Foken and Leclerc, 2004).

2.2 Base dataset selection

The annual dataset that contained the least number of total missing NEE data was selected to maximize the number of data points to develop our gap-filling models. The intermediate site for the year 2009 had the fewest missing data, with ~6200 of 17520 (~ 35%) missing half-hours during the calendar year and was therefore selected as our base dataset.

2.3 Gap-filling approaches

Three different gap-filling approaches were developed and compared (Fig. 1). All approaches utilized non-linear light and temperature response curves commonly found in the literature (Falge et al., 2001; Hollinger et al., 2004; Richardson et al., 2006). For daytime NEE

(when photosynthetically active radiation (PAR) was $\geq 10 \mu\text{mol m}^{-2}\text{s}^{-1}$) a Michaelis-Menten approach was used (Michaelis and Menten, 1913):

$$NEE_{day} = \frac{\alpha f P_{max}}{\alpha f + P_{max}} + R_d \quad \text{Eq. 1}$$

When PAR was $< 10 \mu\text{mol m}^{-2}\text{s}^{-1}$, nighttime NEE data were gap-filled using a Van't Hoff exponential expression (Lloyd and Taylor, 1994; Sierra et al., 2011):

$$NEE_{night} = R_{eco} = R_0 \exp^{bT_{air}} \quad \text{Eq. 2}$$

where, α is the apparent quantum efficiency ($\mu\text{mol CO}_2 \mu\text{mol quanta}^{-1}$), ϕ is PAR ($\mu\text{mol quanta}$), R_d is ecosystem respiration ($\mu\text{mol CO}_2 \text{m}^{-2} \text{s}^{-1}$), P_{max} is the maximum ecosystem CO_2 uptake rate ($\mu\text{mol CO}_2 \text{m}^{-2} \text{s}^{-1}$), R_0 is the base respiration rate when air temperature is 0°C , T_{air} is air temperature ($^\circ\text{C}$), and b is an empirical coefficient (Braswell et al., 2005; Hagen et al., 2006; Loescher et al., 2006a). Micrometeorological convention was used with negative values indicating uptake by the ecosystem.

2.3.1 Traditional, static window method

The first method was a “traditional” non-linear regression method (TRAD) where monthly, static windows (12 non-overlapping windows, each spanning one month during the annual period) of data from each artificial gap dataset were taken to parameterize models of daytime and nighttime NEE functions. In practice, if the coefficients from regression for TRAD method are not significant, predictions are made using annual gap-filling curves (Moffat et al., 2007; Whelan et al., 2013). However, we did not use annual curves here as the goal was to compare the performance of methods rather than estimate carbon dynamics at a particular site.

2.3.2 Moving window method

The second method for gap-filling was the application of non-linear regression in a moving window (MW). The moving window was 31 days in duration, centered on each day of interest (15 days before and 15 days after each day of the year). For instance, if October 16 were the day of interest, the window included data from October 1 through October 31. Using the

moving window, both day and night models were parameterized for each day of interest (January 16-December 16) in each of the artificial gap datasets.

2.3.3 Parameter prediction method

A third method for gap-filling (PP) extended the second technique (MW) by application of parameter prediction. In the parameter prediction method, we attributed the variability of parameters observed in the second method to different environmental variables. The use of environmental variables in place of the constant parameters from daytime and nighttime models (used in the TRAD method) would increase the variability of these parameters, and potentially that of the predictions of missing (gap-filled) NEE observations. The variability of environmental variables thus would contribute to the preservation of variance structures inherent in EC data.

VPD, VWC, and T_{air} have been shown to be environmental drivers of ecosystem functions, and so were assumed to affect the parameters in Eqs. 1 and 2. For instance, VWC changed the light response under drought conditions, and altered ecosystem light efficiency (α) and maximum photosynthetic capacity (P_{max}) (Starr et al., 2016). VPD was also hypothesized to co-vary with these parameters by affecting stomatal opening at the individual plant/leaf level. T_{air} can also affect ecosystem functions through temperature dependent changes in enzymatic rates controlling photosynthesis and respiration processes.

As a preliminary step for parameter prediction, we therefore investigated how the coefficients from each of the day and night models (rf. Eq. 1, 2) varied with additional environmental variables. These preliminary models enabled us to identify other variables with additional explanatory power not accounted by the base gap-filling functions (as suggested by Falge et al. 2001 and others). We first applied the MW method to the base dataset to obtain estimated parameters (α , R_d , P_{max} , R_0 , and b) for each window. Then, multiple regression models were utilized to determine the strength of additional environmental variables to predict these parameter values. The variables we tested were: vapor pressure deficit (VPD, KPa), soil

volumetric water content (VWC, %/100), T_{air} , albedo (%/100), relative humidity (RH, %), barometric pressure (BP), horizontal windspeed (u , $m\ s^{-1}$), and soil temperature (T_{soil} , $^{\circ}C$), utilizing mean quantities in each moving window. The associated sensors, location and transformations (if any) are reported in Whelan et al. (2013, 2015). The variance inflation factor (VIF) was calculated to ensure that co-linearity among variables did not influence the significance of each individual variable; $VIF < 10$ was used as a threshold for variable inclusion in models (O'Brien 2007). Based on the strength of the significance and VIF check, three micrometeorological variables (VPD, VWC, and T_{air}) were determined to have strong relationships with the three daytime and two nighttime NEE parameters, and so were included in further analyses.

To determine the best models to use in the PP method of gap-filling, all possible multiple-linear combinations of VPD, VWC, and T_{air} were used in Eqs 1 and 2 to replace each parameter (α , R_d , P_{max} , R_0 , and b). For instance, α was replaced with $(\alpha_0 + \alpha_1 VPD + \alpha_2 VWC + \alpha_3 T_{air})$. For example, one model that incorporated VPD and T_{air} for daytime NEE was:

$$NEE_{day} = \frac{(\alpha_0 + \alpha_1 T_{air} + \alpha_2 VPD)\phi(P_{max_0} + P_{max_1} T_{air} + P_{max_2} VPD)}{(\alpha_0 + \alpha_1 T_{air} + \alpha_2 VPD)\phi + (P_{max_0} + P_{max_1} T_{air} + P_{max_2} VPD)} + (R_{d0} + R_{d1} T_{air}) \quad \text{Eq. 3}$$

We modeled all possible combinations of variables using one, two, or three environmental variables to replace parameters in Eqs. 1 and 2 individually and combined together. In the case of nighttime NEE, R_0 was replaced with a combination of VPD and VWC only, assuming the base respiration rate at $0\ ^{\circ}C$ would not change with changing temperature (by definition). We refer to these statistical models as 'potential PP models' in the remainder of the text.

Starting parameter values were chosen so that the linear combination of replaced parameters would be within the range of physically defined parameters (α , R_d , P_{max} , R_0 , and b). Multiple combinations of starting values were tested to avoid the risk of a local maxima. In order to satisfy statistical assumptions, residuals from each potential PP model were tested for normality and homoscedasticity. It is important to note that with time series data (particularly

EC data), the assumption of independence is often violated. Because error variances are deflated in proportion to the strength of the autocorrelation (Anderson 1954), p-values are artificially deflated. We therefore present model significance (p-values via a standard t-test) as a measure for relative comparisons among potential PP models. The number of significant parameter estimates ($p < 0.05$), as well as the average root mean square errors (RMSE) over all moving windows were computed to compare among potential PP models,

In addition, potential PP models were also estimated and compared with moving windows of 15 days and 60 days to analyze the efficacy of different window lengths. Based on the p-values and average RMSE values, a 31-day moving window was favored (results not shown).

All gap-filling and other analyses were performed in the RStudio platform (Version 3.2.3, www.rstudio.com/products/rstudio). Non-linear regression parameters and associated fit statistics were estimated via the *nls* (nonlinear least square) function, and AR(1) (first-order autocorrelation coefficient) values were estimated using *acf* function from the *stats* package.

2.4 Datasets with artificial gaps

We created four datasets with artificial gaps. First, a mixed gap file was created using non-missing data from the selected base dataset, with gaps of different sizes based on the frequency of gap sizes in that base dataset. Locations in the dataset where there were ≥ 12 -h of consecutive data points were identified, and out of these locations, 45 12-h gaps (24 consecutive observations; 1080 missing observations in total) were randomly introduced to the dataset. Similarly, locations where > 5 -h of consecutive data points were identified, ensuring that these would not overlap with already introduced 12-h gaps. Out of these locations, 50 5-h gaps (10 consecutive observations; 500 missing observations in total) were randomly introduced to the dataset. Then, 1520 $\frac{1}{2}$ -h gaps (*i.e.*, one $\frac{1}{2}$ -h period) were introduced randomly to the remaining parts of the dataset, ensuring there was at least one non-missing point before and after the gap

(Fig. 1). The mixed artificial gap dataset was used to examine characteristics of the three gap-filling methods on an annual basis, to provide evaluation of the methods in an operational sense.

The proportions and lengths of artificial gaps were created to be similar to those generated operationally in the selected sample site. The total proportion of missing data in the base dataset was ~35%. The original dataset had a 2:1 ratio for nighttime gaps to daytime gaps. However, matching the proportion of daytime to nighttime gaps from original dataset limited the number of retained data points for nighttime conditions. Thus, due to the already reduced numbers of available data for nighttime, introduced artificial gaps were in a ratio of 1:1 daytime:nighttime. Artificial gaps were not introduced within the first 15 days of the year or the last 15 days of the year, as these gaps would need data from the end of the previous year or the start of the following year, respectively, to make the appropriate moving windows complete (31-d).

Next, we created artificial gap datasets to more closely examine the differences among methods with respect to particular gap lengths. We created three separate datasets which had; 199 5-h gaps (1990 missing observations); 100 twelve-hour gaps (2400 missing observations); and 9 3½-d gaps (1512 missing observations). Since there was no 3½-d period without missing observations, these gaps were randomly introduced despite the presence of missing observations in the base dataset. Each of the three gap-filling methods was used on these three gap files (12-h, 5-h and 3½-d) to analyze the performance of individual gap-filling methods specifically at different gap-lengths.

2.5 Orthonormal wavelet transformation (OWT)

Orthonormal wavelet transform (OWT) has been previously used in flux data for spectral analysis and shown the efficiency of decomposing flux signals into time-scale (Katul et al., 2001; Stoy et al., 2005; Torrence and Compo, 1998). Detailed discussion of wavelet analysis and its application in flux research can be found in Stoy et al. (2005), Katul et al. (2001), Torrence and Compo (1998), and Katul and Parlange (1995). Here, we used OWT to compare the spectra of observed versus predicted NEE across gap-filling methods to determine the variance explained at

different time scales. For our analysis, we used a Haar wavelet basis function (a square wavelet) to estimate wavelet coefficients due to its localization in the temporal domain and minimum effects of gaps or discontinuities (rf. Stoy et al., 2009).

To compute the OWT coefficients, we first multiplied all half-hourly measurements in the dataset by a quality control flag (1 for raw measurements and 0 for missing data) and then separated the daytime NEE and nighttime NEE data. For daytime NEE, corresponding nighttime values were multiplied by 0 and vice versa (Stoy et al., 2009). Both daytime and nighttime NEE time series were then normalized to have mean zero and variance one. As the transformation decomposes the signal to half of the number of coefficients at each subsequent scale, the length of the time series must be a power of 2 (Burrus et al., 1998). In order to do so, zeroes were added to each end of the time series for both day and night, which were again normalized to have unit variance. This process was repeated for the base dataset, as well as the four artificial gap datasets.

The wavelet transformation coefficients were calculated using the ‘*wavelets*’ package in RStudio and coefficients for each time scale (from 1 hour to 1 year) were averaged. The number of coefficients at low frequency (6 months or 1 year) was limited and therefore estimates for these time scales are not robust.

2.6 Statistical analysis to compare gap-filling methods

The RMSE, bias, variance, and AR(1) coefficient were computed to compare the performance of each of the three different gap-filling methods. RMSE and bias for each of the gap-lengths were used to analyze the prediction precision and accuracy of the methods at different gap sizes. While it is expected that gap-filled data will have less variance and more autocorrelation than the observed data, the variance and AR(1) coefficient provide a useful comparison to evaluate the degree of variation that was preserved. For each gap location (longer than half hour), the variance after gap-filling with each of three methods was calculated and evaluated against the variance of observed NEE. Since different datasets can have the same

variance and yet still have different patterns and distributions, the AR(1) coefficient of data in gap windows provides an additional measure of the variability of adjacent observations. The AR(1) coefficient calculated in the filled artificial gaps $> \frac{1}{2}$ hour was compared to that of the observed data to quantify how gap-filling changes the autocorrelation structure of the data. In addition, RMSE, variance and bias were calculated separately for day and night predictions to check the consistency of model predictions.

3. Results

3.1 Environmental variables and model parameters response

The preliminary models indicated that the estimated parameters in Eqs. 1 and 2 were strongly related to VPD, VWC, and T_{air} , and thus, these environmental variables explain variability in daytime NEE (Table 1, Fig. 2). For each parameter, however, the strongest relationship was with T_{air} (Table 1, Fig. 2), which showed strong linear relationships with R_d and R_0 . With the nighttime model, this gives evidence for a potential secondary response of air temperature, i.e., a modification of the effect of temperature in the Van't Hoff exponential expression (Eq. 2).

3.2 Selection of day and night models

Based on the analyses of the preliminary multiple regressions, the original Eq. 1 and 2 parameters were replaced with linear combinations of VPD, VWC, and T_{air} to evaluate potential PP models. The performance from these models was evaluated using RMSE, and the proportion of moving windows with significant parameters for the whole year (Table 2 and 3). RMSE values for both daytime light response and nighttime temperature response models were similar in most cases (Tables 2 and 3), and thus the proportion of significant parameters was the main determinant in selecting which variables should be included within the PP method. Several of the more complex potential PP models would not converge, such as a 12-parameter model where all three light response parameters were simultaneously replaced by combinations of all 3 environmental variables (data not presented), and thus these combinations were excluded. A

total of 7 daytime and 8 nighttime potential PP models were compared (Tables 2 and 3). The selected PP models for daytime and nighttime gap-filling were:

$$NEE_{day} = \frac{(a_0 + a_1 VPD + a_2 VWC) f(P_{max_0} + P_{max_1} VPD)}{(a_0 + a_1 VPD + a_2 VWC) f + (P_{max_0} + P_{max_2} VPD)} + (R_{d_0} + R_{d_0} T_{air}) \quad \text{Eq. 4}$$

$$NEE_{night} = R_{eco} = (R_{0_0} + R_{0_1} VPD) \exp^{b T_{air}} \quad \text{Eq. 5}$$

Table 1. Estimates, standard errors, t-values and associated p-values, and variance inflation factors (VIF) for preliminary models relating parameters from light (α , P_{max} and R_d) and temperature (R_0 and b) response curves to environmental variables: vapor pressure deficit (VPD), volumetric water content (VWC) and air temperature (T_{air}). VIF values were calculated to test for multicollinearity among environmental variables.

Parameter	Effect	Estimate	Std. Error	t value	Pr(> t)	VIF
α	Intercept	0.02916	0.0133	2.192	0.0291	
	VPD	0.00864	0.00171	5.039	<0.0002	2.77
	VWC	-0.00117	0.00061	-1.921	0.0556	2.285
	T_{air}	-0.00215	0.00007	-28.93	< 0.0002	2.381
P_{max}	Intercept	-26.5756	6.53287	-4.068	< 0.0002	
	VPD	12.4124	0.84225	14.74	< 0.0002	2.77
	VWC	0.75589	0.30031	2.517	0.0123	2.285
	T_{air}	-0.99997	0.03663	-27.3	< 0.0002	2.381
R_d	Intercept	-18.8592	1.3186	-14.3	< 0.0002	
	VPD	-1.77063	0.17	-10.41	< 0.0002	2.77
	VWC	0.82912	0.06062	13.68	< 0.0002	2.285
	T_{air}	0.4561	0.00739	61.69	< 0.0002	2.381
R_0	Intercept	-8.4083	21.5245	-4.107	0.0056	
	VPD	-17.849	4.598	-3.882	< 0.0002	1.596
	VWC	4.142	0.9425	4.395	0.010	2.287
	T_{air}	1.2871	0.1276	10.09	< 0.0002	2.266
b	Intercept	0.28468	0.05686	5.007	<0.0002	
	VPD	0.02554	0.01215	2.102	0.0363	1.596
	VWC	-0.01029	0.00249	-4.135	<0.0002	2.287
	T_{air}	-0.00485	0.00034	-14.4	< 0.0002	2.266

375
 376 Fig. 2. Monthly average quantum efficiency (α) as a function of (A) vapor pressure deficit (VPD) and (B) volumetric water content (VWC); maximum
 377 photosynthetic capacity ($P_{\max}$) as a function of (C) vapor pressure deficit (VPD); daytime respiration (R_d) as a function of (D) air temperature (T_{air}); and base
 378 respiration (R_o) as a function of (E) air temperature (T_{air}).

379 Table 2. Summary of significant parameter estimates and root mean square error (RMSE) values
380 for evaluating potential parameter prediction (PP) models with various combinations of
381 environmental variables in the light response curve. The numbers in parameter columns
382 represent the percentage of significant parameter values out of the total number of windows
383 (329). The first column for each parameter contains the proportion significant for the intercept
384 of that parameter prediction (e.g., α_0). Bolded row shows the selected PP model to be used in
385 gap-filling method comparisons.

Model	α				P_{max}				R_d		RMSE
	α_0	VPD	VWC	T _{air}	P_{max0}	VPD	VWC	T _{air}	R_{d0}	T _{air}	
1	86	77		62	100				100		3.5582
2	100				87	74		62	100		3.5321
3	60	82	62	62	100				100		3.5248
4	100				59	76	46	52	100		3.5061
5	58	71	57		92	85			96	67	3.5280
6	60	71	53		85	69	55		98	64	3.6238
7	60	63	83		82	85		47	50	64	3.6335

386

Table 3. Summary of significant parameter estimates and root mean square error (RMSE) values for evaluating potential parameter prediction (PP) models with various combinations of environmental variables in parameter prediction in the temperature response curve. The numbers in parameter columns represent the percentage of significant parameter values out of the total number of windows (329). The first column for each parameter contains the proportion significant for the intercept of that parameter prediction. Bolded row shows the selected PP model to be used in gap-filling method comparisons.

Model	R_0				b			RMSE
	R_0	VPD	VWC	T_{air}	b_0	VPD	VWC	
1	15		25		56			3.1942
2	76	53			65			3.1835
3	82			35			29	3.1945
4	73				71	51		3.1860
5	21	19	53		60			3.1693
6	70				33	24	50	3.1728
7	66	35			47	23		3.1770
8	35	45	37		55	37	54	3.1369

3.3 Performance of gap-filling methods

The temperature and light response curves of NEE derived with the mixed-gap dataset using the TRAD method with static monthly windows appeared fairly smooth with low variability for each month's predictions (Figs. 3B and 4B). The MW method progressively captured seasonal changes and provided better prediction with higher variance (closer to the measurements) as seen in its temperature and light response curves (Figs. 3C and 4C). The results from the PP method showed an increase in the variance of the predictions, which is higher than that of MW (3D and 4D). Focusing only on mid-June to mid-August (days of year 175-250), when stable nighttime conditions lead to frequent data gaps, plots of observed NEE (day and night) and predictions from all three gap-filling methods demonstrated the different amount

of variability resulting from these gap-filling approaches (Figs. 5 and 6). TRAD showed consistent but comparatively lower variability at each monthly period while the variability changed among months (Figs. 5 and 6). MW and PP appeared to have higher variances than the TRAD method. Measures such as RMSE and variance, however, could not quantify the degree to which the native variance was preserved.

Fig. 3. Daytime net ecosystem exchange of CO₂ (NEE_{day}) as a function of photosynthetically active radiation (PAR): (A) Observed, and predictions by different gap-filling methods utilizing the mixed gap dataset: (B) TRAD, (C) MW, and (D) PP.

Fig. 4 Nighttime net ecosystem exchange of CO₂ (NEE_{night}) as function of air temperature (T_{air}): (A) Observed, and predictions by different gap-filling methods utilizing the mixed gap dataset: (B) TRAD, (C) MW, and (D) PP.

Fig. 5. Daytime observed net ecosystem exchange of CO_2 (NEE_{day} ; black), and predicted nighttime net ecosystem exchange using traditional method (TRAD, green), moving window (MW, red) and parameter prediction (PP, blue) for days of year (DOY) 175-250 (Summer), utilizing the mixed gap dataset.

Fig. 6 Nighttime observed net ecosystem exchange of CO_2 , (NEE_{night} ; black), and predicted nighttime net ecosystem exchange using traditional method (TRAD, green), moving window (MW, red) and parameter prediction (PP, blue) for days of year (DOY) 175-250 (Summer), utilizing the mixed gap dataset.

3.4. Spectral analyses

The OWT spectra of NEE_{day} showed that the variance of predictions made with the PP method were more similar to that of the measured NEE time series across all time-scales (Fig. 7), as compared to the other methods. With each of the artificial gap scenarios, all models underestimated the variance at the 1 h-to-2 h timestep. For daytime NEE, all models returned variances similar to that of the measured NEE at the 4 h-to-month timestep for the 12 h and 3.5 d gap scenarios (Fig. 7C-F), and at the seasonal-to-year timestep across all artificial gap scenarios. However, variance of MW predictions at the monthly-to-seasonal timestep for the mixed gap

Fig. 7. Orthonormal wavelet spectra of daytime (left panel) and nighttime (right panel) observed and predicted NEE by gap-filling methods at different gap-scenarios. The first row (A,B) for mixed gap-scenarios, second (C,D) for 3.5 d gaps, third (E,F) for 12 h gaps and fourth (G,H) for 5 h gaps.

scenario was overestimated (Fig. 7A). For nighttime NEE, all models over estimated variance from 8 h-to-season across all artificial gap scenarios. While TRAD also showed closer variance to that of other methods at most timesteps, the variance of TRAD predictions was very different from other models at a monthly timestep. At high frequency (1 h-to-8 h timestep), the prediction patterns among the models were similar for all artificial gap lengths. Moreover, the variances of these predictions were lower than that of the observed, and as the frequency of artificial gaps decreased, the variance of the predictions were overestimated at most timesteps, and in particular for nighttime.

3.5 Imputed artificial gaps

When comparing the performance of gap-filling methods for different-sized artificial gaps in the base dataset, the average RMSE of simulated gap lengths of 5-h, 12-h and 3.5-d showed only minor differences among different methods; PP had lower RMSE during the day, whereas TRAD had lower RMSE at night (Table 4, Fig. 8). RMSE values were higher and more variable during the second and third quarter of the year for all gap-filling methods (data not presented). The range of RMSE values was higher with increased gap-length; however, median RMSEs across all methods and sizes were close (Fig. 8). On average, MW had lower estimated variances, but all methods had similar interquartile ranges. In the case of 3½-d gaps, the median PP and MW variances were higher than that of the observed data. The variance of nighttime predictions was much lower than that of the observed, regardless of method (Table 4). In general, at smaller gap-lengths the variance was underestimated (which was also observed in the spectra). The half hour and 5-h gaps were almost evenly distributed throughout the dataset

462 Table 4. Averages value of RMSE, bias, variance, and AR(1) coefficients of gap-filled
 463 predictions during daytime, nighttime, and overall for each of three gap-filling methods for three
 464 datasets with introduced gaps of different lengths. Observed variance and AR(1) were computed
 465 only where predictions were made. The AR(1) coefficient was not computed separately for day
 466 and night observations, as gaps spanned both day and night conditions.

		Gap Length	TRAD	MW	PP	Observed
RMSE	Daytime	5-h	3.03	2.97	2.78	
		12-h	3.01	2.91	2.78	
		3.5d	3.09	2.88	2.75	
	Nighttime	5-h	1.96	1.99	1.98	
		12-h	2.14	2.18	2.18	
		3.5d	2.35	2.56	2.56	
	Overall	5-h	2.86	2.82	2.67	
		12-h	2.76	2.74	2.61	
		3.5d	2.83	2.73	2.66	
Bias	Daytime	5-h	0.10	0.21	0.26	
		12-h	0.25	0.32	0.38	
		3.5d	0.07	0.19	0.46	
	Nighttime	5-h	0.26	0.41	0.41	
		12-h	0.18	0.20	0.20	
		3.5d	0.51	0.59	0.65	
	Overall	5-h	0.16	0.29	0.32	
		12-h	0.18	0.31	0.37	
		3.5d	0.14	0.34	0.53	
Variance	Daytime	5-h	9.12	9.08	10.75	20.64
		12-h	17.42	16.58	19.52	26.48
		3.5d	18.15	16.07	18.58	22.67
	Nighttime	5-h	0.02	0.03	0.03	4.29
		12-h	0.03	0.08	0.10	2.39
		3.5d	0.26	0.13	0.16	5.27
	Overall	5-h	7.87	7.88	9.08	17.94
		12-h	20.96	20.12	22.34	28.34
		3.5d	28.78	26.81	27.36	30.74
AR(1)	Overall	5-h	0.56	0.55	0.55	0.29
		12-h	0.78	0.78	0.79	0.62
		3.5d	0.96	0.96	0.96	0.82

Fig. 8. Distribution of the values of RMSE, variance, and AR(1) coefficients by gap-filling method (TRAD, MW and PP) for three datasets with introduced gaps of different lengths.

while the 12-h gaps were mostly located in the first half of the year due to the lower availability of continuous 12-h NEE observation in the base dataset. Thus, this might have introduced bias in quantifying the preservation of variance for this gap size. The AR(1) coefficients indicated that for all gap-sizes, there was higher autocorrelation in predicted values (for all three methods) than in observed values (Table 4, Fig. 8), especially for shorter duration gaps.

4. Discussion

The comparison of gap-filling approaches showed that the use of a MW allowed more flexibility to change gap-filling model parameters that correspond to ecologically meaningful and changing environmental factors. Thus, finer-scale NEE responses to changes in the environment (such as assimilation rate in response to stochastic rain events) could be captured by the use of MW in contrast to the TRAD method. The MW method was also able to better capture variation in NEE due to major phenological events (a key biological source of variability in NEE estimates), such as NEE responses due to seasonal weather patterns; the timing of needle cast, needle drop, etc. Use of this method (and PP method equally) however, cannot reflect changes in parameter estimates for shorter time periods during which such events may occur. Rather, the estimated parameters represent a mean response during the sampling window and the effect of such phenological activities is dampened, decreasing our ability to estimate the full magnitude of such an event.

Changes in abiotic drivers that are cumulative in nature would hence have an effect that is evenly distributed across the moving window. This has implications for different types of environmental forcings over different ecosystem types. For instance, a single precipitation event could increase VWC for a few days (as in the case of a xeric forest or desert chaparral), and this increase would affect only a small portion of the observations in the data window. Because the

sample for parameterization is much larger than the time period of the event, and there may be a time lag in the ecological response that is not accounted for, the sample may not adequately describe the impact of precipitation events. The magnitude of parameters would likely not reflect the range of this variability, and may explain the inherently lower modeled variance than that observed in the native datasets. However, if precipitation occurred frequently during a particular window, the change in VWC would be significant in PP. The magnitude of parameters in PP might be reduced due to size of the sampling window, and might not fully capture the full range of variability, but it would more closely approximate the *true* variability (as in the case of a wet tropical or Pacific NW forest). In the latter example, The TRAD method could potentially reflect such events equally as well as the MW, given that the event occurrence aligned with a calendar month.

A shorter moving window encompassing fewer days for MW and PP methods could potentially show variability of predictions to be close to the observed. However, some potential PP models that we investigated with a 15 d moving window had higher RMSE values. Given that there is an optimal period of moving window with lower RMSE, variability at monthly time scales, as well as diurnal scales could be predicted to be closer to the observed. Optimization of the moving window length could be tuned for the expected changes in annual abiotic drivers, and would have to be tested further in each ecosystem of interest.

The PP model demonstrated a better ability to preserve the variance and had low overall error; however, it had higher overall bias when compared to the other methods tested here (Table 4). As observed in the spectra (Fig. 7), the low variance of predictions for all three methods at 1 h-to-8 h timesteps could be due to the larger window size (any event responsible for increase in NEE signal strength is distributed over the sampling window). On the other hand, higher

variance at timesteps 8 h-to-annual for nighttime and at timesteps 8 h-to-biweekly for daytime predictions could be due to accounting for the same event in multiple windows (in the case of MW and PP) or accounting for one major event (responsible for higher variability) for a whole month (in the case of TRAD). With site- and time-specific models, adjustments to the moving window length and to further account for the physical and/or physiological controls of NEE variability, we predict that the PP method or its derivative can be useful in preserving variance while also reducing uncertainties.

Variance in observed NEE is attributed to the occurrence of various physical and physiological phenomenon in vegetation (Falge et al., 2001; Falge et al., 2002; Hagen et al., 2006; Jimenez et al., 2012; Loescher et al., 2006; Moffat et al., 2007). For example, environmental factors responsible for variability of NEE at a longleaf pine forest include needle drop, prescribed fire (Starr et al., 2013, Whelan et al., 2015), wind sweep events and morning venting of the canopy. Although our model captured some of the seasonal features via moving window and parameter prediction, the gap-filling methods still seemed to be unable to match larger variances observed in the measured data (at ½-h time scales). As seen in the Figs. 4 and 5, variances of extreme NEE measurements (both during daytime and nighttime) are not captured by any of our models. Preserving the variance that can capture those most extreme observations would be very difficult because of their stochasticity and short time duration in which they occur. However, since NEE observations are computed averages of high frequency (10Hz) measurements, examining spectral characteristics of fluxes at different time scales (Katul et al., 2001) and analyzing the influence of potential environmental factors at the turbulent time scale (using ogive analysis, for example following Loescher et al., 2006) at various averaging time

periods could help reduce uncertainty as well explain physical sources of NEE variability, and may be possible in the future.

While increasing averaging time (1-h or higher) has been proposed to account for all the low-frequency transport of mass and energy (Loescher et al. 2006b), reducing averaging time to account for the high-frequency contributions to understand the variance associated with ecological processes has not been previously suggested. Here, the justification to reduce the averaging time is based on the loss of ecological information. This study suggests that ecological processes occurring at shorter times inform the variability of NEE processes. One of the primary micrometeorological assumptions in the EC technique is of consistent turbulence (stationarity) over the averaging period, *i.e.*, in the control volume when applying the conservation equation (Loescher et al. 2006). Moreover, a chief reason why EC is such a robust measure is its ability to measure the exchange of mass and energy at the interface of the ecosystem and the atmosphere in the well-mixed surface layer. This interaction also embodies the physical influence of the ecosystem structure, particularly those with large flexible canopies, to convey ecological information (scalar processes) to the turbulent environment (vector processes). For example, induced ramp effects or coherent structures can build-up and dissipate wind energy at the frequency of 6-10 sec (Thomas and Foken 2007, Loescher et al., 2006b). These coherent structures contain ecological information that are transported past our sensors in a matter of a few seconds, and further suggest a mechanism that informs the variance structure in our 30-min averages.

Here is the conundrum in the EC approach and at the crux of being able to estimate the sources of the native variance structure and improve the preservation of this variance for model predictions and longer temporal integrals, re. annual budgets (Berger et al. 2001; Loescher et al.

2006b). On one hand, the EC technique requires a standardized averaging time to meet the practical needs of the modeling community, to facilitate merging climate datasets, and to assure some known measure of stationarity, *i.e.*, 30-min. On the other hand, the variance that is embodied in the shorter turbulence timescales contains important ecological or physical information that cannot be retrieved from the 30-min averages. Even though we are able to estimate new sources of variance into gap-filling techniques, our study suggests reexamining the raw high frequency time series (e.g., 10 Hz) for additional ecological and micrometeorological sources of variance at different timescales < 30-min, *e.g.*, ogive analysis (Berger et al. 2001; Loescher et al. 2006b).

Attempts to increase the flexibility of gap-filling models to capture the variance in NEE led to complexity of our model(s), and in some cases, the complex models did not converge. The somewhat complex models in our PP method did not notably improve the estimation of NEE values or the preservation of variance of observed NEE (Table 4) when compared to the comparatively simpler MW method, which also showed more bias. The parameters expressed in Eqs. 1, α , the quantum efficiency, and P_{max} , maximum ecosystem CO₂ uptake rate, inherently and indirectly depend on other ecophysiological factors. For example, α depends on water availability, water use efficiency and stomatal conductance, which in turn can be interpreted by proxy of VPD and VWC (Monteith and Unsworth, 2007). Similarly, P_{max} is also dependent on VPD (Monteith and Unsworth, 2007), and the daytime respiration parameter R_d can be directly proportional to T_{air} . For NEE_{night}, (Eq 2) a model using VPD in a linear form to replace R_0 performed better in terms of RMSE and the percentage of significant parameters (Eq. 5). Interestingly, there has not been a previous approach that explicitly incorporated these environmental variables into our current understanding, *i.e.*, Eq 1, 4. In the non-linear model

used in the PP method, we assumed that the parameters would have a linear response to the environmental variables. However, it is possible that the response to one environmental variable or response to an interaction of multiple variables is non-linear. Further investigation of how the interaction of VPD, VWC and/or temperature (and potentially other environmental variables) changes each of the parameters will be important for parameter prediction methods to improve gap-filling or other modeling approaches. In addition, the size of moving window would potentially have prominent effects on the preservation of variance.

The ability of the environmental variables in the PP method to inform the modeled variance might also have been dampened because of the size of the window. The variability in the environmental variables and NEE differs at different time scales (Stoy et al. 2009). Although we investigated the model predictions from the gap-filling methods at different window lengths (15-d, 31-d and 60-d), we did not thoroughly explore the variability of environmental variables and NEE at these and other time scales. In addition, we also did not analyze long-gap scenarios (>10 d), longer sampling windows could be defensible for such gaps. Because the largest fraction of gaps were <<10-d, however, we focused on the gaps fractions that have the largest utility to the flux community and excluded long-gaps in this analysis. Additionally, such gaps usually occur due to stochastic events such as natural disasters (*e.g.*, lightning or windthrow), loss of a sensor, power outages, etc., which would require the re-parameterization of models.

The predictions for nighttime NEE using the MW and PP methods had higher variability than the TRAD method for 5 h and 12 h gaps, while TRAD had higher variability than MW and PP for 3.5 d gaps (Table 4, Fig. 4, temperature response curve and Fig. 6). The observed nighttime data, however, had higher variability (Fig. 6). This means that nighttime data having higher variability combined with most of the missing data occurs at night (re. filtered due to

insufficient mixing) (Gorsel et al, 2008), limits our ability to estimate the native variance by a gap-filling method. Even though it was for different reasons, Sierra et al. (2011) and Gorsel et al. (2009) also advocate linking the nighttime NEE to soil and plant respiration components as well as other environmental variables to improve nighttime gap-filling predictions.

Environmental variables such as VPD, VWC, T_{air} , and/or wind motions can influence nighttime NEE at different time scales, beyond which we have analyzed in this research. Further analysis of co-spectra of environmental variables and NEE could also shed light on these additional sources of information.

Further examination of the (co-)spectra of other environmental variables on NEE could also elucidate sources of variance in daytime NEE. One interesting observation of daytime NEE at a diurnal scale in our dataset was a hysteresis affect during extreme values of VWC (low: ≤ 0.04 and high: ≥ 0.12). Using NEE data that was binned by PAR (bins of $200 \mu\text{mol m}^{-2} \text{s}^{-1}$, from $200\text{-}2000 \mu\text{mol m}^{-2} \text{s}^{-1}$), we analyzed the carbon assimilation at very high, average, or very low daily average water availability conditions. The graph of PAR-NEE showed peak NEE measurement during midday during high VWC days (Fig. 9), with carbon assimilation increasing with PAR until about midday, and slowly decreasing and shutting down with decreasing PAR when VWC was average. This pattern was also evident on some days with average water availability (data not shown). In contrast, at low water availability, the carbon assimilation increased with PAR until ~10 am and gradually decreased even with the increasing PAR (Fig. 9). Similar hysteresis effect showing departure from normal light response curve was also shown by Pingintha et al. (2010) in peanut crops in same region (Unadilla, GA). This indicates that accounting for diurnal hysteresis can also improve our ability to incorporate the native variance in our gap-filling models.

Having incorporated the major environmental variables that are responsible for the variability of NEE at many different scales (however, in limited time-scale combinations) in our model, we were still unable to preserve all of the variance, although PP and MW were closer than the traditional method. Further research should investigate new approaches that could potentially improve variance preservation, such as: 1) using categorical independent variables in the model (e.g., at different water availability conditions, temperature conditions or other environmental conditions), 2) incorporating time-scale variability of environmental variables and NEE in the model, 3) using non-linear responses of environmental variables in parameter prediction, 4) examining how the variability in the turbulent times series can better inform our ability to model the native variance at 30-min average, 5) accounting for diurnal hysteresis, and 6) examining different averaging times of the NEE data.

Fig. 9. Average NEE by PAR (200 -2000 $\mu\text{mol m}^{-2} \text{s}^{-1}$ with 200 $\mu\text{mol m}^{-2} \text{s}^{-1}$ increments): (A) during days with highest average VWC ($\text{VWC} > 0.12 \text{ m}^3 \text{m}^{-3}$; DOYs 86-88, 91, 93, 94, 244, 289, 359) and (B) during dates with lowest VWC ($\text{VWC} = 0.04 \text{ m}^3 \text{m}^{-3}$; DOYs 70-78).

5. Conclusion

Gap-filling EC data is challenging (especially during nighttime) because of complex response of plant physiology to environmental conditions. Some of the responses may be lagged or stochastic (rainfall) while other demonstrate diurnal or seasonal patterns (PAR, T_{air}). Embracing the factors that influence variability of NEE at different time-scales and magnitudes in these methods with further investigation will reduce uncertainty in flux observation and promote understanding of the complexity of ecosystem physiology in response to variability of environmental forcings.

Analyses of these forcings and responses, and incorporating them into models will improve our predictions and forecasting abilities, while preserving variance at different time-scales. Further investigation can also be extended at different geographical location or to different plant functional types to advance our knowledge of ecosystem complexity with the pretext of changing climate. Hence, we conclude that a gap-filling method (or related assimilation activities) should be flexible enough to capture the variability of ecosystem response at different time scales. In order to quantify carbon budgets and understand physiology at different spatial scales with reduced uncertainty, flexibility of gap-filling methods is essential. Flexibility of gap-filling methods can be achieved by adjusting models for variability in the length of moving window based on the length of missing data as well as inclusion of environmental variables based on climatic variability or occurrence of major synoptic weather patterns. In addition, using the altered averaging times for measured NEE could potentially help attain the variance structures that we have missed by using our gap-filling methods.

669 **Acknowledgements**

670 We would like to thank Staudhammer lab, Starr lab and Scientific Writing and
671 Presentation class at the University of Alabama for review of the manuscript and providing
672 valuable feedback. This research was funded by the University of Alabama, the Joseph W. Jones
673 Ecological Research Center (JJERC) and the Department of Defense – ESTCP (RC#-201467).
674 The authors would also like to acknowledge the excellent logistical support of the JJERC. HWL
675 acknowledges the National Science Foundation (NSF), EF-102980, for their ongoing support.
676 The National Ecological Observatory Network (NEON) is a project sponsored by the NSF and
677 managed under cooperative agreement to Battelle. Any opinions, findings, and conclusions or
678 recommendations expressed in this material are those of the authors and do not necessarily
679 reflect the views of our sponsors. This research would not have taken place without decades of
680 collegial interactions and community engagement with our peers and mentors.

References

- Anderson, R. L. (1954). The problem of autocorrelation in regression analysis. *Journal of the American Statistical Association*, 49(265), 113-129. doi: 10.2307/2281039.
- Baldocchi, D., Valentini, R., Running, S., Oechel, W., & Dahlman, R. (1996). Strategies for measuring and modelling carbon dioxide and water vapour fluxes over terrestrial ecosystems. *Global Change Biology*, 2 (June 1995), 159–68. doi:10.1111/j.1365-2486.1996.tb00069.
- Baldocchi, D. Falge, E. Gu, L. Olson, R., Hollinger, D., Running, S., Anthoni, P.; Bernhofer, Ch., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., & Meyer, S. (2001). FLUXNET: A new tool to study the temporal and spatial variability of Ecosystem-Scale Carbon Dioxide, Water Vapor, and Energy Flux Densities. *Bulletin of the American Meteorological Society*, 82, 2415-2434.
- Berger, B. W., Davis, K. J., Yi, C., Bakwin, P. S., & Zhou, C. (2001) Long-term carbon dioxide fluxes from a very tall tower in a northern forest: Part I. Flux measurement methodology. *Journal of Atmospheric and Oceanic Technology* 18, 529-541.
- Braswell, B. H., Schimel, D. S., Linder, E., & Moore, B. (1997). The response of global terrestrial ecosystems to interannual temperature variability. *Science*, 278, 870–872.
- Braswell, B.H., Sacks, B., Linder, E., & Schimel, D.S. (2005). Estimating ecosystem process parameters by assimilation of eddy flux observations of NEE. *Global Change Biology* 11, 335–355.

704 Burba, G. (2013). *Eddy Covariance Method for Scientific, Industrial, Agricultural and*
705 *Regulatory Applications*. Lincoln, Nebraska: LI-COR Biosciences. pp. 345

706 Burrus, C.S., Gopinath, A.R., & Guo, H. (1998). *Introduction to Wavelets and Wavelet*
707 *Transforms: A Primer*. Upper Saddle River, NJ: Prentice Hall. Print

708 Falge, E., Baldocchi, D., Olson, R., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G.,
709 Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T.,
710 Hollinger, D., Jensen, N. O., Katul, G., Keronen, P., Kowalski, A., Lai, C. T.,
711 Law, B. E., Meyers, T., Moncrieff, J., Moors, E., Munger, J. W., Pilegaard, K.,
712 Rannik, Ü., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T.,
713 Wilson, K., & Wofsy, S. (2001). Gap filling strategies for defensible annual sums
714 of net ecosystem exchange. *Agricultural and Forest Meteorology*, 107(1), 43–69.

715 Falge, E., Baldocchi, D., Tenhunen, J., Aubinet, M., Bakwin, P., Berbigier, P., Bernhofer,
716 C., Burba, G., Clement, R., Davis, K. J., Elbers, J., Goldstein, A. H., Grelle, A.,
717 Granier, A., Gumundsson, J., Hollinger, D., Kowalski, A. S., Katul, G., Law, B.
718 E., Malhi, Y., Meyers, T., Monson, R. K., Munger, J. W., Oechel, W., Paw U, K.
719 T., Pilegaard, K., Rannik, Ü., Rebmann, C., Suyker, A., Valentini, R., Wilson, K.,
720 & Wofsy, S. (2002). Seasonality of ecosystem respiration and gross primary
721 production as derived from FLUXNET measurements. *Agricultural and Forest*
722 *Meteorology*, 113(1-4), 53–74.

723 Foken, T., and Leclerc, M. Y. (2004). Methods and limitations in validation of footprint
724 models. *Agricultural and Forest Meteorology*, 127, 223-234. doi:
725 10.1016/j.agrformet.2004.07.015.

726 Goulden, M. L., Munger, J. W., Fan, S. M., Daube, B. C., & Wofsy, S. C. (1996).
 727 Exchange of carbon dioxide by a deciduous forest: Response to interannual
 728 climate variability. *Science*, 271, 1576–1578.

729 Gove, J. H., and Hollinger, D. Y. (2006). Application of a dual unscented Kalman filter
 730 for simultaneous state and parameter estimation in problems of surface-
 731 atmosphere exchange. *Journal of Geophysical Research: Atmospheres*, 111(8),
 732 D08S07. doi:10.1029/2005JD006021.

733 Gu, L., Falge, E. M., Boden, T., Baldocchi, D. D., Black, T. A., Saleska, S. R., Suni, T.,
 734 Verma, S. B., Vesala, T., Wofsy, S. C., & Xu, L. (2005). Objective threshold
 735 determination for nighttime eddy flux filtering. *Agricultural and Forest*
 736 *Meteorology*, 128, 179–197. doi:10.1016/j.agrformet.2004.11.006.

737 Hagen, S. C., Braswell, B. H., Linder, E., Frohling, S., Richardson, A. D., & Hollinger,
 738 D. Y. (2006). Statistical uncertainty of eddy-flux based estimates of gross
 739 ecosystem carbon exchange at Howland Forest, Maine. *Journal of Geophysical*.
 740 *Research*. (111): D08S03, doi: 10.1029/2005JD006154.

741 Hollinger, D. Y., Aber, J., Dail, B., Davidson, E., Goltz, S. M., Hughes, H., Leclerc, M.
 742 Y., Lee, J. T., Richardson, D., Rodrigues, C., Scott, N., Achuatavarier, D., &
 743 Walsh, J. (2004). Spatial and temporal variability in forest-atmosphere CO₂
 744 exchange. *Global Change Biology*, 10(10), 1689–1706. doi:10.1111/j.1365-
 745 2486.2004.00847.

746 Hollinger, D. Y., and Richardson, A. D. (2005). Uncertainty in eddy covariance
 747 measurements and its application to physiological models. *Tree Physiology*. 25(7)
 748 873–885. doi: 10.1093/treephys/25.7.873.

749 International Standardization Organization (ISO) (1995). *Guide to the Expression of Uncertainty*
750 *in Measurement*. International Organization for Standardization. pp 101

751 Jimenez, K. L., Starr, G., Staudhammer, C. L., Schedlbauer, J. L., Loescher, H. W.,
752 Malone, S. L., & Oberbauer, S. F. (2012). Carbon dioxide exchange rates from
753 short- and long-hydroperiod Everglades freshwater marsh. *Journal of*
754 *Geophysical Research: Biogeosciences*, 117(G4), 1–17.
755 doi:10.1029/2012JG002117.

756 Joint Committee for Guides in Metrology (JCGM) (2008). *Evaluation of measurement*
757 *data – Guide to the expression of uncertainty in measurement*. International
758 Bureau of Weights and Measures. pp. 120

759 Katul, G. and Parlange, M. B. (1995). Analysis of land surface heat fluxes using the
760 orthonormal wavelet approach. *Water Resources Research*. 31(11), 2743-2749

761 Katul, G., Lai, C. T., Schäfer, K., Vidakovic, B., Albertson, J., Ellsworth, D., & Oren, R.
762 (2001). Multiscale analysis of vegetation surface fluxes: from seconds to
763 years. *Advances in Water Resources*, 24(9), 1119-1132.

764 Loescher, H. W., Oberbauer, S. F., Gholz, H. L., & Clark, D. B. (2003). Environmental
765 controls on net ecosystem-level carbon exchange and productivity in a Central
766 American tropical wet forest. *Global Change Biology*, 9(3), 396–412.
767 doi:10.1046/j.1365-2486.2003.00599.

768 Loescher, H. W., Law, B. E., Mahrt, L., Hollinger, D. Y., Campbell, J., & Wofsy, S. C.
769 (2006a). Uncertainties in, and interpretation of, carbon flux estimates using the
770 eddy covariance technique. *Journal of Geophysical Research: Atmospheres*,
771 111(D11), 1–19. doi:10.1029/2005JD006932.

772 Loescher, H. W., Starr, G., Martin, T. A., Binford, M., & Gholz, H. L. (2006b). The
 773 effect of local atmospheric circulations on daytime carbon dioxide flux
 774 measurements over a *Pinus elliottii* canopy. *Journal of Applied Meteorology and*
 775 *Climatology*, 45, 1127-1140.

776 Lloyd, J., and Taylor, J. A. (1994). On the temperature dependence of soil respiration.
 777 *Functional Ecology*. 315–323.

778 Massman, W. J. (2000). A simple method for estimating frequency response corrections
 779 for eddy covariance systems. *Agricultural and Forest Meteorology* 104:185–198

780 Massman, W. J. (2004). Toward an ozone standard to protect vegetation based on
 781 effective dose: a review of deposition resistances and a possible metric.
 782 *Atmospheric Environment*. 38(15):2323–2337.

783 Michaelis, L., and Menten, M.L., 1913. Die Kinetik der Invertinwirkung. *Biochemische*
 784 *Zeitschrift* 49, 333-369.

785 Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A.
 786 G., Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., Falge, E., Gove, J.
 787 H., Heimann, M., Hui, D., Jarvis, A. J., Kattge, J., Noormets, A., & Stauch, V. J.
 788 (2007). Comprehensive comparison of gap-filling techniques for eddy covariance
 789 net carbon fluxes. *Agricultural and Forest Meteorology*, 147(3), 209–232.
 790 doi:10.1016/j.agrformet.2007.08.011.

791 Monteith, J., and Unsworth, M. (2007). *Principles of Environmental Physics*. Academic
 792 Press. pp 440.

793 O'Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation
 794 Factors. *Quality & Quantity*, 41, 673–690. doi:10.1007/s11135-006-9018-6

795 Ocheltree, T. O., & Loescher, H. W. (2007). Design of the AmeriFlux portable eddy-covariance
 796 system and uncertainty analysis of carbon measurements. *J. Atmos. Ocean. Tech.*, 24,
 797 1389-1409.

798 Pingintha, N., Leclerc, M. Y., Beasley Jr., J. P., Durden, D., Zhang, G., Senthong, C., &
 799 Rowland, D. (2010). Hysteresis response of daytime net ecosystem exchange
 800 during drought. *Biogeosciences*, 7(3), 1159–1170.

801 Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W.,
 802 Longdoz, B., Rambal, S., Valentini, R., Vesala, T., & Yakir, D. (2006). Towards a
 803 standardized processing of Net Ecosystem Exchange measured with eddy
 804 covariance technique: algorithms and uncertainty estimation. *Biogeosciences*,
 805 3(4), 571–583.

806 Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P.,
 807 Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., Grünwald, T.,
 808 Havrankova, K., Ilvesniemi, H., Janous, D., Knohl, A., Laurila, T., Lohila, A.,
 809 Loustau, D., Matteucci, G., Meyers, T., Miglietta, F., Ourcival, J. M., Pumpanen,
 810 J., Rambal, S., Rotenberg, E., Sanz, M., Tenhunen, J., Seufert, G., Vaccari, F.,
 811 Vesala, T., Yakir, D., & Valentini, R., 2005. On the separation of net ecosystem
 812 exchange into assimilation and ecosystem respiration: review and improved
 813 algorithm. *Global Change Biology*. 11(9), 1424– 1439.

814 Richardson, A. D., Braswell, B. H., Hollinger, D.Y., Burman, P., Davidson, E.A., Evans,
 815 R.S., Flanagan, L. B., Munger, J. W., Savage, K., Urbanski, S. P., & Wofsy, S. C.,
 816 2006. Comparing simple respiration models for eddy flux and dynamic chamber
 817 data. *Agricultural and Forest Meteorology*, 141(2), 219–234.

818 Schimel, D. (2000). Contribution of increasing CO₂ and climate to carbon storage by
819 ecosystems in the United States. *Science*, 287(5460), 2004–2006.

820 Schimel, D. S., House, J. I., Hibbard, K. a, Bousquet, P., Ciais, P., Peylin, P., & Wirth, C.
821 (2001). Recent patterns and mechanisms of carbon exchange by terrestrial
822 ecosystems. *Nature*, 414(6860), 169–172. doi:10.1038/35102500.

823 Sierra, C. A., Harmon, M. E., Thomann, E., Perakis, S. S., & Loescher, H. W. (2011).
824 Amplification and dampening of soil respiration by changes in temperature
825 variability. *Biogeosciences*, 8(4), 951–961. doi:10.5194/bg-8-951-2011.

826 Smith, M.D., Knapp, A.K., & Collins, S.L. (2009). A framework for assessing ecosystem
827 dynamics in response to chronic resource alterations induced by global change. *Ecology*
828 30:3279-3289.

829 Starr, G., Staudhammer, C. L., Loescher, H. W., Mitchell, R., Whelan, A., Hiers, J. K., &
830 O'Brien, J. J. (2015). Time series analysis of forest carbon dynamics: recovery of
831 *Pinus palustris* physiology following a prescribed fire. *New Forests*, 46(1), 63-90
832 doi:10.1007/s11056-014-9447-3.

833 Starr, G., Staudhammer, C. L., Wiesner, S., Kunwor, S., Loescher, H. W., Baron, A. F.,
834 Whelan, A., Mitchell, R., Boring, L. (2016). Carbon dynamics of *Pinus palustris*
835 ecosystems following drought. *Forests*, 7(5) doi:10.3390/f7050098

836 Stoy, P. C., Katul, G. G., Siqueira, M. B. S., Juang, J. Y., McCarthy, H. R., Kim, H. S.,
837 Oishi, C., & Oren, R. (2005). Variability in net ecosystem exchange from hourly
838 to inter-annual time scales at adjacent pine and hardwood forests: a wavelet
839 analysis. *Tree Physiology*, 25(7), 887–902. doi:10.1093/treephys/25.7.887.

840 Stoy, P. C., Katul, G. G., Siqueira, M. B. S., Juang, J. Y., Novick, K. A., Uebelherr, J.
841 M., & Oren, R. (2006). An evaluation of models for partitioning eddy covariance-
842 measured net ecosystem exchange into photosynthesis and respiration.
843 *Agricultural and Forest Meteorology*, 141(1), 2-18.
844 doi:10.1016/j.agrformet.2006.09.001.

845 Stoy, P. C., Richardson, A. D., Baldocchi, D. D., Katul, G. G., Stanovick, J., Mahecha,
846 M. D., Reichstein, M., Detto, M., Law, B. E., Wohlfahrt, G., Arriga, N., Campos,
847 J., McCaughey, J. H., Montagnani, L., Paw U, K. T., Sevanto, S., & Williams, M.
848 (2009). Biosphere-atmosphere exchange of CO₂ in relation to climate: a cross-
849 biome analysis across multiple time scales. *Biogeosciences Discussions*, 6(10),
850 2297-2312.

851 Thomas, C., Foken, T. (2007) Flux contribution of coherent structures and its
852 implications for the exchange of energy and matter in a tall spruce canopy.
853 *Boundary-Layer Meteorology*, 123, 317-337.

854 Torrence, C., and Compo, G. P. (1998). A Practical Guide to Wavelet Analysis. *Bulletin*
855 *of the American Meteorological Society*, 79(1), 61–78. doi:10.1175/1520-
856 0477(1998).

857 van Gorsel, E., Leuning, R., Cleugh, H. A., Keith, H., Kirschbaum, M. U. F., & Suni, T.
858 (2008). Application of an alternative method to derive reliable estimates of
859 nighttime respiration from eddy covariance measurements in moderately complex
860 topography. *Agricultural and Forest Meteorology*, 148(6), 1174–1180.
861 doi:10.1016/j.agrformet.2008.01.015.

862 van Gorsel, E., Delpierre, N., Leuning, R., Black, A., Munger, J. W., Wofsy, S., Aubinet,
 863 M., Feigenwinter, C., Beringer, J., Bonal, D., Chen, B., Chen, J., Clement, R.,
 864 Davis, J. K., Desai, A. R., Dragoni, D., Etzold, S., Grunwald, T., Gu, L.,
 865 Heinesch, B., Hutya, L. R., Jans, W.W.P., Kutsch, W., Law, B. E., Leclerc, M.
 866 Y., Mammarella, I., Montagnani, L., Noorments, A., Rebmann, C., & Wharton, S.
 867 (2009). Estimating nocturnal ecosystem respiration from the vertical turbulent
 868 flux and change in storage of CO₂. *Agricultural and Forest Meteorology*,
 869 149(11), 1919–1930. doi:10.1016/j.agrformet.2009.06.020.
 870 Wang, H.J., Riley, W.J., & Collins, W.D. (2015) Statistical uncertainty of eddy
 871 covariance CO₂ fluxes inferred using a residual bootstrap approach. *Agricultural*
 872 *and Forest Meteorology* 206 (2015) 163–171.
 873 Whelan, A., Mitchell, R., Staudhammer, C. L., & Starr, G. (2013). Cyclic Occurrence of
 874 Fire and Its Role in Carbon Dynamics along an Edaphic Moisture Gradient in
 875 Longleaf Pine Ecosystems. *PLoS ONE*, 8(1), e54045.
 876 doi:10.1371/journal.pone.0054045.
 877 Whelan, A., Starr, G., Staudhammer, C. L., Loescher, H. W., & Mitchell, R. J. (2015). Effects of
 878 drought and prescribed fire on energy exchange in longleaf pine ecosystems. *Ecosphere*,
 879 6(7), 128. <http://dx.doi.org/10.1890/ES15-00111.1>.