

Research Article

A Comparison of Three Gap Filling Techniques for Eddy Covariance Net Carbon Fluxes in Short Vegetation Ecosystems

Xiaosong Zhao¹ and Yao Huang²

¹Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China
²State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences

²State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

Correspondence should be addressed to Yao Huang; huangyao@ibcas.ac.cn

Received 9 March 2015; Accepted 17 May 2015

Academic Editor: Marcos Heil Costa

Copyright © 2015 X. Zhao and Y. Huang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Missing data is an inevitable problem when measuring CO_2 , water, and energy fluxes between biosphere and atmosphere by eddy covariance systems. To find the optimum gap-filling method for short vegetations, we review three-methods mean diurnal variation (MDV), look-up tables (LUT), and nonlinear regression (NLR) for estimating missing values of net ecosystem CO_2 exchange (NEE) in eddy covariance time series and evaluate their performance for different artificial gap scenarios based on benchmark datasets from marsh and cropland sites in China. The cumulative errors for three methods have no consistent bias trends, which ranged between -30 and $+30 \text{ mgCO}_2 \text{ m}^{-2}$ from May to October at three sites. To reduce sum bias in maximum, combined gap-filling methods were selected for short vegetation. The NLR or LUT method was selected after plant rapidly increasing in spring and before the end of plant growing, and MDV method was used to the other stage. The sum relative error (SRE) of optimum method ranged between -2 and +4% for four-gap level at three sites, except for 55% gaps at soybean site, which also obviously reduced standard deviation of error.

1. Introduction

Eddy covariance technique to measure CO₂, water, and energy fluxes between biosphere and atmosphere is widely spread and used in various regional networks [1]. At present, over 600 tower sites are operating on a long-term and continuous basis around the world, covering different climate conditions and land use and land cover changes, some of them running continuously for more than 10 years (http://fluxnet.ornl.gov/). However, missing or rejected data in these measurements is a unavoidable problem due to equipment failures (system/sensor breakdown), maintenance and calibration, spikes in the raw data, and physical and biological constraints (e.g., storms, hurricanes, and nonoptimal wind directions) [2]. In general, about 17-50% of the observations in net ecosystem CO₂ exchange (NEE) are reported as missing or rejected at FluxNet sites [3]. The gaps in observed data cause at least three problems: (1) difficulty in

annual estimation of NEE, (2) biased relationships between NEE with climatic variables, and (3) low quality data for modeling validation [2].

To accurately calculate annual values of NEE at sites, gap-filling to account for the missing data is imperative. The commonly used methods for filling missing data include mean diurnal variation (MDV) [3], look-up table (LUT) [3], nonlinear regression (NLR) [3–5], marginal distribution sampling [6], multiple imputation model [7], artificial neural network [8–11], and terrestrial biosphere model [12]. This diversity hinders synthesis activities because the biases and uncertainties associated with each technique are unknown [13, 14].

In a comprehensive study, Falge et al. [3] compared three methods including MDV, LUT, and NLR on the annual sum of NEE for 28 datasets from 18 FluxNet sites and found that the differences in annual NEE estimation by different gap-filling methods ranged from -45 to 200 gCm^{-2} per year.

Their study also emphasized the importance of the method of standardization during the data postprocessing phase, so comparable data can be obtained to address intercomparisons across different ecosystems, climatic conditions, and multiple years. Richardson and Hollinger [15] quantified the uncertainties in annual NEE with a simple model using data assimilation techniques that are due both to random measurement error and to gap filling, including the additional uncertainty that can be attributed to long gaps and the relationship between gap length and uncertainty in NEE. The CO₂ flux data come from a coniferous, two deciduous, two mixed species, and two mediterranean sites. Moffat et al. [9] reviewed 15 techniques for estimating missing values of NEE in eddy covariance time series and evaluate their performance for different artificial gap scenarios based on a set of 10 benchmark datasets from six forested sites in Europe which is the same with Richardson and Hollinger [15]. Papale et al. [2] introduced a new standardized set of corrections and assessed the uncertainties associated with these corrections for eight different forest sites in Europe with a total of 12 yearly datasets.

However, most comparison works about gap-filling methods were objected to tall vegetation, that is, forests. Less research focused on short vegetation, that is, croplands or marshlands. The vegetation structure for short vegetation changes more rapidly in the growing season, which may affect ability of gap-filling methods. So, it is important to evaluate the performance of gap-filling methods and search the optimum methods for short vegetation.

In this study, we reviewed three methods (MDV, LUT, and NLR) and applied the techniques to a set of benchmark datasets from marshland and croplands (rice and soybean) in China. Artificial gaps were added to observed NEE time series based on Falge et al. [3], and the ability of different gapfilling techniques to replicate the missing data was evaluated using statistical analysis. The objective of this paper is to find the optimum method for short vegetation.

2. Methods

2.1. Data Basis. For this analysis, we used half-hourly eddy flux measurements of the net ecosystem exchange of CO_2 from three different ecosystem types. As case studies, we chose CO_2 flux data from May to October in 2005 from marshland and agriculture (rice and soybean cropland) sites in the Sanjiang Plain. The marshland site locates at $(47^{\circ}35'N,$ $133^{\circ}31'E)$, the field areas are approximately 105 ha. The rice and soybean site locate approximately 1.5 km west and 500 m north to the marshland, respectively. The field areas are approximately thousands of hectares for rice site and 25 ha for soybean site. The altitude is 55.4–57.9 m. The more detailed information is available in Zhao et al. [16].

The EC system consisted of a triaxial sonic anemometer (CAST3, Campbell Scientific, USA) and a fast response open-path CO_2/H_2O infrared gas analyzer (Li-7500, LiCor Inc., USA). The meteorological parameters including air humidity and air temperature, wind speed, precipitation, soil temperature, and water content were measured [16]. Raw data acquired at 10 Hz were processed using the postprocessing,

TABLE 1: Percentages of original gaps in NEE measurements.

Elux oito	Daytime		Nighttime		All time	
TTUX SILC	N	Gap (<i>P</i> %)	N	Gap (<i>P</i> %)	N	Gap (<i>P</i> %)
Marsh site	4583	18.4%	2953	36.2%	7536	25.4%
Rice site	4804	12.4%	3212	26.7%	8016	18.2%
Soybean site	4804	15.0%	3212	28.7%	8016	20.5%

The N is the total number of data for daytime, nighttime, and all the time. The P% is the percentage of gaps for daytime, nighttime, or all the time.

TABLE 2: The number and percentage of friction velocity (u^*) filtered data in artificial datasets.

Elux cito	u^* correction						
Flux site	35% gap	45% gap	55% gap	65% gap			
Marsh site		3.3%	6.6%	10.7%			
Rice site	2.9%	7.0%	10.8%	13.5%			
Soybean site	2.1%	5.6%	8.6%	13.5%			

including spike removal, frequency response correction [17], sonic virtual temperature correction [18], the performance of the planar fit coordinate rotation [19], and corrections for density fluctuation (WPL correction) [20].

The quality control of the half-hourly flux data was carried out as follows: (i) data from periods of sensor malfunction were rejected (e.g., when there was a faulty diagnostic signal), (ii) data within 1 h before or after precipitation were rejected, (iii) incomplete 30 min data were rejected when the missing data constituted more than 3% of the 30 min raw record, and (iv) data were rejected when the value was larger than mean \pm 3 standard deviation. The information of original gaps in NEE measurements is showed in Table 1. The gap percentages in all time were 25.4%, 18.2%, and 20.5% at the marsh, rice, and soybean site, respectively. Gap percentages at nighttime (ranging from 26.7% to 36.2%) were slightly higher than at daytime (ranging from 12.4% to 18.4%) (Table 1).

For this comparison, four artificial datasets were created, containing 35%, 45%, 55%, and 65% of gaps [3]. Based on random function RAND, sets of data with random distribution were generated. The random dataset corresponded to the dataset of NEE measurements except for original gaps. According to the difference from the number of artificial gaps to original gap, a range of certain numbers were selected from a set of random data, then a new gap was generated by deleting corresponding the NEE dataset. Starting from the original gap percentage, artificial gaps were created separately for daytime and nighttime, until the dataset contained a given percentage of gaps at both daytime and nighttime [3]. To avoid underestimation of CO2 flux during calm conditions at night, the friction velocity (u^*) was applied at nighttime [21, 22]. The data were rejected at night when the u^* was below 0.10 ms^{-1} . Due to the percentage of u^* filtered data was about 10% at three sites, which caused high percentage of original gaps at nighttime. Therefore, the u^* correction was applied to artificial data instead of original data. The percentage of u^* filtered data ranged from 2.1% to 13.5% in different percentages of artificial datasets (Table 2).



FIGURE 1: Seasonal variation of the parameters in Michaelis-Menten functions at (a) marsh, (b) rice, and (c) soybean sites.

After introducing artificial gaps for each of the four datasets, the respective gap-filling methods were parameterized with the remaining data and applied to fill the artificial datasets. The gap-filling error was calculated using the observed fluxes in these artificial gaps to validate the predictions of each filling technique.

2.2. Filling Methods. Three gap-filling methods were applied here, including mean diurnal variation (MDV), look-up tables (LUT), and nonlinear regression (NLR) methods.

2.2.1. Mean Diurnal Variation. MDV is an interpolation technique where the missing NEE value for a certain time period (half-hour) is replaced with the averaged value of the adjacent days at exactly that time of day. Data windows of 7 days during daytime and 14 days during nighttime were chosen for averaging in the application.

2.2.2. Look-Up Tables. In a look-up table, the NEE data are binned by variables such as light and temperature presenting similar meteorological conditions, so that a missing NEE value with similar meteorological conditions can be "looked up" [3]. Tables were created to represent changing environmental conditions based on monthly period, using the photosynthetic photon flux density- (PPFD-) air temperature-(Ta-) sort during day, and the relative humidity- (RH-) Tasort during night. For look-up tables the average NEE was compiled for six monthly periods * 11 PPFD-class * 36 Ta-classes. The PPFD-classes consisted of $200 \,\mu \text{mol m}^{-2} \,\text{s}^{-1}$ intervals from 0 to $2000 \,\mu \text{mol m}^{-2} \,\text{s}^{-1}$. Similarly, Ta-classes were defined through 1°C intervals ranging from -5° C to 31°C. For night day, average NEE was compiled for six monthly periods * 8 RH-classes * 19 Ta-classes. RH-classes range from 20% to 100% with 10% intervals, and Ta-classes were the same as the daytime.

2.2.3. Nonlinear Regression Methods. The nonlinear regressions are based on parameterized nonlinear equations which express (semi-)empirical relationships between the CO_2 flux and environmental variables such as temperature and light.

For filling daytime gaps, the light response function of *Michaelis–Menten* [3, 23] was selected as follows:

NEE = Re -
$$\frac{\alpha \cdot PPFD \cdot F_{GPP,sat}}{F_{GPP,sat} + \alpha \cdot PPFD}$$
, (1)

where NEE is the net ecosystem exchange (mgCO₂ m⁻² s⁻¹) and Re is the ecosystem respiration rate (mgCO₂ m⁻² s⁻¹) during the day. PPFD is the photosynthetic photon flux density (μ mol m⁻² s⁻¹), and α is the ecosystem quantum yield (mgCO₂ μ mol⁻¹ quantum). $F_{GPP,sat}$ is the gross primary productivity at "saturating" light (mgCO₂ m⁻² s⁻¹). The light response function was fitted with window sizes of 15 days from June to middle September, and the seasonal variation of parameters was showed in Figure 1. The parameters in Figure 1 were calculated according to original NEE datasets before artificial gap introduction. The $F_{\text{GPP,sat}}$ at the rice and soybean sites ranged from 0.17 to 2.0 mgCO₂ m⁻² s⁻¹, which were larger than that at the marsh site (from 0.04 to 0.7 mgCO₂ m⁻² s⁻¹). The Re ranged from 0.05 to 0.37 mgCO₂ m⁻² s⁻¹ at all three sites, which were lower than the $F_{\text{GPP,sat}}$. The correlation coefficients (R^2) of observed and simulation value during this period were 0.69, 0.83, and 0.81 (P < 0.01) at the marsh, rice, and soybean sites, respectively. For each artificial dataset, the parameters in *Michaelis–Menten* function were recalculated and applied to artificial gaps.

The net ecosystem CO_2 exchange (NEE) at nighttime represents the ecosystem respiration (R_{eco}) because of no photosynthesis. The ecosystem respiration (R_{eco}) is conceptualized to consist of soil respiration, R_s , and aboveground component attributed to the respiration by various plant components, R_p . For nighttime NEE, the temperature response function was selected based on Wohlfahrt et al. [24]:

$$\begin{split} R_{\rm eco} &= R_s + R_p \\ &= R_{s,T_{\rm ref}} \cdot \exp\left[\frac{E_s}{R \cdot T_{\rm ref}} \left(1 - \frac{T_{\rm ref}}{T_a}\right)\right] + L \cdot R_{p,T_{\rm ref}} \quad (2) \\ &\quad \cdot \exp\left[\frac{E_p}{R \cdot T_{\rm ref}} \left(1 - \frac{T_{\rm ref}}{T_a}\right)\right], \end{split}$$

where R_{eco} is ecosystem respiration at nighttime $(\text{mgCO}_2 \text{ m}^{-2} \text{ s}^{-1})$, which includes soil respiration (R_s) and plant respiration (R_p) . $R_{x,T_{ref}}$ is the respiration rate $(\text{mgCO}_2 \text{ m}^{-2} \text{ s}^{-1})$ at a reference temperature (T_{ref}) , the reference temperature is 10°C, and E_x denotes an activation energy $(J \text{ mol}^{-1})$, the subscript of x as s represent to the soil components, and the subscript of x as p represent to the plant components. R is the universal gas constant, 8.314 J mol⁻¹ K⁻¹. T_a is air temperature (°C), and L denotes leaf area index. The respiration function was fitted for the whole growing season, whose parameters were showed in Table 3. The parameters in Table 3 were calculated according to original NEE datasets before artificial gap introduction. For each artificial dataset, the parameters in temperature response function were recalculated and applied to artificial gaps.

2.3. Error Assessment. To assess the applicability of a standard data filling method at three sites, we examined the potential bias error associated with each method. The bias errors for different methods were calculated as the observed value minus the predicted value for each gap level. For daytime carbon uptake, a positive error therefore indicates an over-estimation and a negative error in underestimation by the respective method.

The statistical sums were calculated using the individual observed NEE data o_i and the predicted value p_i , mean bias

error (MBE), mean absolute error (MAE), and sum relative error (SRE) were as follows:

$$MBE = \frac{1}{N} \sum (o_i - p_i),$$

$$MAE = \frac{1}{N} \sum |o_i - p_i|,$$

$$SRE = \frac{\sum o_i - \sum p_i}{\sum o_i}.$$
(3)

3. Results

3.1. Frequency of Gaps. The gap distribution for benchmark sets showed the number of gaps decreased with gap length (Figure 2). However, the majority of 35% artificial gaps consisted of short gaps (less than 10 half-hours), and very short gaps (less than 2 half-hours) were more than other 3 benchmark sets. Though gap numbers of long gaps (more than 20 half-hours) for 65% gaps of artificial data sets were similar to benchmark sets, gap numbers of short and medium gaps were higher than benchmark sets.

3.2. Error Analysis in Half-Hourly Scale. The most frequent distribution of error for gap-filling methods in half-hourly scale was nearly normal distribution (Figure 3), which indicated an unbiased estimate for gap-filling error. The mean and standard deviation of bias error for gap-filling methods were showed in three sites (Table 4). The count for nighttime data was less than daytime data because of u^* -correction. They were not consistent with negative or positive error for three methods or four gap percentage levels. No trends were found that MBE was larger than gap percent levels during day and nighttime at three sites; moreover, the error for high gap level was small inversely, that is, error of 65% gaps with MDV and LUT methods at marsh site at daytime and error of 65% gaps with three methods at rice site at nighttime. For daytime data, standard deviation for LUT method was the largest and for NLR method was the lowest among three methods for each gap level. For nighttime data, standard deviation for MDV method was the largest, especially at soybean site.

3.3. Seasonal Variation of Error. The seasonal patterns of daily MBE, taking 65% gaps level as an example, showed the difference in each site (Figure 4). These patterns were affected by different methods and stage of growth. In general, all methods have good performance before germination or sawing stage and after entirely wilting or harvesting; in these stages, daily MBE for each method was around zero and MDV method with less fluctuation. The daily MBE was large in the peak of growing season because of strong assimilated CO₂ ability, and spikes often occurred for MDV and LUT methods in this stage at three sites. This result was agreed with large standard deviation for MDV and LUT in Table 4. The significant difference among the methods in fast growth stage of spring (LAI was rapid increase), that is, late May at marshland site, early June at rice site, and middle June at soybean site (Figure 4). However, MDV method has good performance in this stage.



FIGURE 2: Frequency of gap lengths at (a) marsh, (b) rice, and (c) soybean sites. To show clearly, only frequencies of 35% and 65% gaps were plotted. The frequencies of 45% and 55% gaps were distributed between them.

TABLE 3: The parameters in temperature response function for nighttime data.

Sites	$\frac{R_{s,T_{\rm ref}}^{(a)}}{{\rm mgCO}_2{\rm m}^{-2}{\rm s}^{-1}}$	E_s KJ mol ⁻¹	$R_{p,T_{\rm ref}}^{(b)}$ mgCO ₂ m ⁻² s ⁻¹	E_p KJ mol ⁻¹	R^2	Р
Marsh site	0.035	55.3	0.022	2.8	0.54	< 0.01
Rice site	0.033	85.2	0.0013	79.8	0.62	< 0.01
Soybean site	0.068	52.8	0.0026	148.4	0.47	< 0.01

 $\overline{(a)}R_{s,T_{ref}}$: soil respiration at reference temperature of 10°C.

^(b) $R_{p,T_{ref}}$; plant respiration at reference temperature of 10°C and unit leaf area.



FIGURE 3: Frequency distribution of bias error with different methods for 65% artificial gaps at three sites, separated into daytime (a) and nighttime (b) data.

The variation of cumulative error was shown in Figure 5, taking 65% gaps level as an example. The cumulative error at daytime has stronger fluctuation than at nighttime at three sites, as the results of a little data and small error at nighttime. The cumulative error for three methods has no consistent bias at marsh site (Figure 5). However, positive bias errors were observed for three methods during the day at rice site, and negative bias errors were observed at soybean site. This suggests different methods may cause complicated effects at three sites. The huge bias error for LUT method was observed at each site, especially, in spring of rapid growth stage. The large bias error for NLR method was also observed after August at soybean site, whereas this phenomenon has not

occurred in 35% and 45% gaps. Based on cumulative error from May to October, it indicates that MDV method has good performance, especially smooth trend in the end of growing season at three sites. Overall, the cumulative error at three sites ranged between -30 and $+30 \text{ mgCO}_2 \text{ m}^{-2}$.

The SRE was showed in Table 5, and it is convenient to evaluate the performance of gap-filling methods and compare it with other sites. In general, the SRE for 35% and 45% artificial gaps filled by three methods was smaller than 55% and 65% gaps at daytime, while these patterns were not marked at nighttime. The gap-filling methods have distinct different performances at three sites; for example, MDV method showed small SRE at daytime over the rice site, while



FIGURE 4: Seasonal variations of daily mean bias error (MBE) with different methods at (a) marsh, (b) rice, and (c) soybean sites.

				Day time		Night time (u^* corrected)			
Flux site	Gap %	, (b)	MDV	LUT	NLR	1 (b)	MDV	LUT	NLR
		n		$mgCO_2 m^{-2} s^{-1}$		71	$mgCO_2 m^{-2} s^{-1}$		
	35	763	-2.51 (79.04)	-0.75 (88.05)	-4.36 (73.46)	na ^(c)	na	na	na
Marsh site	45	1220	3.42 (73.26)	-0.14 (78.84)	0.66 (69.65)	161	2.21 (47.18)	1.19 (34.85)	0.73 (35.65)
	55	1677	4.27 (78.65)	3.09 (90.18)	2.00 (78.88)	359	-1.26 (47.56)	2.19 (35.97)	2.16 (36.46)
	65	2136	1.66 (84.39)	-0.78 (90.66)	-3.32 (74.42)	532	4.14 (51.87)	4.95 (41.49)	5.49 (41.24)
3 Rice site 5 6	35	1084	-0.10 (144.58)	-4.59 (154.58)	-7.17 (123.62)	175	3.89 (53.96)	2.52 (47.19)	2.16 (46.24)
	45	1565	-0.17 (152.18)	1.79 (165.22)	3.5 (124.08)	366	3.56 (54.27)	-2.04 (48.94)	1.60 (42.13)
	55	2045	-1.14 (145.34)	-2.69 (168.33)	-5.33 (126.43)	564	-5.22 (58.38)	-1.27 (48.69)	0.55 (40.58)
	65	2525	2.76 (149.03)	9.68 (160.66)	11.25 (123.96)	797	-0.59 (57.99)	1.11 (48.94)	0.06 (41.25)
	35	964	-2.17 (115.52)	1.64 (134.87)	-6.9 (114.04)	136	-21.8 (93.51)	-9.42 (44.1)	-8.8 (33.09)
Soybean site	45	1445	-1.88 (114.98)	1.81 (136.21)	-2.27 (113.32)	345	1.91 (74.24)	4.96 (46.71)	3.57 (44.11)
	55	1929	-5.21 (112.73)	-3.2 (132.46)	-9.41 (110.09)	569	1.72 (69.91)	0.19 (45.55)	1.19 (41.73)
	65	2411	-1.7 (112.09)	-2.05 (134.55)	-10.97 (109.28)	732	-3.26 (74.96)	0.30 (47.45)	-2.59 (41.61)

TABLE 4: Mean bias error of half-hourly F_{NEE} and standard deviation from the mean^(a).

^(a)Values in parentheses are standard deviation.

^(b)Number of the artificial gaps.

^(c)No measurements were available due to large gaps.



FIGURE 5: Cumulative error in gap-filled NEE from May to October, separated in daytime, nighttime, and all time contribution, at marsh, rice, and soybean sites.

LUT and NLR methods represented well at daytime over the marsh site. The majority of SRE ranged from -10 to 10% during whole day, except for 55% and 65% artificial gaps filled by NLR method at soybean site, this caused by huge bias after August (showed in Figure 5).

3.4. Error Analysis in Gap Size Class. Gap size and distribution were produced in random, whereas those greatly impacted performance of gap filling methods. The colored surface plots are depicted in Figure 6, 65% gaps were taken as an example, which provide a visual means of qualitatively assessing the impact of gap length on NEE uncertainty. For short vegetation, small MAE was expressed in dormant season (early spring and late autumn), regardless of methods and gap length. This was related to the fact that measured fluxes at this stage tended to be smaller. Large errors for three methods were concentrated to the zone of gaps of less than 5 in growing season, especially in stage of growth rapidly of plants. Among the methods, LUT method resulted in the largest error ranging from short to long gaps and then MDV and NLR methods. Though the patterns did not find that MAE for all methods increased with gap length increasing, long gaps added appreciably to the uncertainty of gap-filling (results were not shown).

3.5. Optimum Gap Filling Method for Short Vegetation. Selection of methods was based on the most stable performance and smallest errors; however, according to the above analysis, no one method was perfect during the measuring stage. To reduce sum bias in maximum, combined gap-filling method



FIGURE 6: Two-dimensional plots illustrating mean absolute error (MAE) in gap-filled NEE, expressed as MAE, varied as a function of day of year (*x axis*) and gap length (*y axis*). Plots in row were shown for marsh, rice, and soybean sites, and plots in column were shown for MDV, LUT, and NLR methods. The different size and color of the circle refer to magnitude of MAE, and unit is $mgCO_2 m^{-2} s^{-1}$.

TABLE 5: The sum relative error (SRE) for NEE in different	gap levels, at marsh, rice, and soybean site
--	--

	Can %	Cap % Daytime			Nighttime (u^* corrected)			All time		
	Gap %	MDV %	LUT %	NLR %	MDV %	LUT %	NLR %	MDV %	LUT %	NLR %
	35	2.83	0.85	4.92	na ^(a)	na	na	2.83	0.85	4.92
March site	45	-3.61	0.15	-0.70	2.51	1.36	0.83	-4.46	-0.02	-0.92
Marsh site	55	-4.98	-3.60	-2.33	-1.46	2.54	2.49	-5.93	-5.23	-3.67
	65	-1.74	0.82	3.48	4.60	5.50	6.11	-3.67	-0.61	2.66
D: :/	35	0.05	3.37	2.28	4.38	2.84	2.43	-0.28	3.41	2.27
	45	0.08	-0.81	-1.59	3.97	-2.27	1.79	-0.33	-0.66	-1.94
Rice site	55	0.58	1.38	2.73	-5.67	-1.38	0.60	1.52	1.79	3.05
	65	-1.40	-4.89	-5.68	-0.69	1.29	0.08	-1.51	-5.86	-6.59
	35	2.21	-1.67	7.01	-22.99	-9.94	-9.28	5.16	-0.37	9.58
Soybean site	45	1.82	-1.75	2.20	1.57	4.09	2.94	1.92	-4.04	1.91
	55	5.80	3.56	10.48	1.47	0.16	1.02	8.49	5.67	16.36
	65	1.74	1.97	10.11	-2.88	0.27	-2.29	4.93	3.6	20.28

^(a)No measurements were available due to large gaps.

TABLE 6: Gap filling strategies at short	vegetation	ecosystems.
--	------------	-------------

		Method of gap filling	
Ecosystem	NLR	LUT	MDV
Marshland	Early Jun to mid Sept		Other time from May to Oct
Rice paddy	Mid Jun to mid Aug		Other time from May to Oct
Soybean		Early Jul to early Sept	Other time from May to Oct



FIGURE 7: The sum relative error (SRE) of optimum gap-filling method for four-gap level at three sites.

was selected for short vegetation. NLR or LUT method was used after plant rapidly growing in spring and after end of plant growth, and MDV method was used to the other stage. In this case, based on the growth stage of different vegetation, the gap filling strategies at three ecosystems were showed in Table 6. The SRE of optimum method reduced to the range of -2 and +4% for four-gap level at three sites (Figure 7), except for 55% gaps at soybean site. The optimum method also reduced standard deviation of error that was around 0.07, 0.11, and 0.12 mgCO₂ m⁻² s⁻¹ at marsh, rice, and soybean site, respectively; there were no significant different within four-gap level.

4. Discussion

4.1. The Response of Error on Environmental and Biological Factors. The performance of gap-filling methods impacted on climatic and biological variables such as PPFD and LAI [6, 23]. The LUT and NLR methods have considered the effect of PPFD, a residual error with NLR method distributed evenly around zero response of PPFD with small magnitude (Figure 8), while residual error with LUT method scattered strongly around zero. This was the reason of high standard deviation at daytime (Table 4). The residual error with MDV method caused positive bias when PPFD was less than 500 μ mol m⁻² s⁻¹.

The residual errors response of LAI with NLR method has even distribution around zero, and MDV and LUT method have more scatter; moreover, significant negative error occurred around LAI = 1 (Figure 8). The large scatters were showed for MDV method that did not consider LAI when filling gaps. Though the LUT method filled gaps per half month, there was weak relationship between LAI and NEE, especially when LAI = 1 (Figure 8). This result must be expected from potential changes in the ecosystem properties, particularly as related to canopy development and senescence [6, 25].

4.2. The Selection of Gap-Filling Methods for Short Vegetations. In this study for short vegetations, error introduced by gap-filling differed between methods at different gap levels (Table 4). The choice of a technique should be based on the application, Moffat et al. [9] considered NLR method can serve well for an annual sum estimate, but an artificial neural network will best reproduce the half-hourly profile of the flux. Falge et al. [3] also commented on semiempirical methods because they preserve the response of NEE to main meteorological conditions. However, the NLR method in our study has good performance in variation of daily NEE (Figure 4) and caused huge bias in cumulative NEE, especially for high gap level (Figure 5), which can explain that great uncertainty was introduced, because little data was available to simulating nonlinear function.

The MDV method had large error in half-hourly NEE (Figure 4) but consistent performance and reliability in sum NEE (Figure 5). For MDV, the method does not make use of the ancillary meteorological data and can be expected to have additional problems filling gaps of more than 3–7 days in length, as synoptic changes in weather are strongly linked to changes in diurnal cycles of photosynthesis and respiration [1, 9]. So, to reduce error in half-hourly and annual NEE, the combined method of MDV and NLR was selected in our study (Figure 7) and performed well for short vegetation.

The methods caused large bias during periods of active change in ecosystem properties (Figure 4), because when the flux data are missing, it is impossible to know the timing of magnitude of the change [3, 15]. The magnitudes of NEE for short vegetation, that is, marsh or cropland, and so forth, were smaller than forests; especially for soybean cropland, high GPP and high Re caused low NEE during the growing season. So little error may cause large bias of cumulative NEE, and underestimating NEE or overestimating Re may change carbon sink to carbon source. In this study, the optimum gapfilling method can resolve partly this problem.

5. Conclusion

The three major gap-filling methods (mean diurnal variation, look-up table and nonlinear regression) for estimating net

LUT MDV NLR 1.2 1.2 1.2 $Error\ (mgCO_2\ m^{-2}\ s^{-1})$ 0.6 0.6 0.6 0.0 0.0 0.0 -0.6-0.6-0.6 -1.2-1.2-1.20 500 1000 1500 2000 0 500 1000 1500 2000 0 500 1000 1500 2000 PPFD (μ mol m⁻² s⁻¹) PPFD (μ mol m⁻² s⁻¹) PPFD (μ mol m⁻² s⁻¹) 1.2 1.2 1.2 Error (mgCO₂ m⁻² s⁻¹) 0.6 0.6 0.6 0.0 0.0 0.0 -0.6 -0.6-0.6-1.2.2 0 2 4 6 0 2 4 6 0 2 4 6 LAI $(m^2 m^{-2})$ LAI $(m^2 m^{-2})$ LAI $(m^2 m^{-2})$

FIGURE 8: The relationship between bias error and PPFD and LAI in day time, taking 65% gaps at rice site as an example.

carbon fluxes (NEE) were reviewed and their gap-filling performance was evaluated based on a set of datasets from three short vegetations (marsh, rice, and soybean sites). The performance of the filling techniques depended on the time scale, gap length, and time of day (day or night). In halfhourly scale, standard deviation for NLR method was the smallest among three methods for each gap level. The MDV method has good performance in seasonal scale, especially before germination or sawing and after entirely wilting or harvesting. Though LUT and NLR methods showed small error for daily mean error during the peak of growing season, the huge bias was observed in cumulative NEE for two methods. The combined gap-filling methods were used for short vegetation, which showed NLR or LUT method was selected after plant rapidly increasing in spring and before end of plant growth and MDV method was used to the other stage. This combined method distinctly reduced sum bias and deviation for gap-filled NEE.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (no. 41471022).

References

- D. Baldocchi, E. Falge, L. Gu et al., "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*, vol. 82, no. 11, pp. 2415–2434, 2001.
- [2] D. Papale, M. Reichstein, M. Aubinet et al., "Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation," *Biogeosciences*, vol. 3, no. 4, pp. 571–583, 2006.
- [3] E. Falge, D. Baldocchi, R. Olson et al., "Gap filling strategies for defensible annual sums of net ecosystem exchange," *Agricultural* and Forest Meteorology, vol. 107, no. 1, pp. 43–69, 2001.
- [4] A. R. Desai, P. V. Bolstad, B. D. Cook, K. J. Davis, and E. V. Carey, "Comparing net ecosystem exchange of carbon dioxide between an old-growth and mature forest in the upper Midwest, USA," *Agricultural and Forest Meteorology*, vol. 128, no. 1-2, pp. 33–55, 2005.
- [5] A. Noormets, J. Chen, and T. R. Crow, "Age-dependent changes in ecosystem carbon fluxes in managed forests in northern Wisconsin, USA," *Ecosystems*, vol. 10, no. 2, pp. 187–203, 2007.
- [6] M. Reichstein, E. Falge, D. Baldocchi et al., "On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm," *Global Change Biology*, vol. 11, no. 9, pp. 1424–1439, 2005.
- [7] D. Hui, S. Wan, B. Su, G. Katul, R. Monson, and Y. Luo, "Gapfilling missing data in eddy covariance measurements using multiple imputation (MI) for annual estimations," *Agricultural and Forest Meteorology*, vol. 121, no. 1-2, pp. 93–111, 2004.

- [8] B. H. Braswell, W. J. Sacks, E. Linder, and D. S. Schimel, "Estimating diurnal to annual ecosystem parameters by synthesis of a carbon flux model with eddy covariance net ecosystem exchange observations," *Global Change Biology*, vol. 11, no. 2, pp. 335–355, 2005.
- [9] A. M. Moffat, D. Papale, M. Reichstein et al., "Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes," *Agricultural and Forest Meteorology*, vol. 147, no. 3-4, pp. 209–232, 2007.
- [10] D. Papale and R. Valentini, "A new assessment of European forests carbon exchanges by eddy fluxes and artificial neural network spatialization," *Global Change Biology*, vol. 9, no. 4, pp. 525–535, 2003.
- [11] A. Schmidt, T. Wrzesinsky, and O. Klemm, "Gap filling and quality assessment of CO₂ and water vapour fluxes above an urban area with radial basis function neural networks," *Boundary-Layer Meteorology*, vol. 126, no. 3, pp. 389–413, 2008.
- [12] W. Knorr and J. Kattge, "Inversion of terrestrial ecosystem model parameter values against eddy covariance measurements by Monte Carlo sampling," *Global Change Biology*, vol. 11, no. 8, pp. 1333–1351, 2005.
- [13] K. Morgenstern, T. A. Black, E. R. Humphreys et al., "Sensitivity and uncertainty of the carbon balance of a Pacific Northwest Douglas-fir forest during an El Niño/La Niña cycle," *Agricultural and Forest Meteorology*, vol. 123, no. 3-4, pp. 201–219, 2004.
- [14] Y.-Y. Chen, C.-R. Chu, and M.-H. Li, "A gap-filling model for eddy covariance latent heat flux: estimating evapotranspiration of a subtropical seasonal evergreen broad-leaved forest as an example," *Journal of Hydrology*, vol. 468, pp. 101–110, 2012.
- [15] A. D. Richardson and D. Y. Hollinger, "A method to estimate the additional uncertainty in gap-filled NEE resulting from long gaps in the CO₂ flux record," *Agricultural and Forest Meteorology*, vol. 147, no. 3-4, pp. 199–208, 2007.
- [16] X. Zhao, Y. Huang, Z. Jia et al., "Effects of the conversion of marshland to cropland on water and energy exchanges in northeastern China," *Journal of Hydrology*, vol. 355, no. 1–, pp. 181–191, 2008.
- [17] W. J. Massman and X. Lee, "Eddy covariance flux corrections and uncertainties in long-term studies of carbon and energy exchanges," *Agricultural and Forest Meteorology*, vol. 113, no. 1–4, pp. 121–144, 2002.
- [18] A. I. J. M. van Dijk and A. J. Dolman, "Estimates of CO₂ uptake and release among European forests based on eddy covariance data," *Global Change Biology*, vol. 10, no. 9, pp. 1445–1459, 2004.
- [19] J. M. Wilczak, S. P. Oncley, and S. A. Stage, "Sonic anemometer tilt correction algorithms," *Boundary-Layer Meteorology*, vol. 99, no. 1, pp. 127–150, 2001.
- [20] E. K. Webb, G. I. Pearman, and R. Leuning, "Correction of flux measurements for density effects due to heat and water vapour transfer," *Quarterly Journal Royal Meteorological Society*, vol. 106, no. 447, pp. 85–100, 1980.
- [21] P. D. Blanken, T. A. Black, H. H. Neumann et al., "Turbulent flux measurements above and below the overstory of a boreal aspen forest," *Boundary-Layer Meteorology*, vol. 89, no. 1, pp. 109–140, 1998.
- [22] L. Gu, E. M. Falge, T. Boden et al., "Objective threshold determination for nighttime eddy flux filtering," *Agricultural* and Forest Meteorology, vol. 128, no. 3-4, pp. 179–197, 2005.
- [23] D. Y. Hollinger, J. Aber, B. Dail et al., "Spatial and temporal variability in forest—atmosphere CO₂ exchange," *Global Change Biology*, vol. 10, no. 10, pp. 1689–1706, 2004.

- [24] G. Wohlfahrt, C. Anfang, M. Bahn et al., "Quantifying nighttime ecosystem respiration of a meadow using eddy covariance, chambers and modelling," *Agricultural and Forest Meteorology*, vol. 128, no. 3-4, pp. 141–162, 2005.
- [25] V. J. Stauch and A. J. Jarvis, "A semi-parametric gap-filling model for eddy covariance CO₂ flux time series data," *Global Change Biology*, vol. 12, no. 9, pp. 1707–1716, 2006.



Journal of Geochemistry





















International Journal of Mineralogy

International Journal of Atmospheric Sciences

Soil Science

Computational Environmental Sciences