Global Change Biology

Global Change Biology (2013) **19**, 1941–1952, doi: 10.1111/gcb.12127

TECHNICAL ADVANCE

Gap filling strategies and error in estimating annual soil respiration

NURIA GOMEZ-CASANOVAS*†, KRISTINA ANDERSON-TEIXEIRA*†[1], MARCELO ZERI‡, CARL J. BERNACCHI*§¶ and EVAN H. DELUCIA*†§

*Energy Biosciences Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, †Institute of Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, ‡Centro de Ciência do Sistema Terrestre, Instituto Nacional de Pesquisas Espaciais, Cachoeira Paulista, SP 12630-000, Brazil, §Department of Plant Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA, ¶Global Change and Photosynthesis Research Unit, Agricultural Research Service, USDA, Urbana, IL 61801, USA

Abstract

Soil respiration (R_{soil}) is one of the largest CO_2 fluxes in the global carbon (C) cycle. Estimation of annual R_{soil} requires extrapolation of survey measurements or gap filling of automated records to produce a complete time series. Although many gap filling methodologies have been employed, there is no standardized procedure for producing defensible estimates of annual R_{soil}. Here, we test the reliability of nine different gap filling techniques by inserting artificial gaps into 20 automated R_{soil} records and comparing gap filling R_{soil} estimates of each technique to measured values. We show that although the most commonly used techniques do not, on average, produce large systematic biases, gap filling accuracy may be significantly improved through application of the most reliable methods. All methods performed best at lower gap fractions and had relatively high, systematic errors for simulated survey measurements. Overall, the most accurate technique estimated R_{soil} based on the soil temperature dependence of R_{soil} by assuming constant temperature sensitivity and linearly interpolating reference respiration (R_{soil} at 10 °C) across gaps. The linear interpolation method was the second best-performing method. In contrast, estimating R_{soil} based on a single annual R_{soil} – T_{soil} relationship, which is currently the most commonly used technique, was among the most poorly-performing methods. Thus, our analysis demonstrates that gap filling accuracy may be improved substantially without sacrificing computational simplicity. Improved and standardized techniques for estimation of annual R_{soil} will be valuable for understanding the role of R_{soil} in the global C cycle.

Keywords: calculation error, carbon cycle, measurement, temporal extrapolation

Received 20 November 2012 and accepted 10 December 2012

Introduction

Soil respiration (R_{soil}) is one of the largest CO_2 fluxes in the global carbon (C) cycle, contributing up to 80 Pg C to the atmosphere annually (Schlesinger, 1977; Raich *et al.*, 2002), and plays an important role in regulating the feedbacks between terrestrial ecosystems and the atmosphere (Raich *et al.*, 2002; Field *et al.*, 2007). Quantifying the contribution of R_{soil} to the global C cycle requires estimates of total annual R_{soil}. However, the large temporal variability in R_{soil} rates presents a methodological challenge to obtaining reliable annual sums.

Measurements of R_{soil} may be made using either manual or automated systems (reviewed in Savage & Davidson, 2003; Savage *et al.*, 2008). Regardless of the methodology used, continuous year-long data records appropriate for constructing annual R_{soil} sums are generally not obtained. Manual systems are limited to low sampling frequency, yielding a record that is mostly gap (>99% of half hour periods throughout a year for biweekly sampling). Measurements typically are made at weekly to monthly intervals for the growing season or an entire year, and over a specific period of the day believed to be representative of the daily R_{soil} flux (typically during daytime hours; e.g. Davidson *et al.*, 1998; Savage & Davidson, 2003). Thus, gap occurrence in manual R_{soil} datasets is nonrandom, with data existing for defined times – generally at approximately the same time of day and with a potentially strong bias toward favorable sampling conditions (e.g. nonstormy days). As of yet, there is no consensus as to the best methodology for extrapolation of these low-frequency measurements to produce annual R_{soil} estimates.

Correspondence: Evan H. DeLucia, tel. +1 217 333 6177, fax +1 217 244 7246, e-mail: delucia@life.illinois.edu

[1]Present address: Smithsonian Conservation Biology Institute (Front Royal, Virginia) and Smithsonian Institution Global Earth Observatory-Center for Tropical Forest Science (Smithsonian Tropical Research Institute, Panama)

Automated systems are designed to sample R_{soil} continuously over long periods of time, often with the purpose of obtaining annual estimates. Measurements typically occur at least once per hour throughout the day. Although these systems can in theory provide a continuous data record for an entire year, the data records usually contain gaps (e.g. Wu *et al.*, 2010). Gaps occur for a variety of reasons, including unrepresentative conditions (e.g. plant growth inside a R_{soil} collar), instrument damage or failure, low quality of the data, and intentional shutdown because of unfavorable conditions (e.g. snow). Data gaps in automated R_{soil} records are semi random, occurring at any time of the day or year, but are more likely to occur under certain conditions (e.g. winter). Gaps vary in length from less than an hour to multiple weeks. As with automated systems, there is no consensus regarding the best methodology for gap filling automated R_{soil} records.

A variety of methods have been employed to gap fill R_{soil} records to produce annual estimates (Fig. 1), including simple algorithms (e.g. linear interpolation) or by prediction of missing values based on the relationship of R_{soil} to known variables (e.g. soil temperature). The linear interpolation method is commonly used (Fig. 1; e.g. Savage & Davidson, 2003). Implicit in this method are the assumptions that (1) R_{soil} changes at a constant rate between one measured record and the next, and (2) measurements are timed so as to characterize all critical points in the time series. Both these assumptions are violated by real R_{soil} time series because measurements do not typically characterize diurnal variation and because R_{soil} varies nonlinearly through time. However, if measurements are timed so as to characterize average daily R_{soil}, this method can be the most practical means of estimating annual R_{soil} when soil micrometeorology data are unavailable. Other simple algorithms, such as cubic interpolation or monthly average methods, may also be viable gap filling methods.

When continuous records of variables that drive R_{soil} are available, missing values may be predicted based upon these variables. The most frequently used variable is soil temperature, which is generally a good predictor of R_{soil} (Raich & Schlesinger, 1992; Raich *et al.*, 2002; but see Davidson *et al.*, 2006; Gomez-Casanovas *et al.*, 2012). The relationship between temperature and R_{soil} has been described using a variety of exponential functions; one of the most accurate and widely used functions is that proposed by Lloyd & Taylor (1994):

$$R_{soil}(T) = R_{ref} e^{E_0 \cdot \left(\frac{1}{T_{ref}-T_0} - \frac{1}{T-T_0}\right)} \quad (1)$$

Here, R_{ref} is respiration at a reference temperature (T_{ref}; often 10 °C), E_0 is a fitted parameter characterizing the

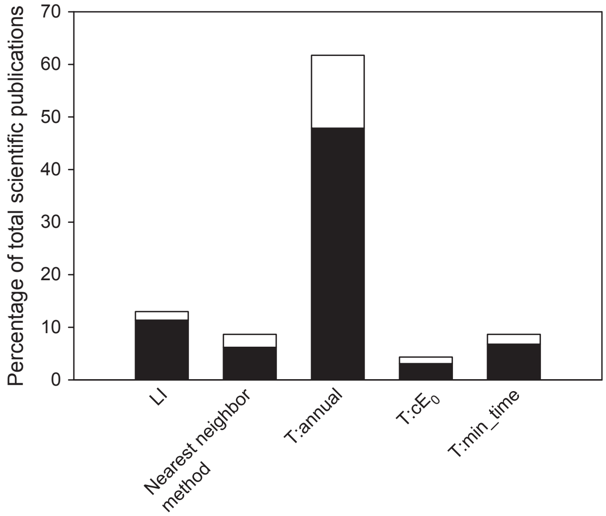

Fig. 1 Published strategies used to gap fill R_{soil} records. A literature search was conducted using Science Citation Index Expanded database from ISI Web of Knowledge, Web of Science with annual sum of R_{soil} rates obtained using either manual (closed bar) or automated (open bar) systems with no restriction on date of publication. A total of 177 scientific publications were found that met these criteria, and these publications reported 196 values of annual R_{soil} sums (Appendix S1). Methods are as described in Table 1, with the caveat that the temperature-dependence models (T : annual, T : cE_0, T : min_time) are more broadly defined to include fits based on Van't Hoff, Arrhenius, or the Lloyd and Taylor respiration functions (Lloyd & Taylor, 1994). Studies are counted as T : min_time if E_0 (or Q_{10}) (temperature sensitivity of R_{soil}) and R_{ref} (temperature-standardized respiration rate) were calculated by fitting more than a single relationship between R_{soil} and soil temperature for the year (e.g. growing season and nongrowing season; Ruehr *et al.*, 2010). From a total of 177 scientific publications, 5% of the soil temperature methods assumed an interaction between soil temperature and moisture, whereas 13% of these methods assumed no interaction between these two variables.

temperature sensitivity (Arrhenius, 1889), and T_0 is the lower temperature limit for R_{soil}. This model has two fitted parameters that may vary substantially over the course of a year, R_{ref} and E_0. R_{ref} is a temperature-standardized respiration rate which may be influenced by many variables including root biomass, the size and composition of the microbial community, belowground carbon allocation by plants, the quantity and quality of C compounds in the soil, and soil moisture. E_0 theoretically characterizes the inherent temperature sensitivity of aerobic respiration; however, the apparent value of E_0 will be altered if R_{ref} varies at the time scale over which the model is applied (Curiel-Yuste *et al.*, 2004; Davidson *et al.*, 2006; Sampson *et al.*, 2007; Anderson-Teixeira *et al.*, 2008). The most commonly used method of extrapolating incomplete R_{soil} records to yield an annual estimate is predicting missing values based on a fitting of Eqn 1 (or a similar function) to the entire annual dataset (Fig. 1; e.g. Irvine & Law, 2002). Implicit in this approach are the assumptions that E_0 and R_{ref} both are constant over the year. However, E_0 and R_{ref} can vary seasonally (Curiel-Yuste *et al.*, 2003; Janssens & Pilegaard, 2003; Sampson *et al.*, 2007; Chen *et al.*, 2010), implying that better models may be obtained by allowing one or both parameters to vary throughout the year. Relatively few studies have employed gap filling techniques that allow E_0 or R_{ref} to vary over the year (Fig. 1; e.g. Anderson-Teixeira *et al.*, 2013; Drake *et al.*, 2011).

Models predicting R_{soil} may be further improved by including soil moisture and correlates of primary productivity (e.g. leaf area index; Reichstein *et al.*, 2003). Including soil moisture in models to gap fill R_{soil} datasets is complex (Reichstein & Beer, 2008) because soil water status can affect several physiological parameters such as enzyme activities, or the turnover of roots and soil microorganisms. In addition, soil water status can affect gas and substrate diffusion thereby affecting substrate availability (Rodrigo *et al.*, 1997). Soil moisture and temperature often interact; E_0 generally decreases with decreases in soil moisture (Reichstein *et al.*, 2002a, b; Janssens & Pilegaard, 2003). The interactive effects of soil temperature and moisture may be described using the following equation (Reichstein *et al.*, 2003):

$$R_{soil} = R_{ref} \cdot e^{E_0(\mathrm{RSWC}) \cdot \left(\frac{1}{T_{ref}-T_0} - \frac{1}{T_{soil}-T_0}\right)} \cdot \frac{\mathrm{RSWC}}{\mathrm{RSWC}_{1/2} + \mathrm{RSWC}} \qquad (2)$$

Here, RSWC is relative soil water content (RSWC = $\mathrm{SWC}/\mathrm{SWC}_{FC}$), where SWC is soil water content in $m^3\ m^{-3}$, SWC_{FC} is soil water content at field capacity and $\mathrm{RSWC}_{1/2}$ is the RSWC where R_{soil} is half its maximal value (at a given temperature). E_0 is assumed to vary linearly with RSWC (E_0 (RSWC) = $a + b \cdot$ RSWC, where a and b are fitted parameters). Although gap filling methods using both soil temperature and moisture are rare (Reichstein *et al.*, 2003), models incorporating temperature-moisture interactions may yield improvements for predicting R_{soil}.

Table 1 Gap filling strategies evaluated in this study. Full methodological details are presented in Appendix S1

Method	Data required	Description
Simple Algorithm Methods	R_{soil}	
Linear Interpolation (LI)		Missing R_{soil} rates extrapolated by linear interpolation to the entire year
Cubic Interpolation (CI)		Missing R_{soil} rates extrapolated by cubic interpolation to the entire year
Monthly Average (MA)		For gap time frames <30 days, missing R_{soil} values replaced by the average of R_{soil} data for the current month
Soil Temperature-Dependence Methods	R_{soil}, T_{soil}	
Temperature dependence, fit annually (T : annual)		E_0 and R_{ref} calculated by fitting a single relationship between R_{soil} and temperature for the entire year
Constant temperature sensitivity, varying R_{ref} (T : cE_0)		E_0 assumed to be constant for the entire year (308.56 K^{-1}); R_{ref} calculated at each time frame for which R_{soil} data was available and then extrapolated to the entire year using a linear interpolation
Temperature dependence, fit over minimum time frame (T : min_time)		E_0 and R_{ref} fitted using the shortest possible time for characterizing the relationship between R_{soil} and soil temperature for which sufficient data are available. Defaults to linear interpolation when fitting fails or is unrealistic
Soil Temperature- and Moisture-Dependence Methods	R_{soil}, T_{soil}, *RSWC*	
Temperature × moisture dependence, fit annually (TxSWC : annual)		E_0 (RSWC) and R_{ref} calculated by fitting a single relationship between R_{soil}, soil temperature and moisture (Eqn 2) for the entire year
Constant temperature sensitivity, R_{ref} moisture-dependent (T+SWC : cE_0)		E_0 assumed to be constant for the entire year (308.56 K^{-1}); R_{ref} calculated by assuming a linear dependency of R_{ref} on the relative water content (*RSWC*)
Temperature x moisture dependence, fit over minimum time frame (TxSWC : min_time)		E_0(*RSWC*)and R_{ref} calculated using the shortest possible time for characterizing the relationship between R_{soil}, soil temperature and moisture (Eqn 2) for which sufficient data are available. Defaults to linear interpolation when fitting fails or is unrealistic

The reliability of various strategies for filling gaps in R_{soil} records has not been assessed (Vargas *et al.*, 2010). In this study, we evaluate the performance of different gap filling methods (Table 1) using measured records with varying gap fractions. Specifically, we evaluate the performance of these different gap filling methods by analyzing the errors introduced when filling artificial gaps in automated R_{soil} datasets obtained over 3 years for various ecosystem types. Introduced gaps ranged from none to >99% gap. In addition, we analyzed how the timing of survey measurements (>99% gap) affected gap filling performance, considering two time frames for measurement (9:00 hours–17:00 hours and 9:00 hours–12:00 hours) and two portions of the year (entire year and growing season only). We use our results to provide guidance as to which gap filling methods should be used to provide accurate annual R_{soil} estimates for both automated and manual R_{soil} systems.

Materials and methods

Study site

High-frequency data for this analysis were collected from research plots at the University of Illinois Energy Farm in Urbana, IL (40°3′46.209″N, 88°11′46.0212″W, ca. 220 m above sea level). Average mean monthly (1979–2009) air temperature was 11.1 °C while the mean accumulated rainfall was 1042 mm (Illinois State Water Survey Historic Climate Data). For the 3 years used in this study (2009, 2010 and 2011), annual mean temperature was 10.4, 11.3, and 11.6 °C, respectively, according to measurements from a weather station located ca. 5 km away at the University of Illinois-Willard Airport. Cumulative rainfall from the same weather station for these 3 years was 1147, 898, and 921 mm, respectively. The study area is flat and soils belong to silt loam Flanagan soil series with small contributions of Drummer soils. The experimental plots at the Energy Farm consisted of four 3.8-ha plots, each planted with miscanthus (*Miscanthus* × *giganteus*), switchgrass (*Panicum virgatum* L), a mix of native prairie species (see Zeri *et al.*, 2011 for species list), or a corn-corn-soybean rotation. Prior to this experiment, these plots were cultivated with annual arable crops for over 100 years. Land conversion into the experimental plots occurred in spring 2008. Details about crop management after land conversion at this site can be found in Zeri *et al.* (2011).

R_{soil}, temperature and moisture measurements

R_{soil} was measured continuously from January 2009 to December 2011 in each experimental plot with an automated system consisting of an infrared gas analyzer (LI-8100-101; Licor, Inc., Lincoln, NE, USA) connected to a multiplexer (LI-8150; Licor, Inc.) controlling four automated 20-cm diameter chambers (LI-8011; Licor, Inc.). Each chamber was located ca. 15 m from the center of the plot. Each chamber closed on a PVC collar permanently inserted ca. 2–5 cm into the soil. R_{soil} fluxes were measured every half hour providing a total of 48 annual R_{soil} records (i.e. 4 R_{soil} records per plot for 3 years). Quality control of the data consisted on deleting R_{soil} values associated with equipment malfunction (e.g. chamber close failures, pump failures, damaged air hoses) and anomalous values (e.g. from plant growth within the chamber space). Only those R_{soil} records with less than 50% of data missing were used in this study, yielding a total of 20 R_{soil} records. On average, 33% of the R_{soil} data was missing for 2009, 2010, and 2011 R_{soil} datasets. The 20 R_{soil} records were evenly distributed among the four ecosystems. Soil temperature at 5 cm and moisture at 10-cm depth were also measured continuously every half-hour within each plot. Quality control of the data consisted on the removal of out-of-range values (±3.5 standard deviations from the mean).

Artificial gap creation and filling

To investigate the performance of each gap filling strategy (Table 1), artificial gaps were created in the 20 records of annual R_{soil} using code written in Matlab® v. 7.8.0. As R_{soil} records already contained real gaps, artificial gaps were only created in periods of existing data. This allowed for the analysis of the performance of each method when existing and predicted R_{soil} data were compared. Artificial gaps were introduced at 12 gap fractions (i.e. length of gap relative to the existing half hour R_{soil} rates): 0%, 5%, 15%, 25%, 35%, 45%, 55%, 65%, 75%, 85%, 95%, and >99% of the existing data. The highest gap scenario (>99% gap) mimicked a typical survey measurement scheme, with bimonthly measurements made between 9:00 hours and 17:00 hours at random dates in the first and second halves of each month.

Because gaps in automated R_{soil} records are nonrandomly distributed, occurring in clusters rather than as half-hour gaps randomly distributed throughout the record, artificial gaps were introduced to simulate the distribution of real gaps. This was critical because the performance of each method depends on the randomness of the gap distribution (Falge *et al.*, 2001). In our 48 automated R_{soil} records, less than 10% of the gaps occurred at a frequency of single half-hours, 60% of the gaps occurred at time frames between half-hours and 30 days, and 40% occurred at >30 days. This frequency distribution of gaps in R_{soil} records was mimicked when creating artificial gaps. As gaps introduced in this manner approached but did not exactly meet the target gap fractions, random half-hour gaps were inserted to make up the difference necessary to achieve the exact introduced gap fraction target.

For each R_{soil} chamber record, the artificial gap construction algorithm was permuted 100 times at each gap fraction, giving a total of 2000 records containing artificial gaps for each gap fraction. Artificial gaps were filled using each of the nine gap filling strategies (Table 1; detailed in Appendix S2). With 100 permutations of artificial gap creation, the standard error of the predicted annual R_{soil} sum obtained from each method at each gap fraction was accurate to within 10%.

In addition to the analysis of gap filling performance as a function of gap fraction, we also compared performance of gap filling method for four different simulated survey respiration schemes: (1) bimonthly measurements over the entire year between 9:00 hours and 17:00 hours (same as in the above analysis), (2) bimonthly measurements over the entire year between 9:00 hours and 12:00 hours – a time frame that has been recommended as representative of the average daily R_{soil} (Davidson *et al.*, 1998), (3) bimonthly measurements over the growing season (April-October) between 9:00 hours and 17:00 hours, and (4) bimonthly measurements over the growing season (April–October) between 9:00 hours and 12:00 hours. Gaps were constructed and filled as described above, with the exception that the T : min_time and TxSWC : min_time methods (detailed in Appendix S2) were not evaluated, as these are theoretically identical to their annual fitting counterparts (T : annual and TxSWC : annual, respectively) or to LI (when curve fitting fails) at the highest gap fraction.

Statistical analysis and annual R_{soil} sum

For statistical purposes, the performance of each method at each gap fraction was evaluated by comparing the filled R_{soil} data with the observed values. The coefficient of determination (R^2) and the root mean square error (RMSE) between the observed and predicted R_{soil} values was obtained. In addition, the potential bias error (BE), annual R_{soil} sum (ΣR_{soil}), and the ratio between the predicted and observed ΣR_{soil} was calculated for each gap fraction for each gap filling method. The coefficient of determination (R^2), RSME, and BE between the predicted (p) and the observed (o) R_{soil} values at any given time were calculated as in Moffat *et al.* (2007).

The ΣR_{soil} was calculated using the following equation:

$$\sum R_{soil} = \sum R_{soil} \cdot \Delta t \qquad (3)$$

Here, R_{soil} is the R_{soil} rate at any given half-hour and Δt is the measurement time interval. In addition, the ratio between the predicted and observed $\sum R_{soil}$ (ΣR_{soil} ratio) was calculated.

The performance of each gap filling strategy at each gap fraction was assessed with one-way analysis of variance (ANOVA) analysis. Gap filling strategy was set as a main factor, with years (2008, 2009, and 2011) and vegetation type (*Miscanthus*, switchgrass, perennial grass native, and a corn-corn-soy rotation) as covariates. To test their performance, the methods were ranked according to each of the statistical metrics (e.g. R^2, RSME, BE, ΣR_{soil} ratio) based on a multiple comparison test as in Moffat *et al.* (2007). These tests were implemented using Statgraphics Plus 5.0 (Statistical Graphics Corporation, Rockville, MD, USA).

The R^2 and RSME tested for the accuracy of each model (in phase correlation, and magnitude and distribution of the individual errors, respectively), and BE and the ΣR_{soil} ratio indicate the bias induced in the estimated half-hour R_{soil} rates and ΣR_{soil}, respectively. For instance, a method with low R^2 and a ΣR_{soil} ratio close to 1 would indicate that whereas the accuracy of the model is low, overestimates of half-hour R_{soil} values are being cancelled by underestimates, yielding reasonable estimates of ΣR_{soil}.

Results

The ability of all gap filling models to accurately predict ΣR_{soil} decreased above 15% and the variability increased with increasing gap fraction (Fig. 2). Inaccuracy and variability in the gap filling methods increased substantially for the simulated survey record (>99% gap fraction) relative to 95% gap fraction. Each method either consistently under- or overestimated BE and the ΣR_{soil} ratio. The performance of the various gap filling methods (Table 1) varied across R_{soil} records and the overall performance of each method was assessed by comparing statistics across all records.

Averaged across 20 R_{soil} records, permuted 100 times, RSME increased with increasing gap fraction (Fig. 3).

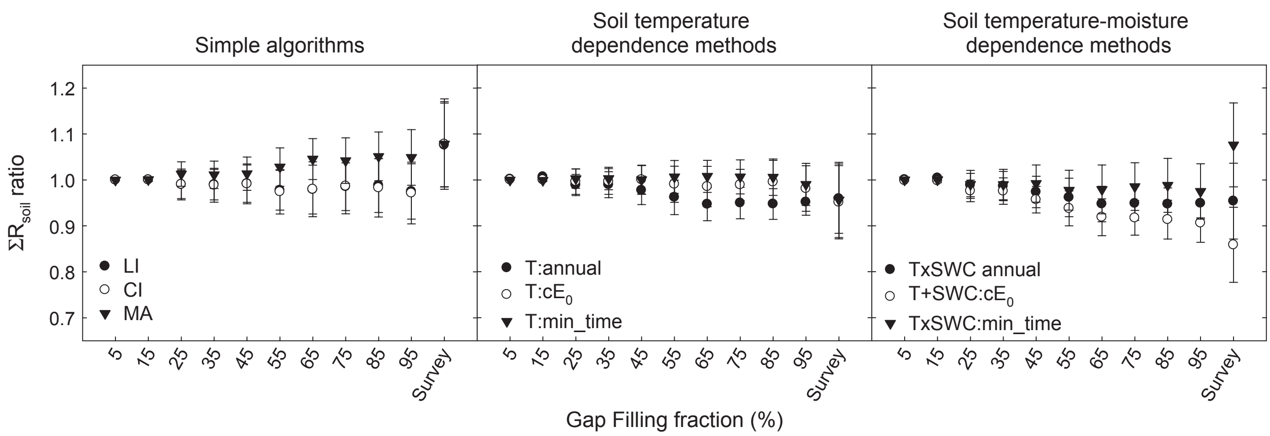

Fig. 2 The performance of three contrasting gap filling methods in estimating ΣR_{soil} across a range of gap fractions. Values are means ± standard deviation of the ratio of gap filled to real ΣR_{soil} (ΣR_{soil} ratio) for 100 artificial gap permutations of one record (switchgrass, 2009). Gap filling methods include the simple algorithms, soil temperature-dependence and soil temperature-moisture-dependence methods and are abbreviated as in Table 1.

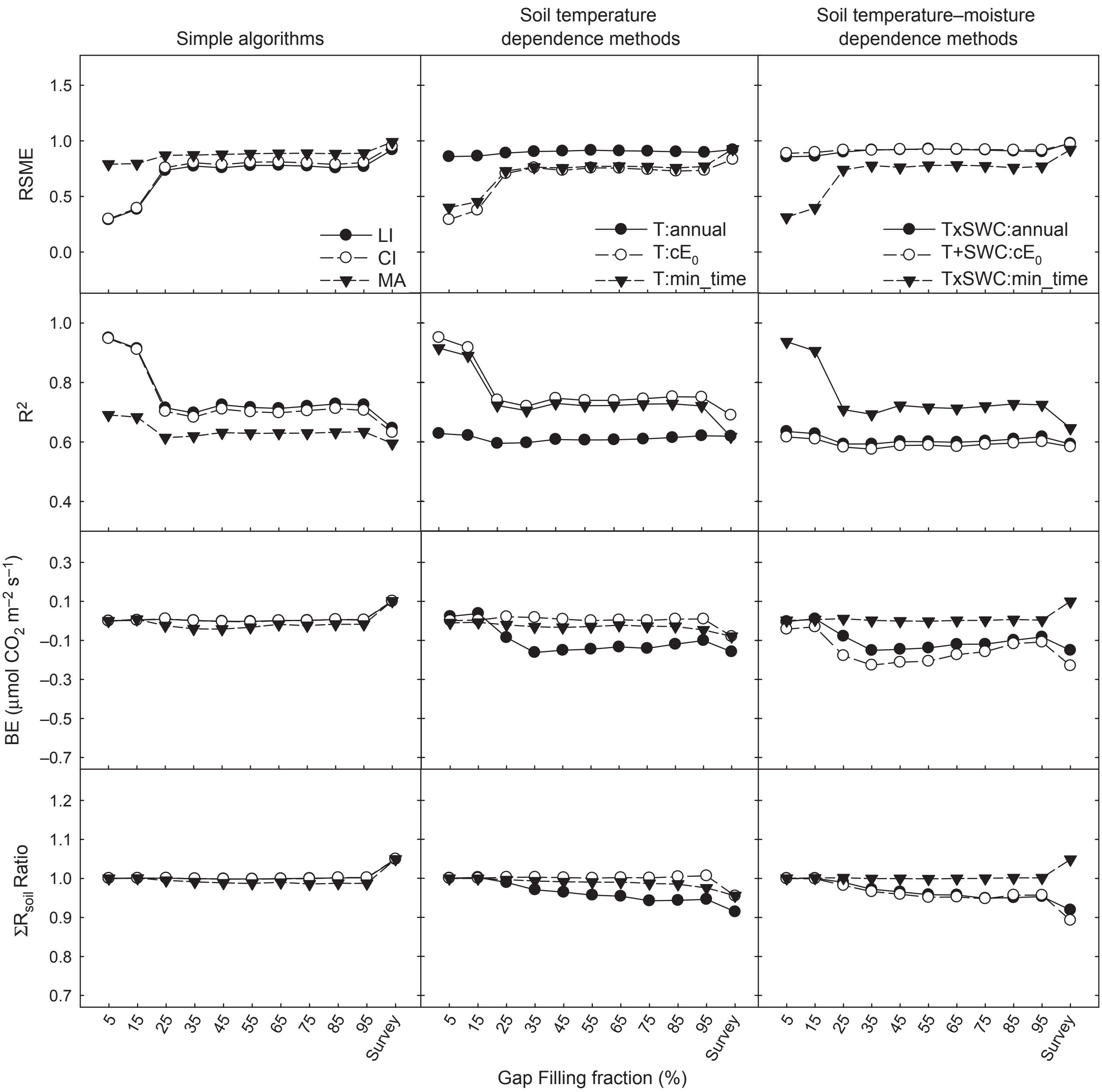

Fig. 3 Average performance of three contrasting gap filling methods across a range of gap fractions. Shown are means ± standard error of RMSE, R^2, bias error (BE), and the ratio of gap filled to real ΣR_{soil} (ΣR_{soil} ratio) for 20 records run 100 times each. Standard error values are smaller than symbols used for means. These metrics represent the average performance of each method over 100 artificial gap permutations as opposed to the expected performance on any one gap scenario (Fig. 2). Gap filling methods include the simple algorithms, soil temperature-dependence and soil temperature-moisture-dependence methods, and are abbreviated as in Table 1.

The statistical comparison of the RSME at 95% confidence level allowed for the evaluation of each method's performance at each gap fraction (Fig. 3; Table 2). At all gap fractions, the soil temperature method with constant temperature sensitivity and varying R_{ref} (i.e. the T : cE$_0$ method) had the lowest RSME (Fig. 3; Table 2). In addition, the RSME for the T : min_time method was the lowest at gap fractions between 25% and 65% (Fig. 3; Table 2). In contrast, methods that filled gaps based upon a single temperature-dependence or temperature- and moisture-dependence function (i.e. T : annual, TxSWC : annual, and T+SWC : cE$_0$ methods) had the highest RSME at all gap fractions evaluated, and at > 99% gap fraction the MA method had the highest RSME (Fig. 3; Table 2). The MAE (data not shown) of the gap fractions from each gap filling

Table 2 Methods performance at each gap fraction according to their RMSE (Root Mean Square Error), R^2, BE (Bias Error) and ΣR_{soil} ratio (ratio between predicted and observed ΣR_{soil}). The 'best performance' refers to methods that ranked significantly better than the others (95% CI), whereas 'poorest performance' refers to methods that ranked lowest based on each statistical metric. For each gap fraction level, methods that performed consistently best or poorest across all performance metrics are in bold. The survey scheme used in this analysis was Annual: 0900 CST-1700 CST

Gap fraction (%)	Best performance				Poorest performance			
	RMSE	R^2	BE	ΣR_{soil} ratio	RMSE	R^2	BE	ΣR_{soil} ratio
5	**LI, CI, T : cE$_0$**	**LI, CI** **T : cE$_0$**	**LI, CI** **T : cE$_0$** TxSWC : min_time	**LI, CI** **T : cE$_0$** TxSWC : min_time	T+SWC : cE$_0$	T : annual T+SWC : cE$_0$	T : annual	T : annual
15	**LI** T : cE$_0$	**LI**, CI TxSWC : min_time	**LI**, CI T : cE0 TxSWC : min_time	**LI**, CI T : cE$_0$ TxSWC : min_time	T+SWC : cE$_0$	T : annual T+SWC : cE$_0$	T : annual	T : annual
25	T : min_time **T : cE$_0$**	LI **T : cE$_0$**	LI, CI **T : cE$_0$** TxSWC : min_time	LI, CI **T : cE$_0$** TxSWC : min_time	**T+SWC : cE$_0$** TxSWC : annual	T : annual **T+SWC : cE$_0$** TxSWC : annual	**T+SWC : cE$_0$**	**T+SWC : cE$_0$**
35	T : min_time **T : cE$_0$** TxSWC : min_time	LI **T : cE$_0$** T : min_time	LI, CI **T : cE$_0$** TxSWC : min_time	LI, CI **T : cE$_0$** TxSWC : min_time	T : annual **T+SWC : cE$_0$** TxSWC : annual	T : annual **T+SWC : cE$_0$** TxSWC : annual	**T+SWC : cE$_0$**	**T+SWC : cE$_0$**
45	T : min_time **T : cE$_0$**	LI **T : cE$_0$**	LI, CI **T : cE$_0$** TxSWC : min_time	LI, CI **T : cE$_0$** TxSWC : min_time	T : annual **T+SWC : cE$_0$** TxSWC : annual	**T+SWC : cE$_0$** TxSWC : annual	**T+SWC : cE$_0$**	**T+SWC : cE$_0$**
55	T : min_time **T : cE$_0$**	**LI, CI** **TxSWC : min_time**	**LI, CI** **T : cE$_0$** **TxSWC : min_time**	**LI, CI** **T : cE$_0$** **TxSWC : min_time**	T : annual **T+SWC : cE$_0$** TxSWC : annual	T : annual **T+SWC : cE$_0$** TxSWC : annual	**T+SWC : cE$_0$**	**T+SWC : cE$_0$**
65	T : min_time **T : cE$_0$**	LI **T : cE$_0$**	LI, CI **T : cE$_0$** TxSWC : min_time	LI, CI **T : cE$_0$** TxSWC : min_time	**T : annual** **T+SWC : cE$_0$** TxSWC : annual	**T : annual** **T+SWC : cE$_0$** TxSWC : annual	**T : annual** **T+SWC : cE$_0$**	**T : annual** **T+SWC : cE$_0$**
75	**T : cE$_0$**	LI **T : cE$_0$**	LI, CI **T : cE$_0$** TxSWC : min_time	LI, CI **T : cE$_0$** TxSWC : min_time	**T : annual** T+SWC : cE$_0$ TxSWC : annual	**T : annual** T+SWC : cE$_0$ TxSWC : annual	**T : annual**	**T : annual**
85	**T : cE$_0$**	LI **T : cE$_0$**	LI, CI **T : cE$_0$** TxSWC : min_time	LI, CI **T : cE$_0$** TxSWC : min_time	**T : annual** T+SWC : cE$_0$ TxSWC : annual	**T : annual** T+SWC : cE$_0$ TxSWC : annual	**T : annual**	**T : annual**
95	**T : cE$_0$**	LI **T : cE$_0$**	LI, CI **T : cE$_0$** TxSWC : min_time	LI, CI **T : cE$_0$** TxSWC : min_time	T+SWC : cE$_0$ TxSWC : annual	T : annual T+SWC : cE$_0$ TxSWC : annual	T : annual	T : annual
Survey (>99)	**T : cE$_0$**	**T : cE$_0$**	T : min_time **T : cE$_0$**	T : min_time **T : cE$_0$**	MA **T+SWC : cE$_0$** TxSWC : annual	MA **T+SWC : cE$_0$** TxSWC : annual	**T+SWC : cE$_0$**	**T+SWC : cE$_0$**

method yielded results in agreement with the RSME (Fig. 3).

The R^2 for all gap filling techniques decreased with increasing gap fraction (Fig. 3). In the range of gap fractions from 5% to 95%, LI and T : cE_0 had the highest R^2 (Table 2). At 99% gap fraction, the soil temperature method with a 308.56 K^{-1} temperature sensitivity and varying R_{ref} had the highest R^2. At all gap fractions evaluated, the R^2 for the T : annual, and T+SWC : cE_0 methods were the lowest, and at >99% gap fraction the MA method had the lowest R^2.

The BE refers to the error in μmol CO_2 m^{-2} s^{-1} introduced to the predicted half-hour R_{soil} data. Error increased with increasing the gap fraction for all the methods evaluated (Fig. 3), and was largest for the T : annual, T+SWC : cE_0 and TxSWC : annual methods (within a range of $\pm$0.02 to $\pm$0.23 μmol CO_2 m^{-2} s^{-1} for 5% and >99% gap fractions, respectively; Fig. 3; Table 2). In contrast, the LI, CI, T : cE_0, and TxSWC : min_time methods had among the lowest BE at all gap fractions compared to the other methods ($\pm$0.0002 and $\pm$0.11 μmol CO_2 m^{-2} s^{-1} for 5% and >99% gap fraction, respectively; Fig. 3; Table 2).

The ratio between the predicted and observed ΣR_{soil} represents the bias on the ΣR_{soil} introduced through gap filling. The statistical comparison between predicted and observed ΣR_{soil} allowed for the evaluation of each method's performance (Table 2). For gaps of all lengths (5–95% gap fraction), the LI, CI, T : cE_0 and TxSWC : min_time methods consistently had ΣR_{soil} ratio closest to 1 compared to the other gap filling techniques (Fig. 3). At >99% gap fractions, the soil temperature methods that filled gaps based upon a constant or subannual temperature-dependence (i.e. T : cE_0 and T : min_time methods) had a ratio similar to 1. In contrast, the T : annual and T+SWC : cE_0 methods ΣR_{soil} ratio departed significantly from 1 at all gap fractions.

For the automated field measurements conducted for this study, the average ΣR_{soil} for the 20 R_{soil} records was 645 gC m^{-2} y^{-1}. The average gap fraction for these records was 34 $\pm$ 2% (ranging from 16% to 48% missing data) depending upon year and ecosystem measured. Using the T : annual method with 100 permutations of artificial gap creation, the predicted uncertainty for the ΣR_{soil} was within a range of $\pm$19 gC m^{-2} y^{-1} (Fig. 3). In contrast, for a gap fraction of 34% the soil temperature method with constant temperature sensitivity and varying R_{ref} estimated ΣR_{soil} within a range of $\pm$1.9 gC m^{-2} y^{-1} (Fig. 3).

The soil temperature method assuming a E_0 of 308.56 K^{-1} and varying R_{ref} yielded the most accurate estimates of ΣR_{soil} when R_{ref} was estimated from 0900 CST to 1700 CST and from 0900 CST to 1200 CST during

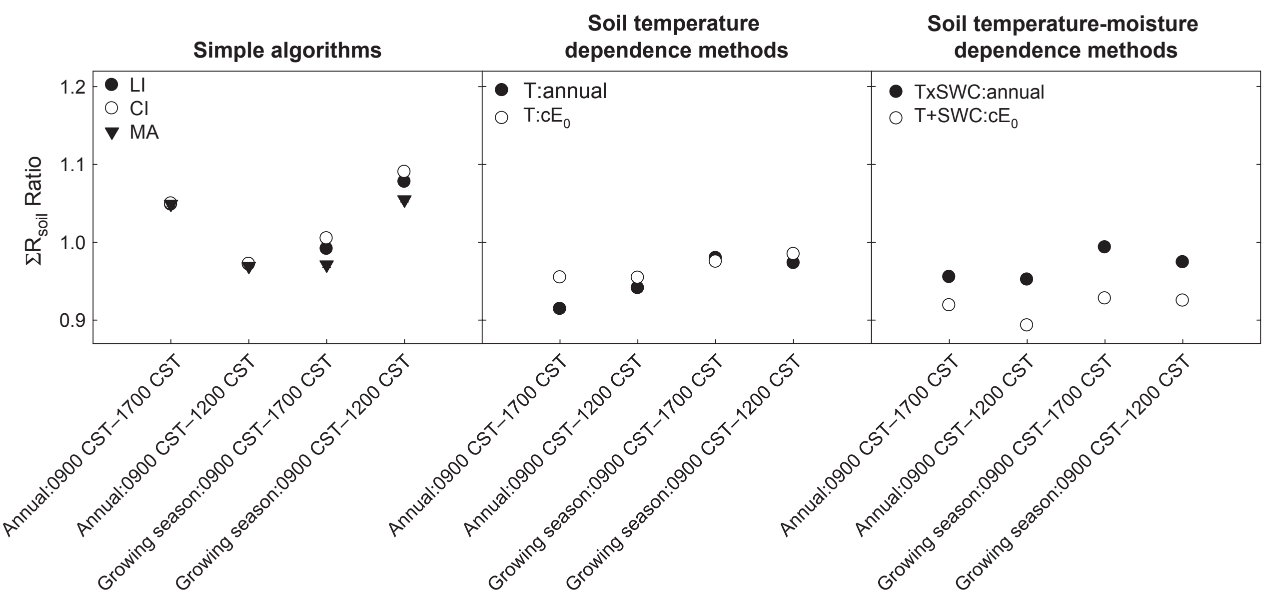

Fig. 4 Average performance of three contrasting gap filling methods for various survey soil respiration schemes. Artificial gaps represent surveys occurring over two different time frames (0900 CST-1700 CST or 0900 CST—1200 CST) and two different portions of the year (entire year-Annual- or from April to October-Growing season). Values are means $\pm$ standard error of the ratio of gap filled to real ΣR_{soil} (ΣR_{soil} ratio) for 20 records run 100 times each. Standard error values are smaller than symbols used for means. It is important to note that this metric represents the average performance of each method over multiple random gap permutations as opposed to the expected performance on any one gap scenario. Gap filling methods include the simple algorithms, soil temperature-dependence and soil temperature-moisture dependence methods and are abbreviated as in Table 1.

the growing season (Fig. 4). The ΣR_{soil} ratio for simple algorithm methods deviated least from 1 when R_{soil} was measured from 0900 CST to 1700 CST during the growing season (Fig. 4; 95% CI). For the growing season survey scheme from 0900 CST to1700 CST, the LI and TxSWC : annual methods had a ratio similar to 1 (Fig. 4). At all survey schemes, the method with a ratio that significantly departed from 1 was the soil temperature-moisture method without interaction (Fig. 4).

Discussion

In this study, we evaluated the performance of nine gap filling methods (Table 1) by analyzing the errors introduced when filling artificial gaps into R_{soil} datasets. The performance of these gap filling methods was assessed using various statistical metrics (RMSE, R^2, BE and ΣR_{soil} ratio). All methods performed best at lower gap fractions, and had relatively high, systematic errors for simulated survey measurements (Figs 2 and 3). For gap fractions between 5 and 95%, methods involving linear interpolation of R_{ref} or R_{soil} performed best (i.e. T : cE_0 and LI methods). For simulated survey measurements (>99% gap fraction), the T : cE_0 method performed best, although it tended to underestimate ΣR_{soil} by an average of 4% (Fig. 4). In contrast, methods that filled gaps based upon a single temperature-dependence or temperature- and moisture-dependence function (i.e. T : annual, TxSWC : annual, and T+SWC : cE_0 methods) performed relatively poorly at all gap fractions (Table 2). Indeed, the predicted uncertainty in ΣR_{soil} at 34% gap fraction using the T : annual method was twofold higher than the uncertainty in the ΣR_{soil} estimates calculated using the T : cE_0 method. These results indicate that the most commonly employed gap filling method (T : annual; Fig. 1) is not optimal, and that gap filling accuracy could be improved using easily applied methods for both automated and manual R_{soil} systems.

The differential performance of these gap filling methods can be understood based on the mechanisms driving soil respiration. Temperature is a primary driver of soil respiration and it is therefore no surprise that the method with best overall performance (T : cE_0) is the one that accounted for temperature. The T : cE_0 method assumed a constant E_0 while R_{ref} varied (see Eqn 1), as opposed to the other two temperature-dependent models that fit both E_0 and R_{ref} at annual or subannual time frames (T : annual and T : min_time). Although the temperature-dependence of R_{soil} often has been described using a constant temperature-sensitivity parameter (Lloyd & Taylor, 1994; Allen *et al.*, 2005; Reichstein *et al.*, 2005), E_0 varies seasonally (Janssens & Pilegaard, 2003; Curiel-Yuste *et al.*, 2004; Suseela *et al.*, 2012). Indeed, the theoretically expected temperature-dependence of aerobic respiration by microbes (e.g. Allen *et al.*, 2005) may be modified by many factors including substrate quality and supply, soil moisture, biophysical constraints, temperature-dependence of resource supply, and secretion of extracellular enzymes with a lower temperature -sensitivity (Davidson & Janssens, 2006; Allison *et al.*, 2010; O'Brien *et al.*, 2010; Wallenstein *et al.*, 2011; Anderson-Teixeira & Vitousek, 2012; Gomez-Casanovas *et al.*, 2012).

In addition, the temperature sensitivity of root respiration can be influenced by changes in plant carbon allocation patterns between the above- and the belowground component, root biomass, soil moisture and thermal acclimation of respiration (Kucera & Kirkham, 1971; Carbone & Trumbore, 2007; Drake *et al.*, 2008; Wang *et al.*, 2010). Our results showed that for the purposes of gap filling, it was almost always preferable to assume a constant E_0 (i.e. T : cE_0) than to attempt statistical estimation of E_0 over the minimum possible time frame (i.e. T : min_time method; detailed in Appendix S2; Figs 2–3, Table 2). This does not necessarily imply that E_0 is constant but that a reasonable approximation of the physiological temperature-dependence of R_{soil} coupled with linear interpolation of R_{ref} provides the best statistical fit. Further research will be required to understand the degree to which E_0 actually matches the assumed value (E_0 =308.56 K^{-1} in Eqn 1 ≈ E_a = 0.62 eV in Arrhenius equation ≈ Q_{10} = 2.4) over short-time frames (i.e. days to weeks), as is required to disentangle the temperature sensitivity of R_{soil} from confounding factors.

When both E_0 and R_{ref} were fit statistically (T : min_time method), gap filling performance should be optimized by minimizing the time frame over which the R_{soil}-temperature relationship (Eqn 1) was fit (T : min_time method; Figs 2–3; Table 2). This result makes sense in light of the fact that both E_0 and R_{ref} can vary seasonally in response to changes in plant phenology, root or microbial biomass, microbial composition, and soil moisture (Janssens & Pilegaard, 2003; Suseela *et al.*, 2012). Therefore, estimates of E_0 and R_{ref} calculated over long periods of time (i.e. the entire year as in the T : annual method) can reflect confounding factors such as phenology effects on soil respiration that may alter these parameters (Reichstein *et al.*, 2005). Gap filling methods that estimate E_0 and R_{ref} over short periods of time should theoretically reduce confounding effects in E_0 and R_{ref} estimates. Particularly at lower gap fractions ($\leq$65%) where there were sufficient data to fit Eqn 1 over short time frames, the T : min_time method was among the most accurate gap filling approaches (Fig. 3, Table 2), indicating its success at representing seasonal variation in E_0 and R_{ref}. For simulated survey records (>99% gap fraction), however, there were insufficient data for this method to create a unique fit; rather, it fit a single relationship for the entire year and defaulted to an LI fit if the fit was unreasonable (this gave this method the advantage over T : annual at the highest gap fraction; Fig. 3, Table 2).

In contrast to the T : min_time method, the T : annual method assumed that the R_{soil}-temperature relationship (Eqn 1) could be represented by a fit over the entire year. Because fitting a temperature-dependence relationship over an entire year can result in a strongly biased estimate of E_0 (Reichstein *et al.*, 2005), it is not surprising that this method was among the poorest performers (Fig. 3, Table 2). Particularly at the highest gap fractions, the T : annual method tended to underestimate R_{soil}. Underestimates in R_{soil} were likely caused by generally high E_0 (E_0 = 300–470 K^{-1}) with consequent low R_{ref} estimates as calculated R_{soil} depends on both E_0 and R_{ref}. The relatively poor performance of this method is notable, given that it has been the most widely used for gap filling in the past (Fig. 1). Our analysis indicates that annual R_{soil} estimates could be significantly improved by using the T : cE_0 method in place of the T : annual method (Fig. 3, Table 2).

Soil moisture plays an important role in the regulation of R_{soil} (e.g. Reichstein *et al.*, 2003); however, methods incorporating *RSWC* (TxSWC : annual, T+SWC : cE_0, and TxSWC : min_time) generally did not perform well (Table 2). There are two potential explanations for this relatively weak performance. First, because our study site is relatively mesic, soil moisture was not strongly limiting R_{soil} most of the time. In fact, the relationship between R_{ref} and *RSWC* at an annual time scale (T+SWC : cE_0 method) often was slightly negative or zero (data not shown), indicating that R_{soil} was more often unaffected or slightly reduced by excessive soil moisture than limited by dry conditions. Second, incorporating another predictive variable into gap filling methods may increase the predictive power when gap fractions are relatively low – as indicated by the superior performance of TxSWC : annual over T : annual and TxSWC : min_time over T : min_time at low gap fractions (Table 2), but this also increases the prevalence of unreasonable fits – particularly when data are limiting. Indeed, at gap fractions of 25% and higher, TxSWC : annual was consistently among the poorest performers in terms of accuracy (RMSE and R^2; Table 2). Particularly at higher gap fractions, the TxSWC : min_time method failed to produce reasonable predictions and defaulted to the LI fit. It is for this reason that TxSWC : min_time closely parallels LI, particularly at higher gap fractions (Fig. 3, Table 2), and, therefore, the success of TxSWC : min_time is generally not derived from the success of this model *per se*.

The strong performance of the gap filling techniques that use interpolation to estimate R_{soil} or R_{ref} (LI, CI, and T : cE_0) indicates that factors in addition to temperature and soil moisture were important predictors of R_{soil}. By assuming that missing values were best predicted by their nearest neighbors these methods effectively integrate several factors that may influence R_{soil} (Brüggemann *et al.*, 2011), including above and below ground partitioning (e.g. Carbone & Trumbore, 2007), labile C inputs from plant roots (Badri & Vivanco, 2009), the quantity, and quality of soil organic carbon (Kirschbaum, 1995), soil nitrogen availability (e.g. Janssens *et al.*, 2010), and measurement bias (e.g. minor, undetected equipment problems). The fact that R_{ref} was better predicted by nearest-neighbor measurements than by models that included soil moisture (i.e. comparing methods T : cE_0 and T+SWC : cE_0; Table 2) indicates that the contribution of soil moisture to R_{ref} was relatively small at our site. The relative performance of techniques predicting R_{ref} based on linear interpolation and *RSWC* (i.e. T : cE_0 and T+SWC : cE_0, respectively) may differ at semiarid or tropical sites where soil moisture more strongly affects R_{soil}.

Because of the extremely low measurement frequency, the most challenging gap filling scenario is estimating annual R_{soil} from a survey measurement record. All gap filling methods had low accuracy and high bias error when applied to >99% gap fractions (Fig. 3). The simple algorithm methods which make no correction based on temperature overestimated R_{soil} by on average ca. 5% (Fig. 3). In contrast, methods that incorporate a temperature-dependence systematically underestimated R_{soil} by on average ca. 5–12% (with the exception of TxSWC : min_time, as it defaulted to LI) (Fig. 3). Underestimates in ΣR_{soil} with soil temperature-dependence methods are likely driven by overestimates of E_0 and consequent underestimation of R_{ref}. This implies that factors other than temperature and moisture may be limiting R_{soil} and thus influencing the temperature sensitivity of R_{soil} (Davidson *et al.*, 2006; Gomez-Casanovas *et al.*, 2012).

The performance of gap filling methods for estimating annual R_{soil} based upon a survey measurement record interacted with the timing of measurements (e.g. 0900 CST-1700 CST or 0900 CST-1200 CST; Fig. 4) and whether measurements are made during the entire year or only during the growing season (April-October for our research sites; Fig. 4). The simple algorithm methods were strongly dependent upon the extent to which survey measurements were representative of average daily R_{soil} (Fig. 4). Therefore, if daily measurements are timed to represent average daily R_{soil}, annual estimates will be quite accurate. The methods that incorporated temperature and moisture-dependence were sensitive to the timing of measurements (Fig. 4). On average, these methods underestimated R_{soil} at all survey time frames. However, the soil temperature method with constant E_0 but varying R_{ref} performed best at survey schemes where R_{ref} was estimated during the growing season from 0900 CST to 1200 CST and from 0900 CST to 1700 CST (Fig. 4). These results imply that R_{ref} calculated over these time frames are representative of the daily R_{ref} estimates yielding good predictions of the daily R_{soil} flux.

Additional research is needed to understand the variables that influence R_{ref}, as this is fundamental to understanding the mechanisms that determine the temperature sensitivity of R_{soil}. This analysis of various gap filling strategies sets the basis for selecting optimal gap filling methods for accurate estimates of ΣR_{soil}. In general, the most accurate gap filling method for automated records among those evaluated here was T : cE_0, which assumed a constant E_0 and filled gaps in R_{ref} using linear interpolation (Fig. 3, Table 2; details in Appendix S2). Another advantage of this method is that it is computationally simple. If a complete soil temperature record is not available, linear interpolation (LI) was the most accurate method among the simple algorithms and was almost as accurate as T : cE_0. In ecosystems where soil moisture is a critical driver of R_{soil}, methods accounting for soil moisture may be most appropriate, particularly if soil moisture varies dramatically on time scales shorter than the average gap length. From this analysis, it appears that the TxSWC : annual method may provide a better fit than T+SWC : cE_0 (Fig. 3, Table 2), and this method would be theoretically expected to perform better in strongly water-limited systems as it captures the previously observed increase in E_0 with RSWC (Eqn 2; Reichstein *et al*., 2003). The TxSWC : min_time method did not provide accurate estimates of ΣR_{soil}, but it may perform better in highly water-limited systems. Determination of the best gap filling method for semiarid and tropical systems will require an analysis of the type presented here.

The time of day and time of year of survey measurements were important for accurate extrapolation to ΣR_{soil} (Fig. 4), and the daily representative R_{soil} fluxes will likely depend upon site and vary over time. Therefore, we recommend carefully scheduling surveys to represent the average daily R_{soil} for the T : cE_0 and LI method (Fig. 4).

This analysis provides guidance as to optimal techniques for estimating reliable ΣR_{soil} and sheds light on the mechanisms driving soil respiration. On the basis of this analysis, we recommend a shift away from the most widely used gap filling method (T : annual; Fig. 1), which was among the most poorly-performing methods evaluated here. We show that alternative methods, particularly T : cE_0 and LI are significantly more reliable. Improved and standardized techniques for estimation of annual R_{soil} will be valuable for comparing across sites and understanding the role of soil respiration in ecosystem- to global-scale C cycles.

Acknowledgements

This project was funded by the Energy Biosciences Institute (EBI; University of Illinois at Urbana-Champaign, Urbana, IL, USA). We thank Michael Masters for assistance in maintaining the automated R_{soil} systems and field sampling. We are also grateful to Tim Mies, Chris Rudisill, Drew Schlumpf and Emily Doherty for operation of the Energy Farm at the University of Illinois, Urbana-Champaign.

References

Allen AP, Gillooly JF, Brown JH (2005) Linking the global carbon cycle to individual metabolism. *Functional Ecology*, **19**, 202–213.

Allison SD, Wallenstein MD, Bradford MA (2010) Soil-carbon response to warming dependent on microbial physiology. *Nature Geoscience*, **3**, 336–340.

Anderson-Teixeira KJ, Vitousek PM (2012) Ecosystems. In: *Metabolic Ecology: A Scaling Approach* (eds Sibley RM, Brown JH, Kodric-Brown A), pp. 99–111. Wiley-Blackwell Publishing Ltd, New Jersey.

Anderson-Teixeira KJ, Vitousek PM, Brown JH (2008) Amplified temperature dependence in ecosystems developing on the lava flows of Mauna Loa. *Hawaii. Proceedings of the National Academy of Sciences*, **105**, 228–233.

Anderson-Teixeira KJ, Masters MD, Black CK, Zeri M, Bernacchi CJ, DeLucia EH (2013). Altered belowground carbon cycling following land use change to perennial bioenergy crops. *Ecosystems*, doi: 10.1007/s10021-012-9628-x.

Arrhenius S (1889) Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. *Zeitchsrift fur Physik Chemique*, **4**, 226–248.

Badri DV, Vivanco JM (2009) Regulation and function of root exudates. *Plant, Cell & Environment*, **32**, 666–681.

Brüggemann N, Gessler A, Kayler Z *et al*. (2011) Carbon allocation and carbon isotope fluxes in the plant-soil-atmosphere continuum: a review. *Biogeosciences*, **8**, 3619–3695.

Carbone MS, Trumbore SE (2007) Contribution of new photosynthetic assimilates to respiration by perennial grasses and shrubs: residence times and allocation patterns. *New Phytologist*, **176**, 1–12.

Chen Q, Wang Q, Han X, Wan S, Li L (2010) Temporal and spatial variability and controls of soil respiration in a temperate steppe in northern China. *Global Biogeochemical Cycles*, **24**, GB2010, doi: 10.1029/2009GB003538.

Curiel-Yuste J, Janssens IA, Carrara A, Meiresonne L, Ceulemans R (2003) Interactive effects of temperature and precipitation on R_{soil} in a temperate maritime pine forest. *Tree Physiology*, **23**, 1263–1270.

Curiel-Yuste J, Janssens IA, Carrara A, Ceulemans R (2004) Annual Q_{10} of soil respiration reflects plant phenological patterns as well as temperature sensitivity. *Global Change Biology*, **10**, 161–169.

Davidson E, Janssens I (2006) Temperature sensitivity of soil carbonn decomposition and feedbacks to climate change. *Nature*, **440**, 165–173.

Davidson EA, Belk E, Boone RD (1998) Soil water and temperature as independent or confounded factors controlling soil respiration in a temperate mixed hardwood forest. *Global Change Biology*, **4**, 217–227.

Davidson EA, Janssens IA, Luo Y (2006) On the variability of respiration in terrestrial ecosystems: moving beyond Q_{10}. *Global Change Biology*, **12**, 154–164.

Drake JE, Stoy PC, Jackson RB, DeLucia EH (2008) Fine-root respiration in a loblolly pine (Pinus taeda L.) forest exposed to elevated CO_2 and N fertilization. *Plant, Cell and Environment*, **31**, 1663–1672.

Drake JE, Gallet-Budynek A, Hofmockel KS *et al*. (2011) Increases in the flux of carbon belowground stimulate nitrogen uptake and sustain the long-term enhancement of forest productivity under elevated CO_2. *Ecology Letters*, **14**, 349–357.

Falge E, Baldocchi D, Olson R *et al*. (2001) Gap filling strategies for defensible annual sums of net ecosystem exchange. *Agricultural and Forest Meteorology*, **107**, 43–69.

Field CB, Lobell DB, Peters HA, Chiariello NR (2007) Feedbacks of terrestrial ecosystems to climate change. *Annual Review of Environment and Resources*, **32**, 1–29.

Gomez-Casanovas N, Matamala R, Cook DR, Gonzalez-Meler MA (2012) Net ecosystem exchange modifies the relationship between the autotrophic and heterotrophic components of soil respiration with abiotic factors in prairie grasslands. *Global Change Biology*, **18**, 2532–2545.

Irvine J, Law BE (2002) Contrasting soil respiration in young and old-growth ponderosa pine forests. *Global Change Biology*, **8**, 1183–1194.

Janssens IA, Pilegaard K (2003) Large seasonal changes in Q_{10} of R_{soil} in a beech forest. *Global Change Biology*, **9**, 911–918.

Janssens IA, Dieleman W, Luyssaert S *et al*. (2010) Reduction of forest soil respiration in response to nitrogen deposition. *Nature Geoscience*, **3**, 315–322.

Kirschbaum MUF (1995) The temperature dependence of soil organic matter decomposition, and the effect of global warming on soil organic C storage. *Soil Biology and Biochemistry*, **27**, 753–760.

Kucera C, Kirkham D (1971) Soil respiration studies in tall-grass prairie in Missouri. *Ecology*, **52**, 912–915.

Lloyd J, Taylor JA (1994) On the temperature dependence of R_{soil}. *Functional Ecology*, **8**, 315–323.

Moffat AM, Papale D, Reichstein M *et al*. (2007) Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *Agricultural and Forest Meteorology*, **147**, 209–232.

O'Brien SL, Jastrow JD, Grimley DA, Gonzalez-Meler MA (2010) Moisture and vegetation controls on decadal-scale accrual of soil organic carbon and total nitrogen in restored grasslands. *Global Change Biology*, **16**, 2573–2588.

Raich JW, Schlesinger WH (1992) The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate. *Tellus B*, **44**, 81–99.

Raich JW, Potter CS, Bhagawati D (2002) Interannual variability in global soil respiration, 1980–94. *Global Change Biology*, **8**, 800–812.

Reichstein M, Beer C (2008) Soil respiration across scales: the importance of a model-data integration framework for data interpretation. *Journal of Plant Nutrition and Soil Science*, **171**, 344–354.

Reichstein M, Tenhunen JD, Roupsard O, Ourcival JM, Rambal S, Dore S, Valentini R (2002a) Ecosystem respiration in two mediterranean evergreen holm oak forests: drought effects and decomposition dynamics. *Functional Ecology*, **16**, 27–39.

Reichstein M, Tenhunen JD, Roupsard O *et al*. (2002b) Severe drought effects on ecosystem CO_2 and H_2O fluxes at three Mediterranean evergreen sites: revision of current hypotheses? *Global Change Biology*, **8**, 999–1017.

Reichstein M, Rey A, Freibauer A *et al*. (2003) Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices. *Global Biogeochemical Cycles*, **17**, 1–15.

Reichstein M, Falge E, Baldocchi D *et al*. (2005) On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *Global Change Biology*, **11**, 1424–1439.

Rodrigo A, Recous S, Neel C, Mary B (1997) Modelling temperature and moisture effects on C-N transformations in soils: comparison of nine models. *Ecological Modelling*, **102**, 325–339.

Ruehr NK, Knohl A, Buchmann N (2010) Environmental variables controlling R_{soil} on diurnal, seasonal and annual time-scales in a mixed mountain forest in Switzerland. *Biogeochemistry*, **98**, 153–170.

Sampson DA, Janssens IA, Yuste JC, Ceulemans R (2007) Basal rates of soil respiration are correlated with photosynthesis in a mixed temperate forest. *Global Change Biology*, **13**, 2008–2017.

Savage K, Davidson EA (2003) A comparison of manual and automated systems for soil CO_2 flux measurements: trade-offs between spatial and temporal resolution. *Journal of Experimental Botany*, **54**, 891–899.

Savage K, Davidson EA, Richardson AD (2008) A conceptual and practical approach to data quality and analysis procedures for high-frequency R_{soil} measurements. *Functional Ecology*, **22**, 1000–1007.

Schlesinger WH (1977) Carbon balance in terrestrial detritus. *Annual Review of Ecology, Evolution, and Systematics*, **8**, 51–81.

Suseela V, Conant RT, Wallenstein MD, Dukes JS (2012) Effects of soil moisture on the temperature sensitivity of heterotrophic respiration vary seasonally in an old-field climate change experiment. *Global Change Biology*, **18**, 336–348.

Vargas R, Carbone MS, Reichstein M, Baldocchi DD (2010) Frontiers and challenges in R_{soil} research: from measurements to model-data integration. *Biogeochemistry*, **102**, 1–13.

Wallenstein M, Allison SD, Ernakovich J, Steinweg JM, Sinsabaugh R (2011) Controls on the temperature sensitivity of soil enzymes: a key driver of in situ enzyme activity rates. In: *Soil Enzymology, Soil Biology* (eds Shukla G, Varma A), pp. 245–258. Springer Berlin Heidelberg, New York.

Wang X, Piao S, Ciais P, Janssens IA, Reichstein M, Peng S, Wang T (2010) Are ecological gradients in seasonal Q_{10} of R_{soil} explained by climate or by vegetation seasonality? *Soil Biology and Biochemistry*, **42**, 1728–1734.

Wu X, Brüggermann N, Gasche R, Shen Z, Wolf B, Butterbach-Bahl K (2010) Environmental controls over soil-atmosphere exchange of N_2O, NO, and CO_2 in a temperate Norway spruce forest. *Global Biogeochemical Cycles*, **24**, GB2012, doi: 10.1029/2009GB003616.

Zeri M, Anderson-Teixeira K, Hickman G, Masters M, DeLucia Bernacchi CJ (2011) Carbon exchange by establishing biofuel crops in Central Illinois. *Agricultural, Ecosystems and Environment*, **144**, 319–329.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Reference list used for the literature search on published strategies used to gap fill R_{soil} records.
Appendix S2. Description of gap filling methods.