Basic and extensible post-processing of eddy covariance flux data with REddyProc.

Thomas Wutzler[1], Antje Lucas-Moffat[2,3], Mirco Migliavacca[1], Jürgen Knauer[1], Kerstin Sickel[1], Ladislav Šigut[4], Olaf Menzer[5], and Markus Reichstein[1]

[1]Max Planck Institute for Biogeochemistry, Hans-Knöll-Straße 10, 07745 Jena, Germany
[2]German Meteorological Service, Centre for Agrometeorological Research, Bundesallee 33, 38116 Braunschweig, Germany
[3]*Also at:* Thuenen Institute of Climate-Smart Agriculture, Bundesallee 65, 38116 Braunschweig, Germany
[4]Global Change Research Institute CAS, Bělidla 986/4a, CZ-60300 Brno, Czech Republic
[5]Department of Geography, University of California, Santa Barbara, CA 93106-4060, USA

Correspondence: Thomas Wutzler
(twutz@bgc-jena.mpg.de)

Abstract.

With the eddy covariance (EC) technique, net fluxes of carbon dioxide (CO_2) and other trace gases as well as water and energy fluxes can be measured at the ecosystem level. These flux measurements are a main source for understanding biosphere-atmosphere interactions and feedbacks by cross-site analysis, model-data integration, and up-scaling. The raw fluxes measured with the EC technique require an extensive and laborious data processing. While there are standard tools[1] available in open source environment for processing high-frequency (10 or 20 Hz) data into half-hourly quality checked fluxes, there is a need for more usable and extensible tools for the subsequent post-processing steps. We tackled this need by developing the `REddyProc` package in the cross-platform language R that provides standard CO_2-focused post-processing routines for reading (half-)hourly data from different formats, estimating the u_* threshold, gap-filling, flux-partitioning, and visualizing the results. In addition to basic processing, the functions are extensible and allow easier integration in extended analysis than current tools. New features include cross-year processing and a better treatment of uncertainties. A comparison of `REddyProc` routines with other state-of the art tools resulted in no significant differences in monthly and annual fluxes across sites. Lower uncertainty estimates of both u_* and resulting gap-filled fluxes with the presented tool was achieved by an improved treatment of seasons during the bootstrap analysis. Higher estimates of uncertainty in day-time partitioning resulted from a better accounting for the uncertainty in estimates of temperature sensitivity of respiration. The provided routines can be easily installed, configured, used, and integrated in further inquiry. Hence the eddy covariance community will benefit from using the `REddyProc` package, allowing easier integration of standard post-processing with extended analysis.

---

[1]fluxnet.fluxdata.org/2017/10/10/toolbox-a-rolling-list-of-softwarepackages-for-flux-related-data-processing

1 Introduction

The availability of ecosystem level observations of net ecosystem exchange (NEE) of carbon dioxide (CO_2) and other gases, latent heat (LE) and sensible heat (H) fluxes measured by the eddy covariance (EC) method (Aubinet et al., 2000) allowed a boost in ecosystem understanding at site to global scales (Baldocchi et al., 2017). The EC method provides half-hourly or hourly records of turbulent fluxes between an entire ecosystem and the atmosphere. These data are derived from high frequency measurements (10 or 20 Hz) of wind speed and direction together with measurements of air scalar characteristics such as CO_2 and water vapor concentration, or temperature. Methods to compute fluxes from high frequency measurements, for the quality check and quality assessment (QA/QC), and for the storage corrections have been consolidated in the last decades (Rebmann et al., 2012; Foken et al., 2012; Aubinet et al., 2012) and are available as open source software[2]. Although measured continuously, the (half)hourly EC data contain gaps due to instrument malfunctioning or records which are not representative of the ecosystem because of micrometeorological conditions under which the assumptions of the EC technique are not met (details in e.g. Foken and Wichura, 1996; Foken et al., 2004; Göckede et al., 2004; Foken et al., 2012). Hence, (half-)hourly records are marked with different quality flags and need further extensive post-processing as described by Papale et al. (2006).

NEE from periods with low friction velocity (u_*) (Aubinet et al., 2012) need to be detected and filtered out to avoid systematic biases in nighttime NEE (Papale et al., 2006). The screened flux time series with gaps need to be filled (Reichstein et al., 2005a) using the available flux data and meteorological measurements. An additional information can be obtained from NEE thanks to flux partitioning methods, that provide model estimates of gross primary production (GPP) and ecosystem respiration (R_{eco}) (Reichstein et al., 2005a). These gross fluxes are important to understand land-atmosphere interactions.

All these post-processing steps need to be performed routinely for EC data. Hence, it is desirable to have automated and reproducible post-processing tools available that can be easily used, extended, and integrated into researcher's own workflow. For this purpose we compiled all routines for the important CO_2-focused post-processing steps in the `REddyProc` package for the cross-platform, free R language. The `REddyProc` package loads time series of quality checked and storage corrected fluxes and the basic set of meteorological variables and provides a software environment to perform u_* threshold detection and filtering, gap filling and partitioning. Furthermore, a series of other functionalities like data import routines and data visualization are provided.

The objectives of the paper are to 1) provide a reference that describes the methodology of the processing used in the `REddyProc` package, and 2) show that the obtained results do not differ systematically from results obtained with standard post-processing implemented in the FLUXNET community (based on Papale et al., 2006; Reichstein et al., 2005a; Lasslop et al., 2010; Pastorello et al., 2017).

Table 1 explains used abbreviations. The first part of the paper (section 2) describes the post-processing methods. The second part (Section 3) presents the benchmarks the `REddyProc` implementation with standard post-processing tools. It details

---

[2] fluxnet.fluxdata.org/2017/10/10/toolbox-a-rolling-list-of-softwarepackages-for-flux-related-data-processing

differences in the implementations and possible consequences in obtained results and aggregated fluxes. An appendix gives an overview of the package with general design, an example of the post-processing, and links resources that get readers started with post-processing their own data.

Table 1. Abbreviations used repeatedly in the paper.

Symbol	Description
EC	eddy covariance
CO_2	carbon dioxide
NEE	net ecosystem exchange towards the atmosphere in $\mu\text{mol}\,CO_2\,\text{m}^{-2}\text{s}^{-1}$ (aggregated in $\text{gC}\,\text{m}^{-2}\text{yr}^{-1}$)
GPP	gross primary productivity (same units as NEE)
R_{eco}	ecosystem respiration (same units as NEE)
H, LE	sensible and latent heat flux in Wm^{-2}
u_*	friction velocity in ms^{-1}
Rg	shortwave incoming global radiation in Wm^{-2}
Tair	air temperature in °C
Tsoil	soil temperature in °C
VPD	vapor pressure deficit in hPa
LUT	look-up table (section 2.2.1)
MDC	mean diurnal course (section 2.2.2)
MDS	marginal distribution sampling (section 2.2.3)
E_0	temperature sensitivity parameter in eq. 1
R_{Ref}	respiration at reference temperature parameter eq. 1
LRC	light response curve (section 2.3.2)
DP06	C-implementation of the u_* threshold estimation by Dario Papale (section 3)
BGC16	2016-version of the online tool provided by the MPI-BGC, Jena, (section 3)

2   Methods of Post-processing

5   The post-processing relies on half-hourly or hourly measurements of NEE and ancillary meteorological data of u_*, global radiation (Rg), air or soil temperature (Tair, Tsoil), and vapor pressure deficit (VPD). The fluxes should be quality checked and, if applicable, storage corrected before their use in the package.

The post-processing follows a specific workflow:

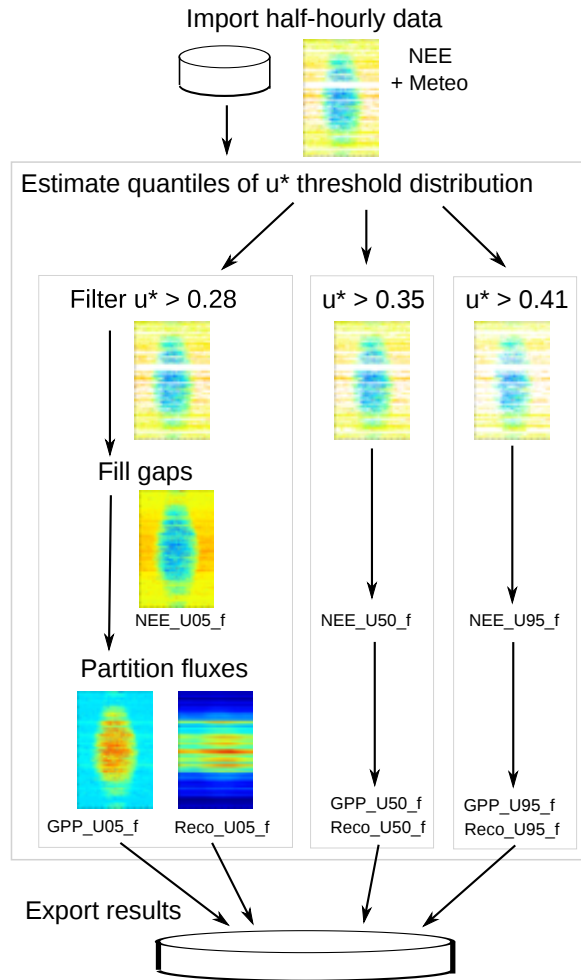



Figure 1. The workflow starts with importing the half-hourly (or hourly) data, in this example the year 1998 of site DE-Tha. Next, a probability distribution of u_* threshold is estimated for each season. Gap-filling and flux-partitioning is performed for several quantiles of this distribution for an estimate of uncertainty. Finally the results are exported.

1. determination and filtering of periods with low turbulent mixing (u_*-filtering),

2. replacing missing data in the half-hourly/hourly records (gap-filling), and

3. partitioning NEE into the gross fluxes GPP and R_{eco} (flux-partitioning).

Usage of the `REddyProc` package follows this data post-processing workflow (Fig.1). The following sections explain the steps
5   in more detail.

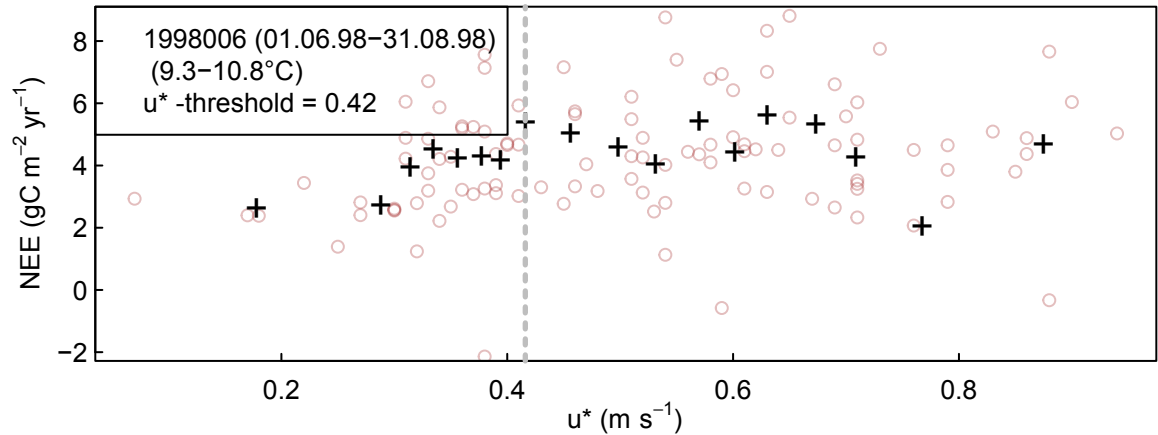



Figure 2. Concept of the u_*-filter for a subset of data from DE-Tha. Night-time NEE at low u_* friction velocities, i.e with low turbulence, is biased towards lower NEE values compared to cases with higher u_*, although other environmental conditions are similar. The u_* value below which this effect is considered significant, i.e. u_* threshold (dashed line), is estimated by a moving point method on u_* bins (crosses) across half-hourly records (circles).

2.1 u_*-filtering

Determining periods with low turbulent mixing is a critical step in the EC data post-processing. Standard steady state and integral turbulence characteristics tests in the initial processing exclude the most problematic records of sensible (H), latent (LE) heat and CO_2 fluxes (Foken and Wichura, 1996). However, it is well known (summarized in Aubinet et al., 2012, chapter 5), that

5   such a quality checking strategy is not sufficient, especially in the case of CO_2. Stable stratification that is present often during the night-time dampens turbulence and leads to an underestimation of the night-time NEE, i.e. the ecosystem respiration (Gorsel et al., 2007). Massman and Lee (2002) proposed that unfavorable conditions could be detected by inspecting the relationship of night-time NEE versus u_*. Within similar time period and similar environmental conditions respiration should not be dependent on the u_*. At low u_* values, a negatively biased respiration is measured. A heuristic class of methods, which is widely accepted,

10   assumes that a threshold of u_* can be established above that nighttime fluxes are considered valid. Hence, the u_* threshold is the minimum u_* above which respiration reaches a plateau (Fig. 2). This threshold is specific for each season of a site year. Uncertainties in the u_* threshold estimate represent one of the largest uncertainty components in the post-processing of NEE.

There are at least two methods of estimating the u_* threshold: the moving point method (Reichstein et al., 2005c; Papale et al., 2006), which is currently more routinely used, and the break-point detection method (Barr et al., 2013).

2.1.1 Moving point method for u_*

The method of Papale et al. (2006) detects a plateau in the relationship of night-time NEE versus u_* among all records within temperature subset by a moving point test of records binned into different u_* bins.

The nighttime data (default: Rg < 10 Wm^{-2}) is split into different times of year, here called seasons, to account for differing surface roughness. Then the data of each season is split into default six temperature subsets of equal size (according to quantiles). Within each temperature subset data is split into 20 about equally sized u_* bins. The default moving point method, called `Forward2`, determines the threshold based on these u_* bins. It checks for each bin if the mean NEE is higher than 0.95 times the mean of the following 10 bins. If this holds true also for the next bin, the mean u_* of the bin is reported as threshold. There are often subsets of data, where no clear threshold can be detected. Hence, there are quality criteria on whether the estimate of a given subset is used in subsequent aggregation. One quality criterion specifies that temperature and u_* should not be correlated within the temperature subset, another requires a minimum number of valid records within a subset. Next, the u_* estimates for different temperature classes and seasons (details in section 3.2.1) are aggregated to derive a robust u_* estimate. Within one season, the median is taken across the estimates of different temperature subsets. Within one year, the maximum is taken across the associated seasons.

Records during night-time with u_* smaller than the estimated threshold are flagged as invalid and are replaced in the subsequent gap-filling processing step. In addition, each half-hour after records with u_* smaller than the threshold is flagged to be invalid.

2.1.2 Breakpoint detection method for u_*

Alternatively, a breakpoint-detection can be applied to the unbinned data, that avoids the sensitivity of the Moving point method to the specifics of the binning schemes (Barr et al., 2013). `REddyProc` provides this method by estimating the breakpoint based on unbinned records within the seasons/temperature subsets using the `segmented` R-package (Muggeo, 2003, 2008). However, `REddyProc` differs from Barr et al. (2013) by keeping the same aggregating scheme of seasonal/temperature estimates to annual thresholds as with the Moving point method.

2.1.3 Bootstrapping uncertainty of the u_* threshold

Estimates of the u_* threshold are often sensitive to the specifics of the combination of methods and the data, e.g. the binning, minimum number of records within a season or temperature subset, criteria in aggregation, etc. Therefore, bootstrap (re-sampling with replacement) is applied to generate 200 artificial replicates of the dataset and for each replicate the threshold is estimated (Efron and Tibshirani, 1986; Davison and Hinkley, 1997). The 5th, 50th, and 95th percentile of the estimates are reported as a range of threshold estimates. The subsequent post-processing steps of gap-filling and partitioning are then repeated using those

different thresholds to propagate the uncertainty of u_* threshold estimation to derived quantities such as annual NEE, GPP and R$_{eco}$

2.2 Gap-filling methods

After quality checks and u_*-filtering, the dataset of half-hourly NEE fluxes may contain up to 50 percent gaps (sometimes
5 this fraction is even higher, depending on the site conditions). For the benchmark datasets used in this manuscript, the percentage of gaps before u_*-filtering was on average 32% and after u_*-filtering 60% and 48% for upper and lower u_* threshold estimate, respectively. Filling of gaps in half-hourly NEE data is necessary to obtain complete time series for the calculation of daily averages or balances such as monthly or seasonal sums. The following three gap-filling methods are implemented in `REddyProc`.

10 ### 2.2.1 Look-up tables

In the look-up table (LUT) approach, the fluxes are binned by the meteorological conditions within a certain time window. Within the chosen time window and respective bin, each meteorological variable deviates less than a fixed margin to ensure similar meteorological conditions. The missing value of the flux is then calculated as the average value of the binned records and its uncertainty estimated from their standard deviation.

15 The original LUT of Falge et al. (2001) consisted of fixed periods over a year, while in `REddyProc` the meteorological conditions are sampled with a moving window around the gap to be filled.

2.2.2 Mean diurnal course

NEE fluxes have a mean diurnal course (MDC) that follows the course of the sun with only respiration during nighttime and a combination of respiration and photosynthesis during daytime. This autocorrelation of the fluxes is exploited by taking the
20 average value at the same time of day within a moving time window of adjacent days (Falge et al., 2001). In `REddyProc`, the same time of day includes also the fluxes of the adjacent hour (±1 hour).

Though the MDC method only showed a medium performance in the gap filling comparison for net carbon fluxes by (Moffat et al., 2007), it has the advantage that this approach can be used even if no meteorological information is available.

2.2.3 Marginal distribution sampling

25 The so-called marginal distribution sampling (MDS) by (Reichstein et al., 2005b) exploits the covariation of the fluxes with the meteorological variables and their temporal autocorrelation based on the two methods LUT and MDC described above.

The filling of each half-hourly NEE with the MDS algorithm depends on the availability of the meteorological data of Rg, Tair, and VPD: 1) If all three meteorological variables are available, LUT will be used with default margins of 50 $\mathrm{Wm^{-2}}$, 2.5 °C, and 5.0 hPa, respectively. 2) If Tair or VPD are missing, only the variable Rg will be used. 3) If no meteorology is available, the gaps are filled with MDC. Following a specific sampling procedure, the MDS algorithm increases the number of days in the vicinity of the gap until there are enough data points (at least two) for gap-filling. A more detailed description with a flow diagram is provided in Supplement 1.

The MDS algorithm is optimized for carbon dioxide and water fluxes and can also be used to estimate the uncertainty of the half-hourly fluxes. In the comparison of gap-filling methods by (Moffat et al., 2007), the MDS algorithm performed well for different artificial gap scenarios ranging from single half-hours to several days. Due to its flexibility in dealing with missing meteorological input data and its fast and highly automated routines available as online tool (BGC16, section 3), the MDS gap-filling method has been widely used.

2.3   Flux-partitioning methods

The gross fluxes of GPP into the land system and R_{eco} out of the land system are the two opposing parts of NEE: NEE = R_{eco} - GPP. Availability of GPP and R_{eco} is pivotal as they are the two biggest terms of the carbon cycle (e.g. Chapin et al., 2006; Jung et al., 2011). Moreover, understanding their sensitivity to environmental drivers (e.g., radiation, temperature, and soil moisture) is important to interpret land-atmosphere interactions and to improve earth systems models (Reichstein et al., 2012). Therefore several methods were developed to partition NEE into these two components (Reichstein et al., 2005c; Lasslop et al., 2010; Moffat, 2012; Wehr and Saleska, 2015; Desai et al., 2008; Stoy et al., 2006).

The two most widely used methods are the so-called night-time partitioning and day-time partitioning (Reichstein et al., 2012). The night-time partitioning (Reichstein et al., 2005c) relies on the temperature response function of nighttime NEE fluxes that are representative of R_{eco}. It assumes that this relationship is applicable also to daytime data. The relationship is then used to predict R_{eco} from measured temperature and GPP is computed as a difference between R_{eco} and NEE. This method is currently the most widely used approach. Alternatively, the day-time partitioning (Lasslop et al., 2010) fits a model to observations of daytime NEE and global radiation, accounting for the effects of radiation and VPD on GPP as well as the effects of temperature on R_{eco}.

2.3.1   Nighttime flux-partitioning

The method of Reichstein et al. (2005c) estimates a temporally varying respiration-temperature relationship from night-time data. First night-time data is selected by a threshold of Rg < 10 $\mathrm{Wm^{-2}}$, which is congruent with the BGC online tool (BGC16, section 3), but differs from the 20 $\mathrm{Wm^{-2}}$ reported in Reichstein et al. (2005c).

Next, temperature sensitivity, E_0 of the Lloyd and Taylor (1994) relationship (1) is estimated by fitting the model to successive 15 day periods of night-time data, and the resulting E_0 series is aggregated to an annual estimate.

$$R_{eco}(T) = R_{Ref} \exp\left[E_0 \left(\frac{1}{T_{Ref} - T_0} - \frac{1}{T - T_0}\right)\right], \tag{1}$$

where T_0 is kept constant at -46.02°C (Lloyd and Taylor, 1994) and where reference temperature T_{Ref} is 15°C, which is congruent with the BGC online tool (BGC16, section 3) but differs from the 10°C reported in (Reichstein et al., 2005c). For robustness each fit is repeated on a trimmed data set excluding records with residuals outside the 5%-95% residual distribution. The annual aggregate is the mean across the 3 valid estimates having the lowest uncertainty in the fit. Single estimates of E_0 are considered valid, if there were minimum of 6 records, temperature ranged across at least 5°C, and estimates were inside the range of 30 to 450 K.

Subsequently, the respiration at reference temperature, R_{Ref}, is re-estimated from night-time data using the annual E_0 temperature sensitivity estimate for 7-day windows shifted consecutively for 4 days. The estimated value is then assigned to the central time-point of the 4 day period and linearly interpolated between periods. Hence, the obtained respiration-temperature relationship varies across time.

Finally, R$_{eco}$ is estimated for both day- and night-time from the temporarily varying R$_{eco}$ temperature relationship and daytime GPP is computed as R$_{eco}$ - NEE.

2.3.2 Daytime flux-partitioning

The method of (Lasslop et al., 2010) models NEE using the common rectangular hyperbolic light-response curve (LRC) (Falge et al., 2001):

$$NEE = \frac{\alpha \beta R_g}{\alpha R_g + \beta} + \gamma, \tag{2}$$

where α ($\mu mol\,CO_2\,J^{-1}$) is the canopy light utilization efficiency and represents the initial slope of the light-response curve, β ($\mu mol\,CO_2\,m^{-2}s^{-1}$) is the maximum CO_2 uptake rate of the canopy at infinite Rg, and γ ($\mu mol\,CO_2\,m^{-2}s^{-1}$) is a term accounting for ecosystem respiration. The hyperbolic light-response curve is modified to account for the temperature dependency of respiration after Gilmanov et al. (2003) by setting respiration γ to the Lloyd and Taylor respiration model (Lloyd and Taylor, 1994) (1). Further, the constant parameter β in (2) is replaced by an exponential decreasing function (Körner, 1995) at higher VPD values (3).

$$\beta = \begin{cases} \beta_0 \exp\left[-k(\text{VPD} - \text{VPD}_0)\right] & \text{if VPD} > 10 \text{ hPa} \\ \beta_0 & \text{otherwise} \end{cases}, \tag{3}$$

where the VPD_0 threshold is 10 hPa in accordance with earlier findings at the leaf level (Körner, 1995) ignoring potential vegetation specific differences.

Parameter T_0 in (1) was fixed as in the night-time partitioning (section 2.3.1). Parameter T_{Ref} was fixed in each window to the median temperature within the window. The other parameters $(E_0, R_{Ref}, \alpha, \beta_0, k)$ of the model are estimated by the following steps. 1) A time varying temperature sensitivity E_0 is estimated from night-time data for a window shifted by two days. 2) The E_0 estimates are smoothed across successive windows by fitting a Gaussian Process (Rasmussen and Williams, 2006; Menzer et al., 2013) using the `mlegp` R-package that also estimates uncertainty of the smoothed E_0. Next, a prior respiration, R_{Ref}, for reference temperature $T_{Ref} = 15$ °C is re-estimated from night-time data for each window with smoothed E_0. 3) Parameters of the rectangular hyperbolic light-response curve (R_{Ref}, α, β_0, k) are fitted using only day-time data and the previously determined temperature sensitivity (E_0) for each window. 4) Finally, for each NEE record, GPP and R_{eco} are estimated with the parameter set of the previous valid window and the parameters of the next valid window, and the two results are interpolated linearly by the time difference to the window centers. Supplement 1 reports necessary technical details to these steps.

Note, that contrary to the night-time based flux-partitioning, both GPP and R_{eco} are model predictions and do not add up exactly to observed NEE.

3 Benchmarking `REddyProc` post-processing steps

The post-processing steps implementations of `REddyProc` were benchmarked with the post-processing tools widely used in the FLUXNET processing. Specifically, `REddyProc` (version 0.8.1) u_*-filtering results were compared with results by a C-implementation from Dario Papale (Papale et al., 2006), here referred to as DP06. Results of `REddyProc` (version 1.1.2) gap-filling and flux-partitioning were compared with results obtained by the 2016-version web based tool provided by the Max Planck Institute for Biogeochemistry, Jena, best described in Reichstein et al. (2005a). The tool was accessed in 2016 (29.07.2016) and is hereafter referred as BGC16. Annually and monthly aggregated values, here, refer to the mean across all valid values in a month or a year, which can differ from real annual or monthly budgets in the presence of large gaps. The first section describes the dataset used for benchmarking for each processing step implemented in the package. Within each of the following sections for the processing steps, subsections describe differences in the code, report the results of benchmarking, and discuss them. Supplement 1, additionally, provides more detailed results and statistics.

3.1 Dataset used for benchmarking

Data of twenty-five sites with open data policy of the LaThuile FLUXNET dataset[3] were used for benchmarking. The sites are located in different climate zones and belong to a variety of plant functional types (Table 2) to guarantee testing different conditions (i.e. presence of snow, management such as cuts and crop rotation, sites disturbed) and ecosystem types (e.g. deciduous vs evergreen forests, grasslands and croplands). For each site the following variables were used: NEE already filtered for quality flags (Foken and Wichura, 1996), despiked and u_* filtered (Papale et al., 2006), random error of NEE computed as described by Reichstein et al. (2005a), Tair, Tsoil, Rg, and VPD. Moreover, NEE time series before the u_*-filtering and the u_* data were downloaded from AMERIFLUX and the European Flux Database to test the u_* threshold estimation. Finally, time series of gap-filled NEE (NEE_f), GPP partitioned with the night-time based method (GPP_{NT}) (Reichstein et al., 2005a) were downloaded from the LaThuile dataset, while GPP partitioned with the daytime method (GPP_{DT}) were computed with BGC16.

3.2 u_*-filtering: Benchmarking with DP06

Estimation of u_* threshold by REddyProc using the default moving point method (section 2.1.1) was benchmarked to estimation based on Papale's DP06 C-implementation (Papale et al., 2006). The benchmark applied a bootstrap sample of size 60 and recorded lower, median, and upper quantiles of 10%, 50% and 90% instead of the default 5% and 95% based on a larger sample size to save computing time.

The different estimates of the u_* threshold have potential consequences for the inferred fluxes. To explore these consequences, we used the different thresholds to mark gaps, gap-fill the data, and compute the annual NEE based on the gap-filled time series. NEE uncertainty was estimated by the difference between NEE based on lower quantile u_* and NEE based on upper quantile u_* estimate.

3.2.1 Differences in code

The biggest difference of REddyProc compared to DP06 is that REddyProc by default employs seasons that can span across years. With the *'Within one year'* classification option, which is employed also by DP06, records of December are associated to the same season as January and February of the *same* year. With the default *'continuous'* classification, seasons start the same as in DP06 by default in March, June, September, and December. However, December is treated in the same season as January and February of the *next* year to avoid discontinuities at year boundaries. The annual u_* threshold is then applied according to those continuous seasons spanning year boundaries. For example, the processing of 2014 data would by default use data from Winter 2014 (starting in December 2013) to Autumn 2014 (ending in November 2014). REddyProc also allows more flexibility with the *'User specified'* classification into seasons as explained below.

---

[3]www.fluxdata.org

Table 2. Description of sites and times used for benchmarking `REddyProc`.

[a] Abbreviations for land cover type from International Geosphere-Biosphere Programme (IGBP) classification: CRO: cropland, DBF: deciduous broadleaf forest, EBF: evergreen broadleaf forest, ENF: evergreen needleleaf forest, GRA: grassland, OSH: open shrubland, WET: permanent wetland, WSA: woody savanna.

[b] Abbreviations for climate from Köppen-Geiger classification: Af: equatorial, rainforest; BSh: hot arid steppe; Cfa: humid, warm temperate, hot summer; Cfb: humid, warm temperate, warm summer; Csa: summer dry, warm temperate, hot summer; Dfb: cold, humid, warm summer; Dfc: cold, humid, cold summer.

Site	Year	Lat	Lon	Land cover[a]	Climate[b]
CA-NS7	2004	56.64	-99.95	OSH	Dfc
CA-TP3	2005	42.71	-80.35	ENF	Dfb
CH-Oe2	2004	47.29	7.73	CRO	Cfb
DE-Hai	2002	51.08	10.45	DBF	Cfb
DE-Tha	1998	50.96	13.57	ENF	Cfb
DK-Sor	2006	55.49	11.65	DBF	Cfb
ES-ES1	2000	39.35	-0.32	ENF	Csa
ES-VDA	2005	42.15	1.45	GRA	Cfb
FI-Hyy	1998	61.85	24.29	ENF	Dfc
FI-Kaa	2001	69.14	27.30	WET	Dfc
FR-Gri	2006	48.84	1.95	CRO	Cfb
FR-Hes	1998	48.67	7.06	DBF	Cfb
FR-Lq1	2006	45.64	2.74	GRA	Cfb
FR-Lq2	2006	45.64	2.74	GRA	Cfb
FR-Pue	2003	43.74	3.60	EBF	Csa
IE-Dri	2004	51.99	-8.75	GRA	Cfb
IL-Yat	2005	31.34	35.05	ENF	BSh
IT-Amp	2004	41.90	13.61	GRA	Cfa
IT-MBo	2005	46.02	11.05	GRA	Cfb
IT-SRo	2001	43.73	10.28	ENF	Csa
PT-Esp	2004	38.64	-8.60	EBF	Csa
RU-Cok	2004	70.62	147.88	OSH	Dfc
SE-Nor	1997	60.09	17.48	ENF	Dfb
US-Ton	2004	38.43	-120.97	WSA	Csa
VU-Coc	2002	-15.44	167.19	EBF	Af

There are further slight differences between `REddyProc` and DP06. Both methods bin in a way such that the number of records in each bin is similar. If there are numerically equal u_* values, they are sorted into the same bin, resulting in bins with unequal record numbers. In DP06 sometimes no records are sorted into the subsequent bins hampering the moving point

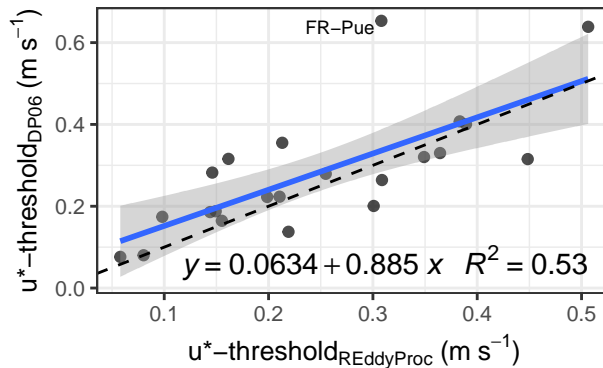



Figure 3. Large scatter but retained relationship between u_* threshold derived with DP06 versus u_* threshold derived with `REddyProc` across site-years as shown by a regression (solid line with shaded uncertainty bound) close to the 1:1 line (dashed).

detection. Contrary, the binning with `REddyProc` ensures that there are a minimum number of records in all bins. This often results in fewer bins than without numerically equal u_* values. Moreover, differing from the Papale C-implementation, `REddyProc` employs several more quality criteria. First, when comparing the threshold bin to NEE in the following bins, it makes sure that there are least 3 bins to infer a plateau in NEE. Next, when aggregating the thresholds of different temperature
5 classes to season, it ensures that a threshold was found in at least 20% of the temperature classes. For those seasons where no threshold could be determined, the annual estimate is used. When there are too few records within one year, a single season comprising all records is used for threshold estimation.

In difference to DP06, `REddyProc` re-samples data only within seasons instead across the entire year during the bootstrap, in order to protect periods of similar u_*-NEE relationship and to avoid seasonal biases in re-sampling.

10 **3.2.2 Benchmark results**

The general relationship in the estimation of the u_* threshold was retained between the two methods (Fig. 3), although there was a big scatter. The exceptionally high threshold value of $> 0.6 \, \mathrm{m\,s^{-1}}$ for site FR-Pue was very probably an overestimate by DP06. However, one has to remember that each estimate has a high uncertainty, and the differences between the two methods were in the range of this uncertainty (suppl 1) The estimate of the uncertainty of the u_* thresholds with `REddyProc` was,
15 however, only half of the uncertainty range estimated by DP06 (suppl 1). This increased precision was mainly due to the modified bootstrapping scheme, that respects the u_* seasons.

When propagating the differences in u_* to differences in annual NEE, there was no bias and very low scatter across sites between all the methods (Fig. 4), despite the differences in u_* threshold. The absolute differences in annual NEE between the methods were small (mostly $< 20 \, \mathrm{gC\,m^{-2}\,yr^{-1}}$), and mostly lower than half of the uncertainty range estimated from the bootstrap (suppl

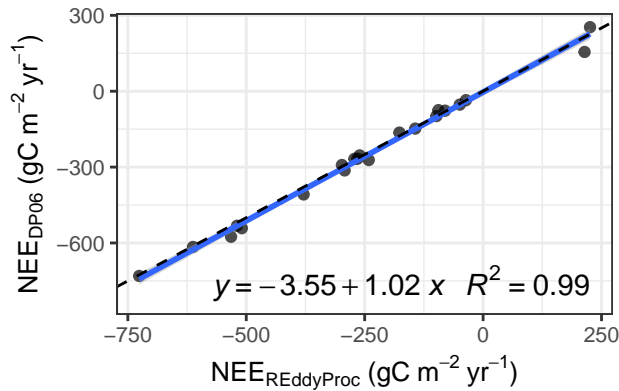



Figure 4. Strong correspondence in NEE based on u_* threshold estimated by `REddyProc` and NEE based on u_* threshold estimated by DP06 across site-years.

1) `REddyProc` estimates u_* thresholds with roughly double the precision compared to DP06, due to protecting seasons during bootstrap (suppl 1).

3.2.3   Discussion of u_* threshold estimation

The agreement between NEE based on u_* estimates of `REddyProc` moving point implementation and current FLUXNET
5 standard post-processing (DP06) (Fig. 4) indicates that the sensitivity of NEE to the u_* threshold estimate in the inferred ranges
is low, which also explains the large uncertainty of the u_* threshold estimate. One reason for the missing effect could be site
selection of this study without many sites affected by advection, which show limited saturation of the NEE - u_* dependence.
Since in such cases the filtering does not work properly anyway, it should not change the conclusions for NEE. Hence, we infer
that u_* estimates of both DP06 and `REddyProc` are appropriate due to the negligible effect on NEE sums. The agreement
10 implies that both methods can be interchanged in studies that are based on aggregated values, such as annual carbon budgets or
for upscaling, without the need to reprocess data.

However, the increase in estimated precision, i.e. lower standard deviation, of the u_* threshold estimate also yields an increase
in estimated precision of the annual NEE (suppl 1). This will lead to improved accuracy and usability of EC measurements and
any downstream, post-processed data products in model-data integration studies.

15 While the default seasons and their aggregation are in line with previous approaches, `REddyProc` allows to use site specific
knowledge to derive better threshold estimates. For example, if there is disturbance such as harvest, the u_* threshold is expected
to change and a different threshold should be applied for filtering before and after the disturbance. In this case the user can
define a season change at harvest date and use season-specific threshold estimates instead of the annually aggregated estimate
(section B7).

3.3 Gap-filling: Benchmark with BGC16

The gap-filling implementation of `REddyProc` was benchmarked with the BGC online tool (BGC16, section 3) that used pvWave-code from Reichstein et al. (2005c).

3.3.1 Differences in code

5   Compared to the BGC16, the new implementation of the MDS algorithm in `REddyProc` was not limited to single years but filled the gaps with a window moving continuously over all years in the input data. This had the advantage of smoother gap-filling over the end of year timestamp and will especially be of interest for sites in which vegetation is not dormant during this time. This new feature led to different, probably more realistic gap-filled NEE values in the beginning and end of the year.

There were also slight differences in the window size between the old and new code. For MDC, the old window size had a few
10   more intermediate day steps than the new implementation which affected longer gaps with missing meteorology. The default meteorological variables and margins for LUT (see Chapter 4.2.2 above) were the same in both implementations.

While `REddyProc` restricts gap-filling to interpolation of gaps, BGC16 also extrapolated missing records in periods without measurements.

3.3.2 Benchmark results and discussion for gap-filling

15   In the benchmark, `REddyProc` gap-filling was run using the same measured NEE as input that passed the QA/QC routines and u_*-filtering. The annually aggregated values comprised both, filled gaps and originally valid records.

`REddyProc` gap-filling results agreed with the results of BGC16. Few discrepancies at half-hourly time scale were mostly at longer gaps due to usage of fewer window sizes, as shown for the DE-Tha case (Fig. 5 top). At annually aggregated time scale, the agreement between methods was strong ($R^2 = 0.99$) (Fig. 5 bottom). The outlier of site RU-Cok is due to the availability of
20   only a few months of data for the whole year. While `REddyProc` filled gaps in the time period with data available, BGC16 extrapolated also into the time before and after. The seasonal cycle was well reproduced at each site (suppl 1).

The good agreement between NEE based on gap-filling by `REddyProc` and gap-filling by BGC16 (Fig. 5) imply that both gap-filling tools can be used interchangeably without the need to reprocess data.

3.4 Nighttime flux-partitioning: Benchmark BGC16

25   The night-time based flux-partitioning was benchmarked to the BGC online tool (BGC16, section 3), that used pvWave code developed by Reichstein et al. (2005c).

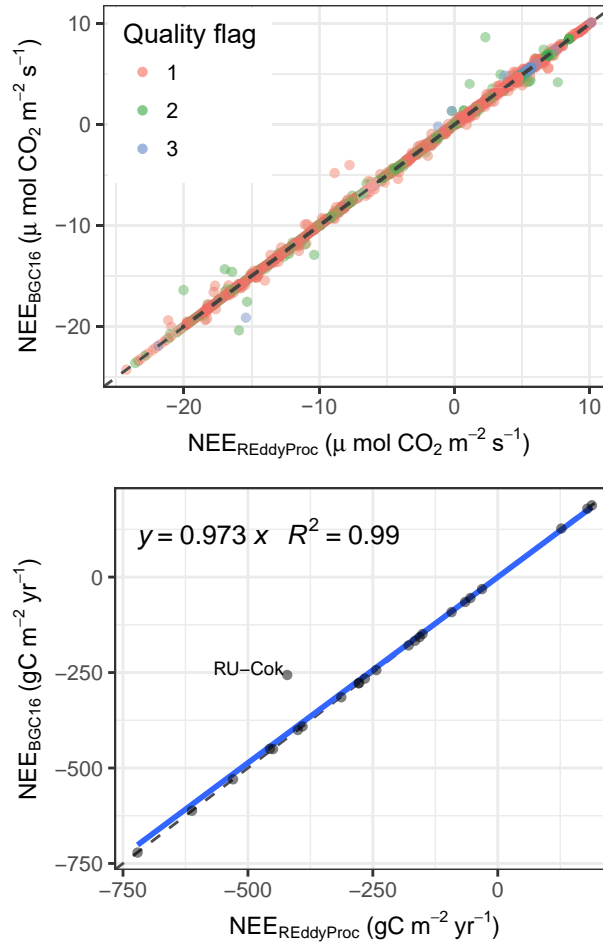



Figure 5. Predictions of NEE by `REddyProc` after gap-filling agree with BGC16 both at half-hourly values (top), shown for the DE-Tha 1998 use case, and annual means across sites (bottom). Larger quality flags are associated with larger window sizes.

3.4.1 Differences in code

The main features of the `REddyProc` implementation of the night-time based partitioning algorithm were very similar to BGC16, with using a reference temperature of 15°C and trimming the estimates of temperature sensitivity E_0 before aggregating them (section 2.3.1). `REddyProc` differed from BGC16 in computing the potential radiation that is used in subsetting the night-time data to derive E_0 and R_{Ref} (Reichstein et al., 2005). While `REddyProc` used the exact solar time for the calculation of the potential radiation, where the sun culminates exactly at noon, BGC16 used the local winter time which differs from the solar time depending on the location within the time zone.

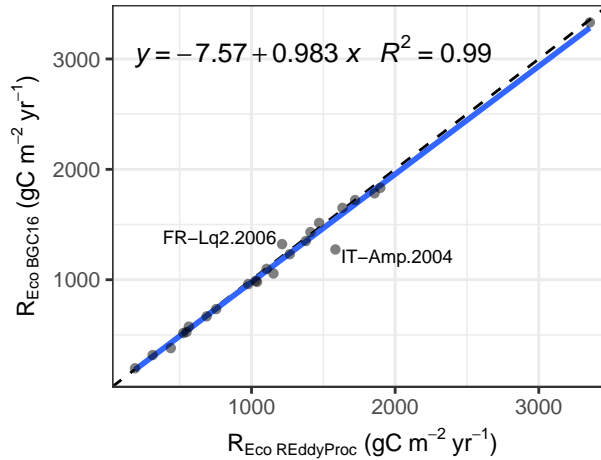



Figure 6. Predictions of annually aggregated ecosystem respiration, R_{eco} from `REddyProc` night-time partitioning agree with the predictions by BGC16.

3.4.2 Benchmark results and discussion for night-time flux-partitioning

Annual aggregated values of R_{eco} predicted by `REddyProc` were in very good agreement ($R^2 = 0.99$; slope ≈ 1) with BGC16 as shown in Fig. 6 and in supplement 1.

In order to evaluate the effects of the differences introduced in the code described above, we also computed R_{eco} by prescribing 5 in `REddyProc` either E_0, selection of night-time data, or both from BGC16 output. Results are reported in supplement 1 and showed that the most important factor affecting the R_{eco} computed with `REddyProc` was the different selection of night-time data, though the differences were almost negligible at annual time scale.

The two implementations agree very well for most sites at annual time-scale. Because of no systematic deviations across sites, the spatial upscaling of fluxes should not be affected by `REddyProc` implementation. However, for some sites, such as IT-Amp, 10 the quite large relative errors indicates problems related to selection of night-time data and problems due to large gaps in the dataset.

3.5 Daytime flux-partitioning: Benchmark with BGC16

The day-time flux-partitioning was benchmarked with results of the BGC online tool (BGC16, section 3), which is based on pvWave code developed by Lasslop et al. (2010) and used in the processing of the 2015 Fluxnet release (Pastorello et al., 2017).

3.5.1 Differences in code

BGC16 differed from `REddyProc` (section 2.3.2) mainly in aspects of separation of nighttime data, estimation of temperature sensitivity from night-time data, uncertainty estimation, treatment of missing values, and optimization library code.

While for separating nighttime data `REddyProc` used exact solar time where sun culminates exactly at noon, BGC16 used the local winter time zone time.

For the estimation of temperature sensitivity E_0 from night-time data BGC16 used a reference temperature of 15 °C, instead of the median temperature inside the window. Hence, it estimated stronger correlations between parameters for windows with a different temperature range. Moreover, it omitted smoothing of the estimated E_0 across time, often leading to large fluctuations of the E_0 estimates across few days (suppl 1), larger estimates of its uncertainty, and differences in subsequent estimation of LRC parameters.

For uncertainty estimation, BGC16 relied on the curvature of the LRC fit's optimum instead of a bootstrap procedure. Hence, it could not take into account the uncertainty of E_0 estimated from night-time data before the day-time LRC fit. Moreover, during interpolation of fluxes based on previous and next valid estimates, the distance weights differed. While `REddyProc` assignes the estimates to the time of the mean of valid record in a window, BGC16 assigned it to the start of the third day, also if there were only valid data for the first day in the window.

For weighting the records in the LRC fit, BGC16 used the raw estimated NEE uncertainty of each record. It did not check for high leverage of spurious low NEE uncertainty estimates. Its estimates, therefore, were in some windows very strongly influenced by a few records, and failed if a NEE uncertainty estimate of zero was provided. Moreover, when there were missing values or values below zero in provided NEE uncertainty, it set all uncertainty to 1, while `REddyProc` fills the gaps by setting the missing uncertainty to the maximum of 20% of respective NEE but at least 0.7 $\mu molCO_2 m^{-2} s^{-1}$.

Treatment of missing values was not considered by BGC16 and assumed to be handled prior to the processing. Hence, it did not handle missing VPD values and did not re-try the LRC fit without the VPD-effect in order to use also records with missing VPD. Moreover, as described above, when there were missing values of NEE uncertainty, weighting records in the LRC fit was omitted.

For compatibility with BGC16, the above code-differences can be disabled in `REddyProc`. But differences in optimization library code and specifically the conditions of non-convergence on scattered data could not be eliminated, which led to differences in results as shown in the following section.

3.5.2 Benchmark results for day-time partitioning

Annually GPP predictions of both implementations showed no significant bias across the test sites (Fig. 7), although, there was some scatter for the individual predictions. A similar scatter was observed when comparing the predictions of the default

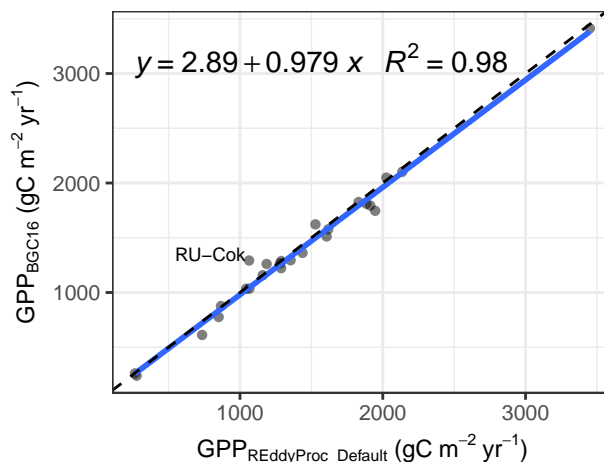



Figure 7. Prediction of annually aggregated GPP from `REddyProc` day-time partitioning agree with BGC16 across sites.

`REddyProc` options to the predictions with compatibility options. Most of the differences were caused by decreasing the unreasonable high influence of NEE records with small NEE uncertainty (suppl 1).

The largest differences in aggregated fluxes between implementations were due to the extrapolation of fitted parameters to periods where no parameter fits were obtained. In many of those cases, there were fits at the boundaries of these periods, whose
5   validity was questionable. Whether those fits passed the quality check or not had a large influence on the extrapolation and hence on the aggregated values. For example, at RU-Cok parameter estimates for valid periods agreed between implementations. However, no valid parameters could be obtained for winter months. While `REddyProc` reported missing values, BGC16 reported GPP values based on summer parameterizations also for periods further away from summer, which in turn lead to higher annual GPP estimates.

10   Uncertainty estimates of gross fluxes were larger with `REddyProc` due to accounting for uncertainty in temperature sensitivity estimates from night-time data (suppl 1).

3.5.3   Discussion of day-time flux-partitioning

Agreement between aggregated fluxes predicted by the day-time method and absence of bias for the test sites (Fig. 7) suggest that the methods can be used interchangeably for upscaling, although differences in results of influential sites can potentially
15   propagate to differences in upscaled estimates. `REddyProc` provides a quality flag for the results of the day-time partitioning, that allows excluding less reliable data in upscaling studies. For the results associated with good quality flags, we set stronger trust in the `REddyProc`-based estimates.

The day-time flux-partitioning is quite sensitive to the details of LRC fit. Small changes in treatment of extreme or missing NEE uncertainty estimates or changes in pre-processing and treatment of missing values cause different estimates of LRC parameters and propagate to predicted fluxes of GPP and R_{eco}. Although, we put much effort in trying to reproduce the results of BGC16, we were not able to eliminate all differences, especially in the subtle details in the parameter optimization library codes. The differences in predicted half-hourly fluxes, however, average out across sites and across time (suppl. 1) making this issue less severe at larger scales.

The estimated uncertainties are even more sensitive. Both implementations occasionally produce unreasonably high outliers that affect the aggregated values. REddyProc, in general, estimates higher uncertainties of predicted fluxes, because it accounts for uncertainty in temperature sensitivity. Note, that the uncertainty introduced to annually aggregated fluxes due to flux-partitioning is smaller than uncertainty due to uncertain u_* threshold estimate. Hence, differences or difficulties in uncertainty estimation caused by flux-partitioning do affect conclusions of the overall uncertainty estimates to a lesser extent.

4 Conclusions

The REddyProc software provides a set of tools for the CO_2-focussed post-processing of eddy covariance flux data including u_*-filtering, gap-filling and flux-partitioning and propagation of the uncertainty from the u_*-filtering to the gap-filled NEE and partitioned GPP and R_{eco}.

The freely available R-package enables researchers to integrate the flux data processing into their own, offline environment or work stream without the need of uploading data. This seamless integration allows to improve overall workflow, to speed up processing routines and to ultimately generate cleaner, reproducible scientific results.

The compatibility of the implemented methods to the available standard tools provides continuity of the data analysis, when adopting REddyProc for processing EC-data. REddyProc can closely reproduce results of the widely used BGC online tool (BGC16, section 3).

A number of enhancements provide more flexibility to the user in the processing of their data. For instance, the new processing allows to treat multi-year data without breaks at annual boundaries that can significantly affect sites in the southern hemisphere or sites characterized by vegetation activity in winter. Another new feature of REddyProc is the flexibility to define different seasons for the application of the u_*-filtering and gap-filling routines, which is important for sites with discontinuous surface cover associated with snow melt, dry seasons, or harvest.

Sensitivity of the results to subtle details of the implementation, however, call for caution when interpreting results. This is especially true for u_* threshold estimation and the day-time flux-partitioning, and especially for data with long gaps.

Continued integration of new methodological developments into the package will support research using EC data. We strive to provide new developments in a basic and extensible manner, while paying attention to compatibility with results of reference implementations.

In summary, research using (half-)hourly eddy covariance data can benefit from building blocks for standardized and extensible
5   post-processing provided by `REddyProc`.

Appendix A: The `REddyProc` package

The `REddyProc` processing tool is freely available in two options: a) online as a web-service[4] with a smaller range of user option, and b) as a package of the open-source R environment with a larger set of user options and with each of the steps and methods available independently.

10 The `REddyProc` package can be installed by typing at the R-terminal:

```
install.packages("REddyProc")
library(REddyProc)
?REddyProc
```

Alternatively, there is a automatically built docker image (Metzger et al., 2017) at docker hub allowing to run RStudio with REddyProc from a browser without any other installation besides docker. Regarding installation issues and docker images we refer the reader to the github project homepage[5].

15 Some general principles and choices in the design of `REddyProc` that lead to trade-offs between robustness and flexibility are explained in supplement 1.

Appendix B: Example application

This section reports an example R session using `REddyProc`. Code is shown in a shaded area and corresponding output with monospace font.

20 **B1   Importing the half-hourly data**

The workflow starts with importing the half-hourly data. The example reads a text file with data of the year 1998 from the DE-Tha site and converts the separate decimal columns year, day, and hour to a POSIX timestamp column. Next, it initializes the `sEddyProc` class.

---

[4] www.bgc-jena.mpg.de/bgi/index.php/Services/REddyProcWeb

[5] github.com/bgctw/REddyProc

```
#+++ load libraries used in this vignette
library(REddyProc)
library(dplyr)
#+++ Load data with 1 header and 1 unit row from (tab-delimited) text file
fileName <- getExamplePath('Example_DETha98.txt', isTryDownload = TRUE)
EddyData.F <- if (length(fileName)) fLoadTXTIntoDataframe(fileName) else
  # or use example dataset in RData format provided with REddyProc
  Example_DETha98
#+++ Add time stamp in POSIX time format
EddyDataWithPosix.F <- fConvertTimeToPosix(EddyData.F, 'YDH',Year.s = 'Year'
    ,Day.s = 'DoY',Hour.s = 'Hour')
#+++ Initalize R5 reference class sEddyProc for post-processing of eddy data
#+++ with the variables needed for post-processing later
EddyProc.C <- sEddyProc$new('DE-Tha', EddyDataWithPosix.F,
    c('NEE','Rg','Tair','VPD', 'Ustar'))
```

B2   Estimating the u_* threshold distribution

The second step is the estimation of the distribution of u_* thresholds to identify periods of low friction velocity (u_*), where NEE is biased low. Discarding periods with low u_* is one of the largest sources of uncertainty in aggregated fluxes. Hence, several quantiles of the distribution of the uncertain u_* threshold are estimated by a bootstrap.

The friction velocity, u_*, needs to be in a column named "Ustar" of the input dataset.

```
uStarTh <- EddyProc.C$sEstUstarThresholdDistribution(
  nSample = 100L, probs = c(0.05, 0.5, 0.95))
#filter(uStarTh, aggregationMode == "year")
select(uStarTh, -seasonYear)

##   aggregationMode  season    uStar        5%       50%       95%
## 1          single    <NA> 0.4162500 0.3658125 0.4414583 0.6197520
## 2            year    <NA> 0.4162500 0.3658125 0.4414583 0.6197520
## 3          season 1998001 0.4162500 0.3658125 0.4414583 0.6197520
## 4          season 1998003 0.4162500 0.3198846 0.4037981 0.5615500
## 5          season 1998006 0.3520000 0.3192476 0.3858571 0.4393708
## 6          season 1998009 0.3369231 0.2195485 0.3906458 0.5235175
## 7          season 1998012 0.1740000 0.1829667 0.4189286 0.6180982
```

The output reports u_* estimates of 0.42 for the orignal data and 0.37, 0.44, 0.62 for lower, median, and upper quantile of the estimated distribution. The threshold can vary between periods of different surface roughness, e.g. before and after harvest. Therefore, there are estimates for different time periods of the year, called seasons, reported as different rows. These season-estimates can be aggregated to entire years or to a single value across years, reported by rows with corresponding aggregation mode.

The subsequent post processing steps will be repeated using the three quantiles of the u_* distribution. They require to specify a u_*-threshold for each season and a suffix to distinguish the outputs related to different thresholds.

For this example of an evergreen forest site, the same annually aggregated u_* threshold estimate will be chosen for each of the seasons within a year. In order to distinguish the automatically generated columns, the column names of the estimation results are written to variable uStarSuffixes.

```
uStarThAnnual <- usGetAnnualSeasonUStarMap(uStarTh)[-2]
uStarSuffixes <- colnames(uStarThAnnual)[-1]
print(uStarThAnnual)
```

```
##     season       U05       U50      U95
## 1 1998001 0.3658125 0.4414583 0.619752
## 2 1998003 0.3658125 0.4414583 0.619752
## 3 1998006 0.3658125 0.4414583 0.619752
## 4 1998009 0.3658125 0.4414583 0.619752
## 5 1998012 0.3658125 0.4414583 0.619752
```

B3 Gap-filling

The second post-processing step is filling the gaps using information of the valid data. In this case, the same annual u_* threshold estimate is used for each season, as described above, and the uncertainty will be computed also for valid records (FillAll).

```
EddyProc.C$sMDSGapFillAfterUStarDistr('NEE',
   UstarThres.df = uStarThAnnual,
   UstarSuffix.V.s = uStarSuffixes,
     FillAll = TRUE
)
```

The screen output (not shown here) already shows that the u_*-filtering and gap-filling was repeated for each given estimate of the u_* threshold , i.e. column in uStarThAnnual, with marking 22% to 38% of the data as gap.

For each of the different u_* threshold estimates, a separate set of output columns of filled NEE and its uncertainty is generated, distinguished by the suffixes given with `uStarSuffixes`. Suffix "_f" denotes the filled value and "_fsd" the estimated standard devation of its uncertainty.

```
grep("NEE_.*_f$",names(EddyProc.C$sExportResults()), value = TRUE)
grep("NEE_.*_fsd$",names(EddyProc.C$sExportResults()), value = TRUE)
```

```
## [1] "NEE_U05_f" "NEE_U50_f" "NEE_U95_f"
## [1] "NEE_U05_fsd" "NEE_U50_fsd" "NEE_U95_fsd"
```

B4   Partitioning net flux into GPP and R_{eco}

The third post-processing step is partitioning the net flux (NEE) into its gross components GPP and R_{eco}. The partitioning algorithm needs a precise criteria between night-time and day-time. Therefore, geographical coordinates and the time zone need to be provided to allow computing exact solar time of sunrise and sunset. Further, missing values in the used meteorological data need to be filled.

```
EddyProc.C$sSetLocationInfo(Lat_deg.n = 51.0, Long_deg.n = 13.6, TimeZone_h.n = 1)
EddyProc.C$sMDSGapFill('Tair', FillAll.b = FALSE)
EddyProc.C$sMDSGapFill('VPD', FillAll.b = FALSE)
```

Now we are ready to invoke the partitioning, here by the night-time approach, for each of the several filled NEE columns.

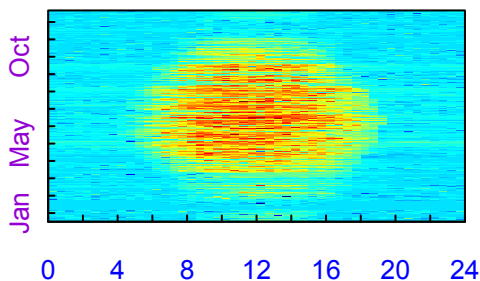
```
#variable uStarSuffixes was defined above at the end of uStar threshold estimation
resPart <- lapply(uStarSuffixes, function(suffix){
                  EddyProc.C$sMRFluxPartition(Suffix.s = suffix)
            })
```

The results are stored in columns `Reco` and `GPP_f` modified by the respective u_* threshold suffix.

```
grep("GPP.*_f$|Reco",names(EddyProc.C$sExportResults()), value = TRUE)
```

```
## [1] "Reco_U05"  "GPP_U05_f" "Reco_U50"  "GPP_U50_f" "Reco_U95"  "GPP_U95_f"
```

The visualizations of the results by a fingerprint plot gives a compact overview.

```
EddyProc.C$sPlotFingerprintY('GPP_U50_f', Year.i = 1998)
```

B5 Estimating the uncertainty of aggregated results

First, the mean of the GPP across all the year is computed for each u_*-scenario and converted from $\mu\text{mol}\,CO_2\,\text{m}^{-2}\text{s}^{-1}$ to $\text{gC}\,\text{m}^{-2}\text{yr}^{-1}$.

```
FilledEddyData.F <- EddyProc.C$sExportResults()
#suffix <- uStarSuffixes[2]
GPPAggCO2 <- sapply( uStarSuffixes, function(suffix) {
    GPPHalfHour <- FilledEddyData.F[[paste0("GPP_",suffix,"_f")]]
    mean(GPPHalfHour, na.rm = TRUE)
})
molarMass <- 12.011
GPPAgg <- GPPAggCO2 * 1e-6 * molarMass * 3600*24*365.25
print(GPPAgg)
```

```
5   ##      U05       U50       U95
    ## 1894.097 1956.090 1985.817
```

The difference between those aggregated values is a first estimate of uncertainty range in GPP due to uncertainty of the u_* threshold.

```
(max(GPPAgg) - min(GPPAgg)) / median(GPPAgg)
```

In this run of the example a relative error of about 4.7% is inferred.

10 For a better but more time consuming uncertainty estimate, specify a larger sample of u_* threshold values, for each repeat the post-processing, and compute statistics from the larger sample of resulting GPP columns. This can be achieved by specifying a larger sequence of quantiles when calling `sEstUstarThresholdDistribution` in section B2.

```
EddyProc.C$sEstUstarThresholdDistribution(
  nSample = 200, probs = seq(0.025,0.975,length.out = 39) )
```

B6 Storing the results in a csv-file

The results still reside inside the `sEddyProc` class. To export them to an R Data.frame, the newly generated columns need to be appended to the columns with the original input data. Then this data.frame is written to a text file in a temporary directory.

```r
FilledEddyData.F <- EddyProc.C$sExportResults()
CombinedData.F <- cbind(EddyData.F, FilledEddyData.F)
fWriteDataframeToFile(CombinedData.F, 'DE-Tha-Results.txt', Dir.s = tempdir())
```

B7 Specifying seasons where u_* thresholds differs

With changing surface roughness, e.g. on harvest or leaf-fall, also the u_*-NEE relationship can change. Therefore the u_* threshold needs to be re-estimated at different times of the year, called seasons. The default uses continuous seasons, for details see section 3.2.1. In order to yield results corresponding to DP06, the user can specify `seasonFactor.v = usCreateSeasonFactorMonthWithinYear( EddyData.C$sDATA$sDateTime, startMonth= c(3,6,9,12))` as an argument to routine `sEstUstarThreshold`. By default the annual aggregate of the season thresholds, i.e. maximum across seasons, is used to identify unfavorable conditions, but the seasonal estimates can also be used instead.

Moreover, the users can specify also other user-defined seasons, e.g. when harvest dates are known (see package vignette DEGebExample). They can create a grouping by specifying exact starting days of the periods by function `usCreateSeasonFactorYdayYear`, or they can provide a column with the data that indicates e.g. the same group for two wet seasons. Each season is associated to the year corresponding to the center day between first and last day of the season.

With all methods, there is a required minimum number of 160 records within a season. If there are too few records, the data of the seasons within one year are combined and the u_* threshold for this seasons is set to the estimate obtained for the data of the entire year.

Acknowledgements. This work used eddy covariance data acquired and shared by the FLUXNET community, including these networks: AmeriFlux, AfriFlux, AsiaFlux, CarboAfrica, CarboEuropeIP, CarboItaly, CarboMont, ChinaFlux, Fluxnet-Canada, GreenGrass, ICOS, KoFlux, LBA, NECC, OzFlux-TERN, TCOS-Siberia and USCCC. The ERA-Interim reanalysis data are provided by ECMWF and processed by LSCE. The FLUXNET eddy covariance data processing and harmonization was carried out by the European Fluxes Database Cluster, AmeriFlux Management Project, and Fluxdata project of FLUXNET, with the support of CDIAC and ICOS Ecosystem Thematic Center, and the OzFlux, ChinaFlux and AsiaFlux offices.

The authors acknowledge Dario Papale, Gilberto Pastorello and Trevor F Keenan for the discussions on the benchmarking of REddyProc and PvWave code. M. M. and M. R. acknowledge the Alexander Von Humbold foundation that funded part of this research activity with the Max Planck Research Preis to Markus Reichstein. M. M. acknowledge the MSCA-ITN project TRUSTEE.

L.Š. was supported by the Ministry of Education, Youth and Sports of the Czech Republic within the CzeCOS program, grant number LM2015061, and within the National Sustainability Program I (NPU I), grant number LO1415.

References

Aubinet, M., Grelle, A., Ibrom, A., Rannik, U., Moncrieff, J., Foken, T., Kowalski, A. S., Martin, P. H., Berbigier, P., Bernhofer, C., Clement, R., Elbers, J., Granier, A., Grünwald, T., Morgenstern, K., Pilegaard, K., Rebmann, C., Snijders, W., Valentini, R., and Vesala, T.: Estimates of the annual net carbon and water exchange of European forests: The EUROFLUX methodology, Advances in Ecological Research, 30, 114–177, 2000.

Aubinet, M., Vesala, T., and Papale, D., eds.: Eddy Covariance, Springer Netherlands, https://doi.org/10.1007/978-94-007-2351-1, 2012.

Baldocchi, D., Chu, H., and Reichstein, M.: Inter-annual variability of net and gross ecosystem carbon fluxes: A review, Agricultural and Forest Meteorology, https://doi.org/10.1016/j.agrformet.2017.05.015, 2017.

Barr, A., Richardson, A., Hollinger, D., Papale, D., Arain, M., Black, T., Bohrer, G., Dragoni, D., Fischer, M., Gu, L., et al.: Use of change-point detection for friction–velocity threshold evaluation in eddy-covariance studies, Agricultural and forest meteorology, 171, 31–45, https://doi.org/10.1016/j.agrformet.2012.11.023, 2013.

Chapin, F. S., Woodwell, G. M., Randerson, J. T., Rastetter, E. B., Lovett, G. M., Baldocchi, D. D., Clark, D. A., Harmon, M. E., Schimel, D. S., Valentini, R., Wirth, C., Aber, J. D., Cole, J. J., Goulden, M. L., Harden, J. W., Heimann, M., Howarth, R. W., Matson, P. A., McGuire, A. D., Melillo, J. M., Mooney, H. A., Neff, J. C., Houghton, R. A., Pace, M. L., Ryan, M. G., Running, S. W., Sala, O. E., Schlesinger, W. H., and Schulze, E. D.: Reconciling carbon-cycle concepts, terminology, and methods, Ecosystems, 9, 1041–1050, 2006.

Davison, A. and Hinkley, D.: Bootstrap methods and their application, Cambridge University Press, Cambridge, 1997.

Desai, A. R., Richardson, A. D., Moffat, A. M., Kattge, J., Hollinger, D. Y., Barr, A., Falge, E., Noormets, A., Papale, D., Reichstein, M., and Stauch, V. J.: Cross-site evaluation of eddy covariance GPP and RE decomposition techniques, Agricultural and Forest Meteorology, 148, 821 – 838, https://doi.org/https://doi.org/10.1016/j.agrformet.2007.11.012, 2008.

Efron, B. and Tibshirani, R.: Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy, Statistical science, pp. 54–75, 1986.

Falge, E., Baldocchi, D., Olson, R. J., Anthoni, P., Aubinet, M., Bernhofer, C., Burba, G., Ceulemans, R., Clement, R., Dolman, H., Granier, A., Gross, P., Grünwald, T., Hollinger, D., Jensen, N. O., Katul, G., Keronen, P., Kowalski, A., Ta Lai, C., Law, B. E., Meyers, T., Moncrieff, J., Moors, E., Munger, J. W., Pilegaard, K., Rannik, U., Rebmann, C., Suyker, A., Tenhunen, J., Tu, K., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: Gap filling strategies for defensible annual sums of net ecosystem exchange, Agric. For. Meteorol., 107/1, 29–43, 2001.

Foken, T. and Wichura, B.: Tools for quality assessment of surface-based flux measurements, Agricultural and forest meteorology, 78, 83–105, https://doi.org/10.1016/0168-1923(95)02248-1, 1996.

Foken, T., Göockede, M., Mauder, M., Mahrt, L., Amiro, B., and Munger, W.: Post-field data quality control, in: Handbook of micrometeorology, pp. 181–208, Springer, 2004.

Foken, T., Leuning, R., Oncley, S. R., Mauder, M., and Aubinet, M.: Corrections and Data Quality Control, pp. 85–131, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-007-2351-1_4, 2012.

Gilmanov, T. G., Johnson, D. A., and Saliendra, N. Z.: Growing season CO2 fluxes in a sagebrush-steppe ecosystem in Idaho: bowen ratio/energy balance measurements and modeling, Basic and Applied Ecology, 4, 167 – 183, https://doi.org/http://dx.doi.org/10.1078/1439-1791-00144, 2003.

Gorsel, E. V., Leuning, R., Cleugh, H. A., Keith, H., and Suni, T.: Nocturnal carbon efflux: reconciliation of eddy covariance and chamber measurements using an alternative to the u.-threshold filtering technique, Tellus B: Chemical and Physical Meteorology, 59, 397–403, 2007.

Göckede, M., Rebmann, C., and Foken, T.: A combination of quality assessment tools for eddy covariance measurements with footprint modelling for the characterisation of complex sites, Agricultural and Forest Meteorology, 127, 175–188, https://doi.org/10.1016/j.agrformet.2004.07.012, 2004.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J.,
5  Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, Journal of Geophysical Research: Biogeosciences, 116, n/a–n/a, 2011.

Körner, C.: Leaf Diffusive Conductances in the Major Vegetation Types of the Globe, pp. 463–490, Springer Berlin Heidelberg, Berlin,
10  Heidelberg, https://doi.org/10.1007/978-3-642-79354-7_22, 1995.

Lasslop, G., Reichstein, M., Papale, D., Richardson, A., Arneth, A., Barr, A., Stoy, P., and Wohlfahrt, G.: Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation, Global Change Biology, 16, 187–208, https://doi.org/10.1111/j.1365-2486.2009.02041.x, 2010.

Lloyd, J. and Taylor, J.: On the temperature dependence of soil respiration, Functional ecology, 8, 315–323, 1994.

15 Massman, W. J. and Lee, X.: Eddy covariance flux corrections and uncertainties in long-term studies of carbon and energy exchanges, Agricultural and Forest Meteorology, 113, 121–144, 2002.

Menzer, O., Moffat, A. M., Meiring, W., Lasslop, G., Schukat-Talamazzini, E. G., and Reichstein, M.: Random errors in carbon and water vapor fluxes assessed with Gaussian Processes, Agricultural and Forest Meteorology, 178-179, 161–172, https://doi.org/10.1016/j.agrformet.2013.04.024, 2013.

20 Metzger, S., Durden, D., Sturtevant, C., Luo, H., Pingintha-Durden, N., Sachs, T., Serafimovich, A., Hartmann, J., Li, J., Xu, K., and Desai, A. R.: eddy4R 0.2.0: a DevOps model for community-extensible processing and analysis of eddy-covariance data based on R, Git, Docker, and HDF5, Geoscientific Model Development, 10, 3189–3206, https://doi.org/10.5194/gmd-10-3189-2017, 2017.

Moffat, A. M.: A new methodology to interpret high resolution measurements of net carbon fluxes between terrestrial ecosystems and the atmosphere, Ph.D. thesis, Friedrich-Schiller-Universität Jena, 2012.

25 Moffat, A. M., Papale, D., Reichstein, M., Hollinger, D. Y., Richardson, A. D., Barr, A. G., Beckstein, C., Braswell, B. H., Churkina, G., Desai, A. R., et al.: Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes, Agricultural and Forest Meteorology, 147, 209–232, https://doi.org/10.1016/j.agrformet.2007.08.011, 2007.

Muggeo, V.: Segmented: An R package to fit regression models with broken-line relationships, R News, 8 (1), 20: 25, 2008.

Muggeo, V. M.: Estimating regression models with unknown break-points, Statistics in medicine, 22, 3055–3071, 2003.

30 Papale, D., Reichstein, M., Aubinet, M., Canfora, E., Bernhofer, C., Kutsch, W., Longdoz, B., Rambal, S., Valentini, R., Vesala, T., and et al.: Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation, Biogeosciences, 3, 571 – 583, https://doi.org/10.5194/bg-3-571-2006, 2006.

Pastorello, G., Papale, D., Chu, H., Trotta, C., Agarwal, D., Canfora, E., Baldocchi, D., and Torn, M.: The FLUXNET2015 dataset: The longest record of global carbon, water, and energy fluxes is updated, Eos Trans. AGU, https://doi.org/10.1029/2017EO071597, 2017.

35 Rasmussen, C. E. and Williams, C. K. I.: Gaussian Processes for Machine Learning, The MIT Press, 2006.

Rebmann, C., Kolle, O., Heinesch, B., Queck, R., Ibrom, A., and Aubinet, M.: Data Acquisition and Flux Calculations, pp. 59–83, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-007-2351-1_3, 2012.

Reichstein, M., Falge, E., Baldocchi, D., Papale, D., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Gilmanov, T., Granier, A., et al.: On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm, Global Change Biology, 11, 1424–1439, https://doi.org/10.1111/j.1365-2486.2005.001002.x, 2005a.

Reichstein, M., Katterer, T., Andren, O., Ciais, P., Schulze, E. D., Cramer, W., Papale, D., and Valentini, R.: Temperature sensitivity of decomposition in relation to soil organic matter pools: critique and outlook, Biogeosciences, 2, 317–321, https://doi.org/10.5194/bg-2-317-2005, 2005b.

Reichstein, M., Subke, J. A., Angeli, A. C., and Tenhunen, J. D.: Does the temperature sensitivity of decomposition of soil organic matter depend upon water content, soil horizon, or incubation time?, Global Change Biology, 11, 1754–1767, 2005c.

Reichstein, M., Stoy, P. C., Desai, A. R., Lasslop, G., and Richardson, A. D.: Partitioning of Net Fluxes, pp. 263–289, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-007-2351-1_9, 2012.

Stoy, P. C., Katul, G. G., Siqueira, M. B., Juang, J.-Y., Novick, K. A., Uebelherr, J. M., and Oren, R.: An evaluation of models for partitioning eddy covariance-measured net ecosystem exchange into photosynthesis and respiration, Agricultural and Forest Meteorology, 141, 2 – 18, https://doi.org/10.1016/j.agrformet.2006.09.001, 2006.

Wehr, R. and Saleska, S.: An improved isotopic method for partitioning net ecosystem–atmosphere CO 2 exchange, Agricultural and Forest Meteorology, 214, 515–531, 2015.