



中国通量观测研究联盟
ChinaFLUX



生态系统通量观测缺失数据插补理论与技术

贾昕（博士、教授、博士生导师）

北京林业大学，水土保持学院

xinjia@bjfu.edu.cn

https://www.researchgate.net/profile/Xin_Jia5



数据插补不是万能药!!!

- 再好的插补算法也不能完全补救数据缺失的损失
- 并不是所有变量都能够被合理插补（如反射辐射、土壤水分，**重复的重要性！**）
- 减小数据插补造成的不确定性（**站点维护的重要性！**）



数据缺失的原因

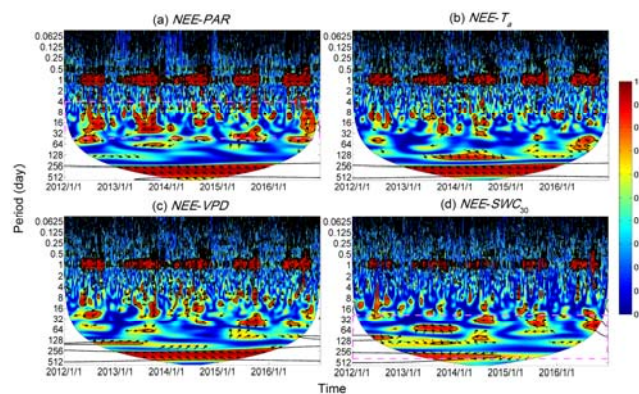
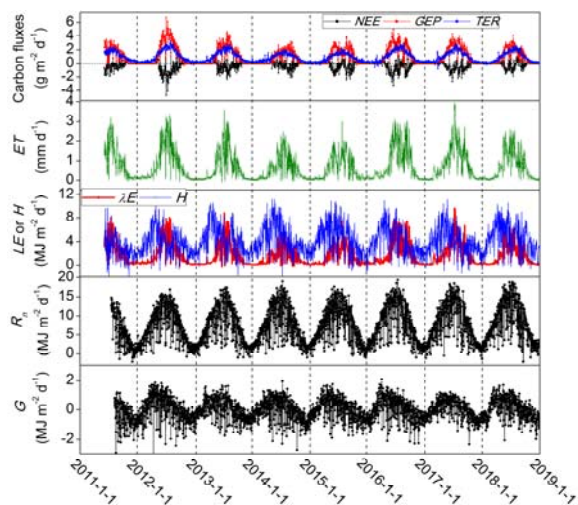
- 仪器故障
- 供电故障
- 不利天气
- 仪器校准
- 人为因素
- 质量控制

EC平均年数据覆盖约60-80% (Falge et al., 2001a,b; Moffat et al., 2007)



为什么要插补缺失值？

1. 估算一段时间（如一年）的通量累计值；
2. 进行特定数据分析（如小波分析）



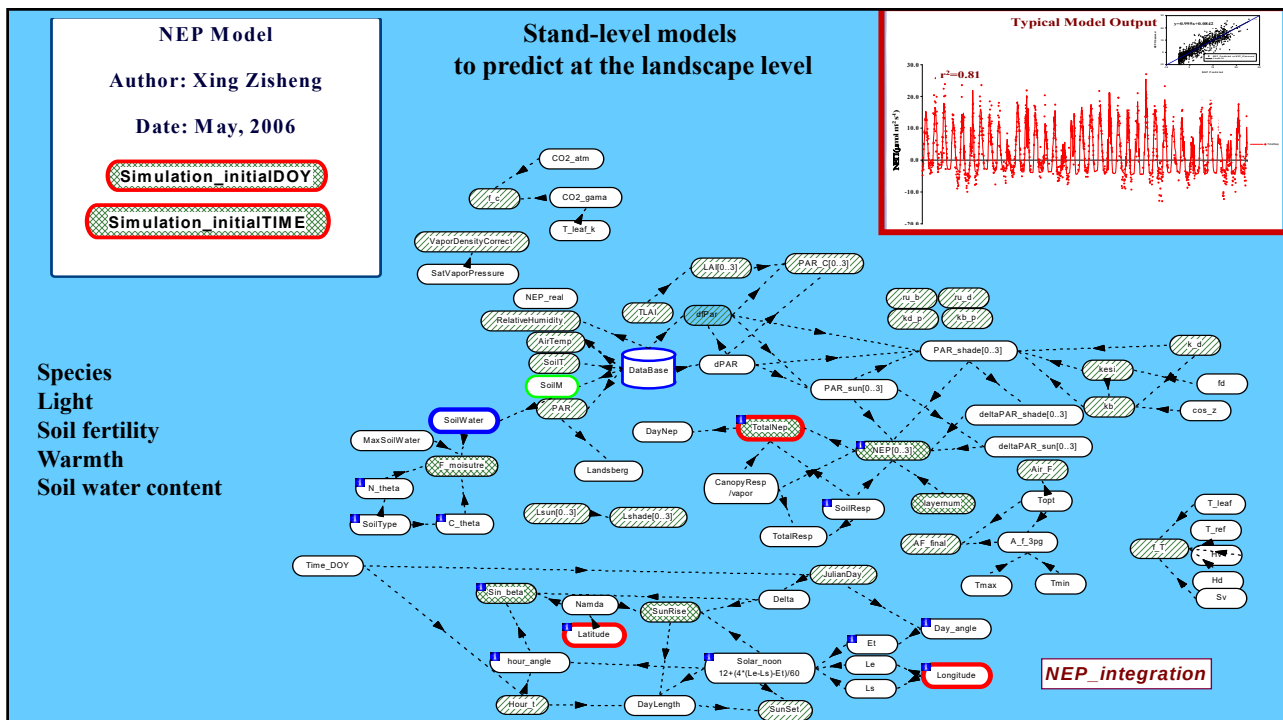
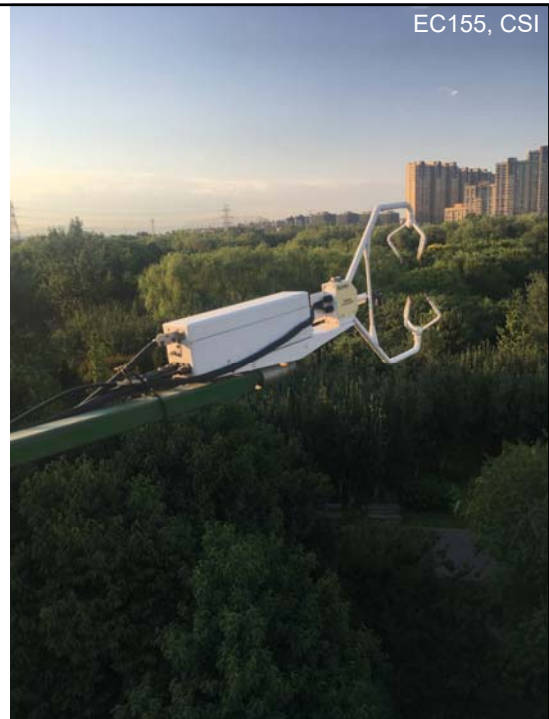
Jia et al. 2018 AFM

数据插补理论

if Gap > 2h

- 根据内在变化规律，如MDV/MDC/MDT, LUT, **Marginal Distribution Sampling (MDS)** (Falge et al. 2001a,b; Reichstein et al. 2005; Moffat et al. 2007; Wutzler et al. 2018)
- 经验模型，如**NLR**、**ANN** (Amiro et al. 2006; Moffat et al. 2007; Jia et al. 2014, 2016a,b, 2018a,b)
- 机理模型，如Xing et al. 2007, 2008a,b

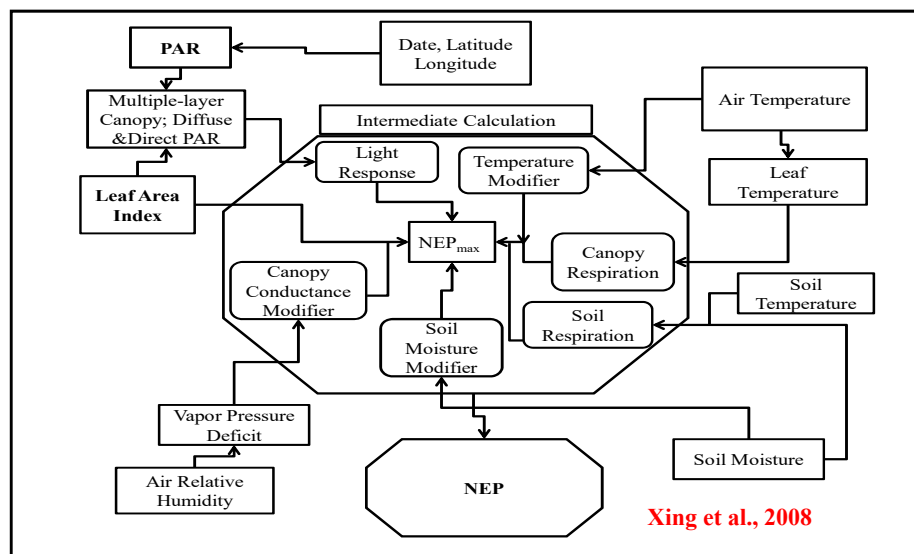
if Gap ≤ 2h
then 线性插值



Evolution of NEP model initially intended for gap filling



Full Model



Model parameters

Model Parameter	Function	Value derivation	Values	Range ^a
Ecoresp_beta	Ecosystem respiration curve	Set according to the literature	0.017825	n/a
Ecoresp_alpha	Ecosystem respiration curve		308	n/a
optimum_soilwater	Field capacity for sandy cla		0.32	n/a
optimum_temperature	Tree response to air temperature	Species-specific and set based on NWL 2006 data	17.5	n/a
alpha	Light use efficiency for multi-layered canopy		0.0016 ^b	0-1
NEP_max	Maximum possible NEP as		34.4	0-100
NEPgama	Light compensation point of balsam fir	Determined by fitting NEP model to data from both sites and for all years	202.9	0-300
krt	Influence of temperature on photosynthesis		- ^c	0-3000
pco2	Rate of CO ₂ exchange between the atmosphere and forest canopy		-	0-5
kvapord	Relates stomatal conductance	Site-specific parameters	-	0-5
R_eco	Maximum ecosystem respiration		-	0-20

n/a= not applicable

^a the range is provided to NEP model to constrain feasible solution

^b determined through pre-calibration of the NEP model (see Results and Discussion section); alpha is in $\mu\text{mol CO}_2 (\mu\text{mol PAR})^{-1}$; NEP_max, in $\mu\text{mol CO}_2 \text{ m}^{-2} \text{ s}^{-1}$; and NEPgama, in $\mu\text{mol PAR m}^{-2} \text{ s}^{-1}$

^c variable parameter values to be set according to a non-linear fit of NEP model to NEP and eco-hydrometeorological data from all sites and years

Gap filling strategies for defensible annual sums of net ecosystem exchange

We review several methods of gap filling and apply them to data sets available from the EUROFLUX and AmeriFlux databases. The methods are based on mean diurnal variation (MDV), look-up tables (LookUp), and nonlinear regressions (Regr.), and the impact of different gap filling methods on the annual sum of F_{NEE} is investigated. The difference between annual F_{NEE} filled by MDV compared to F_{NEE} filled by Regr. ranged from -45 to $+200 \text{ g C m}^{-2}$ per year (MDV-Regr.). Comparing LookUp and Regr. methods resulted in a difference (LookUp-Regr.) ranging from -30 to $+150 \text{ g C m}^{-2}$ per year.

We also investigated the impact of replacing measurements at night, when turbulent mixing is insufficient. The nighttime correction for low friction velocities (u_*) shifted annual F_{NEE} on average by $+77 \text{ g C m}^{-2}$ per year, but in certain cases as much as $+185 \text{ g C m}^{-2}$ per year.

Our results emphasize the need to standardize gap filling-methods for improving the comparability of flux data products

Falge et al. 2001a AFM

Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes

statistical metrics. The performance of the gap filling varied among sites and depended on the level of aggregation (native half-hourly time step versus daily), long gaps were more difficult to fill than short gaps, and differences among the techniques were more pronounced during the day than at night.

The non-linear regression techniques (NLRs), the look-up table (LUT), marginal distribution sampling (MDS), and the semi-parametric model (SPM) generally showed good overall performance. The artificial neural network based techniques (ANNs) were generally, if only slightly, superior to the other techniques. The simple interpolation technique of mean diurnal variation (MDV) showed a moderate but consistent performance. Several sophisticated techniques, the dual unscented Kalman filter (UKF), the multiple imputation method (MIM), the terrestrial biosphere model (BETHY), but also one of the ANNs and one of the NLRs showed high biases which resulted in a low reliability of the annual sums, indicating that additional development might be needed. An uncertainty analysis comparing the estimated random error in the 10 benchmark datasets with the artificial gap residuals suggested that the techniques are already at or very close to the noise limit of the measurements. Based on the techniques and site data examined here, the effect of gap filling on the annual sums of NEE is modest, with most techniques falling within a range of $\pm 25 \text{ g C m}^{-2} \text{ year}^{-1}$.

Moffat et al. 2007 AFM

Gap filling strategies for long term energy flux data sets

At present a network of over 100 field sites are measuring carbon dioxide, water vapor and sensible heat fluxes between the biosphere and atmosphere, on a nearly continuous basis. Gaps in the long term measurements of evaporation and sensible heat flux must be filled before these data can be used for hydrological and meteorological applications. We adapted methods of gap filling for NEE (net ecosystem exchange of carbon) to energy fluxes and applied them to data sets available from the EUROFLUX and AmeriFlux eddy covariance databases. The average data coverage for the sites selected was 69% and 75% for latent heat (λE) and sensible heat (H). The methods were based on mean diurnal variations (half-hourly binned means of fluxes based on previous and subsequent days, MDV and look-up tables for fluxes during assorted meteorological conditions (LookUp), and the impact of different gap filling methods on the annual sum of λE and H is investigated. The difference between annual λE filled by MDV and λE filled by LookUp ranged from -120 to 210 MJ m^{-2} per year, i.e. -48 to $+86 \text{ mm}$ per year, or -13 to $+39\%$ of the annual sum. For annual sums of H differences between -140 and $+140 \text{ MJ m}^{-2}$ per year or -12 to $+19\%$ of the annual sum were found. © 2001 Elsevier Science B.V. All rights reserved.

Falge et al. 2001b AFM

数据插补工具

- 研究人员独立开发代码（MATLAB, Python, Fortran）
- 网页工具
(<https://www.bgcjena.mpg.de/bgi/index.php/Services/REddyProcWeb>)
- R软件REddyProc包



indow of ± 7 days. Similar meteorological conditions are present when R_g , T_{air} and VPD do not deviate by more than 50 W m^{-2} , 2.5°C , and 5.0 hPa , respectively. If no

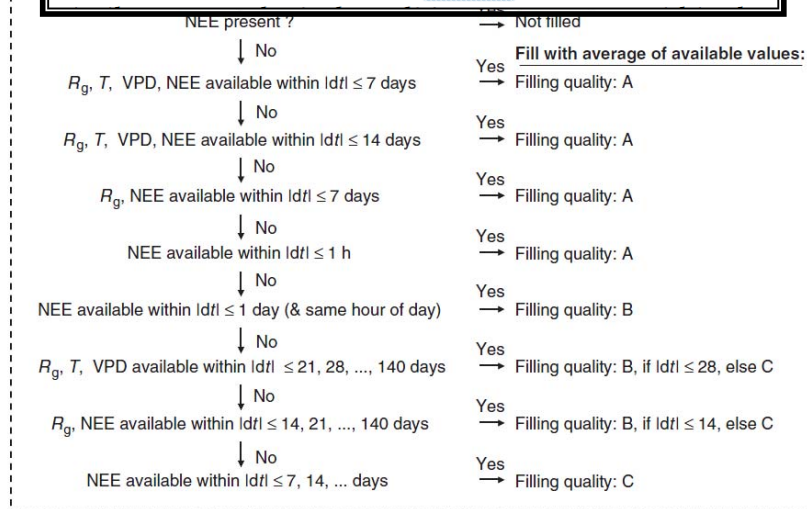
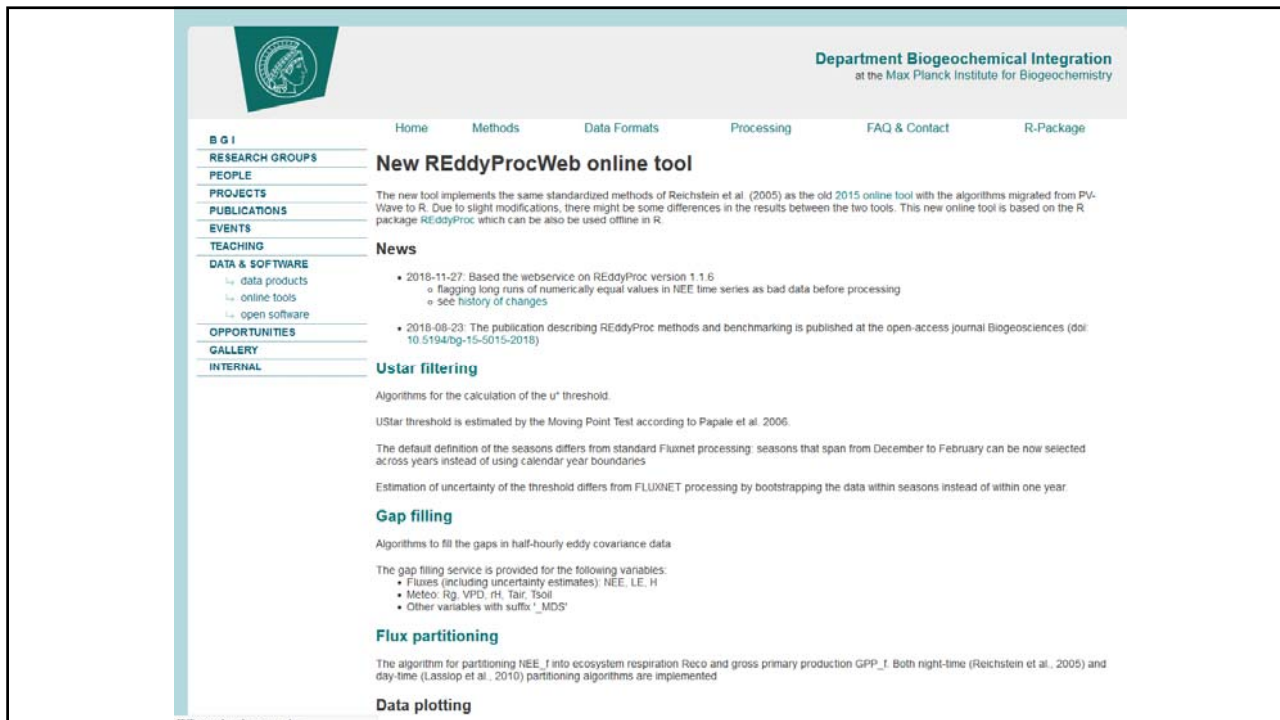


Fig. A1 Flow diagram of the gap-filling algorithm used in this study. Abbreviations: NEE, net ecosystem CO_2 exchange; R_g , global radiation; T , air temperature; VPD, vapour pressure deficit; $|dt|$, absolute difference in time. Filling qualities: A, high; B, medium; C, low.



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

(Wutzler et al. 2018 BG)

Type Package

Version 1.2

Title Post Processing of (Half-)Hourly Eddy-Covariance Measurements

Description Standard and extensible Eddy-Covariance data post-processing (Wutzler et al. (2018) <doi:10.5194/bg-15-5015-2018>)

includes

uStar-filtering, gap-filling, and flux-partitioning.

The Eddy-Covariance (EC) micrometeorological technique quantifies continuous exchange fluxes of gases, energy, and momentum between an ecosystem and the atmosphere.

It is important for understanding ecosystem dynamics and upscaling exchange fluxes. (Aubinet et al. (2012) <doi:10.1007/978-94-007-2351-1>).

This package inputs pre-processed (half-)hourly data and supports further processing.

First, a quality-check and filtering is performed based on the relationship between measured flux and friction

velocity (u^*) to discard biased data

(Papale et al. (2006) <doi:10.5194/bg-3-571-2006>).

Second, gaps in the data are filled based on information from environmental conditions

(Reichstein et al. (2005) <doi:10.1111/j.1365-2486.2005.001002.x>).

Third, the net flux of carbon dioxide is partitioned

into its gross fluxes in and out of the ecosystem by night-time

based and day-time based approaches

(Lasslop et al. (2010) <doi:10.1111/j.1365-2486.2009.02041.x>).

Details

A detailed example of the processing can be found in the [useCase.vignette](#).

A first overview of the REdDyProc functions:

These functions help with the preparation of your data for the analysis:

- Loading text files into dataframes: [fLoadTXTIntoDataFrame](#)
- Preparing a proper time stamp: [fConvertTimeToPosix](#)
- Calculating latent variables, e.g. VPD: [fCalcVPDfromRHandTair](#)

Then the data can be processed with the [sEddyProc-class](#) R5 reference class:

- Initializing the R5 reference class: [sEddyProc_initialize](#)
- Estimating the turbulence criterion, Ustar threshold, for omitting data from periods of low turbulence: Functions [sEddyProc_sEstUstarThreshold](#) and [sEddyProc_sEstUstarThresholdDistribution](#).
- Gap filling: [sEddyProc_sMDSGapFill](#) and [sEddyProc_sMDSGapFillAfterUstar](#).
- Flux partitioning based on Night-Time: [sEddyProc_sMRFluxPartition](#)
- Flux partitioning based on Day-Time: [sEddyProc_sGLFluxPartition](#)

Processing across different scenarios of u^* threshold estimate is supported by

- Estimating the turbulence criterion, Ustar threshold, for omitting data from periods of low turbulence: [sEddyProc_sEstimateUstarScenarios](#) and associated
 - query the thresholds to be used [sEddyProc_sGetUstarScenarios](#)
 - set the thresholds to be used [sEddyProc_sSetUstarScenarios](#)
 - query the estimated thresholds all different aggregation levels [sEddyProc_sGetEstimatedUstarThresholdDistribution](#)
- Gap-Filling: [sEddyProc_sMDSGapFillUstarScen](#)
- Flux partitioning based on Night-Time (Reichstein 2005): [sEddyProc_sMRFluxPartitionUstarScen](#)
- Flux partitioning based on Day-Time (Lasslop 2010): [sEddyProc_sGLFluxPartitionUstarScen](#)
- Flux partitioning based on modified Day-Time (Keenan 2019): [sEddyProc_sTKFluxPartitionUstarScen](#)

Before or after processing, the data can be plotted:

- Fingerprint: [sEddyProc_sPlotFingerprint](#)
- Half-hourly fluxes and their daily means: [sEddyProc_sPlotHHFluxes](#)
- Daily sums (and their uncertainties): [sEddyProc_sPlotDailySums](#)
- Diurnal cycle: [sEddyProc_sPlotDiurnalCycle](#)

A complete list of REdDyProc functions be viewed by clicking on the [Index](#) link at the bottom of this help page.

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

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

REddyProc package (REddyProc)

Post Processing of (Half-)Hourly Eddy-Covariance Measurements

Description

Standard and extendable Eddy-Covariance data post-processing including uStar-filtering, gap-filling, and flux-partitioning (Wutzler <doi:10.5194/bg-15-5015-2018>).

The Eddy-Covariance (EC) micrometeorological technique quantifies continuous exchange fluxes of gases, energy, and moment an ecosystem and the atmosphere. It is important for understanding ecosystem dynamics and upscaling exchange fluxes. (Aub <doi:10.1007/978-94-007-2351-1>).

This package inputs pre-processed (half-)hourly data and supports further processing. First, a quality-check and filtering is performed on the relationship between measured flux and friction velocity (u_{Star}) to discard based data (Papele et al. (2006) <doi:10.5194/2006>).

Second, gaps in the data are filled based on information from environmental conditions (Reichstein et al. (2005) <doi:10.1111/j.1248-2005.00102.x>).

Third, the net flux of carbon dioxide is partitioned into its gross fluxes in and out of the ecosystem by night time based and day approaches (Lasslop et al. (2010) <doi:10.1111/j.1365-2446.2009.02041.x>).

A general description and an online tool based on this package can be found here: <https://www.bgc-jena.mpg.de/bg/index.php/Services/REddyProcWeb>

Details

A detailed example of the processing can be found in the [useCase.vignette](#)

REddyProc typical workflow

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 Tharandt site and converts the separate decimal columns year, day, and hour to a POSIX timestamp column. Next, it initializes the `sEddyProc` class.

```
#### 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 <- if (length(fileName)) fLoadTXTIntoDataFrame(fileName) else
  # or use example dataset in RData format provided with REdDyProc
  Example_DETha98

#### Replace long runs of equal NEE values by NA
EddyData <- filterLongRuns(EddyData, "NEE")
#### Add time stamp in POSIX time format
EddyDataWithPosix <- fConvertTimeToPosix(
  EddyData, 'YDH', Year = 'Year', Day = 'DoY', Hour = 'Hour') %>%
  filterLongRuns("NEE")

#### Initialize R5 reference class sEddyProc for post-processing of eddy data
#### with the variables needed for post-processing later
EProc <- sEddyProc$new(
  'DE-Tha', EddyDataWithPosix, c('NEE', 'Rg', 'Tair', 'VPD', 'Ustar'))
```

A fingerprint-plot of the source half-hourly shows already several gaps. A fingerprint-plot is a color-coded image of the half-hourly fluxes by daytime on the x and and day of the year on the y axis.

```
EProc$sPlotFingerprintY('NEE', Year = 1998)
```

1998

```
install.packages("REddyProc")
library(REddyProc)
vignette("DEGebExample")
```

Crop example demonstrating multiple years and user defined uStar-Seasons

```
#isDevelMode <- TRUE
if (!exists("isDevelMode")) library(REddyProc)
set.seed(0815) # for reproducible results
```

First, the data is loaded. This example uses data that has been downloaded from <http://www.europe-fluxdata.eu> and preprocessed by fLoadEuroFlux16, where the DateTime Column has been created, and the variables renamed to the BGC-convention (e.g. Tair instead of Ta).

```
data(DEGebExample)
summary(DEGebExample)
```

```
##      DateTime              NEE              Ustar
## Min. :2004-01-01 00:30:00 Min. : -49.919 Min. :0.0000
## 1st Qu.:2004-10-01 00:22:30 1st Qu.: -1.864 1st Qu.:0.0640
## Median :2005-07-02 00:15:00 Median : 0.635 Median :0.1490
## Mean :2005-07-02 00:15:00 Mean : -1.935 Mean :0.1884
## 3rd Qu.:2006-04-02 00:07:30 3rd Qu.: 1.834 3rd Qu.:0.2800
## Max. :2007-01-01 00:00:00 Max. : 19.008 Max. :2.0450
##      NA's :21849 NA's :1149
##      Tair              rH              Rg
## Min. : -16.710 Min. : 15.87 Min. : 0.00
## 1st Qu.: 3.360 1st Qu.: 66.61 1st Qu.: 0.00
## Median : 9.970 Median : 79.10 Median : 2.04
## Mean : 9.664 Mean : 75.24 Mean : 124.71
## 3rd Qu.: 15.520 3rd Qu.: 87.07 3rd Qu.: 176.03
## Max. : 34.680 Max. :100.00 Max. :1046.03
##      NA's :1
```

VPD was not given with the original dataset and is calculated from Tair and rH.

```
DEGebExample$VPD <- fCalcVPDfromRHandTair(DEGebExample$rH, DEGebExample$Tair)
EProc <- sEddyProc$new("DE-Geb", DEGebExample, c("NEE", "Rg", "Tair", "VPD", "Ustar"))
EProc$setLocationInfo(LatDeg = 51.1, LongDeg = 10.9, TimezoneHour = 1) #Location of Gebesee
```

The data file (.txt)

Year	DoY	Hour	NEE	LE	H	Rg	Tair	Tsoil	rH	VPD	Ustar	ms-1
-	-	-	umolm-2s-1		Wm-2	Wm-2	Wm-2	degC	degC	%	hPa	
2012	1	0.5	0.062326853	NaN	NaN	0	-1.849	-5.800000191	64.33	1.903025589	0.048973552	
2012	1	1	-0.165436291	-0.778009294	-15.64580659	0	-2.723	-6.199999809	69.2	1.540025377	0.067567371	
2012	1	1.5	-0.126259831	-0.154872185	NaN	0	-3.104	-6.599999905	69.22999	1.495413882	0.113586233	
2012	1	2	0.262742768	0.713130663	NaN	0	-3.553	-6.900000095	72.86	1.275397472	0.139318479	
2012	1	2.5	0.523579766	0.612032201	0.359986048	0	-2.839	-7.300000191	70.02	1.486123646	0.048600718	
2012	1	3	-0.625205804	0.612032201	0.359986048	0	-2.903	-7.599999905	70.86	1.437607024	0.078936704	
2012	1	3.5	-1.360858563	4.661005899	23.37437795	0	-2.683	-7.800000191	69.36	1.536595265	0.104057126	
2012	1	4	0.380567857	-0.307229415	-0.273915455	0	-2.195	-8.100000381	67.59	1.685372314	0.07376325	
2012	1	4.5	1.474014017	-4.085713297	-31.6707434	0	-2.529	-8.300000191	68.36999	1.604523871	0.165340273	
2012	1	5	0.912719987	-6.539702622	-29.79400384	0	-3.325	-8.600000381	70.15	1.426920805	0.083928596	
2012	1	5.5	NaN	-1.095457187	NaN	0	-4.441	-8.699999809	74.21	1.13357506	0.055634092	
2012	1	6	NaN	0.512234438	NaN	0	-3.84	-8.899999619	70.45	1.359041509	0.172926146	
2012	1	6.5	0.178161311	-0.233896668	-7.665083882	0	-4.062	-9.100000381	72.54	1.241993902	0.047280415	
2012	1	7	0.6639848	0.096683429	-21.39934296	0	-3.631	-9.399999619	69.06	1.445487472	0.108012508	
2012	1	7.5	-0.164257392	3.428261776	NaN	0	-4.501	-9.5	71.45	1.249209043	0.0293947	
2012	1	8	-0.03732808	0.74226946	0.113004522	0	-4.157	-9.699999809	70.89	1.307229193	0.042399865	
2012	1	8.5	-0.430982036	1.574053984	NaN	9.08	-2.95	-9.800000191	66.41	1.651344936	0.065911951	
2012	1	9	-0.020893184	-1.683268857	-3.726658827	48.96	-2.876	-10	66.31	1.66542929	0.137036884	
2012	1	9.5	-2.083207238	0.675026328	-4.224244421	111.2	-2.332	-10.10000038	64.45	1.82996385	0.299611888	
2012	1	10	0.805983493	9.987954901	14.41320663	174.2	-2.239	-9.899999619	65.03	1.812572591	0.560396495	
2012	1	10.5	0.37324969	9.773311177	60.3504843	257.2	-1.608	-9.5	61.98	2.064823983	0.633162697	
2012	1	11	-0.274293101	13.59481358	115.5701785	332.8	-1.238	-8.899999619	55.94	2.458959764	0.791812395	
2012	1	11.5	NaN	15.80876768	139.5562811	401.6	-0.834	-8.100000381	51.87	2.766967952	0.786137794	
2012	1	12	NaN	24.05910124	134.0851843	459.2	-0.431	-7.199999809	46.95	3.141082414	0.83634247	
2012	1	12.5	NaN	20.53811245	154.8865028	498.9	0.004	-6.400000095	41.65	3.566222989	0.825796476	
2012	1	13	1.015615595	21.37417242	176.4805345	510.4	0.051	-5.599999905	39.95	3.682695545	0.82946154	
2012	1	13.5	0.695218662	21.55171503	175.0144737	513	0.123	-5	40.9	3.643462459	0.846063536	
2012	1	14	1.333478774	21.39182311	161.366286	492.1	0.299	-4.5	40.27	3.729685657	0.722337688	
2012	1	14.5	1.093555623	27.76241407	160.9788818	455	0.385	-4	38.18	3.884356421	0.727514348	
2012	1	15	0.531958911	25.71434207	144.592734	398	0.332	-3.599999905	37.62	3.904499204	0.785776592	

A code template

```

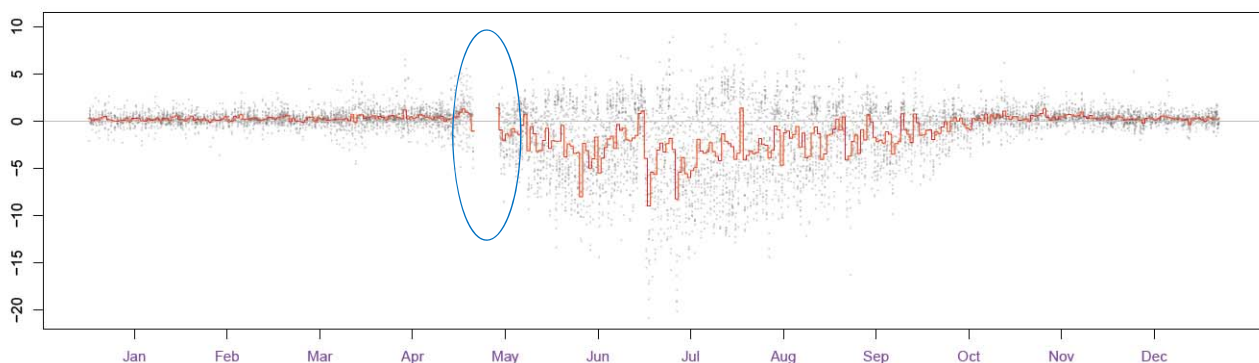
1 library(ReddyProc)
2 library(dplyr)
3 setwd("C:/Users/hp/Documents")
4 EddyData.F<-floadTXTIntoDataFrame("Test data.txt")
5 EddyDataWithPosix.F<-fconvertTimeToPosix(EddyData.F,'YDH',Year='Year',Day='DoY',Hour='Hour')
6 EddyProc.C<-sEddyProc$new('yc1', EddyDataWithPosix.F,c('NEE','Rg','Tair','VPD','Ustar'))
7 uStarTh<- (EddyProc.C$EstUstarThreshold())$uStarTh
8 select(uStarTh, ~seasonYear)
9 uStarThAnnual<-usGetAnnualSeasonUstarMap(uStarTh)
10 uStarThAnnual
11 uStarSuffixes<-colnames(uStarThAnnual)[-1]
12 uStarSuffixes
13 EddyProc.C$MDSGapFillAfterUstar(fluxVar="NEE",uStarVar="Ustar",uStarTh=uStarThAnnual,uStarSuffix=uStarSuffixes,FillAll=TRUE)
14 grep("NEE_.*_fs$",names(EddyProc.C$ExportResults()),value=TRUE)
15 grep("NEE_.*_fsd$",names(EddyProc.C$ExportResults()),value=TRUE)
16 #EddyProc.C$PlotFingerprintY('NEE_uStar_f',Year=2012)
17 EddyProc.C$SetLocationInfo(LatDeg=37.68,LongDeg=107.23,TimeZoneHour=8)
18 EddyProc.C$MDSGapFill('Tair',FillAll=FALSE)
19 EddyProc.C$MDSGapFill('VPD',FillAll=FALSE)
20 EddyProc.C$MDSGapFill('Rg',FillAll=FALSE)
21 EddyProc.C$MRF fluxPartition(Suffix=uStarSuffixes)
22 grep("GPP.*_f$|Reco",names(EddyProc.C$ExportResults()),value=TRUE)
23 #EddyProc.C$PlotFingerprintY('GPP_uStar_f',Year=2012)
24 FilledEddyData.F<-EddyProc.C$ExportResults()
25 FilledEddyData.F
26 CombinedData.F<-cbind(EddyData.F, FilledEddyData.F)
27 fwriteDataFrameToFile(CombinedData.F,'Test data output.xls',Dir="")

```

Half-hourly fluxes and daily means at yc1:
NEE_uStar_orig (umolm⁻²s⁻¹)

`EddyProc.C$PlotHHFluxes("NEE_uStar_orig")`

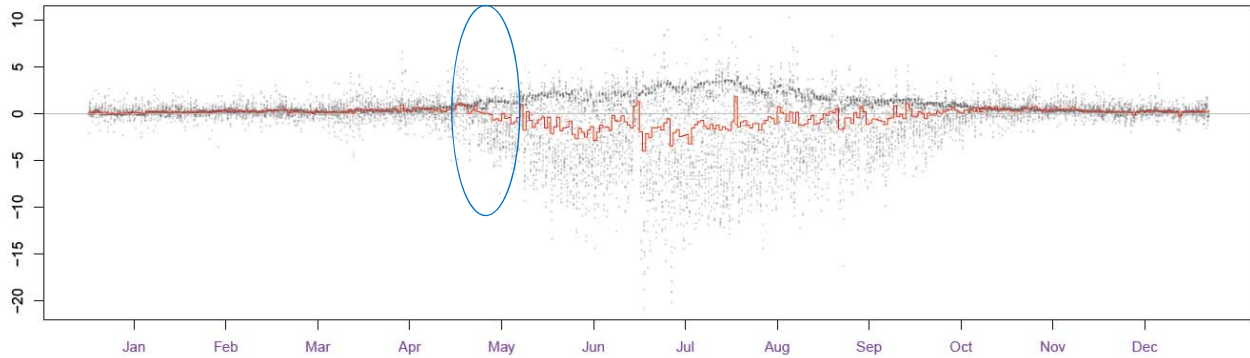
2012



Half-hourly fluxes and daily means at yc1:
NEE_uStar_f ($\mu\text{mol m}^{-2}\text{s}^{-1}$)

```
EddyProc.C$sPlotHHFluxes("NEE_uStar_f")
```

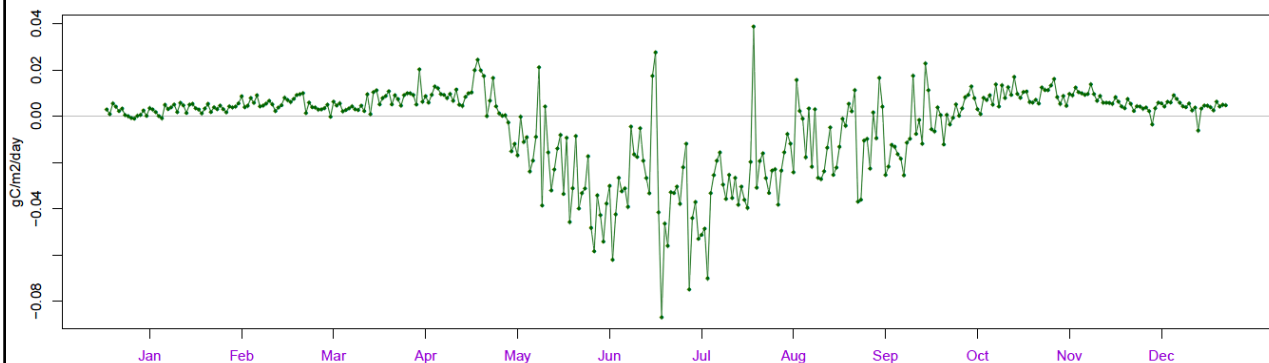
2012



Daily sums at yc1:
NEE_uStar_f ($\text{gC/m}^2/\text{day}$)

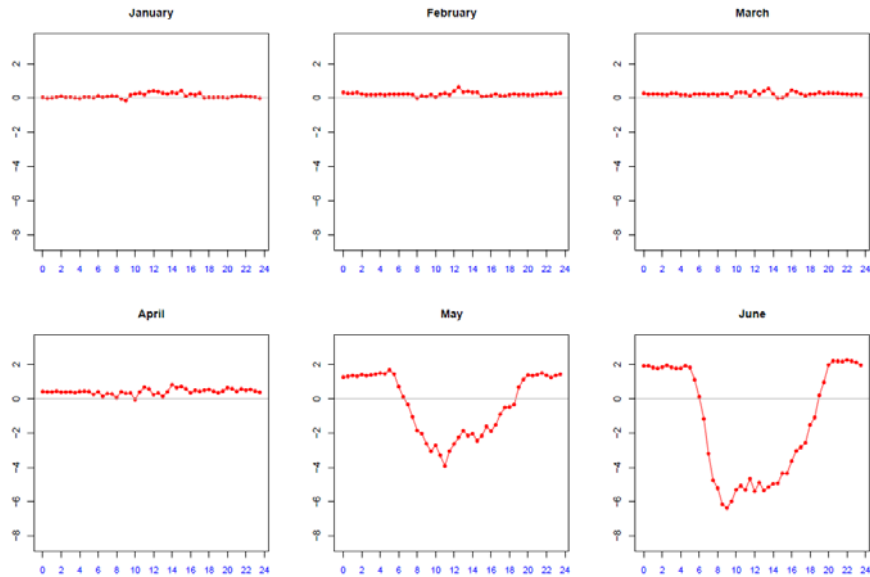
```
EddyProc.C$sPlotDailySums("NEE_uStar_f")
```

2012

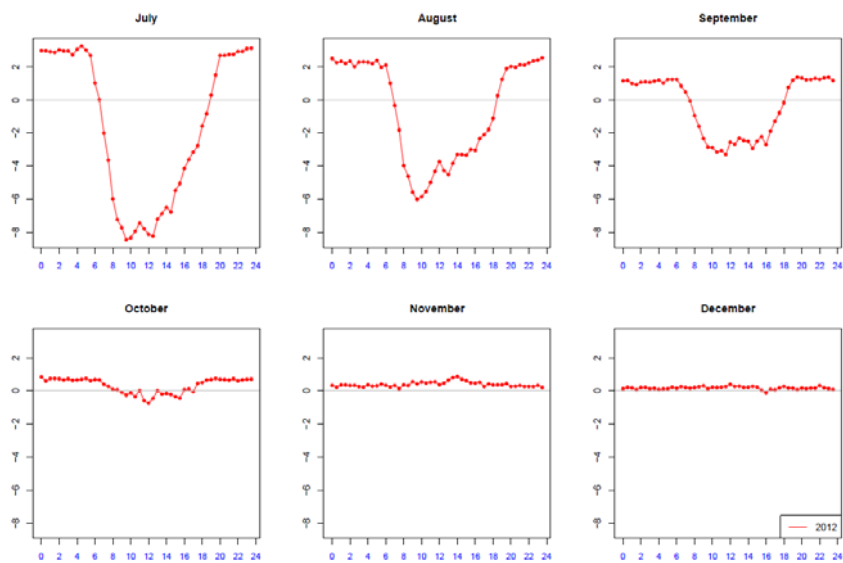


Diurnal cycles at yc1:
NEE_uStar_f (umolm⁻²s⁻¹)

EddyProc.C\$sPlotDiurnalCycle("NEE_uStar_f")

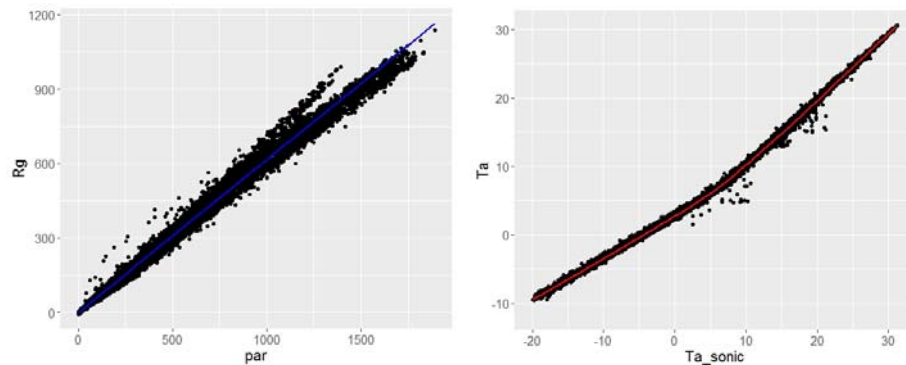


Diurnal cycles at yc1:
NEE_uStar_f (umolm⁻²s⁻¹)



如何插补微气象数据

- “Gaps in environmental variables (e.g., R_g and T_a) were filled either using complementary measurements at the site (e.g., photosynthetically active radiation or soil temperature), or with the MDS method.”(Jia et al. 2019 manuscript)



参考文献

- Falge E et al. 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. *AFM* 107, 43-69.
- Falge E et al. 2001. Gap filling strategies for long term energy flux data sets. *AFM* 107, 71-77.
- Reichstein M et al. 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. *GCB* 11, 1424-1439.
- Moffat AM et al. 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. *AFM* 147, 209-232.
- Wutzler T et al. 2018. Basic and extensible post-processing of eddy covariance flux data with REddyProc. *Biogeosciences* 15, 5015-5030.
- Jia X et al. 2014. Biophysical controls on net ecosystem CO_2 exchange over a semiarid shrubland in northwest China. *Biogeosciences* 11, 4679-4693.
- Jia X et al. 2016a. Carbon and water exchange over a temperate semi-arid shrubland during three years of contrasting precipitation and soil moisture patterns. *AFM* 228:120-129.
- Jia X et al. 2016b. Energy partitioning over a semi-arid shrubland in northern China. *Hydrological Processes*. 30, 972-985.
- Jia X et al. 2018a. Multi-scale dynamics and environmental controls on net ecosystem CO_2 exchange over a temperate semiarid shrubland. *AFM* 259, 250-259.
- Jia X et al. 2018b. Canopy photosynthesis modulates soil respiration in a temperate semi-arid shrubland at multiple timescales. *Plant and Soil* 432, 437-450.
- Amiro BD et al. 2006. Carbon, energy and water fluxes at mature and disturbed forest sites, Saskatchewan, Canada. *AFM* 136, 237-251.
- Lasslop G et al. 2010. Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation. *GCB* 16, 187-208.
- Xing Z et al. 2007. A simple net ecosystem productivity model for gap filling of tower-based fluxes: an extension of Landsberg's equation with modifications to the light interception term. *Ecological Modelling* 206, 250-262.
- Xing Z et al. 2008. Modification of an ecosystem model for filling medium-sized gaps in tower-based estimates of net ecosystem productivity. *Ecological Modelling* 213, 86-97.
- Xing Z et al., 2008. A process-based model designed for filling of large data gaps in tower-based measurements of net ecosystem productivity. *Ecological Modelling* 213, 165-179.

