Gapfilling of EC fluxes

Pasi Kolari
Department of Forest Sciences / Department of Physics
University of Helsinki

EddyUH training course
Helsinki 23.1.2013

Contents

- Basic concepts of gapfilling
- Example of gapfilling a CO_2 flux dataset
- Biological and statistical considerations

The purpose of gapfilling

- Replacement of missing or bad flux records to allow for determining matter or energy budgets over prolonged periods

- The replacement procedure is based on ”good” fluxes

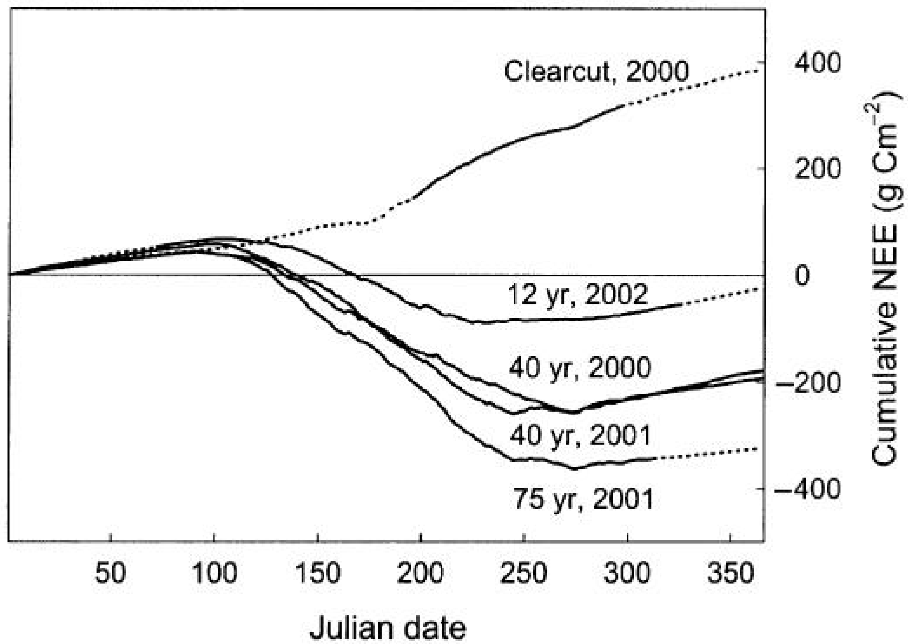
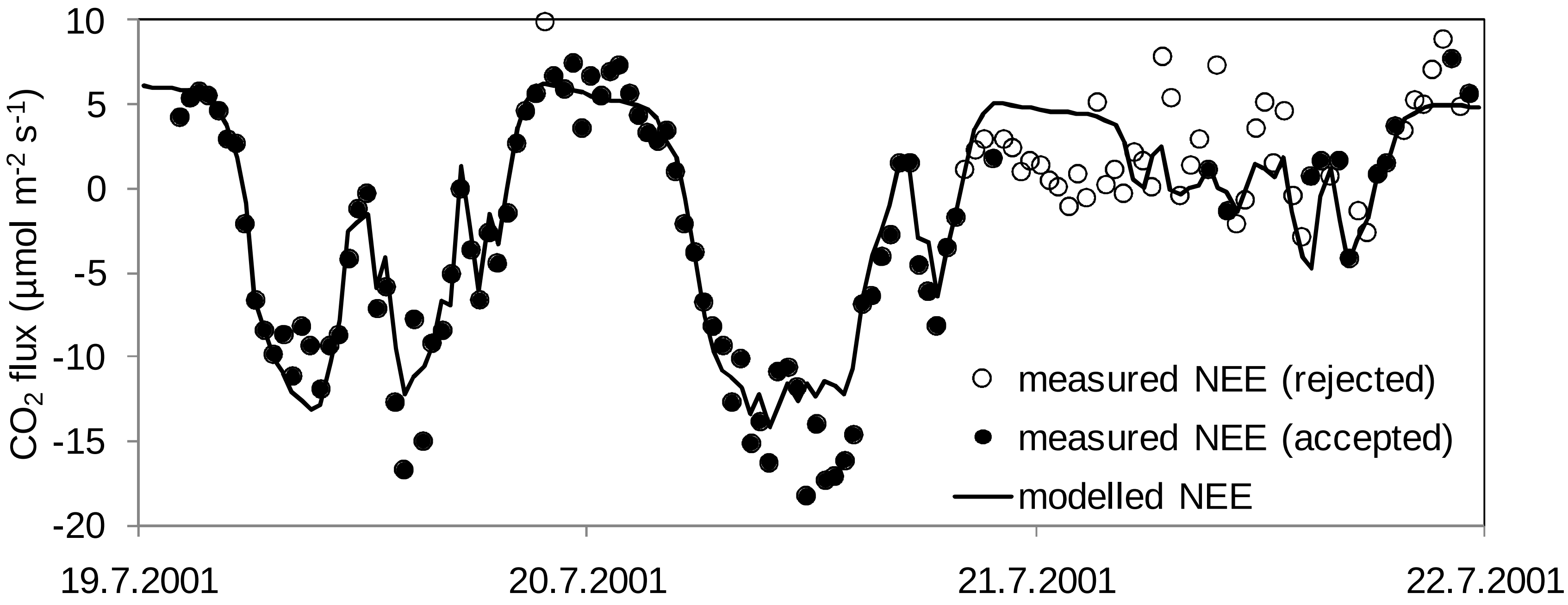


Fig. 4 Cumulative net ecosystem exchange (NEE) and the year of the measurements for Scots pine stands of different ages. Solid lines indicate measured NEE, dotted lines modelled NEE. Positive sign indicates release of carbon by the stand, negative sign uptake of carbon.

Kolari et al. 2004

The purpose of gapfilling

- Especially important with EC as the proportion of ”bad” flux records is considerable

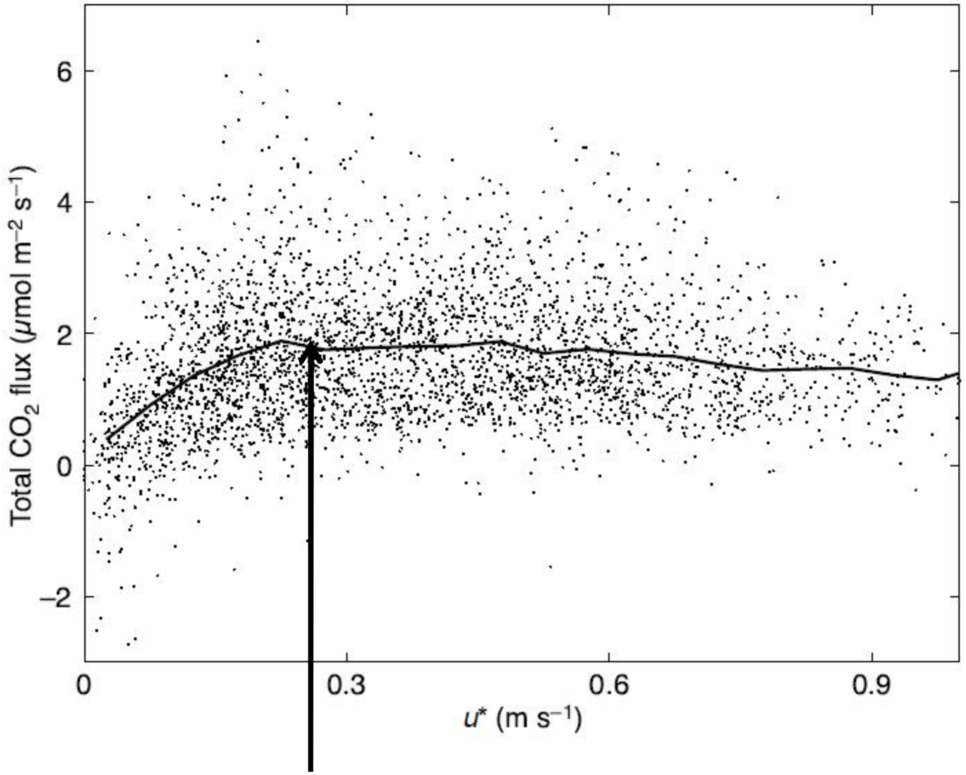


Binning ”good” and ”bad” fluxes

- Definition of ”good” = turbulent transport well representative of total exchange

- Somewhat subjective decision, no universal rules or threshold values exist
  - strength of turbulence (e.g. friction velocity u*, std w), u* vs normalised NEE is frequently used criterion
  - stability (e.g. Monin-Obukhov length)
  - flux stationarity

Nighttime NEE vs u*



Markkanen et al. 2001

Fig. 2. Nocturnal summertime total fluxes (at 23 meters) of the periods with air temperature 2 °C < T_a < 8 °C as a function of friction velocity. Total flux is a sum of flux measured by eddy-covariance and the storage term calculated from the gas concentration gradient. The line indicates a running average of the total flux.

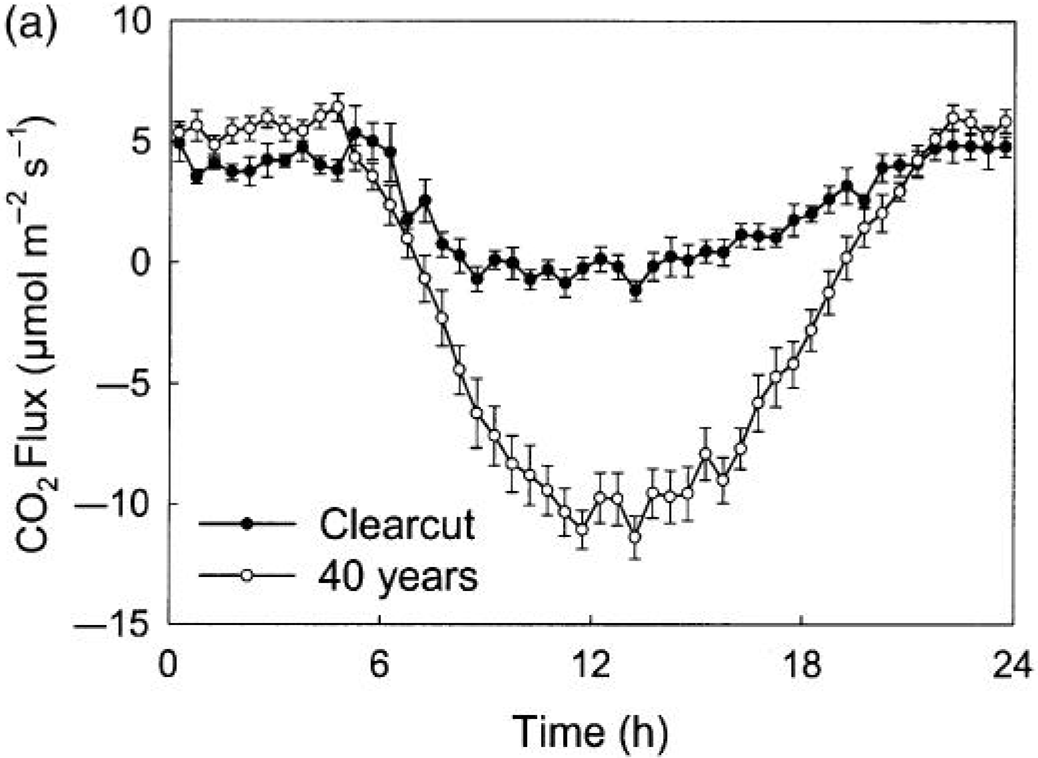
- u* threshold of 0.25 m s^{-1}
- Different acceptance criteria were applied for daytime fluxes

Basic gapfilling methods

- Replacing missing or bad records by utilising regular patterns in the fluxes; accepted fluxes measured in the same time of day, similar conditions etc.
  - Replacement by mean value; very questionable, why?
  - Interpolation
    - very short gaps only, a bit questionable, why?
  - Mean diurnal variation – temporal autocorrelation of fluxes
  - Look-up tables – flux dependence on environment
  - Regressions of fluxes on environmental variables
    - most widely used method
  - Recursive flux estimation algorithms (e.g. Kalman filter)
  - Artificial neural networks

Basic methods: Mean diurnal variation

- Bin-averages of half hours (or hours) on previous and subsequent days
  - time window typically days to weeks



Kolari et al. 2004

Diurnal patterns of accepted NEE over one summer month

Basic methods: Look-up tables

- Utilising other regularities in fluxes than just temporal patterns
- Weather conditions = environmental driving factors of fluxes

- Example:
  Missing flux at {R_{glob}=536 W/m^2, T_{air}=20.5°C}
  is replaced by
  mean accepted flux at {R_{glob}=500...600 W/m^2, T_{air}=20...25°C}

Basic methods: (Nonlinear) regressions

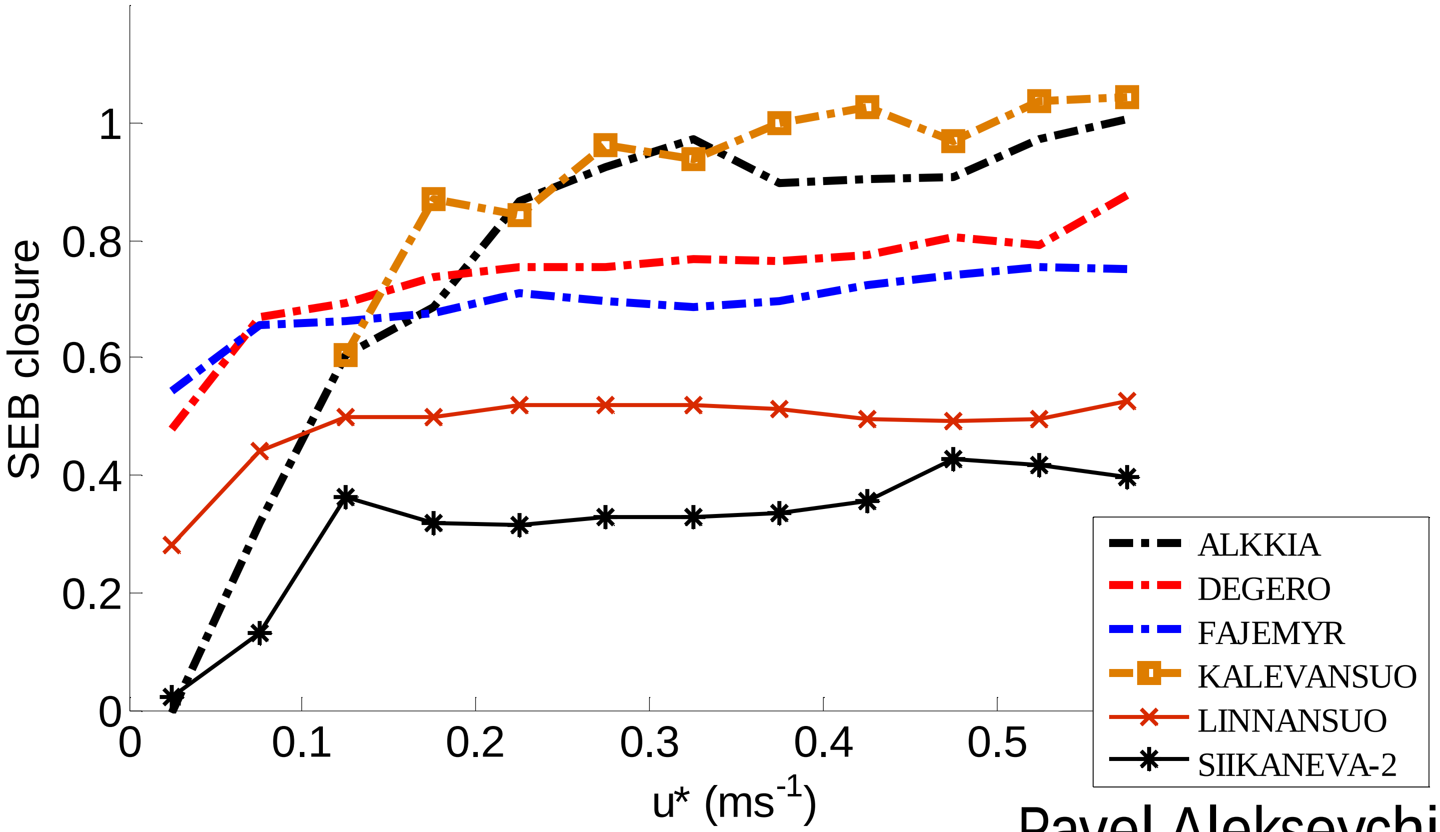
- Purely statistical models
  - In theory models could be formulated in a way that optimises robustness of parameter estimation (e.g. multiple imputation, Hui et al. 2004)
  - In practise arbitrary regressions are not widely used for CO_2 fluxes because simple process models for partitioning the net flux into main components exist
- Simple aggregated process descriptions
  - CO_2 exchange NEE = GPP + R_e
  - Gross CO_2 uptake = Gross Primary Production = ecosystem photosynthesis
    $GPP = f(I, T_{air}, VPD, \ldots)$
  - Gross CO_2 release = ecosystem respiration
    $R_e = f(T_{air}, T_{soil}, W_{soil}, \ldots)$

Other gapfilling methods

- Kalman filter
  - Basically (nonlinear) regressions with parameter estimation procedure more sophisticated than normal least-squares fitting
  - Recursive algorithm with more weight being given to flux estimates with higher certainty
- Artificial neural networks (ANN)
  - ”learning” algorithms, no fixed pre-assumed relationships between explanatory variables and fluxes

Gapfilling of heat fluxes

- Observed surface energy balance is dependent on turbulence



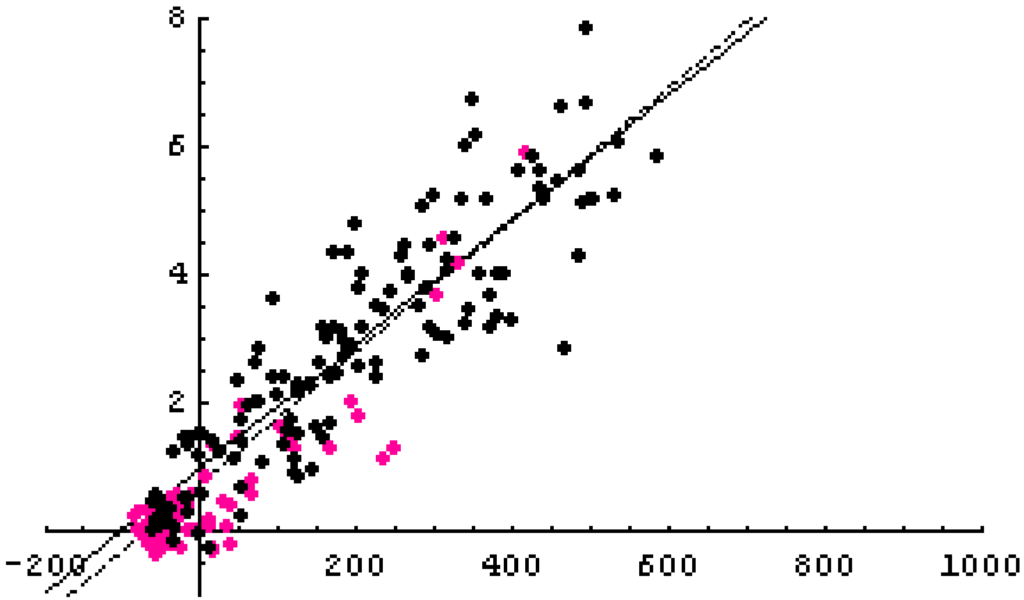
Pavel Alekseychik

Gapfilling of heat fluxes

- Simpler methods than in CO_2 flux gapfilling are sufficient
  - Heat fluxes in conditions of high stability/low turbulence are small
  - Difference in annual LE with and without turbulence filtering (only missing records filled) just few % vs tens of % in NEE

ET (mmol m^{-2} s^{-1}) vs R_{net} (W m^{-2})

black = good data
red = rejected data (low turbulence, storage change missing, errors in raw data,...)



Gapfilling of heat fluxes

- Heat fluxes are largely driven by solar energy
  →Regressions of fluxes on
  available radiative energy (R_{net}, R_{glob})
  - Priestley-Taylor model
  - Penman-Monteith model
  - Simple empirical regressions

$$LE = \frac{\Delta Rn + \rho c_p (e_s - e_a)/r_a}{\Delta + \gamma\left(1 + \frac{r_s}{r_a}\right)}$$

- Energy balance methods

Example of using nonlinear regression models for gapfilling of CO_2 fluxes

- Flux partitioning into GPP and ecosystem respiration

NEE = R_e - GPP

$R_e = R_{ref}\ e^{f(T)}$

$$\mathrm{GPP} = -\frac{\alpha I + P_{\max} - \sqrt{\left(\alpha I + P_{\max}\right)^2 - 4\theta\alpha I P_{\max}}}{2\theta}$$

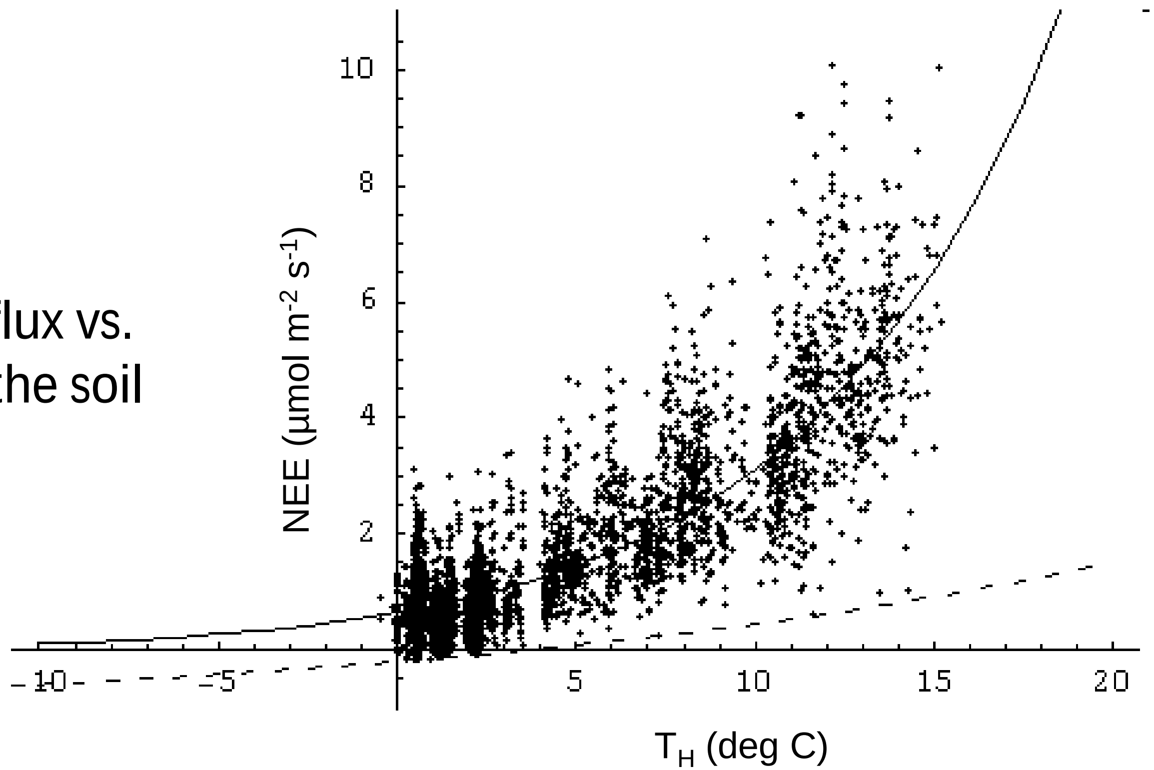
- Θ is fixed,
  α and P_{max} vary over year, estimated in moving time window

- more simple version $\mathrm{GPP} = \dfrac{\alpha I P_{\max}}{\alpha I + P_{\max}}$

Example of using nonlinear regression models for gapfilling of CO_2 fluxes

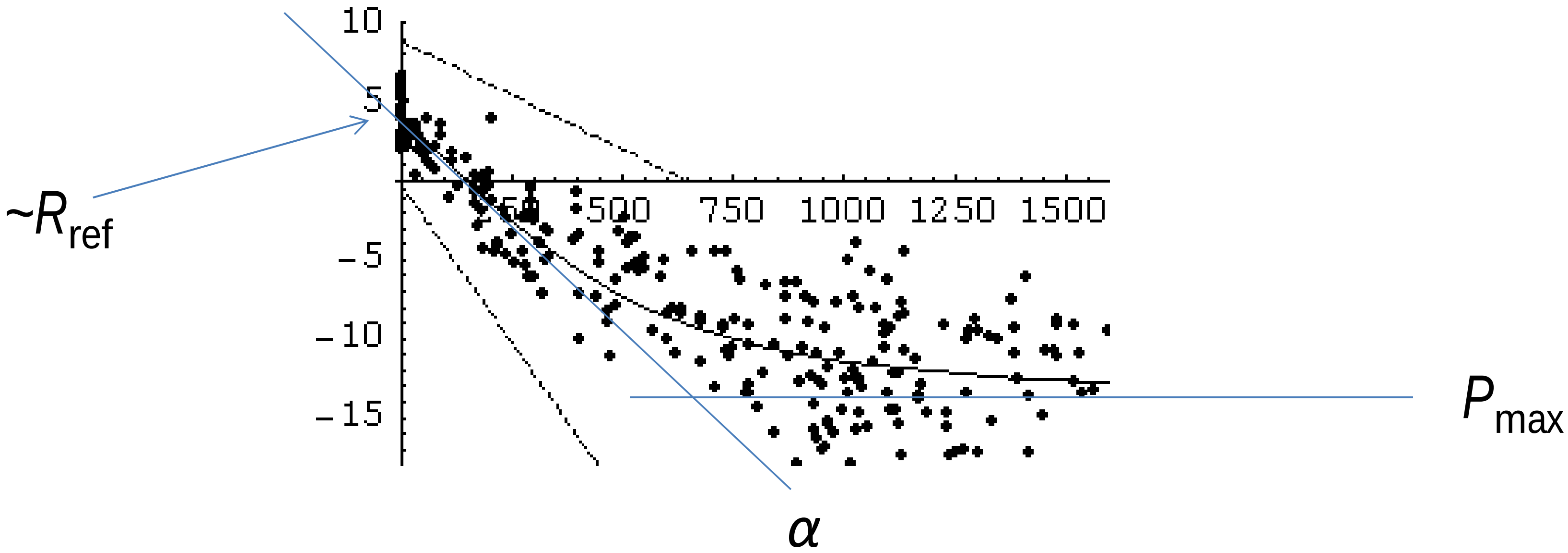
- Estimation of the slope of the *T* response of R_e using night-time fluxes
  - often requires time window of a month or more

Night-time CO_2 flux vs. temperature in the soil organic layer, whole year



Example of using nonlinear regression models for gapfilling of CO_2 fluxes

- Storage-corrected NEE vs PPFD over 9 days of July in Hyytiälä

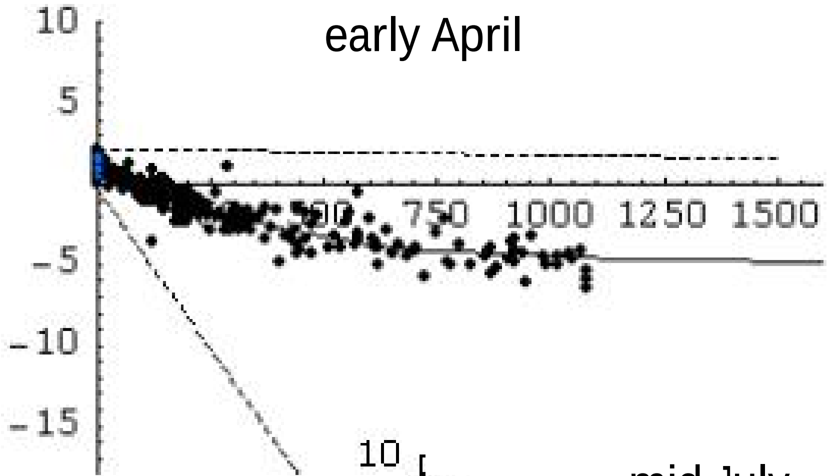


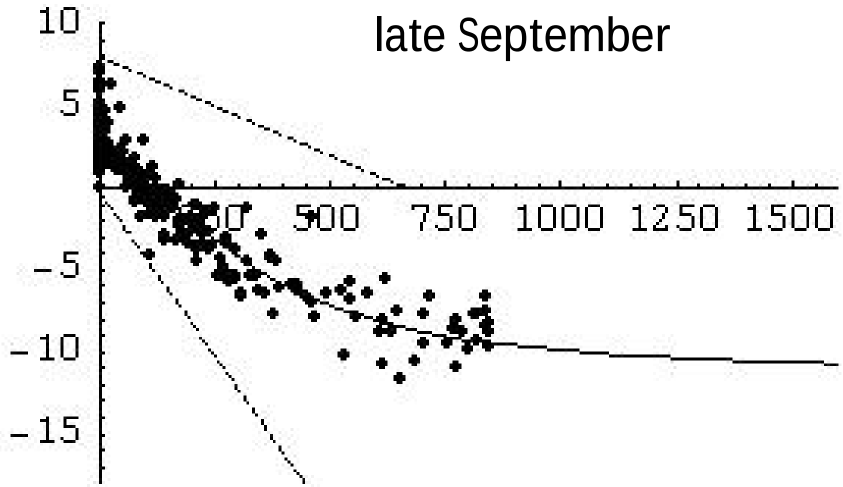
$$R_e = R_{ref}\, Q_{10}^{(T/10)}$$

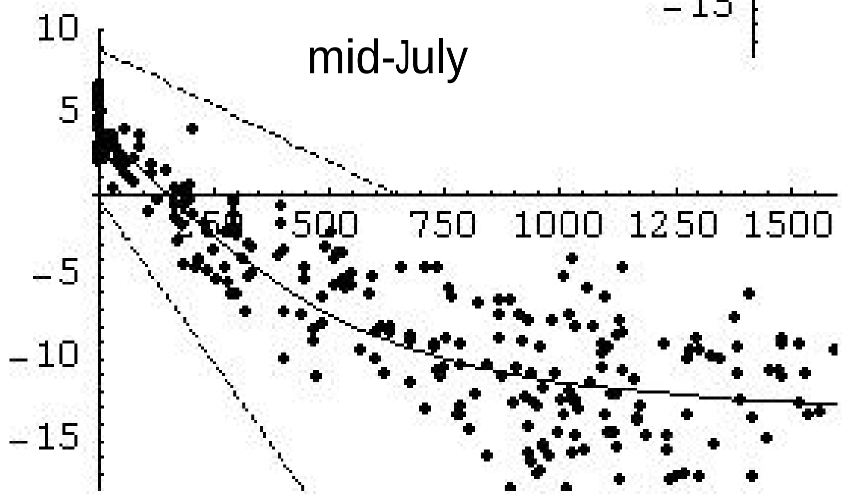
$$\mathrm{GPP} = -\frac{\alpha I + P_{\max} - \sqrt{(\alpha I + P_{\max})^2 - 4\theta\alpha I P_{\max}}}{2\theta}$$

Example of using nonlinear regression models for gapfilling of CO_2 fluxes

- Storage-corrected NEE vs PPFD over 9 days in Hyytiälä



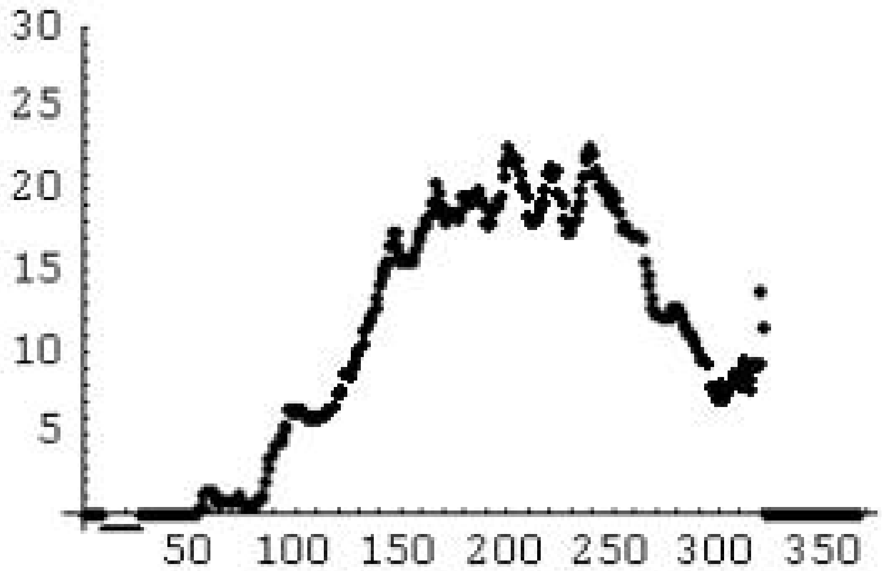




Example of using nonlinear regression models for gapfilling of CO_2 fluxes

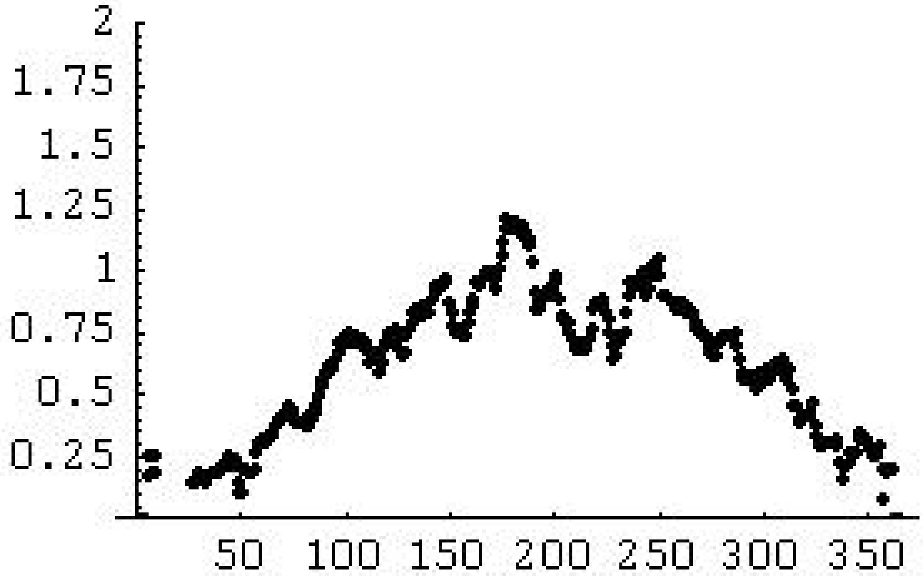
- Temporal patterns of gapfilling model parameters
- Parameters estimated in moving time window of 9 days

P_{max} (µmol m^{-2} s^{-1})



Day of year

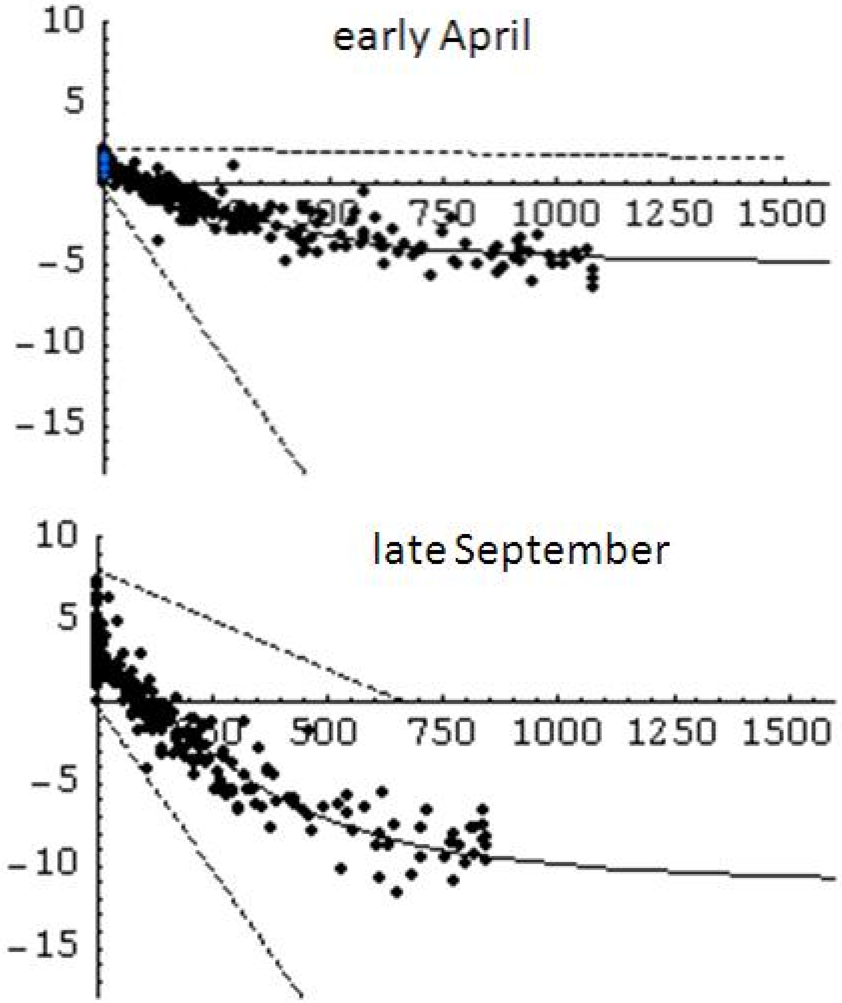
R_{ref} (µmol m^{-2} s^{-1}), T_{ref}=0



Day of year

Interpretation of temporal dynamics in gapfilling

- Instantaneous responses vs seasonal changes in the state of the ecosystem
  - state changes → instantaneous responses change
- Essential to consider at northern latitudes where the seasonality is very strong
  - also important where the seasonal drought affects fluxes
- Estimation of gapfilling model parameters in narrow (sliding) time window

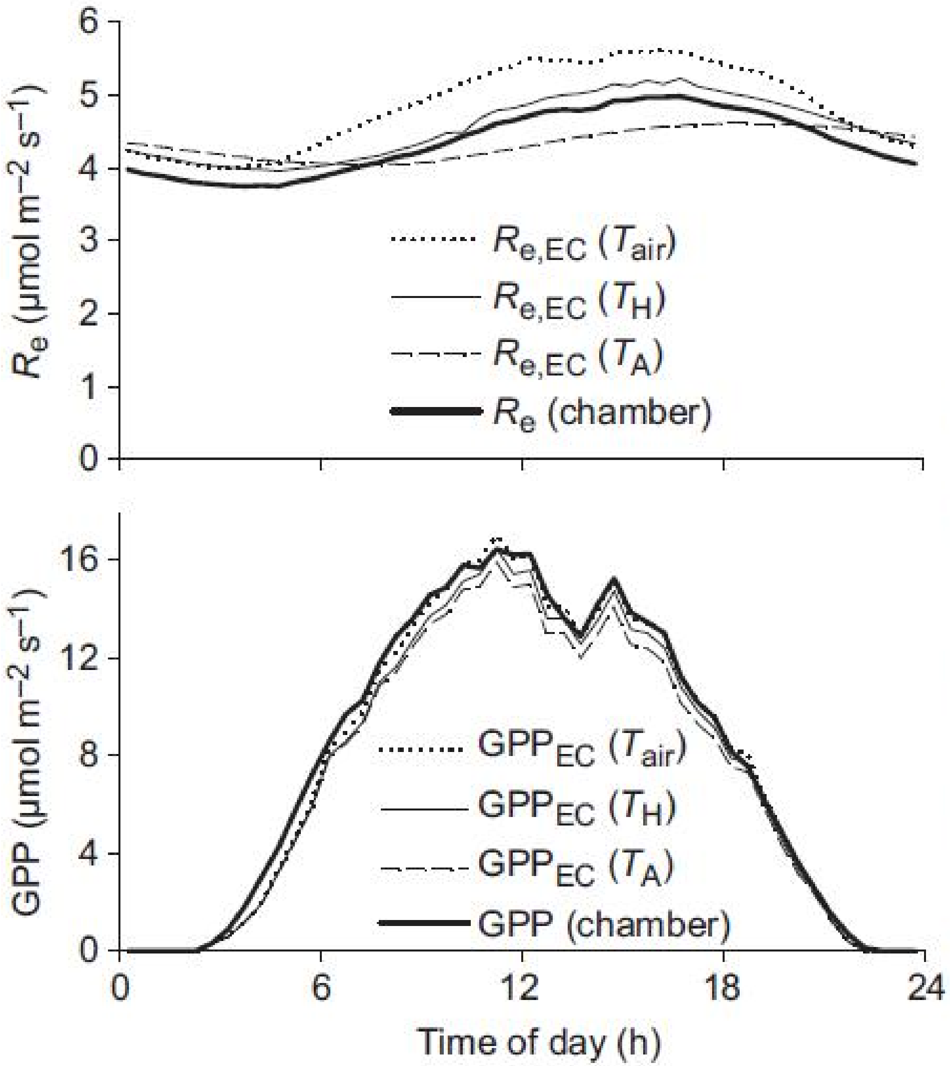


The importance of mechanistic and unbiased CO_2 flux partitioning

- Modellers use EC data (net exchange or component fluxes)
  - for estimating model parameters
  - as independent data for testing models

- GPP is often requested data by ecosystem modellers, usually a by-product of gapfilling
  and happily accepted as is by the modellers!

Diurnal patterns of GPP and R_e with different drivers



Biological relevance of gapfilling model variables is preferable!

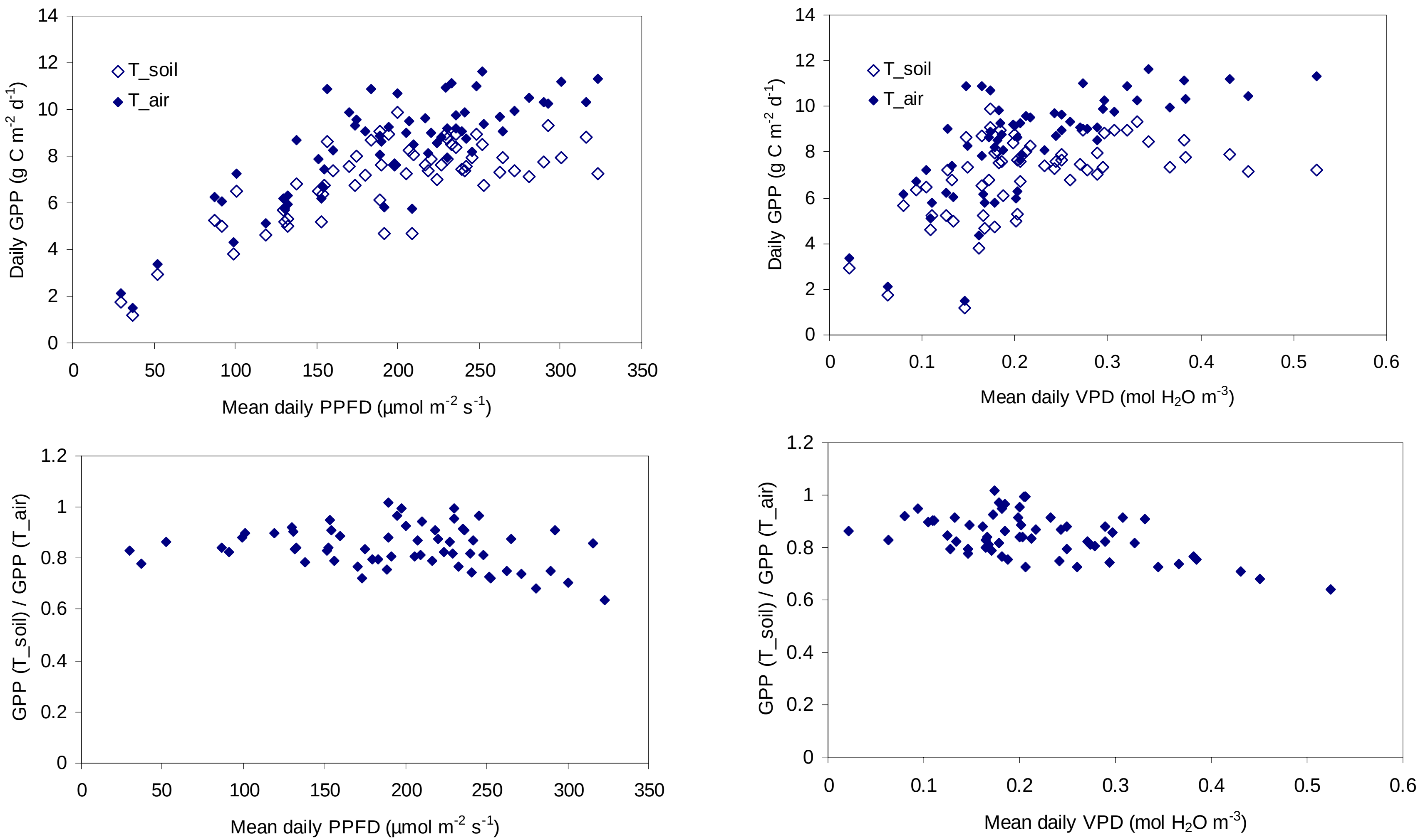
Fig. 11. Diurnal patterns of chamber-based R_e and GPP in comparison with $R_{e,EC}$ and GPP_{EC} calculated using different explanatory factors for the R_e model. The diurnal patterns were averaged for 15 June–15 July 2004. The sensitivity of $R_{e,EC}$ to each type of temperature was first estimated from night-time EC fluxes in June and July. Light-saturated stand GPP ($P_{e,max}$) and the base level of R_e (R_{10}) were then estimated daily in a moving time window of nine days. $R_{e,EC}(T_{air})$ and $GPP_{EC}(T_{air})$ refer to air temperature as the explanatory factor, $R_{e,EC}(T_H)$ and $GPP_{EC}(T_H)$ to the organic layer temperature, and $R_{e,EC}(T_A)$ and $GPP_{EC}(T_A)$ to the temperature at the depth of 5 cm from the mineral soil surface.

Kolari et al. 2009

Parameterising a model of daily forest GPP using eddy-covariance GPP

$GPP = LUE\ f_{Ta}\ f_{PAR}\ f_{VPD}\ f_{SWC}$

Especially VPD modifier is heavily affected by TER submodel selection



Variability of cumulative NEE among different gapfilling methods

- For NEE, selection of method is not critical
  - bias in annual NEE <10% vs tens of % without data filtering

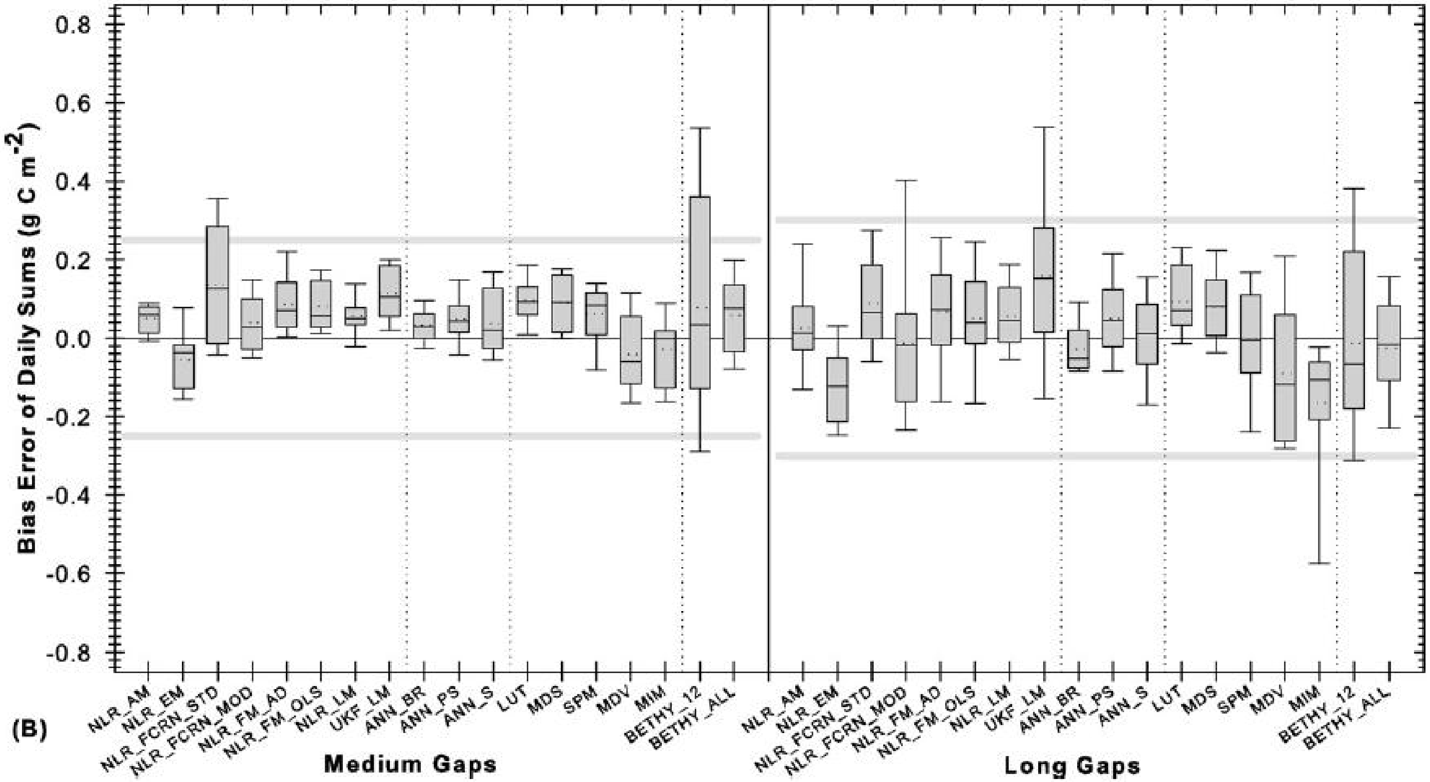
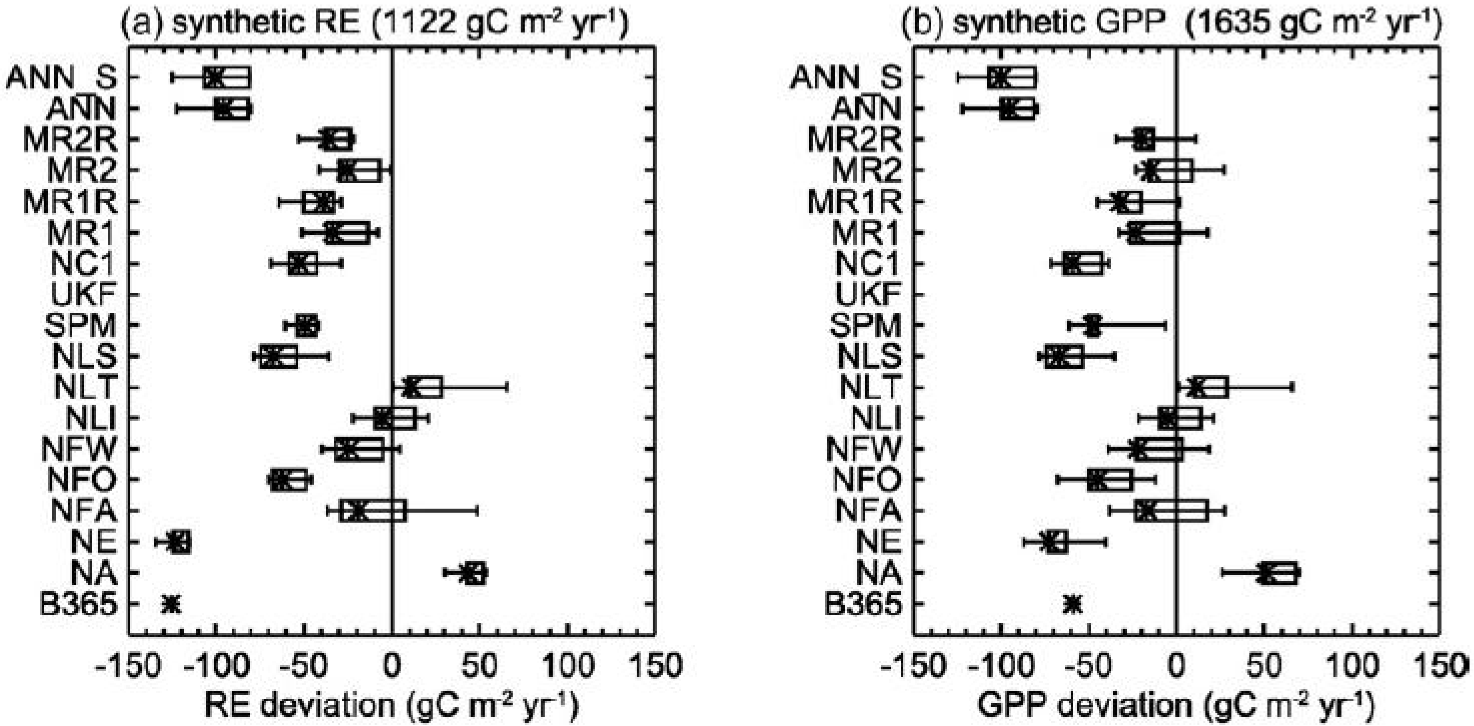


Fig. 7. (A) Bias error of the gap-filling techniques in the prediction of half-hourly NEE: boxplot of the very short gap length scenario calculated for all 10 benchmark datasets and all 10 permutations (100 data points), separated into daytime (left) and nighttime (right) data. The boxplot is drawn as in Fig. 2. (B) Bias error of the gap-filling techniques in the prediction of the daily NEE: boxplot for medium and long gap length scenario calculated for the 10 benchmark datasets (10 data points). The boxplot is drawn as in Fig. 2.

Moffat et al. 2007

Variability of cumulative CO_2 component fluxes among different gapfilling methods

- GPP and R_e are more sensitive than NEE to the type of the gapfilling model (typical example of drawing the right conclusion by false thinking)

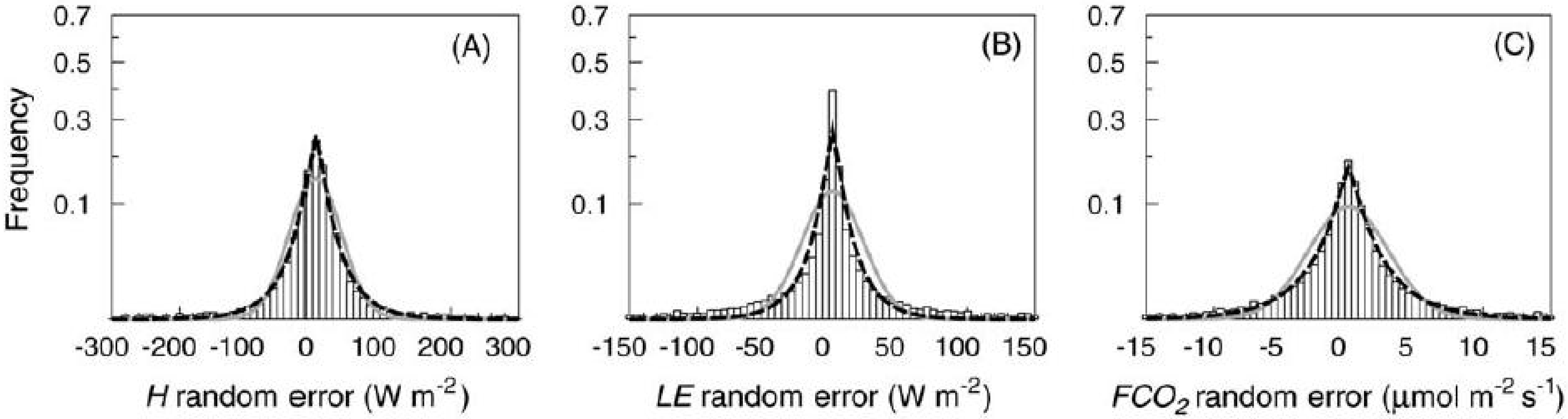


Desai et al. 2008

Statistical considerations

- The random error in fluxes is double-exponentially rather than normally distributed

A.D. Richardson et al./Agricultural and Forest Meteorology 136 (2006) 1–18

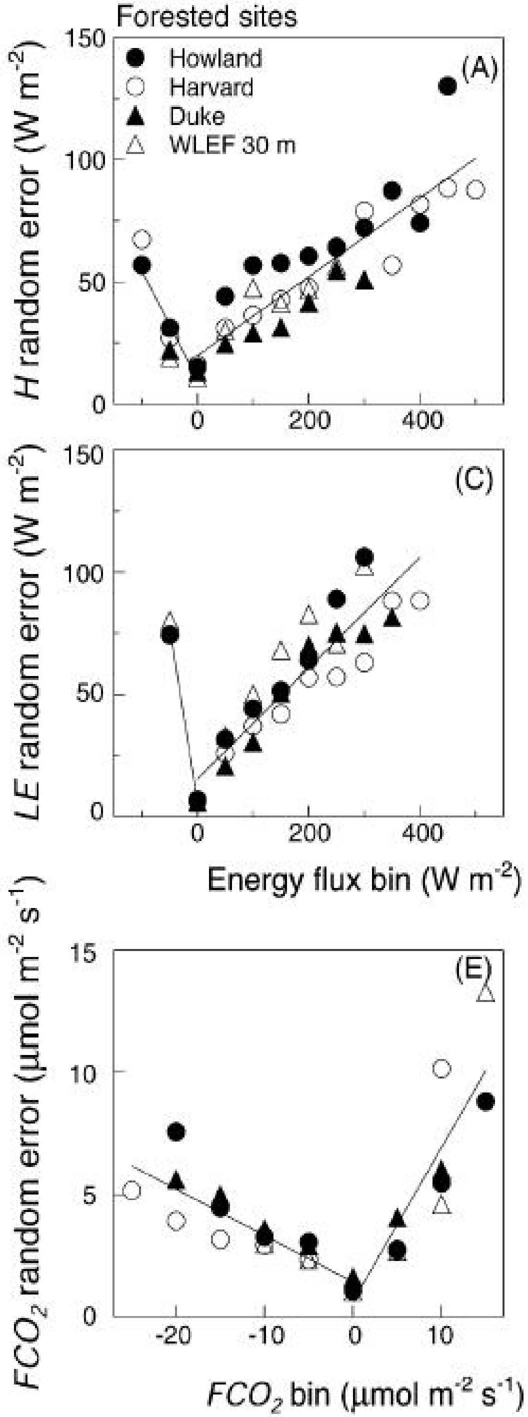


Statistical considerations

- Half-hourly EC fluxes are heteroscedastic, i.e. variance is not constant
  - Absolute error increases but relative error gets smaller as the magnitude of the flux increases

- Least-squares methods that are based on assumptions of constant variance and normal distribution of error do not result in maximum likelihood parameter values

- In practise these issues are ignored



Final remarks

- Always flag the gapfilled fluxes
  - You must not use momentary gapfilled fluxes for other purposes than calculating budgets or analysing the gapfilling method itself!
  - Never underestimate other peoples' ability to misinterpret your data!

- Think why you are gapfilling fluxes
  - Determining long-term budgets: statistical robustness, ”simulation of data”
  - Modelling use: mechanistic basis preferable, ”simulation of processes”

Further reading

- Falge et al. 2001. Gap filling strategies for defensible annual sums of net ecosystem exchange. Agricultural and Forest Meteorology 107, 43–69.

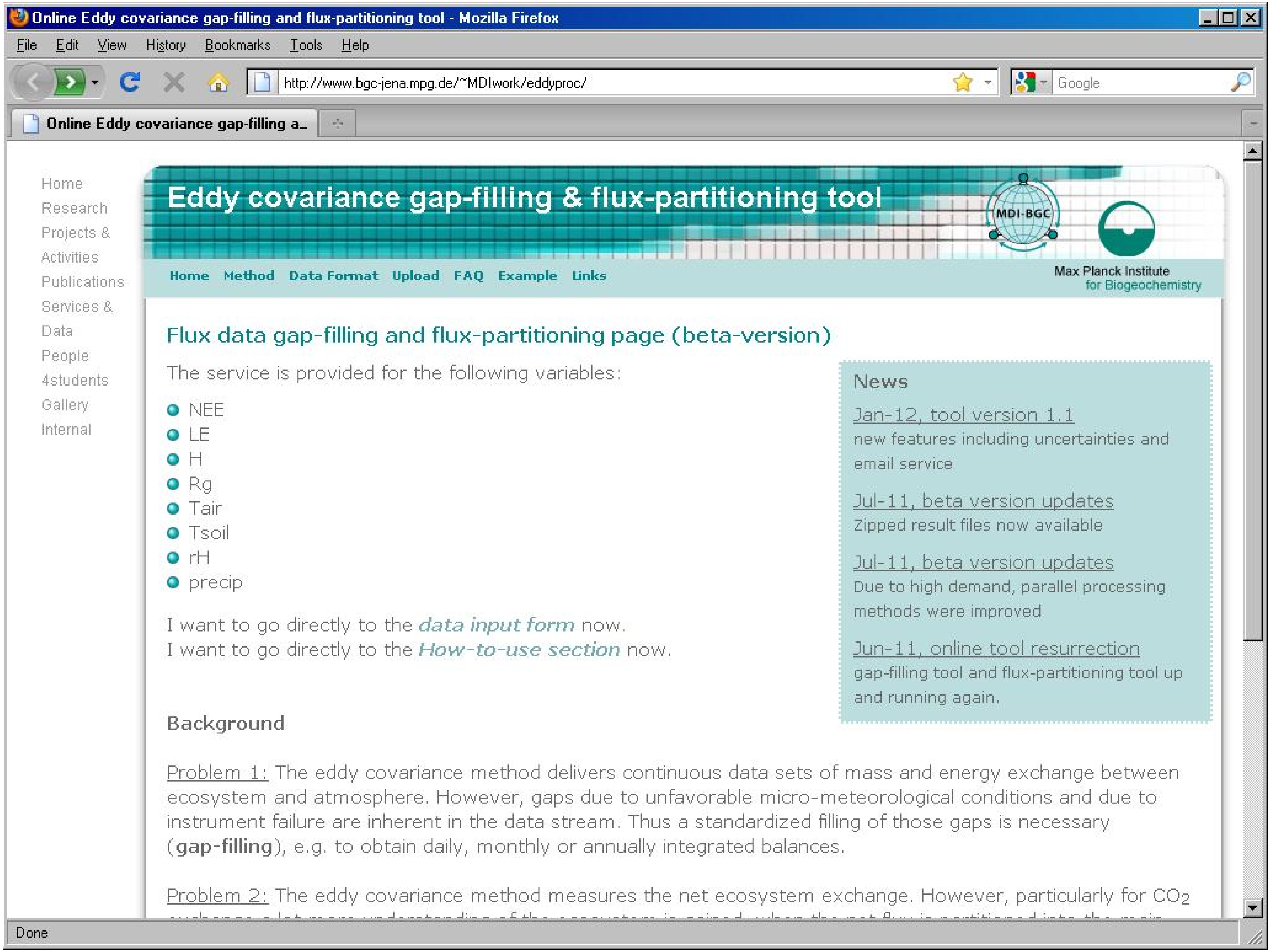
- Falge et al. 2001. Gap filling strategies for long term energy flux data sets. Agricultural and Forest Meteorology 107, 71–77.

- Gu et al. 2005. Objective threshold determination for nighttime eddy flux filtering. Agricultural and Forest Meteorology 128, 179–197.

- Moffat et al. 2007. Comprehensive comparison of gap-filling techniques for eddy covariance net carbon fluxes. Agricultural and Forest Meteorology 147, 209–232.

- Reichstein et al. 2005. On the separation of net ecosystem exchange into assimilation and ecosystem respiration: review and improved algorithm. Global Change Biology 11, 1424–1439.

  Online gapfilling tool http://www.bgc-jena.mpg.de/~MDIwork/eddyproc/



Online Eddy covariance gap-filling and flux-partitioning tool - Mozilla Firefox
File Edit View History Bookmarks Tools Help
http://www.bgc-jena.mpg.de/~MDIwork/eddyproc/
Google
Online Eddy covariance gap-filling a...
Home
Research
Projects & Activities
Publications
Services & Data
People
4students
Gallery
Internal
Eddy covariance gap-filling & flux-partitioning tool
MDI-BGC
Home Method Data Format Upload FAQ Example Links
Max Planck Institute for Biogeochemistry
Flux data gap-filling and flux-partitioning page (beta-version)
The service is provided for the following variables:
NEE
LE
H
Rg
Tair
Tsoil
rH
precip
I want to go directly to the data input form now.
I want to go directly to the How-to-use section now.
News
Jan-12, tool version 1.1
new features including uncertainties and email service
Jul-11, beta version updates
Zipped result files now available
Jul-11, beta version updates
Due to high demand, parallel processing methods were improved
Jun-11, online tool resurrection
gap-filling tool and flux-partitioning tool up and running again.
Background
Problem 1: The eddy covariance method delivers continuous data sets of mass and energy exchange between ecosystem and atmosphere. However, gaps due to unfavorable micro-meteorological conditions and due to instrument failure are inherent in the data stream. Thus a standardized filling of those gaps is necessary (gap-filling), e.g. to obtain daily, monthly or annually integrated balances.
Problem 2: The eddy covariance method measures the net ecosystem exchange. However, particularly for CO2
Done