A Quality Assessment Scheme for Long-term Eddy-Covariance Measurements

Matthias Mauder, Matthias Cuntz, Clemens Drüe, Alexander Graf, Corinna Rebmann, Marius Schmidt, Rainer Steinbrecher

Atmospheric Environmental Research KIT/IMK-IFU

Outline

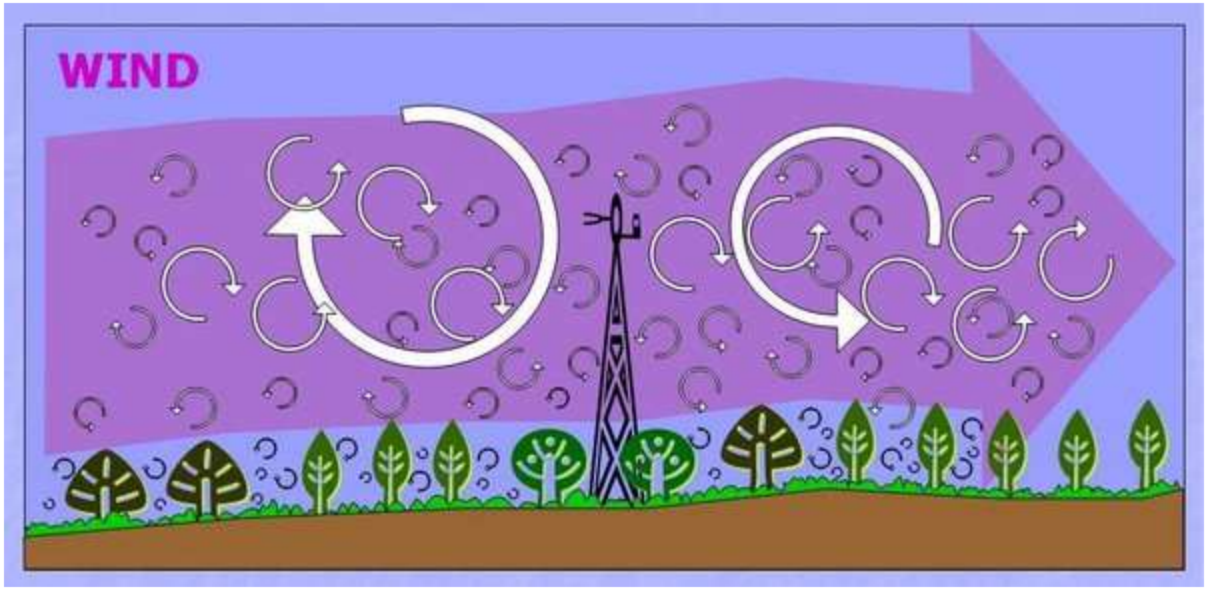
- Introduction
- Quality assessment algorithms
  - Tests on high-frequency data
  - Tests on statistics, flux calculation and corrections
  - Quantification of errors and uncertainty estimates
  - The quality assessment scheme
  - Test data sets
- Results and discussion
  - Quality flags
  - Random errors and noise
  - Systematic error and energy balance ratio
  - Footprint analysis
  - Further tests
  - Effectiveness of the quality assessment scheme
- Conclusions

Introduction – Eddy Covariance

Karlsruher Institut für Technologie

- **Objective:**
  - to measure the transport/flux of
    air properties (e.g. temperature) or air constituents (e.g. GHGs)
    between an ecosystem at the earth‘s surface and the atmosphere
- **Advantages:**
  - non-destructive
  - no disturbance of the exchange conditions
  - area-averaging on the ecosystem-scale
  - quasi-continuous
- **Assumptions (Limitations):**
  - Atmospheric exchange is fully turbulent, zero mean vertical wind
  - Taylor hypothesis: space and time interchangeable
  - Horizontal homogeneity, stationarity/steady state conditions
  - All transporting eddy scales are captured
  - Constant flux layer

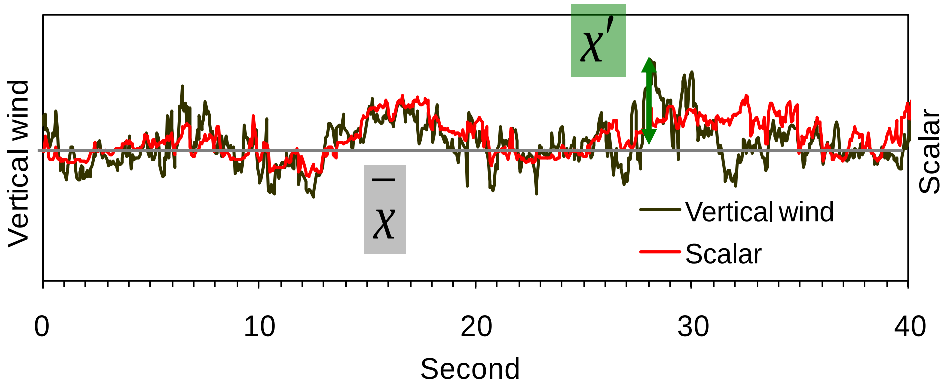
Introduction – Eddy Covariance

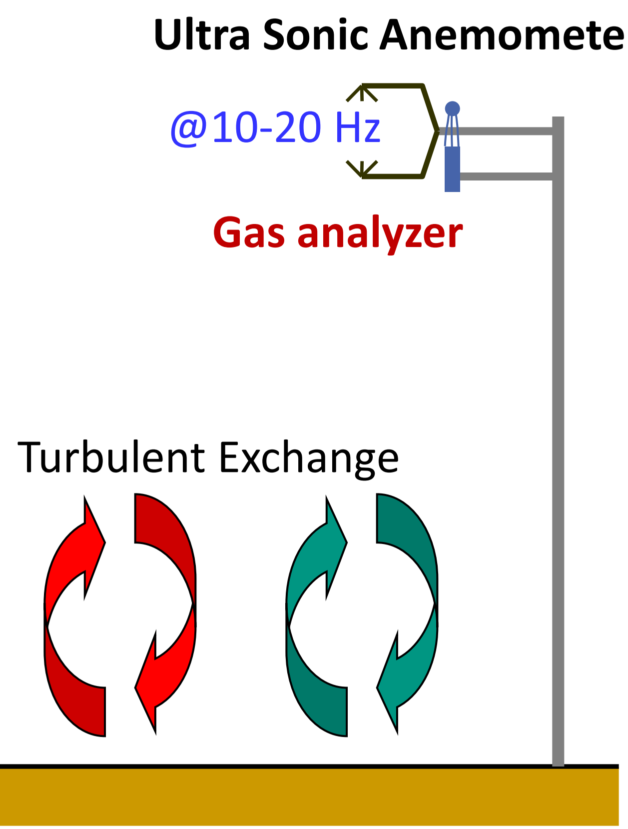


(Burba and Anderson, 2010)

Measurements at a single point can represent the ecosystem flux from an upwind area

Introduction – Eddy Covariance





Flux Calculation (assume stationarity)

$$F = \overline{wq} = \overline{w}\,\overline{q} + \overline{w'q'}$$

$\overline{w} = 0$, assume horizontal homogeneity,
30 – 60 min averaging

Introduction – Challenges for QA scheme

- Extensive amount of data
  - Per site @ 20 Hz: 72000 data lines per hour, 1728000 lines per day, 51840000 lines per month, 1.89216 10^{10} lines per year
  - **> 10 sites** operational or planned within TERENO running **for > 10 years**
- Detection of instrumental failure
  - coverage by rain droplets, dew or frost
- Validity of assumptions for the EC method
  - well-developed turbulence, stationarity, zero mean vertical wind
- Quantification of measurement errors
  - noise, random, systematic
- Representativeness of the flux estimate for the targeted ecosystem
  - footprint

Introduction – State of the art

Karlsruher Institut für Technologie

- Foken and Wichura (1996, FW96), Foken et al. (2004)
  - Tests on stationarity, well-developed turbulence, zero mean vertical wind => flag system 1 – 9
  - Foken T, Göckede M, Mauder M, Mahrt L, Amiro BD, Munger JW (2004) Post-field data quality control. In: Handbook of Micrometeorology. A Guide for Surface Flux Measurements. Lee X, Massman WJ, Law BE, eds. Dordrecht: Kluwer, pp. 181-208.
  - Foken T, Wichura B (1996) Tools for quality assessment of surface-based flux measurements. Agric Forest Meteor 78:83-105.
- Vickers and Mahrt (1997)
  - Various tests on raw data => system of hard and soft flags
  - Vickers D, Mahrt L (1997) Quality control and flux sampling problems for tower and aircraft data. J Atmos Oceanic Technol 14:512-526.
- Papale et al. (2006)
  - Test on fluxes without knowledge of raw data, plus uncertainty due to corrections
  - Papale D, Reichstein M, Canfora E, Aubinet M, Bernhofer C, Longdoz B, Kutsch W, Rambal S, Valentini R, Vesala T, Yakir D (2006) Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences 3:571-583.
- Error estimates by various authors
  - noise, random, systematic
  - Lenschow DH, Mann J, Kristensen L (1994) How long is long enough when measuring fluxes and other turbulence statistics? J Atmos Oceanic Technol 11:661-673.
  - Finkelstein PL, Sims PF (2001) Sampling error in eddy correlation flux measurements. J Geophys Res 106:3503-3509.
  - Billesbach DP (2011) Estimating uncertainties in individual eddy covariance flux measurements: A comparison of methods and a proposed new method. Agric Forest Meteor 151:394-405.

Introduction – Goals

- Provide **comparable** flux data with **simple** quality flags and **quantitative error** estimates (random and systematic)
- Create a **comprehensive** quality assessment **scheme** of tests and algorithms, which can be **applied automatically** to long-term measurements (as strict as necessary and as simple/lean/efficient as possible)
- The selected tests and criteria should be as **fundamental** as possible to allow a **wide applicability** to different site conditions
- Provide **pre-processed additional information** about the fluxes in a **standardised** way, which may **assist the user** in deciding how to use and filter the data for his/her specific application.

Quality assessment algorithms

Quality assessment algorithms can be subdivided into three parts

- Tests on high-frequency data
- Tests on statistics, fluxes and corrections
- Quantification of error/uncertainty estimates

Algorithms - Tests on high-frequency data

- Modern micrometeorological instruments have **internal quality tests** and provide a diagnostic flag, e.g. Campbell CSAT3 or Licor LI-7500.
- **Spike test** based on Median Absolute Deviation (MAD), which is a robust measure for outlier or spike detection
- Limits for **instrumental plausibility** screening of the high-frequency data

High-frequency data points rejected by the above tests were replaced by an error code (NaN)

Alternatively for short data gaps: - repeat the last measured value
- linear interpolation

Algorithms – Tests on statistics

- **Consensus about flux calculation and correction** procedures is a pre-condition for the quality assessment scheme.
  - We follow the recommendations of Lee et al. (2004), which is in accordance with procedure used in the CarboEurope-IP **software comparison** of Mauder et al. (2008)

- **Simple flag system**, which follows the agreement of the 2nd CarboEurope-IP QA/QC workshop (Mauder and Foken, 2004):
  - Flag 0 – high quality data, use in fundamental research possible
  - Flag 1 – moderate quality data, no restrictions for use in long-term observation programs
  - Flag 2 – low data quality, to be discarded

- Lee, X., Massman, W., and Law, B. E. (2004) Handbook of Micrometeorology. A Guide for Surface Flux Measurement and Analysis. Dordrecht: Kluwer Academic Press, 250 pp.
- Mauder M, Foken T, Clement R, Elbers JA, Eugster W, Grünwald T, Heusinkveld B, Kolle O (2008b) Quality control of CarboEurope flux data - Part 2: Inter-comparison of eddy-covariance software. Biogeosciences 5:451-462.

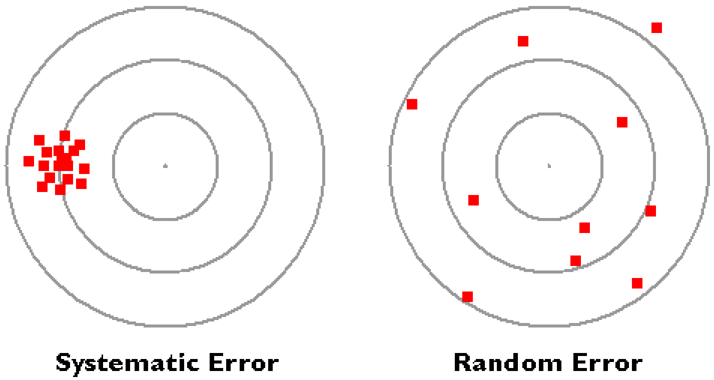
Algorithms – Tests on statistics

- Minimum of 90% of the raw data available per averaging period
- Test the assumptions of the EC method (simplified after FW96):
  - Stationarity (>30% = *flag* 1; >75% = *flag* 2)
  - Well-developed turbulence (>30% then *flag* 1; >75% = *flag* 2)
  - Zero mean vertical wind velocity
    (>0.10 m s-1 then former *flag* +1; >0.15 then *flag* 2)
- Flux conversions and corrections cause an interdependence of the simultaneously measured flux estimates:
  - if *flagλE* == 2 then former *flagH* + 1
    else if *flagH* == 2 then former *flagλE* + 1
    else if *flagλE* == 2 or *flagH* == 2 then former *flagNEE* + 1
- All these thresholds may appear somewhat arbitrary but are based on many years of **experience** of FW96, the authors of this work, and in part have been applied by >500 users of TK2/TK3 worldwide

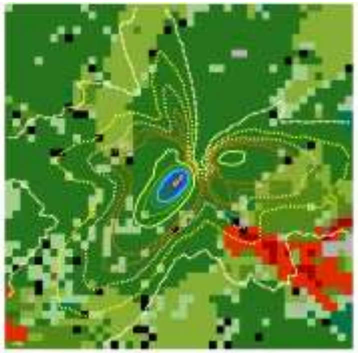
Algorithms –
Quantification of errors and uncertainty

1. **Instrumental noise**

2. **Random error**
3. **Systematic error**



4. **Representativity – Source area**

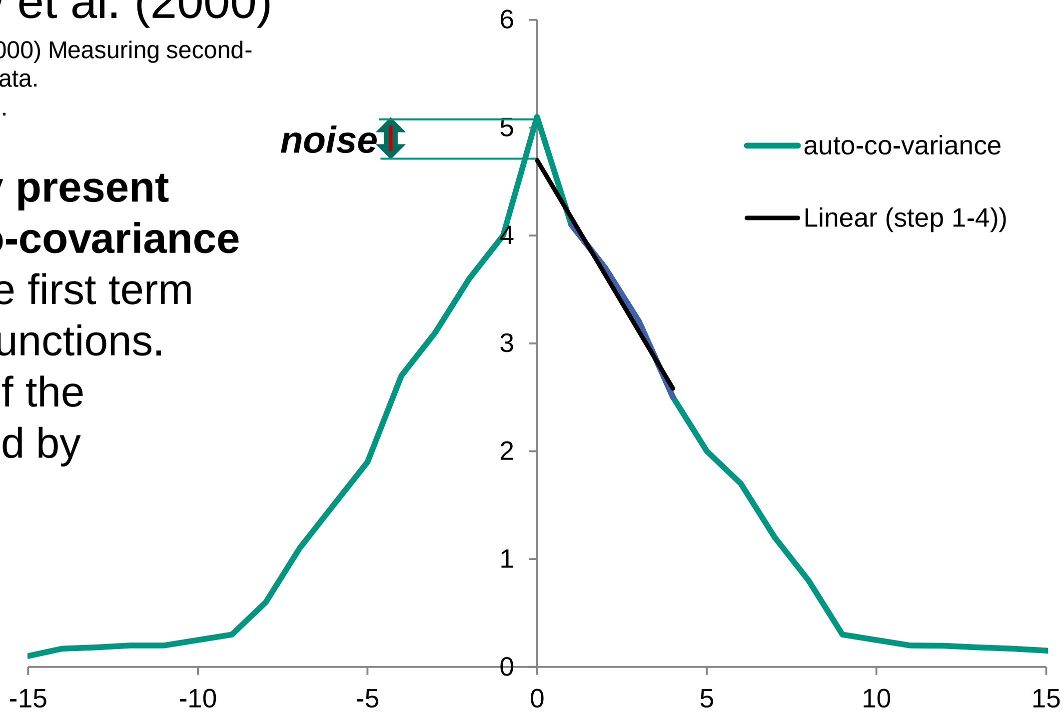


Algorithms – Quantification of errors and uncertainty

- **Instrumental noise**
  - Spectral contribution in the high-frequency range
  - Non-correlated within time series

  - Modified after Lenschow et al. (2000)
    - Lenschow DH, Wulfmeyer V, Senff C (2000) Measuring second-through fourth-order moments in noisy data. J Atmos Oceanic Technol 17:1330-1347.

    - The noise error is **only present** in the first term of **auto-covariance** functions **but not** in the first term of **cross-covariance** functions. Thus, the noise error of the covariance is calculated by **error propagation**.



Algorithms –
Quantification of errors and uncertainty

- **Random error**
  - Generally: ~ 1 / √ # **independent** observations
  - Beware: # independent observations **≠** # observations, because the time series are auto- and cross-correlated.

  - Modified after Finkelstein & Sims (2001)
  - Finkelstein PL, Sims PF (2001) Sampling error in eddy correlation flux measurements. J Geophys Res 106:3503-3509.
    - The statistical variance of a covariance can be expressed as function of its **auto-covariances and cross-covariance**
    - An autocovariance is only meaningful for data without trend => **detrending** through high-pass filter **before** calculation of random error

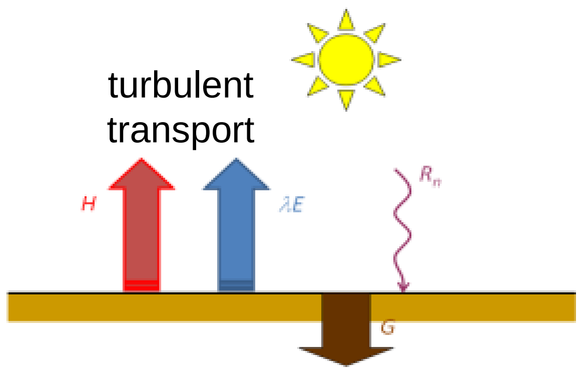
Algorithms – Quantification of errors and uncertainty

- **Systematic error**
  - In the presence of **large eddies**, the covariance at a single point **does not represent** the total surface flux
  - Why?
    - averaging interval too short
    - large eddies often “attached” to surface heterogeneities and do not propagate with the mean wind
  - Capturing all relevant scales of biosphere-atmosphere exchange is subject of current research (among others: Helmholtz Young Investigators Group)
  - Systematic error can only be determined indirectly

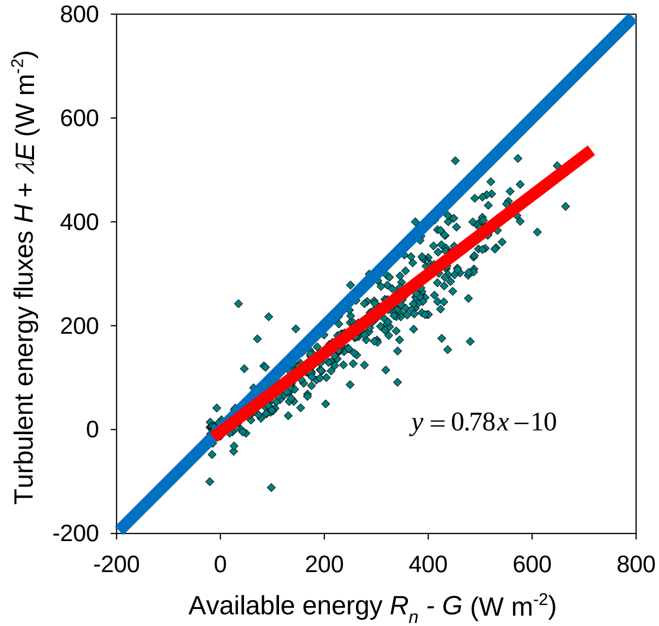
Algorithms –
Quantification of errors and uncertainty

Energy balance closure problem

$$R_n - G = \lambda E + H$$



H: sensible heat flux, R_n*:* net radiation
λE: latent heat flux, *G:* soil heat flux



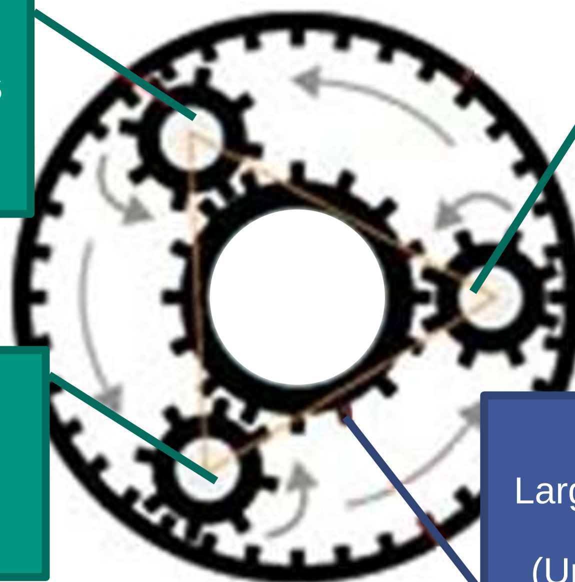
Graswang,
Juli/August
2010

- In-situ measurements worldwide show:
  underestimation of the turbulent transport (*λE* + *H)* **by 10-30%**
- Strong evidence from **180 sites** (Stoy and Mauder, 2011):
  transport on **large scales** (10^2 - 10^4 m) **not captured** by measurements
- Stoy PC, Mauder M (2011) Energy balance closure at global-distributed eddy covariance research sites: the role of landscape-level heterogeneity. In: EGU General Assembly 2011. EGU2011-3066.

Helmholtz Young Investigators Group

Integrated research approach which interlinking modeling and measurements:

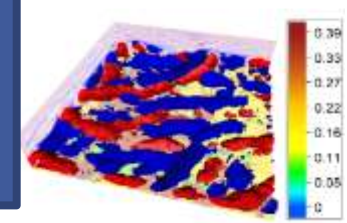
Tower-measurements
Helmholtz-TERENO Observatories



Aircraft-measurements
NRC Twin Otter (Canada)

Ground-based remote sensing
Doppler LiDAR

Modeling
Large-Eddy-Simulation LES
(Uni Hannover, DLR)



Goals

- **Quantification/parameterization** of the missing fluxes
- Improve the **understanding of the effects** of large-scale transport on flux measurements

Algorithms – Quantification of errors and uncertainty

- **Systematic error**
  - Energy balance ratio **EBR**

$$EBR = \sum_{i=1}^{K}(H_i + \lambda E_i) \Big/ \sum_{i=1}^{K}(R_{n,i} - G_i - J_i)$$

  - only applicable for daytime (R_g > 20 W m^{-2}) and for 1-day averaging
  - the error is defined as

$$\sigma_F^{sys} = F \cdot \left(\frac{1}{EBR} - 1\right) \qquad R_g > 20\ \mathrm{W\,m^{-2}}$$

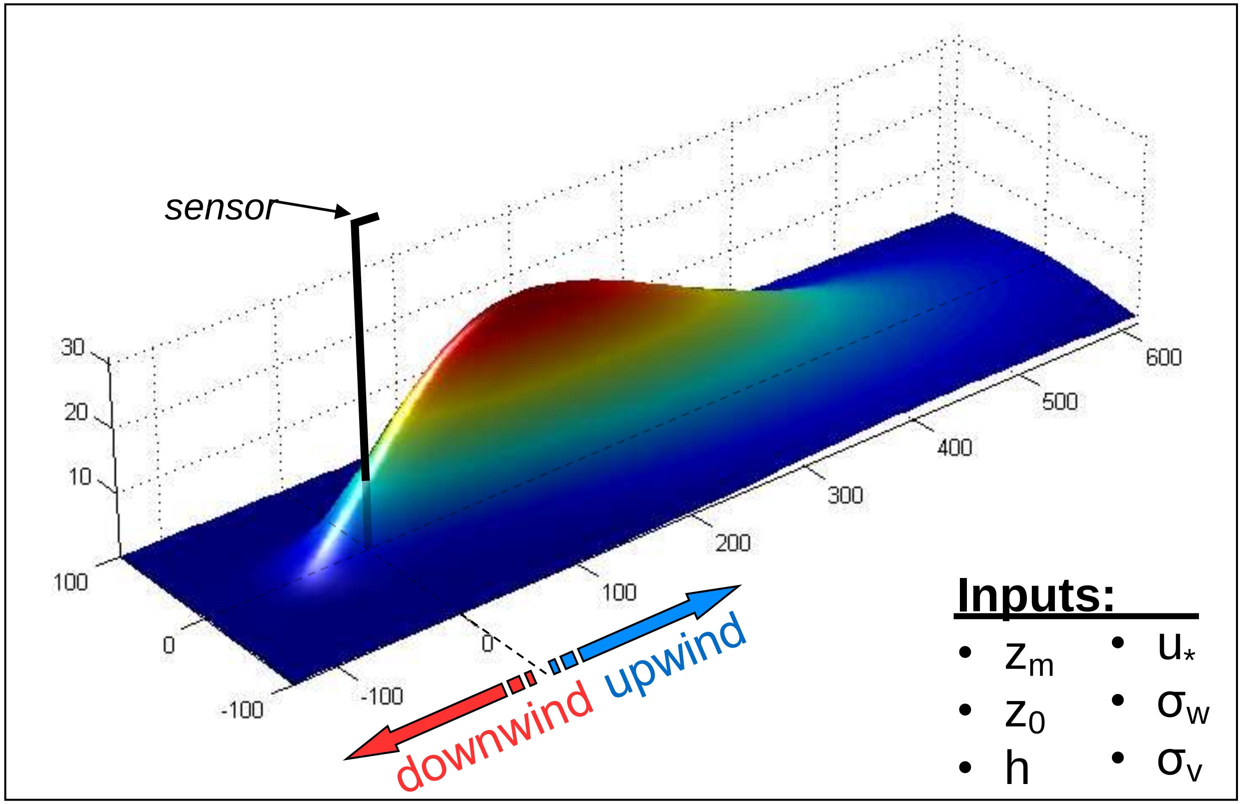
$$\sigma_F^{sys} = n/a \qquad R_g \leq 20\ \mathrm{W\,m^{-2}}$$

Algorithms – Quantification of errors and uncertainty

- **Footprint** = *spatial* ***filter***, *"field of view"*

$$F(\mathbf{x}) = \iint_{\Re} Q_s(\mathbf{x}') \cdot f(\mathbf{x}-\mathbf{x}') \cdot d\mathbf{x}' = Q_s * f$$

(**convolution** of the **source distribution**, ***Q_S***, with the **footprint**, ***f***)



Schmid HP (1994) Source areas for scalars and scalar fluxes. Bound - Layer Meteor 67:293-318.

Algorithms –
Quantification of errors and uncertainty

- **Footprint-Model**
- Kormann R, Meixner FX (2001) An analytical footprint model for non-neutral stratification. Bound -Layer Meteor 99:207-224.

$$f = \frac{1}{\Gamma(\mu)} \frac{\xi^{\mu}}{x^{1+\mu}} e^{-\xi/x}$$

$\xi(z) = \frac{U z^r}{r^2 \kappa}$, flux length scale

$\mu = (1+m)/r$. constant

$r = 2 + m - n$ shape factor

  - Analytical
  - Numerically robust (log-profiles into power laws)
  - Computationally fast and reliable enough to be applied for every 30-min interval in a long-term measurement programme

The complete quality assessment scheme

Tests on high-frequency data

- Instrument diagnostic flags (CSAT3 0-63, LI7500 240-251)
- Instrumental/plausibility limits (site-specific)
- Spike-detection with MAD-test, $z = 7$

Tests on statistics, flux calculation + corrections

- Maximum number of missing values: ≤ 10%: *flag* = 0, >10%: *flag* = 2
- Calculate averages, variances and covariances on 30 min and 5 min basis with automatic lag correction by maximising covariances
- Stationarity test covariances (FW96, < 30%: *flag* = 0, < 75%: *flag* = 1)
- Flux calculation, conversions and corrections applied: Planar fit, Schotanus, spectral losses (Moore or Eugster&Senn), WPL, iteration
- ITC test (FW96, MF04, < 30%: flag = 0, < 100%: flag = 1)
- Combination of flags according to MF04
- $\overline{w}$ after Planar fit > 0.10 m s^{-1}: former *flag* +1, > 0.15 m s^{-1}: *flag* = 2
- Interdependence of flags due to corrections/conversions:
  - if *flagλE* == 2 then former *flagH* + 1
  - else if *flagH* == 2 then former *flagλE* + 1
  - else if *flagλE* == 2 or *flagH* == 2 then former *flagNEE* + 1

Quantification of errors/uncertainty estimates

- Random error: use Finkelstein&Sims (2001), but on high-pass filtered (detrended) time series and on full period. Consider Gaussian error propagation for u_*; take relative error of covariances and multiply with corrected fluxes (neglecting error propagation through corrections)
- Noise error after Lenschow et al. (2000)
- Systematic error: Very low-frequency eddies and non-propagating eddies lead to a flux underestimation and consequent lack of energy balance closure, only applicable for R_g > 20 W m^{-2} for one day; the error is flux*(1/*EBR*−1)
- Footprint: Kormann&Meixner (2001) using measured <u> instead of z_0-dependent wind profile; calculate percentage of flux contribution for several targets of interest; recommended percentage threshold 70% depending on user's requirements

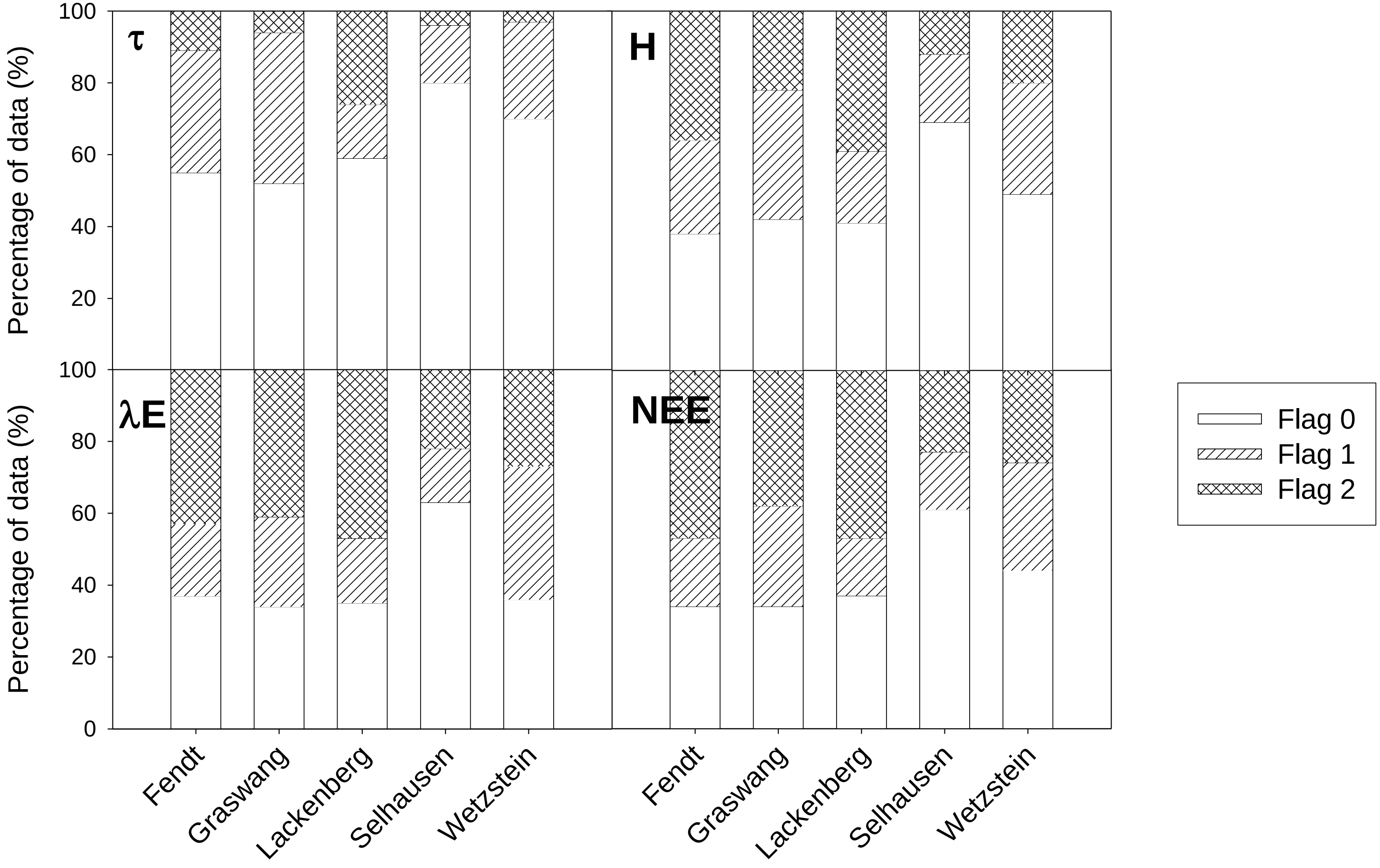
The test data sets

Overview of the test data sets

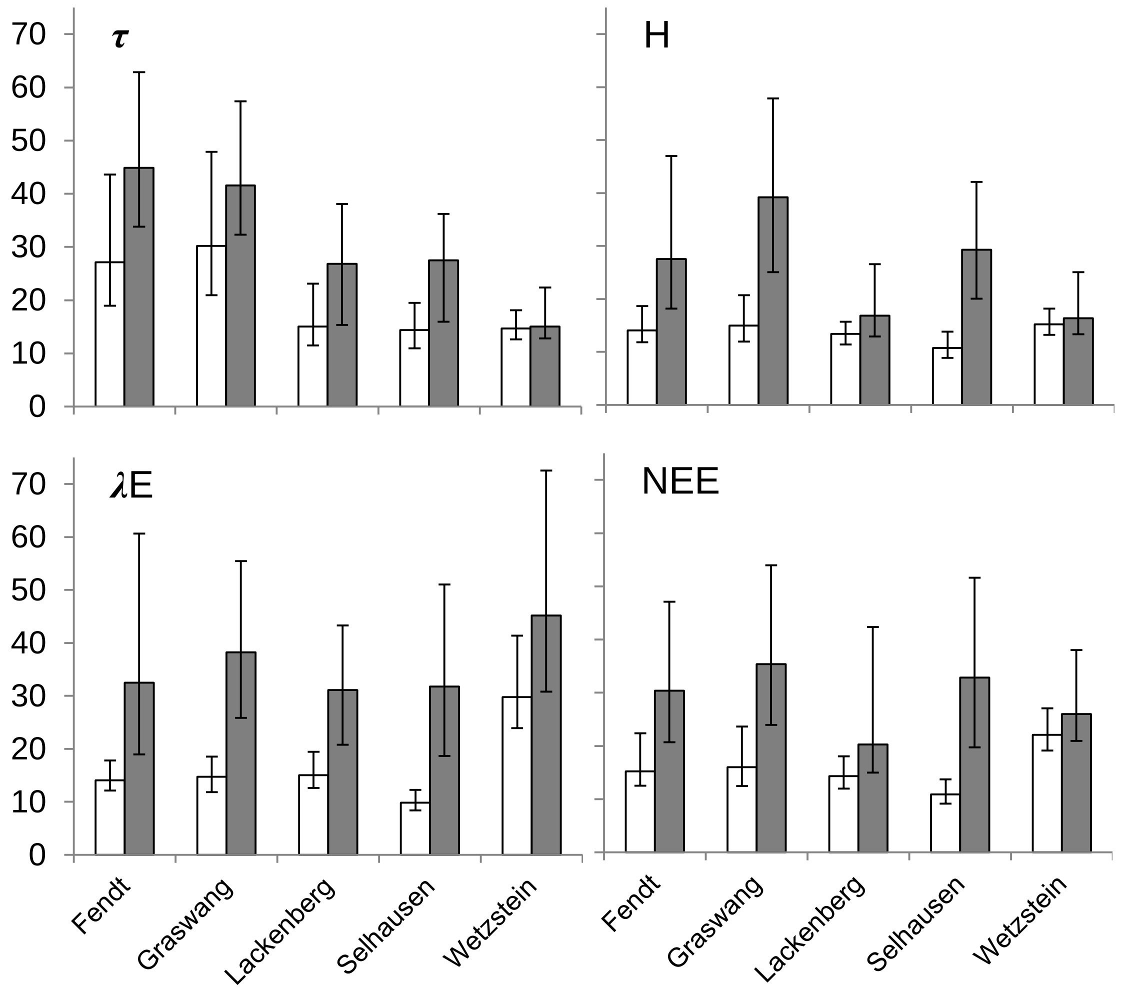
Site name	Operator	Ecosystem	Measurement height (a.g.l.*)	Sensor combination	Data period
Fendt	KIT	grassland in pre-alpine valley	3.5 m	CSAT3/ LI-7500	25/07/2010 – 23/08/2010
Graswang	KIT	grassland in pre-alpine valley	3.5 m	CSAT3/ LI-7500	25/07/2010 – 23/08/2010
Lackenberg	KIT	wind throw on low mountain range	9.0 m	CSAT3/ LI-7500	25/07/2010 – 23/08/2010
Selhausen	FZJ	agricultural land, sugar beet	2.5 m	CSAT3/LI-7500	01/06/2011 – 30/06/2011
Wetzstein	MPI-BGC	Spruce forest on low mountain range	30.0 m	Solent-R3/ LI-6262	15/07/2006 – 13/08/2006

* a.g.l.: above ground level

Results - Percentage of available data with flag 0, 1 and 2



Results –
Relative random flux error (%) vs. quality flag



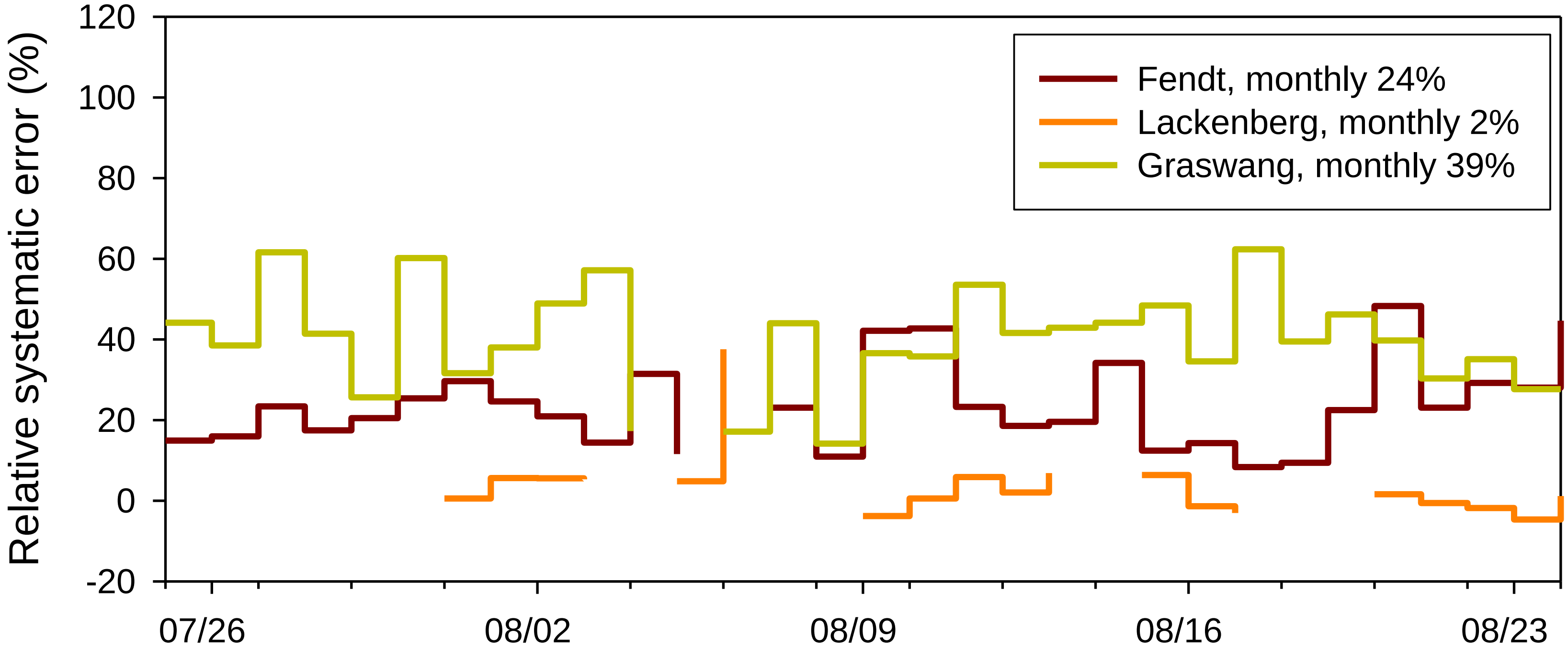
Results – Average noise errors in percent (%)

	Fendt	Graswang	Lackenberg	Selhausen	Wetzstein
τ	0.39	0.38	0.12	0.34	0.08
H	0.54	0.64	0.31	0.95	0.30
λE	0.45	0.46	0.35	0.67	0.53
NEE	0.61	0.47	0.48	0.79	0.41

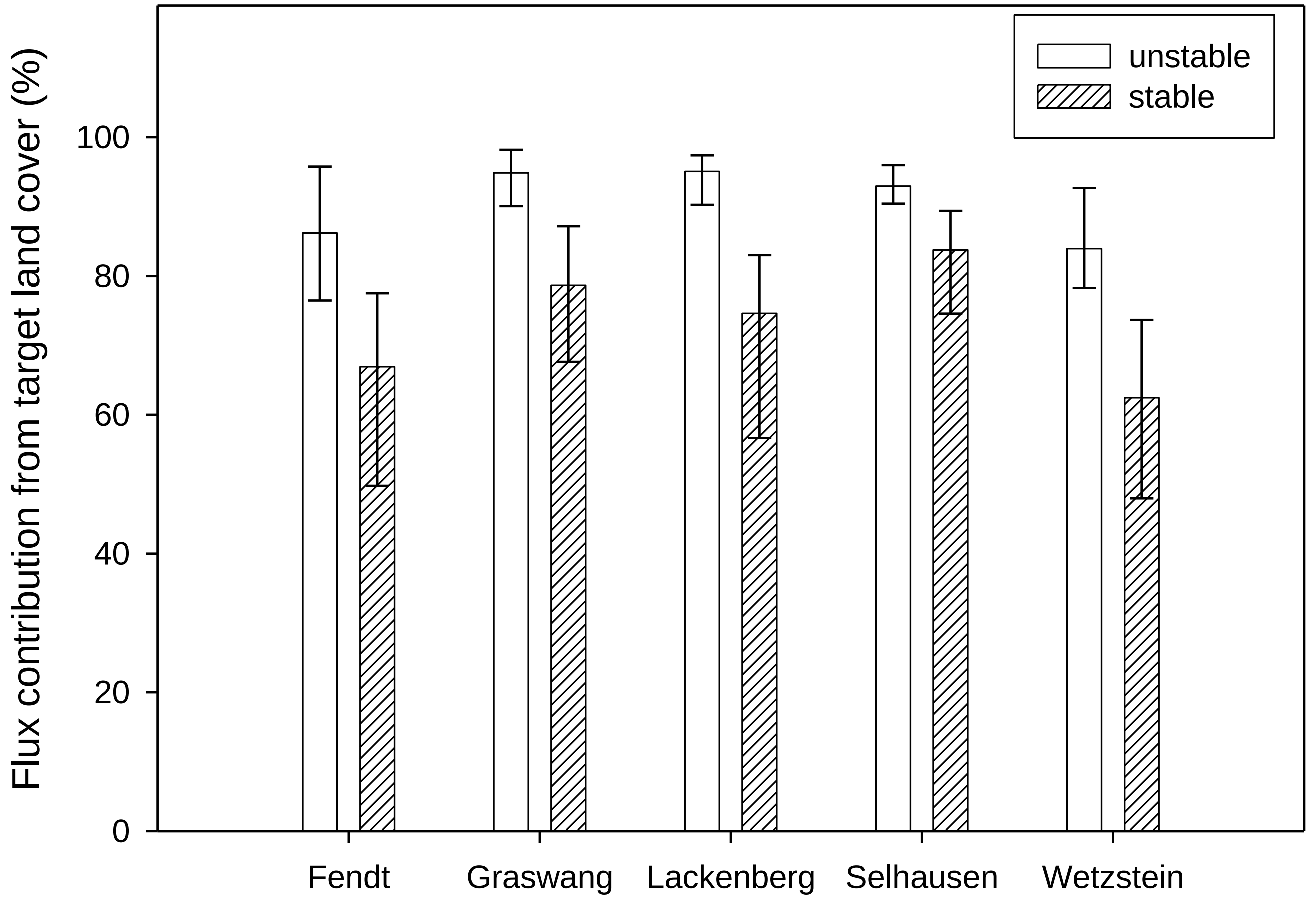
Even, if this error may be small in relation to the magnitude of the flux on average, this can be different for single estimates, e.g. the **maximum relative noise error** that we found for the five test data sets was **9%**.

Sturm P, Eugster W, Knohl A (2012) Eddy covariance measurements of CO_2 isotopologues with a quantum cascade laser absorption spectrometer. Agric Forest Meteor 152:73-82.

Results –
Systematic errors in percent (%)

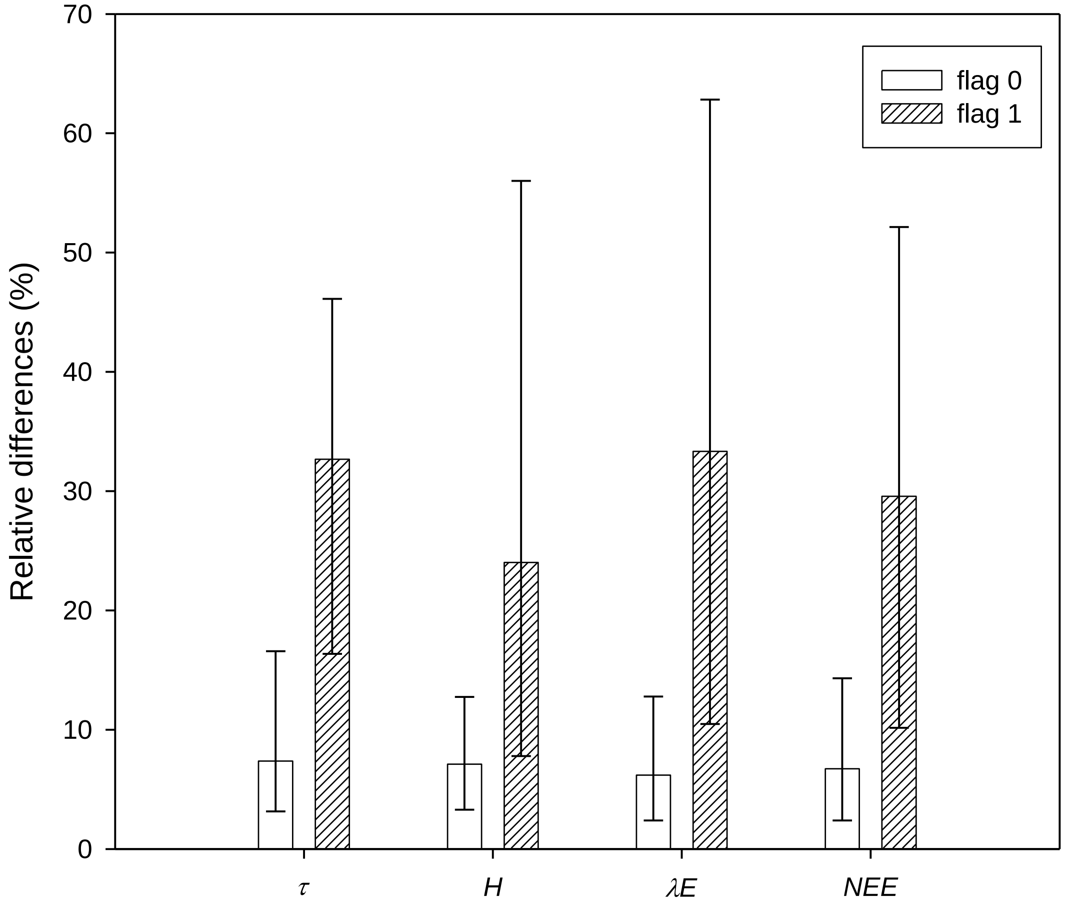


Results –
Footprint: flux contribution from target



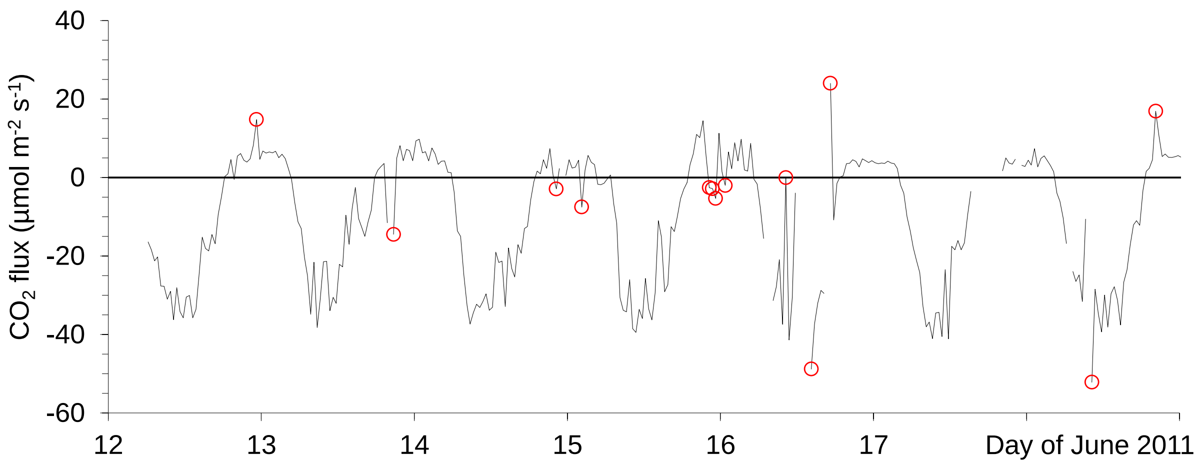
Possible further tests, not implemented in the quality assessment scheme

- Stationarity-test based on detrending vs. block-averaging



Possible further tests, not implemented in the quality assessment scheme

- Skewness and Kurtosis



Excerpt of CO_2 flux time series at the site Selhausen with visually identified strongly suspect fluxes (encircled).

Graf A, Schüttemeyer D, Geiß H, Knaps A, Möllmann-Coers M, Schween J, Kollet S, Neininger B, Herbst M, Vereecken H (2010) Boundedness of turbulent temperature probability distributions, and their relation to the vertical profile in the convective boundary layer. Bound -Layer Meteor 134:459-486
.

Vickers D, Mahrt L (1997) Quality control and flux sampling problems for tower and aircraft data. J Atmos Oceanic Technol 14:512-526.

Possible further tests, not implemented in the quality assessment scheme

- Skewness and Kurtosis

Performance of the quality assessment scheme presented in this paper and several potential additional flagging schemes on visually determined suspect fluxes.

	presented scheme		skewness and kurtosis flags			
	a) Flag 1 + 2	b) Flag 2	c) VM97	d) bimodal	e) K > β + 3	a) + d)
α: detections	3.0%	1.5%	1.7%	0.9%	0.3%	3.1%
β: missed	0.6%	2.1%	1.9%	2.7%	3.3%	0.5%
γ: false detection	24.4%	7.3%	30.8%	5.3%	8.2%	26.7%
δ: correct null	71.9%	89.1%	65.6%	91.1%	88.2%	69.6%
TSS: true skill score	0.59	0.33	0.16	0.20	0.01	0.59

Effectiveness of the quality assessment scheme

Results of the MAD-based outlier test (Papale et al. 2006) after application of the quality assessment scheme (number of detected values/available data after QC).
Before QC, 1440 data were available per test data set.

	Fendt	Graswang	Lackenberg	Selhausen	Wetzstein
τ	1/1277	5/1348	0/1044	1/1383	2/1395
H	1/916	7/1121	21/882	9/1262	19/1153
λE	2/820	5/850	7/762	13/1127	18/1059
NEE	3/757	9/888	8/765	7/1113	2/1064

Papale D, Reichstein M, Canfora E, Aubinet M, Bernhofer C, Longdoz B, Kutsch W, Rambal S, Valentini R, Vesala T, Yakir D (2006) Towards a standardized processing of Net Ecosystem Exchange measured with eddy covariance technique: algorithms and uncertainty estimation. Biogeosciences 3:571-583.

Conclusions

- An effective automatic quality assessment of long-term EC measurements was presented.
- This scheme assesses the data quality based on *fundamental* information contained in the measured data directly: no *a-posteriori* checks required.
- Regular/daily visual inspection of the data is still required.
- Expert knowledge can help to retain more data from the original data set.
- The presented quality assessment scheme will be the standard for EC-measurements within TERENO. However, further development is possible when new scientific findings become evident.
- The algorithms are newly implemented in the software TK3.1 (freeware) and will be available as a pre- and postprocessor for the software package ECpack and as part of EddySoft.

Thank you for your attention!