Chapter 9

POST-FIELD DATA QUALITY CONTROL

Thomas Foken, Mathias Göckede, Matthias Mauder, Larry Mahrt,
Brian Amiro, William Munger
thomas.foken@uni-bayreuth.de

Abstract This Chapter summarizes the steps of quality assurance and quality control of flux measurements with the eddy covariance method. An important part is the different steps of the control for electronic, meteorological and statistical problems. The fulfillment of the theoretical assumptions of the measuring method and the non-steady state test and the integral turbulence test are extensively discussed as well as an overall flagging for data quality and a site specific quality analysis using footprint models. Finally, problems are discussed which are not included yet in the control program, mainly connected with the complicated turbulence structure at a forest site.

1 Introduction

A consistent procedure for quality control of meteorological data is essential for measurement networks and long-term measurement sites. This issue has been extensively addressed for standard meteorological networks. Reliable, automated procedures based on inspection of time series which can reduce quality control efforts and provide a consistent product across measurement networks, have been the focus of several studies. Smith et al. (1996) have constructed automated quality control procedures for slow response surface data that flag questionable data points for visual inspection. Hall et al. (1991) examined the quality assurance of observations from ships and buoys using output from a numerical weather prediction model as a constraint. Lorenc and Hammon (1988) constructed an automated procedure to flag errors from ship reports, buoys and synoptic reports. They concluded that their procedure does not give completely reliable results, and that subjective analysis did better than the automated program during unusual conditions, such

X. Lee et al. (eds.), Handbook of Micrometeorology, 181–208.

as developing depressions. Essenwanger (1969) presented an automated procedure for detecting erroneous or suspicious observational records based on obvious data errors, comparison of adjacent (in time or space) data, and by comparing to prescribed limits of a standard Weibull distribution. Essenwanger (1969) concluded that his automated technique could not unequivocally pinpoint differences between a rare event and an instrument problem. DeGaetano (1997) presents a scheme to quality control wind measurements. Methods to control radiation measurements were discussed by Gilgen et al. (1994), which can be implemented into continuously running systems.

In contrast to standard meteorological measurements there are only a few papers available that discuss quality control of eddy covariance measurements (Foken and Wichura 1996, Vickers and Mahrt 1997). Quality control of eddy covariances should include not only tests for instrument errors and problems with the sensors, but also evaluate how closely conditions fulfill the theoretical assumptions underlying the method. Because the latter depends on meteorological conditions, eddy covariance quality control tools must be a combination of a typical test for high resolution time series and examination of the turbulent conditions. A second problem is connected with the representativity of the measurements depending on the footprint of the measurement. The control of the percentage of the area of interest in the actual footprint is a further issue. It is the aim of the present Chapter to describe a set of possible tests and protocol for data flagging and give practical guidance for use in continuously running eddy covariance systems like the FLUXNET program.

2 Quality Assurance and Quality Control

Quality assurance is one of the most important issues for creation and management of a measuring program. Issues of quality assurance are widely known for routine meteorological measuring programs (Shearman 1992). The present network of carbon dioxide flux sites evolved from an assemblage of individual sites with varying objectives (biological or micrometeorological) and protocols, rather than being designed from the outset as a network. Therefore, the quality assurance of such measuring programs was written after the measurements had started (e. g. Aubinet et al. 2000, Moncrieff et al. 1997). And even now some of the topics are under discussion. A quality assurance (QA) scheme needs the following components:

- Specification of user requirements: The users of the flux data, which may be modelers or policy-makers, who need the information for example in the Kyoto process, need basic information of the measuring program such as accuracy, resolution in time and space (number of sites and surface types). An important task is the development of reliable and feasible measuring programs.

- Specification of the measuring system: A suitable measuring system must be developed according to the requirements and the personal, financial and scientific constraints. This was partly done (Moncrieff et al. 1997), but presently different types of systems are used because of changes and improvements in the measuring technique. This makes the comparability of the results of different sets of instruments difficult and comparison experiments are urgently required.

- Identification of suitable measuring locations: This is a most difficult problem, because several measuring stations were created where research facilities were already in place, rather than being selected according to micrometeorological criteria. Therefore, site characterization tools are needed to ensure data quality (see Section 3.3). Ideally, site selection would be made based on quality testing of data collected from a temporary tower prior to construction of an expensive tower station.

- Definition of necessary calibrations: Calibrations allow comparison of data among sites. The accuracy of any measurement is ultimately limited by the accuracy and frequency of calibration standards that are used. Most of the necessary calibrations and control issues are well described (e. g. Aubinet et al. 2000, Goulden et al. 1996, Moncrieff et al. 1997).

- Definition of quality control (QC): The most important part of quality assurance is quality control. Several tests are discussed in this Chapter. Quality control must be done in realtime or shortly after the measurements to minimize data loss by reducing the time to detect and fix instrument problems.

- Quality evaluation: This topic is similar to QC. The main difference is a description of the data quality to be able to compare data for different periods and sites. This is also a main goal of the present Chapter.

- Corrective actions: Corrective actions refers to corrections caused by calibrations, by the choice of the coordinate system, and the sensor size and separation, etc. Most of the corrections are discussed in the other Chapters of the book and the literature (Aubinet et al. 2000, Moncrieff et al. 1997, etc.).

- Feedback from the user of the data: The database is often the end product of a measuring program. However, the user needs some control of the data and the opportunity to provide feedback to the experimentalist to improve the data quality and to make necessary changes in the program.

3 Quality Control of Eddy Covariance Measurements

A uniform scheme does not exist for quality control of eddy covariance measurements. Only several aspects are discussed in the literature. For the producer of flux data there are a number of specific techniques but no instructions for practical handling of the data. In the following, an overview of different quality control steps is given:

- The first steps of data analysis are basic tests of the raw data (Vickers and Mahrt 1997) such as electrical tests of the amplitude, the resolution of the signal, the control of the electronic and meteorological range of the data and spikes (Højstrup 1993), which are discussed further in Section 3.1.
- Statistical tests must be applied to sampling errors of the time series (Finkelstein and Sims 2001, Haugen 1978, Vickers and Mahrt 1997) and are discussed in Section 3.2. Also abrupt step changes in the time series, or reasons for non-stationarity must be identified (Mahrt 1991, Vickers and Mahrt 1997).
- A main issue for quality control are tests on fulfillment of the requirements for eddy covariance measurements. Steady state conditions and a developed turbulent regime are influenced not from the sensor configuration but from the meteorological conditions (Foken and Wichura 1996). The fulfillment of these conditions is discussed in Section 3.3.
- A system of general quality flagging of the data is discussed in Section 3.4 and a site specific evaluation of the data quality using footprint models is in Section 3.5.

3.1 Basic tests of the raw data

Vickers and Mahrt (1997) developed a framework of test criteria for quality control of fast response turbulence time series data with a focus on turbulent flux calculations. The tests are not framed in terms of similarity theory, nor do they assume that the fields necessarily follow any particular statistical distribution. Many types of instrument malfunctions can be readily identified with simple automated criteria. However, even after tuning the threshold values, the automated tests still occasionally identify behaviors that appears to be physical after visual inspection. Physically plausible behavior and instrument problems can overlap in parameter space. This underscores the importance of the visual inspection step in quality control to either confirm or deny flags raised by the automated set of tests. Data flagged but later deemed physical after graphical inspection are often found to be the most unusual and interesting situations, including intermittent turbulence, downward turbulence bursting, microfronts, gravity waves and other stable boundary layer phenomena. Some automated tests for quality control of turbulence time series are briefly summarized below.

Spikes are typically characterized as short duration, large amplitude fluctuations that can result from random noise in the electronics (Brock 1986). Quality control should include the identification and removal of spikes. For example, correlated spikes in the temperature and vertical velocity from a sonic anemometer can contaminate the calculated heat flux. Spikes that do not influence the fluxes still affect the variances. When the number of spikes becomes large, the entire data period should be considered suspect and discarded. The effect of water collecting on the transducers of some sonic anemometers often appears as spikes. Less than optimum electrical power supplies, which are sometimes necessary at remote measurement sites, can lead to frequent spiking. Unrealistic data values occur for a number of reasons. These data should be detected by comparing the minimum and maximum values to prescribed limits. For example, a vertical velocity in excess of $5\ \mathrm{m\,s^{-1}}$ close to the ground is probably not physical. However, visual inspection is sometimes required due to special circumstances, such as high turbulence levels associated with exceptionally strong surface heating. Højstrup (1993) tested a data screening procedure for application to Gaussian distributed turbulence data. Spikes are absolute quantities of measuring values which are larger than approximately four times of the standard deviation of the time series. This test should be repeated 2 or 3 times with each time series.

Some success identifying instrument problems has been achieved by comparing higher moment statistics to threshold values. Abnormally large skewness often indicates a problem, although care must be taken because, for example, the temperature near the ground during strong surface heating typically has large positive skewness. Unusually small or large kurtosis often indicates an instrument problem. Large kurtosis in the temperature field from a sonic anemometer is sometimes related to spiking associated with water on the transducers. Most despiking algorithms fail to remove this persistent type of spiking, in contrast to short duration high amplitude spikes associated with noise in the electronics. Histograms of values of a single turbulence channel are also useful. A non-typical distribution of the measuring data can indicate averaging errors connected with the digitization. Such errors were found for the Solent sonic anemometers R2 and R3 (Chr. Thomas, University of Bayreuth, 2002, personal communication, problem solved partly by Gill in 2003). In this case for example the R2 measured no vertical wind of -0.01 $\mathrm{m\,s^{-1}}$ but the number of measuring points for 0.00 $\mathrm{m\,s^{-1}}$ was twice as high as the other data. This indicates a small shift to positive vertical wind velocities.

Unusually large discontinuities in the mean can be detected using the Haar transform. The transform is simply the difference between the mean calculated between two adjacent windows. Large values of the transform identify changes in the mean that are coherent on the time scale of the window width. The goal here is to detect semi-permanent changes as opposed to smaller scale fluctuations. A sudden change of offset is one example of an instrument related jump in mean variables. The window size and the threshold values that identify suspect periods may need adjustment for particular datasets. For example, for aircraft data in the convective boundary layer, the mean vertical wind may change significantly as the aircraft enters and exits large scale coherent thermals. However, for tower measurements close to the ground, coherent changes in the mean vertical wind are typically much smaller. Care must be taken with aircraft data over heterogeneous surfaces, where coherent changes in the mean fields are common due to the formation of local internal boundary layers. For example, a sharp change in mean temperature will be found where the aircraft intersects the top of a warm internal boundary layer. In less clear cases, data from other levels and other instruments should be consulted for verification.

Instrument problems can also be detected by comparing the variance to prescribed thresholds. A sequence of variances should be calculated for a sequence of sliding, overlapping windows to detect isolated problems. For example, a brief period with near zero temperature fluctuations could be due to a temporarily non-responding instrument. Visual inspection is sometimes necessary in stable conditions where the true physical variances can become very small, usually due to a combination of strong temperature stratification and weak mean wind shear. Unusually large variance often indicates an instrument malfunction.

In recent years many closed path carbon dioxide analyzers (LiCor 6262) were replaced by open path sensors (LiCor 7500). These sensors are more sensitive to rain and frost. The development of a site-specific test using precipitation, radiation wind and temperature data can help to indicate these situations. This can be done with statistical methods like multiple regressions. Such tests can be important, because interference is not always clearly indicated in the time series.

3.2 Statistical tests

The calculation of means, variances and covariances in geophysical turbulence is inherently ambiguous, partly due to nonturbulent motions on scales which are not large compared to the largest turbulent eddies. As a result of these motions, geophysical time series are normally nonstationary to some degree (Foken and Wichura 1996, Vickers and Mahrt 1997). The physical interpretation of the flux computed from nonstationary time series is ambiguous in that it simultaneously represents different conditions and the computed perturbations for calculation of the flux are contaminated by nonstationarity, which can only be partially removed by detrending or filtering. Nonturbulent motions contaminate the flux calculation in that the flux due to nonturbulent motions may be primarily random error, as found in Sun et al. (1996). Attempts to remove nonstationarity by trend removal or filtering violates Reynolds averaging, although often the errors are small. Attempts to reduce the nonstationarity by reducing the record length increases the random flux error. Techniques for approximately separating random variations and nonstationarity are presented in Mahrt (1998) and Trevino and Andreas (2000). Tests on non-steady state conditions are given in Section 3.3.1.

Systematic errors (flux bias) result from failure to capture all of the turbulent transporting scales (Foken and Wichura 1996, Lenschow et al. 1994, Oncley et al. 1996, Vickers and Mahrt 1997). Such systematic errors occur at either the large scale end where the largest transporting eddies may be excluded from the flux calculation, or at the small scale end where transport by small eddies can be eliminated by instrument response time, pathlength averaging, instrument separation and post-process filtering. With weak winds and substantial surface heating, many flux calculation procedures may exclude larger-scale turbulent flux due to slowly moving boundary-layer scale eddies (Sakai et al. 2001). Increasing the averaging time also captures nonturbulent, mesoscale motions (nonstationarity). With very stable conditions, turbulence quantities may be confined to very short time scales, sometime less than one minute (Vickers and Mahrt 2003). Use of traditional averaging periods of five minutes or more leads to perturbation quantities, which are strongly contaminated by gravity waves, meandering motions and other mesoscale motions (see Mahrt et al. 2001a and references therein). Some of these problems can be identified with the tests given in Section 3.3.2.

The random flux error is the uncertainty due to inadequate record length and the random nature of turbulence (Finkelstein and Sims 2001, Lenschow et al. 1994, Lumley and Panofsky 1964, Mann and Lenschow 1994, Vickers and Mahrt 1997). Once perturbation quantities are computed and products are taken to compute variances, fluxes and other turbulence moments, the turbulence quantities can be averaged over a longer time period to reduce random sampling errors. The latter is sometimes referred to as the "flux-averaging time scale" to distinguish it from the shorter averaging time scale used to define the perturbations. The time scale for averaging the flux normally should be longer than that used to compute the perturbations themselves. Reynolds averaging can still be satisfied as long as the averaging is unweighted (no filtering or detrending) (Mahrt et al. 2001b). For example, one might choose an averaging time of 2 minutes for very stable conditions but wish to average the 2-minute fluxes over 30 minutes or one hour to reduce random flux errors.

With very stable conditions where the turbulence is intermittent, reduction of the random error to acceptable levels may require a prohibitively long averaging time (e. g. Haugen 1973). The flux for a one-hour period can be dominated by one or two events and therefore a much longer averaging time is required. Howell and Sun (1999) choose the record length by attempting to objectively maximize the flux and minimize the random flux error.

The above results also apply to analysis of turbulence quantities from moving platforms such as aircraft, except that one must determine the averaging length from which to compute perturbations (often chosen to be 1 km) and choose the flux averaging length, sometimes chosen as the flight path length. In convective conditions with deep boundary layers, such an averaging length may exclude significant flux (Betts et al. 1990, Desjardins et al. 1992). The nonstationarity problem above becomes the heterogeneity problem for moving platforms (e. g. Desjardins et al. 1997). Reduction of random flux errors is facilitated by long flight paths for homogeneous surfaces or many repeated passes over heterogenous surfaces (Mahrt et al. 2002).

The autocovariance analysis is widely used to determine the time lag for closed-path gas analyzers (Leuning and Judd, 1996), because the concentration signal is measured some seconds later than the wind signal. Even data from open-path gas analyzer may have a small time offset between the measuring time and the position of the value in the data file because of electronic delays in recording and storing the data and finite signal processing times. If this is not known and not corrected in the logger program, it must be included in calculation of the fluxes. It is important to check the whole measuring system with an autocovariance analysis to identify time shifts between the signals.

3.3 Tests on fulfillment of theoretical requirements

The widely used direct measuring method for turbulent fluxes is the eddy covariance method, which involves a simplification of turbulent conservation equations for momentum and scalar fluxes, e. g., the flux of a scalar, c

$$F_c = \overline{w'c'} = \frac{1}{N-1}\sum_{k=0}^{N-1}\left[(w_k - \overline{w})(c_k - \overline{c})\right] \tag{9.1}$$

where w is the vertical wind component. This equation implies steady-state conditions. The choice of averaging length depends on the cospectra of the turbulence and steady state conditions. With an ogive test (Oncley et al., 1990)

$$\mathrm{Og}_{w,c}(f_o) = \int_{-\infty}^{f_o} \mathrm{Co}_{w,c}(f)\mathrm{d}f \tag{9.2}$$

where Co is the cospectra of the vertical wind velocity and the concentration. The convergence of Og at low frequencies indicates that all relevant eddies are collected. On the other hand an excessive measuring length may include nonsteady-state conditions (see Chapters 2 and 5). Therefore, these conditions should be tested for each time series, because they can influence the data quality significantly (see Section 3.3.1). However, in most cases, convergence occurs within a 30-minute period.

The integral turbulence characteristics in the surface layer may depend on the latitude (Johansson et al. 2001); this may be relevant for tests on eddy covariance measurements. The influence of density fluctuations can be corrected (see Chapters 6 and 7). Conditions of horizontal homogeneity must also be fulfilled in order to avoid significant advection, which can be influenced by the choice of the coordinate rotation (see Chapters 3 and 10).

Of greater importance is whether developed turbulent conditions exist, with very weak turbulence the measuring method and methods based on surface layer similarities may not be valid. Examination of normalized standard deviations (integral turbulence characteristics, see Section 3.3.2) provides an effective test for adequately developed turbulence. These tests are also sensitive to other influences on the data quality like limitations of the surface layer height, gravity waves, internal boundary layers, flow distortion, high frequency flux loss (see Chapter 4). For example, internal boundary layers and flow distortion problems of the sensors and towers can indicate higher standard deviations of turbulence parameters. For situations with gravity waves the correlation coefficient between the vertical wind velocity and scalars can be high, resulting in unusually large fluxes. Such situations, often during the night and under stable conditions, must be indicated and the wave and the turbulent signal must be separated (Handorf and Foken 1997).

Foken and Wichura (1996) applied criteria to fast-response turbulence data to test for non-stationarity and substantial deviations from flux-variance similarity theory, whether due to instrumental or physical causes. These are described below.

3.3.1 Steady state tests

Steady state conditions means that all statistical parameters do not very in time (e. g., Panofsky and Dutton, 1984). Typical non-stationarity is driven by the change of meteorological variables with the time of the day, changes of weather patterns, significant mesoscale variability, or changes of the measuring point relative to the measuring events such as the phase of a gravity wave. The latter may occur because of changing footprint areas, changing internal boundary layers (especially internal thermal boundary layers in the afternoon), or by gravity waves. Presently there are two main tests used to identify non-steady state conditions. The first is based on the trend of a meteorological parameter over the averaging interval of the time series (Vickers and Mahrt, 1997) and the second method indicates non-steady state conditions within the averaging interval (Foken and Wichura, 1996).

Vickers and Mahrt (1997) regressed the meteorological element x over the averaging interval of a time series and determined the difference of x between the beginning and the end of the time series according to this regression, δx. With this calculation they determined the parameter of relative non-stationarity, mainly for wind components

$$\mathrm{RN}_x = \frac{\delta x}{\overline{x}} \tag{9.3}$$

Measurements made over the ocean exceeded the threshold ($\mathrm{RN}_x > 0.50$) 15 % of the time and measurements over forest exceeded the threshold

55 % of the time. A more rigorous measure of stationarity can found in Mahrt (1998).

The steady state test used by Foken and Wichura (1996) is based on developments of Russian scientists (Gurjanov et al., 1984). It compares the statistical parameters determined for the averaging period and for short intervals within this period. For instance, the time series for the determination of the covariance of the measured signals w (vertical wind) and x (horizontal wind component or scalar) of about 30 minutes duration will be divided into $M = 6$ intervals of about 5 minutes. N is the number of measuring points of the short interval ($N = 6{,}000$ for 20 Hz scanning frequency and a 5 minute interval):

$$(\overline{x'w'})_i = \frac{1}{N-1}\left[\sum_j x_j w_j - \frac{1}{N}\sum_j x_j \sum_j w_j\right]$$

$$\overline{x'w'} = \frac{1}{M}\sum_i (\overline{x'w'})_i \tag{9.4}$$

This value will be compared with the covariance determined for the whole interval:

$$(\overline{x'w'})_o = \frac{1}{M(N-1)}\left[\sum_i\Big(\sum_j x_j w_j\Big)_i - \frac{1}{MN}\sum_i\Big(\sum_j x_j \sum_j w_j\Big)_i\right] \tag{9.5}$$

The authors proposed that the time series is steady state if the difference
n covariances

$$\mathrm{RN_{cov}} = \left|\frac{(\overline{x'w'}) - (\overline{x'w'})_o}{(\overline{x'w'})_o}\right| \tag{9.6}$$

is less than 30%. This value is found by long experience and is in a good agreement with other test parameters also of other authors (Foken and Wichura, 1996).

3.3.2 Test on developed turbulent conditions

Flux-variance similarity is a good measure to test the development of turbulent conditions. This similarity means that the ratio of the standard deviation of a turbulent parameter and its turbulent flux is nearly constant or a function of stability. These so-called integral turbulence characteristics are basic similarity characteristics of the atmospheric turbulence (Obukhov 1960, Wyngaard et al. 1971) and are routinely discussed in boundary layer and micrometeorology textbooks (Arya 2001, Foken 2003, Kaimal and Finnigan 1994, Stull 1988). Foken and Wichura

Table 9.1. Coefficients of the integral turbulence characteristics (Foken et al. 1997, Foken et al. 1991, Thomas and Foken 2002).

Parameter	z/L	c_1	c_2
σ_w/u_*	$0 > z/L >$ -0.032	1.3	0
	-0.032 $> z/L$	2.0	1/8
σ_u/u_*	$0 > z/L >$-0.032	2.7	0
	-0.032 $> z/L$	4.15	1/8
σ_T/T_*	0.02 $< z/L <$1	1.4	-1/4
	0.02 $> z/L >$-0.062	0.5	-1/2
	-0.062$> z/L >$-1	1.0	-1/4
	-1 $> z/L$	1.0	-1/3

Table 9.2. Coefficients of the integral turbulence characteristics for wind components under neutral conditions (Thomas and Foken 2002).

Parameter	$-0.2 < z/L < 0.4$
σ_w/u_*	$0.21\ln(\frac{z_+ \times f}{u_*}) + 3.1,\ z_+ = 1$ m
σ_u/u_*	$0.44\ln(\frac{z_+ \times f}{u_*}) + 6.3,\ z_+ = 1$ m

(1996) used functions determined by Foken et al. (1991). These functions depend on stability and have the general form for standard deviations of wind components

$$\frac{\sigma_{u,v,w}}{u_*} = c_1 (\frac{z}{L})^{c_2} \tag{9.7}$$

where u is the horizontal or longitudinal wind component, v the lateral wind component, u_* the friction velocity and L the Obukhov length. For scalar fluxes the standard deviations are normalized by their dynamical parameters (e. g., the dynamic temperature T_*)

$$\frac{\sigma_x}{X_*} = c_1 (\frac{z}{L})^{c_2} \tag{9.8}$$

The constant values in Equations 9.7 and 9.8 are given in Table 9.1. For the neutral range the external forcing assumed by Johansson et al. (2001) and analyzed for the integral turbulence characteristics by Thomas and Foken (2002) was considered in Table 9.2 with the latitude (Coriolis parameter f). The parameters given for the temperature can be assumed for most of the scalar fluxes. It must be mentioned that under nearly neutral conditions the integral turbulence characteristics of the scalars have extremely high values (Table 9.1) and the test fails.

Table 9.3. Typical values for the correlation coefficient of the momentum and sensible heat flux.

Author	r_{uw}	r_{wT}
Hicks (1981)	-0.32	0.35 ($z/L \rightarrow -0.0$) 0.6 ($z/L \rightarrow -2.0$)
Kaimal et al. (1990)	-0.3	0.5 ($z/L < 0.0$)
Kaimal and Finnigan (1994)	-0.35	0.5 ($-2 < z/L < 0$) -0.4 ($0 < z/L < 1$)
Arya (2001)	-0.15	0.6 ($z/L < 0.0$)

The test can be done for the integral turbulence characteristics of both parameters used to determine the covariance. The measured and the modeled parameters according to Equations 9.7 or 9.8 will be compared according to

$$\mathrm{ITC}_\sigma = \left| \frac{(\sigma_x/X_*)_{\text{model}} - (\sigma_x/X_*)_{\text{measurement}}}{(\sigma_x/X_*)_{\text{model}}} \right| \qquad (9.9)$$

If the test parameter ITC_σ is < 30 %, a well developed turbulence can be assumed.

A similar parameter is the correlation coefficient between the time series of two turbulent parameters. If this correlation coefficient is within the usual range (Table 9.3) a well-developed turbulence can be assumed (Kaimal and Finnigan, 1994).

3.4 Overall quality flag system

To be useful, the results of data quality checking must be made available in the final data archive. Measurements are normally flagged according to their status such as uncontrolled, controlled, corrected, etc. The quality tests given above open the possibility to flag also the quality of a single measurement. Foken and Wichura (1996) proposed to classify the tests according to Equations 9.6 and 9.9 into different steps and to combine different tests. An important parameter, which must be included in the classification scheme, is the orientation of the sonic anemometer, if the anemometer is not an omnidirectional probe and the measuring site does not have an unlimited fetch in all directions. For these three tests the definition of the flags is given in Table 9.4. Further tests, like an acceptable range of the mean vertical wind velocity, can be included into this scheme.

The most important part of a flag system is the combination of all flags into a general flag for easy use. This is done in Table 9.5 for the

Table 9.4. Classification of the data quality by the steady state test according to Equation 9.6 and the integral turbulence characteristics according to Equation 9.9 and the horizontal orientation of a sonic anemometer of the type CSAT3 (Foken 2003).

a		b		c	
class	range	class	range	class	range
1	0-15%	1	0-15%	1	±0-30°
2	16-30%	2	16-30%	2	±31-60°
3	31-50%	3	31-50%	3	±61-100°
4	51-75%	4	51-75%	4	±101-150°
5	76-100%	5	76-100%	5	±101-150°
6	101-250%	6	101-250%	6	±151-170°
7	251-500%	7	251-500%	7	±151-170°
8	501-1000%	8	501-1000%	8	±151-170°
9	>1000%	9	>1000%	9	> ±171°

a: State-state test according to Equation 9.6.

b: Integral turbulence characteristics according to Equation 9.9.

c: Horizontal orientation of the sonic anemometer.

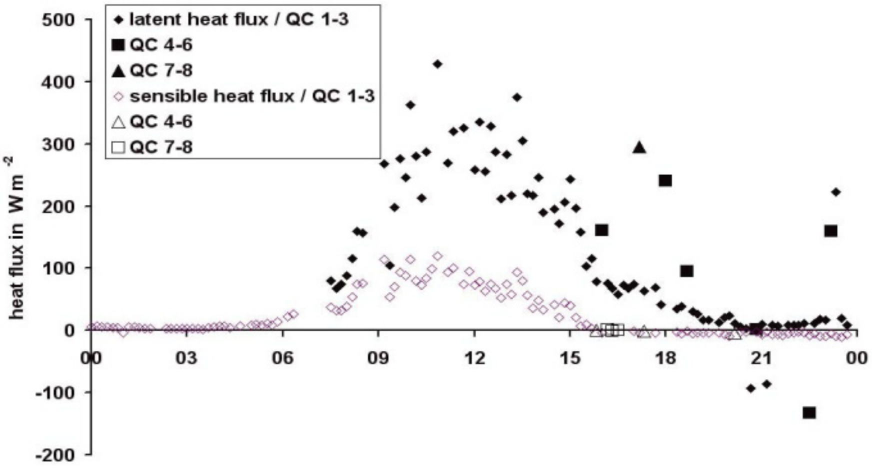

Figure 9.1. Daily cycle of the sensible and latent heat flux with quality classes measured by the University of Bayreuth during the LITFASS-1998 experiment (Beyrich et al. 2002) on June 02, 1998 in Lindenberg/Germany over grassland.

Table 9.5. Proposal for the combination of the single quality flags into a flag of the general data quality (Foken 2003).

a	b	c	d
1	1	1-2	1-5
2	2	1-2	1-5
3	1-2	3-4	1-5
4	3-4	1-2	1-5
5	1-4	3-5	1-5
6	5	$\leq$5	1-5
7	$\leq$6	$\leq$6	$\leq$8
8	$\leq$8	$\leq$8	$\leq$ 8
9	*	*	*

a: Flag of the general data quality

b: Steady state test according to Equation 9.6

c: Integral turbulence characteristics according to Equation 9.9

d: Horizontal orientation of the sonic anemometer

$*$: One or more of flags b, c and d equals 9

flags given in Table 9.4. The user of such a scheme must know the appropriate use of the flagged data. The presented scheme was classified (by micrometeorological experiences) so classes 1 to 3 can be used for fundamental research, such as the development of parameterizations. The classes 4-6 are available for general use like for continuously running systems of the FLUXNET program. Classes 7 and 8 are only for orientation. Sometimes it is better to use such data instead of a gap filling procedure, but then these data should not differ significantly from the data before and after these data in the time series. Data of class 9 should be excluded under all circumstances. Such a scheme gives the user a good opportunity to use eddy covariance data. Finally the data can be presented together with the quality flag like in Figure 9.1. Most of the unusual values can be explained by the data quality flag. At night, other reasons can influence the measurements. For analysis of integrated fluxes rejected data will need to be filled in. Obviously, investigations to infer process relationships should exclude both flagged data and the gap-filled values.

3.5 Site dependent quality control

Besides the quality classification of a single measurement series, classification of the site-specific data quality is needed to compare different sites within a network like FLUXNET for a better interpretation of experimental and modeled data. The data quality differs because of topography and this must be taken into account by comparison of the data quality. This quality check was developed to include footprint information (Foken et al. 2000). There are two different points of interest: The first point is the area of interest (e. g. a spruce forest) in the footprint of the measurements. The second points concerns the question: for which footprint areas can good data quality be assumed?

A program package has been developed (Göckede et al. 2003) and used for 18 CarboEurope eddy covariance measuring sites. The land use information of the surrounding is given by input matrices. Together with necessary meteorological input parameters, the main iteration loop of the program starts with a footprint calculation employing a user-defined start value for the roughness length z_0. The integrated Schmid (1997) model produces characteristic dimensions defining the two-dimensional horizontal extension of each so-called effect-level ring. Using these dimensions, which sketch a discrete version of the source weight function, it is possible to assign a weighting factor to each of the cells of the roughness matrix. A new roughness length z_0-final is calculated as the mean value of all the cells within the source area under consideration of the weighting factors. The iteration loop starts again with the improved value of z_0-final as the input value for the footprint routine. In the next step, the land use structure within the computed source area is analyzed. The weighting factors of the last source weighting function results are used to calculate the contribution of each type of land use (which can be up to 20, as defined by the user) to the total flux. Due to certain restrictions of the footprint model concerning the necessary input parameters, a portion of the input data set cannot be processed. Most of the time, these problems occur during stable stratification, when the computed source area grows to an extent that makes the numerical algorithms unstable. Finally figures like Figure 9.2 for the Weidenbrunnen/Waldstein site near Bayreuth/Germany (50°08'N, 11°52'E, 775 m a.s.l.), can be constructed that give a flux distribution over a four month measuring period that depends on the footprint. The color of the grid elements characterize the part of the area of interest to the flux. Such pictures can help find the best wind directions and the best positions of the tower to link the fluxes with the underlying surface.

To produce the overall performance of the flux data quality for a specific site, the results of all the footprint calculations are combined with the data quality assessment. The products of the procedure are two-dimensional matrices and graphs that form a combination of all the footprint analyses for the specific site. These matrices show, for example, the dominating data quality class for each of the grid cells (mean value) of the matrix surrounding the tower, in combination with its contribu-

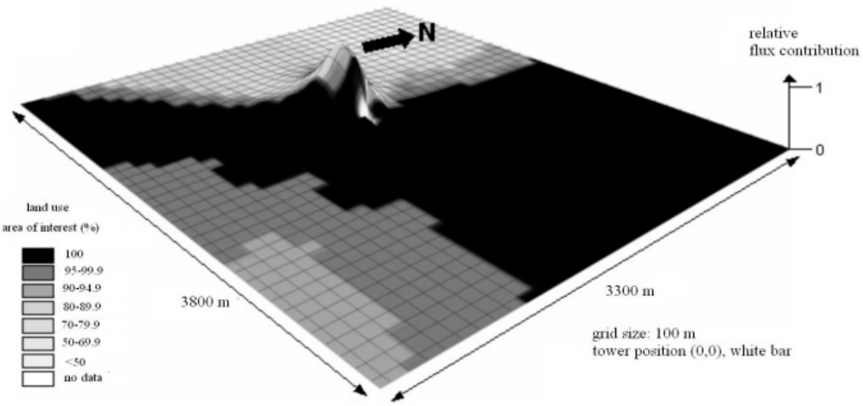

Figure 9.2. Quality analysis for the land use evaluation with flux contribution. Results were obtained with data from the Weidenbrunnen/Waldstein site for the period 01.05. – 31.08.1998 (Göckede et al. 2002).

tion to the total flux. This can be done for all types of fluxes. Only for scalar fluxes the quality flag of internal turbulence characteristics must be excluded in the near neutral case. As an example, the data quality distribution for the latent heat flux of Weidenbrunnen/Waldstein site is given in Figure 9.3. The lower data quality in western wind directions is caused by a clearing, which can also be indicated from the land use distribution (Figure 9.2). The low data quality in SWS direction (for stable stratification) is caused by the Waldstein mountain at a distance of 1.5 km. The possibility to bring data quality and possible influencing factors together is an application of the footprint model. Using the limit settings, the user of the program package can restrict the analysis to certain quality classes or a range of values for specific meteorological parameters, allowing a more detailed analysis under special conditions. The variation of these input parameters can also be performed automatically in a sequence mode with user defined upper and lower limits at specific increments.

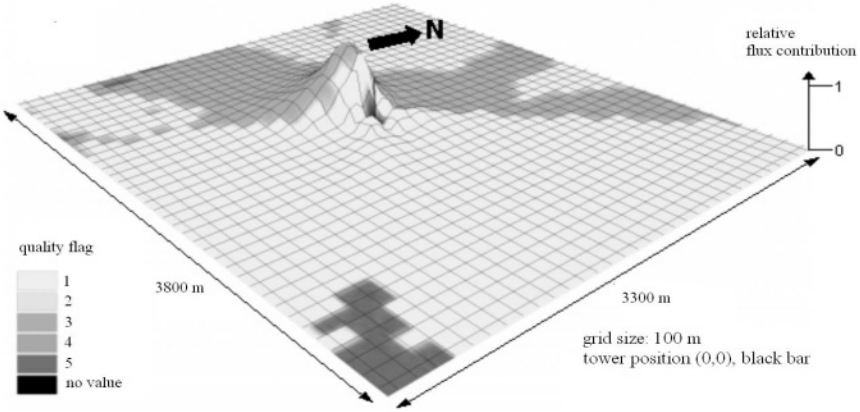

Figure 9.3. Quality flags for special distribution of the contribution to the latent heat flux. Results were obtained with data from the Weidenbrunnen/Waldstein site for the period 01.05.–31.08.1998 (Göckede et al. 2002).

4 Further Problems of Quality Control

Energy balance closure has often been used to identify the quality of eddy covariance measurements (Aubinet et al. 2000). For most of the sites a closure of the energy balance equation

$$R_n - H - \lambda E - G \pm \Delta S = \text{Res} \tag{9.10}$$

with R_n net radiation, H sensible heat flux, λE latent heat flux, G ground heat flux, ΔS heat storage, is not zero but has a residual Res of approximately 10-20%. In some investigations of the energy balance closure problem (Culf et al. 2003, Foken and Oncley 1995, Oncley et al. 2002), the main reasons for this problem are errors of the sensors. For example the influence of net radiometers is significant because of the large part of net radiation in the energy balance. Measuring problems also exist of heat storage especially in the soil layer above the heat flux plates. Another reason is that mesoscale fluxes are not measured (Chapter 5. These reasons for the residual of the energy balance closure do not allow an energy balance closure as a correction factor for all turbulent fluxes or the use of energy balance closure as a measure of the data quality. However, there are many other studies where energy balance closure is consistently underestimated, without an identifiable cause. This has created some disparity among the methods employed by different groups. Some researchers use the energy balance closure as a further check, and adjust the CO_2 flux in the same proportion as the loss in the other turbulent fluxes (e. g., Amiro 2001, Barr et al. 2002). Some other researchers do not account for this turbulent loss, and consensus has not been reached in the research community. As an additional problem different instruments have different footprints.

The method of coordinate rotation also influences the data. Such rotations are necessary to align the x-axis with the mean wind (first rotation), to define a z axis so that the mean vertical wind component is zero (second rotation) and to rotate the system on the third axis so that the lateral momentum flux is zero (third rotation). This method was discussed by McMillen (1988) with a running mean as the reference coordinate system. Presently a rotation for each averaging interval (30 minutes) without the third rotation is proposed (Aubinet et al. 2000). This method is widely criticized because single events like convection, gusts, coherent structures etc., which have nothing to do with the coordinate system, are the reason for a significant rotation for a particular averaging interval. Even over low vegetation and flat terrain, rotation angles of 20-40° can be detected in the night and early morning hours. Therefore the planar-fit method (Wilczak et al. 2001) has been suggested (see Chapter 3) which rotates according the mean streamlines (Paw U et al. 2000). This streamline dependent coordinate system must be determined for one site and changes only with changes in the mounting of the sensor, with the time of the year (deciduous forest), with the wind speed (two classes) and the wind direction in heterogeneous and hilly terrain. The rotation angles are small and on the order of 2-5° and can be more with significant slope. After the rotation the data quality analysis as described in Sections 3.3 and 3.4 produces significant differences in the data quality especially for low wind velocities. As shown in Figure 9.4, the data quality is significantly lower for double rotation in comparison to planar fit in the classes of low friction velocity. The first method has low quality data typically for $u_* < 0.3\ \mathrm{m\,s^{-1}}$ whereas the planar fit corresponds to approximately $u_* < 0.2\ \mathrm{m\,s^{-1}}$. This influence must be recognized, because it can influence the so-called u_*-criteria to correct nighttime carbon dioxide fluxes (Goulden et al., 1996).

Quality control procedures identify periods of unsuitable data, leaving non-random gaps in the dataset. The quality control procedures, instrument malfunctions, maintenance and calibration periods often remove 20 to 40% of the data. These gaps need to be filled for applications where long-term integrations are needed, though gaps should not be

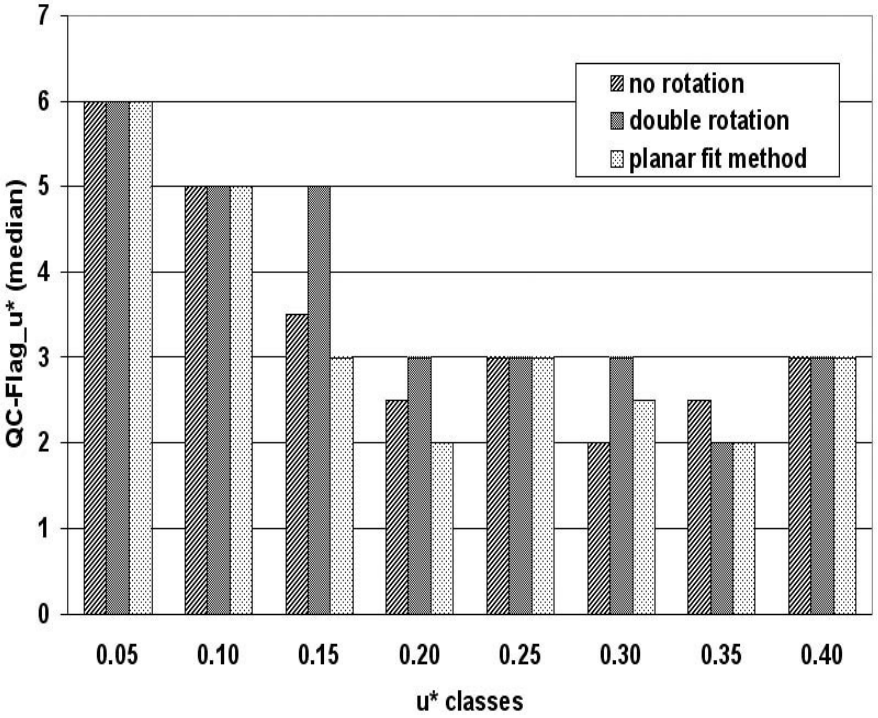

Figure 9.4. Data quality analysis for double rotation (Aubinet et al. 2000) and planar fit rotation (Wilczak et al. 2001) for measurements over an irrigated cotton field during EBEX-2000 (Oncley et al. 2002).

filled for process studies. Gap-filling creates additional uncertainty in the data, and there will always be a compromise between the use of possibly questionable flux data and replacement with values generated from a gap-filling algorithm. Confidence in gap-filling increases with knowledge and experience at any given flux site.

Falge et al. (2001a, 2001b) provide reviews of gap-filling strategies for energy flux and net ecosystem exchange (NEE) measurements. A variety of methods need to be applied, depending on the reasons for the gap creation. Nighttime gaps in NEE are best filled by either using soil respiration chambers or through developing a site-specific relationship between the respiration flux (mostly soil) and environmental variables such as soil temperature and moisture. Missing daytime NEE data can be estimated using physiological relationships that typically incorporate air temperature and light measurements. Short (e. g., a single half-hour period) gaps are usually filled through interpolation, whereas longer gaps may be estimated using the average of some period of good data for the same time of day. Gap-filling by averaging also needs to consider that gaps are often created by environmental conditions differing from the average, such as instrument malfunctions during heavy precipitation. The implications of gap-filling can be substantial, and in the case of NEE, can change the conclusions on the magnitude of annual carbon sequestration. Falge et al. (2001a) compared some gap-filling methods for NEE for 18 sites, illustrating that different methods could alter the annual sum of NEE by -45 to +200 $\mathrm{g\,C\,m^{-2}}$, a significant portion of the total flux for some ecosystems. The conclusion is that quality-procedures need to focus on truly incorrect data since there is still a large uncertainty in filling gaps, and that the estimation of long-term fluxes can best be improved with good knowledge of the site processes.

Over a forest site the turbulence structure is very complicated (Amiro, 1990) sometimes with ramp structures mainly at daytime and wave structures (gravity waves) at nighttime (Chapter 8). The contribution of coherent structures to the whole flux is generally unknown. Well-organized ramp structures may be measured with the eddy covariance method. The determination of the flux due to ramp structures with the surface renewal method (Snyder et al., 1996) compares well with eddy covariance measurements (Rummel et al. 2002). In contrast, single coherent structures can indicate non-stationary conditions and be identified falsely as low quality data. We need continuously running procedures to calculate and control fluxes under these circumstances.

The decoupling of the atmosphere from the forest also needs to be considered. This is a typical situation during stable stratification at night. One must also consider the possibility of a mixing layer immediately above the forest canopy (Finnigan 2000, Raupach et al. 1996), caused by the high wind shear above the forest. The similarity analysis of the length scales of the shear layer and the coherent structures show that the forest and the atmosphere are often only coupled at daytime, often with strong coherent structures (Wichura et al. 2002).

One must also consider the mean transport at the upper boundary of a control volume (Chapter 10). Also the horizontal and vertical advective transport must be taken into account to interpret the vertical flux. An adequate choice of the coordinate system, for instance by planar fit rotation can help to interpret the vertical advection. Nevertheless, these site specific phenomena are difficult to check through automatic quality control procedures.

Plant physiological tests and ecosystem level measurements of carbon or water budgets can also be very useful in verifying the quality of the flux data. For example, soil chambers can give nighttime estimates of respiration during periods of weak turbulence when micrometeorological conditions fail. Plant leaf chambers can confirm the response of plants to certain conditions when turbulent flux measurements are questioned. Biomass inventories (Curtis et al. 2002) provide additional checks on annual integrals of flux data. The best possible estimate of net ecosystem exchange should combine a consistent set of independently determined quantities.

5 Conclusion

The quality assurance and quality control are outstanding problems that are incompletely fulfilled in most of the FLUXNET networks. For new stations a complete quality assurance plan can help provide a measuring system that can run within a short time on a high quality level. The quality control is always a combination of different levels of control and some very site-specific tests. Although an absolute uniform tool is impossible, a set of minimum standards is essential to ensure data comparability between sites in a network and over time for long-term measurements. Nevertheless some tools for electrical, meteorological and statistical tests are available. Not only the tests but the correction of the data are necessary to produce high quality data. Very important are tests on the fulfillment of the theoretical basis of the eddy covariance method as in the non-stationarity tests and the integral turbulence characteristic test. Important is the combination of all test results in an overall quality flag for the user of the data. A proposal is given in this Chapter, but only standardization makes flux measurements comparable. This Chapter included a footprint dependent quality analysis in the CarboEurope flux program. Such analysis helps to assess the data quality of different stations. Nevertheless, the data quality is only one part of the problem. Ecological reasons make stations with a lower quality important, if the investigated ecosystem does not allow better data qualities due to hilly terrain etc. The presented quality control tools work under most of the meteorological conditions especially over low vegetation. The measurement of nighttime fluxes, when the theoretical basis of the eddy covariance method fails, is not yet included in this procedure and the complicated turbulence structure over forests needs more investigation to find adequate algorithms to check the data.

Quality control and quality assurance tests are a fundamental part of the protocols used to arrive at good estimates of turbulent fluxes and

NEE. Many of the methods have been derived through experience by an ensemble of researchers. Although there is often a good reason for site-specific procedures, most of the scientific community has similar issues to address. Hence, networks are developing prescriptive procedures to achieve a basic level of data quality. Objective methods of removing spikes and identifying appropriate turbulent conditions, instrument malfunctions, non-stationary conditions, and appropriate fetch are common to all measurement sites. Decisions regarding coordinate rotation schemes, averaging periods, energy balance closure and gap-filling are less straightforward, and need to be further investigated to arrive at standard techniques. With the wide experience being gained through international FLUXNET collaborations, consensus on all of these procedures may be reached in the near future.

6 Acknowledgment

For developing the system of quality control procedures for eddy covariance measurements several scientists supported this work, mainly Corinna Rebmann, Christoph Thomas and Bodo Wichura from the University of Bayreuth, and Dean Vickers form the Oregon State University. The authors also thank Hans Peter Schmid (Indiana University) for using his footprint model. The work was supported by the European CARBOEUROFLUX and the INTAS program and the German afo-2000, VERTIKO and DEKLIM, EVA-GRIPS program, and the AMERIFLUX program.

7 References

Amiro, B. D.: 1990, ‘Comparison of turbulence statistics within three boreal forest canopies’, *Bound.-Layer Meteorol.* **51**, 99-121.

Amiro, B. D.: 2001, ‘Paired-tower measurements of carbon and energy fluxes following disturbance in the boreal forest’, *Global Change Biology* **7**, 253-268.

Arya, S. P.: 2001, *Introduction to Micrometeorology*, Academic Press, San Diego, 415 pp.

Aubinet, M., Grelle, A., Ibrom, A., Rannik, Ü., 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., Vesala, T.: 2000, ‘Estimates of the annual net carbon and water exchange of forests: The EUROFLUX methodology’, *Adv. Ecol. Res.* **30**, 113-175.

Barr, A. G., Griffis, T. J., Black, T. A., Lee, X., Staebler, R. M., Fuentes, J. D., Chen, Z., Morgenstern, K.: 2002, ‘Comparing the carbon budgets of boreal and temperate deciduous forest stands’, *Can. J. For. Res.* **32**, 813-822.

Betts, A. K., Desjardins, R. L., MacPherson, J. I., Kelly, R. D.: 1990, ‘Boundary-layer heat and moisture budgets from FIFE’, *Bound.-Layer Meteorol.* **50**, 109-138.

Beyrich, F., Herzog, H.-J., Neisser, J.: 2002, ‘The LITFASS project of DWD and the LITFASS-98 Experiment: The project strategy and the experimental setup’, *Theor. Appl. Climat.* **73**, 3-18.

Brock, F. V.: 1986, ‘A nonlinear filter to remove impulse noise from meteorological data’, *J. Atmos. Oceanic Techn.* **3**, 51-58.

Culf, A. D., Foken, T., Gash, J. H. C.: 2003, ‘The energy balance closure problem’, in: P. Kabat et al. (Editors), *Vegetation, Water, Humans and the Climate. A New Perspective on an Interactive System*, Springer, Berlin, Heidelberg, pp. 159-166.

Curtis, P. J., Hanson, P. J., Bolstad, P., Barford, C., Randolph, J. C., Schmid, H. P., Wilson, K. B.: 2002, ‘Biometric and eddy-covariance based estimates of annual carbon storage in five eastern North American deciduous forests’, *Agric. Forest Meteorol.* **113**, 3-19.

DeGaetano, A. T.: 1997, ‘A quality-control routine for hourly wind observations”, *J. Atmosph. Oceanic Techn.* **14**, 308-317.

Desjardins, R. L., Hart, R. L., MacPherson, J. I., Schuepp, P. H., Verma, S. B.: 1992, ‘Aircraft- and tower-based fluxes of carbon dioxide, latent, and sensible heat’, *J. Geophys. Res.* **97**, 18,477-18,485.

Desjardins, R. L., MacPherson, J. I., Mahrt, L., Schuepp, P., Pattey, E., Neumann, H., Baldocchi, D., Wofsy, S., Fitzjarrald, D., McCaughey, H., Joiner, D. W.: 1997, ‘Scaling up flux measurements for the boreal forest using aircraft-tower combination’, *J. Geophys. Res.* **102D**, 29125-29133.

Essenwanger, O. M.: 1969, ‘Analytical procedures for the quality control of meteorological data’, *Proc. Amer. Meteor. Soc. Symposium on Meteorological Observations and Instrumentation, Meteorol. Monogr.* **11(33)**, 141-147.

Falge, E., et al.: 2001a, ‘Gap filling strategies for defensible annual sums of net ecosystem exchange’, *Agric. Forest Meteorol.* **107**, 43-69.

Falge, E., et al.: 2001b, ‘Gap filling strategies for long term energy flux data sets’, *Agric. Forest Meteorol.* **107**, 71-77.

Finkelstein, P. L., Sims, P. F.: 2001, ‘Sampling error in eddy correlation flux measurements’, *J. Geophys. Res.* **D106**, 3503-3509.

Finnigan, J.: 2000, ‘Turbulence in plant canopies’, *Ann. Rev. Fluid Mech.* **32**, 519-571.

Finnigan, J. J., Clement, R., Malhi, Y., Leuning, R., Cleugh, H. A.: 2003, ‘A re-evaluation of long-term flux measurement techniques, Part I: Averaging and coordinate rotation’, *Bound.-Layer Meteorol.* **107**, 1-48.

Foken, T.: 2003, *Angewandte Meteorologie, Mikrometeorologische Methoden*, Springer, Berlin, Heidelberg, 289 pp.

Foken, T., Jegede, O. O., Weisensee, U., Richter, S. H., Handorf, D., Gösdorf, U., Vogel, G., Schubert, U., Kirzel, H.-J., Thiermann, V.: 1997, ‘Results of the LINEX-96/2 Experiment”, *Deutscher Wetterdienst, Forschung und Entwicklung, Arbeitsergebnisse* **48** 75 pp.

Foken, T., Mangold, A., Rebmann, C., Wichura, B.: 2000, ‘Characterization of a complex measuring site for flux measurements’, *14th Symposium on Boundary Layer and Turbulence*, Am. Meteorol. Soc., Boston, 388-389.

Foken, T., Oncley, S. P.: 1995, ‘Results of the workshop “Instrumental and methodical problems of land surface flux measurements”’, *Bull. Am. Meteorol. Soc.* **76**, 1191-1193.

Foken, T., Skeib, G., Richter, S. H.: 1991, ‘Dependence of the integral turbulence characteristics on the stability of stratification and their use for Doppler-Sodar measurements’, *Z. Meteorol.* **41**, 311-315.

Foken, T., Wichura, B.: 1996, ‘Tools for quality assessment of surface-based flux measurements’, *Agric. Forest Meteorol.* **78**, 83-105.

Gilgen, H., Whitlock, C. H., Koch, F., Müller, G., Ohmura, A., Steiger, D., Wheeler, R.: 1994, ‘Technical plan for BSRN data management’, *WRMC, Techn. Rep.* **1**, 56 pp.

Gö *Characterization of a Complex Measuring Site for Flux Measurements*, Univ. Bayreuth., Abt. Mikrometeorologie, Arbeitsergebnisse 20, 20 pp. [Available from University of Bayreuth, Dept. of Micrometeorology, D-95440 Bayreuth]

Göckede, M., Rebmann, C., Foken, T.: 2003, ‘Use of footprint models for data quality control of eddy covariance measurements, *Agric. Forest Meteorol.* submitted.

Goulden, M. L., Munger, J. W., Fan, F.-M., Daube, B. C., Wofsy, S. C.: 1996, ‘Measurements of carbon sequestration by long-term eddy covariance: method and critical evaluation of accuracy’, *Global Change Biol.* **2**, 159-168.

Gurjanov, A. E., Zubkovskij, S. L., Fedorov, M. M.: 1984, ‘Mnogokanalnaja avtomatizirovannaja sistema obrabotki signalov na baze’, *EVM. Geod. Geophys. Veröff., R. II* **26**, 17-20.

Hall, C. D., Ashcroft, J., Wright, J. D.: 1991, ‘The use of output from a numerical model to monitor the quality of marine surface observation’, *Meteorol. Mag.* **120**, 137-149.

Handorf, D., Foken, T.: 1997, ‘Analysis of turbulent structure over an Antarctic ice shelf by means of wavelet transformation’, *12th Symosium on Boundary Layer and Turbulence*, Am. Meteorol. Soc., 245-246.

Haugen, D. A.: 1978, ‘Effects of sampling rates and averaging periods on meteorological measurements’, *Fourth Symp. on Meteorol. Observ. Instr.*, Am. Meteorol. Soc., 15-18.

Haugen, D. H. (Editor): 1973, *Workshop on Micrometeorology*, Am. Meteorol. Soc., 392 pp.

Hicks, B. B.: 1981, ‘An examination of the turbulence statistics in the surface boundary layer’, *Bound.-Layer Meteorol.* **21**, 389-402.

Højstrup, J.: 1993, ‘A statistical data screening procedure’, *Meas. Sci. Techn.* **4**, 153-157.

Howell, J. F., Sun, J.: 1999, ‘Surface layer fluxes in stable conditions’, *Bound.-Layer Meteorol.* **90**, 495-520.

Johansson, C., Smedman, A., Högström, U., Brasseur, J. G., Khanna, S.: 2001, 'Critical test of Monin-Obukhov similarity during convective conditions', *J. Atm. Sci.* **58**, 1549-1566.

Kaimal, J. C., Finnigan, J. J.: 1994, *Atmospheric Boundary Layer Flows: Their Structure and Measurement*, Oxford University Press, New York, NY, 289 pp.

Kaimal, J. C., Gaynor, J. E., Zimmerman, H. A., Zimmerman, G. A.: 1990, 'Minimizing flow distortion errors in a sonic anemometer', *Bound.-Layer Meteorol.*, **53**, 103-115.

Lenschow, D. H., Mann, J., Kristensen, L.: 1994, 'How long is long enough when measuring fluxes and other turbulence statistics?' *J. Atmosph. Oceanic Techn.* **11** 661-673.

Leuning, R., Judd, M. J.: 1996, 'The relative merits of open- and closed-path analyzers for measurements of eddy fluxes', *Global Change Biology*, **2**, 241-254.

Lorenc, A. C., Hammon, O.: 1988, 'Objective quality control of observations using Bayesian methods: theory, and a practical implementation', *Quart. J. Roy. Meteorol. Soc.* **114**, 515-544.

Lumley, J. L., Panofsky, H. A.: 1964, *The Structure of Atmospheric Turbulence*, Interscience Publishers, New York, 239 pp.

Mahrt, L.: 1991, 'Eddy asymmetry in the sheared heated boundary layer', *J. Atm. Sci.* **48**, 472-492.

Mahrt, L.: 1998, 'Flux sampling errors for aircraft and towers', *J. Atmosph. Oceanic* . **15**, 416-429.

Mahrt, L., Moore, E., Vickers, D., Jensen, N. O.: 2001a, 'Dependence of turbulent and mesoscale velocity variances on scale and stability', *J. Appl. Meteorol.* **40**, 628-641.

Mahrt, L., Vickers, D., Sun, J.: 2002, 'Spatial variations of surface moisture flux from aircraft data', *Adv. Water Res.* **24**, 1133-1142.

Mahrt, L., Vickers, D., Sun, J., Jensen, N. O., Jørgensen, H.: 2001b, 'Determination of the surface drag coefficient', *Bound.-Layer Meteorol.* **99**, 249-276.

Mann, J., Lenschow, D. H.: 1994, 'Errors in airborne flux measurements', *J. Geophys. Res.* **99**, 14519-14526.

McMillen, R. T.: 1988, 'An eddy correlation technique with extended applicability to non-simple terrain', *Bound.-Layer Meteorol.* **43**, 231-245.

Moncrieff, J. B., Massheder, J. M., DeBruin, H., Elbers, J., Friborg, T., Heusinkveld, B., Kabat, P., Scott, S., Søgaard, H., Verhoef, A.: 1997, 'A system to measure surface fluxes of momentum, sensible heat, water vapor and carbon dioxide', *J. Hydrol.* **188-189**, 589-611.

Obukhov, A. M.: 1960, 'O strukture temperaturnogo polja i polja skorostej v uslovijach konvekcii', *Izv. AN SSSR*, ser. Geofiz., 1392-1396.

Oncley, S. P., Businger, J. A., Itsweire, E. C., Friehe, C. A., LaRue, J. C., Chang, S. S.: 1990, 'Surface layer profiles and turbulence measurements over uniform land under near-neutral conditions, *9th Symp. on Boundary Layer and Turbulence*,Am. Meteorol. Soc., 237-240.

Oncley, S. P., Foken, T., Vogt, R., Bernhofer, C., Kohsiek, W., Liu, H., Pitacco, A., Grantz, D., Ribeiro, L., Weidinger, T.: 2002, 'The energy balance experiment

EBEX-2000', *15th Symp. on Boundary Layer and Turbulence*, Am. Meteorol. Soc., 1-4.

Oncley, S. P., Friehe, C. A., Larue, J. C., Businger, J. A., Itsweire, E. C., Chang, S. S.: 1996, 'Surface-layer fluxes, profiles, and turbulence measurements over uniform terrain under near-neutral conditions', *J. Atm. Sci.* **53**, 1029-1054.

Panofsky, H. A., Dutton, J. A.: 1984, *Atmospheric Turbulence - Models and Methods for Engineering Applications*, John Wiley and Sons, New York, 397 pp.

Paw U, K. T., Baldocchi, D., Meyers, T. P., Wilson, K. B.: 2000, 'Correction of eddy covariance measurements incorporating both advective effects and density fluxes', *Bound.-Layer Meteorol.* **97**, 487-511.

Raupach, M. R., Finnigan, J. J., Brunet, Y.: 1996, 'Coherent eddies and turbulence in vegetation canopies: the mixing-layer analogy', *Bound.-Layer Meteorol.* **78**, 351-382.

Rummel, U., Ammann, C., Meixner, F. X.: 2002, 'Characterizing turbulent trace gas exchange above a dense tropical rain forest using wavelet and surface renewal analysis', *15th Symp. on Boundary Layer and Turbulence*, Am. Meteorol. Soc., 602-605.

Sakai, R., Fitzjarrald, D., Moore, K. E.: 2001, 'Importance of low-frequency contributions to eddy fluxes observed over rough surfaces', *J. Appl. Meteorol.* **40**, 2178-2192.

Schmid, H. P.: 1997, 'Experimental design for flux measurements: matching scales of observations and fluxes', *Agric. Forest Meteorol.* **87**, 179-200.

Shearman, R. J.: 1992, 'Quality assurance in the observation area of the Meteorological Office', *Meteorol. Mag.* **121**, 212-216.

Smith, S. R., Camp, J. P., Legler, D. M.: 1996, *Handbook of Quality Control, Procedures and Methods for Surface Meteorology Data*, Center for Ocean Atmospheric Prediction Studies, TOGA/COARE, Technical Report, 96-3: 60 pp. [Available from Florida State University, Tallahassee, FL, 32306-3041.]

Snyder, R. L., Spano, D., Paw U, K. T.: 1996, 'Surface renewal analysis for sensible and latent heat flux density', *Bound.-Layer Meteorol.* **77**, 249-266.

Stull, R. B.: 1988, *An Introduction to Boundary Layer Meteorology*, Kluwer Acad. Publ., Dordrecht, Boston, London, 666 pp.

Sun, J., Howell, J. F., Esbensen, S. K., Mahrt, L., Greb, C. M., Grossman, R., LeMone, M. A.: 1996, 'Scale dependence of air-sea fluxes over the Western Equatorial Pacific', *J. Atm. Sci.* **53**, 2997-3012.

Thomas, C., Foken, T.: 2002, 'Re-evaluation of integral turbulence characteristics and their parameterizations', *15th Conference on Turbulence and Boundary Layers*, Am. Meteorol. Soc., 129-132.

Trevino, G., Andreas, E. L.: 2000, 'Spectral analysis of nonstationary turbulence', *Bound.-Layer Meteorol.* **95**, 231-247.

Vickers, D., Mahrt, L.: 1997, 'Quality control and flux sampling problems for tower and aircraft data', *J. Atmosph. Oceanic Techn.* **14**, 512-526.

Vickers, D., Mahrt, L.: 2003, 'The cospectral gap and turbulent flux calculations', *J. Atmosph. Oceanic Techn.* in press.

Wichura, B., Buchmann, N., Foken, T.: 2002, 'Carbon dioxide exchange characteristics above a spruce forest, *25th Symp. Agric. Forest Meteor.*, Am. Meteorol. Soc., 63-64.

Wilczak, J. M., Oncley, S. P., Stage, S. A.: 2001, 'Sonic anemometer tilt correction algorithms', *Bound.-Layer Meteorol.* **99**, 127-150.

Wyngaard, J. C., Cot, O. R., Izumi, Y.: 1971, 'Local free convection, similarity and the budgets of shear stress and heat flux', *J. Atm. Sci.* **28**, 1171-1182.