THE ENERGY BALANCE CLOSURE PROBLEM: AN OVERVIEW

THOMAS FOKEN 1

Department of Micrometeorology, University of Bayreuth, D-95440 Bayreuth, Germany

Abstract. This paper gives an overview of 20 years of research on the energy balance closure problem. It will be shown that former assumptions that measuring errors or storage terms are the reason for the unclosed energy balance do not stand up because even turbulent fluxes derived from documented methods and calibrated sensors, net radiation, and ground heat fluxes cannot close the energy balance. Instead, exchange processes on larger scales of the heterogeneous landscape have a significant influence. By including these fluxes, the energy balance can be approximately closed. Therefore, the problem is a scale problem and has important consequences to the measurement and modeling of turbulent fluxes.

Key words: Bowen ratio; carbon dioxide flux; energy balance closure; energy storage; latent heat flux; net radiation; scalar similarity; sensible heat flux; soil heat flux; turbulent flux.

INTRODUCTION

During the late 1980s, it became obvious that the energy balance at the earth’s surface could not be closed with experimental data (Foken and Oncley 1995). The available energy, i.e., the sum of the net radiation and the ground heat flux, was found in most cases to be larger than the sum of the turbulent fluxes of sensible and latent heat. This was a main topic of a workshop held in 1994 in Grenoble (Foken and Oncley 1995). In most of the land surface experiments (Leuning et al. 1982, Tsvang et al. 1991, Kanemasu et al. 1992, Bolle et al. 1993), and also in the carbon dioxide flux networks (Aubinet et al. 2000, Wilson et al. 2002), a closure of the energy balance of approximately 80% was found. The residual is

\[ \text{Res} = Q^S - Q_G - Q_H - Q_E \]

where \( Q^S \) is net radiation, \( Q_G \) is soil heat flux, \( Q_H \) is sensible heat flux, and \( Q_E \) is latent heat flux. The problem cannot be described only as an effect of statistically distributed measuring errors because of the clear underestimation of turbulent fluxes or overestimation of the available energy. In the literature, several reasons for this incongruity have been discussed, most recently in an overview paper by Culf et al. (2004).

An experiment designed to investigate this problem, the EBEX-2000, took place in the summer of 2000 near Fresno, California, USA. The EBEX-2000 results confirming these findings are now available (Onclety et al. 2007). Furthermore, in recent papers it was found that time-averaged fluxes (Finnigan et al. 2003) or spatially averaged fluxes including turbulent organized structures (Kanda et al. 2004) can close the energy balance.

The aim of the following paper is to summarize the given problems and all available findings. The explanations for the energy balance will be recapitulated. Furthermore, the consequences to near-surface modeling are discussed.

THE ENERGY BALANCE CLOSURE PROBLEM

As stated above, the available energy \( (Q^S + Q_G) \) was found in most of the experiments to be larger than the turbulent fluxes of sensible and latent heat. In the past, two reasons for this energy abundance were mainly discussed: the nonidentical balance layers of these measurements and possible measuring errors. Typical errors and scales for the measurements are given in Table 1 and Fig. 1 illustrates the measurement conditions.

While the measuring height of the net radiation is approximately 2 m and has only a small influence on the upwelling radiation components, the measuring height of the turbulent fluxes, usually 2–5 m, has a significant influence on the footprint (Schmid 1997) and the size of

<table>
<thead>
<tr>
<th>Component</th>
<th>Error (%)</th>
<th>Energy (W/m²)</th>
<th>Horizontal scale (m)</th>
<th>Height (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latent heat flux</td>
<td>5–20</td>
<td>20–50</td>
<td>100</td>
<td>2–10</td>
</tr>
<tr>
<td>Sensible heat flux</td>
<td>5–20</td>
<td>10–30</td>
<td>100</td>
<td>2–10</td>
</tr>
<tr>
<td>Net radiation</td>
<td>5–20</td>
<td>20–100</td>
<td>10</td>
<td>1–2</td>
</tr>
<tr>
<td>Ground heat flux without storage</td>
<td>20–50</td>
<td>20–50</td>
<td>0.1</td>
<td>−0.02 to −0.1</td>
</tr>
<tr>
<td>Storage term</td>
<td>20–50</td>
<td>20–50</td>
<td>0.1–1</td>
<td>−0.02 to −0.1</td>
</tr>
</tbody>
</table>
the underlying surface. This depends, furthermore, on the stability. The ground heat flux is measured some centimeters below the surface. The horizontal influences on the measurements are not much larger than the size of the heat flux plate, but the heat storage in the layer between the surface and the plate and the heterogeneity of the soil can have a significant influence on the results (Liebethal et al. 2005).

The typical residual of the energy balance closure in daytime is found to be 50–300 W/m$^2$. This can be easily explained by the errors given in Table 1, but according to recent studies (see The findings), these errors can be reduced. Because the closure problem is always characterized by low turbulent fluxes, either the sensible and latent heat flux is underestimated or the net radiation and the ground heat flux are overestimated. Because of the complicated data analysis of the eddy covariance method that was used to determine the turbulent fluxes, an underestimation of the turbulent fluxes was often argued as the reason for the problem. In some papers, even the Bowen ratio method was favored (Doran et al. 1989, Fritschen et al. 1992, Brotzge and Crawford 2003), but in these methods the energy balance is closed by definition (see The findings: Experimental results). On the other hand, no arguments could be found to prove that the net radiation was overestimated. It is known that the net radiation was underestimated in the past due to less accurate sensors. The energy balance closure problem was found about 15 years ago when more precise net radiometers were available (Halldin and Lindroth 1992). Additionally, the soil heat flux is often underestimated because of missing or insufficient calculation of the storage term (Liebethal and Foken 2007).

The findings

In the following section, the relevant findings on the energy balance closure problem are summarized, mainly on the basis of recent investigations. This is done more or less as an enumeration, while in the following chapter these findings will be brought together to find a possible solution.

Experimental results

Low vegetation.—Many investigations above low vegetation are available in the literature (many old investigations are reviewed in Laubach and Teichmann [1996]). Here, the discussion is concentrated on a selected set of experiments (Table 2), because these were done by a group of scientists with similar devices and calculation procedures and compared to each other. The last four data sets were measured on the boundary-layer field site of the German Meteorological Service (Meteorological Observatory Lindenberg), for which a recent study of the LITFASS-2003 experiment (Beyrich and Mengelkamp 2006, Mengelkamp et al. 2006) underlined these findings with an averaged residual of 20–30% for different agricultural fields (Mauder et al. 2006).

These data sets, except for the two LITFASS experiments in 1998 and 2003, were investigated by Panin et al. (1998) and Foken (1998a), respectively, in different ways. Panin et al. (1998) found a correlated increase of the residual with an increase of the heterogeneity of the underlying surface in the vicinity of the measuring place with less heterogeneities during FIFE-89, moderate at the boundary-layer field site of the German Weather Service and high during KUREX and TARTEX. They introduced a heterogeneity factor $k$ to correct the energy balance closure,

$$Q_S = Q_G = k(Q_H + Q_E) \quad (2)$$

while the reasons could be advection or larger turbu-

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Measurement height and horizontal scale of the measurement of the energy balance components (Foken 2008). The bar on the right of the figure is a tower; the cone with a black top is a radiation sensor showing the radiation footprint; arrows show the direction of flux. $Q_S$ is net radiation, $Q_G$ is soil heat flux, $Q_H$ is sensible heat flux, $Q_E$ is latent heat flux, and $\Delta Q_S$ is heat storage.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Reference</th>
<th>Residual (%)</th>
<th>Surface</th>
</tr>
</thead>
<tbody>
<tr>
<td>Müncheberg 1983 and 1984</td>
<td>Koitzsch et al. (1988)</td>
<td>14</td>
<td>winter wheat</td>
</tr>
<tr>
<td>KUREX-88</td>
<td>Tsvang et al. (1991)</td>
<td>23</td>
<td>different agricultural fields</td>
</tr>
<tr>
<td>FIFE-89</td>
<td>Kanemasu et al. (1992)</td>
<td>10</td>
<td>steppe</td>
</tr>
<tr>
<td>TARTEX-90</td>
<td>Foken et al. (1993)</td>
<td>33</td>
<td>barley and bare soil</td>
</tr>
<tr>
<td>KUREX-91</td>
<td>Panin et al. (1998)</td>
<td>33</td>
<td>different agricultural fields</td>
</tr>
<tr>
<td>LINEX-96/2</td>
<td>Foken et al. (1997)</td>
<td>20</td>
<td>high grass</td>
</tr>
<tr>
<td>LINEX-97/1</td>
<td>Foken (1998b)</td>
<td>32</td>
<td>short grass</td>
</tr>
<tr>
<td>LITFASS-98</td>
<td>Beyrich et al. (2002b)</td>
<td>37</td>
<td>bare soil</td>
</tr>
<tr>
<td>LITFASS-2003</td>
<td>Mauder et al. (2006)</td>
<td>20–30</td>
<td>different agricultural sites</td>
</tr>
</tbody>
</table>
fluence structures. On the other hand, Foken (1998a) argued on the basis of investigations by Kukharets et al. (1998) that the residual is smaller for a plant-covered surface (FIFE, LINEX-96/2) and lower for sites with a higher exposure of the soil (most of the other experiments). The energy storage in the upper soil layer and its correct determination was seen as a main reason for the problem. Therefore, they rewrite Eq. 1 in the form

\[
\text{Res} = \frac{Q^\pm}{C_0} - \frac{Q_G}{C_0} - \frac{Q_H}{C_0} - \frac{Q_E}{C_0} \pm \Delta Q
\]  

with the storage term \(\Delta Q\). This was also one of the main reasons given in the overview paper by Culf et al. (2004).

Recent investigations have shown that an accurate determination of this term is possible (see Storage terms).

Tall vegetation.—The energy balance closure for tall vegetation was mainly investigated by Aubinet et al. (2000) and Wilson et al. (2002) in addition to other studies like Oliphant et al. (2004). Both found energy balance closures for most of the sites of about 80% of the available energy. The scatter plot for Aubinet’s investigation is shown in Fig. 2.

Recently, for most of the European FLUXNET sites (Baldocchi et al. 2001), a detailed footprint analysis was conducted (Rebmann et al. 2005) based on the method described by Góckede et al. (2004). This method was updated with a Langrangian footprint model (Góckede et al. 2006) and again applied on the measuring sites (Góckede et al. 2008). The stations were classified using a footprint threshold, which was defined as 80% of the flux is coming from the specific target area of the station. In the case of the highest class, 1, more than 90% of the data are within this footprint threshold and for the next class, 2, 60–90% are in the threshold. The contribution of fluxes not coming from the target area is a measure of the heterogeneity of the landscape due to other land use classes. Comparing the residual and the footprint class (Table 3), it was found that, for the footprint class 1, the energy balance closure is better than 90%. Therefore, the phenomena should be related to the heterogeneity of a larger landscape. Because of the different instrumentation of the stations this comparison can only be presented for orientation.

Measuring errors

Measuring errors were mainly discussed as the reason for the residual of the energy balance closure. Because of the nonstatistically distributed residual values, a significant underestimation of the turbulent fluxes with the eddy covariance method was assumed, as well as an occasional overestimation of the net radiation.
Turbulent fluxes.—Turbulent fluxes are measured with the eddy covariance method, first described more than 50 years ago (Montgomery 1948, Obukhov 1951, Swinbank 1951), which is now widely used (Moncrieff et al. 1997, Aubinet et al. 2000, Lee et al. 2004). Because the method is a stochastic one, typical errors like sampling errors, which are negligible for typical sampling frequencies of 20 Hz and a typical measuring length of 30 minutes (Haugen 1978), and random errors (Lenschow et al. 1994) occur. In more recent studies, random errors based on the comparison of measurements made with at least two systems installed within a short distance were investigated (Finkelstein and Sims 2001, Richardson et al. 2006), but these errors are not of a size to solve the closure problem. These investigations are closely connected to data quality analysis (Foken and Wichura 1996) and the comparison of sensors; a study including all these aspects is still missing.

Comparisons of sonic anemometers and fast response sensors for scalars are now available for most of the recently used sensor types; unfortunately, these comparisons have been mainly published in the grey literature and only a few in reviewed papers (Loescher et al. 2005, Mauder et al. 2006, 2007b). Generally it was found that the accuracy of the sensors, or better, that the agreement of the results is very good, but dependent on the type of the sonic anemometer and the data quality. Foken and Oncley (1995) already classified the anemometers into a more scientific type (e.g., Campbell CSAT3, Solent HS; see Plate 1) and a type for more general use (omnidirectional types). Using the data quality tool for the eddy covariance method (Foken and Wichura 1996, Foken et al. 2004), the data quality was also found to influence the results of the sensor comparison. Mauder et al. (2006) combined both aspects of the intercomparison and found (Table 4) that in most of the cases the fluxes can be measured with an acceptable accuracy, which cannot explain the residual of the energy balance closure.

The eddy covariance method needs several corrections (Aubinet et al. 2000, Lee et al. 2004); most of them increase the turbulent flux. Such corrections or transformations are: the determination of the time delay of all additional sensors through the calculation of cross correlations; cross wind correction of the sonic temperature according to Liu et al. (2001), if not already implemented in sensor software; the “planar fit” method for coordinate transformation (Wilczak et al. 2001); a correction of oxygen cross sensitivity of krypton-hygrometers (Tanner et al. 1993, van Dijk et al. 2003); spectral corrections according to Moore (1986) using the spectral models by Kaimal et al. (1972) and Højstrup (1981); conversion of fluctuations of the sonic temperature into fluctuations of the actual temperature according to Schotanus et al. (1983); density correction of scalar fluxes of H2O and CO2 according to Webb et al. (1980); iteration of the correction steps because of their interacting dependence and data quality analysis according to Foken and Wichura (1996) and Vickers and Mahrt (1997) in the updated version by Foken et al. (2004).

The most significant changes of the heat fluxes are the transformation from the buoyancy flux into the “exact” sensible heat flux (Schotanus et al. 1983), as well as corrections for both spectral losses (Moore 1986, Eugster and Senn 1995) and the effect of density fluctuations on the latent heat flux (Webb et al. 1980). The effects of all these steps are shown in Fig. 3. Overall, a careful data correction reduces the residual of the

<table>
<thead>
<tr>
<th>Station</th>
<th>Surface</th>
<th>Closure (%)</th>
<th>Footprint</th>
</tr>
</thead>
<tbody>
<tr>
<td>BE1 Vielsalm</td>
<td>mixed <em>Fagus sylvatica</em> and <em>Pseudotsuga menziesii</em></td>
<td>93</td>
<td>class 1</td>
</tr>
<tr>
<td>GE2 Tharandt</td>
<td><em>Picea abies</em></td>
<td>92</td>
<td>class 1</td>
</tr>
<tr>
<td>FR2 Les Landes</td>
<td><em>Pinus pinaster</em></td>
<td>89</td>
<td>class 2</td>
</tr>
<tr>
<td>GE1 Waldstein-Weidenbrunnen</td>
<td><em>Picea abies</em></td>
<td>77</td>
<td>class 2</td>
</tr>
<tr>
<td>FR1 Hesse</td>
<td><em>Fagus sylvatica</em></td>
<td>71</td>
<td>class 2</td>
</tr>
</tbody>
</table>

Table 4. Accuracy of turbulent fluxes measured with the eddy covariance method (Mauder et al. 2006).

<table>
<thead>
<tr>
<th>Anemometer</th>
<th>Quality class (Foken et al. 2004)</th>
<th>Sensible heat flux</th>
<th>Latent heat flux</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type A, e.g., CSAT3</td>
<td>high</td>
<td>5% or 10 W/m²</td>
<td>10% or 20 W/m²</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>10% or 20 W/m²</td>
<td>15% or 30 W/m²</td>
</tr>
<tr>
<td>Type B, e.g., R3</td>
<td>high</td>
<td>10% or 20 W/m²</td>
<td>15% or 30 W/m²</td>
</tr>
<tr>
<td></td>
<td>medium</td>
<td>15% or 30 W/m²</td>
<td>20% or 40 W/m²</td>
</tr>
</tbody>
</table>

Note: The classes “high” and “medium” correspond to the classes 1–3 and 4–6, respectively, according to Foken et al. (2004).
energy balance closure by about 10–20%, but flux corrections cannot explain the magnitude of the residual and nowadays the eddy covariance method and the correction steps can be classified as well established (Moncrieff 2004).

There are several papers that report that flow distortion problems and the so-called angle of attack have a significant influence on the energy balance closure (Gash and Dolman 2003, van der Molen et al. 2004, Nakai et al. 2006), with larger effects on forest sites than on low vegetation sites. The flow distortion error results from the imperfect (co)sine response of the sonic anemometer. But according to recent findings, the closure is often better on forest sites than on low vegetation sites, because the forest is often decoupled from the atmosphere (Thomas and Foken 2007), storage terms have no significant influence and forests are more homogeneous than agricultural fields. Because of the nearly laminar flow in the wind tunnel, the transformation to the turbulent field is difficult and the eddies seem to be less affected by flow distortion problems than assumed from wind tunnel measurements (Högström and Smedman 2004). Therefore, the angle of attack may have an influence but is less significant to the energy balance closure problem.

Net radiation measurements.—That net radiation sensors could be a reason for the unclosed energy balance was evident after the comparison study by Halldin and Lindroth (1992). But the errors tend more to an underestimation (the residual was not able to be found) than to an overestimation, which is necessary to explain the energy balance closure problem. In the last 15 years, much has been done to increase the accuracy of the radiation measurements (Ohmura et al. 1998, Halldin 2004), mainly due to the activities of the Basic Surface Radiation Network (BSRN) of the World

![Graph showing energy balance fluxes](image)

**TABLE 5.** Increase of the accuracy of the radiation sensors due to the BSRN activities (Ohmura et al. 1998).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Sensor</th>
<th>Accuracy 1990 (W/m²)</th>
<th>Accuracy 1995 (W/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global radiation</td>
<td>pyranometer</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>Solar radiation</td>
<td>actinometer, sun photometer</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Diffuse radiation</td>
<td>shaded pyranometer</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Down-welling longwave radiation</td>
<td>pyrgeometer</td>
<td>30</td>
<td>10</td>
</tr>
</tbody>
</table>
Climate Research Program. In Table 5, the significant increase in the accuracy is documented. For the experiment EBEX-2000, all net radiometers were compared (Kohsiek et al. 2007) and their high accuracy was confirmed (Table 6). It cannot be assumed that the accuracy of the shortwave components better than 2% and of the long wave components better than 5 W/m² have a significant influence on the energy balance closure problem. The shortwave sensors have a very high accuracy, which is classified by the World Meteorological Organization as “secondary standard” (Brock and Richardson 2001).

Storage terms

As stated above, the energy storage in the upper soil layer above the heat flux plate may have a significant influence on the residual, if this energy cannot be exactly determined. But there are also other storage terms in the air below the flux measurements and in the plants. In addition, the necessary energy for photosynthesis can be discussed as a storage problem. The storage may be a reason for asymmetric residuals in relation to the daily or annual cycle. But such results were not always found (Oliphant et al. 2004).

Two recent studies confirm this: Meyers and Hollinger (2004) analyzed the storage in the soil as well as the storage in the canopy and in photosynthetic products, while Heusinkveld et al. (2004) exclusively concentrated on the soil heat storage. Both studies agree that including storage terms is very important for correct ground heat flux determination at the surface and for good energy balance closure.

According to a sensitivity analysis by Liebethal et al. (2005), the most reliable method to determine the ground heat flux for the data of the LITFASS-2003 experiment recorded over a maize field turned out to be a combination of two methods. The gradient approach was applied at a depth of 0.20 m; the change in the heat storage in the soil layer above this reference level was added to it:

$$ Q_G = -\lambda_s(z_r) \frac{\partial T_s}{\partial z} \bigg|_{z_r} + \int_{0}^{z_r} c_v \frac{\partial T_s}{\partial t} dz $$

where $z_r$ is the reference depth, $T_s$ is the soil temperature, $t$ is time, $\lambda_s$ is the thermal conductivity of the soil according to Fourier’s law of heat conduction, and $c_v$ is the volumetric soil heat capacity (calculated from the volumetric fractions of soil constituents according to de Vries [1963], where organic compounds are neglected).

Even with this high-quality data set for the ground heat flux, the energy balance can only be closed within the error margins of flux determination during nighttime (Fig. 4). During daytime, a considerable residual of several tens to over 100 W/m² still exists (Liebethal and Foken 2007). If the ground heat flux is not well determined, it can be a significant influence on the energy balance closure problem.

For the irrigated cotton field of the EBEX-2000 experiment, all the energy terms were investigated (Oncley et al. 2007). As shown in Fig. 5, the storage in the air can be negligible and also the storage in the biomass is very small. The largest storage term is the energy storage in the ground. The photosynthesis was found to be 3.8 W/m² with a maximal value of 12 W/m². Similar studies of storage terms are also available from other authors, e.g., McCaughey and Saxton (1988), Mayocchi and Bristow (1995), and Oliphant et al. (2004).

Influences of long-wave eddies

An energy transport with large eddies which cannot be measured with the eddy covariance method is assumed as one of the main reasons of the closure problem. The heterogeneity of the landscape was seen as a reason for such eddies. In the literature, several methods are discussed to investigate this problem (Sakai et al. 2001, Finnigan et al. 2003, Foken et al. 2006b, Sun et al. 2006), which are discussed briefly.

The ogive function.—About 15 years ago, the ogive function was introduced into the investigation of turbulent fluxes (Desjardins et al. 1989, Oncley et al. 1990, Friehe 1991). This function was proposed as a test to check if all low frequency parts are included in the turbulent flux measured with the eddy covariance method (Foken et al. 1995, 2004). The ogive is the cumulative integral of the co-spectrum starting with the highest frequencies:

<table>
<thead>
<tr>
<th>Type</th>
<th>Sensor</th>
<th>Accuracy (%)</th>
<th>Accuracy (Wm²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shortwave radiation</td>
<td>Eppley PSP</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kipp&amp;Zonen CM11, CM 21</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Longwave radiation</td>
<td>Eppley PIR</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Net radiation</td>
<td>Kipp&amp;Zonen CNR1</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>REBS Q*7</td>
<td>20†</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Schulze-Däke</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

† For Q*7, a large scatter between the sensors was found.
\[ \text{og}_{w,x} (f_0) = \int_{f_0}^T \text{Co}_{w,x} (f) \text{df} \tag{5} \]

where \( \text{Co}_{w,x} \) is the co-spectrum of a turbulent flux, \( w \) is the vertical wind component, \( x \) is the horizontal wind component or scalar, and \( f \) is frequency. In this study, co-spectra for all relevant combinations of time series were calculated over integration times of up to four hours. For the LITFASS-2003 experiment, only frequency values higher than approximately \( 1.39 \times 10^{-4} \) Hz that correspond to periods of two hours and shorter were used for the test (Foken et al. 2006); an underlying interval of four hours improves the statistical significance. Longer periods were not investigated due to nonstationary conditions related to the diurnal cycle of the fluxes and high non-steady-state conditions. Because the ogive test must be applied to time series without any...
gaps, only 121 series for the whole experiment were available. The convergence of the ogive was analyzed as follows.

In the ideal convergent case, the ogive function increases during the integration from high frequencies to low frequencies until a certain value is reached and remains on a more or less constant plateau before a 30-minute integration time. If this condition is fulfilled, the 30-minute covariance is a reliable estimate for the turbulent flux, because we can assume that the whole turbulent spectrum is covered within that interval and that there are only negligible flux contributions from longer wavelengths (Case 1). But it can also occur that the ogive function shows an extreme value and decreases again afterwards (Case 2) or that the ogive function doesn’t show a plateau but increases throughout (Case 3). Ogive functions corresponding to Case 2 or 3 indicate that a 30-minute flux estimate is possibly inadequate.

An overview of the number of measuring series consisting of these cases is given in Table 7. It can be concluded that a 30-minute averaging interval appears to be sufficient to cover all relevant flux contributions in roughly five out of six cases (85%). For the remaining cases, the eddy covariance method does not measure the total flux within the 30-minute interval. The 30-minute flux may be reduced because the flux in one direction was already reached in a shorter time period (Case 2) and an integration of up to 30 minutes reduces the fluxes due to non-steady state conditions or long-wave trends, or because significant flux contribution can be found for integration periods longer than 30 minutes (Case 3). A simplified correction of the turbulent fluxes by the ratio of the ogive function for 30 minutes and the maximum ogive function (extreme or convergence) shows a reduced residual by 5–10%.

Increase of the averaging period.—Finnigan et al. (2003) proposed a site-specific extension of the averaging time of up to several hours to close the energy balance. This was also done for the LITFASS-2003 experiment (Fig. 6) and underlines the finding that, in the first hours, the effect is small. If the averaging over longer time periods is from the statistical point of view acceptable, the energy balance can be closed over heat flux. More investigations about steady state conditions and the interpretation of the data are necessary to apply this method. But it can be seen from the results that probably 24 hours are responsible for the closure problem for this data set, mainly due to an increase of the sensible larger turbulent structures.

Advection.—Advection is also discussed as a reason for the energy balance closure. Up to now, advection has been mostly investigated in connection with the carbon dioxide advection in tall vegetation (Aubinet et al. 2003, 2005, Staebler and Fitzjarrald 2004). The accurate determination needs an optimal choice of the coordinate system and an expanded experimental setup, because the net advection is often a small difference of the horizontal and vertical advection in a sloping terrain. These studies investigated mainly katabatic flows. For the EBEX-2000 experiment (Onclley et al. 2007) the setup of several profile towers was used to determine the horizontal advection, which was found to be up to 30 W/m² and was discussed as one of the reasons of the closure problem. The simple measurement of the divergence or convergence term due to advection in flat terrain with an acceptable accuracy is still an outstanding problem.

Area-averaging flux measurements.—If larger eddies have a remarkable contribution to the energy exchange, these eddies cannot only be detected with time-averaging of a measuring series, but also with area-averaging measuring systems like aircraft-based turbulence systems. Stationary area-averaging measurement systems are large aperture scintillometer (LAS) for the sensible heat flux (Beyrich et al. 2002) and a microwave scintillometer (MWS) for the latent heat flux (Meijninger et al. 2006). Such systems were used during the LITFASS-2003 experiment with a path length of approximately 5 km.

The combination of a (near-infrared) LAS and a (94 GHz) microwave scintillometer (known as the two-wavelength method) make it possible to measure the

### Table 7. Number and percentage of convergent ogives (Case 1), ogives with an extreme value (Case 2), and non-convergent ogives (Case 3) of momentum (og\(_m\)), sensible heat (og\(_h\)), and latent heat (og\(_l\)) flux.

<table>
<thead>
<tr>
<th>Type</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>og(_m)</td>
<td>103 (85%)</td>
<td>13 (11%)</td>
<td>5 (4%)</td>
</tr>
<tr>
<td>og(_h)</td>
<td>100 (83%)</td>
<td>14 (12%)</td>
<td>7 (6%)</td>
</tr>
<tr>
<td>og(_l)</td>
<td>100 (83%)</td>
<td>17 (14%)</td>
<td>4 (3%)</td>
</tr>
</tbody>
</table>

Note: The numbers in parentheses are the percentages of the data set of 121 series for the whole period (Foken et al. 2006b).
Fluxes of sensible heat and latent heat flux directly at scales of several kilometers (Meijninger et al. 2002b, 2006). Applying Obukhov’s similarity relations (Obukhov 1960), the surface fluxes can be derived from the path-averaged structure parameter data. A footprint analysis of the set-up performed by Meijninger et al. (2006) showed that more than 85% of the source area of the scintillometer represents farmland (for all wind directions) and can be easily compared with a composite of 11 flux towers over agricultural and grassland sites. The fluxes measured at these stations were combined to a so-called flux composite, taking into account the data quality of the individual measurements and the relative occurrence frequency of the different types of low vegetation (Beyrich et al. 2006) in relation to the typical footprint area. For 25 May 2003, the area-averaged fluxes were compared with the composite of the surface layer fluxes (Fig. 7). The sensible and latent heat fluxes estimated with the scintillometer are approximately 20–50 W/m² larger than the eddy covariance data and can nearly compensate the residual with a maximum of approximately 100 W/m².

**Summary and Hypothesis**

Summarizing all these findings, the typical reasons for the energy balance closure discussed in the past can nowadays be excluded. The data quality and accuracy of the measurements of the net radiation and the turbulent fluxes has increased significantly in the last ten years and cannot be an argument for the energy balance closure. Additionally, the ground heat flux, including the storage term in the upper soil layer, can be determined with high accuracy.

But results such as a better closure for an extension of the averaging time or for area-averaged fluxes give a hint that larger turbulence structures may have a significant influence on the energy balance closure. Such turbulence structures must be in relation to the structures of the underlying surface. Therefore, Mauder et al. (2007a) compared satellite pictures of four energy balance studies (Fig. 8). The landscapes of the EBEX-2000 and

![Figure 7](image_url)
the LITFASS-2003 experiments with typical residual of 10–15% and 25–35%, respectively, are very heterogeneous. In contrast, the energy balance closure for a desert (Unland et al. 1996, Heusinkveld et al. 2004) or the African bush land (Mauder et al. 2007a) is nearly ideal.

If larger eddies have a significant contribution to the energy exchange, they must be generated at boundaries between different land uses that are excluded from flux measurements due to their influence on the footprint and on the generation of internal boundary layers. From some selected experiments (e.g., Klaassen et al. 2002), it is known that the turbulent fluxes increase near the forest edge, also found with parallel modeling studies (Klaassen and Sogatchev 2006). Similar results were also found by model studies in an artificial heterogeneous landscape (Friedrich et al. 2000).

Kanda et al. (2004) found with Large Eddy Simulation (LES) studies that turbulent organized structures have a contribution to the energy exchange which can close the energy balance. Similar results were found for secondary circulations (Inagaki et al. 2006, Steinfeld et al. 2007). For the LITFASS-2003 experiment, the parallelized LES model (PALM; Raasch and Schröter

![Fig. 8](image-url)  The landscape of a $20 \times 20$ km$^2$ area around the measuring points of different experiments, according to Mauder et al. (2007a): (a) EBEX-2000, California (Oncley et al. 2007), residual 10–15%; (b) LITFASS-2003, Germany (Beyrich and Mengelkamp 2006), residual 25–35%; (c) NIMEX-1, Nigeria (Mauder et al. 2007a), residual nearly 0%; (d) Negev Desert, Israel (Heusinkveld et al. 2004), residual 0%.
2001) was applied to investigate the fluxes at approximately 40 m height for 30 May 2003; unfortunately, it was a day where the microwave scintillator did not work. As a result of the model, Fig. 9 was generated, which shows the secondary circulations for 30 May 2003, a day with weak mean horizontal wind. The secondary circulation structures were found to be very stable in relation to the underlying surface. Along the investigated path, 240 virtual towers of 40 m height were built up with the LES model. The data of the LES simulation of these towers with a sampling frequency of 2 Hz were used for an eddy covariance calculation in two ways: determination of the fluxes of all towers and averaging of these fluxes and spatial calculation of the fluxes (similar to an aircraft flight along the towers). It was assumed that these towers were within the constant...
flux layer. The results are given in Fig. 10. The spatially calculated flux is approximately 20 W/m² larger than the averaging of the fluxes of the towers but significantly larger than the measured fluxes of the flux stations, and partially larger than those of the scintillometer as well. Foken et al. (2006a) used these findings from the LITFASS-2003 experiment to attempt to close the energy balance. Unfortunately, the LES modeling and the scintillometer measurements are only available for very short periods. But other findings (Finnigan et al. 2003, Kanda et al. 2004) seem to support their assumption that not only small eddies, which can be measured by the eddy covariance method, have a contribution to the energy balance, but also larger eddies in the lower boundary layer, which do not touch the surface or are steady state. These can only be modeled or measured with area-averaging methods. The long-term averaging of eddy covariance data may be also a possibility, because these frequently steady-state eddies move in the transition times in the morning and afternoon. Therefore, this long-term averaging has only a significant effect for very long averaging times and the typical ogive test is unable to close the energy balance.

These findings lead to the hypothesis that the turbulent fluxes have a contribution of smaller eddies, which can be measured with the eddy covariance method, and a contributions of larger eddies, which are related to the heterogeneous structure of the landscape. These structures are quite large and not identical with those smaller ones assumed by Panin et al. (1998). Obviously there is a spectral gap between both parts, probably comparable with the so-called mesoscale minimum. Therefore, different scales must be taken into account. The energy balance equation can be written in the following form:

\[ 0 = Q_s^H - Q_G - (Q_H)_s - (Q_H)_l - (Q_E)_s - (Q_E)_l \]  

(6)

with the sensible and latent heat fluxes for smaller (s) and larger (l) eddies. It is assumed that the net radiation does not differ for smaller and larger scales on average and the ground heat flux is also assumed as identical. Near the surface, the smaller eddies \( (Q_{H,E})_s \) are measured with micrometeorological methods and the long-wave part \( (Q_{H,E})_l \) is either not available or part of the advection, as probably measured by Oneley et al. (2007). Such a possible schema is illustrated in Fig. 11.

When only the fluxes \( (Q_{H,E})_s \) can be measured with the eddy covariance technique and measuring methods or LES-simulations for \( (Q_{H,E})_l \) are not available for typical experiments, the question remains as to how to determine the fluxes of large eddies. As a first guess we can apply the simplified correction of the energy balance closure, which some authors have already used in the past (Lee 1998, Twine et al. 2000) to distribute the residual according to the Bowen ratio to the sensible and latent heat flux.

Under the conditions defined in Eq. 6, this means that the Bowen ratios for small- and large-scale eddies are similar or even equal. Such a similarity was found for scalar concentrations, \( c \) (Gao 1995, Pearson et al. 1998, Katul and Hsieh 1999), and can be defined with the correlation coefficient between a scalar and a proxy scalar:

\[ r_{c,proxy} = \frac{c \sigma_{proxy}}{\sigma_c \sigma_{proxy}} \]  

(7)

FIG. 11. Schematic figure of the generation of secondary circulations and the hypothesis of turbulent fluxes in different scales based on small eddies (s) and large eddies (l). See Eq. 6.
with the covariance in the numerator and the standard deviations in the denominator. For smaller eddies, this similarity is given (Pearson et al. 1998), but not for the long-wave part of the turbulence spectra (Ruppert et al. 2006). This similarity is even different for different scalars and different times of the day (Fig. 12). Therefore, the accurate distribution of the residual to the large-eddy part of the sensible and latent heat flux is still an outstanding problem.

**The Consequences**

The modified form of the energy balance equation (Eq. 6) has dramatic consequences to all measuring and modeling techniques which use this equation to determine turbulent fluxes from the available energy. These techniques distribute the turbulent fluxes according the Bowen ratio into the sensible and latent heat flux. In experiments, the Bowen-ratio similarity, the ratio of the sensible and latent heat flux is proportional to the ratio of the temperature and humidity difference of two heights, is used. Because the Bowen-ratio method (Bowen 1926) closes the energy balance, the turbulent fluxes are related to the landscape and not to small-scale fluxes near homogeneous measuring fields. A similarity of the small and large scale eddies is assumed.

The same problem is relevant for simple models to determine the potential evaporation with the Priestley-Taylor approach (Priestley and Taylor 1972) or the actual evaporation with the Penman-Monteith equation (Monteith 1965). Both methods use the energy balance equation and distribute the sensible and latent heat flux according to the Bowen ratio. If these methods are calibrated against experimental data measured over small-scale homogeneous surfaces (e.g., the Priestley-Taylor coefficient), e.g., the latent heat flux and the other turbulent flux will be overestimated.

For other trace gas measurements, like the carbon dioxide flux, the problem is similar. Within the FLUXNET network for nearly all stations, only the flux of the small eddies is measured. It must be assumed that also a portion of the flux is exchanged with larger eddies. Already in the past the percentage of the unclosed energy balance was proposed as useful for a correction of the carbon flux (e.g., Twine et al. 2000). Such an assumption is only correct if a similarity of the sum of the sensible and latent heat flux and the carbon dioxide flux is given.

**Conclusions**

All findings from experiments and modeling studies show that measuring errors of terms of the energy balance equation or storage terms cannot explain the problem of the unclosed energy balance and have no significant influence on the residual if the measurements are carefully done. As a hypothesis the energy balance closure problem can be assumed as a scale problem and a closure is only possible on a landscape scale, including the turbulent exchange of the smaller eddies with the
classical eddy covariance method and exchange of larger eddies, which can be up to now only measured with area-averaging measuring systems like scintillometers or airborne sensors. The larger eddies are missing in a landscape without any heterogeneity, like deserts and uniform bush lands.

The energy balance closure has a significant influence on the calibration and validation of models which are made for larger scales like weather prediction or climate models, because the data for comparison are measured on smaller homogeneous scales and typically underestimate the fluxes.

Measurements and models for smaller scales work well on this scale if they are not combined with the energy balance equation. This means that the Monin-Obukhov similarity theory (Monin and Obukhov 1954) is valid and also that measuring methods like the modified Bowen-ratio method (Businger 1986, Meyers et al. 1996, Liu and Foken 2001) and SVAT models which are not based on the energy balance equation.

For the parameterization of the contribution of the larger eddies from measurements of the available energy and the turbulent fluxes of smaller eddies, the problem of the scalar similarity must be studied urgently. As a first guess, a distribution of residual according to the Bowen ratio can only be a temporary method. The same must be said for the correction of other scalar fluxes.

Because in the past the research on the energy balance closure was mainly directed on measuring errors, only a few results underline the scale hypothesis. It should be a subject of further research to recalculate former experiments again and to plan experiments with a specialized measuring setup for this problem. LES modeling studies, especially, should be done to support this research.

Therefore, this overview cannot be a final paper about the energy balance closure problem but a guideline for further steps to come to a final solution.

Acknowledgments


LITERATURE CITED


