

# On the Measurement of Turbulence Over Complex Mountainous Terrain

Ivana Stiperski<sup>1</sup> · Mathias W. Rotach<sup>1</sup>

Received: 18 February 2015 / Accepted: 20 October 2015 / Published online: 7 November 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** The theoretical treatment of turbulence is largely based on the assumption of horizontally homogeneous and flat underlying surfaces. Correspondingly, approaches developed over the years to measure turbulence statistics in order to test this theoretical understanding or to provide model input, are also largely based on the same assumption of horizontally homogeneous and flat terrain. Here we discuss aspects of turbulence measurements that require special attention in mountainous terrain. We especially emphasize the importance of data quality (flux corrections, data quality assessment, uncertainty estimates) and address the issues of coordinate systems and different post-processing options in mountainous terrain. The appropriate choice of post-processing methods is then tested based on local scaling arguments. We demonstrate that conclusions drawn from turbulence measurements obtained in mountainous terrain are rather sensitive to these post-processing choices and give suggestions as to those that are most appropriate.

**Keywords** Atmospheric boundary layer · Coordinate systems · Eddy covariance · Flow separation · Local scaling · Mountainous terrain · Turbulence measurements

## 1 Introduction

The complexity of the atmospheric boundary layer (ABL) stems from the complexity of the underlying surface, both in terms of terrain (determined by height and slope angle) and land-surface characteristics. Mountainous terrain with its pronounced steep slopes and complex topographic features produces an intrinsic spatial inhomogeneity in the boundary-layer response to surface forcing (Rotach and Zardi 2007). Giving an exact definition of mountainous (i.e., truly complex) terrain is a difficult task and generally includes information on different combinations of slope and elevation (Barry 2008). When taking observations

---

✉ Ivana Stiperski  
ivana.stiperski@uibk.ac.at

<sup>1</sup> Institute of Atmospheric and Cryospheric Sciences, University of Innsbruck, Innsbruck, Austria

in mountainous terrain, however, it may happen that a site, for example on a valley floor, is not ‘sufficiently sloped’ to qualify as mountainous even if it is clearly influenced by mountain effects. We therefore refer to ‘mountainous terrain’ when a relevant area qualifies as mountainous (e.g., according to [Barry 2008](#)) even if the local site characteristics do not. It is generally assumed that mountains cover approximately 25 % of the Earth’s land surface.

The ABL in mountainous terrain, though a subject of growing research (e.g. [Turnipseed et al. 2003](#); [Rotach et al. 2004](#); [Park and Park 2006](#); [Hammerle and Haslwanter 2007](#); [Sun 2007](#); [Hiller et al. 2008](#); [de Franceschi et al. 2009](#); [Martins et al. 2009](#); [Nadeau et al. 2013](#); [Večenaj and De Wekker 2014](#); [Medeiros and Fitzjarrald 2015](#)), nonetheless remains elusive to a large degree, mainly due to limited data availability. The lack of consensus on boundary-layer characteristics thus constrains, for example, operational weather, hydrological or climatological modelling efforts in mountainous terrain, which still rely on physical parametrizations developed over flat, horizontally homogeneous, terrain. Determination of turbulent fluxes plays a crucial role in air pollution modelling in complex terrain (e.g., [Szintai et al. 2009](#)), runoff modelling (e.g., [Zappa et al. 2008](#)) or the assessment of the mass balance of a glacier (e.g., [Nicholson et al. 2013](#)) as well as surface energy ([Leuning et al. 2012](#)) or CO<sub>2</sub> balance and ecosystem exchange (e.g., [Sun et al. 2010](#); [Rotach et al. 2014](#)). Many of these applications, furthermore, are by definition relevant only in mountainous terrain.

Turbulent fluxes can directly be measured (or modelled) through the eddy-covariance (EC) approach ([Aubinet et al. 2012](#)) or estimated from mean vertical gradients using Monin-Obukhov similarity theory (MOST, [Monin and Obukhov 1954](#)). Even as the EC approach becomes more widespread, the high experimental costs still result in a virtual absence of routine turbulence measurements in operational networks of, e.g., meteorological or hydrological services. Nonetheless, the EC technique does in principle directly yield the desired turbulent flux, at the single point of observation. Much research has been performed on the technical aspects, challenges and limitations of EC measurements over heterogeneous surfaces especially in terms of data quality and its influence on budget assessments (e.g., [Vickers and Mahrt 1997](#); [Mahrt 1998, 2010](#); [Paw et al. 2000](#); [Baldocchi et al. 2000](#); [Finnigan 2008](#)). Whereas the considerations therein translate also to complex terrain, and especially to gentle hills (e.g., [Taylor et al. 1987](#); [Kaimal and Finnigan 1994](#); [Goulden et al. 1996](#); [Baldocchi et al. 2000](#); [Barford et al. 2001](#), to name just a few), comparatively little work exists so far on the applicability and demands of EC measurements in mountainous terrain, challenging both in terms of surface heterogeneity and steep topography (e.g., [Geissbühler et al. 2000](#); [Turnipseed et al. 2002, 2003](#) and [Turnipseed et al. 2004](#); [Hammerle and Haslwanter 2007](#); [Sun 2007](#); [Hiller et al. 2008](#)). [Turnipseed et al. \(2003\)](#) measured turbulence statistics (e.g., vertical profiles of turbulent quantities, spectral peaks, turbulence length scales) at a forested site in mountainous terrain (slope angle 5°–7°) and concluded that the canopy effect is more significant than the effect of topography. For steeper slopes (24° and 25°, respectively) [Hammerle and Haslwanter \(2007\)](#) and [Hiller et al. \(2008\)](#) concluded, on the basis of energy balance closure considerations (which for their sites was as under-closed as for flat inhomogeneous terrain), that the EC method is appropriate in mountainous terrain.

Employing MOST to infer turbulent fluxes from mean meteorological variables, on the other hand, is not justified in mountainous terrain, since MOST has been developed for, and is hence a priori valid only in, horizontally homogeneous and flat (HHF) terrain. While a substantial literature exists on the application of MOST over inhomogeneous (but flat) terrain, including the requirement for appropriate fetch conditions or footprints (e.g., [Kljun et al. 2002, 2004](#)), again relatively little can be found concerning the applicability of MOST in mountainous terrain. The few studies that do address the subject use different post-processing options and quality criteria that limit the possible universality of their conclusions. For exam-

ple, [de Franceschi et al. \(2009\)](#) presented data from an Alpine valley site, but from a tower on the valley floor and used a recursive filter to eliminate the low frequency oscillations. Data in [Martins et al. \(2009\)](#) are for a very steep site (nearly  $40^\circ$  slope) in Brazil, and prior to analysis they were detrended, and double rotation was applied for the coordinate rotation (see Sect. 4), no flux corrections (see Sect. 3) are mentioned, but non-stationary periods and periods with weak winds were excluded from the analysis by examining only data for which the wind speed was above  $1 \text{ m s}^{-1}$ . [Nadeau et al. \(2013\)](#) presented a test of MOST in complex terrain using data from a steep slope site in the Swiss Alps. No flux corrections were applied in that study as well, and statistical uncertainty was enforced by excluding transition periods, i.e. periods when the sensible heat flux  $|H| < 10 \text{ W m}^{-2}$ . A sectorial planar-fit method (see Sect. 4) was used for coordinate rotation. The results of all of these studies yielded scaling coefficients different than for HHF surface conditions, however, [Martins et al. \(2009\)](#) concluded that “normalized standard deviations of the vertical velocity component ... follow Monin–Obukhov similarity” and [de Franceschi et al. \(2009\)](#) reached a similar conclusion. [Nadeau et al. \(2013\)](#) attribute the difference between their best-fit curves and MOST to the failure of the constant-flux assumption at their site. [Kral et al. \(2014\)](#) examined the validity of MOST for a station at the bottom of a wide arctic fjord and found the results dependent on meteorological conditions. The non-dimensional vertical wind shear was found to be most sensitive to the topographic influence. No dependence of non-dimensional gradients on coordinate rotation (see Sect. 4) was observed. The results of these studies cannot be generalized until we can specify more objective quality criteria (e.g. not arbitrarily chosen limits for the heat flux or minimum wind speed) and until we quantify the uncertainties introduced by different post-processing choices they employ.

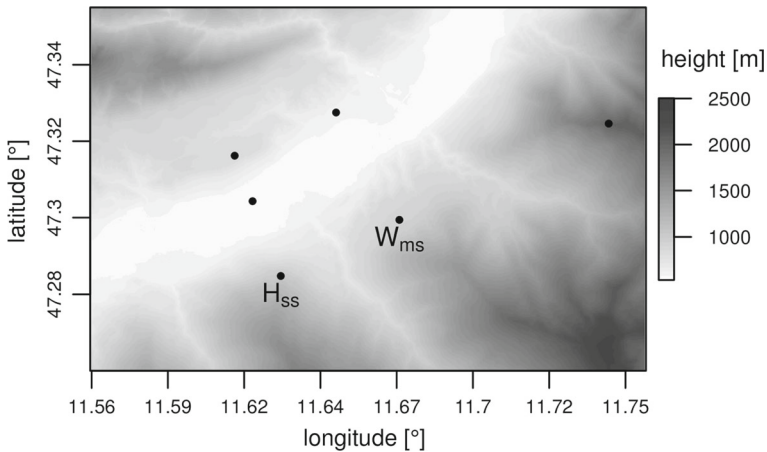
In the present paper we first address the technical problems that arise when directly measuring turbulent fluxes in mountainous terrain. Secondly we test different post-processing choices and their influence on turbulence statistics, so as to determine the best practice for measurements in mountainous terrain. For this purpose we evaluate data from mountainous sites in terms of scaling relations, which we choose as an objective criterion for the performance of different post-processing methods. These scaling relations do not necessarily have to comply with MOST (i.e., yield identical empirical relations, including the numerical values for the involved parameters) but rather present a best-fit to a curve in terms of local scaling (cf. [Nadeau et al. 2013](#)). Post-processing options that give the smallest scatter (i.e., the best fit) will consequently be considered the preferred choice. Thus we use local similarity as a guiding principle, but do not address the applicability of MOST in mountainous terrain (the latter will be a topic of a follow-up paper).

In Sect. 2 we introduce the datasets and methods used, Sect. 3 is devoted to assessing the impact of quality control and different (technical) choices when applying eddy covariance in truly complex mountainous terrain, while Sect. 4 looks at coordinate systems and coordinate rotation options in mountainous terrain. In Sect. 5 we test different post-processing options by determining the best fit to local scaling, and finally, conclusions and a further outlook are provided in Sect. 6.

## 2 Datasets and Methods

### 2.1 Datasets

Four datasets based on observations over complex mountainous terrain are used. All the datasets originate from the i-Box project ([Stiperski et al. 2012](#)). This is a platform for studying



**Fig. 1** Topographic map of the i-Box area in the Inn Valley, Austria. Locations of i-Box sites are shown by black dots. Data from north facing sites  $H_{ss}$  and  $W_{ms}$  are analyzed here

exchange processes in truly complex mountainous terrain, comprising six sites in the Inn Valley, Austria (Fig. 1), a moderately wide (less than 20 km ridge separation distance), steep valley (ranging from roughly 600 m a.s.l. at the valley floor to about 2300 m a.s.l. at mountain ridges to the north and 2700 m a.s.l. to the south) with complex topographical features such as plateau-like foothills, side valleys and slopes of different exposition and steepness. The sites are situated in terrain of varying complexity both in terms of local and global slope angles, vegetation cover, and exposition. Full details and the scope of the i-Box project will be described elsewhere (Rotach et al. in preparation), here we only use data from two sites.

We focus on 11 months of data (August 2013 to June 2014) from two north-facing stations: a very steep site Hochäuser (denoted  $H_{ss}$  in the following with a  $27^\circ$  slope) and a moderately steep site Weerberg (denoted  $W_{ms}$  with a  $10^\circ$  slope). The measurement height is 6.8 m above the surface for the former, and 7 m for the latter site. The footprint of observations at both sites can be characterized as an alpine meadow, quite homogeneous in all directions, especially for site  $H_{ss}$ . Site  $H_{ss}$  is situated on a semi-convex undulating 400 m long slope, surrounded by a line of individual trees. The lateral distance to the tree line is approximately 70 m with a forest located 250 m uphill of the station. Station  $W_{ms}$  is located on a uniform, smooth 500 m long slope, with a locally more complex footprint affecting the flow from the sector  $135^\circ$ – $160^\circ$ . This however does not influence the dominant flow regimes (along-slope and along-valley). A forest is located approximately 300 m uphill of the tower where the terrain slope increases significantly. Wind and temperature data are collected with Campbell CSAT3 sonic anemometers and humidity fluctuations with a fast-response Krypton hygrometer (KH20 of Campbell Scientific) at 20-Hz resolution and analyzed using EdiRe data software from the University of Edinburgh (Clement and Moncreif 2007).

Additionally, data from two three-week long campaigns, organized as preliminary test phases at site  $H_{ss}$ , are used. During the first campaign (denoted  $H_{ss\_T1}$ ), conducted before the fixed tower was set-up, a mobile 2-m high tower with two CSAT3 sonic anemometers was operated with one sonic anemometer oriented in the vertical direction and the other slope-normal. The full energy balance was also assessed during this time, and the energy balance closure was found to be very poor: on average on the order of 50 %. During the second field campaign ( $H_{ss\_T2}$ ) two additional mobile 2-m tall towers were located along

the  $H_{ss}$  slope (at approximately 100-m intervals) with one sonic anemometer and two levels of temperature measurements (at 1 and 3 m above the ground).

### 2.2 Methods

The temporal averaging can be done using block averaging (i.e., simple arithmetic mean) or the data might be filtered (detrended or using higher level filters) prior to averaging to eliminate the non-turbulent low frequency contributions to the flux. Although it has been claimed that filtering violates Reynolds averaging (Richardson et al. 2012, based on Rannik and Vesala 1999), its use is still encouraged by the same authors due to the reduction of random errors. Filtering is often applied (e.g., de Franceschi et al. 2009; Martins et al. 2009) in complex terrain, particularly under stable stratification. As such we test here its influence on turbulence statistics (see Sect. 5).

Before applying Reynolds decomposition (Stull 1988) and computing integral turbulence statistics, the measurements have to be transformed from the instrument coordinate system into a Cartesian coordinate system, usually the streamline coordinates, where the  $x$ -axis is directed into the mean wind direction. Traditionally, the coordinate transformation is performed using the so-called double rotation, by which the mean lateral and vertical velocity components are brought to zero for each averaging interval (e.g., McMillen 1988). For completeness we also mention the triple rotation though this third rotation is usually discouraged (e.g., Rebmann et al. 2012). In complex terrain the planar-fit approach (Wilczak et al. 2001), which essentially fits a climatological plane through the streamlines of the local flow, is commonly used (cf. Mauder et al. 2013; Nadeau et al. 2013; Kral et al. 2014).

Similarity theory (e.g., Stull 1988) is often invoked to relate turbulent statistics to mean variables that are more easily measured. For the surface layer, under the assumption of HHF surface conditions, local scaling (and more generally MOST) states that any suitably scaled mean variable is a function of only one non-dimensional group  $z/\tilde{L}$

$$\frac{\bar{a}}{a_*} = f_a \left( z/\tilde{L} \right), \tag{1}$$

where  $a_*$  is the corresponding scaling variable for  $a$ ,  $z$  is the height above the surface, and the local Obukhov length  $\tilde{L}$  is defined as

$$\tilde{L} = -\frac{1}{k} \frac{\tilde{u}_*^3}{w'\theta'} \left( \frac{g}{\theta_v} \right)^{-1}, \tag{2}$$

where  $k = 0.4$  is the von Karman constant, and the local friction velocity is defined as

$$\tilde{u}_* = \left[ \left( \overline{u'w'} \right)^2 + \left( \overline{v'w'} \right)^2 \right]^{1/4}. \tag{3}$$

Using the full definition of friction velocity is essential in mountainous terrain where the lateral component of the momentum flux can be of the same order as the longitudinal component (Rotach et al. 2008). Apart from the HHF terrain requirement in the derivation of MOST, the assumption of constant fluxes (i.e.,  $\overline{w'\theta'} \approx \overline{w'\theta'_0}$  throughout the surface layer, and correspondingly for the momentum fluxes) has to be made. The general form of the scaling for the standard deviation of the vertical velocity component,

$$\frac{\sigma_w}{\tilde{u}_*} = \phi_w \left( \frac{z}{\tilde{L}} \right) = a \left( 1 + b \left| \frac{z}{\tilde{L}} \right| \right)^c, \tag{4}$$

is commonly used as a test of MOST (cf. Nadeau et al. 2013) or, more generally, local scaling. Here we fit the data to this functional form (Eq. 4) and use the scatter of the data as a measure of the goodness of the post-processing options, to be defined in the following sections. The post-processing method with the smallest scatter around the best-fit curve should be the most appropriate for use in mountainous terrain.

Monin-Obukhov type similarity is plagued by self-correlation since  $\tilde{u}_*$  is found both on the abscissa and the ordinate (Klipp and Mahrt 2004). Grachev et al. (2013) mention that this spurious correlation leads to deviations from  $z$ -less scaling in the stable ABL. Although this problem is known for surface-layer scaling relations it is not fully addressed here, since it affects different post-processing methods in the same manner and therefore will not be an impediment in determining the difference between them. We return to this topic in Sect. 5.

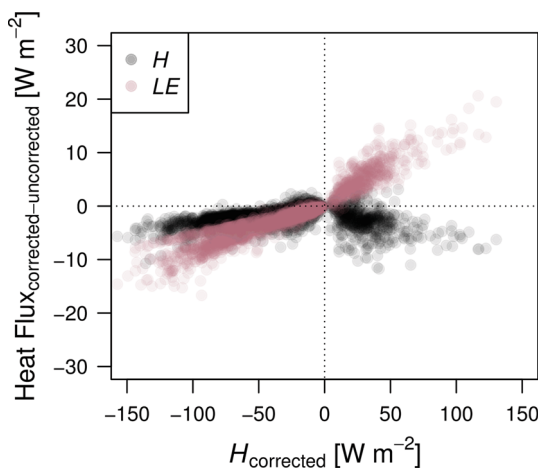
### 3 Data Quality

#### 3.1 Flux Corrections, Data Quality Assessment and Post-processing Options

Turbulent statistics obtained by standard EC instrumentation need to be corrected for several sources of systematic errors. Sensible heat flux measured by sonic anemometers has to be corrected for humidity effects (and cross-wind effects for some sonic anemometers) following Schotanus et al. (1983), latent heat flux (using open-path analyzers) for oxygen effects (van Dijk et al. 2003, in the case of krypton-based analyzers) and density effects (Webb et al. 1980). All fluxes, including momentum fluxes, require spectral corrections due to sensor characteristics and sensor separation (e.g., Moore 1986). Additionally recently Horst et al. (2015) have proposed a flow distortion correction for CSAT3 sonic anemometers, this, however, was not applied here. Still, due to the fact that this correction is not dependent on stability and wind direction it is not expected to affect the results of Sect. 5 in a decisive manner.

Flux corrections can be applied a posteriori and are independent of surface complexity. Figure 2 shows the difference between the corrected and uncorrected sensible and latent heat fluxes as a function of sensible heat flux at site  $H_{ss}$ . Latent heat flux and the positive sensible fluxes are particularly sensitive to corrections, so that for sites with low Bowen ratio (e.g. site  $H_{ss}$ ) the contribution of corrections becomes non-negligible (on average 10–20 %). Note that this estimate is slightly too low due to the neglect of the flow distortion correction (Horst

**Fig. 2** Difference between the corrected and uncorrected sensible ( $H$ ) and latent ( $LE$ ) heat fluxes as function of corrected sensible heat flux  $H$ , from 11 months of data at site  $H_{ss}$ . Corrections are detailed in Sect. 3



et al. 2015). Failure to apply corrections can thus lead to erroneous energy and scalar budgets and Bowen ratio estimates.

Before any further analysis of the data is performed, data quality should be assured. There have been a number of studies dealing with data quality assessment (e.g., Foken and Wichura 1996; Vickers and Mahrt 1997; Foken 2008; Foken et al. 2012; Mauder et al. 2013). Issues of data quality are particularly pronounced in the very stable ABL (e.g., within a cold pool at the valley floor) where non-turbulent contributions and non-Kolmogorov turbulence can significantly contaminate turbulence statistics and affect conclusions based on these (Klipp and Mahrt 2004; Grachev et al. 2013; Sanz Rodrigo and Anderson 2013). Although of paramount importance, the issue of data quality is not specific to mountainous terrain, we therefore refer to the above publications for a thorough overview.

In the present paper the following quality control is performed on the computed fluxes to ensure good data quality. When one of the following occurred averaging intervals were discarded from further analysis:

1. The sonic diagnostic flag was set high (malfunctioning of the instrument) inside the averaging period.
2. KH20 voltage fell below 5 mV (indication of condensation occurring on the KH20 window).
3. Skewness of temperature and wind components fell outside the  $[-2, 2]$  range, following Vickers and Mahrt (1997).
4. Kurtosis of temperature and wind components was  $>8$ , following Vickers and Mahrt (1997).

Additionally, the interdependence of fluxes, introduced through correction procedures, is taken into account (e.g., the erroneous latent heat flux influences the sensible heat flux through the correction, after Schotanus et al. 1983). Thus, the corrected heat flux is discarded during periods with low quality latent heat flux and vice versa.

The above list comprises the *minimum quality criteria* for data to be used herein and up to 30 % of heat fluxes are usually discarded for both sites. The large amount of data not passing the minimum quality control is mainly due to heavy rain, and issues with humidity measurements using KH20 instruments over prolonged periods of time in rough environmental (mountainous) and wet conditions. If corrections are not applied and cross-contamination of errors is not introduced, less than 10 % of heat-flux data on average would fail this minimum quality control.

A test for integral turbulence (Foken and Wichura 1996) is not performed here, since both sites examined have significant turbulence throughout the day and night (due to katabatic flows, not shown). Additionally, since integral turbulence characteristics are used as a tool to assess the suitability of different post-processing methods (see Sect. 5) performing the test for integral turbulence as part of the quality control could bias the conclusions by forcing the data to comply with the scaling.

Apart from data that pass the minimum quality control, we define the *high quality* (HiQ) dataset (to be used in Sects. 4 and 5) as the one that additionally satisfies the quality criteria for stationarity and statistical uncertainty (see the following section for full definition).

Stationarity of each averaging interval can be assessed using an a posteriori test, for which purpose several are available (e.g. Bendat and Piersol 1986; Foken and Wichura 1996; Vickers and Mahrt 1997; Mahrt 1998). In principle, stationarity of all turbulence statistics under consideration should be tested, but most often only second-order statistics (i.e., variances and covariances) are examined. Večenaj and De Wekker (2014) have performed a thorough analysis of different stationarity tests and conclude that there is no single test that is able

to recognize non-stationarity under all conditions. They, however, do single out the [Foken and Wichura \(1996\)](#) test as the single most rigorous. This test is therefore used here with its standard level of 30 (up to 30 % difference between 5-min and 30-min statistics) for data to be declared stationary. If the time series is filtered prior to the analysis, as is done here for certain processing options (see below), at least first-order statistics (i.e., mean variables) are stationary by construction, and most often this also applies to higher-order statistics.

Post-processing of turbulence time series requires a number of choices (e.g. averaging time, linear detrending vs. more sophisticated filtering or none, de-spiking) to be made. These decisions are hardly influenced by the presence of mountainous terrain, so the reader is referred to [Foken \(2008\)](#) and [Aubinet et al. \(2012\)](#) where comprehensive discussions concerning all these aspects can be found. A quite common choice for post-processing options (e.g., [Finnigan et al. 2003](#); [Mauder et al. 2013](#)) is to apply block averaging over 30 min in conjunction with planar-fit coordinate rotation. We discuss the applicability of these choices in Sects. 4 and 5.

### 3.2 Estimating Uncertainty

Any measurement of turbulence statistics involves a degree of statistical uncertainty or sampling (representation) error. Only if the measurements under consideration were repeated for an infinite number of times under identical conditions, so as to obtain the true ensemble average, this error would tend towards zero. Different estimates of this uncertainty are available (see e.g. [Mahrt 1998, 2010](#); [Finkelstein and Sims 2001](#); [Vickers et al. 2010](#); [Billesbach 2011](#); [Mauder et al. 2013](#)). Most methods require elaborate measurement techniques, an a priori knowledge of integral time scales or an iterative approach. Here we propose an alternative simple method for determining the uncertainty of each averaging period from the measurements themselves.

[Wyngaard \(1973\)](#) estimated the averaging interval required for each statistical moment to meet a given uncertainty level, from a relation that can be viewed as the signal-to-noise ratio, noting that higher-order turbulence statistics need a longer averaging interval. Additionally, several authors have shown a need for a variable, stability-dependent averaging interval (e.g., [Vickers and Mahrt 2003](#); [Acevedo et al. 2006](#)). In applications, and especially in long-term measurements, as herein, a fixed averaging interval is usually employed for the entire dataset (e.g., [Rebmann et al. 2012](#); [Foken et al. 2012](#); [Nadeau et al. 2013](#); [Mauder et al. 2013](#)) since variable averaging intervals are impractical to implement. Still, following [Wyngaard \(1973\)](#) the information on the appropriate averaging interval is retained in the information on uncertainty. Therefore, following, for example, [Forrer and Rotach \(1997\)](#), we invert the procedure suggested by [Wyngaard \(1973\)](#), and estimate the statistical uncertainty of turbulent variances and covariances in each averaging period for a given fixed averaging interval. Here we predominately use 30 min (although we also test a 60-min averaging interval in Sect. 5). Choosing the level of uncertainty that we find acceptable for our data thus ensures that the choice of the averaging interval (e.g., 30 min) is appropriate.

The statistical uncertainty in momentum and heat fluxes, respectively, and for the variances is assessed following [Wyngaard \(1973\)](#) using

$$a_{uw}^2 = \frac{z}{\tau_a U} \left[ \frac{\overline{(u'w')^2}}{\tilde{u}_*^4} - 1 \right], \quad (5a)$$

$$a_{vw}^2 = \frac{z}{\tau_a U} \left[ \frac{\overline{(v'w')^2}}{\tilde{u}_*^4} - 1 \right], \quad (5b)$$



$$a_{w\theta}^2 = \frac{z}{\tau_a U} \left[ \frac{\overline{(w'\theta')^2}}{\overline{(w'\theta')^2}} - 1 \right], \tag{6}$$

$$a_{x^2}^2 = \frac{4z}{\tau_a U} \left[ \frac{\overline{x'^4}}{\overline{(x')^2}^2} - 1 \right], \tag{7}$$

where  $a_{xy}$  is the uncertainty for the indexed turbulence (co)variance,  $\tau_a$  is the averaging interval,  $U$  is the mean wind speed at height  $z$ ,  $\tilde{u}_*$  is the friction velocity as defined in Eq. 3 and  $x$  is any turbulence variable.

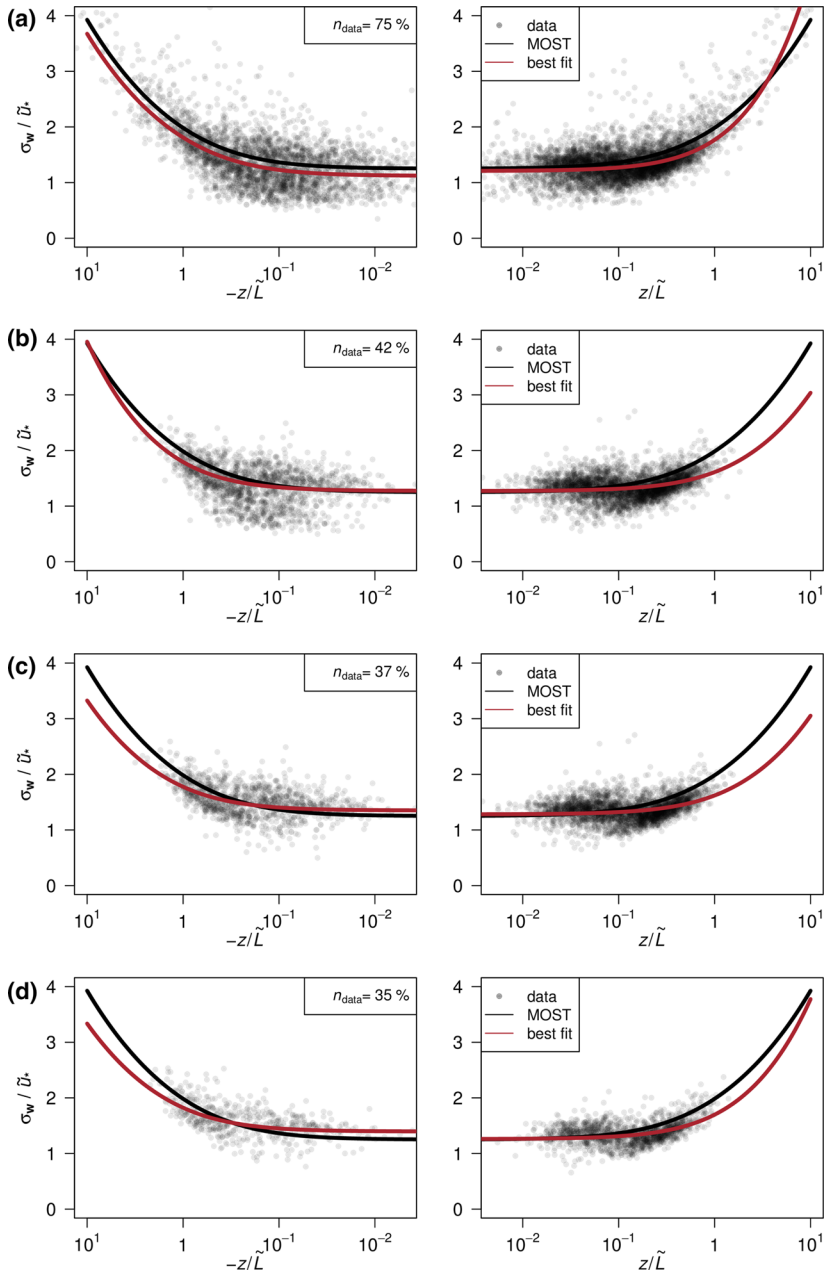
Traditional ‘micrometeorological wisdom’ would set the uncertainty of a flux measurement (outside the transition times, say) at  $\approx 20\%$ , e.g. [Vickers et al. \(2010\)](#). In complex terrain, and generally at non-ideal sites, however, this uncertainty can be substantially larger. Whilst data with high uncertainty are not necessarily non-stationary (or vice versa), periods declared non-stationary according to [Foken and Wichura \(1996\)](#) on average have uncertainty levels  $> 20\%$ .

Non-stationary periods in mountainous terrain cover more than 30 % of all the data for site  $H_{ss}$  that passed the minimum quality control (see Sect. 3), and more than 60 % for site  $W_{ms}$ . The high number of non-stationary periods appears to be typical for complex terrain and corresponds to those reported in [Večenaj and De Wekker \(2014\)](#). Still, depending on the limit chosen, uncertainty is the more stringent criterion on data quality than non-stationarity, and at a 20 % uncertainty level can identify twice as many periods of dubious quality as does non-stationarity. For a 50 % limit, uncertainty and non-stationarity identify a similar number of periods with poor quality (some 40 %). It is interesting to note that the uncertainty in variances is on average  $< 20\%$  for a 30-min averaging interval, whereas the covariances have a larger uncertainty.

Figure 3a, b shows the scaled standard deviation of vertical velocity fluctuations ( $\sigma_w/\tilde{u}_*$ ) as a function of stability ( $z/\tilde{L}$ ) from site  $H_{ss}$ , with data that pass the minimum quality control but have different levels of statistical uncertainty (data that pass the minimum quality control, i.e. uncertainty is not limited, are shown in Fig. 3a and for uncertainty below 50 % are shown in Fig. 3b). Choosing only data that satisfy a more restrictive criterion on statistical uncertainty largely reduces the scatter around the best-fit curve defined by Eq. 4, stressing the need to include the information on uncertainty when assessing, for example, local scaling properties. However, it also reduces the total amount of available data. This is especially so in the near-neutral range and in the strongly convective and strongly stable limits, respectively. At the uncertainty limit of 20 %, as much as 80 % of data from sites  $H_{ss}$  and  $W_{ms}$  would be declared to have low quality (not shown). The criterion for uncertainty and stationarity employed here thus eliminates the need to discard data with low fluxes by choosing arbitrary thresholds (cf. [Klipp and Mahrt 2004](#); [Nadeau et al. 2013](#); [Sanz Rodrigo and Anderson 2013](#)).

Fitting a functional form to the data (e.g. of the form given in Eq. 4), in order to investigate for example the degree of local scaling in complex terrain is therefore rendered extremely sensitive to the availability of sufficiently long datasets (longer than a year) that cover the relevant stability range, and a sound assessment of data quality. For energy balance closure assessment, choosing only data with small statistical uncertainty (below 20 %, say) would be prohibitive.

We define our *high quality dataset* (HiQ) as that for which the data satisfy the minimum quality control (Sect. 3), are stationary (according to [Foken and Wichura 1996](#)) and the uncertainty of which is smaller than 50 % (cf. Fig. 3c, d). We choose 50 % as an optimal compromise between relatively high data quality and relatively moderate data loss (still retaining data in the very (un)stable and near-neutral range).



**Fig. 3** Scaled standard deviation of vertical velocity fluctuations,  $\sigma_w / \tilde{u}_*$ , as a function of stability  $z/\tilde{L}$ . Data are from 11 months of observations at site H<sub>SS</sub> with double rotation (coordinate rotation, see Sect. 4) and block-averaging with **a–c**  $\tau_a = 30$ -min, and **d**  $\tau_a = 60$ -min averaging interval. Different panels show different levels of quality control: **a** minimum quality control defined in Sect. 3, **b** uncertainty of variances and co-variances according to Eqs. (5)–(7) below 50 %, **c** uncertainty below 50 % and stationary (HiQ), **d** HiQ data but for 60-min averaging period. *Red line* best fit to the data points according to Eq. 4, *black line* MOST fit according to Panofsky and Dutton (1984).  $n_{data}$ : percentage of averaging periods (out of the total data number) in both stable and unstable regime that fulfil the above mentioned quality criteria

## 4 Coordinates

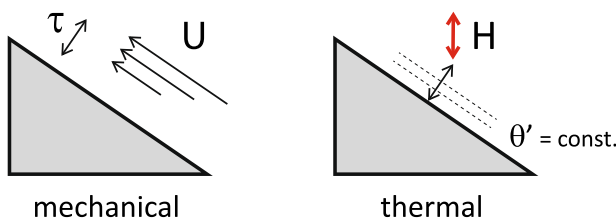
Once the topography on which measurements are performed is no longer flat, the question of coordinate systems becomes important. This is especially valid for mountainous terrain where steep slopes of various complexities are the norm.

### 4.1 Coordinate System

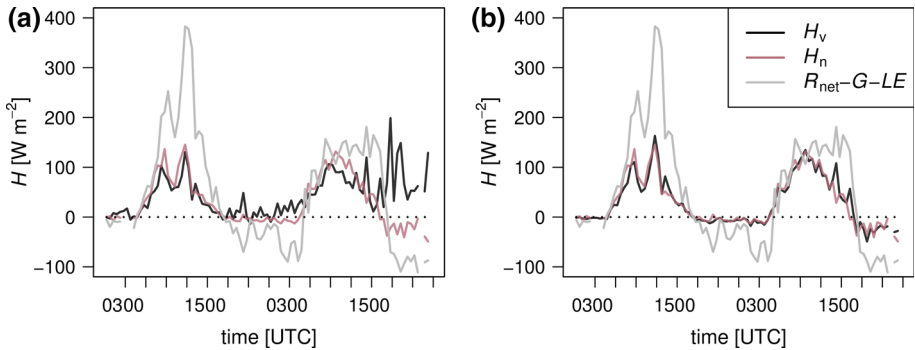
Measured turbulent fluctuations must first be brought into a Cartesian coordinate system, usually defined as natural coordinates, i.e. with its  $x$ -axis pointing into the mean flow direction and the  $z$ -axis normal to the surface. Under HHF surface conditions the  $z$ -axis corresponds to the vertical direction and thus to that of dominant turbulent exchange. Similarly, on a sloped surface, friction will lead to a shear stress normal to the surface (Fig. 4, left) so that the local normal seems to be the appropriate coordinate direction for the ‘vertical flux’. If we assume the perturbation isentropes to be parallel to the local surface on a slope (Fig. 4, right), very close to the surface the local normal again is appropriate for the sensible heat flux. Indeed, multiple temperature measurements along the  $H_{ss}$  slope during the  $H_{ss\_T2}$  campaign, show very weak along-slope potential temperature gradients at 1- and 3-m heights, suggestive of isentropes that are almost terrain-following close to the surface (not shown).

Heat flux, however, is driven by buoyancy that is directed along the geopotential, i.e. in the vertical direction defined by the Earth’s geoid, irrespective of the underlying slope. At a certain height, therefore, the dominant sensible heat flux is ‘vertical’ rather than normal to the local slope.

Finnigan (2004) has pointed out that when comparing vertical to slope-normal fluxes the apparent surface portion must be taken into account in a similar fashion as for radiative fluxes (e.g., Matzinger et al. 2003). This reduces the vertical flux in proportion to the cosine of the slope angle. Figure 5 shows turbulent sensible heat fluxes in the slope-normal and vertical reference frame over two radiatively-driven days during campaign  $H_{ss\_T1}$ . The slope-normal and vertical fluxes were measured by sonic anemometers at 2-m height, installed in the slope-normal and vertical direction. Since obtaining an accurate slope-normal orientation of the sonic anemometer is very difficult, especially for non-uniform slopes, we here assume that the  $z$ -coordinate obtained by coordinate rotation in post-processing is normal to the slope. Therefore, the slope-normal heat flux in Fig. 5a is the one rotated by double rotation (see next section). The rotation angles obtained by double rotation are, however, small compared to the slope angle, and the difference between the rotated heat flux and that measured directly in the slope-normal orientation is small, justifying our set-up. To ensure that both sonic



**Fig. 4** Definition of local coordinates over (ideally) sloped terrain,  $\tau = -\rho u'w'$  corresponds to the longitudinal component of shear stress (frictional stress),  $U$  is the longitudinal wind component,  $\theta'$ -lines denote isolines of potential temperature perturbations (around a mean state) and  $H = \rho C_p \overline{w'\theta'}$  is the sensible heat flux



**Fig. 5** Daily cycle of sensible heat flux,  $H = \rho C_p \overline{w'\theta'}$  at 2-m height for site  $H_{ss}$  during  $H_{ss\_T1}$ , when a mobile tower with two sonic anemometers was operated: one installed vertically ( $H_v$ ) and one slope-normal ( $H_n$ ). Data are from two consecutive days with weak synoptic forcing ('purely' thermally driven valley and slope wind regime). **a** Values of heat flux from both sonic orientations, rotated only into the mean wind. Coordinate rotation for  $w$  was only applied to slope-normal orientation. **b** Values of heat flux from both installations, after double rotation was applied to both. The remaining terms of the energy balance closure (net radiation  $R_{net}$  minus ground heat flux,  $G$  and latent heat flux,  $LE$ ) are also plotted as a reference

anemometers were indeed measuring the same footprint and produced meaningful results, we have rotated both measurements using double rotation and obtained a good match between the heat fluxes (Fig. 5b).

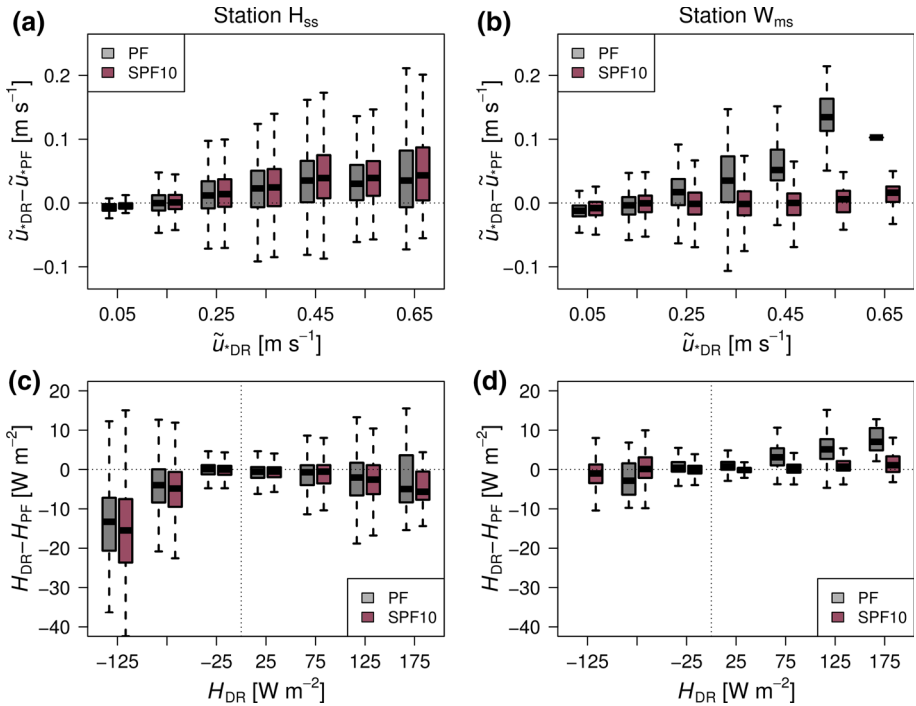
The results for radiatively-driven days (two of which are shown in Fig. 5a) show that the slope-normal sensible heat flux dominates during the day, and is on average about 12 % larger than the vertical. This is in line with Finnigan (2004) attributing some 10–15 % of the difference (according to the local slope, equal to  $27^\circ$  at site  $H_{ss}$ ) to the apparent surface. The daytime difference between the normal and the vertical frame of reference although small is statistically significant for radiatively-driven days. Still, the statistical significance is lost for other weather patterns during the three-week long campaign at site  $H_{ss}$ . During nighttime the sensible heat flux in the vertical framework is likely erroneous (i.e., positive) and works in the opposite direction, away from closing the energy balance (Fig. 5a). Thus, apart from daytime fluxes during radiatively-driven days, the apparent surface argument of Finnigan (2004) is not supported by our data.

Under the assumption of local isentropes parallel to the surface (i.e., turbulent exchange normal to the local gradient) the vertical reference frame therefore misses a substantial part of the turbulent exchange (which is then attributed to  $\overline{u'\theta'}$ ). It is, however, quite unclear at present, up to which height this systematic difference prevails.

## 4.2 Coordinate Rotation

Finnigan et al. (2003) have disentangled in detail the various contributions to the total flux in different coordinate systems (note that they were in fact interested in the necessary length of the averaging period in order to capture the total flux, of mass in their case, normal to the surface). We take their long-term averaging as a surrogate for the planar-fit approach (Wilczak et al. 2001) where appropriate for the present purpose. Adapting to our notation, planar-fit (PF) and double rotation (DR) fluxes (kinematic sensible heat flux in this example) for a long enough averaging period are related through

$$\overline{w_{PF}''\theta_{PF}''} + \overline{w'\theta'}_{PF} = \overline{w'\theta'}_{DR} \quad (8)$$



**Fig. 6** Difference between double rotation (DR) and different versions of planar-fit rotation (single plane: PF and ten different planes: SPF10) for friction velocity ( $\tilde{u}_*$ ; **a, b**) and sensible heat flux ( $H$ ; **c, d**), for site  $H_{ss}$  (**a, c**) and site  $W_{ms}$  (**b, d**). Only high quality data (see Sect. 3) are shown

where the double prime denotes the departure of the mean in the corresponding averaging period from the climatological reference state (i.e., for planar-fit, the climatological normal defining the vertical coordinate, and a base state defining the mean potential temperature field). Even over perfectly HHF surfaces, therefore, for a single averaging interval planar-fit may yield a different flux than double rotation since the momentary vertical coordinate in the double rotation is not exactly normal to the climatological surface as defined by planar-fit (and hence  $\bar{w}'' \neq 0$  compensating for local horizontal flux divergence). When considering an ensemble average (over different realizations/periods) under identical (boundary) conditions these differences will average out if the flow is truly one-dimensional. If we consider a uniform infinite slope to approach mountainous terrain this picture should, in principle, not change, at least as long as the one-dimensionality of the flow is preserved (e.g., in a physical or numerical model). Hence any systematic difference between double rotation and planar-fit fluxes will be an indication of the importance of, what could be considered a surrogate of, local advection, despite the fact that infinite slopes are not very common in mountainous terrain.

For sites with larger local terrain complexity several climatological planes can be fitted according to different wind sectors as in Ono et al. (2008) and Yuan et al. (2011). This method is called directional or sectorial planar-fit. Here we tested several planar-fit options: fitting a single plane, four planes chosen according to dominant wind directions and terrain complexity, and ten planes with equal size of the sectors. The planar-fit angles were calculated from 30-min averages from the 11-month dataset.

Figure 6 shows the difference between double rotation and different types of planar-fit rotations for friction velocity and sensible heat flux at sites  $H_{ss}$  and  $W_{ms}$ . The data shown

satisfy the HiQ quality criteria (see Sect. 3), which mostly eliminates data around the near-neutral regime ( $H \approx 0$ ). The magnitude of the momentum fluxes at both sites  $H_{ss}$  and  $W_{ms}$  (Fig. 6a, b) are generally larger with double rotation than with planar-fit options. This can mostly be attributed to differences in the lateral momentum flux between the two methods, as also observed by Kral et al. (2014) over an Arctic fjord, suggesting the existence of  $\bar{v}''\bar{w}''$ , or the ‘local advection’ of higher momentum air. In other words, planar-fit underestimates the magnitude of momentum fluxes. Apparently, a sectorial planar-fit largely resolves the underestimation of the planar-fit approach at the site with moderate slope ( $W_{ms}$ ), caused by direction-dependent irregularities in the different upwind sectors.

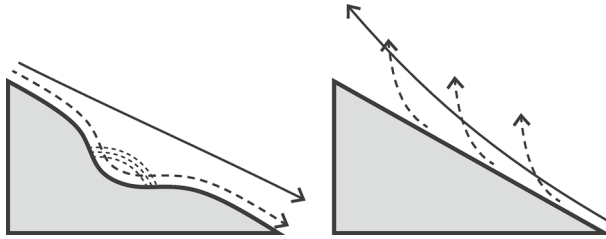
The results for the sensible heat flux are site dependent. We can first note that the nighttime sensible heat fluxes at site  $H_{ss}$  can be extremely negative (Fig. 6c), coincident with periods of foehn winds (e.g. Lothon et al. 2003) at this site. Additionally, nighttime heat fluxes computed with double rotation are systematically more negative (larger absolute value) than those computed with planar-fit methods. On the other hand, during daytime the planar-fit rotations yield slightly larger positive sensible heat fluxes than double rotation. There also does not appear to be a significant difference between fitting a single plane, fitting ten planes or four planes (not shown). At the less steep site  $W_{ms}$  sensible heat flux (Fig. 6d) is on average larger for double rotation than for planar-fit, irrespective of stability. Contrary to site  $H_{ss}$ , there is a systematic difference between fluxes obtained with single plane (PF) and ten planes (SPF10), so that sectorial planar-fit with ten planes yields similar results as double rotation, the same as for the momentum flux. Comparable results were obtained from the MAP-Riviera project (unpublished M.Sc. thesis, Brugger 2011) where no difference between the double rotation and planar-fit rotation methods was found at a valley floor site, whilst on a 40° slope fluxes obtained with the double rotation approach were found to be significantly larger (some 20 %) than with different versions of planar-fit.

All these findings are generally at odds with those of Turnipseed et al. (2003) who showed that the difference in fluxes calculated by different rotation methods (double rotation vs. different types of planar-fit) for their dataset (a relatively gentle forested slope) were not significant. Kral et al. (2014) also observed no systematic bias for sensible heat flux obtained by both methods, however, they note a wind-direction dependence that could point to the need for a sectorial planar-fit. Obviously the differences between the rotation methods depend on the local slope angle but even more so on the local slope complexity.

Although the differences between different rotation methods are comparatively small, they can still be larger than the influence of flux corrections (see Sect. 3, Fig. 2), especially for strong negative heat fluxes at site  $H_{ss}$ . The differences would be especially large during transition periods (not shown), but these are eliminated here with the stationarity and uncertainty quality criteria.

Given the discussion above, the difference between sensible heat fluxes obtained with double rotation and planar-fit point to surface inhomogeneity and hence ‘local advection’ (cf. Eq. 8) at our sites, and can mostly be attributed to slope flows. It is important to note that for both of our north-facing sites, the prevalence of downslope wind exhibits the dominant influence on determining the single climatological plane (planar-fit).

The contribution of  $\bar{w}''\bar{\theta}''$  in Eq. 8 seems to be site specific. At the relatively uniform site  $W_{ms}$  the discrepancy between double rotation and planar-fit methods is most pronounced for large positive sensible heat fluxes (unstable upslope flow), when  $\bar{w}''\bar{\theta}''$  is small but positive. These results can be attributed to warm-air advection ( $\bar{\theta}'' > 0$ ) in the upslope flow and thermals that separate from the surface ( $\bar{w}'' > 0$ ; cf. Hocut et al. 2015) thus crossing the climatological plane (Fig. 7a) that is mostly determined by terrain-following downslope flow. At the same time the difference between double rotation and sectorial planar-fit with



**Fig. 7** Schematic diagram for (left) downslope flow at site  $H_{ss}$  with undulating slope. Dashed arrow indicates terrain-following weak flow, full arrow indicates high-speed flow that separates from the slope (separation zone is shown with thin dashed lines); (right) upslope flow at site  $W_{ms}$  with a uniform slope. Full arrow indicates upslope flow, dashed arrows indicate thermals separating from the surface

ten planes is comparatively negligible because the uniformity of the slope enables clear separation between the different planes for upslope and downslope flow.

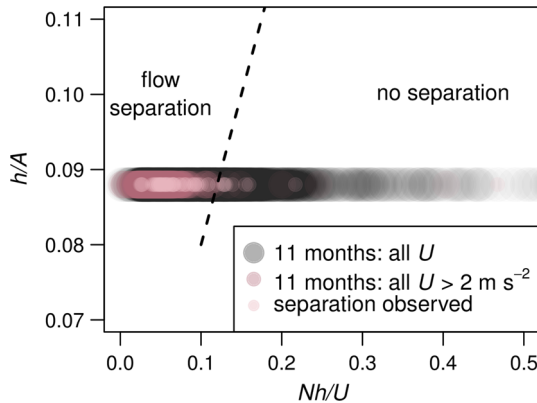
For the less uniform site  $H_{ss}$  the influences are more complex. Both the upslope flow where the heat flux for planar-fit ( $H_{PF} > H_{DR}$ ), and the downslope flow where the magnitudes of  $H_{DR} > H_{PF}$  are associated with a negative  $\bar{w}''\bar{\theta}''$ . The upslope flow (i.e., warm advection,  $\bar{\theta}'' > 0$ ) at site  $H_{ss}$  is associated with sinking motion compared to the climatological plane ( $\bar{w}'' < 0$ ) because the climatological plane, dominated by the downslope flow, is steeper than the upslope flow itself (Fig. 7b).

For the downslope flow (i.e. cold air advection,  $\bar{\theta}'' < 0$ ) negative  $\bar{w}''\bar{\theta}''$  points to rising motion compared to the climatological plane ( $\bar{w}'' > 0$ ). The reason for this may be flow separation from undulations in the terrain (Fig. 7b). For the weak downslope flow ( $U < 2 \text{ m s}^{-1}$ ), which dominates the climatology (more than 80 % of data fall into this category), the flow is terrain following. However, faster flow ( $U > 2 \text{ m s}^{-1}$ ) can separate from the surface and approach the station from a less steep angle relative to the slope thus leading to apparent rising motion through the climatological plane. The largest difference between planar-fit and double rotation (Fig. 6c) is indeed observed for higher wind speeds (not shown). The same is confirmed by wind-speed measurements at three locations along the slope during the  $H_{ss\_T2}$  campaign. Downslope flow with higher wind speeds was observed to separate from the slope, i.e. the wind speed at the top-most site was higher than the wind speed at the middle site located on a flatter part of the terrain (not shown).

To test this hypothesis within a theoretical framework we examine the conditions that lead to flow separation according to the concept of bluff-body separation (Baines 1995; cf. the regime diagram in his Fig. 5.8). The bluff-body separation is the closest analogy to our undulating slope. Whether flow separates from the slope depends on the non-linearity of the flow ( $Nh/U$ ) and steepness of the terrain ( $h/A$ ). Here we use the height ( $h \approx 2.6 \text{ m}$ ) and half-width ( $A \approx 3 \text{ m}$ ) of individual undulations in the slope and the locally determined Brunt–Väisälä frequency ( $N$ ) and mean wind speed ( $U$ ) to position ourselves on the regime diagram shown in Fig. 8. Given the lack of temperature measurements at multiple levels for determining  $N$  during the 11 months of measurements, we employ stable boundary-layer similarity theory (Stull 1988) and compute  $N$  from

$$N = \sqrt{\frac{g}{\bar{\theta}} \frac{\theta_*}{kz} \left( 0.74 + 4.7 \frac{z}{L} \right)}, \tag{9}$$

where  $k$  is the von Karman constant,  $g$  is the acceleration due to gravity,  $\bar{\theta}$  is the mean potential temperature and  $\theta_*$  is the surface-layer temperature scale. The results for 11 months of data at



**Fig. 8** Regime diagram for lee-side flow separation according to Baines (1995), (part of) his Fig. 5.8.  $N$  is the Brunt–Väisälä frequency,  $U$  is the mean wind speed,  $h$  is the relative height of the undulations in the slope,  $A$  is the half width of the undulations. *Black line* as in Baines (1995) dividing the regimes with and without flow separation. Data shown are 11 months at site  $H_{ss}$  for downslope stable conditions and HiQ quality criteria. *Black circles* all 30-min intervals for any wind speed, *purple circles* intervals with mean wind speed  $U > 2 \text{ m s}^{-2}$ . *Small pink dots* data from  $H_{ss\_T2}$  (spanning 5 days) when flow separation was observed along the  $H_{ss}$  slope

site  $H_{ss}$  for stable downslope flow (determined by  $H < 0$  and the appropriate wind direction) are shown in Fig. 8. We see that the flow spans both the regime for which flow-separation is possible and that for which the flow does not separate from the surface. Data with higher wind speeds ( $U > 2 \text{ m s}^{-1}$ ), however, are all contained in the flow-separation part of the diagram, as well as the data from the  $H_{ss\_T2}$  campaign when flow separation was actually observed. Thus our results confirm that higher wind speeds lead to flow separation whereas the lower do not. This behaviour is impossible to capture with the usual planar-fit coordinate rotation.

These results thus suggest that, apart from direction-dependent planar-fit in complex mountainous terrain, a wind-speed dependent planar-fit is also needed in order to eliminate the influence of wind-speed dependent flow separation that might occur for non-ideal slopes. It also shows that the sensitivity to the rotation method, as well as the representativeness of the measurements itself (Medeiros and Fitzjarrald 2015), will be largely determined by the exact positioning of the station on the non-uniform slope.

## 5 Finding the Optimal Post-processing Procedure

In the previous sections we have highlighted several decisions that need to be made when analyzing turbulence data in mountainous terrain. Which method (e.g., averaging, coordinate rotation) is the best, however, remains elusive due to the lack of objective criteria and to date has been based only on theoretical considerations (e.g., Finnigan et al. 2003). Most authors generally accept planar-fit as the preferred rotation method for turbulence analysis in complex terrain (e.g. Finnigan et al. 2003; Sun 2007; Mauder et al. 2013; Nadeau et al. 2013). Sensitivity tests, however, are scarce and usually focus only on one, preferably ideal, site (e.g., Turnipseed et al. 2003; Kral et al. 2014).

Here we attempt to quantify the effect of the aforementioned choices and to objectively determine the best practice for measurements in complex terrain by testing local similarity on



the example of the vertical velocity variance (Eq. 4) for the two mountainous sites of different slope angle and terrain complexity (sites  $H_{ss}$  and  $W_{ms}$ ). The hypothesis is that the post-processing combination with the smallest scatter best supports a local similarity approach, whilst less appropriate post-processing adds additional uncertainty and thus scatter.

For the present purpose the data processed with different processing options were first fitted to Eq. 4 using robust regression, with parameters  $a$ ,  $b$  and  $c$  left free (cf. Fig. 3c, d). Robust regression results were tested for different values of the first guess of  $a$ ,  $b$  and  $c$ , in order to cover the relevant part of parameter space (e.g., the first guess of parameter  $b$  varied between 1 and 5, and parameter  $c$  between 0.1 and 0.5), since the results proved to be sensitive to the choice of parameter first guesses. The scatter of data from the best-fit curve was determined by computing the mean absolute deviation (*MAD*) due to its insensitivity to outliers. The combination of first guess parameters that produced the smallest *MAD* value was later used as the final *MAD* value of the particular post-processing method under examination.

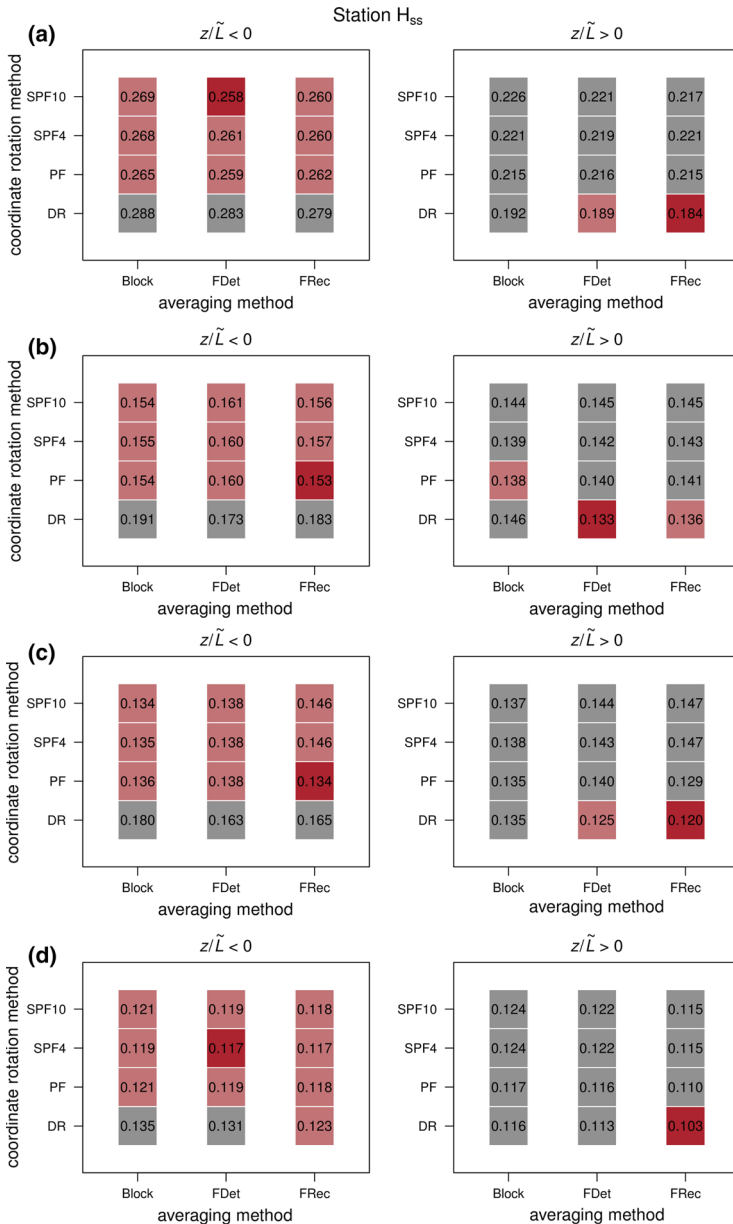
We test the following combinations of post-processing choices for the two i-Box sites,

1. Filtering: no filtering (i.e., block average, denoted Block in the following), linear detrending (FDet) and high-pass recursive filtering with filter length of 200 s (FRec).
2. Coordinate rotation: double rotation vs. planar-fit with a single plane, four planes and ten planes.
3. Data quality control: all data that pass the minimum quality control vs. the best data that at the same time fulfil stationarity requirements for all variances and covariances (limit of 30 %) and data uncertainty of variances and covariances is below 50 % (HiQ).
4. Averaging period: 30 vs. 60 min.

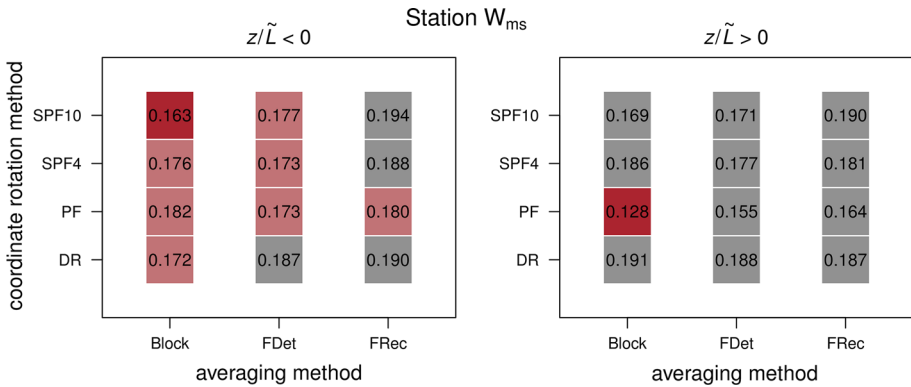
Figures 9 and 10 show *MAD* values from the robust best fit to Eq. 4 for different post-processing combinations, in the stable ( $z/\bar{L} > 0$ ) and unstable ( $z/\bar{L} < 0$ ) regimes, respectively. The post-processing option with the smallest *MAD* values for the given stability, site and quality criteria (i.e., the best method), is shown in red. The post-processing options in pink are those that are not statistically significantly different (according to the student t-test with a 0.1-confidence level) from the best method, whilst the options in grey are the options that are statistically significantly different.

Figure 9 shows the results for site  $H_{ss}$  obtained for data with minimum quality and 30-min averaging period (Fig. 9a), HiQ quality and 30-min averaging period (Fig. 9b) and HiQ quality and 60-min averaging period (Fig. 9c) corresponding to Fig. 3a, c, d, respectively. The results show that the scatter around the best-fit curve (i.e. *MAD*) is significantly reduced for data with HiQ criteria compared to those that only satisfy the minimum quality, as also visible when comparing Figs. 3a, c, d. Thus we reiterate the need to examine only data of the highest quality, in line with the results of previous sections. A 60-min averaging time further reduces *MAD* values while still retaining full coverage of the stability range, both in stable and unstable regimes (cf. Fig. 3d). It thus appears that a longer averaging time would be preferred in complex terrain over the more commonly used 30 min or even shorter averaging times (cf. Martins et al. 2009; Nadeau et al. 2013).

As for the combination of post-processing options, in the stable regime for site  $H_{ss}$ , the combination of filtering and double rotation consistently emerges as the best method for all quality criteria and averaging times and is statistically significantly better than different planar-fit options. These results are in line with those obtained in Sect. 4.2 and can be attributed to slope complexity and flow separation at site  $H_{ss}$ , which cannot be reproduced by fitting one or more direction-dependent climatological planes (not even 10). Thus horizontal and vertical wind components are inappropriately differentiated, which leads to larger scatter in the scaling relations. On the unstable side, however, double rotation is the worst method,



**Fig. 9** Mean absolute deviation (*MAD*) as a measure of scatter of  $\sigma_w/\tilde{u}_*$  around the best-fit curve defined by Eq. 4 obtained by robust regression. Data base: 11 months of data from  $H_{ss}$  that fulfil **a** the minimum quality criterion (see Sect. 3) and corresponds to Fig. 3a. **b** HiQ quality criterion (see Sects. 3 for definition) with 30-min averaging time (cf. Fig. 3c). **c** HiQ quality criterion with 60-min averaging time (cf. Fig. 3d). **d** HiQ quality criterion simultaneously satisfied by all the methods and 60-min averaging time. *Left panels* unstable, *right panels* stable stratification. The best method (smallest *MAD*) shown in red, methods that are not statistically significantly different from the best method (student *t* test for **a–c** and pair-wise *t* test for **d**; 0.1-confidence limit) shown in pink; methods statistically significantly different from the best shown in grey. Average number of data points (stable/unstable panels, respectively): **a** 5500/2800, **b** 2930/1020, **c** 1480/490, **d** 930/230)



**Fig. 10** As in Fig. 9c but for site  $W_{ms}$ . Number of data points used in the best fit calculation are 1350 (stable) and 430 (unstable)

whilst different planar-fit options perform similarly. This is due to low wind speeds in this sector and wind-direction changes associated with them that aggravate the determination of rotation angles in the double rotation method.

Given that different numbers of data satisfy the HiQ criteria for each post-processing method (e.g., filtering increases the stationarity of the time series so less data is eliminated by this criterion) and that filtering reduces the random error, the comparison between the different methods could be considered statistically unfair. In order to make a fair comparison we would therefore need to examine only those data points that simultaneously pass the HiQ quality criteria for all the post-processing methods. The results are presented in Fig. 9d for a 60-min averaging time. This quality criterion is extremely stringent and only a fraction of data (less than 2 % in the unstable range for site  $H_{ss}$  and less than 0.5 % for site  $W_{ms}$ ) passes it. Now we can use a pair-wise  $t$  test to examine the statistical significance of the differences between the methods. The large reduction of data points leads to the results that are slightly different to Fig. 9c. Still, the general conclusions hold and double rotation and filtering is still the best method on the stable side, whereas on the unstable side there is no preferred best method, but double rotation still appears to be the worst.

Figure 10 shows the results for site  $W_{ms}$  for HiQ quality and 60-min averaging interval, corresponding to Fig. 9c. The results are opposite to those of site  $H_{ss}$ : block-averaged results produce the smallest MAD values in both stability regimes. In the stable regime planar-fit is the single best method, whereas on the unstable side there is no statistically significant difference between the methods as long as block averaging is applied.

Although applying a filter to eliminate low frequency non-turbulent motions does reduce the scatter around the best fit curve for stable stratification at site  $H_{ss}$ , it appears to have no systematic advantage elsewhere. Additionally, the rotation method has a stronger (statistically significant) influence on the scatter than the averaging (filtering) method.

Overall, we conclude that demanding the highest quality in the data (HiQ) is a must when studying turbulence statistics in mountainous terrain, and that averaging interval longer than 30 min appears to have advantages. At the same time, we found no best combination of post-processing options for all sites and stabilities. Whilst there is a clear best rotation method for the stable stratification at site  $H_{ss}$  the same does not apply for a different site. Even the best method for stable stratification at a given site might not be the best for unstable stratification. The planar-fit method, most often used in studies in non-ideal terrain (e.g.

Mauder et al. 2013; Nadeau et al. 2013; Kral et al. 2014), appears to be on average superior to double rotation, but only for relatively uniform slopes. Interestingly, choosing four sectors according to dominant wind directions (cf. Nadeau et al. 2013) and site complexity does not appear to have a systematically beneficial effect, neither does choosing 10 sectors. For more undulating, concave or convex slopes, double rotation appears to still outperform any planar-fit option or a wind-speed dependent planar-fit might be necessary. Additionally, we have shown the need for extremely long observation periods (spanning significantly more than a year) in order to have a statistically sound amount of high quality data for analysis.

The increase of  $\sigma_w/\tilde{u}_*$  with increasing  $z/\tilde{L}$  in the stable boundary layer might be attributed to self-correlation (e.g. Klipp and Mahrt 2004; Grachev et al. 2013). In order to test the sensitivity of our results to self-correlation we have attempted to avoid it by substituting  $\tilde{u}_*$  with  $\sigma_u$  in the definition of  $\tilde{L}$  (according to Grachev, personal communication, 2015). The new results now show a very weak dependence on  $z/\tilde{L}$  for stable conditions suggestive of  $z$ -less scaling. However, the scatter around the best-fit curve (i.e., calculated *MAD* values) does not change so that the conclusions on the most appropriate method drawn from it are not altered.

## 6 Discussion and Conclusions

We have attempted to demonstrate that data quality as well as specific post-processing choices play an important role when assessing turbulence characteristics in complex mountainous terrain, more so than for ideal sites over horizontally homogeneous and flat (HHF) surfaces. Similarly Grachev et al. (2012, 2013) showed that the method of data analysis and the prerequisites imposed on the data (e.g. excluding data for the Richardson number below a certain threshold) have a significant impact on the conclusions drawn from turbulence statistics. As a prerequisite we have shown that turbulence integral statistics must be corrected for the known errors (see Sect. 3), as over HHF surfaces, and that these corrections can be quite important.

Over non-ideal mountainous sites, data quality requires special attention. Statistical uncertainty in turbulence integral statistics, assessed through Eqs. 5–7 in the present study, can be substantial. Overall, data quality requirements seem to render it difficult to obtain statistically significant results from short field campaigns (duration of a few weeks or even months) and call for long-term observational sites with datasets spanning more than a year, when using traditional EC analysis techniques. Even with a long-term observational program (11 months of data in the present study), a 20 % statistical uncertainty threshold (as might be appropriate for HHF sites) proves too stringent in order to cover the full stability range. A 50 % uncertainty threshold combined with stationarity tests (usual threshold) appears to be a viable compromise.

The choice of the post-processing approach and especially the choice of the coordinate system in mountainous terrain has a significant impact on the results: Sect. 4 shows that there are substantial differences in sensible heat fluxes between a slope-normal and a vertical frame of reference, with the slope-normal approach being more physically sound (at least close to the surface). The question then arises how to define this frame of reference, i.e., through the double rotation or different types of planar-fit approaches. To answer this we first examined the magnitudes of heat and momentum fluxes obtained by both methods and found that the results are site dependent. On average, planar-fit methods (both single plane or sectorial planar-fit) underestimate the magnitude of the fluxes compared to double rotation.

The difference can be attributed to different characteristics of slope flows, which are not taken into account in planar-fit. For complex slope configurations (e.g., undulations) a wind-speed dependent planar-fit would be required to account for the occurrence of flow separation and the fact that flow of different speed follows a different plane.

Given that the above analysis of the best method is not entirely objective, we have furthermore used the scatter around a best-fit curve through the data as a criterion for the appropriateness of the post-processing approach. In other words we have tested the hypothesis that if one particular post-processing approach (coordinate rotation approach plus averaging method) yielded improved local scaling results (smaller scatter), this would be an argument in favour of this post-processing choice. The results, first of all, confirm the importance of data quality in that quite generally high quality data yield smaller scatter for local scaling characteristics. The same is true for longer averaging periods (60 vs. 30 min), thus confirming the conclusions of [Finnigan et al. \(2003\)](#) for less complex (but still non-ideal) conditions.

As for the post-processing approach itself the results are less clear. Our analysis does not unequivocally single out one combination of post-processing options as the best to be used for any site in complex mountainous terrain, but points to site and stability dependence. The most appropriate choice for stable stratification is clear for both sites, although it is not the same. This could be attributed to the more robust statistics (larger data sets) for stable flow, due to the northerly orientation of both sites. In earlier studies it has often been argued that planar-fit would be the more reliable rotation approach in non-ideal terrain (e.g., [Finnigan 2004](#)). Our results, on the basis of the scatter from local scaling, at least demonstrate that this is not generally the case for complex mountainous terrain. Our results for two sites show that planar-fit and double rotation have equal likelihood to be the best or the worst method.

In general, the decision on (sectorial) planar-fit vs. double rotation, and whether to filter or not to filter, depends of course, on one's goal with the turbulence measurements. [Finnigan \(2004\)](#)'s mathematically substantiated arguments and the results of [Sun \(2007\)](#) in favour of planar-fit refer to the optimal (robust) measurement of surface mass balance (within the FLUXNET community, measuring CO<sub>2</sub> exchange at many sites around the globe). In the present case, this is not substantiated at our sites and planar-fit methods generally underestimate the total flux leading to additional energy balance under-closure. Similarly, inappropriate filtering might underestimate the total flux, especially for unstable stratification. However, if the local response of turbulence statistics to the flow conditions is used as a criterion (as would be appropriate when testing MOST in complex terrain, e.g. [Martins et al. 2009](#); [Nadeau et al. 2013](#)), our analysis shows that the planar-fit approach, using a coordinate system with a climatological reference, may better reflect the response of the system to the momentary forcing than does a local (and instantaneous) coordinate system such as double rotation. This latter finding is, however, only valid for the uniform slope in our sample.

In summary we have demonstrated the importance of data quality (flux corrections, assessment of statistical uncertainty) and duration of the averaging period on the measurement of integral turbulence statistics in mountainous terrain. As for post-processing, the coordinate rotation is more important than the data filtering but no single best approach could be found for all sites and conditions. Double rotation coordinate rotation yields larger fluxes (magnitude) than planar-fit at both the relatively uniform site and that with small-scale undulations. Based on this criterion double rotation should be preferred for budgeting purposes. When turbulence characteristics are assessed—as in our local scaling example where fluxes and variances are combined—the best post-processing approach clearly depends on the local terrain characteristics, thus calling for a careful assessment of flow conditions for different stabilities at each site. Data from other i-Box (or elsewhere) sites, and possibly using other local scaling relations, may give rise to a further generalization of these findings. The degree

to which an ‘advection correction’, as proposed by Paw et al. (2000), can alleviate some of the remaining uncertainties in complex mountainous terrain remains to be investigated.

Proper site selection in mountainous terrain simplifies the elaborate analysis needed for finding the most appropriate post-processing method. However, ideal sites might not give the most representative results for turbulence characteristics in mountainous terrain, where ideal conditions are rather the exception.

**Acknowledgments** We would like to thank A. Grachev and an anonymous reviewer for valuable comments. Special thanks go to Reto Stauffer, Markus Dabernig and Manuel Presser of the University of Innsbruck for their valuable advice and help in statistical analysis in Sect. 5. We are happy to acknowledge the contributions of students in the analysis of the turbulence data: René Schubert (University of Kiel) initially helped with part of the statistics leading to Fig. 2, and Masters’ students of the Alpine Meteorological Field course are thanked for their contribution to field campaign measurements that led to Fig. 5 and Sect. 4. The i-Box infrastructure has been established through funding of the University of Innsbruck. All data used herein are available from the authors.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Acevedo OC, Moraes OLL, Degrazia GA, Medeiros LE (2006) Intermittency and the exchange of scalars in the nocturnal surface layer. *Boundary-Layer Meteorol* 119:41–55
- Aubinet M, Vesala T, Papale D (2012) Eddy covariance. A practical guide to measurements and data analysis, Springer, Dordrecht 438 pp
- Baines PG (1995) Topographic effects in stratified flows. Cambridge University Press, Cambridge 482 pp
- Baldocchi D, Finnigan J, Wilson K, Paw UKT, Falge E (2000) On measuring net ecosystem carbon exchange over tall vegetation on complex terrain. *Boundary-Layer Meteorol* 96:257–291. doi:[10.1023/A:1002497616547](https://doi.org/10.1023/A:1002497616547)
- Barford CC, Wofsy SC, Goulden ML, Munger JW, Pyle EH, Urbanski SP, Hutrya L, Saleska SR, Fitzjarrald D, Moore K (2001) Factors controlling long- and short-term sequestration of atmospheric CO<sub>2</sub> in a mid-latitude forest. *Science* 294:1688–1691
- Barry RG (2008) Mountain weather and climate. Cambridge University Press, New York 506 pp
- Bendat JS, Piersol AG (1986) Random data. Analysis and measurement procedures, Wiley Interscience, New York 640 pp
- Billesbach DP (2011) Estimating uncertainties in individual eddy covariance flux measurements: a comparison of methods and a proposed new method. *Agric For Meteorol* 151:394–405
- Brugger H (2011) Spatial variability of turbulence statistics in complex terrain: a study based on the MAP Riviera project. Thesis, Institute of Meteorol and Geophys, University of Innsbruck, Innsbruck, M.sc
- Clement R, Moncreif J (2007) EdiRe Software, University of Edinburgh. <http://www.geos.ed.ac.uk/abs/research/micromet/EdiRe>
- de Franceschi M, Zardi D, Tagliazuca M, Tampieri F (2009) Analysis of second-order moments in surface layer turbulence in an Alpine valley. *Q J R Meteorol Soc* 135:1750–1765. doi:[10.1002/qj.506](https://doi.org/10.1002/qj.506)
- Finkelstein PL, Sims PF (2001) Sampling error in eddy correlation flux measurements. *J Geophys Res* 106:3503–3509
- Finnigan JJ, Clement R, Malhi Y, Leuning R, Cleugh HA (2003) A re-evaluation of long-term flux measurement techniques. Part I: Averaging and coordinate rotation. *Boundary-Layer Meteorol* 107:1–48
- Finnigan JJ (2004) A re-evaluation of long-term flux measurement techniques. Part II: Coordinate systems. *Boundary-Layer Meteorol* 113:1–41. doi:[10.1023/B:BOUN.0000037348.64252.45](https://doi.org/10.1023/B:BOUN.0000037348.64252.45)
- Finnigan JJ (2008) An introduction to flux measurements in difficult conditions. *Ecol Appl* 18:1340–1350. doi:[10.1890/07-2105.1](https://doi.org/10.1890/07-2105.1)
- Foken T (2008) *Micrometeorology*. Springer, Berlin 308 pp
- Foken T, Wichura B (1996) Tools for quality assessment of surface-based flux measurements. *Agric For Meteorol* 78:83–105. doi:[10.1016/0168-1923\(95\)02248-1](https://doi.org/10.1016/0168-1923(95)02248-1)

- Foken T, Leuning R, Oncley SR, Mauder M, Aubinet M (2012) Corrections and data quality control. In: Aubinet M et al (eds) Eddy covariance. A practical guide to measurements and data analysis, Springer, Dordrecht, pp 85–132
- Forrer J, Rotach MW (1997) On the turbulence structure in the stable boundary layer over the Greenland ice sheet. *Boundary-Layer Meteorol* 85:111–136. doi:[10.1023/A:1000466827210](https://doi.org/10.1023/A:1000466827210)
- Geissbühler P, Siegwolf R, Eugster W (2000) Eddy covariance measurements on mountain slopes: the advantage of surface-normal sensor orientation over a vertical set-up. *Boundary-Layer Meteorol* 96:371–392
- Goulden MJ, Munger JW, Fan SM, Daube BC, Wofsy SC (1996) Measurement of carbon sequestration by long-term eddy covariance: methods and critical evaluation of accuracy. *Glob Change Biol* 22:169–182
- Grachev AA, Andreas EL, Fairall CW, Guest PS, Persson POG (2012) Outlier problem in evaluating similarity functions in the stable atmospheric boundary layer. *Boundary-Layer Meteorol* 144:137–155. doi:[10.1007/s10546-012-9714-9](https://doi.org/10.1007/s10546-012-9714-9)
- Grachev AA, Andreas EL, Fairall CW, Guest PS, Persson POG (2013) The critical Richardson number and limits of applicability of local similarity theory in the stable boundary layer. *Boundary-Layer Meteorol* 147:51–82. doi:[10.1007/s10546-012-9771-0](https://doi.org/10.1007/s10546-012-9771-0)
- Hammerle A, Haslwanter A (2007) Eddy covariance measurements of carbon dioxide, latent and sensible energy fluxes above a meadow on a mountain slope. *Boundary-Layer Meteorol* 122:397–416. doi:[10.1007/s10546-006-9109-x](https://doi.org/10.1007/s10546-006-9109-x)
- Hiller R, Zeeman MJ, Eugster W (2008) Eddy-covariance flux measurements in complex terrain of an alpine valley in Switzerland. *Boundary-Layer Meteorol* 127:449–467. doi:[10.1007/s10546-008-9267-0](https://doi.org/10.1007/s10546-008-9267-0)
- Hocut CM, Liberzon D, Fernando HJS (2015) Separation of upslope flow over a uniform slope. *J Fluid Mech* 775:266–287. doi:[10.1017/jfm.2015.298](https://doi.org/10.1017/jfm.2015.298)
- Horst TW, Semmer SR, Maclean G (2015) Correction of a non-orthogonal, three-component sonic anemometer for flow distortion by transducer shadowing. *Boundary-Layer Meteorol* 155:371–395
- Kaimal JC, Finnigan JJ (1994) Atmospheric boundary layer flows: their structure and measurement. Oxford University Press, Oxford 304 pp
- Klipp CL, Mahrt L (2004) Flux-gradient relationship, self-correlation and intermittency in the stable boundary layer. *Q J R Meteorol Soc* 130:2087–2103. doi:[10.1256/qj.03.161](https://doi.org/10.1256/qj.03.161)
- Kljun N, Rotach MW, Schmid HP (2002) A three-dimensional backward lagrangian footprint model for a wide range of boundary layer stratifications. *Boundary-Layer Meteorol* 103:205–226. doi:[10.1023/A:1014556300021](https://doi.org/10.1023/A:1014556300021)
- Kljun N, Calanca P, Rotach MW, Schmid HP (2004) A simple parameterisation for flux footprint predictions. *Boundary-Layer Meteorol* 112:503–523. doi:[10.1023/B:BOUN.0000030653.71031.96](https://doi.org/10.1023/B:BOUN.0000030653.71031.96)
- Kral ST, Sjöblom A, Nygård T (2014) Observations of summer turbulent surface fluxes in a High Arctic fjord. *Q J R Meteorol Soc* 140:666–675. doi:[10.1002/qj.2167](https://doi.org/10.1002/qj.2167)
- Leuning R, van Gorsel E, Massman WJ, Isaac PR (2012) Reflections on the surface energy imbalance problem. *Agric For Meteorol* 156:65–74
- Lothon M, Druilhet A, Bénech B, Campistron B, Bernard S, Saïd F (2003) Experimental study of five föhn events during the Mesoscale Alpine Programme: from synoptic scale to turbulence. *Q J R Meteorol Soc* 129:2171–2193. doi:[10.1256/qj.02.30](https://doi.org/10.1256/qj.02.30)
- Mahrt L (1998) Flux sampling errors for aircraft and tower. *J Atmos Ocean Technol* 15:416–429
- Mahrt L (2010) Computing turbulent fluxes near the surface: needed improvements. *Agric For Meteorol* 150:501–509. doi:[10.1016/j.agrformet.2010.01.015](https://doi.org/10.1016/j.agrformet.2010.01.015)
- Martins CA, Moraes OLL, Acevedo OC, Degrazia GA (2009) Turbulence intensity parameters over a very complex terrain. *Boundary-Layer Meteorol* 133:35–45. doi:[10.1007/s10546-009-9413-3](https://doi.org/10.1007/s10546-009-9413-3)
- Mauder M, Cuntz M, Drüe C, Graf A, Rebmann C, Schmid H, Schmidt M, Steinbrecher R (2013) A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements. *Agric For Meteorol* 169:122–135. doi:[10.1016/j.agrformet.2012.09.006](https://doi.org/10.1016/j.agrformet.2012.09.006)
- Matzinger N, Andretta M, van Gorsel E, Vogt R, Ohmura A, Rotach MW (2003) Surface radiation budget in an alpine valley. *Q J R Meteorol Soc* 129:877–895. doi:[10.1256/qj.02.44](https://doi.org/10.1256/qj.02.44)
- McMillen RT (1988) An eddy correlation technique with extended applicability to non-simple terrain. *Boundary-Layer Meteorol* 43:231–245. doi:[10.1007/BF00128405](https://doi.org/10.1007/BF00128405)
- Medeiros LE, Fitzjarrald DR (2015) Stable boundary layer in complex terrain. Part II: Geometrical and sheltering effects on mixing. *J Appl Meteorol Climatol* 54:170–188
- Moore CJ (1986) Frequency response correction for eddy correlation systems. *Boundary-Layer Meteorol* 37:17–35. doi:[10.1007/BF00122754](https://doi.org/10.1007/BF00122754)
- Monin A, Obukhov AM (1954) Basic laws of turbulent mixing in the ground layer of the atmosphere. *Tr Akad Nauk SSSR* 151:163–187
- Nadeau DF, Pardyjak ER, Higgins CW, Parlange MB (2013) Similarity scaling over a steep alpine slope. *Boundary-Layer Meteorol* 147:401–419. doi:[10.1007/s10546-012-9787-5](https://doi.org/10.1007/s10546-012-9787-5)

- Nicholson LI, Prinz R, Mölg T, Kaser G (2013) Micrometeorological conditions and surface mass and energy fluxes on Lewis Glacier, Mt Kenya, in relation to other tropical glaciers. *Cryosphere* 7:1205–1225. doi:[10.5194/tc-7-1205-2013](https://doi.org/10.5194/tc-7-1205-2013)
- Ono K, Mano M, Miyata A, Inoue Y (2008) Applicability of the planar fit technique in estimating surface fluxes over flat terrain using eddy covariance. *J Agric Meteorol* 64:121–130
- Panofsky HA, Dutton JA (1984) Atmospheric turbulence. Models and methods for engineering applications, Wiley, New York 389 pp
- Park MS, Park SO (2006) Effects of topographical slope angle and atmospheric stratification on surface layer turbulence. *Boundary-Layer Meteorol* 118:613–633. doi:[10.1007/s10546-005-7206-x](https://doi.org/10.1007/s10546-005-7206-x)
- Paw UKT, Baldocchi DD, Meyers TP, Wilson KB (2000) Correction of eddy-covariance measurements incorporating both advective effects and density fluxes. *Boundary-Layer Meteorol* 97:487–511. doi:[10.1023/A:1002786702909](https://doi.org/10.1023/A:1002786702909)
- Rannik Ü, Vesala T (1999) Autoregressive filtering versus linear detrending in estimation of fluxes by the eddy covariance method. *Boundary-Layer Meteorol* 91:259–280
- Rebmann C, Kolle O, Heinesch B, Queck R, Ibrom A, Aubinet M (2012) Data acquisition and flux calculations. In: Aubinet M et al (eds) Eddy covariance. A practical guide to measurements and data analysis, Springer, Dordrecht, pp 59–83
- Richardson AD, Aubinet M, Barr AG, Hollinger DY, Ibrom A, Lasslop G, Reichstein M (2012) Uncertainty quantification. In: Aubinet M et al (eds) Eddy covariance. A practical guide to measurements and data analysis, Springer, Dordrecht, pp 173–210
- Rotach MW, Zardi D (2007) On the boundary-layer structure over highly complex terrain: key findings from MAP. *Q J R Meteorol Soc* 133:937–948. doi:[10.1002/qj.71](https://doi.org/10.1002/qj.71)
- Rotach MW, Calanca P, Graziani G, Gurtz J, Steyn DG, Vogt R, Andretta M, Christen A, Cieslik S, Connolly R, De Wekker SFJ, Galmarini S, Kadyrov EN, Kadyrov V, Miller E, Neininger B, Rucker M, van Gorsel E, Weber H, Weiss A, Zappa M (2004) Turbulence structure and exchange processes in an alpine valley: The Riviera project. *Bull Am Meteorol Soc* 85:1367–1385. doi:[10.1175/BAMS-85-9-1367](https://doi.org/10.1175/BAMS-85-9-1367)
- Rotach MW, Andretta M, Calanca P, Weigel AP, Weiss A (2008) Boundary layer characteristics and turbulent exchange in highly complex terrain. *Acta Geophys* 56:194–219
- Rotach MW, Wohlfahrt G, Hansel A, Reif M, Wagner J, Gohm A (2014) The world is not flat: implications for the global carbon balance. *Bull Am Meteorol Soc* 95:1021–1028
- Sanz Rodrigo J, Anderson PS (2013) Investigation of the stable atmospheric boundary layer at Halley Antarctica. *Boundary-Layer Meteorol* 148:517–539. doi:[10.1007/s10546-013-9831-0](https://doi.org/10.1007/s10546-013-9831-0)
- Schotanus P, Nieuwstadt FTM, de Bruin HAR (1983) Temperature measurement with a sonic anemometer and its application to heat and moisture fluxes. *Boundary-Layer Meteorol* 26:81–93. doi:[10.1007/BF00164332](https://doi.org/10.1007/BF00164332)
- Stiperski I, Rotach MW, Gohm A (2012) Boundary layer measurements in complex terrain: Innsbruck-Box. Session AS2.1, Air-Land Interactions, Geophys Res Abstracts, EGU2012-8863, 2012
- Stull RB (1988) An introduction to boundary layer meteorology. Kluwer Acad Pup, Dordrecht 666 pp
- Sun J (2007) Tilt corrections over complex terrain and their implication for CO<sub>2</sub> transport. *Boundary-Layer Meteorol* 124:143–159. doi:[10.1007/s10546-007-9186-5](https://doi.org/10.1007/s10546-007-9186-5)
- Sun J, Oncley SP, Burns SP, Stephens BB, Lenschow DH, Campos T, Watt AS, Monson RK, Moore SJP, Hu J, Tschudi M, Schimel DS, Aulenbach S, Sacks WJ, De Wekker SFJ, Lai CT, Lamb B, Allwine E, Coons T, Ojima D, Ellsworth PZ, Sternberg LSL, Zhong S, Clements C, Anderson DE (2010) A multiscale and multidisciplinary investigation of ecosystem-atmosphere CO<sub>2</sub> exchange over the Rocky Mountains of Colorado. *Bull Am Meteorol Soc* 91:209–230
- Szintai B, Kaufmann P, Rotach MW (2009) Deriving turbulence characteristics from the COSMO numerical weather prediction model for dispersion applications. *Adv Sci Res* 3:79–84. doi:[10.5194/asr-3-79-2009](https://doi.org/10.5194/asr-3-79-2009)
- Taylor PA, Mason PJ, Bradley EF (1987) Boundary layer flow over low hills. *Boundary-Layer Meteorol* 39:107–132
- Turnipseed AA, Blanken PD, Anderson DE, Monson RK (2002) Surface energy balance above a high-elevation subalpine forest. *Agric For Meteorol* 110:177–201
- Turnipseed AA, Anderson DE, Blanken PD, Baugh W, Monson RK (2003) Airflows and turbulent flux measurements in mountainous terrain. Part 1: Canopy and local effects. *Agric For Meteorol* 119:1–21
- Turnipseed AA, Anderson DE, Burns SP, Blanken PD, Monson RK (2004) Airflows and turbulent flux measurements in mountainous terrain. Part 2: Mesoscale effects. *Agric For Meteorol* 125:187–205. doi:[10.1016/j.agrformet.2004.04.007](https://doi.org/10.1016/j.agrformet.2004.04.007)
- van Dijk A, Kohsiek W, de Bruin HAR (2003) Oxygen sensitivity of krypton and Lyman-alpha hygrometers. *J Atmos Ocean Technol* 20:143–151. doi:[10.1175/1520-0426\(2003\)0202.0.CO;2](https://doi.org/10.1175/1520-0426(2003)0202.0.CO;2)
- Večenaj Ž, De Wekker SFJ (2014) Determination of nonstationarity in the surface layer during the T-REX experiment. *Q J R Meteorol Soc*. doi:[10.1002/qj.2458](https://doi.org/10.1002/qj.2458)



- Vickers D, Mahrt L (1997) Quality control and flux sampling problems for tower and aircraft data. *J Atmos Ocean Technol* 14:512–526. doi:[10.1175/1520-0426\(1997\)014<0512:QCAFSP>2.0.CO;2](https://doi.org/10.1175/1520-0426(1997)014<0512:QCAFSP>2.0.CO;2)
- Vickers D, Mahrt L (2003) The cospectral gap and turbulent flux calculations. *J Atmos Ocean Technol* 20:660–672. doi:[10.1175/1520-0426\(2003\)20<660:TCGATF>2.0.CO;2](https://doi.org/10.1175/1520-0426(2003)20<660:TCGATF>2.0.CO;2)
- Vickers D, Göckede M, Law BW (2010) Uncertainty estimates for 1-h averaged turbulence fluxes of carbon dioxide, latent heat and sensible heat. *Tellus B* 62:87–99. doi:[10.1111/j.1600-0889.2009.00449.x](https://doi.org/10.1111/j.1600-0889.2009.00449.x)
- Webb EK, Pearman GI, Leuning R (1980) Correction of flux measurements for density effects due to heat and water vapour transfer. *Q J R Meteorol Soc* 106:85–100. doi:[10.1002/qj.49710644707](https://doi.org/10.1002/qj.49710644707)
- Wilczak JM, Oncley SP, Stage SA (2001) Sonic anemometer tilt correction algorithms. *Boundary-Layer Meteorol* 99:127–150. doi:[10.1023/A:1018966204465](https://doi.org/10.1023/A:1018966204465)
- Wyngaard JC (1973) On surface layer turbulence. In: Haugen DA (ed) *Workshop on micrometeorology*. American Meteorological Society, Boston, pp 101–150
- Yuan R, Kang M, Park SB, Hong J, Lee D, Kim J (2011) Expansion of the planar-fit method to estimate flux over complex terrain. *Meteorol Atmos Phys* 110:123–133. doi:[10.1007/s00703-010-0113-9](https://doi.org/10.1007/s00703-010-0113-9)
- Zappa M, Rotach MW, Arpagaus M, Dorninger M, Hegg C, Montani A, Ranzi R, Ament F, Germann U, Grossi G, Jaun S, Rossa A, Vogt S, Walser A, Wehrhan J, Wunram C (2008) MAP D-PHASE: real-time demonstration of hydrological ensemble prediction systems. *Atmos Sci Lett* 2:80–87. doi:[10.1002/asl.183](https://doi.org/10.1002/asl.183)