Master Thesis, Department of Geosciences

Applying the Eddy Covariance Method Under Difficult Conditions.

Lessons from the northernmost FLUXNET site: Bayelva, Svalbard (79°N).

Kristoffer Aalstad

UNIVERSITY OF OSLO

FACULTY OF MATHEMATICS AND NATURAL SCIENCES

Applying the Eddy Covariance Method Under Difficult Conditions.

Lessons from the northernmost FLUXNET site: Bayelva, Svalbard (79°N).

Kristoffer Aalstad

Master Thesis in Geosciences

Discipline: Meteorology and Oceanography

Department of Geosciences

Faculty of Mathematics and Natural Sciences

University of Oslo

01.06.2015

Abstract

We assess how reliable the Eddy-Covariance (EC) method is in estimating surface fluxes under the difficult conditions that occur in the high Arctic. Emphasis is placed on stable stratification and the breakdown of EC assumptions that may occur in such a regime.

To investigate these difficulties we developed an EC processing module from scratch, providing an extensive and transparent overview of the EC method. Raw data was obtained from an open path EC system located in the Bayelva catchment near Ny Ålesund (79°N), Svalbard, Norway.

Our flux estimates are in reasonable agreement with those found from the standardized EC package TK2. Strong relative non-stationarity represents the greatest hindrance to data quality at Bayelva, occurring for 11% of the data period. Overall, average relative flux uncertainties were found to be 20% for both the sensible (SH) and latent heat (LH) flux. Under stable stratification these uncertainties were considerably higher, 27% on average. Through Ogive classification we found that the traditional 30 minute SH and LH fluxes converged (resolved the turbulent cospectrum) 70% of the time. Here too the stable regime stands out, with low convergence fractions of 41% and 48% for LH and SH, respectively. To our knowledge it is the first time such an analysis has been carried out in the Arctic.

Concluding, while usually successful for neutral and unstable conditions, the traditional 30 minute flux averaging period is, more often than not, poorly suited for the stable regime. We attribute this to the observed and predicted shift in cospectral peaks towards lower periods under stable stratification, along with an erosion of the cospectral gap. An apparently simple fix of reducing the averaging period is not generally a valid solution. The required reduction could introduce unacceptable levels of flux uncertainty.

Acknowledgements

First and foremost I would like to thank my supervisor, Terje, for his guidance and feedback. Especially with regards to the structure of my thesis and the importance of highlighting the underlying motivation. I am particularly grateful that I was given the freedom to pursue this somewhat unusual, but fascinating topic. Next I wish to express a thanks to my unofficial co-supervisor, Sebastian, who provided me with the raw data, was always available to answer any questions I might have had about the Bayelva site and helped me out with a grant application. I am also grateful to Kjetil, who introduced me to the problem of surface energy balance closure almost two years back. I also appreciate the Cryomet group allowing me to tag along for a meeting at Finse during March/April of last year. In addition, I feel indebted to UNIS for holding a boundary layer meteorology course where I was introduced to the practice of eddy covariance.

I am also beholden to The Research Council of Norway that awarded me an Arctic Field Grant, through the Svalbard Science Forum, allowing me to conduct field work in the Bayelva catchment and ultimately see the eddy covariance system for myself. I am thankful to the amazing place that is Ny Ålesund, particularly Kings Bay, AWIPEV and the Sverdrup station for receiving me. As the trip in March this year would not have been possible without a field assistant, I am greatly indebted to Ole Henrik for joining me on such short notice. I also thank Kjersti and Bernd for helping me get access to instrumentation before travelling north.

Next I am grateful to Anne for helping me install DFFTPACK, without which this work in its current form would not have been possible. I would also like to express a special thanks to two of my co-students: Marta for helping me out with NetCDF and Eivind for many fruitful discussions over lunch. In addition I would like to thank all the other cohabitants of the MetOs study room, past and present, for providing a great social environment over the last two years. I would like to thank all the other professors that lectured me during my masters degree: Lars Petter, Joe and Jon Egil, not only for the courses they held, but for being generally available and open to answering all kinds of questions.

Last but not least I would like to thank my family and friends for their encouragement and putting up with me being somewhat of a recluse for the last couple of months. My sister, Elin, deserves a big thanks for proof reading my thesis and putting up with my 'wordiness'.

Contents

Chapter 1

Introduction

1.1 Prelude

The ensuing work presents an extensive and transparent analysis of raw high frequency data from a surface energy balance measurement campaign. Located in the Bayelva catchment near Ny Ålesund in the Svalbard archipelago, this campaign along with a complete description of the site is detailed in Westermann (2010). Our work is motivated by three widely recognized and related problems outlined in Section 1.3: the shortcomings of Monin-Obukhov Similarity Theory (MOST) in stable stratification (Mahrt, 1999), underestimation of nocturnal CO_2 respiration (Aubinet et al., 2012) and the lack of surface energy balance closure (Foken, 2008a). Common to all these problems is that they have been identified through observations in the turbulent atmospheric surface layer (ASL). Therefore, our focus is on the estimation of surface exchange based on turbulent fluxes in the ASL, particularly under difficult conditions that may occur under stable stratification (Mahrt, 1999).

Our attention is restricted to the Eddy-Covariance (EC) method, which with the onset of rapid response instrumentation has matured into the most widely used technique (Lee et al., 2006) for direct estimation of the exchange of constituents[1] between the atmosphere and the Earth's surface. The method is employed both continuously in long term global monitoring networks such as FLUXNET (Baldocchi et al., 2001) and as part of short term campaigns such as EBEX-2000 (Oncley et al., 2007). Direct estimates of surface exchange provided by the EC method are valuable beyond diagnosing local budgets for the exchange of water, carbon or energy. Estimates from FLUXNET are widely used by land surface schemes in climate models for: validation, constraining model output and developing parametrizations (Leuning et al., 2012). In addition, results from shorter EC campaigns have have been used to increase our understanding of the turbulent atmospheric surface layer, as in e.g. Kaimal et al. (1972), and provide evidence of the success of MOST (Monin and Obukhov, 1954) through the recovery of universal functions (Högström, 1996). These in turn provide first-order closure (Stull, 1988) such as the flux-profile relationships used in a variety of different models.

To date several well documented and widely used packages for processing eddy covariance data are available online. Some examples are TK2 (Mauder and Foken, 2004), as used by Westermann et al. (2009), the updated TK3 (Mauder and Foken, 2011), as used by Lüers et al. (2014), ECPACK (Van Dijk et al., 2004) and EddyPro package (Burba, 2013). These packages are typically developed and revised as 'in-house' tools at various institutions by micrometeorologists for research purposes (Mauder et al., 2013). In later years, however, due to many of these packages becoming open-access, it is increasingly non-micrometeorologists that have begun to employ the packages in conjunction with measurements as tools in process oriented studies. The canonical example of such studies has perhaps become the diagnosis of the net ecosystem exchange of CO_2 as described in e.g. Lee (1998). This is the case for a recent, albeit micrometeorological, publication concerning the Bayelva site found in Lüers et al. (2014).

Since our interest lies in the EC method itself we wish to have all the data handling open to us and minimize the use of any so-called black boxes[2], so as to ensure transparency. Moreover, the unusual location of the Bayelva site prompts a revision of a few of the typical approaches to EC data processing. Hence we opted for creating our own EC-processing package in the form of a module. This, in turn, allowed great freedom in customizing which steps to include in the data processing and analysis. In fact, which steps to include in processing varies significantly between investigators and poses a serious problem for intercomparisons in

[1] Be it momentum, energy, water vapor, or carbon dioxide (or other trace gases).

[2] A black box is a system which provides (not necessarily correct) output given user input, but what occurs in-between, inside the 'box', is hidden to the user.

the field of micrometeorology (see e.g. Mauder et al. (2008) and Fratini and Mauder (2014)). Accordingly particular weight is given to our Method; all the steps employed in our EC module are described and motivated in detail. For the most part we try to build on well-established routines for processing EC data as outlined in e.g. Aubinet et al. (2012). In addition, we take some unusual steps such as: estimating the uncertainty in flux estimates, calculating autostatistics and applying extensive cospectral analysis. Many of these steps are not included in other packages such as TK2. Furthermore, an extensive quality control procedure is implemented so as to be able to identify difficult conditions in which the EC method is strictly not applicable to the measurement of surface exchange (Foken and Wichura, 1996).

It will be shown that our particular study site provides both a challenge and an opportunity with regards to its location; due to the latitude, topography and land surface type. Moreover, the occurrence of very stable stratification during the polar night means that the winter half of the data set is of particular interest. Measurements in such a regime are few and far between; whether it be in the high Arctic or anywhere else for that matter. Applying the EC method in such a regime proves to be particularly difficult as many of the underlying assumptions may be violated (Mahrt, 2010). Our aim is to implement a methodology that allows us to identify such instances and the mechansisms that lead to their occurence.

Before proceeding with the method in Chapter 2 we introduce the theory on which the EC method is built. We also provide a brief overview of the Bayelva site and the instrumentation employed in the EC system. Furthermore, we introduce the three outstanding problems that provide the motivation for our work. The theory, and in particular the assumptions therein, provides the fuel for much of the discussion that follows in the remainder of this thesis.

1.2 Eddy Covariance

1.2.1 Theory

Fundamental Equation of Eddy Covariance

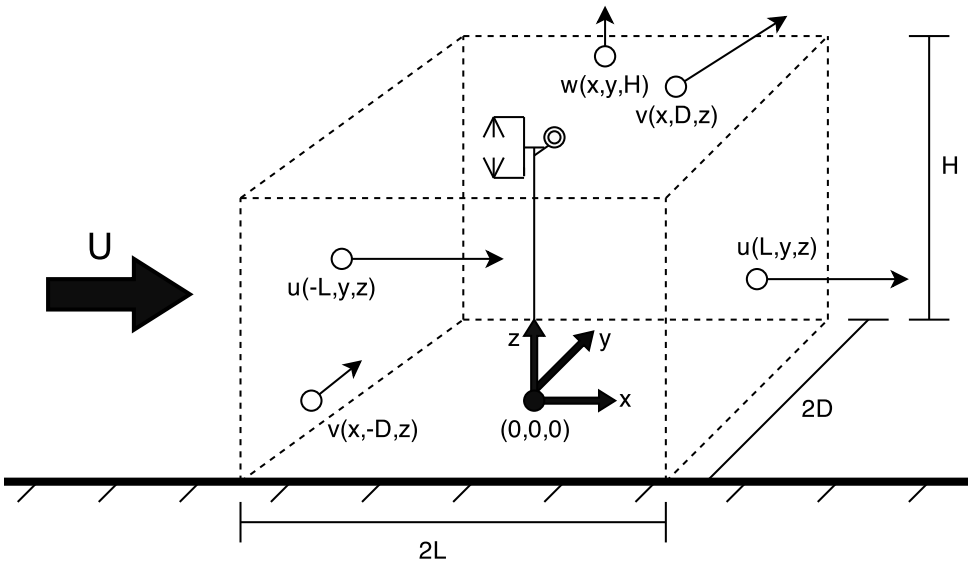


Figure 1.1: *An idealized rectangular control volume $\mathcal{V}_c = 4LDH$ centered about a mast on which a leveled eddy covariance system is mounted at a height $z = H$ above flat ground, with the sonic anemometer facing into the mean wind U. Normal components of the velocity vector are depicted at arbitrary positions on each aerial face of the control volume.*

A natural way of introducing the EC method is in the context of a generalized scalar conservation equation (cf. Fuehrer and Friehe (2002), Lee et al. (2006), Foken (2008b) and Aubinet et al. (2012))

$$\frac{\partial \xi}{\partial t} + \frac{\partial}{\partial x_j}\left(v_j \xi\right) = \Sigma_\xi + \nu_\xi \frac{\partial^2 \xi}{\partial x_j^2} \tag{1.1}$$

where ξ is the scalar in question, v_j is the wind vec-

tor, Σ_ξ is the net source term (sum of sources minus sinks) and ν_ξ is the molecular diffusivity (viscosity for momentum) coefficient which we take to be a constant. In (1.1) we have adopted indicial notation, also known as summation notation (Stull, 1988), which we clarify, for those unfamiliar wih the concept, in Appendix A.1. The presented conservation equation is generalized in that with only slight adjustments, accounting for relevant sources and sinks, it can be applied to any scalar quantity such as absolute humidity or enthalpy. The same can even be said in the case of a vector quantity such as momentum. In essence the equation states that a scalar is conserved under advection in the absence of sources/sinks and molecular diffusion.

In the subsequent derivation where our aim is to arrive at an equation for the surface exchange budget of the scalar ξ we begin by constructing a control volume about the local patch of surface that is of interest. For simplicity, as in Finnigan et al. (2003), we construct a rectangular control volume, $\mathcal{V}_c$, in a Cartesian coordinate system as depicted in Figure 1.1. As shown in Finnigan (2004) the simplified derivation with a flat surface and rectangular control volume is readily extended to more realistic control volumes where the topography and surface type are typically heterogeneous to some degree.

Next we introduce the concept of the block time average: for any continuous series, $\xi(t)$, the m-th block time average is defined as

$$\overline{\xi}^{(m)} = \frac{1}{\tau_A}\int_{t_m}^{t_m+\tau_A} \xi(t)\,dt\,,$$

where τ_A is the duration of the block (the averaging period). The superscript (m) is usually omitted and made implicit. That is, when refering to block averages it is understood that the average of an instantaneous variable is equal for all instances within a given block, that is for $t \in [t_m, t_m + \tau_A]$, but not (necessarily) for instances in adjacent blocks. Moreover, any instantaneous value in the series within a block can be decomposed into the sum of the block average and a deviation, i.e. $\xi(t) = \overline{\xi} + \xi'(t)$. Further, we note that the high pass filtering operation involving the mean removal of the block average satisfies *all* of Reynolds averaging rules (these are outlined in e.g. Stull (1988)). For example,

$$\overline{\xi'} = \frac{1}{\tau_A}\int_{t_m}^{t_m+\tau_A} \xi(t)\,dt - \frac{\overline{\xi}}{\tau_A}\int_{t_m}^{t_m+\tau_A} dt = \overline{\xi} - \overline{\xi} = 0\,,$$

the block average of the fluctuation about the block average is always zero. Now consider what happens upon introducing the complimentary variable χ when we consider the block average of the product of the fluctuations

$$\begin{aligned}\overline{\xi'\chi'} &= \frac{1}{\tau_A}\int_{t_m}^{t_m+\tau_A} \left(\xi(t)-\overline{\xi}\right)\left(\chi(t)-\overline{\chi}\right)\,dt \\ &= \frac{1}{\tau_A}\int_{t_m}^{t_m+\tau_A} \xi(t)\chi(t)\,dt + \frac{\overline{\xi}\,\overline{\chi}}{\tau_A}\int_{t_m}^{t_m+\tau_A} dt \qquad (1.2)\\ &\quad - \frac{\overline{\chi}}{\tau_A}\int_{t_m}^{t_m+\tau_A} \xi(t)\,dt - \frac{\overline{\xi}}{\tau_A}\int_{t_m}^{t_m+\tau_A} \chi(t)\,dt \\ &= \overline{\xi\chi} + \overline{\xi}\,\overline{\chi} - \overline{\xi}\,\overline{\chi} - \overline{\xi}\,\overline{\chi} = \overline{\xi\chi} - \overline{\xi}\,\overline{\chi} \qquad (1.3)\end{aligned}$$

clearly this is generally *not* zero; even if $\xi = \chi$. We refer to $\overline{\xi'\chi'}$ as the (block) eddy covariance between ξ and χ, whereas $\overline{\xi'^2}$ is the (block) eddy variance of ξ. We emphasize that we are working with discrete block averages and subseqent mean removal when defining fluctuations. We do so because other filtering operations do not satisfy the traditional Reynolds averaging rules as demonstrated in Lee et al. (2006). The former authors show that this is even the case for the still relatively simple linear detrend where each block in the series is decomposed into a linear trend and fluctuations about the trend. So working with mean removal is simplest not only in its implementation but also in the fact that Reynolds averaging rules are obeyed making the subsequent derivation much clearer. It is straightforwards to recover the total covariance $\overline{\xi\chi}$ provided that the eddy covariance $\overline{\xi'\chi'}$ along with the respective means are stored using (1.3).

Before continuing with the block average of (1.1) we digress to clarify some concepts and terminology. The term $v_j\xi$ is the kinematic advective flux density of the scalar ξ in direction $\widehat{\mathbf{i}}_j$. To understand what we mean by kinematic flux density it is easiest to first consider what we mean by dynamic flux density by following the discussion in Stull (1988). A dynamic flux density has dynamic units (e.g. energy, mass or momentum) per unit area (hence the term density) per unit time (hence the term flux). So, for example, ρuw is the amount of horizontal momentum passing a unit horizontal surface area[3] (to which w is the surface normal velocity component) per unit time with corresponding units $\left[\mathrm{Nm^{-2}}\right] = \left[\mathrm{kg\,m\,s^{-1}}\right]\left[\mathrm{m^{-2}s^{-1}}\right]$. Unfortunately ρu (ρ in particular) is a difficult quantity to measure, hence why it is practical to operate with the kinematic momentum flux density uw with units of momentum per unit mass times velocity $\left[\mathrm{ms^{-1}}\right]^2 = \left[\mathrm{kg\,m\,s^{-1}}\right]\left[\mathrm{m\,s^{-1}kg^{-1}}\right]$. So converting from kinematic to dynamic units is 'simply' a matter of multiplication by air density. As a corollary then w is the vertical kinematic advective mass flux density, since ρw with units $[\mathrm{kg}][\mathrm{m^{-2}s^{-1}}]$ is clearly the vertical dynamic advective mass flux density. This 'multiply

[3] Due to the symmetry inherent in this term it can equally be interpreted as $(\rho w)\,u$ the amount of vertical momentum passing a unit vertical area (to which u is the surface normal velocity component) per unit time.

by ρ' rule holds *provided* that the kinematic flux has the right units. For example, as outlined in Businger (1982) kinematic sensible heat flux should be expressed in terms of enthalpy c_pT, not temperature alone, for multiplication by ρ to yield the correct dynamic units $\left[\mathrm{Wm}^{-2}\right] = [\mathrm{J}]\left[\mathrm{s}^{-1}\mathrm{m}^{-2}\right]$. Understandably, and as is the convention[4], we will abreviate flux densities as fluxes and omitt the prefix kinematic/dynamic as these are implicit in the units of the flux. The context should always make it clear if we are really refering to the net flux out of the boudning surfaces of some control volume *or* the flux density.

Next we express the diffusive term in (1.1) as a flux divergence by expanding the Laplacian operator (recall that ν_ξ is assumed constant),

$$-\nu_\xi \frac{\partial^2 \xi}{\partial x_j^2} = \frac{\partial}{\partial x_j}\left(-\nu_\xi \frac{\partial \xi}{\partial x_j}\right) .$$

Where the term in the brackets is the (kinematic) diffusive flux (density) of the scalar ξ. Note that the diffusive flux runs *against* (down) the scalar gradient. So we could just as well have expressed (1.1) in the following form

$$\frac{\partial \xi}{\partial t} = -\frac{\partial}{\partial x_j}\left(v_j \xi - \nu_\xi \frac{\partial \xi}{\partial x_j}\right) + \Sigma_\xi \,. \tag{1.4}$$

This corresponds to the 'flux form' of (1.1) with the term in the brackets representing the total (sum of advective and diffusive) flux of ξ.

As a thought experiment to familiarize ourselves with the concepts of diffusive and advective flux let us for simplicity restrict ourselves to the case of a one dimensional flow and scalar profile as depicted in Figure 1.2. Assuming that there are no sources or sinks of the scalar then (1.4) reads

$$\frac{\partial \xi}{\partial t} = -\frac{\partial}{\partial x}\left(u\xi - \nu_\xi \frac{\partial \xi}{\partial x}\right) . \tag{1.5}$$

The respective flux vectors appearing inside the bracket are depicted in Figure 1.2 at 5 points in the case of a semi-circular scalar profile. The advective flux is positive, flowing in the postive x-direction, at each point. It takes its maximum magnitude where the magnitude of the scalar is at its maximum. Conversely the diffusive flux flows in the negative (positive) x-direction where the scalar gradient is positive (negative) consistent with down-gradient diffusion. It takes its maxmimum absolute value where the scalar *gradient* is at its maximum. Crucially the tendency of the scalar profile, as dictated by (1.5), is determined not by the fluxes themselves but by their *convergence*. That is to say the scalar will increase (decrease) at a point if the sum of the fluxes converges (diverges) at that point.

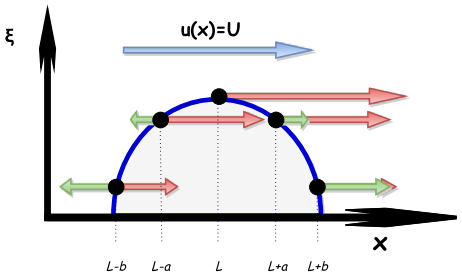


Figure 1.2: *Snapshot of an idealized one dimensional scalar profile $\xi(x)$ (blue line) superimposed on a steady and homogeneous wind field $u(x,t) = U > 0$ (light blue arrow). The colored arrows represent the advective (red) and diffusive (green) flux vectors at 5 points centered about the point $x = L$ where the scalar is at its maximum.*

It is natural make a conjecture as to what the future fate of the scalar ξ in Figure 1.2 will be. Well to the right of $x = L$ the advective flux is converging, whereas to the left it is diverging. Actually in this simple example in the absence of diffusion the scalar profile would simply be advected along in the x direction at a velocity u all the while conserving its shape with a peak at $x = L+ut$ where t is the time since the snapshot in Figure 1.2. The diffusive flux in Figure 1.2, however, is diverging everywhere (except where the scalar is zero). Of the points considered the divergence is greatest at $x = L \pm b$ where the slope of the diffusive flux is large. At the discontinuity in the scalar gradient where the scalar goes to zero, however, the diffusive flux *is* converging. As such in this case diffusion will reduce the magnitude of the scalar where it is non-zero whilst increasing the scalar magnitude where it was previously zero (more generally where the flux is convering). The result of this diffusive, yet non-dissipative, process is that over time the spatial variance of the scalar is reduced whilst conserving the spatial integral ($\int \xi \, dx$) of the scalar profile as outlined in detail in Røed (2013). When we combine advection and diffusion the relative importance of the two processes is generally determined by the wind field, the advective flux is zero if the isotachs are parallel to the scalar isolines, as well as the magnitude of the molecular diffusivity ν_ξ. For orientation if the scalar is temperature then the molecular diffusivity ν_T is typically on the order $2 \times 10^5 \left[\mathrm{m}^2\mathrm{s}^{-1}\right]$ in air (Stull, 1988). Thus, advection is typically the dominant transport mechanism (above the micro-layer). Now despite the simplicity of this thought experiment it is instructive in that the same principles are readily applied without loss of generality to the more realistic case of three dimensional flow. In such a case, however, the velocity components are unlikely to be homogeneous in space such that their gradients also play an important role for the evolution of the scalar.

Having clarified some terminology we are ready to proceed with the derivation. In the context of surface ex-

[4] See for example the AMS glossary: `http://glossary.ametsoc.org/wiki/Flux`.

change it is the value of the source term Σ_ξ that we are interested in diagnosing, specifically near the surface itself (see extensive discussion in Sun et al. (1995)). The instantaneous values are in themselves not very meaningful (Lenschow et al., 1994) representing only a part of the ensemble average flux. So it is more instructive to arrive at a block averaged equation, making the ergodic hypothesis that time and ensemble averages are equivalent (Lee et al., 2006), expressed with respect to the scalar source term

$$\overline{\Sigma}_\xi = \frac{\partial \overline{\xi}}{\partial t} + \frac{\partial}{\partial x_j}\left(\overline{v_j \xi}\right) - \nu_\xi \frac{\partial^2 \overline{\xi}}{\partial x_j^2}\,. \tag{1.6}$$

We also define the following shorthand notation: $\oint_{\mathcal{V}_C} d\mathcal{V} = \int_0^H \int_{-D}^D \int_{-L}^L dx\,dy\,dz$ as the closed integral over the control volume as well as $\iint_{\mathcal{A}_x} d\mathcal{A} = \int_0^H \int_{-D}^D dy\,dz$, $\iint_{\mathcal{A}_y} d\mathcal{A} = \int_0^H \int_{-L}^L dx\,dz$ and $\iint_{\mathcal{A}_z} d\mathcal{A} = \int_{-D}^D \int_{-L}^L dx\,dy$ as the area integrals over the faces of the control volume where x,y and z respectively are held constant. We are now ready to integrate over our control volume. In doing so we expand the advective flux divergence term into its spatial components and use (1.3) on the mean of the products (the total covariance). The subsequent volume integral of (1.6), upon applying Reynolds averaging rules, reads

$$\overbrace{\oint_{\mathcal{V}_c} \overline{\Sigma}_\xi \, d\mathcal{V}}^{\text{Bulk source}} = \overbrace{\iint_{\mathcal{A}_z} \left[\left(\overline{w}\overline{\xi}\right)\big|_{z=H} - \left(\overline{w}\overline{\xi}\right)\big|_{z=0}\right] d\mathcal{A}}^{\text{Vertical advective flux}} + \overbrace{\iint_{\mathcal{A}_z} \left[\left(\overline{w'\xi'}\right)\big|_{z=H} - \left(\overline{w'\xi'}\right)\big|_{z=0}\right] d\mathcal{A}}^{\text{Vertical eddy flux}}$$
$$+ \overbrace{\iint_{\mathcal{A}_x} \left[\left(\overline{u}\overline{\xi}\right)\big|_{x=L} - \left(\overline{u}\overline{\xi}\right)\big|_{x=-L}\right] d\mathcal{A} + \iint_{\mathcal{A}_y} \left[\left(\overline{v}\overline{\xi}\right)\big|_{y=D} - \left(\overline{v}\overline{\xi}\right)\big|_{y=-D}\right] d\mathcal{A}}^{\text{Horizontal advective flux}}$$
$$+ \overbrace{\iint_{\mathcal{A}_x} \left[\left(\overline{u'\xi'}\right)\big|_{x=L} - \left(\overline{u'\xi'}\right)\big|_{x=-L}\right] d\mathcal{A} + \iint_{\mathcal{A}_y} \left[\left(\overline{v'\xi'}\right)\big|_{y=D} - \left(\overline{v'\xi'}\right)\big|_{y=-D}\right] d\mathcal{A}}^{\text{Horizontal eddy flux}}$$
$$+ \overbrace{\oint_{\mathcal{V}_c} \frac{\partial \overline{\xi}}{\partial t} d\mathcal{V}}^{\text{Storage}} + \overbrace{\iint_{\mathcal{A}_z} \left(\left[-\nu_\xi \frac{\partial \overline{\xi}}{\partial z}\big|_{z=H}\right] - \left[-\nu_\xi \frac{\partial \overline{\xi}}{\partial z}\big|_{z=0}\right]\right) d\mathcal{A}}^{\text{Vertical diffusive flux}}$$
$$+ \underbrace{\iint_{\mathcal{A}_x} \left(\left[-\nu_\xi \frac{\partial \overline{\xi}}{\partial x}\big|_{x=L}\right] - \left[-\nu_\xi \frac{\partial \overline{\xi}}{\partial x}\big|_{x=-L}\right]\right) d\mathcal{A} + \iint_{\mathcal{A}_y} \left(\left[-\nu_\xi \frac{\partial \overline{\xi}}{\partial y}\big|_{y=D}\right] - \left[-\nu_\xi \frac{\partial \overline{\xi}}{\partial y}\big|_{y=-D}\right]\right) d\mathcal{A}}_{\text{Horizontal diffusive flux}}\,. \tag{1.7}$$

The above represents, without approximation, the time averaged budget for the scalar ξ in our idealized control volume. It is considered to be the fundamental equation of eddy covariance. Often presented in different forms, compare (1.7) to that in Gu et al. (2012) or Lee and Massman (2011), the physics (see braces) behind each term appearing in the equation remains the same. We use the term bulk source to emphasize that the (point) source term is integrated over the entire control volume. Gauss' theorem is used to express the volume integral of the advective and diffusive flux divergences as the sums of the advective and diffusive fluxes *out* of each face of the volume.

Observant readers will perhaps wonder why we have dubbed the integrals of the $\overline{v}_j\overline{\xi}$ terms the 'advective' fluxes whereas we denote the $\overline{v_j'\xi'}$ integrals 'eddy' fluxes when both terms are in fact advective processes. This is just a convention in micrometeorology where advection is usually associated with the larger scale mean flow and eddy (turbulent) transport is associated with the smaller scale fluctuations. Such a convention is adpoted because the 'advective' terms can be treated deterministically in models whilst the 'eddy' terms are usually unresolved (subgrid scale) and must be parametrized (Stull, 1988).

Traditional Approach

Traditionally a set of assumptions are made to simplify the analysis. That is to say the fields being measured are assumed to approximately satisfy a rather long and stringent list of criteria. These are (see e.g. Foken and Wichura (1996) and Finnigan et al. (2003))

1. Statistical stationarity whereby the mean storage term is zero, i.e. $\frac{\partial \overline{\xi}}{\partial t} = 0$.
2. Horizontal homogeneity of first and second order statistics, in other words no horizontal advective or eddy flux.
3. No *mean* subsidence, that is $\overline{w} = 0$.
4. The point source term is negligible except at the surface, so in the surface exchange budget we may replace it with $\overline{\Sigma}_\xi(x, y, 0)\delta(z)$ where $\delta(z)$ is the Dirac delta function (Finnigan et al., 2003).

Complimenting the above assumptions we make use of two convenient conditions. Frist in line is the diffusive term, whos the treatment varies between investigators. For example, in Lee (1998) the term is explicitly ignored without justification, while in Finnigan et al. (2003), Kowalski and Serrano-Ortiz (2007) and Gu et al. (2012) (amongst others) it is proposed that the term should act as a resistance in the source term near the surface. As our first condition we follow the latter approach in taking molecular diffusion to be negligible in the budget except in the so-called micro-layer, the lowest few millimeters above the surface (Stull, 1988), where we absorb it into the source term. Further, as our second condition we enforce a lower kinematic boundary condition which dictates that there can be no through-flow at the solid boundary represented by the surface, i.e. $w|_{z=0} = 0$.

For clarity we will apply the assumptions and complimentary conditions on a term by term basis. We will also, following Finnigan et al. (2003), divide through by the area of the surface patch $\mathcal{A}_z$ to yield units of flux density. First we introduce the surface exchange term (surface flux), denoted S_0, which we define as the sum of the bulk source minus the integrated diffusive flux. Applying the conditions of horizontal homogeneity, sources/sinks restricted to the surface and negligible diffusive flux outside the microlayer, then upon integration this term becomes

$$S_0 = < \overline{\Sigma}_\xi \big|_{z=0} >_{\mathcal{A}_z} - \frac{\partial \overline{\xi}}{\partial z}\big|_{z=0} .$$

By definition this term corresponds to the horizontally averaged ($<>_{\mathcal{A}_z}$ operator) block time averaged source strength of constituent ξ across the surface patch with area $\mathcal{A}_z$ plus the diffusive flux at the surface. Thereby S_0 represents the average exchange of ξ between the surface and the atmosphere which is ultimately what we are interested in diagnosing.

Next we consider the total vertical covariance as the sum of the vertical advective flux and the eddy covariance, i.e. $\overline{w\xi} = \overline{w}\overline{\xi} + \overline{w'\xi'}$. Due to the no mean subsidence approximation $\overline{w} = 0$ the vertical advective flux term conveniently falls out and we are left with $\overline{w\xi} = \overline{w'\xi'}$. So using the former, horizontal homogeneity and the bottom boundary condition the *total* (mean advective+eddy) vertical advective flux term becomes the vertical eddy flux at the measurement height, i.e.

$$\frac{1}{\mathcal{A}_z} \iint_{\mathcal{A}_z} \left[\overline{w\xi}\right]_0^H \, d\mathcal{A} = \overline{w'\xi'}|_{z=H} .$$

As for the *net* horizontal advective flux density terms, by virtue of the horizontal homogeneity assumption these are both zero. That is

$$\frac{\partial \overline{u\xi}}{\partial x} = 0 \longleftrightarrow \overline{u\xi}|_{z=L} - \overline{u\xi}_{z=-L} = 0 ,$$

and

$$\frac{\partial \overline{v\xi}}{\partial y} = 0 \longleftrightarrow \overline{v\xi}|_{z=D} - \overline{v\xi}_{z=-D} = 0 .$$

Subsequently the total horizontal advective flux terms make no contribution to the surface exchange budget. The storage term is next in line which on account of the stationarity assumption is taken to be zero. As such we are left with the following surface exchange budget

$$S_0 = \overline{w'\xi'}\big|_{z=H} , \tag{1.8}$$

where the eddy covariance term on the right hand side is readily measured using fast-response instrumentation such as a sonic anemometer paired with an open-path gas analyzer, which is the case for our study. Hence, provided the assumptions hold, it is possible to diagnose the surface exchange terms (surface fluxes) by sampling the eddy covariance at some height within the surface layer.

Equation represents the essence of the traditional eddy covariance method[5]. This leads to a *first order* definition of the dynamic surface fluxes (Fuehrer and Friehe, 2002). Considering two familiar examples: for sensible heat flux the first order definition is simply Stull (1988)

$$Q_H = c_p \overline{\rho}\, \overline{T'w'} , \tag{1.9}$$

[5]A word of caution is in order. As pointed out by Foken (2008b) the terms eddy covariance and eddy correlation are often used interchangeably, but this will invariably lead to confusion as the indirect (based on flux variance similarity) eddy correlation method is distinct from the eddy covariance method discussed here.

where T is temperature (units K), c_p the specifc heat at constant pressure (units $JK^{-1}kg^{-1}$) and ρ the air density (units kgm^{-3}). For latent heat flux the first order definition is (Stull, 1988)

$$Q_E = \lambda_v \overline{\rho_v' w'}\,, \tag{1.10}$$

where ρ_v is absolute humidity (units kgm^{-3}) and λ_v is the specific heat of evaporation (units Jkg^{-1}). The operative word in these definitions is first order; there is no guarantee that these provide accurate estimates of the corresponding surface fluxes (Fuehrer and Friehe, 2002).

Modern Approach

The modern approach, as outlined in for example Aubinet et al. (2012) and Mauder et al. (2013), provides more accurate second order estimates of the surface exchange. The complimentary conditions, zero vertical velocity at the surface and negligible molecular diffusion above the micro-layer, remain in the modern method as these are based on sound underlying physical assumptions (Finnigan et al., 2003). We will discuss the modern approach as implemented for a single EC measurement system, in the absence of vertical or horizontal profiles, which is the setup for many of the FLUXNET sites (Baldocchi et al., 2001); including the site we are studying.

Three of assumptions in the traditional approach are unavoidable also in the modern approach. These are: 1. Stationarity , 2. Horizontal homogeneity and 4. Sources/sinks restricted to the surface. It is, however, possible to verify if the underlying assumptions are clearly violated through careful application of despiking and quality control procedures, introduced by Foken and Wichura (1996) and Vickers and Mahrt (1997).

It is widely considered that *not* implementing the assumption of zero block average subsidence, $\overline{w} = 0$, marked the transition from the traditional to the modern EC methodology (Lee and Massman, 2011). It turns out that although block averaged vertical velocities are often too small to be measured acurately (Fuehrer and Friehe, 2002) they are by no means negligible and may have a significant impact, on the surface exchange. To alleviate this the block averaged vertical velocity is diagnosed through the 'WPL' terms (after Webb et al. (1980)) based on the permise of zero dry air mass flux and the fluxes are consequently corrected. The former is far from the only flux correction in the modern methodology. Additionaly, cospectral attenuation is compensated for through methods pioneered by Moore (1986). The sonic heat flux measured by the EC system is converted to the sensible heat flux through the 'SND' correction (after Schotanus et al. (1983)). The effects of sensor separation, gas analyzers being some distance away from velocity sensors, are corrected for by considering cross-correlations (Nordbo et al., 2012). In addition the effects of tilt, arguably the largest source of systematic error (Mahrt, 2010), are corrected for by rotating the horizontal measurement frame into a longterm ensemble streamline plane using the planar fit method proposed by Wilczak et al. (2001). Finally the magnitude of flux sampling uncertainty is quantified using methods such as that of Finkelstein and Sims (2001).

At this stage the brief overview may not be very insightful, but we will explain all of the mentioned steps and how we implemented these in Chapter 2. It is worth keeping in mind that the overall goal of the modern EC method is to provide more accurate estimates of the surface exchange by improving on the deficiencies of the traditional approach (Mahrt, 2010).

Monin-Obukhov Stability Parameter

Being a widely used concept in the EC method the Monin-Obukhov (M-O) stability parameter (after Monin and Obukhov (1954)), denoted ζ, deserves a brief description . It is defined as $\zeta = z/L_*$ where L_* is the Obukhov length first presented in[6] Obukhov (1971) and z is the height in the surface layer at which L_* is evaluated. The Obukhov length in turn is defined as (Foken, 2008b)

$$L_* = -\frac{u_*^3}{\kappa \frac{g}{\theta_v} \overline{\theta_v' w'}}\,,$$

where $u_* \geq 0$ is the friction velocity, $\kappa = 0.4$ is the von Karman constant and θ_v is the virtual potential temperature and $\overline{\theta_v' w'}$ is the buoyancy flux. In the context of EC campaigns it is usually assumed that the measured sonic temperature, T_s, is approximately equal to θ_v (Kaimal and Finnigan, 1994). According to Stull (1988) L_* can be interpreted physically as being proportional to the height where buoyancy effects begin to dominate over wind shear in the turbulent kinetic energy budget (see (1.16)). Thereby for stable stratification where buoyancy inhibits turbulence ($\overline{\theta_v' w'} < 0$) L_* is positive and tends to zero as the stability increases. Converesely under unstable stratification where buoyancy produces turbulence ($\overline{\theta_v' w'} > 0$) L_* is negative tending to zero as stratification becomes more unstable. In perfectly neutral conditions L tends to $\pm\infty$ in that the buoyancy flux is by definiton zero. From the discussion of L_* it is clear that for unstable stratification $\zeta < 0$, for stable stratification $\zeta > 0$ and for neutral stratification $\zeta = 0$.

[6] Our reference is a translation to English of the original Russian version which was published in 1946 (Foken, 2006)

1.2.2 Study Site

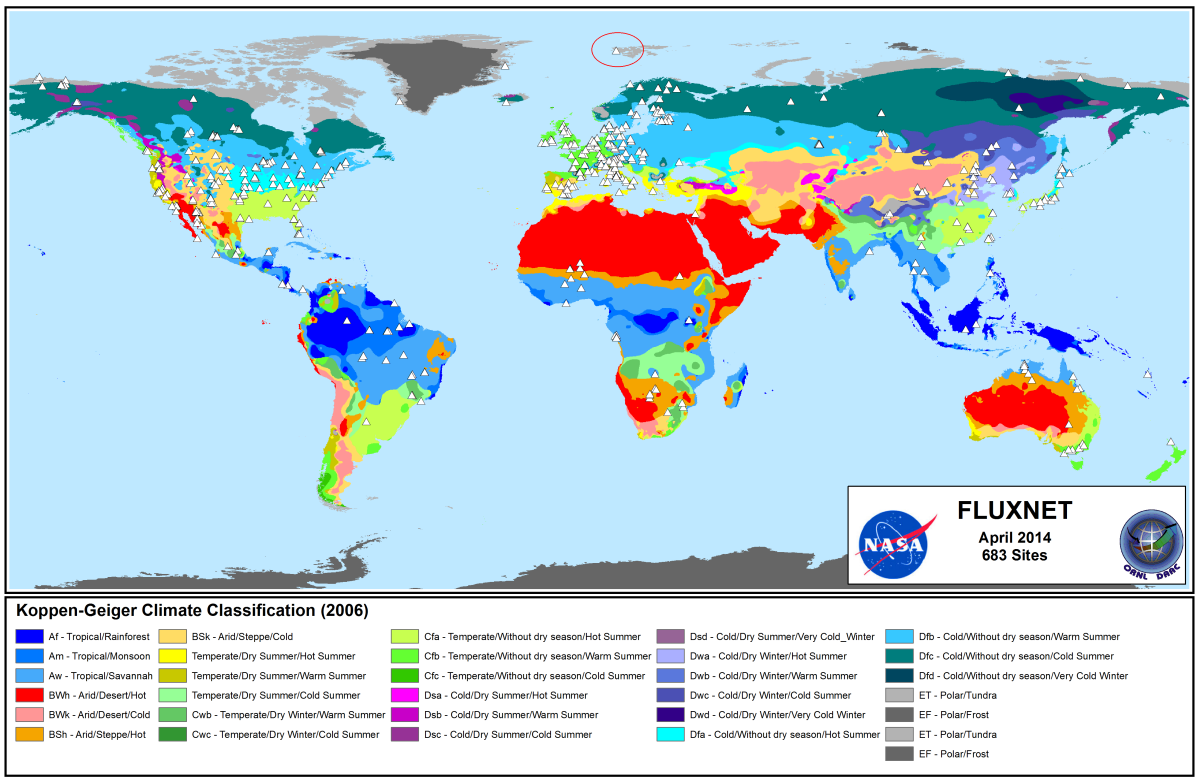


Figure 1.3: Overview of the sites (indicated by white triangles) that form FLUXNET; a component of the NASA Oak Ridge National Laboratory (ORNL) Distributed Active Archive Center for Biogeochemical Dynamics (DAAC). The Bayelva site is circled in red. Source: `http://fluxnet.ornl.gov/maps-graphics`

In the following we provide an overview of what we consider to be the salient points concerning the measurement site. More details can be found in Westermann (2010) where a complete description (on which this overview is based) of the site and the annual surface energy budget in the context of permafrost temperature is given for the period 03.15.2008-03.15.2009. The EC system from which our raw data was sourced is run and owned by the Alfred Wegener Institute (AWI) and is part of the greater FLUXNET global network of EC sites as shown in Figure 1.3[7]. The instrumentation employed, a sonic anemometer and an infrared gas analyzer, will be surveyed in Section 1.2.3.

The EC system is mounted on a mast on the north western slope (inclination of $\gamma < 5°$) of the Leirhaugen hill at 78°55′15”N, 11°49′53”E in the Bayelva river catchment about 2 km south west from Ny Ålesund. The researc settlement that is Ny Ålesund lies on the southern shore of Kongsfjorden on the Brøgger peninsula on the island of Spitsbergen in the north west of the Svalbard archipelago as depicted in the left panels of Figure 1.4. Lierhaugen hill peaks at 25 m ASL and lies at the foot of two large glaciers: Vestre Brøggerbreen to the west-south west and Austre Brøggerbreen to the south. Other significant topographical features include Zeppelinfjellet (566 m ASL) 1.9 km to the south east, Scheteligtoppen (719 m ASL) 3.5 km to the west-north west and Brøggerfjellet (653 m ASL) 3.35 km to the south west separating the two glaciers. Kongsfjorden lies 1.5 km to the north-east. The fjord runs along a north-west (outer) to south-east (inner) axis and was almost entirely free of ice during the entire study period (see Westermann et al. (2009) and references therein). Positions, horizontal distances and peak elevations are based on the free interactive online topographical resource TopoSvalbard (TopoSvalbard, 2015) produced by the Norwegian Polar Institute (NPI) where aerial, satellite and ’3D’ images of the Brøgger peninsula are also available.

[7]Details at `http://fluxnet.ornl.gov/site/4115`

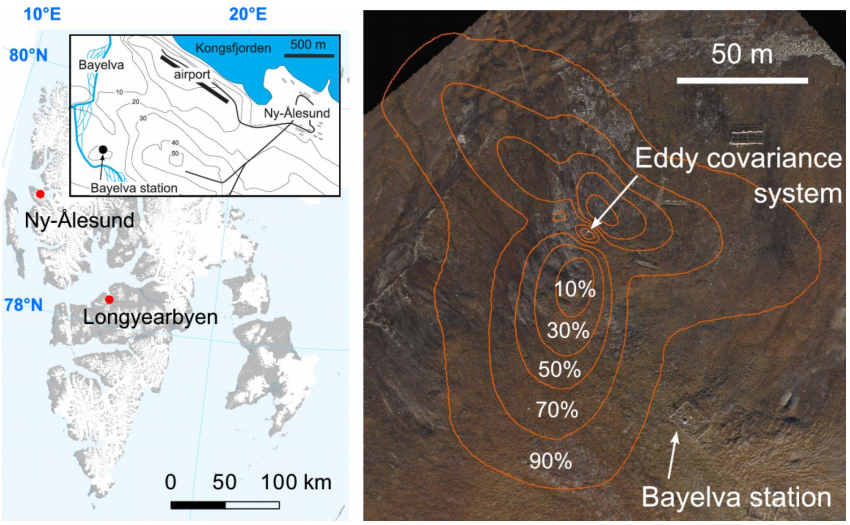
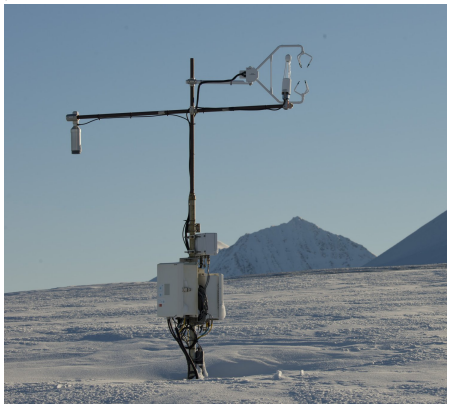


Figure 1.4: *Location of the Eddy Covariance system with a cumulative flux footprint (orange contours) depicted in the right panel. Leirhaugen hill is shown by the closed height contour intercepting the Bayelva climate station in the small left panel. Figure from Lüers* et al. *(2014) where it was adapted from the original Figure in Westermann* et al. *(2009).*

A footprint analysis based on the model of Schmid (1994) was carried out in Westermann et al. (2009) for the snow free period July-September 2008 to determine the average flux source area, i.e. the fetch of the instruments. This analysis, depicted by the orange contours in the right panel of Figure 1.4, indicates that the fetch of the EC system is practically unobstructed by man-made structures[8]. If we take the Bayelva climate station as an example we see that it falls well outside the 90% *cumulative* footprint contour. Most of the contributing surface flux originates from the tundra along the main wind directions. According to this analysis these main wind directions, that is where the footprint contours are stretched, are east-south easterly (from the inner part of Kongsfjorden), southerly (from the Austre Brøggerbreen glacier) and north-westerly (from the outer part of Kongsfjorden). Note that these directions are somewhat in agreement with our own analysis presented later in the form of a wind rose (Figure 2.5). There is nonetheless a slight discrepancy; in our analysis by far the most frequent wind direction is south westerly whereas there is a relatively small contribution directly from the south. We surmise that this south-westerly contribution could be wind that is channeled by Brøggerfjellet as it enters the catchment from either the Vestre or Austre Brøggerbreen. Since the instruments are mounted on the north western slope of a hill the south westerly direction is not as blocked as wind from the south. The discrepancy can also be explained by the fact that we base our analysis on observations from a longer period to what was done in the footprint model. The model of Schmid (1994) is based on an estimate of the field of view, or surface area of influence, of the fast responding sensors. This is in turn dependent on the local highly transient pattern of turbulent transport. Of course, there is no reason to expect that the statistics of the turbulence in the two periods are identical. Yet, the remaining wind directions with significant occurrence that we identify, east-south easterly and north-westerly, are consistent with the footprint analysis. Therefore, we consider the footprint contours to be a useful representation of the fetch of the instruments, at least to a first order, in our study period. At the very least we can safely make the assumption that the footprint area is undisturbed by man-made structures.

Figure 1.5: *View of the Bayelva EC system, mounted on Leirhaugen hill, when facing east. For orientation the northern ridge of Zeppelinfjellet is glimpsed to the far right of the image. Photograph taken downwind of the system on the 23.03.2015 courtesy of Nathalie Grenzhaeuser.*

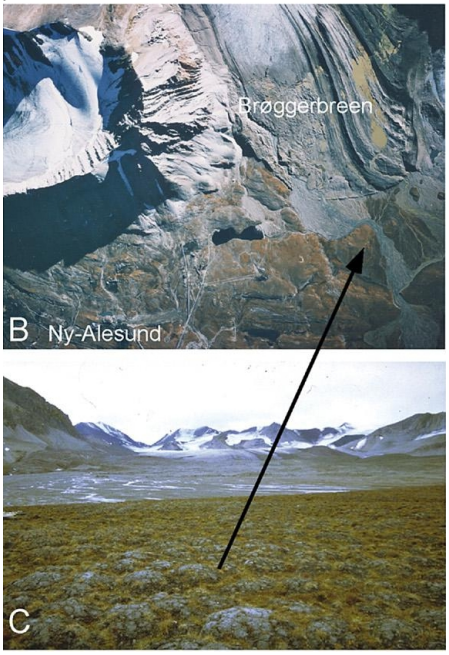


At Leirhaugen the ground is permafrost with a maximum active layer depth, the maximum depth of the $0°$C isotherm below the surface, reported as ~ 1.5 m in 2008. The permafrost is relatively warm with a mean annual temperature of $-2.5°$ at 1.5m as reported by

[8] Apart from the tower structure on which the instruments are mounted.

continuous measurements at the Bayelva climate station (since 1998). The soil content is rich in minerals and has a low organic content. A distinctive feature of the seasonal freeze-thaw pattern of the periglacial landscape is the occurrence mud boils which are non-sorted circles[9] characterized by a trough covered with vegetation in the form of low plants, moss and lichens and a bare mineral soil center with a diameter on the order of 1 m as seen in the bottom of panel C of Figure 1.6. These mud boils, which are surveyed in detail in Boike et al. (2008), introduce micro-scale surface heterogeneity in the sense that the vegetation acts as a buffer both to water and CO_2, while the bare center of the boils are exposed to fluxes of significant magnitude. Due to the small scale, however, these discontinuities in surface exchange are smoothed out by turbulent mixing in a typical flux averaging period. The vegetation cover is estimated to be 60% (Lüers et al., 2014), but as the vegetation is low the surface is quite smooth with a roughness length $z_0 \simeq 7\,\text{mm}$ in the absence of snow. We are spared from forest canopies and roughness sublayers that occur within them; a fascinating but complicating feature of many EC sites at lower latitudes (see e.g. Lee et al. (2006)).

Figure 1.6: *Panel B: Aerial photograph of the area to the south-east of Ny-Ålesund with Leirhaugen hill indicated by the arrow. Panel C: View from the top of Leirhaugen hill looking south (August 1998). Adapted from Boike* et al. *(2008).*

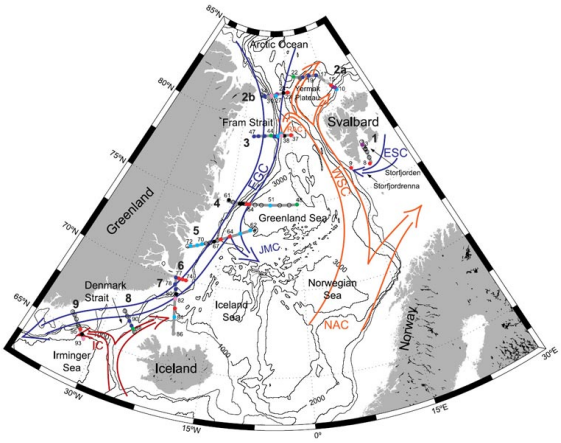


We do, however, have to contend with changing patterns in the surface type (compare Figures 1.6 and 1.5) as a result of snow-fall/freeze/drift/thaw/melt and rare rain-on-snow events. Locally the latter events are very significant as they mark a rapid shift in surface albedo, surface roughness as well as measurement height. Moreover these events have become noticeably more frequent in the so-called 'shoulder' months (spring and fall) (Nowak and Hodson, 2013). In fact, the former authors show that in May,September and October rain has become the dominant form of precipitation in the Bayelva catchment over the course of the last decade. More generally the 'snow cycle' is key in modulating the surface energy balance. Take the effect of changing surface albedo on the net radiation budget as an example. At this latitude the incoming shortwave radiation is largely determined by the annual transition between polar night and day, along with cloud cover, but the *portion* that is reflected is modulated by the whiteness of the surface. As noted in Westermann et al. (2009) there is a fascinating interplay between the properties of the surface and the seasons that determines the surface radiation budget. On the one hand the timing of snow-melt, with the snow disappearing typically somewhere between the end of May to the beginning of July (Winther et al., 2002), marks a significant drop in the albedo and coincides with the period in which the incoming shortwave radiation is at its maximum. On the other hand the timing of the first seasonal snow fall and subsequent snow formation is not as significant; even if it marks a significant increase in the albedo the sun is already low on the horizon at this time, usually late September (Winther et al., 2002), so the effect on net radiation is limited compared to that during the onset of the snow-melt.

The varying snow cover also poses a problem in terms of changing the height of the instruments in relation to the surface. The height varies from as low as 1 m above the snow covered surface (March-May 2008) to 2.75 m in the absence of snow. This poses a challenge for the tilt correction algorithm described in section 2.3.2 as the method requires that the instruments are stationary relative to the surface; whereas at our site over a long enough time scale (months or more) the surface moves significantly from the instruments' perspective (as shown in Figure 2.10). Moreover the M-O stability parameter $\zeta = z/L_*$ requires an accurate measurement of the height AGL, z, of the instruments. Fortunately we had access to data from the Campbell SR50 sonic ranging sensor (SR50, 2007) used to detect changes in the surface below the sonic anemometer which we complimented with snow depth measurements from the nearby Bayelva Climate station as outlined in Section 2.4.2. As such we were able to track changes in measurement height leading to a fairly accurate estimation of the M-O stability parameter.

[9] As opposed to sorted-circles characteristically bordered by rocks.

Figure 1.7: *Ocean currents in the Nordic seas. Currents near Svalbard: EGC=East Greenland Current, ESC=East Spitsbergen Current, WSC=West Spitsbergen Current, RAC=Return Atlantic Current and NAC=North Atlantic Current Blue (red/orange) colors represent cold (warm) currents. Figure adopted from Rudels* et al. *(2005)*



For the sake of context we also provide a short description of the climatology of the Ny Ålesund area which has a solid permanent record of both air temperature and precipitation stretching back to 1969 (Førland and Hanssen-Bauer, 2000). The climate of Svalbard can be classified as ET-Polar Tundra (Figure 1.3), following the Köppen-Geiger Climate Classification (Peel et al., 2007) as the climatological average of the warmest month is in the range $0°\text{C} < T < 10°\text{C}$. This criterion is later shown to be satisfied locally for the Ny Ålesund area in Figure 2.2. The (1981-2010) average annual temperature is $-5.2°$C, with average winter air temperatures of $-12°$C and average summer air temperatures of $3.8°$C (Førland et al., 2012). Temperatures which are relatively mild given the latitude; for comparison in Alert, Nunavut, Canada ($82°30'00''$N, $62°19'59''$W) the average annual air temperature is more than ten degrees lower at $-18.1°$C (Smith et al., 2005). In fact such a zonal temperature asymmetry is also present on the scale of the Spitsbergen island itself. On the west coast, where Ny-Ålesund is situated, the relatively warm West-Spitsbergen Current is the major source of both heat and moisture (Esau et al., 2012) while on the east coast the colder East-Spitsbergen current brings colder polar water masses as well as sea ice in from the north as depicted in Figure 1.7. The (1981-2010) average annual precipitation is 427 mm yr^{-1} (Førland et al., 2012), where on average most (almost twice as much) falls in autumn and winter when compared to spring and summer. Typically only 25% of the precipitation falls as rain with the remainder as snow, or sleet or a combination of all the former. Precipitation can fall as both rain and snow in any given month of the year. As noted in Førland and Hanssen-Bauer (2000) the precipitation is considerably higher over the glaciers as a result of orographic enhancement. This in turn could lead to significant differences between the annual precipitation locally at Leirhaugen hill when compared to Ny Ålesund due to the proximity to the Brøgger glaciers. A point to bear in mind later as both of the instruments are of limited use when precipitation occurs. Increases in both annual precipitation and annual average air temperature have been reported in Ny Ålesund. Annual precipitation has increased by 5% per decade and average annual air temperature has increased by $0.73°$C per decade with a particularly alarming increase of $1.36°$C per decade in average winter time temperatures. Both these decadal temperature trends are based on data from 1975-2011 as detailed in Førland et al. (2012). The former authors note that this warming trend recognized at various sites in the Svalbard region are consistent with the large-scale warming observed in the Arctic over the last decades (see Hartmann et al. (2013)).

1.2.3 Instrumentation

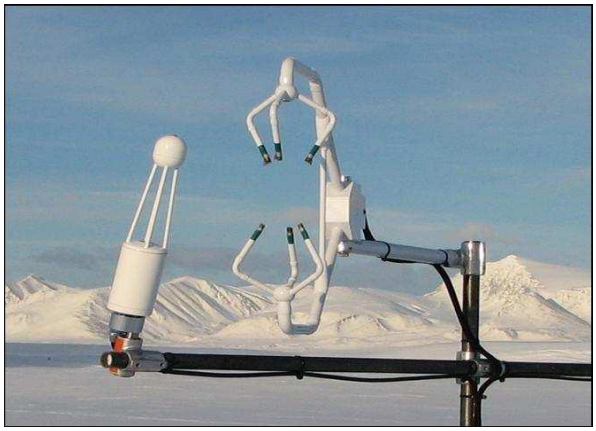
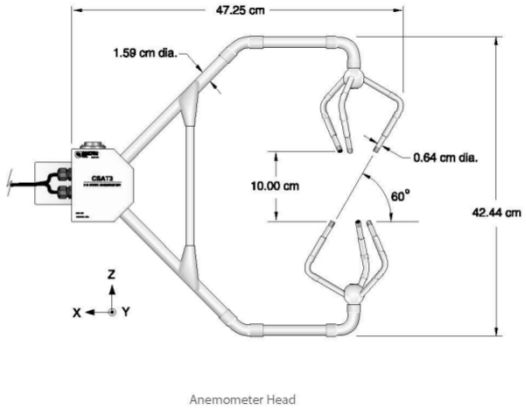


Figure 1.8: *Close up view (facing north) of the Bayelva EC system mounted on Leirhaugen hill with the sonic in the center and the IRGA on the left side of the image. Photograph adopted from Westermann (2010).*

In the following section we will briefly review the instrumentation employed in the Bayelva EC system: a CSAT3 (CSAT3, 2014) sonic anemometer (sonic for short) and a LI-7500 (LI-7500, 2001) open path infrared gas analyzer (abbreviated IRGA). These are mounted on a mast at a height of 2.75m above snowfree ground on the northwestern slope of Leirhaugen hill. Concerns about the orientation of the sonic, the separation between the IRGA and the sonic as well as the frequency response of the instruments will be adressed in the method (Chapter 2). Herein for the respective instruments particular attention is directed to the *basic* operating procedure, the resolution as well as any known issues not adressed elsewhere. More details are avaiable in the respective manuals CSAT3 (2014) and LI-7500 (2001).

CSTA3 Sonic Anemometer

Figure 1.9: *Schematic of the nonorthogonal geometry of a CSAT3 sonic anemometer head. There are six seperate transducers each with a diameter of* 0.64 *cm; when paired these transducers form three acoustic paths. Adopted from (CSAT3, 2014)*



The CSAT3 is a three dimensional nonorthogonal sonic anemometer used to measure the local three dimensional wind vector and sonic temperature both sampled at the same high frequency (CSAT3, 2014). From Figure 1.9 it is evident why this sonic is nonorthogonal: each of the three acoustic paths, the paths separating any two facing transducers, are tilted $30°$ from the instruments vertical axis. So the path vectors have a mutually parallel component, namely this vertical axis. Such a geometry is very different from the earlier orthogonal sonic models (cf. Kaimal and Businger (1963)) which have three orthogonal acoustic paths including the instrument's vertical axis. Crucially, the nonorthogonal geometry of the CSAT3 minimizes the effects of flow distortion that results from transducer shadowing of the wind field (CSAT3, 2014). Even so, the basic operating procedure is much the same as the earlier orthogonal sonics.

Along each of the three acoustic paths, henceforth sonic axes denoted with the subscript[10] $a \in 1(1)3$, two ultrasonic waves are transmitted in opposite directions. Following the discussion in CSAT3 (2014) for a given sonic axis a the flight time of the outgoing ultrasonic wave, $t_{a,1}$, here defined for a given transducer pair as the wave transmitted from the lower to the upper transducer is given by

$$t_{a,1} = \frac{d}{c_{s,a} + u_a} \,, \tag{1.11}$$

and the flight time of the incoming wave, from the upper to to the lower transducer, $t_{a,2}$, is given by

$$t_{a,2} = \frac{d}{c_{s,a} - u_a} \,. \tag{1.12}$$

In (1.11) and (1.12) d is the constant path length ($\simeq 0.115\,\mathrm{m}$ for the CSAT3), $c_{s,a}$ is the speed of sound along the axis and u_a is the component of the wind vector blowing along sonic axis a defined as positive if the wind is blowing from the lower to the upper transducer. Combining (1.11) and (1.12) we have that u_a is given by

$$u_a = \frac{d}{2}\left[\frac{1}{t_{a,1}} - \frac{1}{t_{a,2}}\right] .$$

This allows the CSAT3 to calculate u_a internally for each axis based on the measured flight times. The three non-orthogonal wind components are then rotated into an orthogonal frame internally in the CSAT3 firmware. As such, near instantaneous samples of three orthogonal velocity components are calculated and stored. These samples are the x, y and z components of the velocity vector in the instrument frame (cf. Figure 1.9) denoted as u, v and w respectively.

In addition to the three dimensional wind vector, the CSAT3 also diagnoses the speed of sound along each of the three sonic axes. Once more this is achieved by combining (1.12) and (1.11) to arrive at (CSAT3, 2014)

$$c_{s,a} = \frac{d}{2}\left[\frac{1}{t_{a,1}} + \frac{1}{t_{a,2}}\right] . \tag{1.13}$$

At this stage a correction is made 'online' (internally) in the CSAT3 firmware to account for the influences of crosswind (see Liu et al. (2001)) on the measured speed of sound along each acoustic path (CSAT3, 2014). Subsequently the measured speed of sound is averaged over the three axes via $c_s = \frac{1}{3}\sum_{a=1}^{3} c_{s,a}$. Having sampled and computed the axes averaged speed of sound c_s, the sonic temperature T_s is calculated through the following expression (see e.g. Kaimal and Gaynor (1991))

$$T_s = \frac{c_s^2}{\gamma_d R_d} \,.$$

In the above, $R_d = 287.04\,\mathrm{JK^{-1}kg^{-1}}$ is the gas constant for dry air and $\gamma_d = c_{p,d}/c_{v,d}$ is the ratio of the specific heats of dry air at constant pressure ($c_{p,d}$) and constant volume ($c_{v,d}$). The sonic temperature is practically equivalent to the virtual temperature which is defined as the temperature required for a hypothetical dry air parcel to attain the same density, at the same pressure, as the given moist air parcel (e.g. Stull (1988)). In micrometeorology the distinction between sonic and virtual temperature, pursued in Appendix A.2, between the two is usualy kept (Kaimal and Gaynor, 1991) and so we will also keep the two distinct herein.

The CSAT3 returns the synchronized diagnosed velocities u, v, w (in $\mathrm{ms^{-1}}$) and sonic temperature (in $°\mathrm{C}$) at a sampling frequency, f_s, specified by the user;

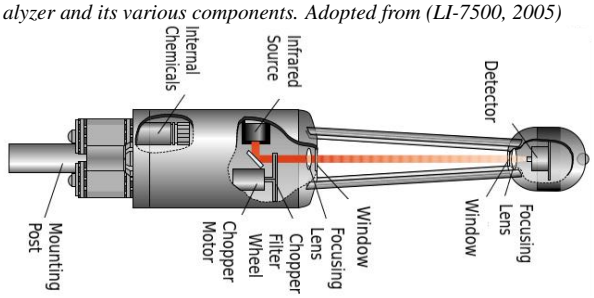
[10] We will use this notation throughout, we take $a \in 1(1)3$ to mean that index a runs from 1 to 3 with the increment of 1 given in the brackets.

typically either $f_s = 10$ or $f_s = 20$ Hz (CSAT3, 2014). At Bayelva the CSAT3 is set to return samples at $f_s = 20\,\text{Hz}$, that is twenty times per second corresponding to a response time (discrete timestep) $\Delta t = 1/f_s = 0.05\,\text{s}$. There is inherently some noise in the measured signals and consequently the CSAT3 has a measurement resolution, standard deviation with respect to a constant signal, of $1 \times 10^{-3}\,\text{ms}^{-1}$ for u and v, 5×10^{-4} for w and $2 \times 10^{-3}\,^{\circ}\text{C}$ for T_s (CSAT3, 2014). Despite its relatively old age, the CSAT3 is classed in the top tier of the sonic anemometers currently available on the market in terms of overall data quality (Mauder et al., 2006).

There are some issues with the CSAT3 and sonic anemometers in general that are worth highlighting. First of all the CSAT3 measurements are averages over the CSAT3 axes as opposed to strictly point values the implications of which we will return to in Section (2.5.2). Secondly, depending on the wind direction, the somewhat bulky CSAT3 structure may perturb the local wind field considerably causing undesired flow distortion (Aubinet et al., 2012), which we will adress in Section 2.3.1. In addition, if the acoustic paths of the CSAT3 are blocked in any way, for example as a result of heavy precipitation or frozen transducers, the sonic anemometer measurement principle fails completely (CSAT3, 2014). Events where the acoustic paths are blocked can usually be identified and removed via despiking routines such as those outlined in Section 2.2. Luckily for the case of Ny Ålesund the local climate is quite dry (Førland et al., 2012) so severe blockage events should be relatively infrequent. Still, as discussed there is a risk of orographically enhanced precipitation events at the Bayelva site.

LI-7500 Infrared Gas Analyzer

Figure 1.10: *Schematic of the LI-7500 open path infrared gas analyzer and its various components. Adopted from (LI-7500, 2005)*



The LI-7500 is a robust fast responding open path infrared gas analyzer that samples the densities, either as number density (mol m^{-3}) or as mass density (g m^{-3}), of CO_2 and water vapor in its measurement path (LI-7500, 2001). Here the term open path stems from the fact that the $12.5\,\text{cm}$ long infrared path, between the two windows in Figure 1.10, is open to the atmosphere. Herein we are primarily concerned with water vapor as opposed to CO_2 and so we will only discuss how the LI-7500 measures the absolute humidity ρ_v (i.e. water vapor mass density). Yet, the basic operating principle is much the same for CO_2 mass density (LI-7500, 2001).

Following LI-7500 (2005) we review the basic operating principle of the LI-7500. As shown in Figure 1.10 the infrared source emits an infrared radiation beam of a given intensity through a chopper filter wheel which is then focused by a lens before passing through a window and entering the open path. Having traveled across the open infrared path the infrared radiation passes through another window after which it is refocused by another lens and directed onto the detector which detects the attenuation (due to scattering and absorption) of the infrared beam across the open path. The rotating (9000 rpm) chopper filter wheel ensures that the peak intensity of the infrared beam alternates between four central wavelength bands (LI-7500, 2005): the reference band for CO_2 centered on $3.95\,\mu\text{m}$, the reference band for water vapor centered on $2.4\,\mu\text{m}$, the CO_2 absorption band centered on $4.26\,\mu\text{m}$ and the water vapor absorption band centered on $2.59\,\mu\text{m}$. By comparing the measured transmisivity of corresponding reference and absorption bands the effects of secondary beam attenuation[11], i.e. scatering and absorption from other sources, are accounted for so that the absorptivity, α_i, of the gas species in question is recovered. The correction for secondary attenuation sources also accounts for the cross sensitivity between the CO_2 and water vapor absorption bands. Having estimated the water vapor absorptivity, α_v, in the open path the absolute humidity is calculated internally in the LI-7500 firmware through (LI-7500, 2001)

$$\rho_v = M_v P f_v \left(\alpha_v s_v / P\right) . \tag{1.14}$$

In the above $M_v = 18.02\,[\text{g mol}^{-1}]$ is the molar mass of water vapor, P is the air pressure, s_v is a span adjustment term and f_v is a third order polynomial calibration function with unique coefficients for a given LI-7500 instrument provided by the manufacturer. More details on the origins of the function f_v are provided in the manual (LI-7500, 2001). The absolute humidity is sam-

[11] It is assumed that other particles responsible for scattering or absorption (e.g. dust or precipitation) are stationary with respect to the optical path for one rotation period of the chopper filter wheel (LI-7500, 2001). It is also assumed that due to their proximity the effects of secondary beam attenuation are the same for corresponding reference and absorption bands.

pled at a frequency set by the users of the EC system, for the case of Bayelva this was the same as for the CSAT3 namely $f_s = 20\,\text{Hz}$.

Being the most widely used open path IRGA model worldwide for the measurement of the turbulent fluxes of CO_2 and water vapor (Burba, 2013) the advantages and limitations of the LI-7500 are well known. In terms of advantages the design of the LI-7500 minimizes flow distortion through its small size and relatively smooth shape. In addition, contrary to closed path gas analyzers, an open path IRGA such as this is well suited for remote long term monitoring sites in that there is less need for maintenance (Leuning and Judd, 1996). Furthermore, the corrections for spectral attenuation of turbulent fluxes due to path length averaging are less extreme than for closed path gas analyzers (Lee et al., 2006) and so an open path EC system is well suited for the simpler analytical attenuation corrections of Massman (2000) which we cover in Section 2.5.2. The LI-7500 also functions well in conjunction with a CSAT3; at Bayelva the CSAT 3000 data logger provides synchronized, albeit not colocated (see Section 2.5.1), measurements from the two instruments.

There are a range of issues with the LI-7500 that are worth considering. First of all, as with the CSAT3, the measurement principle works poorly in the case of severe blockages of the measurement path, due to e.g. heavy precipitation. In such instances multiple assumptions in the LI-7500 beam attenuation correction breakdown and the samples are prone to spike. This effect is particularly severe in the case that dirt, droplets or frost covers the windows of the instrument over longer periods. Secondly, the zero reference levels of LI-7500 measurements are prone to drift over extended periods with large changes in the ambient temperature without maintenance (LI-7500, 2001). Both path blockage and zero drift lead to unphysical measurement values, which we will adress in Section 2.2. In that all open path analyzers effectively measure the density of gas species it is also necessary to apply a correction for the effects of fluctuating dry air density on the mean vertical wind speed (Webb et al., 1980) adressed in Section 2.5.4. The last concern is that the LI-7500 is not a truly passive instrument; it alters the properties of the air that is sampled as a result of the absorption of infrared radiation and subsequent heating. Thereby the LI-7500 is itself a source of convection and fluctuating densities, an effect that may be particularly severe in cold and stable conditions (Lüers et al., 2014). The correction proposed by Burba et al. (2008) may potentially alleviate this undesired heating effect. We will not pursue this correction since Lüers et al. (2014) found that the resulting correction yielded unphysically large flux values at the Bayelva site. It is worth pointing out that by tilting the LI-7500 45° from the vertical, as is done at Bayelva (cf. Figure 1.5), the effects of infrared path blockage and instrument generated convection are reduced (Lüers et al., 2014). The LI-7500 should not be tilted much more than this otherwise the flow distortion induced by the instrument, both on its own path and that of the CSAT3, becomes increasingly severe.

1.3 Motivation

The motivation behind our study are three problems that have been recognized through shorter term EC measurement campaigns (e.g. FINTUREX Sodemann and Foken (2005), EBEX-2000 Oncley et al. (2007),LIFTASS-2003 Mauder et al. (2006), and ARCTEX Lüers and Bareiss (2011)) as well as the long term global flux monitoring network FLUXNET (Baldocchi et al., 2001). The problems are: 1) The underestimation of nocturnal CO_2 respiration (Finnigan, 2008), 2) A failure to close the surface energy balance (Foken, 2008a) and 3) The shortcomings of MOST under stable stratification (Mahrt, 1999). These problems have potentially severe consequences for the universal applicability of the EC method in long term global networks for monitoring surface exchange (Finnigan, 2008).

More generally the problems are symptomatic of an oversimplifiied treatment of the atmospheric boundary layer, particularly the stable boundary layer which is relatively prevalent in the Arctic, with many of the issues extending into the domain of modeling. This is the case for both large scale numerical weather prediction and climate models (Holtslag et al., 2013), as well as more regional models (e.g. Aas et al. (2015)) and local turbulence resolving models such as Large Eddy Simulations (LES) (e.g. Beare et al. (2006)). Land surface schemes in all the former examples build heavily on the accumulated knoweledge from EC campaigns (Leuning et al., 2012). All of these 'schools' of models rely on observationally constrained boundary conditions and/or the validity of MOST. In the following we will outline the 3 problems presented above complimented with theory as well as proposed mechanisms leading to their occurence.

1.3.1 Shortcomings of Monin Obukhov Similarity Theory

We will introduce some observed defficiencies of Monin-Obukhov Similarity Theory, MOST, (after Monin and Obukhov (1954)), in the stably stratified surface layer following the discussion in Mahrt (1999). Before doing so it is worth demonstrating why MOST is an extremely attractive and widely employed concept. To do so we will consider only the turbulent kinetic energy (TKE) budget, however near identical arguments apply to the budgets of scalar variance (see e.g. Högström (1996)). It is woth keeping in mind that MOST in its most general form only applies to the surface layer, or constant flux layer, where vertical fluxes deviate by less than 10% from their surface value (Wyngaard, 2010).

For a horizontally homogeneous and statistically stationary surface layer, required for the application of MOST (Högström, 1996), with no subsidence and u aligned with the mean horizontal wind, the TKE budget becomes (Stull, 1988)

$$0 = \frac{g}{\overline{\theta}_v}\overline{\theta_v' w'} - \overline{u'w'}\frac{\partial \overline{u}}{\partial z} - \frac{\partial \overline{w'e}}{\partial z} - \frac{1}{\overline{\rho}}\frac{\partial \overline{p'w'}}{\partial z} - \epsilon\,, \quad (1.15)$$

where $e = \frac{1}{2}v_i'^2$ is the TKE per unit mass. From left to right the terms are the: buoyant production of TKE, mechanical production of TKE through the mean wind shear, turbulent transport of TKE, pressure transport of TKE and finally the viscuous dissipation of TKE into heat. Here it is assumed for friction velocity that $u_*^2 = -\overline{u'w'} = |\overline{u'w'}|$, i.e. that the wind stress is aligned with the mean wind, which is very nearly satisfied over land (Wyngaard, 2010). Next, to recover the M-O stability parameter, ζ, we multiply (1.15) by $\kappa z/u_*^3$, keeping the terms in the same order, yielding (Stull, 1988)

$$0 = -\zeta + \frac{\kappa z}{u_*}\frac{\partial \overline{u}}{\partial z} - \frac{\kappa z}{u_*^3}\frac{\partial \overline{w'e}}{\partial z} - \frac{\kappa z}{u_*^3\overline{\rho}}\frac{\partial \overline{p'w'}}{\partial z} - \frac{\kappa z}{u_*^3}\epsilon\,, \quad (1.16)$$

or more compactly (Högström, 1996)

$$0 = -\zeta + \phi_m(\zeta) - \phi_{tt}(\zeta) - \phi_{pt}(\zeta) - \phi_\epsilon(\zeta)\,. \quad (1.17)$$

When MOST is satisfied all the nondimensional ϕ terms are universal functions of ζ. As for practical implications this means that, with MOST satisfied and the universal functional forms known, we can diagnose the entire TKE budget using a single EC system. In fact only a sonic anemometer is necessary under the assumptions that: $T_s \simeq \theta_v$ and the $\phi_{pt} \simeq 0$. It is important to point out that these forms are not given by MOST itself, they must be determined by fast responding measurements; including vertical gradients (Kaimal and Finnigan, 1994). Previously, and still now to some extent, finding these universal functional forms was (is) the main concern of EC measurement campaigns (Foken, 2006). An extensive review of proposed universal forms is given in Högström (1996).

The value of MOST extends beyond diagnosing the TKE budget from EC measurements. If we focus on one term, ϕ_m, the nondimensional mean wind shear[12] which is defined as (cf. (1.17) and (1.16))

$$\phi_m(\zeta) = \frac{\kappa z}{u_*}\frac{\partial \overline{u}}{\partial z}\,. \quad (1.18)$$

Similarly, for a generic scalar ξ, for which the associated surface layer scale is $\xi_* = -\overline{\xi' w'}/u_*$, the definition of the nondimensional vertical gradient is (Wyngaard, 2010)

$$\phi_\xi(\zeta) = \frac{\kappa z}{\xi_*}\frac{\partial \overline{\xi}}{\partial z}\,.$$

The above relationships are invaluable provided that 1) the universal form is known and 2) either the surface layer scale or the vertical profile are known. Typically in a model, given a known univeral form, either the surface layer scale (being simply related to the surface flux) is given by a boundary condition or the vertical profile is sufficiently well resolved. In the case that the surface layer scale is known this makes it possible to solve, via integration, for the vertical profile below the lowest grid point, provided that this point is in the surface layer. From the point of view of weather forecasting this is invaluable in that the first model grid level is often at several 10s (or even 100s) of meters above the ground and one may wish to obtain a reliable estimate of 2 meter air temperature or wind speed. Conversely, if the vertical profile is well resolved, it is straightforwards to combine this with the univeral form of the nondimensional gradient to solve for the surface layer scale and thus retrieve the surface flux.

These so called flux-profile relationships that we have just described are among the most widely used versions of first order closure (Stull, 1988) in modeling. The term first order refers to the fact that the relationships provide the means to parametrize second order terms (e.g. $\overline{u'w'}$) that are usually not solved for by the prognostic equations in models, but are nonetheless needed by the prognostic equations of the first order terms (Wyngaard, 2010). Additionaly, these flux profile relationships can be used in measurement campaigns to indirectly estimate the surface flux using slower measurements provided that both the horizontal wind speed and scalar(s) in question are measured as vertical profiles, at least two near lying levels, in the surface layer

[12] Including the von Karman constant κ in the definition of ϕ_m is due both to tradition and for convenience (Högström, 1996). For neutral conditions $\phi_m(\zeta = 0) = 1$ and the classic 'law of the wall' formula $\overline{u}(z) = \frac{u_*}{\kappa}\ln\left(\frac{z}{z_0}\right)$ is recovered where z_0 is the roughness length at which $\overline{u}(z_0) = 0$ by defintion (Stull, 1988).

(Foken, 2008b). Since we do not have access to any vertical profiles we can not control the universal functional forms proposed in e.g. Högström (1996). The exception is the nondimensional dissipation ϕ_ϵ in that ϵ may be estimated in the inertial subrange identified through spectral methods or structure functions. MOST also extends beyond flux-profile relationships. We can and will investigate other variants of MOST namely flux variance similarity (Section 2.6.5) (Wyngaard et al. (1971) and Tillman (1972)) as well as spectral and cospectral similarity theory (Kaimal et al. (1972) and Wyngaard and Coté (1972)).

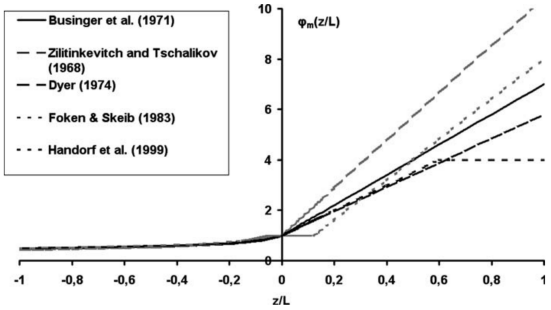


Figure 1.11: *Universal functional forms for nondimensional mean wind shear, $\phi_m(\zeta)$, in the stability domain $\zeta = z/L \in [-1, 1]$ proposed by various investigators (references in the legend). Figure adopted from Foken (2006).*

Having demonsrated the appeal and widespread use of MOST we are now prepared to deal with its observed shortcomings under stable stratification. Evidence of problems in stable stratification is provided through the large scatter between the many proposed universal functional forms in this regime (Högström (1996), Mahrt (1999), Klipp and Mahrt (2004) and Foken (2006)). An example of such scatter is shown for $\phi_m(\zeta)$ in Figure 1.11 where on the unstable side ($\zeta < 0$) the proposed forms for $\phi_m(\zeta)$ are all in good agreement collapsing into a single curve whereas on the stable side ($\zeta > 0$) there is considerable disagremeent. Furthermore, from this Figure it is clear that the disagreement becomes increasingly severe as ζ approaches very stable stratification ($\zeta > 1$).

This begs the question: why is there such disagreement in the proposed universal functions in stable, and particularly very stable, stratification? Essentially the disagremeent boils down to the underlying assumptions in the MOST formulation often being invalid in stable regimes. The three violations along with example physical mechanisms can be summarized as follows:

- Surface layer turbulence can be highly intermittent (i.e. nonstationary) (Andreas et al., 2008). For example long calm periods may be interrupted by sporadic downwards bursts of turbulence from aloft after mean wind shear builds up during the formation of a low level jet (Stull, 1988).
- Flows in the surface layer may become horizontally heterogeneous due to for example: gravity wave propogation (Mahrt, 2010) and drainage flows (Aubinet, 2008).
- The surface layer may become so shallow that it exists well below the observation level. Thereby there is signifcant vertical flux divergence at the observation level and the MOST formulation breaksdown. Frequently a shrinking of the surface layer occurs through longwave radiative flux divergence (Mahrt, 1999).

Several more mechanisms leading to these violations are discussed in Mahrt (1999).

Still, MOST can be applied succesfuly to the stable surface layer so long as the above mechanisms are absent. Care must be taken interpreting results, however, as MOST can appear to be succesful due to the self-correlation inherent in its formulation as discussed in Mahrt (1999) and Klipp and Mahrt (2004). To elaborate we continue with the nondimensional shear ϕ_m as our example. From its definition in (1.18) we note that it is inversely proportional to the friction velocity u_*. Now as ζ, which ϕ_m is a function of when MOST is satisfied, has u_*^3 in its denominator this presents somewhat of a problem. The problem lies in the fact that even in the case where $\frac{\partial \overline{u}}{\partial z}$ and the heat flux (in the nominator of ζ) share no physical correlation, as we might expect in very stable conditions with weak turbulent mixing, ϕ_m and ζ will still be correlated due to the occurence of u_* as a common divisor. As shown in Klipp and Mahrt (2004) accounting for this self correlation, and removing instances where the physical (residual) correlation is small, reduces the scatter in flux-profile relationships for the stable surface layer considerably. Nonetheless, this means that periods when turbulence is weak or highly intermittent which were traditionally included in the formulation of MOST universal forms, leading to apparent success in the stable regime, must now be removed.

The fact that MOST demands an idealized horizontally homogeneous and statistically stationary surface layer turns out to be an advantage for us in that we can use deviations from MOST predictions to identify departures from this idealization. Since such a departure means that using the EC method to diagnose surface layer fluxes is strictly not possible, this becomes a valuable quality control tool (Foken and Wichura, 1996). It is nonethelesss somewhat alarming that most models still, for lack of a better alternative, use MOST formulations in the very stable surface layer where it may not be valid (Mahrt, 1999). As will be shown, departures from this idealized surface layer will be a recurring theme in our motivation.

1.3.2 Nocturnal CO_2 respiration

One of the primary goals of FLUXNET is to provide a global database containing reliable estimates of the exchange of CO_2 between the biosphere and the atmosphere at the ecosystem scale across a range of different ecosystems (Baldocchi et al., 2001). A fairly general form (cf. Lee and Massman (2011) and Gu et al. (2012)) of the block average CO_2 budget for the air column between the surface and a fixed measurement height H is given by

$$\overline{\varrho_c w}\big|_{z=H} + \int_0^H \frac{\partial \overline{\varrho_c}}{\partial t} + \nabla_H \cdot (\overline{\varrho_c \mathbf{u}})\, dz = \int_0^H \frac{\overline{S}_c}{M_c}\, dz\,. \tag{1.19}$$

Where $M_c = 0.044\,\mathrm{kgmol}^{-1}$ is the molar mass of CO_2, ϱ_c (units mol m^{-3}) is the molar density of CO_2 related to the CO_2 mass density ρ_c via $\varrho_c = \rho_c/M_c$, $\overline{S}_c$ (units $\mathrm{kgm}^{-3}\mathrm{s}^{-1}$) is the point source strength of CO_2 and $\nabla_H \cdot (\overline{\varrho_c \mathbf{u}})$ is the horizontal advective flux divergence of CO_2. The vertical integral on the right hand side is the so-called Net Ecosystem Exchange (NEE) (Aubinet et al., 2012) of CO_2 with units of molar flux density (mol $\mathrm{m}^{-2}\mathrm{s}^{-1}$). Effectively the NEE is a measure of the ecosystem balance between the CO_2 that is released through respiration and taken up by photosynthesis below the measurement height. If the NEE is positive there is a net release of CO_2 by the biosphere (including the soil) to the atmosphere and vice versa if the NEE is negative. Typically (e.g. Lüers et al. (2014)) the block averaged NEEs are considered over the course of a year which, when gaps in the data are filled, are combined to yield an annual NEE estimate (units $\mathrm{molm}^{-2}\mathrm{yr}^{-1}$).

In line with our discussion in Section 1.2.1 we note that in the special case of a horizontally homogeneous and statistically stationary surface layer (1.3.2) becomes simply

$$\overline{\varrho_c}\,\overline{w}\big|_{z=H} + \overline{\varrho_c' w'}\big|_{z=H} = \int_0^H \frac{\overline{S}_c}{M_c}\, dz\,. \tag{1.20}$$

Crucially in such a case it is possible to diagnose the NEE based on measurements from a single EC system. In the case that vertical profiles of the CO_2 density are also available, statistical stationarity is no longer required in that it is possible to estimate the change in CO_2 storage in the column, i.e. the $\int_0^H \frac{\partial \overline{\varrho_c}}{\partial t}\, dz$ term in (1.3.2). Nonetheless, from a single EC tower the horizontal advection term, $\int_0^H \nabla_H \cdot (\overline{\varrho_c \mathbf{u}})\, dz$ in (1.3.2), is immeasurable (Lee and Massman, 2011).

Since most FLUXNET sites (cf. Baldocchi et al. (2001)), including Bayelva, consist of measurements from a single tower structure the immeasurability of the horizontal advection term poses a serious problem. In particular at sites located on sloping terrain where strong horizontal advection may occur through drainage flows (Aubinet, 2008). Moreover many sites do not have vertical profile measurements of ϱ_c available and so diagnosing the change in the CO_2 storage is not possible. It turns out that both horizontal advection and the change in storage become particularly important in the persence of stable stratification, typically occuring at night, where turbulent mixing is weak (Massman and Lee (2002) and Finnigan (2008)). In such a regime it becomes impossible to arrive at a reliable estimate of the block averaged NEE via single tower EC measurements as the dominant terms in the NEE balance are not accounted for. This leads to an underestimation of the nighttime CO_2 respiration and is one of the greatest obstacles in using the EC method to arrive at long term estimates of NEE (Aubinet et al., 2012).

As an example consider the case of a stably stratified night at a vegetated site with positive NEE at the foot of a mountain with weak turbulent mixing and initially no horizontal advection. Over the course of the night CO_2 will accumulate in the column below the measurement complex. If at some point a katabatic wind forms, accelerating down the mountain, the CO_2 accumulated in the column below the measurement height would be advected downwind. In such a case any NEE estimated through (1.20) would be zero even if the true NEE given by may be considerable. Note that this is just one of a myriad of examples that can be constructed to highlight the problem with (1.20) under stable stratification. As pointed out in Aubinet et al. (2012), if a horizontal advection event never occurs during the night then as soon as we transition into a more unstable daytime regime and turbulence becomes more developed the accumulated CO_2 in the column below would rapidly mix upwards. In such a case the diurnal NEE would not be significantly underestimated.

It is not advisable to simply assume that no advection occurs during stable nocturnal regimes (Aubinet, 2008). A common but criticized workaround (e.g. Finnigan (2008)) is to combine a filter that discards periods with low turbulent mixing. Typically low values of the friction velocity, $u_* = \left(\overline{u'w'}^2 + \overline{v'w'}^2\right)^{1/4}$, are used to identify such periods. NEE estimates are then discarded for blocks where u_* falls below some minimum threshold. Subsequently gaps in the block averaged NEE estimates are filled by selecting an appropriate gap filling procedure (see Aubinet et al. (2012)). Such a method is nonetheless prone to bias the estimates since the u_*-filter may fail to isolate periods with strong horizontal advection (Finnigan, 2008), the u_* threshold is highly site specific (Massman and Lee, 2002) and gap filling will at best provide a reasonable approximation of the true nighttime NEE evolution.

Finishing off this section it is worth underlining that

the nocturnal CO_2 flux problem is not our primary concern. Yet the mechanisms involved, the nonstationarity and horizontal heterogeneity of surface layer flows in stable stratification, also lead to problems closing the surface energy balance (Foken, 2008a) and these issues are our primary focus. For example Wilson et al. (2002) found, across a range of sites, that the nocturnal CO_2 respiration estimated via the EC method was signifcantly smaller the larger the surface energy imbalance discussed in the next section. As such, much of our methodology and findings may also be applicable to the nighttime CO_2 respiration problem.

1.3.3 Closure of the Surface Energy Balance

As of yet FLUXNET does not, to our knowledge, include any seaborne towers (cf. Figure 1.3) so we will restrict our attention to the surface energy balance (SEB) over land. To start off we consider the energy balance of a thin ground layer of thickness δz that extends from below the surface $z = -\delta z$ to the surface[13] itself ($z = 0$), which reads as follows (e.g. Stull (1988), Foken (2008b))

$$Q_R^* = Q_G + Q_E + Q_H + \Delta Q_S + Q_A \,. \qquad (1.21)$$

Each term in (1.21) has units of energy flux (Wm^{-2}) and all terms, except ΔQ_S, represent a flux of energy across one of the vertical boundaries of our layer δz. Q_R^* is the global (or net downward) radiative flux at the surface defined as the difference between the net downward shortwave radiation, $S_{net}^{\downarrow}$, and the net upward longwave radiation, $L_{net}^{\uparrow}$, i.e.

$$Q_R^* = S_{net}^{\downarrow} - L_{net}^{\uparrow} = S^{\downarrow} + L^{\downarrow} - S^{\uparrow} - L^{\uparrow} \,. \quad (1.22)$$

In particular: $S^{\uparrow}$ and $S^{\downarrow}$ are the upward (reflected) and downward shortwave (solar) radiative fluxes at the surface, while $L^{\downarrow}$ and $L^{\uparrow}$ are, respectively, the downward longwave flux emitted from the atmosphere to the surface and the upward longwave radiative flux emitted by the surface. Q_E and Q_H are the latent and sensible heat fluxes *at the surface* $z = 0$ where latent ('hidden') heat is associated with phase changes of water in the processes of evaporation/condensation/freezing/melting/-sublimation/deposition occuring at the surface, while sensible ('felt') heat is associated with the molecular conduction of heat from (or to) the surface (see (Sun et al., 1995)). Moreover, Q_G is the ground (be it soil, snow or permafrost) heat flux at $z = -\delta z$ related to the conduction of heat in the ground from above or below (Westermann et al., 2009). With the exception of the global radiation, all the heat fluxes are defined as positive when the flux is directed *away* from the layer δz. So Q_E and Q_H are positive when these transport heat from the surface to the atmosphere and Q_G when it transports heat further down into the ground.

The storage term ΔQ_S is principally associated with heating or cooling of the ground layer (Oncley et al., 2007) as well as the air (and canopy if present) when the layer is extended into the atmosphere. If the surface is vegetated then the radiative energy that is expended through photosynthesis or the heat that is released through respiration is often included in the storage term (Leuning et al., 2012). The Q_A term accounts for additional usually minor sources/sinks of energy unaccounted for by the other fluxes or the storage term. Some examples are anthropogegenic heat fluxes Q_F from buildings (Nordbo et al., 2012) or instruments (Burba et al., 2008), warming/cooling of the surface by precipitation Q_P, warming/cooling of the layer through the runoff (Aas et al., 2015) of water Q_W as well as energy expended in melting snow Q_M (Westermann et al., 2009). The snow melt can have a considerable influence on the surface energy balance in the melting season. During the snowmelt in June 2008 at Bayelva Westermann et al. (2009) estimated that the energy expenditure was the equivalent of an average flux of $27\,\mathrm{Wm}^{-2}$ away from the ground layer in the same period.

It is worth emphasizing that the SEB (1.21) is a reformulation of the first law of thermodynamics (Wilson et al., 2002). In particular the SEB is aquired via vertical integration (from $z = -\delta z$ to $z = 0$) of the thermodynamic energy equation with all horizontal heat transport terms in the layer neglected. A detailed derivation is given in Sun et al. (1995); albeit for a layer from the surface $z = 0$ to a height $z = \delta z$ in the ASL for which several more terms must be included. In words (1.21) is simply an expression for the conservation of energy in the ground layer: if there is a net convergence of energy into the layer ($Q_R^* - Q_G - Q_E - Q_H - Q_A > 0$) energy must be stored in the layer ($\Delta Q_s > 0$), whereas if there is a net divergence of energy out of the layer ($Q_R^* - Q_G - Q_E - Q_H - Q_A < 0$) then energy must be released by the layer ($\Delta Q_s < 0$).

Often (e.g. Stull (1988)) the SEB is also presented conceptually for a layer that is infinitisemally thin ($\delta z \to 0$), resulting in a near zero heat capacity and, consequently, a storage term that is negligible $\Delta Q_S \simeq 0$. For such a layer the SEB becomes, upon ignoring additional sources/sinks of heat ($Q_A = 0$), simply

$$Q_R^* = Q_G + Q_E + Q_H \,. \qquad (1.23)$$

Unfortunately the representation in (1.23) is not very

[13] To dispell any doubt we consider the surface to be the point above which we transition to the atmosphere from a mainly solid or liquid medium (or vice versa).

useful in that all the terms on the right hand side are not directly measurable. As are both the surface latent heat and sensible heat fluxes as well as the ground heat flux in (1.21) unless several simplifying assumptions are satisfied.

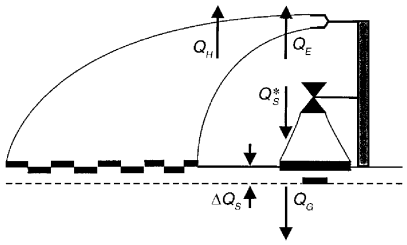


Figure 1.12: *A typical SEB measurement complex consisting of an EC system to measure the turbulent fluxes of sensible heat Q_H and latent heat Q_E in the surface layer, a ground heat flux plate to measure Q_G below the surface and a radiation sensor to measure global radion Q_R^* (denoted Q_S^* in the figure) below the EC system. The footprints of the EC system and the upwards radiation are shown by the leftmost horizontal bars and the thick rightmost horizontal bar respectively. Figure adapted from Foken (2008*b*).*

In practice the components of the SEB are measured by sensors removed from the surface (Figure 1.12). For radiation this is rarely a problem in that, in the absence of fog, the thin (typically $0.5\,\mathrm{m}$) layer between the radiometer and the surface should be practically transparent to both shortwave and longwave radiation. Thereby the measured global radiation is, in the absence of large measurement errors, for all practical purposes equal to Q_R^* in (1.21) or (1.23) (Foken, 2008a).

For the ground heat flux the temperature gradient in the ground is calculated either via a heat flux plate (Figure 1.12) or a vertically displaced array of temperature sensors (Westermann et al., 2009) below the surface. The thermal properties of the ground are assumed to be constant in time and space such that Fourier's law of heat conduction, with the vertical profile of ground temperature, can be solved numerically to calculate Q_G at the uppermost ground temperature measurement level (see Appendix A in Westermann et al. (2009) for details). As underlined in both Wilson et al. (2002) and Westermann et al. (2009), the thermal properties of the ground do vary in both time and space due to changes in the ground temperature, soil moisture and properties of snow. Nevertheless Westermann et al. (2009) points out that for the March 2008-March 2009 campaign at Bayelva the contribution of the error in ground heat flux estimation towards the total error in the SEB estimation is small relative to other terms as a result of the generally small magnitude of this flux. It is worth emphasizing that the uppermost ground temperature measurement is not the surface itself and consequently there will be a storage component in the ground above (Westermann et al., 2009). Therefore, a direct application of (1.23) is questionable particularly on time scales of less than a day when ground heat storage can be large (Leuning et al., 2012).

The estimation of the fluxes of sensible and latent heat at the surface via measurements is arguably the most complicating factor in diagnosing the SEB in (1.21) or (1.23) (cf. Wilson et al. (2002), Foken (2008a) and Leuning et al. (2012)). In practice these fluxes are estimated through the EC method via fast responding measurements in the surface layer that diagnose their 'turbulent' counterparts. In particular for a single EC measurement system, as is the most frequent setup in FLUXNET (Wilson et al., 2002) and the case at Bayelva, after a series of corrections the fluxes of sensible and latent heat at the measurement level are directly equated to those at the surface. Recall that this is done by assuming an idealized 'constant flux' surface layer that is both statistically stationary and horizontally homogeneous (Foken and Wichura, 1996). Due to the underlying assumptions on surface layer turbulece (Leuning et al., 2012), limitations of the instruments (van Dijk, 2002), complexity of corrections (Mahrt, 2010) and changes in the flux footprint (Schmid, 1994) errors, both random and systematic, are present (Finkelstein and Sims, 2001). These exist not only in raw data but are introduced and propogated through EC data processing before the final flux estimates are produced (Billesbach, 2011).

As a result of the many difficulties and assumptions involved in estimating the components of (1.21) there is almost always (Wilson et al., 2002) a residual in local SEB budgets derived from measurements. Thereby, in the context of an experimentally estimated SEB budget with each component determined at a measurement height removed from the surface (1.21), ignoring additional sources/sinks, should instead be expressed as (Foken, 2008a)

$$Res + Q_H + Q_E = Q_R^* - Q_G - \Delta Q_S\,, \qquad (1.24)$$

where Res is the SEB residual that is required to close the measured SEB budget. From the point of view of a turbulent surface layer one interpretation of (1.24) is that the left hand side of the equation, less the residual, represents the realized energy while the right hand side represents the available energy (Leuning et al., 2012). So ideally, in the case of a balanced measured surface energy budget, the sum of the turbulent fluxes should equal the available energy. Conversely if a residual is present then either the available energy is overestimated or the realized energy is underestimated. Typically the relative magnitude of the SEB residual is by no means negligible. On average at FLUXNET sites the energy imbalance, defined as the residual normalized by the available energy, is on the order of 20% (Wilson et al., 2002). As will be shown (Left panel Figure 1.14) for Bayelva we find that the imbalance is above 50% on average for hourly fluxes from March 2008-March 2009.

The widespread existence of an energy imbalance, the so-called SEB closure problem (Foken, 2008a), is arguably the most long-standing problem in the field of micrometeorology (Foken (2008a), Leuning et al. (2012), Foken et al. (2011) and Stephens et al. (2012)). A variety of mechanisms explaining such an imbalance have been proposed. Following the reviews in Leuning et al. (2012) and Foken (2008a) the main mechanisms proposed in explaining the lack of SEB closure in a typical 30 minute flux averaging period are:

1. Errors in the global radiation measurements (Foken, 2008a).
2. Lacking or erroneous diagnosis of the storage terms and the ground heat flux (Leuning et al., 2012).
3. Deviations from the statistically stationary ASL that is assumed when employing the EC method for an isolated measurement complex (Foken and Wichura, 1996).
4. Significant horizontal advection that is immeasurable from a single EC system (Aubinet (2008) and Paw et al. (2000)).
5. Low frequency, i.e. large eddy, contributions to the turbulent transport that may not be captured in a typical flux averaging period (Lenschow et al. (1994), Finnigan et al. (2003) and Charuchittipan et al. (2014)).
6. Loss of flux through cospectral attenuation (Moore (1986), Horst (1997) and Massman (2000)).
7. A reference frame that is poorly suited for diagnosing surface exchange (Wilczak et al. (2001), Finnigan (2004) and Mahrt (2010)).
8. Incosistent measurement footprints. As shown in 1.12 the footprint of each of the measured flux components in the SEB may not only be different in scale and position but also vary in time (Schmid, 1994).

Mechanisms 1. and 2. are related to measurement problems concerning the SEB components that are not estimated via the EC method and it is acknoweldged (Foken, 2008a) that generally these mechanisms are too small to explain the energy imbalance. As noted in Leuning et al. (2012) this especially the case when the averaging time is extended from the typical half hourly scale to the diurnal scale. Mechanisms 3., 4. and 5. revolve around the meteorology of the turbulent ASL not satisfying the basic assumptions allowing for the use of the EC method from single towers in particular the assumption of *ergodicity* (Lee et al., 2006). Mechanisms 6. and 7. are arguably the largest potential sources, even in ideal conditions, of systematic error when using the EC method to diagnose local surface exchange (Leuning et al., 2012). The final mechanism (8.) related to the footprint of measurements is a significant barrier in evaluating a local SEB budget. Nonetheless, given that horizontal homogeneity, a regime under which a near uniform footprint is assured, is required for horizontal advection to be negligible the origin of mechanisms 4. and 8. are closely related (Foken, 2008a). Herein we are primarily concerned with mechanisms 3. through 7. (and 8. indirectly) as it is these that we can identify, and at times rectify, via careful application of the EC method. On account of the above discusion in general it is widely held (Foken (2008a) and Leuning et al. (2012)) that as opposed to the magnitude of the available energy being overestimated it is the magnitude of the turbulent fluxes that are underestiamted. Identifying the reasons for such an underestimation is one of the main concerns of the ensuing work.

As a starting point we analyzed the results of Westermann et al. (2009) in an attempt to diagnose the SEB closure at Bayelva for the period March 2008-March 2009. These results are based on hourly fluxes of: latent heat and sensible heat computed through the TK2 package (Mauder and Foken, 2004), global radiation from the nearby Ny Ålesund Baseline Surface Radiation Network (BSRN) site (Ohmura et al., 1998) and ground heat computed through either the bulk or conduction method (see Westermann et al. (2009) for details). The storeage term ΔQ_S was not measured directly in this campaign, neither were any additional sources/sinks Q_A. However, as previously mentioned Westermann et al. (2009) estimated that the energy expenditure in melting snow amounted to an average flux $Q_M = 27\,\mathrm{Wm}^{-2}$ in the month of June. Consequently in our analysis we combined the residual with the unresolved storeage term and ignored any additional sources apart from the snow melt term. In line with typical presentations of SEB regimes we calculated the diurnal SEB averaged over four seasons consisting of the three months of December-January-February (DJF), March-April-May (MAM), June-July-August (JJA) and September-October-November (SON) respectively. The result of this preliminary analysis of these four 'seasons' is displayed in the respective panels of Figure 1.13. It is worth emphasizing that this definition of seasons may not be particularly relevant to this high Arctic site; six different seasonal regimes are presented in Westermann et al. (2009) and these are also employed in Aas et al. (2015). Even so, we felt that a more traditional presentation of the seasons would aid the reader in attaining a grasp of the regimes of the SEB at Bayelva when compared to lower latitudes.

From the upper left panel in Figure 1.13 we note that the polar night (DJF) SEB is dominated by a longwave radiative cooling of the surface, with $L^{\uparrow}_{net} < 0$ and

$S^{\downarrow}_{net} = 0$, which is mainly compensated by a negative sensible heat flux, $Q_H < 0$, in conjunction with a weaker negative ground heat flux heating the surface. On account of the low temperatures during the polar night the air is necessarily very dry and the latent heat contribution towards the SEB is (on average) negligible. The polar night regime is predominately stably or neutraly stratified, on account of $Q_H < 0 \longrightarrow \zeta > 0$, and a diurnal cycle is practically nonexistant.

With the onset of spring (MAM) shown in the upper right panel of Figure 1.13 a diurnal cycle becomes clear with the shortwave radiative heating of the surace, $S^{\downarrow}_{net}$, being the dominant component near noon local time. This heating is offset by longwave radiative cooling throughout the day. Conversely throughout the day the sensible heat flux heats the surface, compensating for longwave cooling during the night. Moreover the ground heat flux compensates for the global radiation throughout the diurnal cycle: cooling the surface during sunlight hours and warming the surface during the night. The latent heat flux is still negligible. During these spring months the stability is on average either stable or near neutral throughout the day.

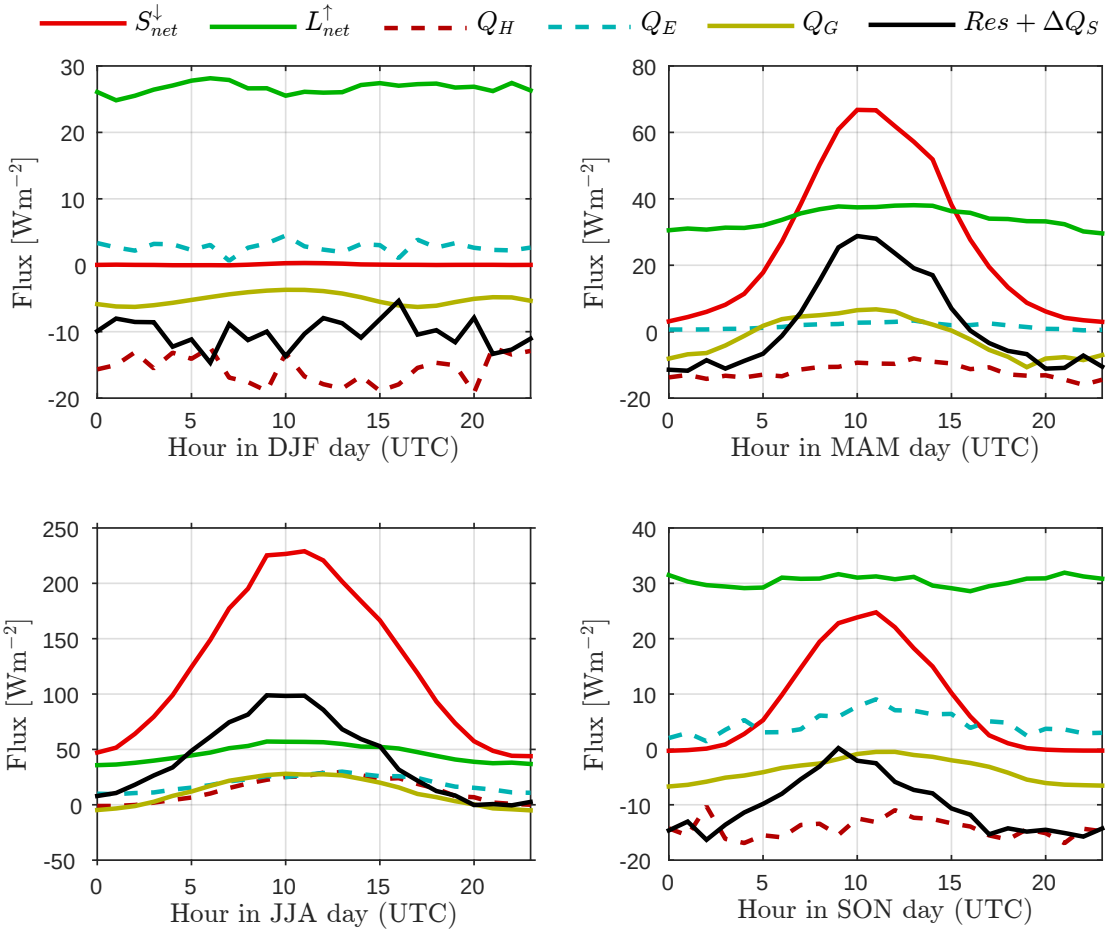


Figure 1.13: *Diurnal SEB regimes in the four traditional seasons at Bayelva from March 2008-March 2009. The panels display the hourly fluxes (in Coordinated Universal Time: UTC) for the course of a day averaged over all available hours in the months: DJF (top left), MAM (top right), JJA (bottom left) and SON (bottom right). Color coding given in the legend with symbols as defined in the main text.*

Once summer (JJA), or polar day, along with snowmelt arrives the diurnal cycle is even more visible as seen in the lower left panel of Figure 1.13 with an even larger peak in the net downwards shortwave radiation. By definition during polar day the sun never sets. Thus, at Bayelva the global radiation is on average positive throughout the polar day. During this regime the nonradiative fluxes are all (nearly) positive throughout the 24 hour period, compensating for the radiative heating of the surface, with these three fluxes attaining a maximum early in the afternoon after the global radiation has attained its maximum. Clearly there is a slight lag in the response of the nonradiative fluxes to the global radiation. The stratification is, on average, either unstable or neutral throughout the polar day ($Q_H \geq 0$) and uniquely during this season the Bowen ratio $Bo = Q_H/Q_E$ (e.g. Stull (1988)) is near unity.

Finally with the arrival of autumn (SON) depicted in the lower right panel of Figure 1.13 the diurnal cycle is no longer as pronounced and the global radiation is, on average, negative throughout a full day. Longwave radiative cooling and a positive latent heat flux make up for shortwave radiative and sensible heating of the surace during daylight hours while the sensible heat flux and ground heat flux heat the surface compensating for longwave cooling during the night. Already during autumn (SON) snow begins to accumulate which continues through winter (DJF) usually reaching a maximum during the spring months (MAM) (Winther et al., 2002). As with the polar night and spring the stratification is, on average, stable or neutral throughout a 24 hour cycle in autumn.

We have purposely not yet discussed the black curve depicting the sum of the SEB residual and storeage term, $Res + \Delta Q_S$, for any of the panels in Figure 1.13. By inspection this residual storage term is far from negligible, on average, in any of the seasons. Following the phase lag conjecture in Leuning et al. (2012) that any energy stored in the soil layer during the day should be released by late afternoon then if the residual were near zero we would in the presence of phase lag expect the sum of the residual storage term to be zero over the course of a full day. By inspection of Figure 1.13 we see that this is indeed the case during spring (MAM) so here we can perhaps attribute the energy imbalance, with the storeage term unresolved, to this phase lag mechanism. During the remaining months, however, the residual storage term is practically single signed throughout the day, negative in DJF and SON but positive in JJA, so such a mechanism, with the storeage term unresolved, does not provide a compelling explanation for the energy imbalance in these seasons. We propose that the unseemliness of a phase lag explanation at our site can be directly attributed, in particular during autumn and winter, to the unusual prevelance of stable stratification at Bayelva on top of a weak diurnal cycle during the polar night and its shoulder months.

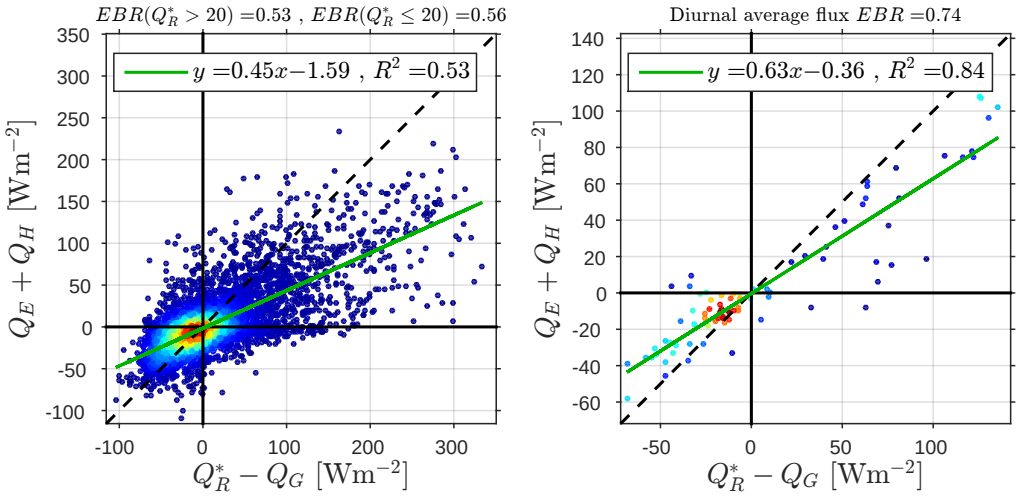


Figure 1.14: *Scatter plots of net turbulent flux, $Q_E + Q_H$, versus available energy, $Q_R^* - Q_G$, both in units of Wm^{-2} for the period March 2008-March 2009 at Bayelva. Left Panel: Hourly averaged fluxes: 6156 hourly averages in total . Right panel: Diurnal average of the hourly averaged fluxes: only days with at least 20 hourly averaged fluxes are included amounting to a total of 81 diurnal averages. In both panels the dots represent the individual flux values with reds indicating a high density of dots and blues a scarce density, the dashed black line is the $1:1$ line corresponding to zero residual and zero storage, the green solid line is the linear best fit. The results of the linear regressions, the equation for the linear best fit (slope and intercept) as well as the correlation factor, are displayed in the legends.*

Diving a bit deeper into an analysis of the local SEB closure at Bayelva we followed the approach of Wilson et al. (2002) in applying a linear regression to the results of Westermann et al. (2009) as displayed in Figure 1.14. Through the regression we calculate how much of the turbulent flux can be described as a linear function of the available energy, less the unresolved storage, of the form $y = mx + c$ where m is the slope and c is the intercept of the turbulent flux as a least squares fit to the available energy. For the case of a closed SEB, ignoring the storage term, the form is $y = x$; i.e. a slope of one with zero intercept (Wilson et al., 2002). In addition we calculate the correlation coefficient, R, to diagnose how linear the relationship between the sum of the realized energy (sum of turbulent fluxes) and the available energy is. In the case $R^2 = 1$ the two share a perfect linear relationship (Stull, 1988) in this case we expect $R \to 1$, i.e. a strong positive linear correlation, for a closed SEB.

As in Wilson et al. (2002) we also calculate the energy

balance ratio (EBR) defined as

$$EBR = \frac{\sum_{i=1}^{I} (Q_E + Q_H)_i}{\sum_{i=1}^{I} (Q_R^* - Q_G)_i}$$

where I is the number of flux estimates considered. So the EBR is simply the ratio of the accumulated realized energy to the accumulated available energy where in our case we ignore the unresolved storeage. In particular, the absolute deviation of the EBR from unity serves as a good estimate for the overall systematic error at a given site in terms of SEB closure (Mauder et al., 2013). Following the suggestion of Mauder et al. (2013) for hourly averaged fluxes we separate the EBR into one 'nocturnal' ($Q_R^* \leq 20$) and one 'daylight' ($Q_R^* > 20$) value. The purpose of this seperation is to eliminate, on the premise that the realized energy is underestiamted, the inadvertent cancelation of realized energy biases occuring when adding together the negative biases during the day and the positive biases during the night. To clarify, if the nocturnal and daytime biases are near equal in magnitude but opposite in sign, as may be the case, the returned EBR would falsely be near unity when we are actually far off reliable estimates at the hourly resolution. When calculating the EBR we also subtracted the average flux of energy expended by melting snow, Q_M, as estimated by Westermann et al. (2009) from each entry in June.

From the left panel in Figure 1.14 we note that although the intercept is near zero the slope of the least squares fit (green line) indicates that for (6156) hourly averaged fluxes the realized energy is typically less than half the available energy. Though this may seem an extereme underestimation it is well within the bounds observed at other FLUXNET sites (Wilson et al., 2002). Moreover, both from the overall scatter and the value of the square correlation coefficient we note that the two are far from being perfectly correlated, at times the realized energy greatly exceeds the available energy. Still, the positive correlation is considerable indicating that there is a systematic underestimation of the magnitude of the turbulent fluxes with regards to the available energy at Bayelva. An underestimation is coroborated by the fact that, even when the snow melt is accounted for, both the nocturnal and daytime EBRs are only on the order of 50%.

Pursuing the phase lag conjecture of Leuning et al. (2012), we considered diurnal averages of the hourly fluxes as shown in the right panel in Figure 1.14. It is worth emphasizing that in an attempt to calculate representative diurnally averaged fluxes we only considered days in which at least 20 hourly flux estimates were available. Enforcing this constraint we had a total of only 81 diurnal averages, as opposed to (roughly) 250 estimates we would get without the constraint. As such, the resultant regression is by no means representative of an entire year at Bayelva. By default the days with few flux estimates, due to quality issues in the data, where we would expect a poor SEB closure are not shown in Figure 1.14. So the regression presented is somewhat optimistic. Nonetheless we note that, at least for the days with enough hourly flux values, the effect of considering the diurnally averaged SEB greatly improves the closure compared to that for hourly averaged fluxes. Not only is there signifcantly less scatter, with a square linear correlation $R^2 = 0.84$, the slope of the least squares fit has increased from 0.45 to 0.63 while the intercept is still near zero. Furthermore, the diurnally averaged EBR is on the order of 70% so we are undoubtedly nearer a closed SEB when fluxes are diurnally averaged consistent with the results of Leuning et al. (2012). So perhaps the phase lag mechanism can lend an explanation for the large SEB residual observed at Bayelva. Even so, a relatively large residual still remains for the diurnally averaged fluxes so the phase lag is only a partial explanation: there must be other mechanisms at work. It is also worth reemphasizing that the diurnally averaged fluxes presented in Figure 1.14 do not represent an entire year of flux data and that we probably mask some of the residual when calculating the EBR of diurnally averaged fluxes due to a cancelation of night and daytime biases in the diurnal averaging procedure.

Closing off this section then it is worth keeping in mind that most of the mechanisms proposed in explaining the SEB closure problem are related to the EC flux methodology itself. These mechanisms introduce the widely acknowledged uncertainties (e.g. Hartmann et al. (2013)) in the quantification of the non radiative components of the global energy budget at the Earths surface. As pointed out in Stephens et al. (2012) it is particularly worrying that the uncertainty in the global net surface energy balance is on the order of $17\,\mathrm{Wm}^{-2}$ which is one order of magnitude larger than change in the net surface fluxes as a result of warming associated with the increasing concentrations of greenhouse gases in the atmosphere. Thereby, accurately attributing measured changes in the global SEB to anthropogenically induced climate change is at present far from trivial. Unfortunately this will continue to be the case so long as large residuals are present in experimentally determined SEB budgets. On the flip side, upon reducing the uncertainty in SEB estimates, through an improved and consistent EC methodology (as presented in Mauder et al. (2013)), it may be possible to add changes in observed SEB budgets to the many observationally based symptoms attributed to anthropogenically induced climate change (Hartmann et al., 2013).

Chapter 2

Method

In the ensuing chapter we will describe the steps taken towards providing a consistent EC methodology as outlined in e.g. Mauder et al. (2013). At the same time we account for the peculiarities of the Bayelva EC site. Since most of our work involved coding up an EC package from scratch, which required a thorough review of the underlying concepts, we devote the majority of this thesis to describing our methodology. First of all the number of points in time at our disposal was on the order of 3×10^8 samples per variable due to the high temporal resolution, twenty samples per second, and the 6 month duration of the complete data set. Processing the raw time series of the 5 variables of interest (u,v,w,T_s and ρ_v) to yield useful results is thus a relatively exhaustive effort. As such, the module was coded in `Fortran 90` in an attempt to keep computational time to a minimum. Due to the length of the module (> 7000 lines of code) it is not included as an appendix, but it is of course available upon request. In the coming sections we cover the sequential steps undertaken in the module to process and quality control the data. These steps involve how the module succesively: 1) Despikes data using a three step approach; 2) Applies various rotations to the measured velocities; 3) Obtains and calculates complimentary variables not measured by the EC system; 4) Applies a series of flux corrections; 5) Implements an extensive quality control procedure where blocks are quality flagged and flux uncertanties are estimated; 6) Calculates 30 minute block averaged fluxes; 7) Produces turbulence cospectra. For clarity we will also supplement the methodology with both examples and underlying theory where applicable.

2.1 Input

An excerpt of the input files is included as a reference in Figure 2.1. These files were produced with by a CSAT3000 data-logger which synchronized the measurements from the LI-7500 open path IRGA and the CSAT3 sonic anemometer and were kindly provided by S. Westermann through AWI.

Excerpt

```
"TOA5","1549","CR3000","1549","CR3000.Std.04","CPU:AWI EC V1_4.CR3","15620","ts_data"
"TIMESTAMP","RECORD","Ux","Uy","Uz","Ts","co2_mm_m3","h2o"
"TS","RN","m/s","m/s","m/s","C","","g/m^3"
"","","Smp","Smp","Smp","Smp","Smp","Smp"
"2007-06-12 12:12:56.6",126881922,-0.67075,2.44925,0.14175,1.2,16.82,4.57
"2007-06-12 12:12:56.65",126881923,-0.76775,2.41175,0.06500001,1.167,16.83,4.6
"2007-06-12 12:12:56.7",126881924,-0.775,2.31275,0.09900001,1.096,16.83,4.613
"2007-06-12 12:12:56.75",126881925,-0.7582501,2.47475,0.03375,1.093,16.83,4.625
```

Figure 2.1: *Excerpt from one of the ASCII-formatted raw data files, `TOA5_070601_070703_0002.dat`, showing the standard four header lines and the first four (out of 10^7) lines of raw data in the file.*

The module reads in and processes the raw data on a file-by-file basis. In these files (cf. Figure 2.1) the variables of interest are the longitudinal velocity u (`"Ux"`), the lateral velocity v (`"Uy"`), the vertical velocity w (`"Uz"`), the sonic temperature T_s (`"Ts"`) and the absolute humidity ρ_v (`"h2o"`), as well as the associated timestamp t (`"TIMESTAMP"`). Here it is worth emphasizing that the timestamp is in Coordinated Universal Time (UTC) as opposed to local time (UTC +1/+2 winter/summer). Any missing records are la-

beled as "NaN" in the file and these are simply set to a placeholder missing value which is conveniently set to -1000 in the module. This value is used throughout the module to indicate missing or faulty data as none of the variables should ever take such a value.

At our disposal we had a total of 20 files of raw data provided in two batches. The first batch, consisting of 18 files, spanned the period from 02.04.2007 to 14.09.2007, that is early spring to early autumn. There were some significant gaps, most notably we have no data from the period 02.05.2007 - 18.05.2007. The second batch, consisting of 2 files, spanned the early winter period from 09.12.2008 to 17.01.2009 with a gap of 3.5 hours between the files on the 28.12.2008. The shortest file in the two batches spanned 20 hours (1.5×10^6 lines of data) while the longest spanned 3.4 weeks (41.2×10^6 lines of data). Only files with a length corresponding to 5 days or longer were used in the data analysis due to the requirements of the planar fit rotation to be explained in more detail in Section 2.3.2.

2.2 Despiking

Once the variables are read into the module the first step in data processing involves screening data for spikes mainly due to instrument malfunction. This is done using a three step approach. First, the data is confined to lie within a physically reasonable range of values. Next, a median absolute deviation (MAD) despiking algorithm, following Mauder et al. (2013), is employed to flag and remove remaining instantaneous spikes in the data that would otherwise contaminate the turbulenct statistics. Last, nondimensional higher moments characterising the shape of turbulent probability distributions are tested against empirical thresholds (following Vickers and Mahrt (1997)) and blocks are flagged accorindlgy. Flagged spikes will be dealt with later in the Quality Control (Section 2.6). Screening data for spikes is a very delicate procedure (Vickers and Mahrt, 1997) and so we were as lenient as possible in all the steps involved. This because it is often extreme, yet completely physical, situations that are of interest to investigators and these can unwittingly fall under the definition of "spikes" in automated screening routines. As such, it was prudent to also visually inspect periods with a high frequency of spike occurrence. Upon such an inspection the spike detection was tuned accordingly. None of the files considered contained a fraction of spikes above 10%.

2.2.1 Plausibility Limits

In setting the upper and lower limits of physical plausibility for each of the measured variables it is of course necessary to keep in mind the location of the site and the local climatology. This is especially the case for the sonic temperature where we set the limits using monthly statistics for (2 m) air temperature from the nearby ($78°55'28''$N,$11°55'42''$E) Ny Ålesund automatic weather station (AWS) provided through eKlima (2014). The limits are based on the monthly extreme (minimum and maximum) temperatures as measured in the period January 1984 to December 2013. We chose a 30 year period as this meets the somewhat arbitrary (see Arguez and Vose (2011)) yet classical definition of a climate normal period set by the World Meteorological Organization. We could have chosen synchronous temperature measurements, that is from exactly the period we are studying, but in choosing a longer period we hope to make up for the fact that the EC system is not colocated with the AWS. Moreover these limits could be applied to periods not studied here. The limits along with the associated 30 year average monthly temperature statistics are depicted in Figure 2.2. We set the monthly upper (lower) limit to the 30 year maximum (minimum) monthly maximum (minimum) temperature plus (minus) the 30 year standard deviation of the monthly maximum (minimum) temperature for each month. To clarify, symbolically we compute the discrete monthly upper limits for temperature as follows

$$T_{xx,i} = \max_j \left(T_{x,i,j}\right) + \sigma^{(j)}_{T_{x,i,j}}\,, \tag{2.1}$$

and similarly for the monthly lower limits for temperature

$$T_{nn,i} = \min_j \left(T_{n,i,j}\right) - \sigma^{(j)}_{T_{n,i,j}}\,, \tag{2.2}$$

where $i \in 1(1)12$ is the month index and $j \in 1(1)30$ is the year index. The respective standard deviations hopefully account for the possible occurrence of different maxima/minima in the sonic temperature measurements due to the faster responding instrumentation and the difference in location. The discrepancy owing to the CSAT3 measuring sonic temperature and the limits being based on absolute temperature should be of little consequence due to the small difference between the two temperature measures.

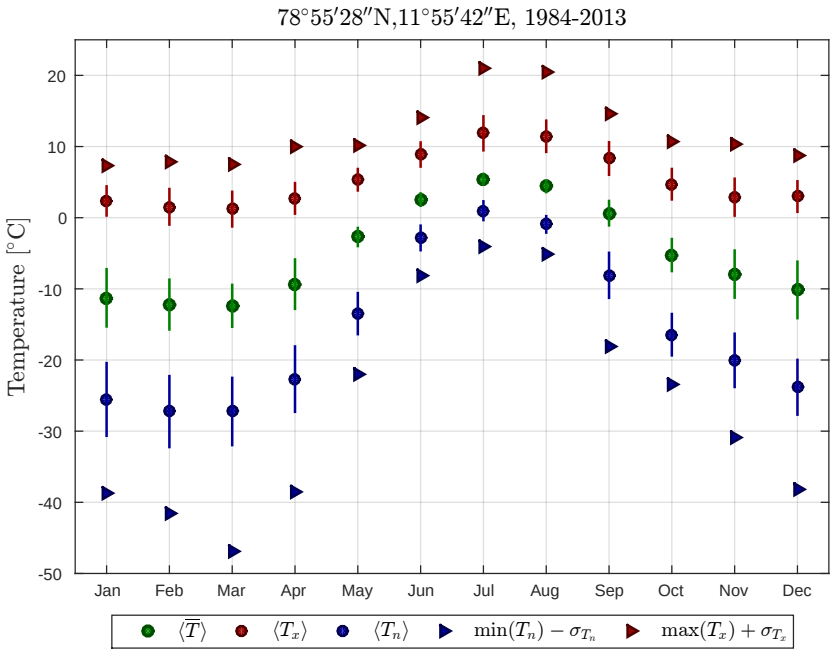


Figure 2.2: *30 year (1984-2013) monthly surface air temperature statistics for Ny Ålesund based on data from eKlima (2014).* $\overline{T}$ *is the monthly mean,* T_x *the monthly maximum and* T_n *the monthly minimum while* $\langle\rangle$ *denotes a 30 year average. The vertical lines span two standard deviations and are centered on the respective* 30 *year averages. The upper (lower) monthly temperature limits used in the module are depicted with red (blue) triangles.*

For the absolute humidity $\rho_v = 0\,[\mathrm{kgm}^{-3}]$ is a naturally defined lower limit. It is worth noting that measured values may be (falsely) negative, as they indeed were at times due to a drift in the zero reference value of the open path IRGA. Such a drift is expected after long measurement periods without maintenance and proper calibration (LI-7500, 2001). Attempting to calibrate periods where negative values are recorded using synchronous humidity measurements from another instrument would be the ideal solution, however, we had no such luxury. In the absence of another instrument we noted based on a visual inspection that periods with $\rho_v < 0$ usually coincided with measurements exceeding the *upper* limit and are as such associated with general instrument malfunction, likely due to precipitation remaining in the path of (or on) the IRGA. Moreover, we note that a slow long-term drift in the zero value should have little effect on the turbulence statistics which we base on fluctuations about means taken over local block periods which are short by comparison. The upper limits for ρ_v were set simply as the estimated saturated absolute humidity, ρ_v^*, at the upper limit for temperature in the given month based on Tetens' equation (see e.g. Foken (2008b))

$$\rho_{v,xx,i} = \rho_v^*\left(T_{xx,i}\right) = \frac{d_1}{T_{xx,i}}\exp\left\{\frac{d_2\left(T_{xx,i}-d_3\right)}{T_{xx,i}-d_4}\right\}, \tag{2.3}$$

where the constants are $d_1 = 1.2933\ \left[\mathrm{kg\ m}^{-3}\mathrm{K}\right]$, $d_2 = 17.6294$, $d_3 = 273.16\ [\mathrm{K}]$ and $d_4 = 35.86\ [\mathrm{K}]$ with the monthly upper limit for temperature $T_{xx,i}$ in Kelvin.

Measured variables	Lower limit	Upper limit	Units (SI)	Instrument
Sonic temperature: T_s	$T_{nn,i}$ see (2.2)	$T_{xx,i}$ see (2.1)	K	CSAT3
Absolute humidity: ρ_v	0	$\rho_{v,xx,i}$ see (2.3)	kgm^{-3}	LI-7500
Horizontal velocities: u,v	-32.768	32.768	ms^{-1}	CSAT3
Vertical velocity: w	-5	5	ms^{-1}	CSAT3

Table 2.1: *Physical plausibility limits for the measured variables as set in the module. The values are based on monthly extremes for Ny Ålesund from eKlima (2014) as well as thresholds outlined in CSAT3 (2014) and Aubinet* et al. *(2012).*

The horizontal velocity limits are set to those outlined in CSAT3 (2014) for the high resolution wind mode (as opposed to full scale wind). This corresponds to hurricane force horizontal winds as the limit for u and v. The vertical velocity limit is set to $\pm 5\text{ms}^{-1}$ in accordance with common practice (Aubinet et al., 2012). All the limits are outlined in Table 2.1 above. Any measurement not satisfying the limits dictated by the table above was flagged as a spike and set to the `"NaN"` placeholder value of -1000 as was done with the missing records when reading in the file. We reiterate that both the temperature limits as well as the upper limit on absolute humidity are assigned[1] on a monthly basis, but are independent of the year of the EC measurements.

2.2.2 Median Absolute Deviation Test

Ultimately the act of confining the measurements to obey a set of physical limits will (should) only isolate the extreme cases of instrument malfunction. Most spikes occur within the range of physically plausible values. These occur simply as a result of glitches in the electronics, instrument vibration and from meteorological factors such as hydrometeors (rain or snow) in the measurement path of, or deposited on, either of the two instruments (Aubinet et al., 2012). Such pikes will typically manifest as instantaneous high amplitude deviations from neighboring points, hence the term spike. Of course, the signals we are recording are turbulent in nature and are as such composed of a multitude of 'spikes' of varying amplitude. Since turbulence itself manifests as perturbations about a local mean it also contributes directly to the variance of the signal. In fact it is precisely the variance we use to quantify the intensity of turbulence. Subsequently it is somewhat of a challenge to disentangle and isolate spikes related to instrument malfunction from physical turbulence.

Luckily the amplitude of spikes is usually large relative to the standard deviation (square root of variance), σ, of the signal in a given segment. As such many automated despiking algorithms, for example the one in Vickers and Mahrt (1997), define spikes as absolute deviations about the mean which exceed a threshold factor of σ in a segment. The segment defining σ is either defined for a block interval, introducing arbitrariness as to where to place each interval, or in terms of segments centered on each point (i.e. a running σ) which is more rigorous but introduces a heavier computational burden. Both methods suffer from a considerable sensitivity to the length of the segment itself and of course to the value that is chosen as the threshold factor. Perhaps of greatest detriment to both methods is that the spikes have a significant influence on the value of σ, which is itself defined with respect to deviations about the mean. That is: the largest spikes lead to a high σ value such that smaller spikes may fall below the threshold. Hence even if one spike is identified multiple iterations may be required to remove all spikes present in a given segment.

Due to the above, Mauder et al. (2013) propose a more robust method, according to Fratini and Mauder (2014), which is both less sensitive to the choice of segment length and less affected by the choice of threshold factors as spikes have less influence on the threshold parameter. As an added bonus the method is relatively cheap computationally and uses simple block segments. The method uses the so-called median absolute deviation (MAD) as the threshold parameter; it is this parameter we adopt in the module with simple 30 minute block segments. For a given measured variable ξ, we define $\xi_{i,m}$ as the i-th element of the m-th block segment of our time series length $N = M \cdot I$, where I is the number of samples per block and M is the number of blocks in the series. Denoting the block median with $\{\}_m$ the median absolute deviation of block m is defined as

$$\text{MAD}_k = \{|\xi_{i,m} - \{\xi\}_m|\}_m \,, \tag{2.4}$$

or in words as the median of the absolute values of the deviations about the median for a given block. A great advantage of the above statistic compared to the standard deviation is that it does not rely on the mean. As such the thresholds are relatively insensitive to the magnitude and frequency of spikes. We use the MAD to test each sample within a block against the criterion

$$|\xi_{i,m} - \{\xi\}_m| \leq \mathcal{T} \cdot \text{MAD}_m \,, \tag{2.5}$$

where $\mathcal{T}$ is a threshold factor which, through trial and error on sample data, we set to 10 in close concurrence with Mauder et al. (2013). Any entry *not* passing the MAD test, i.e. not satisfying the crieterion in (2.5), was flagged as a spike and set to -1000; the placeholder value used for erroneous or missing data in the module. Moreover the MAD routine also spread spikes: if any variable was flagged as a spike at an instant in time all other variables were flagged as spikes for the same instant *even if* they passed the MAD test and the plausibility test. This was done because a spike in one variable is symptomatic of instrument or meteorological problems which can affect other variables. As an added bonus the cross-covariant and cross-spectral calculations which follow become a lot easier to implement. An overview of the number of points flagged as spikes is given in Section 3.1.

[1] A text file containing the monthly temperature limits is read into the module.

2.2.3 Skewness & Kurtosis Test

In the final step of the despiking routine we diagnose the shape of the block probability distributions in an attempt to screen for problems not identified by the previous two procedures. We do so by considering the values of normalized higher moments, the skewness and the kurtosis, with regards to thresholds outlined in Vickers and Mahrt (1997).

To estimate these normalized moments from samples within a given block a linear detrend (LDT) is employed internally in this routine as the first step. Such a detrend is used to eliminate the potential contributions of trends to the sensitive higher moments (Vickers and Mahrt, 1997). Trends are associated with unwanted depatures from stationary conditions which are later identified and flagged in Section 2.6. Now to be able to apply a LDT we first need to compute the linear trend, ignoring flagged spikes, within a given block. We do so by adapting[2] the simple least squares algorithm proposed in Gash and Culf (1996) to not weight flagged entries. Having found the trend the block time series for a given variable ξ is readily detrended via simple subtraction of the trend component, i.e.

$$\xi''(t) = \xi(t) - \underline{\xi}(t)\,, \tag{2.6}$$

where $\underline{\xi}$ is the time dependent trend component and ξ'' is the instantaneous deviation from the trend component. Note that in the special case of a block with no trend the LDT is identical to the mean removal of a block average; in such a case $\underline{\xi} = \overline{\xi}$. More generally, as with mean removal, the block average of the deviation ξ'' is zero. In the following it is the shape of the distribution of ξ'', i.e. of the LDT samples, that we seek to diagnose.

Next, following Graf et al. (2010) two normalized moments, the skewness and the kurtosis, are found from the estimated sampled third and fourth moments of ξ'' upon normalization by the standard deviation $\sigma_{\xi''}$ to the corresponding power. So the skewness of the ξ'' sample, $S_{\xi''}$, is given by

$$S_{\xi''} = \frac{\overline{\left(\xi'' - \overline{\xi''}\right)^3}}{\sigma^3_{\xi''}} = \frac{\overline{(\xi'')^3}}{\sigma^3_{\xi''}}\,, \tag{2.7}$$

and the kurtosis of the ξ'' sample, $K_{\xi''}$, is given by

$$K_{\xi''} = \frac{\overline{\left(\xi'' - \overline{\xi''}\right)^4}}{\sigma^4_{\xi''}} = \frac{\overline{(\xi'')^4}}{\sigma^4_{\xi''}}\,. \tag{2.8}$$

Note that in both (2.7) and (2.8) we have chosen to emphasize the fact that in this particular case of a LDT sample time series the mean of the detrended sample is zero, $\overline{\xi''} = 0$. As a result the programatic implementation of both calculations in the module is somewhat less cumbersome then what it would be in the absence of a LDT.

These normalized moments are invaluable for quantitatively classifying the shape of a given probability distribution. First we note that relative to the variance (the second moment) the skewness and the kurtosis give considerably more weight to the extreme fluctuations due to the increase in the value of the exponent. Moreover we note that the kurtosis is positive definite whereas the skewness is not. The skewness quantifies the asymmetry, the 'tilt' about $\xi'' = 0$, of the distribution (Graf et al., 2010). The more positive the skewness the more the peak of the distribution tilts towards negative deviations ($\xi'' < 0$) and the positive tail is relatively more pronounced. Conversely the tilt is towards increasingly poisitive deviations ($\xi'' > 0$) with a relatively more pronounced negative tail the more negative the skewness. The kurtosis quantifies the 'peakedness' of the distribution (Graf et al., 2010): if the kurtosis is large, as in the red distribution in the top left panel of Figure 2.3, the peak is sharp and the wings are long, conversely if it is small the peak is broader and the wings are short.

Here we use the descriptive terms such as 'large' and 'small' with respect to the Gaussian distribution which has zero skewness and a kurtosis of 3. We will use the Gaussian distribution as a reference throughout; not because it is especially applicable but because it is a distribution which many are familiar with. Graf et al. (2010) point out that in the turbulent statistics of the surface layer the Gaussian distribution is found to be 'the exception rather than the rule'. Based on sample data for Bayelva in a 20 day period, we find on the grounds of a Kolmogorov-Smirnov test that the null-hypothesis that the distributions were Gaussian was rejected for over 99% of the blocks at the 5% significance level ($p \leq 0.05$) for both sonic temperature and instrument vertical velocity corrobarating the claim of Graf et al. (2010). This is seen in the lower panel of Figure 2.3 where the fraction $N(p \leq 0.05) > 0.99$ for both w'' and T_s''. As can be noted from the description of skewness and kurtosis the two moments are somewhat related: large absolute skewness and large kurtosis both lead to the elongation of the wings of the distribution; although for skewness only one of the wings is elongated due to the asymmetry. Thus, it is often seen (Graf et al., 2010) in $S - K$ plots that kurtosis has a square dependence on the kurtosis. We verified this using example data from Bayelva, as shown in the bottom panel of Figure 2.3, where the actual kurtosis and

[2] A detailed description of a more involved least squares method that we also employed will follow in Section 2.3.2.

the parametrized (best fit square dependence on skewness) kurtosis are shown to have a considerable linear correlation ($R > 0.5$) for both the variables we investigated. Nonetheless our square linear correlation values are appreciably lower than to those found in Table 2 of Graf et al. (2010). As of yet it is unclear exactly as to why this would be, but it is possible to stipulate that Graf et al. (2010) testing two conserved scalars, water vapor mixing ratio and potential temperature, whereas we tested the sonic temperature (not a conserved scalar) and vertical velocity could explain the discrepancy.

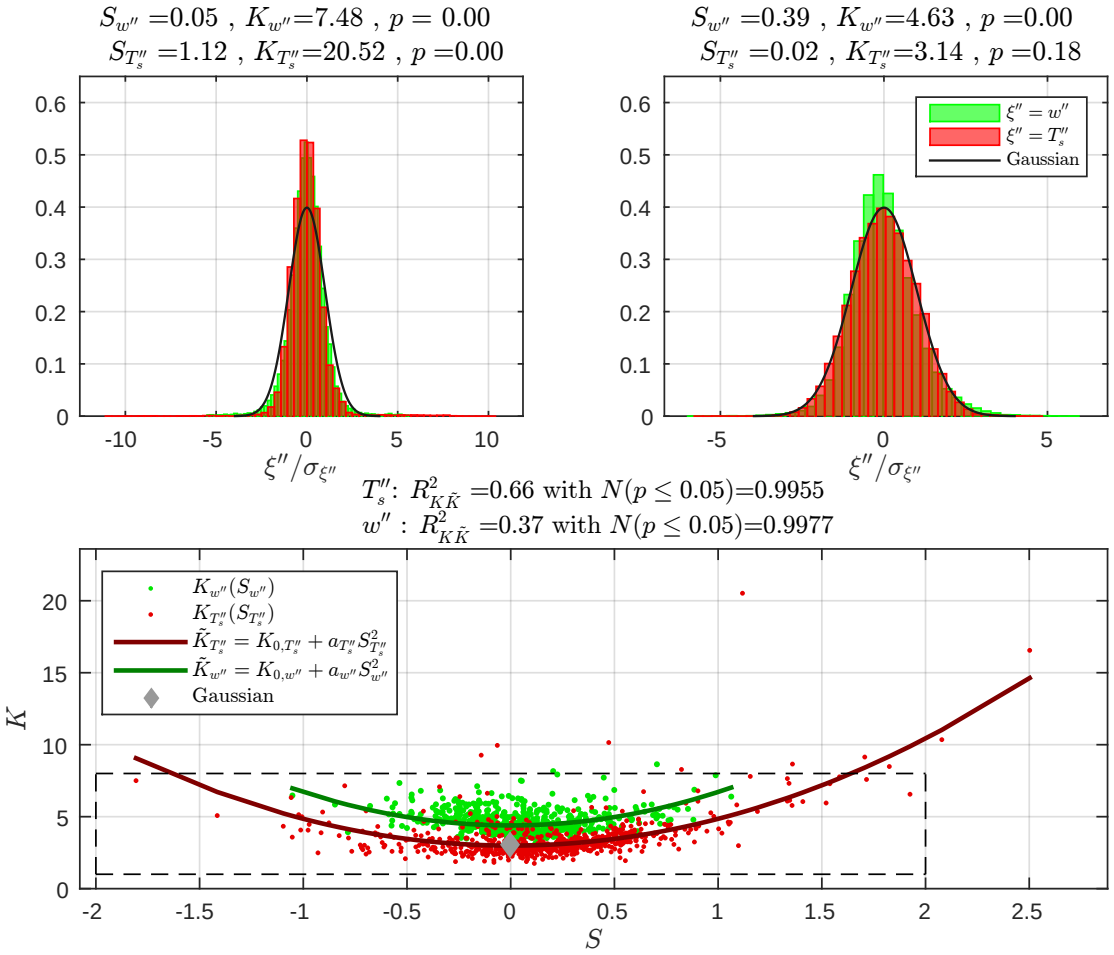


Figure 2.3: *Top panels: Two examples (Left:10.12.2008 18:30-19:00, Right: 27.12.2008 09:00-09:30) of block probability density for LDT nondimensionalized sonic temperature (red), $T_s''/\sigma_{T_s''}$, and LDT nondimensionalized sonic vertical velocity (green), $w''/\sigma_{w''}$. S-K values and the p-values from a Kolmogorov-Smirnov test are included at the top of each panel. Bottom panel: S-K plot for all 30 minute blocks in the example file (time interval: 09.12.2008-28.12.2008). Diagnosed values are given by red/green dots for sonic temperature/vertical velocity. Parametrized kurtosis values, $\tilde{K}$, are shown as lines with the corresponding color but in a darker shade. Values for the square correlations, $R^2_{K\tilde{K}}$, and the fractions $N(p \leq 0.05)$ are included at the top of the panel (see text for details). Blocks that were not hard flagged, i.e. within the thresholds $-2 \leq S \leq 2$ and $1 \leq K \leq 8$, lie within bounding box depicted by the dashed black lines.*

In terms of the turbulent surface layer it has been found (e.g. Tillman (1972) and Graf et al. (2010)) for both vertical velocity and scalar distributions that under unstable stratification ($\zeta < 0$) the skewness is positive and conversely under stable stratification ($\zeta > 0$) the skewness is typically negative. Physically Tillman (1972) attributed the positive skewness of the temperature distribution in unstable stratification to downwards advection of cold well mixed air (nearly uniform temperature) from the mixed layer aloft truncating the cold end of the distribution combined with sporadic plumes bringing much warmer parcels up from the surface resulting in an elongation of the warm tail of the distribution. One could, to a certain extent, turn this argument on its head to explain the occurence of negative skewness in the stable regime. As for kurtosis large values are associated either with instrument related problems, such as condensation remaining on the sensors (Vickers and Mahrt, 1997), or significant horizontal heterogeneity (Graf et al., 2010). Low kurtosis on the other hand is usually associated with the distribution splitting into two distinct peaks due to nonstationarity in the form

of sudden transition in the block; often indicative of the passage of a front (Graf et al., 2010). So both high and low kurtosis values are symptomatic of problems with the application of the EC method; either due to invalid underlying assumptions or instrument problems. Nonetheless, neither skewness nor kurtosis have the clear cut dependence on ζ that variance does (Section 2.6.5); as is evident in the relatively large scatter of the results of both Tillman (1972) and Graf et al. (2010). This scatter is in large part attributed to the exponential increase in statistical error (Graf et al., 2010) involved in the estimation of these normalized third and fourth moments based on samples with a finite (block) duration; with the error being largest for the kurtosis.

We should keep these errors in mind when setting thresholds for these higher moments by giving considerable leeway to account for the possible apparent increase in the values of the moments. Furthermore, as the expectation is *not* a Gaussian distribution we should of course not set the thresholds on the moments as if it were (Gaussian). We found the non-Gaussian, yet admitedly empirical thresholds, outlined in Vickers and Mahrt (1997) satisfactory for our purposes. Symbolically the block skewness-kurtosis flag, f_{SK}, is in accordance with the procedures in Section 2.6 defined as an integer value from 0 (high quality flag) to 2 (hard flag: discard) equal to the maximum value of the integer kurtosis ($\mathrm{f}_{K,\xi''}$) and integer skewness ($\mathrm{f}_{S,\xi''}$) flags for both the LDT sonic temperature ($\xi'' = T_s''$) and the LDT vertical velocity ($\xi'' = w''$) as follows

$$\mathrm{f}_{SK} = \max\left(\left[\mathrm{f}_{S_{T_s''}}, \mathrm{f}_{S_{w''}}, \mathrm{f}_{K_{T_s''}}, \mathrm{f}_{K_{w''}}\right]\right) , \quad (2.9)$$

where the skewness and the kurtosis flags are set according to the thresholds suggested in Vickers and Mahrt (1997) and namely

$$\mathrm{f}_{S_{\xi''}} = \begin{cases} 0 & \text{if } |S_{\xi''}| \leq 1 \\ 1 & \text{if } 1 < |S_{\xi''}| \leq 2 \\ 2 & \text{otherwise} \end{cases} , \; \mathrm{f}_{K_{\xi''}} = \begin{cases} 0 & \text{if } 2 \leq K_{\xi''} \leq 5 \\ 1 & \text{if } 1 \leq K_{\xi''} < 2 \text{ or } 5 < K_{\xi''} \leq 8 \\ 2 & \text{otherwise} \end{cases} . \quad (2.10)$$

Recall that the block kurtosis and block skewness are computed through (2.8) and (2.7) respectively after the block trend is found using the method of Gash and Culf (1996) and the block LDT given by (2.6) has been applied. The thresholds were empirically judged to represent 'normal physical expectations' (Vickers and Mahrt, 1997) and were purposely set to be lenient: accounting for both statistical error and rare, yet completely physical, phenomena that occur towards their outer limits. For economy we only considered the block kurtosis and skewness of despiked (through the plausibility and MAD tests) LDT sonic temperature and LDT instrument vertical velocity when assigning these flags. This because the result did not change considerably in the sample data when more of the measured variables were included, also problematic blocks were most readily identified through these two variables. For the example period in the lower panel of Figure 2.3 only 1% of blocks were hard flagged ($\mathrm{f}_{SK} = 2$) soley[3] on the grounds of either the skewness or the kurtosis exceeding the thresholds of Vickers and Mahrt (1997); these are the points that lie outside the dashed bounding box. In summary, each block was flagged if the kurtosis or the skewness exceeded the accepted ranges for either the sonic temperature or the vertical velocity. If a block was hard flagged by (2.9) then *all variables* in that block are set to -1000 and discarded from further analysis. The value of the skewness-kurtosis flag for each block is stored for bookkeeping, which we will return to in Section 3.1.

[3]That is blocks that were not also hard flagged by the tests in Section 2.6.

2.3 Rotations

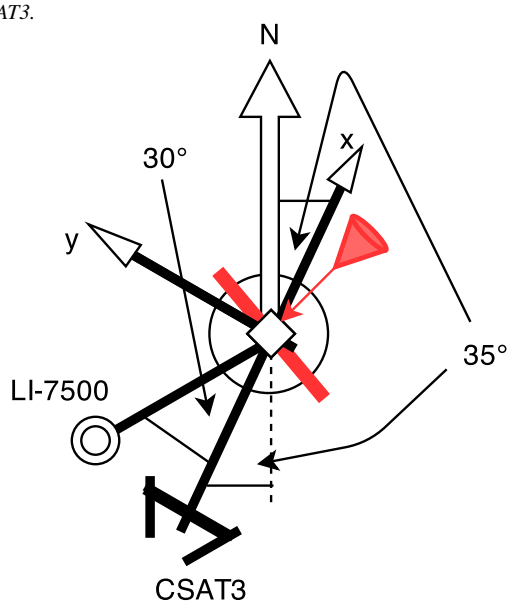
In the following we describe three separate rotation procedures undertaken in the module. Each procedure has distinct objectives. The first two procedures both aid in dealing with sources of systematic error; the orientation procedure enables the identification of problematic wind directions prone to flow distortion whereas the planar fit of Wilczak et al. (2001) minimizes tilt errors. The final procedure, a rotation into the natural ensemble streamline frame (NESF), is used both for convenience and to bring results into a frame which is close to (but not the same as) the natural wind frame employed in classical micrometeorological experiments (e.g. Haugen et al. (1971)).

2.3.1 Orientation

The CSAT3 sonic anemometer was leveled using the instruments' internal leveling mechanism such that the sonic head lay nearly level in a plane of constant geopotential as opposed to the plane defined by the local surface. In other words, the vertical coordinate of the CSAT3 is very close to truly vertical. As such a simple rotation about the vertical axis, approximately parallel to local gravity, will bring the frame into a geographically referenced so-called 'seemingly local earth coordinate' or SLEC for short (Sun, 2007). By determining the orientation of the CSAT3 wind axes relative to the geographically referenced SLEC frame the identification of wind directions influenced by significant flow distortion becomes possible, an important step in the quality control process. Moreover, we can easily test if the system is properly mounted: the sonic anemometer should be facing into the most frequently occurring wind direction if possible.

According to Aubinet et al. (2012), and perhaps not surprisingly, any EC system should be intentionally set up in an attempt to face it *into* the long-term predominant wind direction if such a direction exists at a given site. The aim of such an orientation is to reduce the effects of flow distortion in the 'closed sector' (Li et al., 2012) caused partly by the boom and tower structure on which the instruments are mounted, but mainly by the head of the CSAT3 being upwind of the CSAT3 measurement path. In so doing, unfavorable wind directions rarely occur and the fraction of raw data that must be discarded due to flow distortion is considerably reduced. Hoping to achieve such an optimal setup the CSAT3 was pointed 215° clockwise from north and the LI-7500 245° clockwise from north (S. Westermann personal communication). The orientation is sketched schematically in Figure 2.4. To clarify, this means that a measured wind $\mathbf{v} = u\hat{\mathbf{x}}$ with $u > 0$, where $\hat{\mathbf{x}}$ is the unit vector defining the main sonic axis, corresponds to a wind direction of 215°. On the other hand, a measured wind $\mathbf{v} = v\hat{\mathbf{y}}$ with $v > 0$, where $\hat{\mathbf{y}}$ is the unit vector defining the cross sonic axis, corresponds to a wind direction of 125°.

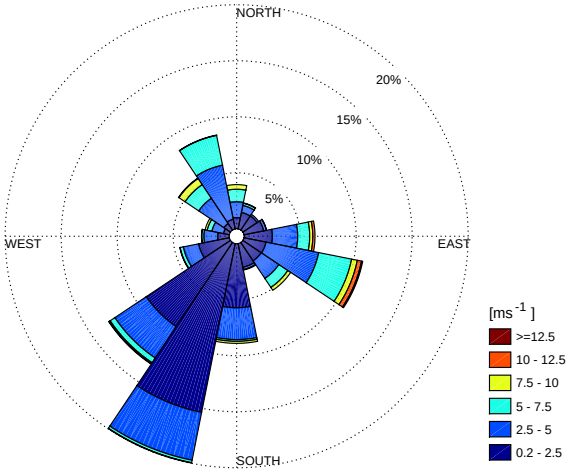
Figure 2.4: *The orientation of the instruments relative to the horizontal SLEC frame. The CSAT3 x-axis and y-axis as well as the SLEC (geographic) north axis are included for reference. The red bar represents the boom attached to the tower on which the instruments are mounted and the red cone represents the wind directions for which to expect significant flow distortion from the boom as well as the head of the CSAT3.*



When computing the corresponding wind directions based on the despiked 'horizontal' wind components our module rotates the coordinate system into SLEC in a few straight forwards steps. Defining U_c as the zonal velocity component and V_c as the meridional velocity component it is readily shown that $U_c = u\cos(\alpha) - v\sin(\alpha)$ and $V_c = u\sin(\alpha) + v\cos(\alpha)$, where α is the angle between due east and the sonic x-axis. α is computed based on the CSAT3 bearing; once computed it is stored in the module for use in subsequent processing steps. The wind direction (bearing) is then computed as $\varphi = \arctan\left(\frac{-U_c}{-V_c}\right)$. Note the sign reversal in the expression, as is the convention in meteorology we define wind direction as where the wind is blowing from as opposed to where it is blowing to. Subsequently the mean wind direction and speed are computed us-

ing mean wind vectors defined for blocks with lengths corresponding to the averaging period set by the user; which in our runs was 30 minutes. As such, both the wind direction and wind speed are synchronized with later calculated block averaged statistics.

Figure 2.5: *Wind rose displaying mean wind statistics at Bayelva in the periods 02.04.2007-14.09.2007 and 09.12.2008-17.01.2009.*



The steps described above enables simple flagging of periods where the (mean) wind direction is in the unfavorable closed sector. For the Bayelva site this corresponds to wind directions in the range 15° to 55°, red arrow in Figure 2.4, in such cases the boom structure and the sonic head are directly upwind of the sonic anemometer measurement path and thus cause significant flow distortion. That is to say the man made obstacle causes synthetic turbulence which contributes to calculated variances and covariances, hence spoiling the estimation of the true surface exchange. Note that the user may specify more than one wind direction bin to exclude, if required by the instrument configuration. Blocks with unfavorable (mean) wind directions are flagged and removed in the quality control routine as discussed in Section 2.6.

At Bayelva, Westermann et al. (2009) noted that the mean wind directions in the closed sector occurred in only 1.5% of the flux averaging blocks in the period 15.03.2008-15.03.2009. An analysis of the two batches of data we considered corroborates the fact that these directions are unusual, as visualized in Figure 2.5. By far the most common wind direction for this period was in the range $195^\circ - 225^\circ$, from Vestre Brøggerbreen, which occurred 37% of the time. For comparison, the unfavorable wind direction occurred only 1.7% of the time in the same period. As such we conclude that the orientation of the instruments was indeed optimized for this particular study site as recommended in Aubinet et al. (2012).

2.3.2 The Planar Fit

To be able to arrive at eddy fluxes that are truly representative of the local surface it is imperative that tilt errors be dealt with early on in the processing steps. As pointed out in e.g. Lee et al. (2006) there are two types of tilt error that we must be wary of: sensor tilt and slope tilt. Sensor tilt involves the fact that, despite manual leveling of a sonic anemometer using its internal leveling mechanism, we can at best hope that the instruments horizontal plane (x, y) is on the order of a degree from the true (geopotential) horizontal plane. Slope tilt accounts for the fact that the local terrain has a slope, which we will denote with the angle γ, relative to the true horizontal plane. As such, the normal to the surface patch that represents the fetch of the sonic will not be parallel to the geopotential gradient or the sonic vertical axis (z). Recall that it is the flux along a representative[4] surface normal that we are interested in measuring. The problem is depicted schematically in Figure 2.6.

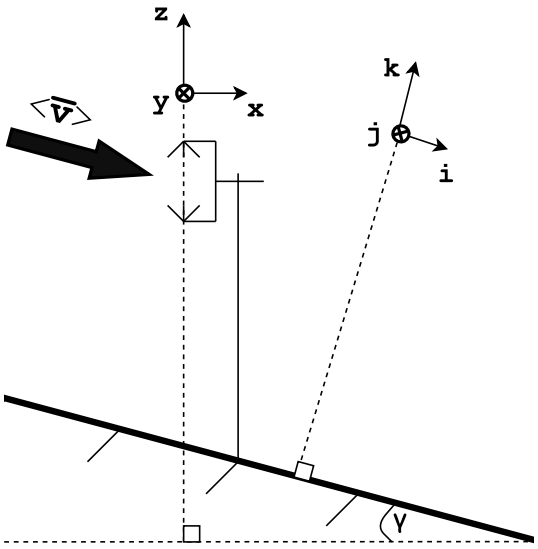
Both types of tilt error are present in the raw data for Bayelva. Westermann et al. (2009) point out that the instruments are mounted on the slightly inclined, $\gamma < 5^\circ$, slope of the Leirhaugen hill; this slope tilt is by far the largest source of tilt error. Secondly the sonic is closely, but still only approximately leveled relative to the true horizontal as previously explained. So, we need a method that can simultaneously correct for these two types of tilt error, otherwise we risk that any computed eddy fluxes are contaminated by the much larger magnitudes of the along (and cross) slope components of the flow leading to significant bias. In Wilczak et al. (2001) MOST is used to show that the contamination is most severe for the Reynolds stress, $\overline{u'w'}$, which is not surprising in that this stress contains the products of velocity components. In the unstable case the fractional tilt error in the stress is found to be as much as 64% in the case of $\zeta = -1$ with a tilt of only 1° relative to the surface. Nonetheless the error is still shown to be significant for scalar fluxes. For example, the heat flux fractional error is approximately 9% per degree of tilt in neutral stratification ($\zeta \simeq 0$).

Before proceeding we follow Lee et al. (2006) in making the important distinction between a coordinate system, which is a global property of the flow, and a unit

[4]The surface normal is likely to be a function of space, and even time (due to the snow cycle), for the surface patch within the footprint of the sonic anemometer.

vector basis that is generally a local property of the flow. The distinction is necessary in that we are making single point measurements of the wind field and can only use these to define a local unit vector basis. In the course of this section we define a set of 'ensemble' streamline planes based on the unit vector $\widehat{\boldsymbol{k}}$ that is normal to these planes. These should not be confused with actual streamline surfaces as we can only hope to define the normal locally. So bear in mind that when we are referring to frames we mean frames dictated by a local set of unit vectors. The only meaningful coordinate system we can define with information from one local point in the flow is a Cartesian coordinate system. So the frame we arrive at is *not* a curvlinear streamline coordinate. MIn anticipation of possible confusion with regards to the use of the term 'ensemble': exclusively in the current section of the thesis, Section 2.3, the term ensemble is a reference to the block ensemble (to be defined) and *not* the 'true' ensemble average.

Figure 2.6: *A sonic anemometer set on a slope which is tilted an exaggerated angle γ relative to the horizontal. The instrument frame, $[\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}]$, is approximately leveled relative to the horizontal. The planar frame $\left[\widehat{\boldsymbol{\imath}}, \widehat{\boldsymbol{\jmath}}, \widehat{\boldsymbol{k}}\right]$ defines the ensemble streamline plane such that the 'ensemble' of the block averaged wind $\langle\overline{\mathbf{v}}\rangle$, with vertical offset taken into account, lies in the plane spanned by $\widehat{\boldsymbol{\imath}}$ and $\widehat{\boldsymbol{\jmath}}$.*



Now we are prepared to describe in full the steps taken in the module to rotate the wind components into an ensemble streamline frame through a planar fit algorithm. We do so in considerable depth because this particular procedure is one of the key steps in the module to ensure that the measured eddy covariance flux properly represents the surface exchange. As opposed to Wilczak et al. (2001), our planar fit method excludes the explicit use of rotation angles, but still yields the same result. We follow a method similar to that outlined in Appendix B in Chapter 3 of Lee et al. (2006) using relatively simple and easy to follow unit vector operations. Defining the two orthogonal sets of triplets of unit vectors $[\widehat{\mathbf{x}}, \widehat{\mathbf{y}}, \widehat{\mathbf{z}}]$ and $\left[\widehat{\boldsymbol{\imath}}, \widehat{\boldsymbol{\jmath}}, \widehat{\boldsymbol{k}}\right]$ where the first set defines the instrument (sonic) frame and the second the planar frame to be determined. We may work entirely in the instrument frame in this derivation, expressing vectors in the planar frame in terms of the instrument basis. So, for example, we may express the planar normal vector as $\widehat{\boldsymbol{k}} = k_x\widehat{\mathbf{x}} + k_y\widehat{\mathbf{y}} + k_z\widehat{\mathbf{z}}$ in the instrument basis. We note based on the above that the block averaged wind vector can be expressed in both frames as follows (Wilczak et al., 2001)

$$\overline{\mathbf{v}} \equiv \overline{u}_s\widehat{\mathbf{x}} + \overline{v}_s\widehat{\mathbf{y}} + (\overline{w}_s - b_0)\,\widehat{\mathbf{z}} = \overline{u}_p\widehat{\boldsymbol{\imath}} + \overline{v}_p\widehat{\boldsymbol{\jmath}} + \overline{w}_p\widehat{\boldsymbol{k}}\,. \tag{2.11}$$

Where we have let the subscript s symbolize the instrument frame and the subscript p the planar frame. Here b_0 is the instrument offset which is a result of possible biases in the measured vertical wind speed related to zeroing issues with the sonic anemometer[5]. Of course, an offset may also be present in the horizontal velocities, but as demonstrated by Wilczak et al. (2001) the effect is negligible as these velocities are typically orders of magnitude larger than the vertical velocity. Hence we only consider the possible influence of a vertical offset.

Subsequently, we isolate the $\widehat{\boldsymbol{k}}$-component of (2.11), using that $\overline{w}_p = \widehat{\boldsymbol{k}} \cdot \overline{\mathbf{v}}$, to yield the following expression for the planar vertical velocity

$$\overline{w}_p = k_x\overline{u} + k_y\overline{v} + k_z\left(\overline{w} - b_0\right)\,,$$

where we have dropped the s subscript for economy. Rearranging the expression in terms of the sonic vertical velocity the expression reads

$$\overline{w} = b_0 + b_1\overline{u} + b_2\overline{v} + \frac{\overline{w}_p}{k_z}\,, \tag{2.12}$$

where we have defined the ratios $b_1 = -k_x/k_z$ and $b_2 = -k_y/k_z$. To proceed we define the ensemble block average, as in Finnigan et al. (2003), but here specifically as the average over all the M block averages in the time series. That is the ensemble block average of the u velocity in the sonic frame is defined as

$$\langle\overline{u}\rangle = \frac{1}{M}\sum_{m=1}^{M}\overline{u}_m\,,$$

with analogous definitions for $\langle\overline{v}\rangle, \langle\overline{w}\rangle$ and $\langle\overline{w}_p\rangle$. For clarity we have now also made the block subscript m explicit. As in Paw et al. (2000) we make the ansatz that the planar vertical velocity in (2.11), which according to Lee (1998) represents 'random-noise', should go to zero when averaged over a sufficient number of blocks, i.e. $\langle\overline{w}_p\rangle \to 0$ as $M \to \infty$. The *maximum* measuring height for the EC system at Bayelva is 2.75 m AGL; this corresponds to the height of the sonic anemometer in the absence of the snow. So we may safely assume that we are measuring within the surface layer; at least if we ignore the cases of extremely stable stratification

[5]The offset is assumed constant for each planar fit.

discussed in (Mahrt, 1999). Hence the ansatz is well grounded as within the surface layer there should exist a long term streamline plane that is approximately aligned with the local surface (Sun, 2007). So provided that a sufficient number of blocks M are included allowing a proper identification of an ensemble streamline plane, (Mauder and Foken, 2004) suggest M corresponding to at least 5 days, the ansatz is physically reasonable.

It is the plane in which this ansatz is satisfied that defines the ensemble streamline plane and hence the planar frame. Or stated explicitly: *the ensemble streamline plane is defined as the plane surface whose ensemble average surface normal velocity is zero.* So taking the block ensemble average of (2.12) rearranging in terms of $\langle\overline{w}_p\rangle$ and applying the planar fit ansatz we seek coefficients b_0, b_1 and b_2 such that

$$0 = \frac{1}{M}\sum_{m=1}^{M}\left(\overline{w}_m - b_0 - b_1\overline{u}_m - b_2\overline{v}_m\right) , \quad (2.13)$$

is satisfied. Note that it is not the trivial and *unphysical*[6] solution with offset $b_0 = M\langle\overline{w}\rangle$ and remaining coefficients $b_1 = b_2 = 0$ that we seek. Nor is the naive approach where we ignore offset and calculate the ensemble averaged wind $\langle\overline{\mathbf{v}}\rangle$ helpful. Even if the latter approach, upon normalization, determines one unit vector which lies in the ensemble streamline plane we need one more unit vector that lies in the same plane with a component orthogonal to $\langle\overline{\mathbf{v}}\rangle$ to be able to find the surface normal to the plane; or vice-versa we need the surface normal to be able to find the unit vector pair spanning the plane. So we need the surface normal to be able to uniquely define the plane. This surface normal, in turn, is uniquely determined by the coefficients that satisfy (2.13). In seeking non-zero coefficients b_0, b_1 and b_2 that satisfy (2.13), which is underspecified, we first decompose the above into M equations (one for each block). In matrix form such a system reads

$$\begin{bmatrix} r_1 \\ \vdots \\ r_M \end{bmatrix} = \begin{bmatrix} \overline{w}_1 \\ \vdots \\ \overline{w}_M \end{bmatrix} - \begin{bmatrix} 1 & \overline{u}_1 & \overline{v}_1 \\ \vdots & \vdots & \vdots \\ 1 & \overline{u}_M & \overline{v}_M \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} , \quad (2.14)$$

where we have introduced the block residual terms $r_m = \overline{w}_{p,m}/k_z$ in the vector $\mathbf{R} = [r_1, \ldots, r_M]$. Note that (2.13) and (2.14) are equivalent if the sum of the residuals is zero. Now *assuming* that $\mathbf{R}$ is a random error vector with no preferred sign as in Lee (1998), then minimizing the sum of the residuals can be achieved through a reduction of the magnitude of the vector $\mathbf{R}$ by optimizing the choice of non-zero coefficients b_i. As pointed out by Wilczak et al. (2001) this is tantamount to minimizing the function

$$S = \sum_{m=1}^{M}\left(\overline{w}_m - b_0 - b_1\overline{u}_m - b_2\overline{v}_m\right)^2 ,$$

in the least squares sense where $S = |\mathbf{R}|^2$. That is we seek coefficients satisfying $\frac{\partial S}{\partial b_0} = \frac{\partial S}{\partial b_1} = \frac{\partial S}{\partial b_2} = 0$. Performing the derivatives, rearranging and dividing through by a factor 2 yields three normal equations. These can be expressed compactly in matrix form, $\mathcal{A}\mathbf{b} = \mathbf{c}$, as

$$\underbrace{\begin{bmatrix} M & \sum\overline{u}_m & \sum\overline{v}_m \\ \sum\overline{u}_m & \sum\overline{u}_m^2 & \sum\overline{u}_m\overline{v}_m \\ \sum\overline{v}_m & \sum\overline{u}_m\overline{v}_m & \sum\overline{v}_m^2 \end{bmatrix}}_{\mathcal{A}} \underbrace{\begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}}_{\mathbf{b}} = \underbrace{\begin{bmatrix} \sum\overline{w}_m \\ \sum\overline{u}_m\overline{w}_m \\ \sum\overline{v}_m\overline{w}_m \end{bmatrix}}_{\mathbf{c}} , \quad (2.15)$$

where the summation range $m \in 1(1)M$ is implied.

The above system is readily solved for the unknown coefficients $\mathbf{b} = \{b_0, b_1, b_2\}$. The first step is computing the inverse of the matrix $\mathcal{A}$, *if it exists*, to get an expression for the vector of unknowns $\mathbf{b}$. Since the matrix $\mathcal{A}$ is symmetric ($\mathcal{A} = \mathcal{A}^T$) with large positive diagonal entries we suspect that it is always positive definite. This is a useful property in that all positive definite matrices are invertible. Recall that given a real non-zero column vector $\mathbf{q}$ a matrix is positive definite *if and only if* the condition $\mathbf{q}^T\mathcal{A}\mathbf{q} > 0$ holds for any choice of $\mathbf{q}$; that is if its quadratic form is always positive (Lay, 2010). Carrying out the calculation of the quadratic form of the matrix in (2.15) the condition reads

$$\begin{aligned} 0 < &\left(q_1^2 M + q_2^2\sum\overline{u}_m^2 + q_3^2\sum\overline{v}_m^2\right) \\ &+ 2\left(q_1q_2\sum\overline{u}_m + q_1q_3\sum\overline{v}_m + q_2q_3\sum\overline{u}_m\overline{v}_m\right) . \end{aligned} \quad (2.16)$$

In the above expression the terms in the first bracket are all positive definite. In particular the terms in the summations involve the squares of block averaged components of the measured wind speeds which will always be positive. Conversely, the terms in the summations in the second bracket can take either sign in that the block averaged horizontal wind can blow in any direction. Moreover, for the last term we do not expect any significant correlation in the horizontal wind components so this too should be low. Hence the terms in the second bracket should be small relative to the first regardless of $\mathbf{q}$. As such, we expect $\mathcal{A}$ to be positive definite and hence invertible provided that we have a sufficient number of averaging periods M. To demonstrate positive definiteness more rigorously we consider the conditions under which a Cholesky decomposition (see e.g. Brezinski and Tournès (2014)), and hence positive definiteness, of the scaled matrix $\mathcal{G} = \frac{1}{M}\mathcal{A}$ is satisfied. Using the definition of the ensemble block average and that the decomposition takes the form $\mathcal{G} = \mathcal{L}\mathcal{L}^T$ where

[6] Unless the ensemble block averaged measured vertical velocity is identically zero, in which case the planar frame is already defined, this solution results in an instrument offset that is unreasonably large.

$\mathcal{L}$ is a lower triangular matrix, we find the following as the only possible analytical solution for $\mathcal{L}$

$$\mathcal{L} = \begin{bmatrix} 1 & 0 & 0 \\ \langle \overline{u} \rangle & \sqrt{\langle \overline{u}'^2 \rangle} & 0 \\ \langle \overline{v} \rangle & \frac{\langle \overline{u}'\overline{v}' \rangle}{\sqrt{\langle \overline{u}'^2 \rangle}} & \sqrt{\langle \overline{v}'^2 \rangle - \frac{\langle \overline{u}'\overline{v}' \rangle^2}{\langle \overline{u}'^2 \rangle}} \end{bmatrix} . \quad (2.17)$$

Here we have used Reynolds averaging rules along with the ensemble notation presented in Finnigan et al. (2003) to write $\overline{u}'_m = \overline{u}_m - \langle \overline{u}_m \rangle$ as the deviation of the block average of block m from the ensemble average[7] to arrive at the simplified form $\langle \overline{u}'^2 \rangle = \langle \overline{u}^2 \rangle - \langle \overline{u} \rangle^2$ and similarly for the $\langle \overline{v}'^2 \rangle$ and $\langle \overline{u}'\overline{v}' \rangle$ terms. Now $\mathcal{L}$ only provides a valid decomposition so long as the entries are real and the diagonals are non-zero. So we require that all the radicands in (2.17) be positive. For the non-negative ensemble variances $\langle \overline{u}'^2 \rangle$ and $\langle \overline{v}'^2 \rangle$ this is satisfied unless these are zero, which is improbable as this would imply a completely steady block averaged wind field over the entire ensemble. Thus, it is the sign of the radicand involving the ensemble covariance, $\langle \overline{u}'\overline{v}' \rangle$, that determines if $\mathcal{G}$ has a Cholesky decomposition which in turn would imply that it is positive definite and hence invertible. In requiring this radicand to be positive the necessary condition for $\mathcal{G}$ to be positive definite boils down to

$$\langle \overline{u}'^2 \rangle \langle \overline{v}'^2 \rangle > \langle \overline{u}'\overline{v}' \rangle^2 . \quad (2.18)$$

We explained earlier that no significant correlation between the horizontal wind components is to be expected over a long period. So the ensemble covariance term is likely to be small relative to the product of the ensemble variances. Hence the necessary condition is likely to always be satisfied for $\mathcal{G}$. Furthermore, as $\mathcal{G}$ is just a scaled version of $\mathcal{A}$ this implies that $\mathcal{A}$ is also positive definite and hence invertible.

Having found the inverse, $\mathcal{A}^{-1}$, (which we are now fairly sure exists) we can simply apply (left) matrix multiplication of the inverse to both sides of (2.15) to get the expression $\mathbf{b} = \mathcal{A}^{-1}\mathbf{c}$; by virtue of the definition of the matrix inverse. In the module the inverse is found using a lower-upper Doolittle factorization algorithm also described in Brezinski and Tournès (2014). After the computation we control that the result is the inverse by checking that the product of the matrix output by the algorithm and the matrix $\mathcal{A}$ is the identity matrix. If this is not the case the module stops and returns an error message[8]. As no such errors occurred for any of our runs we conclude that the algorithm worked properly and the matrix was always found to be invertible implying that the necessary condition for positive definiteness (2.18) was always satisfied.

Once $\mathbf{b}$ is computed since $b_1 = -k_x/k_z$, $b_2 = -k_y/k_z$ and $|\widehat{\boldsymbol{k}}| = 1$ we can easily find the elements of the planar unit vector $\widehat{\boldsymbol{k}}$ as defined in the system frame. For completeness, upon solving the simultaneous equations for k_z we get

$$k_z = \left(b_1^2 + b_2^2 + 1\right)^{-1/2} ,$$

from which k_x and k_y are readily found by insertion of k_z into the definitions of the now known coefficients b_1 and b_2. Having computed $\widehat{\boldsymbol{k}}$ the ensemble streamline *plane* is now defined, but the unit vectors $\widehat{\boldsymbol{\imath}}$ and $\widehat{\boldsymbol{\jmath}}$ must also be determined to set the planar frame. Recalling that α, the angle between due east and the sonic x-axis, has already been computed in the orientation routine we may conveniently incorporate this angle when defining an intermediate geographically referenced planar frame. We do so by noting that the unit vector $\widehat{\mathbf{E}} = \cos(\alpha)\widehat{\mathbf{x}} - \sin(\alpha)\widehat{\mathbf{y}}$ points due east in the system frame and hence define a (right-handed) orthonormal basis, the planar frame, by letting

$$\widehat{\boldsymbol{\jmath}} = \frac{\widehat{\boldsymbol{k}} \times \widehat{\mathbf{E}}}{|\widehat{\boldsymbol{k}} \times \widehat{\mathbf{E}}|} , \quad (2.19)$$

and

$$\widehat{\boldsymbol{\imath}} = \widehat{\boldsymbol{\jmath}} \times \widehat{\boldsymbol{k}} .$$

Note the normalization required in (2.19) as a consequence of $\widehat{\boldsymbol{k}}$ typically *not*[9] being orthogonal to due east. Now that the unit vectors that set the planar frame are defined in the system basis we can easily rotate our velocities into the planar frame from the instrument frame by using the dot product to pick out the velocity components, i.e.

$$\begin{aligned} u_p &= \widehat{\boldsymbol{\imath}} \cdot \mathbf{v} = i_x u + i_y v + i_z \left(w - b_0\right) , \\ v_p &= \widehat{\boldsymbol{\jmath}} \cdot \mathbf{v} = j_x u + j_y v + j_z \left(w - b_0\right) , \\ w_p &= \widehat{\boldsymbol{k}} \cdot \mathbf{v} = k_x u + k_y v + k_z \left(w - b_0\right) . \end{aligned} \quad (2.20)$$

Having computed and performed the controls and steps from (2.11) to (2.20) the module has effectively implemented a planar fit, rotating the instantaneously measured velocities into an ensemble streamline frame without the need for any rotation angle (apart from α).

To summarize, we have now successfully defined a planar frame using a least squares plane regression. In our new frame the principal normal $\widehat{\boldsymbol{k}}$ is perpendicular to the mean streamline plane in which the ensemble averaged plane normal velocity is minimized (approximately zero). It is this plane normal velocity that we adopt as our new vertical velocity, where we have simultaneously accounted for surface tilt, sensor tilt *and* sensor offset in the regression. Moreover, the pair of

[7] Not to be confused with $\overline{u'} = 0$: the block averaged deviation from the block average in a given block is always identically zero.

[8] The error message was verified by temporarily setting $\mathcal{A}$ to the zero matrix which is singular.

[9] In the rare case that it is (orthogonal to east) the instrument frame and the planar frame actually coincide and the above calculations are unnecessary apart from as a control.

plane unit vectors, $\widehat{\boldsymbol{\imath}}$ and $\widehat{\boldsymbol{\jmath}}$, are mutually orthogonal to each other and $\widehat{\boldsymbol{k}}$ such that the pair lie in the ensemble streamline plane with components parallel to due east and due north respectively. We highlight the fact that here, as with all remaining processing, any instant with velocities flagged as spikes and/or winds with the tower and boom structure upstream were excluded from the planar fit to avoid skewing what is intended to be the ensemble streamline plane. In fact, any block consisting of $> 10\%$ faulty data is ignored (not weighted) in the planar fit as the block is considered suspect and is later discarded from any subsequent analysis. Henceforth we drop the subscript p for the planar frame, making it implicit unless otherwise specified. Furthermore, when referring to horizontal or plane in the observational context we no longer allude to a plane of constant geopotential, but rather the ensemble streamline plane.

We finish off the section by demonstrating the effects of applying the planar fit algorithm to eddy covariance data. Figure 2.7 clearly demonstrates a strong negative correlation, square correlation factors $R^2_{\overline{U}\overline{w}}$ are displayed in the figure, between the block averaged horizontal velocity component in the downslope facing direction[10] and the vertical velocity in the instrument frame. This is as expected in that the leveled sonic anemometer is facing *upslope* as previously depicted in Figure 2.6. The corresponding correlation is eliminated by rotation into the planar frame. In addition, both the composite ensemble average $\langle\overline{w}\rangle$ and variance $\sigma^2_{\overline{w}}$ of the block averaged vertical velocity are reduced by over an order of magnitude in the planar frame. The reason the composite ensemble average is not exactly zero in the planar frame is that the planar fits are implemented before the quality controls (QC) in which entire blocks are eliminated based on a set of quality flags outlined in section 2.6. Before the QC each ensemble average vertical velocity is identically zero as can be verified in the log files for each run.

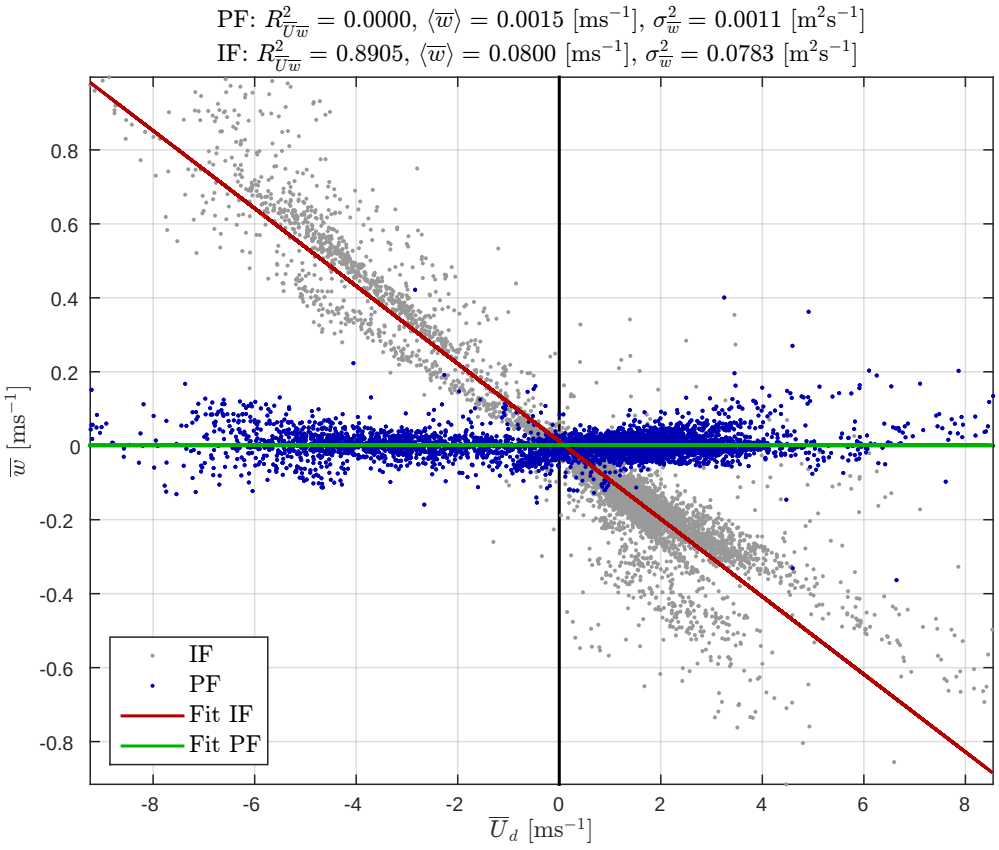


Figure 2.7: *Block averaged vertical velocity, $\overline{w}$, as a function of the downslope, wind direction (blowing from) $\varphi_D = 173.5°$, components of the horizontal block averaged wind, $\overline{U}_d$, in the planar frames (PF) and the instrument frame (IF). A composite of* 11 *planar fits with a total of* 6067 *blocks, each with a length of* 30 *minutes, of processed EC data from the Bayelva site. The red/green lines represent the IF/PF linear best fit.*

[10]The downslope direction (bearing) is computed based on the average (over all 11 planar fits) of the normalized plane downslope vectors $\widehat{\mathbf{D}} = -\nabla\eta/|\nabla\eta|$ where $\eta = -xk_x/k_z - yk_y/k_z y + H$ is the parametric equation for a given ensemble streamline plane with H as the height of the sonic anemometer above ground.

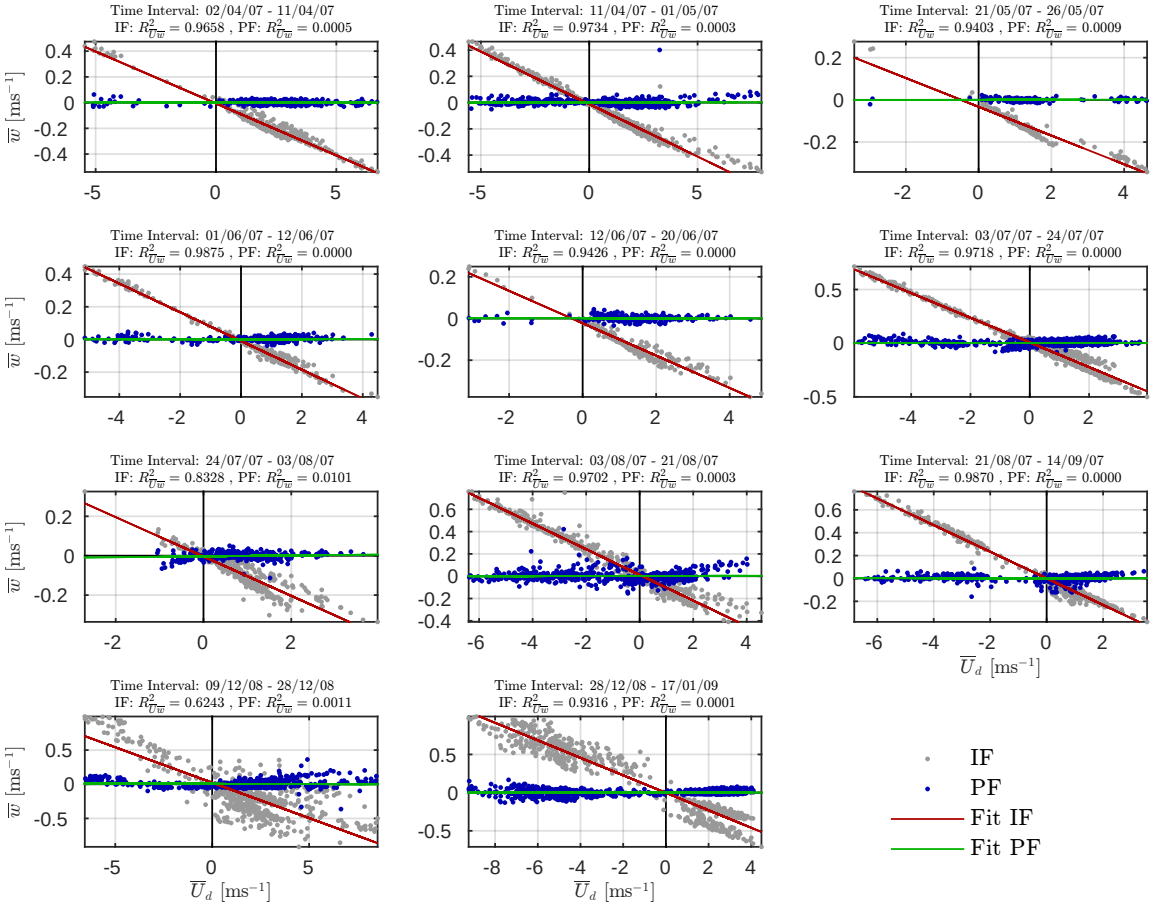


Figure 2.8: *Visualization of the* 11 *planar fits on which the composite in Figure 2.7 is based with the same axes and labels but on a file by file (fit by fit) basis with the time interval of each fit included as a reference.*

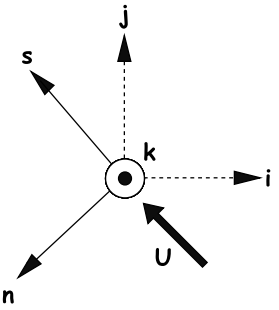
The results indicate that we have indeed identified well defined ensemble streamline *planes* in the local surface layer appropriate for diagnosing eddy covariance fluxes representative of the local surface. We emphasize planes in that the composite scatter in Figure 2.7 based on data from multiple files. Each file with a distinct planar fit (distinct $\widehat{\boldsymbol{k}}$) as depicted in Figure 2.8, where only files spanning at least 5 days were included following the recommendation of Mauder and Foken (2004). Despite being distinct, due to differences in the distributions of the measured wind fields in each file, the planes share the common characteristic that they all mainly follow the slope of the local terrain. That is none of the planes are far off the span of $[\widehat{\imath}, \widehat{\jmath}]$ previously depicted in the idealized Figure 2.6. Recall that for the Bayelva site $\gamma < 5°$ (Westermann et al., 2009). For comparison, the mean slope angle, the angle between slope and the true geopotential horizontal, of the planes in Figure 2.8 is $\gamma = 8°$ with a standard deviation of $1°$. The discrepancy between the actual surface slope and the plane slope is likely due to several factors: 1) the instrument itself being slightly off-level (order of one degree), 2) the planes are defined at a (maximum) height of 2.75 m AGL above a surface which is not completely smooth and changing in time and 3) differences in synoptic scale forcing in each run. Similarly we can attribute the standard deviation in the plane slope angles to factors 2) and 3).

Finally, we reiterate that we have opted for performing the planar fit on a file by file basis both due to ease of implementation, that is files can be read in and processed individually, and due to the location of our study site. Specifically due to the possibility of snowfall (or snow-melt) changing the slope of the local surface on which the planar frame is based. If we instead were to opt for periods longer than those in each file, the maximum raw data file lengths are of the order of three weeks, we would stand the risk of introducing significant changes in the measurement height. A prerequisite for using the planar fit method is that the instruments not move (or be moved) relative to the local surface (Wilczak et al., 2001). Ideally a so-called sectorial planar fit would have been implemented as in Siebicke et al. (2012) where the planar fit is applied for individual (block averaged) wind direction bins. Such a sectorial

planar fit provides a means of investigating the effects of flow distortion in the so-called 'closed sector' (Li et al., 2012) for non-omnidirectional sonic anemometers like the CSAT3 that was used for this study. This proved difficult for our study site for the same reason that fits over longer periods were risky. Moreover, if we chose to do a sectorial planar fit on a file by file basis where we can assume that changes in measurement height are small we would end up with too few measurements in many of the wind direction bins to create the representative ensembles suggested in Mauder and Foken (2004). As such, we did not weight any of the blocks where the wind blew from the 'closed sector' where the sonic head and mast were upwind of the measurement area. We will pick up on how these wind directions are identified in Section 2.6.

2.3.3 Natural Ensemble Streamline Frame

Figure 2.9: *Schematic of the geographically referenced planar frame* $\left[\widehat{\boldsymbol{\imath}}, \widehat{\boldsymbol{\jmath}}, \widehat{\boldsymbol{k}}\right]$ *and the streamwise planar frame* $\left[\widehat{\mathbf{s}}, \widehat{\mathbf{n}}, \widehat{\boldsymbol{k}}\right]$ *where* $\widehat{\mathbf{s}}$ *is the streamwise unit vector pointing along the segment plane mean wind* $\mathbf{U}$.



A final rotation is available in which the plane unit vectors are rotated such that the first, the streamwise unit vector, points along the segment plane mean wind and the second, the cross-stream unit vector, points along the segment plane cross-wind. At first glance such a rotation seems identical to the more traditional natural wind frame. The classical micrometeorological natural wind frame, a so-called 'double rotation', detailed in McMillen (1988) rotates the frame into the segment mean wind. Here, on the other hand, we are rotating into the segment *plane*[11] mean wind. The difference is subtle but crucial. The final rotation brings us into a natural ensemble streamline frame (NESF), where $\overline{v} = 0$, not the classical natural frame, where $\overline{v} = \overline{w} = 0$ for each block. As discussed at length in Finnigan (2004), the classical natural wind coordinate is poorly suited for long term (months or more) studies in sloping terrain. In part due to the risks of the unphysical overrotation that occurs when forcing $\overline{w}$ to zero in each block, but also due to the difficulty in calculating consistent long term ensemble statistics when the orientation of the 'vertical' changes for each block. In our case the long-term fit to an ensemble streamline plane eliminates the risk of over rotation. We have included the final rotation in a separate routine. This due to the fact that segment sizes used in the subsequent analysis tends to vary. For example, block statistics are typically produced for 30 minute blocks while spectra are produced for 3 hour segments. In both the example outputs the final rotation can be used, but it is meaningless for spectra if the rotation is based on half hour means or for block statistics if the rotation is based on 3 hour mean winds. Hence the segment length on which the rotation is based is output dependent and thus not an intrinsic part of the planar fit algorithm. Instead it can be used in the various output routines.

For completeness we document how the final rotation is carried out analytically. Let

$$\mathcal{U}_m = \sqrt{\overline{u}_m^2 + \overline{v}_m^2}\,,$$

denote the segment plane mean wind speed, i.e. the magnitude of $\mathbf{U}$ in Figure 2.9, for the m-th segment. Then the streamwise unit vector for a given segment is given by

$$\widehat{\mathbf{s}}_m = \frac{\overline{u}_m}{\mathcal{U}_m}\widehat{\boldsymbol{\imath}} + \frac{\overline{v}_m}{\mathcal{U}_m}\widehat{\boldsymbol{\jmath}}, \tag{2.21}$$

and the cross-stream, or bi-normal, unit vector for a given segment is given by

$$\widehat{\mathbf{n}}_m = \widehat{\boldsymbol{k}} \times \widehat{\mathbf{s}}_m\,. \tag{2.22}$$

Having computed the streamwise and cross stream unit vectors for a given segment we can easily rotate the plane instantaneous velocities into streamwise and cross-stream components as follows

$$u_n = \widehat{\mathbf{s}}_m \cdot \mathbf{v} \quad \text{and} \quad v_n = \widehat{\mathbf{n}}_m \cdot \mathbf{v}\,,$$

where we have let the subscript n denote the plane velocity components in the natural ensemble streamline frame. Note the explicit use of the subscript m on the natural unit vectors. This is to remind us that the streamwise direction changes for each segment (block or window) in line with the plane mean wind changing between segments, that is the direction of $\mathbf{U}$ in Figure 2.9 is *not* constant in time. The geographically referenced planar frame $\left[\widehat{\boldsymbol{\imath}}, \widehat{\boldsymbol{\jmath}}, \widehat{\boldsymbol{k}}\right]$, in contrast, is fixed for all segments in a given series. Which of these two frames to use is more or less a matter of taste. In the present study the results from the data analysis will be displayed in the natural as opposed to the geographically referenced planar frame unless otherwise specified as this

[11] The plane to which $\widehat{\boldsymbol{k}}$ is the principal normal: the ensemble streamline plane.

is the common approach (Lee et al., 2006), which will aid in a consistent comparison of our results with those from other campaigns. Before closing the discussion on the final rotation we wish to highlight that the natural frame has one major drawback: the 'streamwise' direction is not always well defined. That is mathematically, provided $\mathcal{U}_m \neq 0$, we can always use (2.21) and (2.22) to define a natural frame. Physically, however, it does not always make sense to refer to directions as 'streamwise' and 'cross-stream' within a given segment. This particularly in the case of segments where the instantaneous plane winds are rapidly veering and/or have low magnitudes. We will come back to these considerations in section 2.6.

2.4 Ancillary Variables

In addition to the variables measured by the EC system, the module employs a set of ancillary (supporting) variables when computing fluxes and the M-O stability parameter. Of these variables the measurement height z_m and the air pressure P are obtained directly from secondary measurement sources; that is not from the EC system itself. Conversely, the absolute temperature T, the virtual temperature T_v, the air density ρ and the specific humidity q are diagnosed from the measured variables through simple considerations of the ideal gas law and the definition of sonic temperature. In the following we briefly descibe the source, implementation and diagnosis of the complete set of ancillary variables employed in the module.

2.4.1 Pressure

No fast responding pressure measurements were available from the EC system at Bayelva and, consequently, a calculation of the fluctuating pressure p' was not possible. This is somewhat unfortunate as strictly speaking an estimate of the turbulent vertical pressure flux $\overline{p'w'}$ is required when applying the WPL correction (Section 2.5.4). The lack of fast pressure measurements is not surprising as these are notoriously difficult (Wyngaard, 2010) and by no means a standard component in EC systems. As noted by Wilczak and Bedard Jr (2004) the difficulty is in large part due to the fluctuations being on the order of one part in 10^5 of the mean requiring highly sensitive instrumentation. For comparison, the resolution of the CSAT3 is one part in 10^4 for horizontal wind fluctuations with respect to their mean[12] (CSAT3, 2014) which is more than adequate when estimating typical velocity fluctuations which are on the order of one part in 10^2 of the mean. Now despite being difficult fast responding barometery is possible with microbarometers discussed in e.g. Wilczak and Bedard Jr (2004) and Zhang et al. (2011), however, these microbarometers require high maintenance and are thus not ideal for a largely autonomous EC system like that at Bayelva.

In the absence of a microbarometer we turned to the next best thing namely slow pressure measurements from a nearby source. A suitable source of slow pressure was, as with the monthly temperature climatology (Section 2.2.1), the Ny Ålesund AWS. Thereby, through eKlima (2014) we downloaded station level pressure from the AWS for the period 01.01.2007-31.12.2010. The AWS reports 1-minute average pressure four times a day; at 00:00, 06:00, 12:00 and 18:00 UTC. To facilitate allocation of the slow pressure in the module we converted the text file provided through eKlima (2014) into a NetCDF file via a seperate Fortran routine. Subsequently the entire slow pressure series along with the associated timestamps could be read into the module quickly after which the pressure series was synchronized (discarding superfluous data) with the period under consideration for EC calculations. The synchronization was carried out by assigning the sampled slow pressure to all data blocks within $+/-$ 3 hours of the sample time. In the following we will denote the slow pressure P to differentiate it from instantaneous pressure p. We will also assume that the (30 minute) block averaged pressure $\overline{p}$ is approximately equal to the slow pressure P in the corresponding 6 hour interval. We do so on the grounds of the relative proximity between the AWS and the EC system (2 km apart), the claim in Wilczak and Bedard Jr (2004) that pressure fluctuations are typically on the order 1 [Pa] within a typical block averaging period and also under the assumption that the block averaged pressure field varies slowly. The latter assumption does not hold in general; nonetheless since near statistical stationarity is a prerequisite for the EC method (e.g. Foken and Wichura (1996)) then if the assumption is clearly violated the applicability of the EC method is in itself questionable. At a later stage (Section 2.6) we quantify the stationarity of the fast responding variables, which are all coupled with pressure, and discard instances that are clearly non-stationary. Thereby, at least for the data that passes the quality control, the assumption of a slowly varying mean pressure field is well grounded.

[12] Given a mean wind speed of 10 [ms^{-1}] and the resolution of 10^{-3} [ms^{-1}] listed in CSAT3 (2014)

2.4.2 Measurement Height

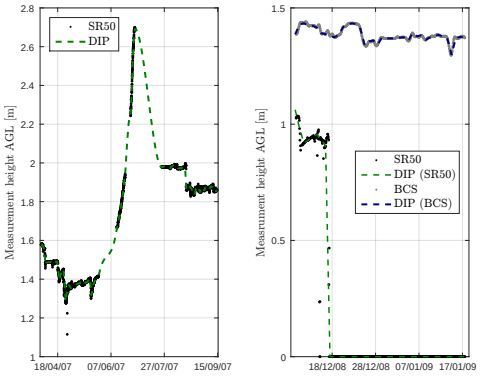


Figure 2.10: *Measurement height above ground (*snow*) level for the two batches of data (left panel: 02.04.2007-14.09.2007, right panel: 09.12.2008 to 17.01.2009). Measurements are displayed from both the SR50 (black dots) and the BCS (grey dots). The dashed lines are the result of diurnal interpolation using* `MATLAB`*'s Piecewise Cubic Hermite Interpolating Polynomial (the* `pchip()` *function) abbreviated DIP for the SR50 data (green) and the BCS data (blue).*

The measurement height, z_m, of the EC system was tracked using a Campbell SR50 (SR50, 2007) sonic ranging sensor. Having access to hourly averaged data from the SR50 we were able to read in and allocate the z_m values to the module allowing for a proper estimation of the M-O stability parameter $\zeta = z_m/L_*$ at the appropriate height. For the sake of convenience before reading in the heights these were low pass filtered in a seperate `MATLAB` routine through the application of a diurnal interpolation (DIP) such that the module only had to contend with one z value per day as opposed to one value per hour (see Figure 2.10) . This DIP procedure also enabled us to estimate the measurement height in the periods where data from the SR50 was missing. Of course such an estimation is not ideal; when interpolating we assume that the accumulation (or melting of snow) is linear as opposed to a sudden increase or decrease that is more often the case in reality. Nonetheless, the interpolated values, where data was previously missing, hopefully provide a better estimate for z_m than a wild guess or a constant value.

Unfortunately the SR50 was not working properly during most of the second batch of data (see right panel in Figure 2.10), reporting a constant measurement height of zero meters above ground level. If this really were the measurement height then both the CSAT3 and the LI-7500 would be covered by snow. As such, the measurement paths would be blocked and the respective output raw data would be completely bogus. Instead both these instruments reported perfectly valid aerial data during the period; so the measurement height was clearly not zero. Fortunately we found a secondary source for the measurement height during this period namely the nearby Bayelva climate station (BCS)[13] where the local snow depth, d_s, was recorded. Using the snow depth d_s the measurement height was estimated through $z_m = z_r - d_s$ where $z_r = 2.75$m is the measurement height of the CSAT3 in the absence of snow. This height was also low pass filtered using the DIP as shown in the right panel of Figure 2.10. Of course this height will only be an estimate in that we do not expect the local snow depth to be the same at BCS as it is below the EC system. After all as displayed in Figure 1.4 the BCS is situated on flat ground on top of Leirhaugen hill approximately 100 meters to the south-south west of the EC system which we recall is mounted on the northwestern slope of the same hill. As such there may at times be a significant difference in the accumulation of snow at the two locations as governed in part by the wind field and subsequent snow drift. This is examplified by the difference between the BCS estimate and the SR50 estimate in the early stages of the second batch of data in Figure 2.10 where the snow is 0.4 m deeper below the EC system. We did not 'correct' for this difference as we are unsure how much to trust the SR50 data in this period considering that it suddenly drops out (to zero) soon after. Still, it is encouraging from a rough eyeball of the start of the two curves in Figure 2.10 that the two snow depth measurements are correlated to a certain extent. Furhtermore, the use of the BCS snowdepth to estimate z_m is a better option than either having to discard almost the entire second batch, which we would have to do if the measurement height really were zero, or alternately employing a constant measurement height.

2.4.3 Diagnosed Variables

The sequential diagnosis of ancillary variables required in the subsequent processing steps is completed in the presented order and relies soley on the measured variables at our disposal. For brevity we will simply present the equations and refer to Appendix A.2 for a detailed derivation and a discussion of the underlying assumptions. First the instantaneous nondimensional specific humidity (q) is estimated via

$$q = \left(\frac{P}{R_d \rho_v T_s} - 0.1 \right)^{-1} . \qquad (2.23)$$

where $R_d = 287\,[\mathrm{JK^{-1}kg^{-1}}]$ is the specific gas con-

[13] Data recovered from the url:`http://doi.pangaea.de/10.1594/PANGAEA.805533` where it was uploaded as a supplement to Westermann (2010).

stant for dry air. Having found q using (2.23) the instantaneous air density (ρ) is readily diagnosed through $\rho = \rho_v/q$. Next by adapting the relation in Kaimal and Businger (1963) the instantaneous absolute temperature is calculated as

$$T = \frac{T_s}{1 + 0.51q} \,. \tag{2.24}$$

Subsequently, following Kaimal and Gaynor (1991) the instantaneous virtual temperature (T_v) is found from Kaimal and Gaynor (1991)

$$T_v = T\left[1 + 0.61q\right] \,. \tag{2.25}$$

In accordance with the remaining procedures in the module wherever the measaured variables are flagged as spikes we also flag all the ancillary variables as spikes and set these to -1000 accordingly. Finishing the section we point out that after being diagnosed only two of these variables (T and ρ) are used in the remaidner of the module. These are only used in their block averaged form in the flux corrections. Nonetheless, all four of the diagnosed ancillary variables are part of the module output and have been verified to satisfy physical expectations. For example, virtual temperature is found to always be higher than the corresponding absolute temperature as expected from the discussion in Appendix A.2. Now the stage is set to perform the necessary flux corrections, an extensive quality control of the results, cross-spectral analysis and cross-statistical analysis.

2.5 Flux Corrections

Before the module can produce reliable estimates of dynamic fluxes in relation to, for example, the surface energy budget a series of flux corrections need to be applied. These corrections account for the effects of sensor separation (Vickers and Mahrt (1997); Eugster et al. (1997)), high and low frequency cospectral attenuation of the measured signals (Moore (1986); Horst (1997); Massman (2000)), sonic thermometry through the so-called'SND' correction (Schotanus et al. (1983); Liu et al. (2001)) and density fluctuations through the so-called 'WPL' correction (Webb et al. (1980); Lee and Massman (2011)). Due to the interdepdence of many of these corrections it is standard practice[14] to apply these iteratively (e.g. Mauder and Foken (2004); Nordbo et al. (2012); Burba (2013)). In the following subsections we will briefly explain the underlying theory, motivation and application of each correction after which the procedure in our modules' combined flux correction approach is laid out. Before the blockwise flux corrections are undertaken the instantaneous plane winds u, v are rotated for convenience, via the routine described in Section 2.3.3, so that u points along the block plane mean wind.

2.5.1 Sensor separation

As discussed in section 1.2.3 the EC system was setup such that the CSAT3 and the LI-7500 were separated a horizontal distance of approximately $d = 0.22\,\mathrm{m}$ (S.Westermann personal communication). Now this separation poses somewhat of a problem for the water vapor flux ($\overline{\rho_v' w'}$) as the theory assumes that the measurements are colocated in space and synchronized in time. If we were to simply accept $\overline{\rho_v' w'}$ in its separated form, that is with synchronzied measurements but at different points in space, we would undoubtedly underestimate the magnitude of this flux. The reason being that when turbulence is well developed the fluctuations of a scalar and the velocity are usually nearly in (or completely out) of phase at a given point in space and time (Stull, 1988). This effect will be more pronounced for small eddies (with a length scale $\ell << d$) when compared to the large integral scale eddies (length scale ℓ_0). As such the relative effect, with respect to a colocated measurement, is small but certainly not negligible especially in stable stratification, as is relatively common at Bayelva, when the integral length scale is shifted towards smaller wavelengths (Kaimal and Finnigan, 1994). This is in fact the main and logical conclusion reached by Kristensen et al. (1997): the loss of vertical flux due to sensor separation depends mainly on the ratio d/ℓ_0; the smaller the ratio the smaller the loss. So how do we account for the flux loss due to sensor separation?

To our knowledge, as noted in Massman (2000), there is to date no straight forwards way to correct for the *vertical* separation between the sensors. Yet, the former author emphasizes the recommendations from the model study of Kristensen et al. (1997) who find that a minimization of scalar flux loss is achieved if the scalar sensor (LI-7500 in this case) is placed just *below* the velocity sensor (CSAT3). A compelling example is provided in Kristensen et al. (1997) where a vertical separation of $D = 0.2$ m is shown to lead to a flux loss of as much as 18% if the scalar sensor is placed above

[14] The exact form of the corrections often differs between processing packages (cf. Van Dijk et al. (2004) and Mauder and Foken (2004)); although the order is usually the same as that presented here.

the velocity sensor, whereas with a corresponding separation the flux loss is only 2% if the scalar sensor is the lower sensor. Fortuitously, the EC system employed in our study was setup with this recommendation in mind. So, with the LI-7500 below the CSAT3, the effects of vertical separation on flux loss are likely to be neglible with respect to other loss factors.

What remains then is to account for the flux loss introduced by the horizontal separation of the sensors. One way of doing this is by calculating the cross-covariance between the absolute humidity and the vertical velocity within each averaging period by moving the two series against eachother (Nordbo et al., 2012); that is lagging one series with respect to the other by a consant lag factor. With the previous claim in mind, that in the case of developed turbulence the fluctuations of a scalar and vertical velocity are typically either in phase or completely out of phase, we should be able to estimate the 'true' flux by finding the lag that maximizes the square of the cross-correlation. In doing so we are also making the implicit assumption that Taylors' Hypothesis (Taylor, 1938) holds, i.e. that small scale turbulence is 'frozen' (conserved) under advection by the mean wind. To the degree that this assumption holds we should be able to recover the flux loss due to sensor separation by finding the lag at which we would effectively be sampling the turbulent signals at the same point in time *and* space. This is the idea behind the so-called maximum cross-correlation method proposed in the context of EC measurements by Vickers and Mahrt (1997) and Eugster et al. (1997).

To proceed with this method in a computationally economical fashion we first need to make use of some geometry. We recall that the boom on which the LI-7500 was mounted pointed into a bearing of $\varphi_L = 245°$ while the CSAT3 was pointed into a bearing of $\varphi_C = 215°$; where both bearings are defined clockwise from North in the SLEC horizontal frame. Using these angles it is straightforwards to compute a separation unit vector $\widehat{\mathbf{d}}$ that points from the CSAT3 to the LI-7500. This is done by taking the bearing of the midpoint between these two bearings $\varphi_{MP} = \varphi_C + (\varphi_L - \varphi_C)/2$, rotating the midpoint bearing by a further $90°$, i.e. $\varphi_d = \varphi_{MP} + 90°$. Next the resultant bearing is converted into a polar angle in the SLEC frame via a simple transformation $\varphi_d \rightarrow \alpha$ from which the unit vector is defined $\widehat{\mathbf{d}} = \cos(\alpha)\,\widehat{\mathbf{E}} + \sin(\alpha)\,\widehat{\mathbf{N}}$ where $\widehat{\mathbf{E}}$ and $\widehat{\mathbf{N}}$ are the unit vectors pointing due east and due north respectively in the SLEC frame. This calculation need only be done once within the orientation routine (section 2.3.1) from which $\widehat{\mathbf{d}}$ can be passed at leisure to this flux correction routine on a block-by-block basis. To compliment this unit vector the normalized block averaged horizontal wind vector $\widehat{\mathbf{u}}$, henceforth wind unit vector, as defined in the SLEC frame using the wind direction φ is calculated for each block.

The purpose of having these two unit vectors is to determine to what degree the block averaged wind blows from one instrument to the other, and crucially to determine which of the two instruments (if any) is downwind of the other. By definition if $\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}} > 0$ then, on average in the given block, the IRGA is (to a certain degree) downwind of the CSAT3 and vice versa if $\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}} < 0$. Nonetheless, in the case that $|\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}}| < 0.5$ the component of the mean wind vector that blows between the two instruments is less than half the magnitude of the component that blows across the separation path. Geometrically this corresponds to cases where the smallest angle $\phi = \arccos(|\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}}|)$ between the axis that spans the separation vector and the axis that spans the wind unit vector exceeds $60°$. Due to horizontal gradients in turbulent velcoties it does not make sense to apply any separation correction in such cases since the small scale 'frozen' turbulence is not advected from one instrument to the other. Instead it is mainly advected across the path that is separating the instrument and so the small scale turbulence measured at one instrument will not be cross-correlated with that measured at the other instrument. We have chosen the threshold of $\phi \leq 60°$ based on sensitivity tests on example data where we noticed that for angles much higher than the threshold the cross-correlation rarely showed a distinct maximum. Conversely, if it was much lower we would underestimate the flux in many instances where a distinct non-zero lag cross-correlation maximum was present. Unavoidably the exact threshold is somewhat arbitrary, but the order of magnitude is physically justifiable. To summarize then the ensuing correction only deals with instances where the mean wind has a considerable component paralell to the path separating the instruments.

Now some definitions are needed to clarify the previous allusions to cross-statistics and to enable an algorithmic presentation of the correction. As a first step a LDT is applied to the block segments of both w and ρ_v. The reason for the LDT is that cross-statistics are, as mentioned in Finkelstein and Sims (2001), highly sensitive to low frequency trends. Thereby it is advisable to remove these (Mauder et al., 2013) via a high pass filter such as a LDT. As usual the LDT is internal to this routine. Next, with similar notation as in Fuller (1996) we define the block cross-covariance between the LDT two variables of interest (w'' and ρ_v''), denoted $\gamma_{w\rho_v}$, as

$$\gamma_{w\rho_v}(\tau_L) = \overline{w''(t)\rho_v''(t+\tau_L)}\,.$$

In the above the averaging operator is the block average as usual with primed quantities as the instantaneous deviations from the block averages while τ_L is the lag time which we define positive when the LI-7500 samples are lagged with respect to (considered at later times than)

the CSAT3 samples. When normalized by the product of block standard deviations we obtain the block cross-correlation

$$R_{w\rho_v}(\tau_L) = \frac{\gamma_{w\rho_v}(\tau_L)}{\sigma_w \sigma_{\rho_v}} .$$

As indicated both the cross-correlation and cross-covariance are functions of the lag time τ_L. In particular, the block covariance and block correlation are recovered in the case that the lag time is zero. In our case of a discreteley sampled time series the cross-covariance is computed, as in Finkelstein and Sims (2001), for a given 30 minute block with I total samples, N_F faulty (spike) samples and $N = I - N_F$ clean samples through the weighted block average:

$$\gamma_{w\rho_v,j} = \begin{cases} \frac{1}{N}\sum_{i=1}^{I-j} w_i'' \rho_{v,i+j}'' \delta_i \delta_{i+j} & \text{if } j \geq 0 \\ \frac{1}{N}\sum_{i=1}^{I+j} w_{i-j}'' \rho_{v,i}'' \delta_{i-j} \delta_i & \text{if } j < 0 \end{cases} , \tag{2.26}$$

where the index j (can be negative) is related to the discrete lag time $\tau_{L,j} = j\Delta t$ where $\Delta t = 1/f_s = 0.05\,\text{s}$ is the sampling time (inverse sampling frequency). The indexed weights δ_k are defined such that they are zero if the measurement with the corresponding index k is faulty (-1000) and one otherwise. We note also how (2.26) is consistent in both cases it converges to the block covarance as $j \to 0$. Moreover, from (2.26) it can also be seen that, contrary to the autocovariance, the crosscovariance is not an even function i.e. $\gamma_{w\rho_v}(\tau_L) \neq \gamma_{w\rho_v}(-\tau_L)$ (nor is it odd $\gamma_{w\rho_v}(\tau_L) \neq -\gamma_{w\rho_v}(-\tau_L)$). So the crosscovariance is not symmetric about zero lag.

Due to this asymmetry the choice of the sign of τ_L, i.e. which instrument to lag with respect to the other, is key in determining whether or not we are likely to find a maximum in the crosscovariance with a one-way lag (single signed τ_L). One-way lag is advantageous in the module as it will cut the number of computations in half. Fortunately we know from the sign of the dot-product $\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}}$ which instrument is upwind of the other and can thus safely set the sign of τ_L accordingly as positive if $\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}} \geq 0.5$ (LI-7500 downwind) and negative if $\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}} \leq -0.5$ (LI-7500 upwind). Furthermore, the magnitude of the dot product can be combined with the block plane mean wind speed $\overline{u}$ (as defined in Section 2.3.3) to estimate the travel time between the block averaged horizontal travel-time, τ_{tr}, between the two sensors as

$$\tau_{tr} = \frac{d}{|\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}}|\,\overline{u}} . \tag{2.27}$$

Inserting a typical value of $\overline{u} = 5\,\text{ms}^{-1}$ when the horizontal wind is blowing along the separation path yields $\tau_{tr} \simeq 0.05\,\text{s}$ which is very close to the sampling time Δt. When $\overline{u} \gtrsim 0.2\,\text{ms}^{-1}$ we are near the calm wind threshold (elaborated in section 2.6) and τ_{tr} is on the order of one second. So it is clear that a lag time larger than a second or two is not necessary when determining the maximum crosscorrelation. To set an upper limit on the lag time in each block the module rounds τ_{tr} up to the nearest integer (usually one second) and multiplies by two as well as the sampling frequency f_s to get the corresponding upper limit on the absolute value of lag index j in (2.26). The factor two is included to check that a distinct maximum is identified in the square cross-correlation. Maxima are discared if the square cross-correlation continues to increase for increasing lag time when $|\tau_L| > \tau_{tr}$, in such cases the cross-correlation technique is not well defined (Nordbo et al., 2012). If on the other hand a distinct maximum is found (as is typically the case) the lag index of this maximum, j_m, for the given block is passed to the module. In such a way the corresponding flux correction factor accounting for sensor separation is readily computed as

$$CF_s = \frac{\gamma_{w\rho_v,j_m}}{\gamma_{w\rho_v,0}} ,$$

from which the *non LDT* flux, $\overline{\rho_v' w'}$, is corrected through

$$\left(\overline{\rho_v' w'}\right)_C = CF_s \overline{\rho_v' w'} .$$

The separation correction factor is passed on a block by block basis to the module output to help track the magnitudes of these corrections.

To finish off we provide an example of this correction by including a cross-correlation analysis for one of the input files as depicted in Figure 2.11. Aiding in a visual understanding of the concept in this particular case a two-way lag has been carried out; while in the module only a one way-lag is needed with $\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}}$ known. This is clearly seen by the fact that in all cases the cross correlation maxima occur on the side of the zero lag line dictated by the sign of $\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}}$ (see figure text for details). For this example the correction resulted in a 4% median increase in flux for the case with negative lag flux maxima while for the positive lag maxima we had a 2% median increase in flux. It is somewhat encouraging that these flux correction factors are on the order of magnitude of those found in Nordbo et al. (2012) using a LI-7500 at a site in similarly heterogeneous (albeit urban) terrain. Furthermore, the median *absolute* lag for the cross-correlation maximum was only Δt (the sampling time) in both cases as is expected for a fast-responding open path IRGA (Leuning and Judd, 1996), but there are multiple weak-wind blocks ($\mathcal{U} \gtrsim 0.2\,\text{ms}^{-1}$) with $|\tau_{L,j_m}|$ exceeding $5\Delta t$ in both panels. The latter examplifies the need for a varying the upper limit on j_m in accordance with variations in τ_{tr}.

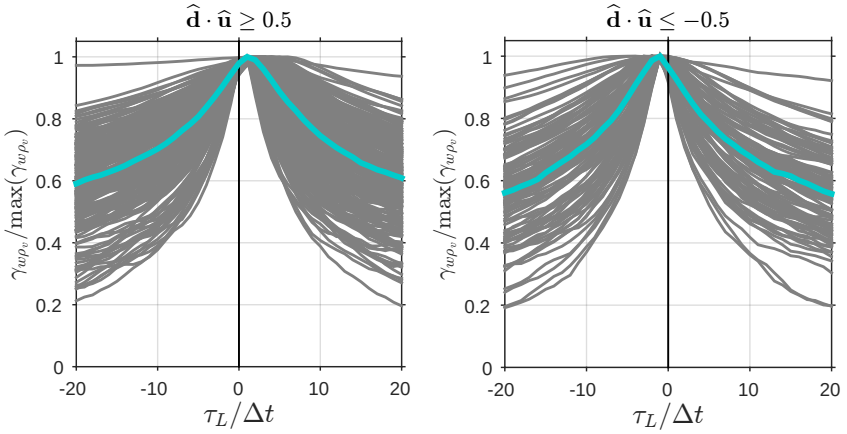
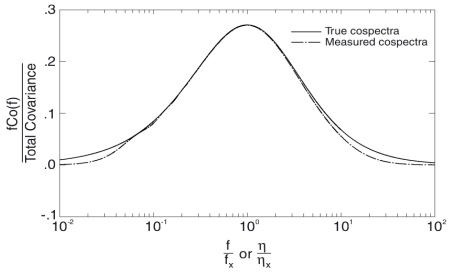


Figure 2.11: *Example of a two-way lag cross-correlation analysis for the period 03.07.2007-24.07.2007. Block cross-covariances $\gamma_{w\rho_v}$ normalized by maximum in each block (dark grey) as functions of lag time τ_L normalized by sampling time Δt, the thick turqouise line is the median normalized block cross-covariance. In the left panel $\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}} \geq 0.5$ implying that the LI-7500 is downwind of the CSAT3 so all maxima occur with $\tau_L > 0$ whereas in the right panel $\widehat{\mathbf{u}} \cdot \widehat{\mathbf{d}} \leq -0.5$ so the reverse is true. In both panels the zero lag line (black) is included as a reference.*

2.5.2 Cospectral Attenuation

Next in line are the cospectral attenuation corrections which account for the undesired attenuation (damping) of the measured flux at high and low frequencies in spectral space. These methods were pionereed by Moore (1986) and later revised and simplified by (among others) Horst (1997) and Massman (2000). After introducing the concept we will follow the approach of the latter due to its relative simplicity, ease of implementation and relevance for open path EC-systems (Lee et al., 2006).

Figure 2.12: *Idealization of the normalized true (solid line) and attenuated measured cospectra (dash-dotted line) as functions of frequency normalized by peak frequency. Figure adopted from Chapter 4 in Lee* et al. *(2006).*



The factors contributing to cospectral attenuation can be split into a high and low frequency component. For the high frequency component the major factors are path length averaging and sensor separation[15]. Conversely, for the low frequency component the attenuation factors are not related to instrumentation but rather the flux sampling in data processing (Kaimal et al., 1989) related to any detrending employed be it in the form of a linear detrend or mean removal. In summary, as noted by Lee et al. (2006), the instruments typically act as low pass filters limiting the resolution of small eddies (right of peak in Figure 2.12) whereas the flux sampling approaches in processing act as high pass filters limiting the resolution of the larger eddies (left of peak in Figure 2.12).

As in the previous section we will attempt to clarify the above with a brief discussion of the underlying theory. We begin as in Massman (2000) by noting that by definition the true (ensemble averaged) flux can be defined with respect to the integral

$$\overline{\xi' w'} = \int_0^\infty Co_{\xi w}(f)\, df\,, \tag{2.28}$$

where $Co_{\xi w}(f)$ is the true one sided (folded) cospectrum as a function of natural frequency f[Hz]. For now we will not dwell on the definition or physical interpretation of cospectra (or spectra) any further than noting

[15] Which we have already dealt with using a more direct method in the separation correction.

[16] Once more a word of caution. As pointed out in Stull (1988) the cospectrum is *not* the spectrum of the product of ξ' and u'.

that the cospectral density is a measure of the contribution to the flux (covariance) at a given frequency[16]. So the relationship between cospectral density and turbulent flux is, for example, analogous to the relationship between monochromatic (spectral) irradiance and total irradiance in radiative transfer[17]. We will denote the measured flux $\left(\overline{\xi' w'}\right)_a$ to emphasize (subscript a) the cospectral attenuation at high and low frequencies. Following Moore (1986), the measured flux may be expressed it in terms of the true cospecturm as

$$\left(\overline{\xi' w'}\right)_a = \int_0^\infty H(f) Co_{\xi w}(f)\, df\,, \tag{2.29}$$

where $H(f)$ is the transfer function of the system used to measure ξ and w, henceforth system transfer function, consisting of the product of the combined high and low pass transfer functions. That is $H = H_{LF} H_{HF}$ where $H_{LF} = \Pi_j H_{LF,j}$ is the product of all high pass transfer functions attenuating cospectral density at low frequencies and $H_{HF} = \Pi_j H_{HF,j}$ is the product of all low pass transfer functions attenuating flux at high frequencies. As implied by the above we assume that the system response is such that there is negligible attenuation in the (logarithmic) mid-frequency range i.e. where the frequency is close to that of the frequency of the cospectral peak (see Figure 2.13). The use of transfer functions comes naturally from the previously alluded to fact that the high/low frequency cospectral attenuation factors act as low/high pass filters. That is to say measured cospectra at frequencies where the transfer function magnitude is unity experience no attenuation wheareas complete attenuation is the case where the transfer function is zero. Using (2.28) and (2.29) we define the flux attenuation factor following Moore (1986)

$$a = \frac{\left(\overline{\xi' w'}\right)_a}{\overline{\xi' w'}} = \frac{\int_0^\infty H_{\xi w}(f) Co_{\xi w}(f)\, df}{\int_0^\infty Co_{\xi w}(f)\, df}\,, \tag{2.30}$$

such that a is the inverse of the correction factor for flux attenuation: $CF_a = a^{-1}$. From which it follows that the measured (attenuated) flux can be corrected for cospectral attenuation through the seemingly simple relation

$$\overline{\xi' w'} = \frac{\left(\overline{\xi' w'}\right)_a}{a}\,. \tag{2.31}$$

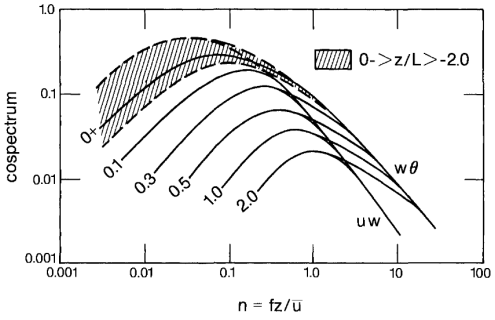
Here the crux of the problem lies in the fact that the exact form of $H_{\xi w}(f)$ and $Co_{\xi w}(f)$ are not generally known. Of course if $Co_{\xi w}(f)$ were known *a priori* the entire exercise of eddy covariance measurements would be redundant. So the question remains: how do we estimate the cospectral attenuation so as to attempt to recover the true flux?

There are in fact many related approaches to this problem each with varying degrees of complexity. According to Massman (2000) the methods can be divided into two schools: experimental and theoretical. For an extensive review of different cospectral attenuation correction methods we refer the reader to Clement (2005). Our focus will be on the theoretical methods which all employ transfer functions in some form or another. Now, all the theoretical methods assume some form of spectral similarity; in fact so do most of the experimental methods. Spectral similarity theory, as presented in Kaimal et al. (1972) and Wyngaard and Coté (1972), assumes that properly nondimensionalized spectra collapse into universal functions of non-dimensional frequency[18] $n = zf/\overline{u}$ and the M-O stability parameter $\zeta = z/L_*$. In the case of cospectra the appropriate choice of nondimensionalization factor is the associated flux (integrated cospectrum) normalized by natural frequency. As such, according to spectral similarity we have that the properly nondimensionalized cospectrum satisfies the universal functional relationship

$$\frac{f Co_{\xi w}(f)}{\overline{\xi' w'}} = \psi_{\xi w}\left(\zeta, n\right) \tag{2.32}$$

where $\psi_{\xi w}$ is the universal functional form of the nondimensional ξw cospectrum.

Figure 2.13: *The 'Kaimal model' nondimensional cospectra of the vertical fluxes of heat, $w\theta$ ($T_s w$ in our notation), and momentum, uw, as a function of nondimensional frequency, n, for various stabilities ζ (numbers to the left of each line). Both axes are logarithmic. Adopted from Kaimal and Finnigan (1994).*



Now, to introduce the most popular cospectral model as used in Moore (1986), Horst (1997) and Massman (2000); namely the flat-terrain Kaimal model cospectra[19](Kaimal et al., 1972). Here (for brevity) we only consider one numerical example namely $\xi = T_s$ in the case of unstable stratification ($\zeta < 0$) where a good approximation for the universal form proposed by Kaimal et al. (1972) given in Kaimal and Finnigan (1994) is

$$\frac{f Co_{T_s w}(f)}{\overline{T_s' w'}} = \begin{cases} \frac{11n}{(1+13.3n)^{7/4}} & \text{for } n \le 1.0 \\ \frac{4n}{(1+3.8n)^{7/3}} & \text{for } n \ge 1.0 \end{cases}. \tag{2.33}$$

[17] As defined in e.g. Wallace and Hobbs (2006)

[18] z is measurement height and $\overline{u}$ average horizontal wind.

[19] Based on the now famous Kansas 1968 experiment detailed in Haugen et al. (1971).

We note from (2.33) and Figure 2.13, as shown on dimensional grounds in Wyngaard and Coté (1972), that cospectra for the vertical flux of heat (as well as momentum) have a $-7/3$ power law dependence on frequency in the inertial subrange as opposed to the $-5/3$ dependence found in the intertial subrange of temperature and velocity spectra. Recall that this is the range $n_d >> n >> n_x$ where n_x is the nondimensional frequency at which the cospectrum attains its peak, i.e. the frequency of the main flux containing eddies, and n_d is the nondimensional frequency of the dissipative eddies near or just above the nondimensional sampling frequency of the instruments' (Foken, 2008b). Moreover, (2.33) implies that the nondimensional heat flux cospectrum is independent of the magnitude of ζ in the unstable case. This can be seen in Figure 2.13 where the nondimensional cospectra in unstable stratification (hatched area) all "*crowd into a narrow band that straddles the neutral cospectrum*" (Kaimal and Finnigan, 1994); these are all randomly yet only marginally scattered about the $0+$ (Horst, 1997) line which is the neutral limit approached from the stable side ($\zeta \gtrsim 0$). Hence, for convenience, Kaimal and Finnigan (1994) parametrize all unstable cospectra according to this $0+$ line; which corresponds to (2.33) for the nondimensional vertical heat flux cospectrum. Conversely, in stable stratification there is a systematic dependence on the magnitude of ζ with the cospectral peaks shifting towards higher frequencies with increasingly stable stratification in accordance with the strengthening of the bouyant destruction of turbulence . As an aside in Kaimal and Finnigan (1994) it is also pointed out that heat flux cospectra peaking at higher frequencies than momentum flux cospectra at the same stability highlights the need for faster responding instruments when seeking heat fluxes (and scalar fluxes in general) as opposed to momentum fluxes alone.

Having familiarized ourselves with spectral similarity theory and a cospectral model we can proceed with the implementation of the theoretical method. Rearranging (2.32) to obtain an expression for the true cospectrum $Co_{\xi w}(f)$ and inserting into (2.30), using that the true flux is independent of frequency, yields

$$\begin{aligned} a &= \frac{\int_0^\infty H(f)\psi_{\xi w}(\zeta, n) f^{-1}\, df}{\int_0^\infty \psi_{\xi w}(\zeta, n) f^{-1}\, df} \\ &= \int_0^\infty H(f)\psi_{\xi w}(\zeta, n) f^{-1}\, df\,. \end{aligned} \tag{2.34}$$

where we have used that by definition (cf. (2.28) and (2.32)) $\int_0^\infty \psi_{\xi w} f^{-1}\, df = \frac{1}{\overline{\xi' w'}} \int_0^\infty Co_{\xi w}\, df = 1$. With the integral method of Moore (1986) (2.34) would be solved numerically using an estimated form of the system transfer function and a cospectral model such as (2.33). The integration method becomes a relatively expensive computation overall given that the frequency range of the model cospectra span multiple, typicaly ≥ 4, decades (powers of ten) and that within a block each flux must be corrected individually as both the model cospectrum and system transfer function will vary depending on the type of flux considered. Contrarily, the analytical approach of Horst (1997) and Massman (2000) has the advantage of being cheap computationally with negligible difference in the corrections when compared to the more extensive method of Moore (1986) in the case of an open path system (Lee et al., 2006).

The analyical approach is based on a simplification of the Kaimal cospectral model and the system transfer function that allows for an analytical solution of (2.34). Horst (1997) used the fact that $n/n_x = f/f_x$ where f_x and n_x are the natural and nondimensional frequencies of the cospectral peak, henceforth peak frequencies, to propose the following simple approximation of the cospectral model

$$\psi_{\xi w}(\zeta, f, \overline{u}, z) = \frac{f Co_{\xi w}(f)}{\overline{\xi' w'}} = \frac{2}{\pi} \frac{f/f_x}{1 + (f/f_x)^2}\,, \tag{2.35}$$

The natural peak frequencies are recovered from the relation $f_x = n_x \overline{u}/z$, where the nondimensional peak frequency dependence on stability is parametrized. For the nondimensional vertical momentum flux[20] cospectra the nondimensional frequency peaks, n_{x,u_*^2}, are parametrized as in Massman (2000)

$$n_{x,u_*^2} = \begin{cases} 0.079 & \text{for } \zeta \leq 0 \\ 0.079\,(1 + 7.9\zeta)^{3/4} & \text{for } \zeta \geq 0 \end{cases}\,. \tag{2.36}$$

The nondimensional vertical scalar flux cospectra are all assumed identical to the nondimensional heat flux cospectrum (Moore, 1986) . Accordingly these non dimensional frequency peaks, $n_{x,\overline{s'w'}}$, are all parametrized following Horst (1997)

$$n_{x,\overline{s'w'}} = \begin{cases} 0.085 & \text{for } \zeta \leq 0 \\ 2.0 - 1.915/\,(1 + 0.5\zeta) & \text{for } \zeta \geq 0 \end{cases}\,. \tag{2.37}$$

Subsequently the peak frequency relations (2.36) and (2.37) can be used in conjunction with (2.35) to model all relevant cospectra across the physical range of stabilities. An analytical solution to (2.34) will be possible, but first the form of the system transfer function must be determined.

For the form of the system transfer function in the original analytical approach of Horst (1997) only the low pass (high frequency) transfer functions were considered. In this approach it is assumed that any low pass

[20] We will use the subscript $u_*^2 = \sqrt{\overline{u'w'}^2 + \overline{v'w'}^2}$ for kinematic momentum flux as opposed to τ/ρ to avoid confusion with the remaining time constant notation.

transfer function, $H_{HF,j}$, can be reasonably well approximated by $\widetilde{H}_{HF,j}$ which takes the form

$$\widetilde{H}_{HF,j}(f) = \frac{1}{1 + (2\pi f \tau_j)^2}\,, \tag{2.38}$$

where τ_j is the first order time constant of the given transfer function which is set such that the half power point, $f_{1/2}$, of the approximated transfer function matches that of the true transfer function (Massman, 2000). That is $f_{1/2}$ is found from the frequency satisfying $H_{HF,j}(f_{1/2}) = 0.5$ and in requiring $\widetilde{H}_{HF,j}(f_{1/2}) = 0.5$, we consequently set $\tau_j = 1/\left(2\pi f_{1/2}\right)$. Massman (2000) approximated the product of all low pass transfer functions yet further by baking the effect of all the approximated low pass transfer functions into a single combined transfer function $\widetilde{H}_{HF}(f)$ of the same form as (2.38). This was done by making use of a combined equivalent first order time constant, τ_e, for which the most adequate form[21] was found to be

$$\tau_e = \sqrt{\sum_{j=1}^{J} \tau_j^2}\,. \tag{2.39}$$

So in summary of the above the product of all low pass transfer functions is approximated as

$$H_{HF}(f) \simeq \Pi_{j=1}^{J} \left[\frac{1}{1 + (2\pi f \tau_j)^2}\right] \simeq \frac{1}{1 + (2\pi f \tau_e)^2}\,, \tag{2.40}$$

where the final approximation corresponds to the $\widetilde{H}_{HF}(f)$ term with τ_e given by (2.39).

Appropriate forms for the first order time constants τ_j relevant to each flux, dependent on the instrumentation, are outlined in Table 1 of Massman (2000) based on the results of references therein. In our case separation is already dealt with so we only have to contend with the effects of sensor path length averaging. Attenuation due to sensor path length averaging is associated with the fact that the sampled variables are not exact point values, but averages over a finite path (Burba, 2013), which acts as a low pass filter such that the contribution of very small eddies to the cospectrum is damped. Kaimal et al. (1968) provided a rough rule of thumb for the onset of significant cospectral attenuation due to averaging over a path length d as the streamwise wavenumber κ_1, related to f through Taylors' hyptohesis $\kappa_1 = 2\pi f/\overline{u}$, where $\kappa_1 d = 1$ with increasing attenuation for larger streamwise wavenumbers (Kaimal and Finnigan, 1994). Inserting a weak wind value of $\overline{u} = 2\,\mathrm{ms}^{-1}$ and a path length of $d = 0.1\,\mathrm{m}$ the onset of attenuation occurs at $f \simeq \overline{u}/(2\pi d) = 3.2\,\mathrm{Hz}$. This is well below the Nyquist frequency of 10 Hz so we can expect a small yet significant attenuation of the measured cospectrum (and hence the measured flux) as a result of path length averaging.

Note that path length averaging affects both the instruments (CSAT3 and LI-7500) employed in this study; both output signals are based on the behavior of waves (acoustic and infrared) transmitted over finite paths. As path length averaging is the only high frequency cospectral attenuation factor we have to concern ourselves with then for each flux only two low pass first order time constants, one for each path averaged component of the flux, are required to form the equivalent first order time constant. For the sake of transparency we will outline the time constants as employed in the module. These are based on those given in Massman (2000), Clement (2005) as well as our own values based on the results of van Dijk (2002) adapted to the instrumentation used in the Bayelva EC system. In the case of the kinematic heat flux, $\overline{T_s' w'}$, we used the discrete values of the numerically computed scalar flux cospectral transfer function for the CSAT3[22] given in Table 1 of van Dijk (2002) to approximate the exact half power point and arrive at an equivalent first order time constant given by

$$\tau_{e,\overline{T_s' w'}} = \frac{l_V}{6.9\overline{u}}\,,$$

where $l_V = 0.1\,\mathrm{m}$ is the vertical acoustic path length of the CSAT3 (CSAT3, 2014). To clarify; we arrived at $\kappa_{1/2} = 6.9/l_V$, the half-power point wavenumber, through linear interpolation of the cospectral transfer function of van Dijk (2002). The half power point wavenumber was then related to the half-power point frequency through the relation $f_{1/2} = \kappa_{1/2}\overline{u}/(2\pi)$ from which the first order time constant is given by $\tau_e = 1/(2\pi f_{1/2})$, so we have the simple relation $\tau_e = 1/(\kappa_{1/2}\overline{u})$. This was somewhat of a special case since we had an exact (numerical) form of the cospectral transfer function related to path length averaging of the same instrument employed at Bayelva, which explains why the equivalent first order time constant is given by the above equation as opposed to a combination of two first order time constants. Now in the case of kinematic water vapor flux, $\overline{\rho_v' w'}$, the equivalent first order time constant is given by

$$\tau_{e,\overline{\rho_v' w'}} = \sqrt{\left[\frac{l_L}{4.0\overline{u}}\right]^2 + \left[\frac{l_V}{6.9\overline{u}}\right]^2}\,,$$

where the term in the first brackets is the time constant for the LI-7500 scalar path length averaging of ρ_v with $l_L = 0.12\,\mathrm{m}$ being the path length of the LI-7500 (LI-7500, 2001). The term in the second brackets is the time constant for the CSAT3 path length averaging for scalar flux as previously used for the heat flux. For the kinematic momentum flux, $u_*^2 = \sqrt{\overline{u'w'}^2 + \overline{v'w'}^2}$, the equivalent first order time constant is given by

$$\tau_{e,u_*^2} = \sqrt{\left[\frac{l_H}{2.8\overline{u}}\right]^2 + \left[\frac{l_V}{5.7\overline{u}}\right]^2}\,,$$

[21] That is for $\widetilde{H}_{HF}(f)$ to best approximate $\Pi_{j=1}^{J}\,\widetilde{H}_{HF,j}(f)$

[22] Recall that there are three acoustic paths, each an angle of 30° from the vertical axis (CSAT3, 2014).

where the term in the first bracket is the time constant for the CSAT3 path length averaging of u for momentum flux with $l_{SH} = 0.058\,\text{m}$ being the horizontal path length of the CSAT3 (see CSAT3 (2014)) and the term in the second bracket is the time constant for the CSAT3 path length averaging of w for momentum flux. Note that the time constants for the CSAT3 sonic anemometer are based on on the work of Kaimal et al. (1968) and Kristensen and Fitzjarrald (1984) in which scalar *and* momentum flux cospectral transfer functions for orthogonal sonic anemometers with one axis aligned with the vertical were derived. van Dijk (2002) extended this work to nonorthogonal sonic anemometers, including the CSAT3, with no axes aligned with the vertical by deriving the associated cospectral transfer function for scalar fluxes and showed that the effects of path length averaging on high frequency cospectral attenuation was slightly more severe for these newer non-orthogonal type sonic anemometers. We adapted this transfer function for the nonorthogonal sonic path length averaging transfer function to the analytical method in the case of both the heat flux and the water vapor flux. For the path length averaging related to momentum flux we had to base the equivalent time constant based on the transfer functions of orthogonal sonic anemometers; so the approximation will not be as accurate. Nonetheless, we do not expect this discrepancy to be significant for the flux correction especially considering that, as previously discussed, cospectra follow $f^{-7/3}$ in the inertial subrange and path length averaging usually only becomes evident well into this range. As such the portion of the true flux contained in these frequencies is small relative to the peak of the cospectrum.

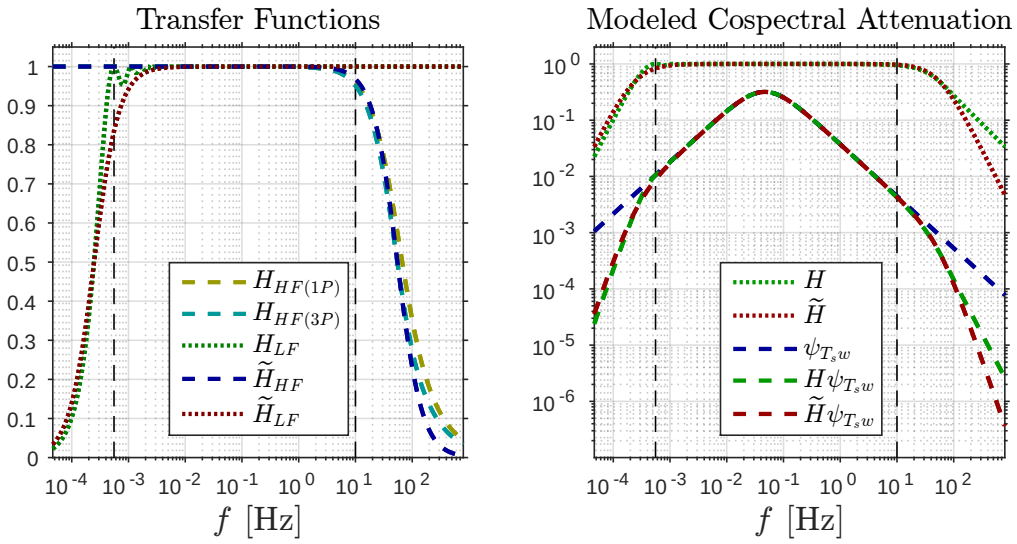


Figure 2.14: *Visualization of the natural frequency (f in Hz) dependence of the cospectral attenuation affecting sonic heat flux $\overline{T_s w'}$ measured with a CSAT3 sonic anemometer when $\overline{u} = 5\,[ms^{-1}]$ and $z = 2.75\,m$ in the unstable-neutral case ($\zeta \leq 0$). Left panel: Forms of the high and low pass transfer functions considered. The low pass transfer functions for path length averaging of scalar flux for nonorthogonal anemometers with three contributing acoustic paths at 30° from the vertical (such as the CSAT3), $H_{HF(3P)}$, (turquoise dashed line) and one contributing acoustic path (included as a reference), $H_{HF(1P)}$, (yellow dashed line) are based on the work of van Dijk (2002) The high pass transfer function associated with block averaging, H_{HF}, (green dotted line) after Kaimal* et al. *(1989). Approximated transfer functions using the half-power point method of Horst (1997) and Massman (2000) for high frequency attenuation, $\widetilde{H}_{HF}$, (blue dashed line) which approximates $H_{HF(3P)}$ and for low frequency attenuation, $\widetilde{H}_{LF}$, (red dotted line) which approximates H_{LF}. Right Panel: System transfer function, $H = H_{HF(3P)} H_{LF}$,(green dotted line) the approximate system transfer function, $\widetilde{H} = \widetilde{H}_{HF}\widetilde{H}_{LF}$, (red dotted line), the unstable-neutral model nondimensional cospectrum, $\psi_{T_s w}$, (blue dashed line) as given by* (2.35)*, the attenuated model nondimensional cospectrum $H\psi_{T_s w}$ (green dashed line) and the approximate attenuated model nondimensional cospectrum $\widetilde{H}\psi_{T_s w}$ (red dashed line). Finally the Nyquist frequency of the system, $f_{ny} = 10\,Hz$, as well as the inverse of the 30 minute block averaging period $\tau_A^{-1} = 5.56 \times 10^{-4}$ [Hz] are included as dashed black lines in both panels for orientation.*

The form of the combined high pass transfer function also needs to be determined. In our case there is only one such high pass transfer function as only one high pass filter is employed in the flux sampling namely that associated with the removal of the block average. In this case the precise form of the associated high pass filter, H_{LF}, in spectral space, as shown in the Left Panel of Figure 2.14, is given by (Kaimal et al., 1989)

$$H_{LF}(f) = 1 - \left[\frac{\sin(\pi f \tau_A)}{\pi f \tau_A}\right]^2$$

where τ_A is the averaging period. Using a similar to approach to that which Horst (1997) used for low pass filters Massman (2000) extended the analytical approach

by including the effects of high pass filters whos form was approximated through the relation

$$\widetilde{H}_{LF,j}(f) = 1 - \frac{1}{1 + (2\pi f \tau_j)^2} \,,$$

where τ_j is the first order time constant associated with high pass filter j found in the same way as the low pass filter time constants. For the block averaging high pass filter the corresponding first order time constant is $\tau_b = \tau_A/2.8$ (Massman, 2000). As a control it is readily verified that $f_{1/2} = 1/(2\pi\tau_b)$ satisfies $H_{\rm b}(f_{1/2}) = 0.50$ rounded to the second significant figure. So with block averaging being the only high pass filter employed we approximate the high pass filter function through (also shown in Figure 2.14)

$$\widetilde{H}_{LF} = 1 - \frac{1}{1 + (2\pi f \tau_b)^2} \,. \qquad (2.41)$$

Examples of the spectral form of both the (near) exact and approximated high and low pass transfer functions considered in the module in the case of heat flux are displayed in Figure 2.14 as a reference. Note that the high pass transfer functions (both exact and approximate) are the same for all the fluxes considered, whereas the low pass transfer functions depend on the instrumentation employed to calculate the corresponding fluxes. Furthermore, in Figure 2.14 for all three displayed high frequency tansfer functions $\overline{u} = 5\,\mathrm{ms}^{-1}$ is used in the conversion from κ to f (Taylors Hypothesis). The position and magnitude of these three example high pass filters relative to each other is independent of $\overline{u}$, but the position of the half power point for all three functions will shift towards higher frequencies with increasing $\overline{u}$ and vice versa for decreasing $\overline{u}$ since $f \propto \kappa\overline{u}$.

Now we finally have the approximate form of the system transfer function given by the product of (2.41) and (2.40) where as discussed the equivalent first order time constant depends on the flux considered. We also have a model for true cospectrum (2.35) as a function of stability through the parametrization of the peak frequency f_x for each flux. Inserting these approximations into (2.34) yields

$$a \simeq \frac{2}{\pi}\int_0^\infty \left[1 - \frac{1}{1 + (2\pi f\tau_b)^2}\right]\left[\frac{1}{1 + (2\pi f\tau_e)^2}\right]\left[\frac{1/f_x}{1 + (f/f_x)^2}\right] df \,. \qquad (2.42)$$

The solution to (2.42) is found and discussed in depth in both (Massman, 2000) and (Massman, 2001). Nonetheless, the form of our solution is not exactly the same as that of Massman. We re-emphasize that we have extended the scalar flux cospectral path-length averaging transfer functions taking into account the 3D nature of the CSAT3 based on the work of van Dijk (2002). As such the form of the first order time constants for scalar flux employed in our module are different and more accurate to that of Table 1 in Massman (2000), while the momentum flux first order time constants are the same. To the point in our case, where the only high pass filter used is associated with the block average, the relevant general solution is given by equation (6) in Massman (2001) namely

$$a = \left[\frac{b^\alpha}{b^\alpha + 1}\right]\left[\frac{b^\alpha}{b^\alpha + p^\alpha}\right]\left[\frac{1}{p^\alpha + 1}\right] \,. \qquad (2.43)$$

where $p = 2\pi f_x \tau_e$ and $b = 2\pi f_x \tau_b$. Here the exponent α accounts for the shape of the cospectra (Massman, 2000): for the stable cospectra ($\zeta > 0$) where the peak is relatively sharp and narrow $\alpha = 1$, whereas in the neutral and unstable range ($\zeta \leq 0$) where the cospectra are broader $\alpha = 0.925$ (cf. Figure 2.13). This exponent was orginally employed by Horst (1997) to improve the simplified cospectral model(2.35) after which it was revised in Massman (2000) upon the inclusion of low frequency cospectral attenuation. Finally we may now use α along with the relevant form of the equivalent first order time constant, τ_e, for each flux in conjunction with the cospectral peak frequency ((2.36) for momentum and (2.37) for scalar fluxes) for each stability to arrive at a close approximation for the cospectral attenuation factor on a block by block basis. After this factor is computed the respective fluxes are readily corrected. Note that the relevant calculations each involve only one simple computation (with some conditionals) as opposed to integrals over multiple frequency decades. So the analytical approach is clearly invaluable for long term analysis. This especially considering many of the transfer functions are even more accurate than in the original approach of Moore (1986) (Massman, 2000). We finish off this somewhat involved yet vital correction with respect to proper estimation of fluxes by including a visualization of the typical magnitudes of the associated flux correction factors across various stabilities and block average wind speeds. These are dispalyed in Figure 2.15 from which it is clear that the correction for cospectral attenuation not only always results in an increase in flux but is also approaching the order of magnitude of typical surface energy balance residuals outlined in Foken (2008a). The correction is largest for the combination of slow wind speeds and small heights z (height is constant in Figure 2.15) for all ζ as in such a case the peak natural frequency shifts towards higher values exarcebating the effects of path length averaging. As with the separation correction the corresponding correction factor (here for each flux) is passed on a block by block basis to the module output to track the magnitude of each correction.

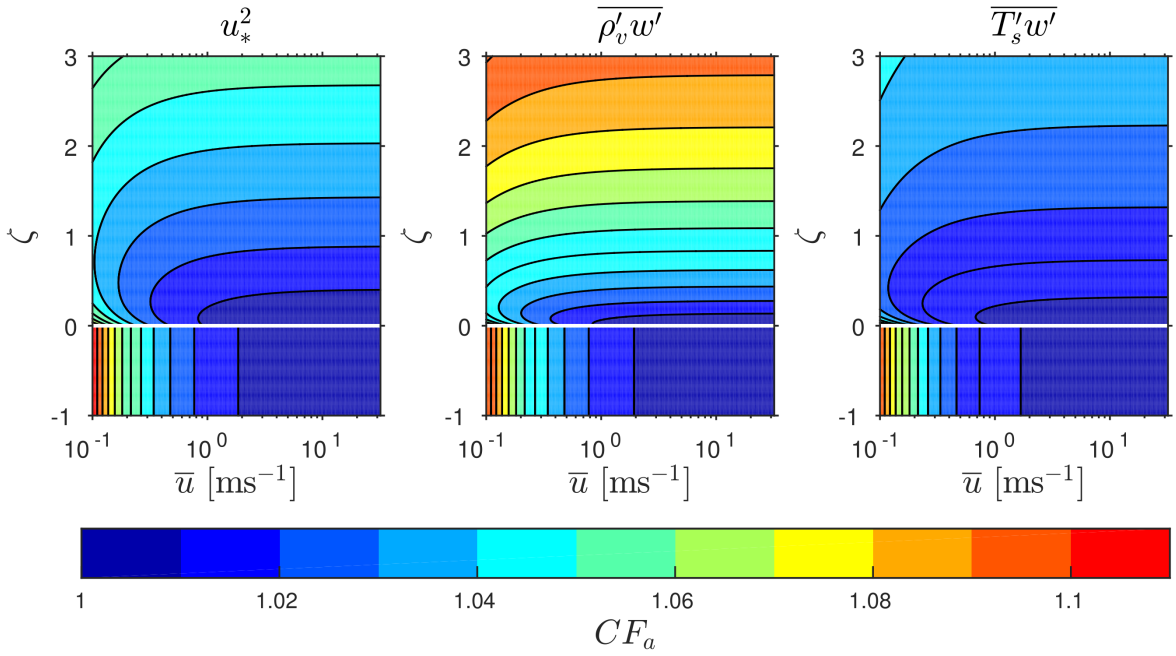


Figure 2.15: *Cospectral attenuation correction factors, CF_a, as functions of the M-O stability parameter and block plane mean wind $\overline{u}$ at a height $z = 2.75m$ for the vertical turbulent fluxes of momentum u_*^2 (Left Panel), water vapor $\overline{\rho_v' w'}$ (Middle Panel) and sonic heat $\overline{T_s' w'}$ (Right Panel). Correction factors span from near no correction $CF_a = 1$ (blues) to nearly an 11% correction $CF_a = 1.11$ (reds) with the same color code applying to all three panels. The horizontal white lines mark $\zeta = 0$ at which there is a transition in the peak frequency parametrization.*

2.5.3 SND

The so-called (Aubinet et al., 2012) 'SND' correction originally proposed by Schotanus, Nieuwstadt and De Bruin (hence the abbreviation) in Schotanus et al. (1983) converts the sonic heat flux, $\overline{T_s' w'}$, measured by a sonic anemometer into the sensible heat flux, $\overline{T' w'}$. Of course the sensible heat flux is the more desireable quantity in that it is a component of the surface energy balance (Foken, 2008a) justifying the application of this correction.

Now the original SND correction was revised by Liu et al. (2001) to account for modern sonic aneometer geometry, such as that of the CSAT3, where sonic temperature is obtained from the average value over three off-vertical acoustic paths. This in contrast to the original correction of Schotanus et al. (1983) which was developed for the earlier types of sonics where the sonic temperature was obtained from a single vertical acoustic path. The revised correction is as follows (Liu et al., 2001)

$$\overline{T' w'} = \overline{T_s' w'} - 0.51 \overline{q' w'}\, \overline{T} + \frac{2\overline{T}}{c^2} \left(\overline{u' w'}\, \overline{u} A + \overline{v' w'}\, \overline{v} B \right) .$$

where $q = \rho_v / \rho$ is the specific humidity (with ρ the air density), c is the speed of sound. A and B are combined (instrument dependent) constant factors (Liu et al., 2001) that appropriately scale the effects of crosswind on sonic temperature according to the sonic geometry (particulary the angle between the acoustic paths and the vertical axis). For the CSAT3 $A = B = 7/8$, however, as pointed out in both Mauder and Foken (2011) and CSAT3 (2014) the crosswind correction is already applied 'online' internally in the CSAT3 firmware. As such, correcting for crosswind effects offline in processing would result in an unwarranted double correction and significant error in the computed fluxes (CSAT3, 2014). So in our case the remaining SND correction that must be applied is

$$\overline{T' w'} = \overline{T_s' w'} - 0.51 \overline{q' w'}\, \overline{T} . \tag{2.44}$$

We adapted (2.44) further into a form that was more tractable both to the measurements at our disposal and the required iterations of the flux corrections. The form we arrived at is

$$\overline{T' w'} = \overline{T_s' w'} \left(1 - \frac{0.51 R_d \overline{\rho_v}\, \overline{T}}{P} \right) - \frac{0.51 R_d \overline{T_s}\, \overline{T}}{P} \overline{\rho_v' w'} . \tag{2.45}$$

A complete derivation of (2.45) (as well as (2.44)) is provided in Appendix A.3. At first glance the imple-

mentation of this adapted correction could seem more demanding than the original form of Schotanus et al. (1983) due to the expansion of the specific humidity flux. Of course this is far from the case in fact all the terms on the right hand side of (2.45) are already stored in the module at this point with all the relevant previous flux corrections applied. After the correction is performed for a given block then, as with all the flux corrections, the corresponding block flux correction factor, CF_{SND}, is calculated through the relation

$$CF_{SND} = \frac{\overline{T'w'}}{\overline{T'_s w'}}\,, \tag{2.46}$$

after which it is passed on a block by block basis to the module output to track the magnitude of this correction.

Now that we have demonstrated the SND correction one may legitimately wonder why we don't just compute $\overline{T'w'}$ directly based on the diagnosed instantaneous absolute temperature and vertical velocity values. The main reason we instead go the usual way (e.g. Mauder and Foken (2011)) of the SND correction is that when accounting for cospectral attenuation in Section 2.5.2 it is the sonic heat flux that is corrected. As such we do not know the exact cospectral attenuation factor for the sensible heat flux $\overline{T'w'}$. We correct the sonic heat flux for cospectral attenuation since it is more accurate to consider approximate high frequency transfer functions from a single instrument[23], especially when (as is the case) these are based on the detailed work of van Dijk (2002) for the very same instrument. Moreover, the scalar model cospectrum of Kaimal and Finnigan (1994) is based on the sonic heat flux cospectrum so we avoid the additional assumption[24] that the nondimensional sensible heat flux cospectrum is equivalent to the nondimensional sonic heat flux cospectrum. So if we want the most reliable analytical cospectral attenuation correction for heat flux it is the sonic heat flux that must be considered. Hence as the next step after Section 2.5.2) in the flux corrections it is necessary, as we do, to apply the SND correction to this flux and arrive at the sensible heat flux, which will consequently also have been corrected for cospectral attenuation.

2.5.4 WPL

A fundamental concern and long lasting problem with EC measurements is the fact that the block averaged vertical velocity, $\overline{w}$, is typically so small (on the order $10^{-4}\,[\mathrm{ms}^{-1}]$) that it is *immeasurable* (Webb et al. (1980), Lee and Massman (2011)). The CSAT3 that we employ, for example, can only resolve vertical velocities on the order $10^{-3}\,\mathrm{ms}^{-1}$ (CSAT3, 2014). To see exactly why this is a problem consider the total vertical flux, F_ξ, of a generic scalar quantity ξ; we recall from Section 1.2.1 that in the surface layer where diffusion (or viscosity depending on the choice of ξ) is negligible the block averaged vertical flux at the measurement height is given by the average of the instantaneous flux

$$F_\xi = \overline{\xi w} = \overline{\xi}\,\overline{w} + \overline{\xi' w'}\,. \tag{2.47}$$

From the above, the total flux is only equal to the eddy flux if $\overline{w} = 0$. Early on this problem was 'solved' by forcing $\overline{w}$ to zero in each block via rotation procedures such as that described in McMillen (1988). This is a prime example of why the planar fit of Wilczak et al. (2001) is to be prefered over double or triple rotation procedures that force $\overline{w}$ to zero; physically there is no reason to expect the 'vertical' velocity to be zero on average in a given block. In fact when the vertical velocity is normal to the surface patch of interest the block averaged vertical velocity is typically small yet rarely zero (Mauder et al., 2013). Thereby, the product of means term in (2.47) can be just as large (and of opposite sign) to the eddy flux (Fuehrer and Friehe, 2002) in that for scalars in particular $\overline{\xi} >> |\xi'|$. The problem with the mean vertical velocity was long recognized in the EC community and from what we can gather appropriate solutions started appearing in the late 1970s (see reviews in Fuehrer and Friehe (2002) and Lee and Massman (2011)). Of these the widely accepted solution (e.g. Mauder and Foken (2004),Van Dijk et al. (2004) and Burba (2013)) was provided by Webb, Pearman and Leuning in Webb et al. (1980). This 'WPL' correction, to be discussed, is considered to be the cornerstone or 'principle underpinning' of modern EC theory (Lee and Massman, 2011) . Ita has propelled the EC method beyond research into the structure of surface layer turbulence (such as Kaimal et al. (1972)) towards what is now perhaps its primary concern, namely diagnosing long-term exchange of energy and mass, of trace gases in particular, across the surface-atmosphere interface (e.g. Baldocchi et al. (2001)).

To solve the problem, Webb et al. (1980) introduced a governing constraint without which the averaged vertical velocity, $\overline{w}$, and thereby vertical fluxes in general are indeterminate. To arrive at this constraint and its implications we begin following Fuehrer and Friehe (2002) by noting that the block averaged equation for the conservation of mass for dry air in the absence of sources

[23] In the case of the buoyancy flux both components of the flux are meaured by the sonic anemometer; whereas for the sensible heat flux we need to also make use of the instanteneous absolute humidity measurements from the gas analyzer as well as (slow) pressure to diagnose the instantaneous absolute temperature.

[24] This assumption is unavoidable for the water vapor flux for which a specific model cospectrum is not available.

or sinks (as emphasized in Leuning (2007)) reads

$$\frac{\partial \overline{\rho_d}}{\partial t} = -\frac{\partial}{\partial x_j}\left(\overline{\rho_d w} - \nu_{\rho_d}\frac{\partial \overline{\rho_d}}{\partial x_j}\right), \tag{2.48}$$

where ν_d is the diffusivity of dry air (units s^{-1}), which we assume to be constant without loss of generality. In a statistically stationary and horizontally homogeneous flow, as required for the point EC method, then

$$\frac{d}{dz}\left(\overline{\rho_d w} - \nu_{\rho_d}\frac{d\overline{\rho_d}}{dz}\right) = 0\,, \tag{2.49}$$

where we have now included the diffusive flux in the the total flux so as to extend the definition to the thin micro-layer that is just below the surface layer and adjacent to the surface. Contrary to the surface layer, in the micro-layer the diffusive flux is the dominant term in the total flux (Stull, 1988). Now in words (2.49) states that, in the absence of sources or sinks, the *total* flux of dry air is constant with height in the surface layer as well as the micro layer. Note that this claim also holds both for the total mass flux and the water vapor (or any other constiuent) mass flux under the same conditions (Fuehrer and Friehe, 2002), which is why the surface layer is sometimes called the 'constant flux layer'. As of yet it may not be clear how this is helpful in determining $\overline{w}$, however, upon integration of (2.49) from the surface ($z = 0$) to the measurement height ($z = z_m$) abbreviating the total dry air mass flux in the brackets as F_{ρ_d} it follows that

$$F_{\rho_d}(z_m) = F_{\rho_d}(0)\,. \tag{2.50}$$

Once more this holds in general for the mass flux of any other constituent, including the total mass flux, provided the same conditions hold between $z = 0$ and $z = z_m$. Now the governing constraint of Webb et al. (1980) is that the total dry air mass flux is zero at the surface. More recently Leuning (2007) emphasized that properly stated the constraint of Webb et al. (1980) is the assumption of no sources or sinks of dry air below the measurement height z_m including the surface itself. As noted in Lee and Massman (2011) this constraint was arrived at through the process of deduction in that, outside of the small contribution of biogenic photosynthesis and respiration at the surface, we do not expect there to be any sources of dry air mass flux at the surface; at least if we neglect the occurence of dramatic events such as sinkholes. Now this constraint dictates that $F_{\rho_d}(0) = 0$ whereby from (2.50) $F_{\rho_d}(z_m) = 0$. As such, provided that z_m is in the surface layer where molecular diffusion is negligible then the constraint reduces to that given in Webb et al. (1980) namely

$$\overline{\rho_d w} = 0 \iff \overline{w} = -\frac{\overline{\rho_d' w'}}{\overline{\rho_d}}\,. \tag{2.51}$$

As demonstrated in Appendix A.4 through manipulation (2.51) the following expression for $\overline{w}$ can be recovered:

$$\overline{w} = (1+\mu\sigma)\left[\frac{\overline{T'w'}}{\overline{T}} - \frac{\overline{p'w'}}{\overline{p}}\right] + \mu\sigma\,\frac{\overline{\rho_v' w'}}{\overline{\rho_v}}\,, \tag{2.52}$$

where following the notation of Webb et al. (1980) $\mu = M_d/M_v = 1.61$ is the ratio of the molar mass of dry air to that of water vapor and $\sigma = \overline{\rho_v}/\overline{\rho_d}$. Thereby the total vertical flux of a generic scalar quantity ξ may be expressed as

$$\overline{\xi w} = \overline{\xi}\left[(1+\mu\sigma)\left[\frac{\overline{T'w'}}{\overline{T}} - \frac{\overline{p'w'}}{\overline{p}}\right] + \mu\sigma\,\frac{\overline{\rho_v' w'}}{\overline{\rho_v}}\right] + \overline{\xi' w'}\,. \tag{2.53}$$

By inspection of (2.53) the term involving the turbulent flux of pressure, $\overline{p'w'}$, is problematic in that p' is not measured at Bayelva (Section 2.4). This is the case for most EC systems and a variety of workarounds have been proposed. For example, in Van Dijk et al. (2004) Bernoulli's law, which is strictly not applicable to the viscuous flows in the atmospheric boundary layer, is used to approximate the fluctuations in pressure and in Zhang et al. (2011) a MOST parametrization of $\overline{p'w'}$ is proposed based on a year long EC campaign in China. Hoping to avoid such parametrizations we noted that upon application of the ideal gas law, in terms of virtual temperature, and Reynolds averaging then $\overline{p'w'}$ may be expressed as

$$\overline{w'p'} \simeq R_d\,\overline{T_v}\,\overline{\rho' w'} + R_d\,\overline{\rho}\,\overline{T_v' w'}\,. \tag{2.54}$$

If the flow is near incompressible, as is often the case in the boundary layer (see e.g. Mahrt (1986)), we expect not only a coincident eddy flux of warm (cold) air and light (heavy) air, but a near cancelation of the two terms on the right hand side of (2.54). With this carrying over to absolute temperature T then the pressure covariance term in (2.53) should be relatively small, albeit not generally neglible, with respect to remaining terms. As such we can approximate, for lack of a better option, (2.53) through

$$\overline{\xi w} = \overline{\xi}\left[(1+\mu\sigma)\,\frac{\overline{T'w'}}{\overline{T}} + \mu\sigma\,\frac{\overline{\rho_v' w'}}{\overline{\rho_v}}\right] + \overline{\xi' w'}\,. \tag{2.55}$$

Note that this is the expression given for $\overline{w}$ in Webb et al. (1980) that is consequently employed in most EC processing software such as TK2 (Mauder and Foken, 2004) as well as our module. It turns out that the WPL correction has implications mainly for the surface exchange of trace constitutents such as water vapor and CO_2; as shown at great length in Van Dijk et al. (2004),Sun et al. (1995) and Fuehrer and Friehe (2002) the mean vertical velocity does not affect the sensible heat flux. Therefore, in the context of the surface energy balance we need only consider the WPL correction for the vertical flux of water vapor which takes the form

$$\overline{\rho_v w} \simeq (1+\mu\sigma)\left[\overline{\rho_v' w'} + \frac{\overline{\rho_v}}{\overline{T}}\,\overline{T' w'}\right]. \tag{2.56}$$

This is the only WPL correction applied in our module as we do not presently investigate fluxes of CO_2

where the WPL correction is also important (Lee and Massman, 2011). As with the remaining flux corrections the corresponding correction factor, for the water vapor flux, computed via

$$CF_{WPL} = \frac{\overline{\rho_v w}}{\overline{\rho_v' w'}},$$

is stored for bookkeeping.

2.5.5 Iterations

First off we note from (2.26) that the sensor separation correction is independent of the other corrections. As such it should be applied first, and then only once, to colocate and sychronize the vertical velocity and water vapor fluctuations. Furthermore, by comparing the cospectral attenuation correction (2.31) (where a is found from (2.43)), the SND correction (2.45) and the WPL correction (2.56) these corrections are clearly codependent. That is to say, the WPL correction depends on both the SND and the cospectral attenuation correction; while the SND correction depends on the cospectral attenuation correction. Moreover the cospectral attenuation correction relies on itself; in that it corrects ζ, which is in turn used to estimate the true cospectra in the very same correction. As such, these corrections need to be applied iteratively as noted in Mauder and Foken (2004), Nordbo et al. (2012), Mauder et al. (2013) and Burba (2013); though exactly how to perform such a procedure was not made clear. Nonetheless, from the previously noted codependence we surmised that for each iteration the corrections should be applied in the order of: the attenuation correction, followed by the SND correction and finally the WPL correction. So in the iterative flux corrections we performed the corrections in the presented order for each iteration and the only value that was updated based on the previous iteration was the M-O stability parameter ζ to improve the accuracy of the cospectral attenuation correction. For each iteration the *raw* (uncorrected) fluxes were corrected in the presented order and only the result of the correction for the previous and current iteration were stored. The module jumped out of the iterations as soon as the fluxes converged; which we defined to be the case when the corrected fluxes varied by less than 0.1% from one iteration to the next.

2.6 Quality Control

In the previous sections we have shown how the raw data must be despiked, rotated, supplemented and corrected for us to be able to calculate representative turbulent statistics as well as dynamic fluxes. All these steps are crucial in quality assuring the output data, but a final step in quality assesment (QA) is needed before the data can be critically analyzed with respect to the underlying assumptions of EC theory. This final step is the quality control (QC) routine which checks the quality of the block averaged statistics. Our routine builds heavily on the tests of Foken and Wichura (1996) and Vickers and Mahrt (1997) which are still recongnized as the QC benchmarks or 'state of the art' in the field of EC according to the QA strategies outlined in Mauder et al. (2013).

The routine is based on a sequence of tests that assign quality flags to the block statistics according to a set of flagging rules. These culimnate in the combined block quality flag, described in Section 2.6.6, where each block in the time series is assigned an overall quality flag ranging from 0 to 2 with f $= 0$ high quality, f $= 1$ medium quality ('soft flag') and f $= 2$ poor quality ('hard flag') similar to what is done in TK2 (Mauder and Foken, 2004). Any block that is hard flagged as being of poor quality (f $= 2$) is discarded from subsequent analysis. This combined flag is assigned after the block undergoes tests for the: fraction of faulty data, degree of flow distortion, deviation from stationarity and so-called integral turbulence characteristiscs. For each of these four tests individual quality flags are assigned so as to help track the origins of quality issues in the data. In the following we describe in detail each of the block quality flags as well as the combined quality flag. As with the despiking routine we should be weary of not being overly strict in our QC as we do, after all, want to end up with as much useable data as possible. Before any flags are implemented in this routine the plane velocities (u,v) are rotated such that u points into the longitudinal (block averaged plane) wind direction and v the lateral wind direction using the routine described in section 2.3.3.

2.6.1 Faulty

The simplest of the individual flags we employ is the faulty block flag which we will denote f_F. Here we follow the suggestion in Mauder et al. (2013) and apply a hard flag, that is set $\mathrm{f}_F = 2$, to any averaging block where the number of spikes exceeds 10% of the block length. Recall that instantaneous measurements are set to our placeholder faulty data value of -1000 if these are flagged as `"NaN"` internally by the instrument, exceed the physical plausibility limits *or* are diagnosed as spikes in the MAD routine. The skewness-kurtosis test (Section 2.2.3) has already hard flagged entire blocks that do not satisfy the given thresholds proposed by Vickers and Mahrt (1997) and the f_{SK} flag is stored for bookkeeping. Thereby, if the block consists entirely of spikes it is of course faulty but we do *not* flag it with this faulty flag; this enables us to track the exact origin of quality issues in a given block. So for a given block m with $i \in 1(1)I$ entries the faulty flag is set according to the criterion

$$\mathrm{f}_F = \begin{cases} 2 & \text{if } 0.1 < \frac{N_F}{I} < 1 \\ 0 & \text{otherwise} \end{cases}, \tag{2.57}$$

where N_F is the number of faulty entries for any of the despiked instantaneous variables in block m. We reiterate that N_F is the same for all variables in a given block as a result of the spike spreading in the MAD routine (see section 2.2). As such it is only neccessary for the module to scan *one* of the despiked variables to test the $> 10\%$ criterion which greatly reduces the computational cost of this flag routine. Of course the threshold of 10% is arbitrary. As noted in Westermann (2010) and as will become clear such arbitrary thresholds are typical in EC QC. More often than not these thresholds are determined by the investigators that first proposed them and remain unchanged as a tradition as opposed to being based on underlying physics as one might hope. As a result the thresholds should, when necessary, take into consideration the specifics of the site and instruments employed. Here at least we agree subjectively with the judgement of Mauder et al. (2013) in that 10% is a reasonable compromise for a spike limit. If the tolerance were much higher the *effective* averaging period would often be too low to resolve all the turbulent scales when computing fluxes and conversely if it were much lower we would risk having to discard larger amounts of data.

2.6.2 Flow Distortion

Next in line is the flow distortion flag f_D where we hard flag, that is set $\mathrm{f}_D = 2$, blocks where the block average wind direction is in the 'closed sector' where the sonic head is upwind of the measurement path of the sonic anemometer. In passing it is worth mentioning explicitly that in the module when block averaged wind direction is computed (Section 2.3.1) an exception is raised when the vector averaged (horizontal) wind speed is *calm*; that is when $|\overline{\mathbf{u}}| < 0.2$. In concurence with this calm wind threshold outlined in WMO (2008) we do not consider wind direction to be well defined in calm conditions and accordingly set $\varphi = -1000$ in such cases. Note that in the special case of wind direction -1000 corresponds to 'undefined' as opposed to a spike or missing value. Moreover the value automatically falls outside the closed sector so instances with an undefined wind direction are *not* hard flagged. Having accounted for calm conditions implementing the distortion flag should just be a matter of excluding blocks where the average wind direction falls in the bin $15^\circ \leq \varphi \leq 55^\circ$ (the closed sector).

So at this point the distortion flag seems as simple to implement as the faulty block flag, but only if we accept the computed wind directions without further reflection. As previously alluded to in the planar fit discussion it does not always make sense to talk about a streamwise (mean) wind direction for a given block as the wind direction can be far from steady over a block averaging period. A useful metric to identify the steadiness of the wind direction is the so-called constancy ratio discussed in Mahrt (1999)[25]. The constancy ratio for a given block with $i \in 1(1)I$ entries is defined as

$$\mathrm{CR} = \frac{|\overline{\mathbf{u}}|}{\overline{|\mathbf{u}|}} = \frac{\sqrt{\left(\sum_{i=1}^{I} u_i\right)^2 + \left(\sum_{i=1}^{I} v_i\right)^2}}{\sum_{i=1}^{I} \sqrt{u_i^2 + v_i^2}},$$

or in words as the ratio of the vector averaged wind speed to the averaged wind speed. As put rather eloquently in Mahrt (1999); conceptually the ratio can be thought of as a measure of the 'flip-flop' of the wind vector in a given period. In the typical case of a nearly steady wind direction the value of the ratio is near its maximum value of unity. Conversely, when the wind is veering or backing excessively[26] in an averaging period the ratio tends to zero. Now that we are armed with this new metric we can identify and exclude cases where the mean wind direction alone would be problematic but where the wind direction is so unsteady as to not warrant discarding the block due to flow distortion. Of course this in turn means we need to set a threshold constancy ratio that is low enough for the wind direc-

[25] In which it is used in a different context: as an indicator of mesoscale motions such as drainge flows or gravity waves that can occur in stably stratified boundary layers

[26] In this discussion we refer to backing (veering) as a counterclockwise (clockwise) turning of the wind vector with time as opposed to with height.

tion to be classified as 'unsteady enough'. As opposed to the thresholds set for the third flag first proposed by Foken and Wichura (1996) we do not have the luxury of justifying our thresholds based on 'long-term experiences'. Instead we base the threshold CR value on the fact that the majority of the time, for our 30 minute averaging periods, CR > 0.9 *and* in an idealized unsteady backing scenario where the wind vector has constant magnitude but is rotated (at constant angular velocity) counter-clockwise from easterly to westerly the value is CR $= 0.64$. So it would make sense to put the threshold for unsteady wind near the $180°$ backing scenario as it is precisely these strongly backing/veering scenarios we wish to ignore when flagging for flow distortion. When we allow the magnitude of the wind vector to vary from one quadrant to the other, a slightly more realistic scenario, we find CR $= 0.75$ to be a reasonable limit for unsteady wind direction. This corresponds to a similar backing scenario only that the wind vector is halved in magnitude with a step-like function when the wind vector enters the second quadrant. Now that we have derived a threshold for the constancy ratio that corresponds to an 'unsteady enough' wind direction we may apply the flow distortion flag to each flag. Blocks are hard flagged if the wind direction falls within the flow distoriton range *and* the constancy ratio is *above* the unsteady limit. That is

$$\mathrm{f}_D = \begin{cases} 2 & \text{if } 15° \leq \varphi \leq 55° \text{ and CR} > 0.75 \\ 0 & \text{otherwise} \end{cases},$$

so in words a block is hard flagged if the wind is such that the sonic head is on average upwind of the sonic path *and* the wind direction is sufficiently steady in time. As a final note on this flag we mention that for consistency the same constancy ratio requirement is implemented in the planar fit where blocks are ignored (not weighted) if they are flagged for flow distortion.

2.6.3 Vertical Velocity

After having applied the planar fit of Wilczak et al. (2001) (Section 2.3.2) the residual block averaged vertical velocity term $\overline{w}$ should be small and typically near the instrument resolution of the CSAT3, i.e. $|\overline{w}| \sim 10^{-3}\,\mathrm{ms}^{-1}$. As discussed in Section 2.5.4 such a block averaged vertical velocity is usually so small that it is immeasurable and so we apply the correction of Webb et al. (1980) to correct for its influence on the total mass flux by considering the physics of a surface layer with zero dry air mass flux. Nonetheless, at times $\overline{w}$ turns out to be far from small and perfectly measurable.

Ironically, as with immeasurability of low $|\overline{w}|$, measurements of relatively high $|\overline{w}|$ poses a problem for the EC method. To demonstrate this we consider a near incompressible surface layer, a reasonable approximation discussed at great length in Mahrt (1986), such that the block averaged velocity field is approximately divergenceless. Following Lee (1998) we use no-through flow kinematic boundary condition at the surface, $w(z = 0) = 0$, and assume a linear vertical velocity profile such that

$$\frac{\partial \overline{w}}{\partial z} \simeq \frac{1}{z_m}\overline{w}\big|_{z=z_m} \iff w(z) \simeq \frac{z}{z_m}\overline{w}\big|_{z=z_m},$$

to arrive at the following expression for the vertically integrated divergence of the *horizontal* velocity field

$$\int_0^{z_m} \nabla_H \cdot \overline{\mathbf{u}}\, dz \simeq -\overline{w}\big|_{z=z_m} \tag{2.58}$$

where ∇_H is the horizontal gradient operator and $\overline{\mathbf{u}}$ is the block averaged horizontal velocity vector. From (2.58) it is clear that if the block averaged vertical velocity is relatively large then the wind field below the sensor is far from horiontally homogeneous. Anticipating that horizontal inhomogeneity of the wind field is symptomatic of horizontal advection of not only momentum, but also mass and energy it is prudent to be wary of blocks where the absolute block averaged vertical velocity is large even after application of the planar fit. Recall that for the point EC method to be able to diagnose surface exchange an underlying assumption is statistical (i.e. block averaged) horizontal homogeneity of both the wind and scalar fields. As such Foken and Wichura (1996) as well as Mauder et al. (2013) propose a flaging procedure if the residual absolute block averaged vertical wind, after the application of the planar fit, is abnormally high. Following the recommendations of Mauder et al. (2013) we assigned a block average vertical velocity flag, f_V, according to the following thresholds on absolute block averaged vertical velocity $|\overline{w}|$ (units ms^{-1})

$$\mathrm{f}_V = \begin{cases} 0 & \text{if } |\overline{w}| \leq 0.1 \\ 1 & \text{if } 0.1 < |\overline{w}| \leq 0.15 \\ 2 & \text{otherwise} \end{cases}. \tag{2.59}$$

It is worth emphasizing that this particular flag can be considered as a second iteration of the limits of physical plausibility (Section 2.2.1). This because upon the application of the planar fit within the atmospheric surface layer, blocks where $|\overline{w}| > 0.15\,\mathrm{ms}^{-1}$ become a highly unusual occurence. As such only extreme cases of vertical advection or subsidence will be hard flagged by this test. For the remaining cases we assume that the WPL correction accurately gauges the block averaged vertical velocity along with its contribution to the total flux and that any departure from the underlying assumption of horizontal homogeneity can be neglected.

2.6.4 Stationarity

The next flag we employ is a stationarity flag which is a measure of the stationarity (or lack thereof) of the respective (vertical) eddy covariances within a given block averaging period. We flag based on a relative non-stationarity factor, henceforth abreviated RNF, derived from the steady-state test proposed for EC QC by Foken and Wichura (1996). This factor considers the absolute value of the normalized difference between the block covariance and the mean of subsegment covariances. The idea is that if the block is statistically stationary then the block covariance should not differ substantially from the covariances of subsegments in the same block. Symbolically the RNF, denoted $\mathrm{RN}_{\xi w}$, for the eddy covariance between some variable ξ and the vertical velocity w for a given block with $I = J \times K$ entries where J is the number of subsegments with K entries is defined as follows

$$\mathrm{RN}_{\xi w} = \left| \frac{\overline{\xi' w'} - \overline{\overline{\xi' w'}_j}^{j}}{\overline{\xi' w'}} \right| \tag{2.60}$$

where $\overline{\xi' w'}$ is the covariance between ξ and w for the entire block while $\overline{\overline{\xi' w'}_j}^{j}$ is the mean over the covariances in J non-overlaping subsegments with equal length into which the block is divided. To clarify the possibly confusing notation and wording $\overline{\xi' w'}_j$ is the covariance of a given subsegment j in the block while $\overline{\overline{\xi' w'}_j}^{j}$ is the average of over all the J non-overlaping subsegments that together make up the block. Note from (2.60) the RNF is a positive definite quantity. Subsequently, we apply the stationarity flag to the eddy covariances of longitudinal wind, sonic temperature and absolute humidity (i.e. $\xi = u, T_s, \rho_v$) according to the following stationarity factor thresholds:

$$\mathrm{f}_{S,\xi} = \begin{cases} 2 & \text{if } \mathrm{RN}_{\xi w} > 1 \\ 1 & \text{if } 0.3 < \mathrm{RN}_{\xi w} \leq 1 \\ 0 & \text{if } \mathrm{RN}_{\xi w} \leq 0.3 \end{cases} . \tag{2.61}$$

As previously alluded to these thresholds are defined by Foken and Wichura (1996) and justified in the TK2 documentation (Mauder and Foken, 2004) as being the result of long term experiences in the field. We can not boast the same experience and thus trust the judgement of the former authors. In using the same thresholds as TK2 for this particular test a comparison between the results of our module and TK2 becomes more straightforwards. Moreover, we set $J = 6$ which with a block averaging period of $\tau_A = 30$ minutes means that the test is based on 5 minute subsegments as was done in the TK2 runs of Westermann et al. (2009).

Having duely noted that these thresholds are arbitrary we also mention that they, along with the choice of J and τ_A, can easily be tuned as desired. Nonetheless, applying the hard flag ($\mathrm{f}_{S,\xi} = 2$) for instances where $\mathrm{RN}_{\xi w} > 1$ is a reasonable choice of threshold for non-stationary blocks as in such instances the mean of subsegment covariances is either twice the block covariance or completely negligible relative to the block covariance. In either case the mean of subsegment covariances differs substantially from the block covariance which is indicative of a highly transient vertical flux of quantity ξ. Recall that such unwanted transience can be due to a number of factors such as trends introduced by diurnal variation, a passing front, varying cloud cover (Andreas et al., 2008), mesoscale variability such as a changing land-sea breeze pattern (for Kongsfjorden see Esau and Repina (2012)) or intermittent ('patchy') turbulence caused by sub-mesoscale gravity wave-turbulence interactions (Mahrt, 2010) in stable conditions. Likewise the soft flag ($\mathrm{f}_{S,\xi} = 1$) for instances where $\mathrm{RN}_{\xi w}$ is in the intermediate range is indicative of conditions that are still transient, but stationary enough for the block statistics to be considered for the purposes of diagnosing surface exchange budgets. Soft flagged blocks should not be used in 'fundamental research' (Foken and Wichura, 1996) such as testing similarity hypotheses based on stationary and horizontally homogeneous conditions. The category that remains unflagged ($\mathrm{f}_{S,\xi} = 0$) where $\mathrm{RN}_{\xi w} < 0.3$ can be considered to be quasi-stationary and thus invaluable for all the purposes of this study.

The stationarity flags for the kinematic vertical fluxes of longitudinal-momentum ($\mathrm{f}_{S,u}$), buoyancy (f_{S,T_s}) and water vapor (f_{S,ρ_v}) are combined into an overall block stationarity flag f_S by taking the median value (operator $\{\}$) of the respective stationarity factors, i.e.

$$\{\mathrm{RN}\} = \{[\mathrm{RN}_{uw}\ ,\ \mathrm{RN}_{T_s w}\ ,\ \mathrm{RN}_{\rho_v w}]\}\ , \tag{2.62}$$

and subjecting the median RNF, $\{\mathrm{RN}\}$, to the same thresholds as the individual relative nonstationarity factors. So for a given block at least two of the RNFs have to be hard flagged ($\mathrm{f}_{S,\xi} = 2$) for the block itself to be hard flagged and discared due to transient conditions.

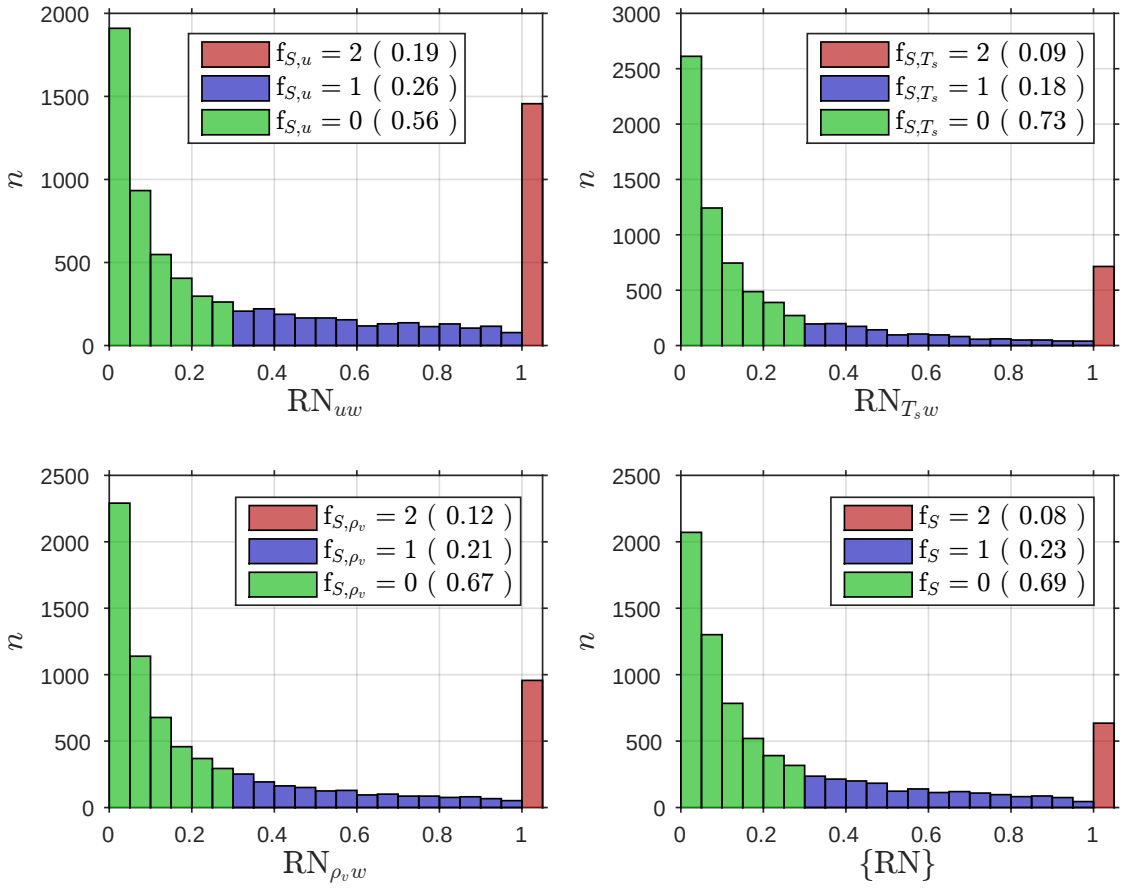


Figure 2.16: *Binned occurences (number of blocks n) of relative nonstationarity factors with corresponding color coded flags as indicated in the legends with the (rounded) fraction of blocks flagged in brackets. The 8082 blocks that were not hard flagged by the previous three flag categories (faulty,distortion and vertical velocity flag) are included in each panel. The last bin includes all relative nonstationarity factors > 1 hence the jump in n. Symbols as defined in the text.*

Again this choice is somewhat arbitrary and we could have instead taken the mean or the maximum of the three RNFs; but we take the median in hope of discarding as little data as possible whilst still identifying periods that clearly defy the underlying assumption of stationary block statistics. We note from Figure 2.16 that only 8% of the blocks are discarded on the sole grounds of this steady-state test, which still leaves an ample 7355 30-minute blocks for analysis. Our values for the individual stationarity flags are largely in agreement with other eddy-covariance campaigns in the area. For example Lüers and Bareiss (2011) report the highest-quality flag ($f_S = 0$) occured 92% of the time for the momentum flux and 73% of the time for the buoyancy flux during the short ARCTEX campaign (May 7th-May 19th 2006) closer to Ny Ålesund. For the Bayelva site over a longer period (March 2008 -March 2009) Westermann et al. (2009) report that about 15% of the values for both the buoyancy and vapor fluxes had to be discarded due to hard flagging. It is interesting to note that in a more recent analysis of the same period, at the same site and with the same software (TK2) Lüers et al. (2014) report that the water vapor flux was only hard flagged in 10% of the time. The main difference between the two analyses was that in Lüers et al. (2014), as with our analysis, the planar fit method was employed for the rotation wheareas in Westermann et al. (2009) the classic natural wind frame (double rotation method) was used.

2.6.5 Integral Turbulence Characteristics

Use of so-called integral turbulence characteristics (after Tillman (1972)), henceforth ITCs, was proposed in the context of EC QC by Foken and Wichura (1996). These ITCs are based on an offspring of MOST, namely flux-variance similarity theory (see Tillman (1972) and Wyngaard et al. (1971)). According to flux-variance similarity in the surface layer the variance of a quantity ξ, $\sigma_\xi^2 = \overline{\xi'^2}$, when properly nondimensionalzied

becomes a univeral function of an appropriate dimensionless scaling factor, $\tilde{\pi}$,

$$\frac{\sigma_\xi}{\xi_*} = \psi_{\sigma_\xi}\left(\tilde{\pi}\right) . \tag{2.63}$$

In the above ψ_{σ_ξ} is the universal function and the normalization factor ξ_* is the surface layer scale for the vertical turbulent flux associated with ξ and σ_ξ is the standard deviation (square root of the variance) of ξ. The surface layer scale is friction velocity u_* for velocities and the dynamic temperature $T_* = -\overline{T_s' w'}/u_*$ for sonic temperature (e.g. Wyngaard et al. (1971)). It is worth emphasizing that flux variance similarity is applicable beyond QC of EC data. From (2.63) it is clear that if the variance of ξ and the value of the dimensionless number $\tilde{\pi}$ is known then ξ_* can be diagnosed from which the turbulent flux of ξ is readily estimated indirectly. This is the essence of the so-called Eddy *Correlation* method (Foken, 2008b) which has an advantage over the EC method in that it may potentially require fewer measurements. However, in being indirect relying on the fullfilment of flux variance similarity it is not universally applicable for diagnosing turbulent fluxes. The EC method, on the other hand, is universally applicable in this regard as turbulent fluxes can be measured directly.

Recall that whether or not these turbulent fluxes are simply related to the surface exchange depends on the approximate fullfilment of the underlying assumptions including the ideal horiontally homogeneous and statistically stationary surface layer (Section 1.2.1). It was precisely with these assumptions in mind that Foken and Wichura (1996) proposed ITCs as a quality control for EC data. That is, flux-variance theory is based on MOST, which in turn requires statistical stationarity and horizontally homogeneous conditions (e.g. Högström (1996)). As such, the more the measured ITC deviates from the 'universal' functions given by (2.63), the further we depart from the ideal surface layer in which Foken and Wichura (1996) refer to turbulent conditions as 'well-developed'. An updated list of proposed ITCs for $\xi = u, T, w$ across a range of stabilites are given in table 1 of Thomas and Foken (2002) based on the results from a number of field experiments. For the most part these are campaigns conducted at lower latitudes. e.g. EBEX-2000 (Oncley et al., 2007) in the San Joaquin Valley (California, USA), with one campaign at a similar (absolute) latitude as Bayelva, namely FINTUREX (Sodemann and Foken, 2005) conducted in the austral summer of 1994 at the Neumayer Station ($70°40'$S) on Antarctica.

Now one would assume that in well-developed turbulent conditions the constants in the universal functions given by Thomas and Foken (2002) are indeed universally applicable, however, as shown in the review of Högström (1996) there is considerable scatter in the universal constants given by various investigations into MOST. A prime example is the controversy surrounding the Von Karman 'constant', κ, which has been reported to have values covering the range $0.32 - 0.65$; though the widely accepted value is $\kappa = 0.4$ (Högström, 1996). For κ, Högström (1996) concludes that these variations stem soley from experimental uncertainty and that κ is probably a true constant. Nonetheless, we felt it prudent to consider the fact that the scatter in constants could also be due to permanent local effects such as the distribution of obstacles as briefly mentioned in Thomas and Foken (2002), and so we tuned the constants in ITC parametrizations accordingly while keeping the functional form the same. The role of such local effects is also discussed in Foken and Wichura (1996) where it is proposed that systematic deviations in the measured ITCs from the parametrized ITCs could be used to detect mechanical turbulence generated by the instruments, mast or other obstacles. As we did not want to discard large chunks of data on account of a *systematic* deviation from parametrized ITCs we revised these by tuning the constants to account for local effects whilst keeping the functional form and in particular the asymptotic slopes on which there is greater agreement in the literature.

We ended up only considered ITCs for the (planar) vertical velocity w in that we were uncomfortable with the scatter in both the proposed and revised ITCs for u and T_s. For the unstable range $-3 \leq \zeta < -0.2$ the parametrized ITC employed was that originally proposed by Panofsky et al. (1977), where $\tilde{\pi} = \zeta$, which takes the form

$$\left(\frac{\sigma_w}{u_*}\right)_{\mathcal{P}} = 1.3\left(1 + c_U|\zeta|\right)^{1/3} \tag{2.64}$$

where we have used, and will continute to use, the subscript $\mathcal{P}$ to emphasize parametrized ITCs as opposed to the measured ITCs. The parametrization in (2.64) is constrained by the conditions that $\frac{\sigma_w}{u_*} \to 1.3$ in the neutral limit (as $\zeta \to 0$) and that $\frac{\sigma_w}{u_*}$ varies as $|\zeta|^{1/3}$ in the free convection limit ($\zeta \to -\infty$) as predicted by Priestley (1954) and later verified by field campaigns (e.g. Wyngaard et al. (1971). The constant c_U on the other hand is reviseable with $c_U = 2$ proposed in Thomas and Foken (2002) compared to the originally proposed $c_U = 3$ in Panofsky et al. (1977). For the near neutral range $-0.2 \leq \zeta \leq 0.4$ the form of the paramtrized ITC for the vertical velocity variance follows that proposed in Thomas and Foken (2002), with $\tilde{\pi} = \ln\left(1/\mathrm{Ro}_*^+\right)$, namely

$$\left(\frac{\sigma_w}{u_*}\right)_{\mathcal{P}} = c_{N1}\ln\left(1/\mathrm{Ro}_*^+\right) + c_{N2} . \tag{2.65}$$

In the above c_{n1} and c_{n2} are constants and $\mathrm{Ro}_*^+ = u_*/f_c z^+$ is a height independent surface layer Rossby number where $z^+ = 1\,\mathrm{m}$ is defined for dimensional

convenience (Thomas and Foken, 2002) and f_c is the Coriolis parameter. This parametrization is based on the work of Högström (1990) in which it was found that a revised similarity formulation employing the surface layer Rossby number $\mathrm{Ro}_* = u_*/zf$ in the scaling factor $\ln(1/\mathrm{Ro}_*)$ was more appropriate than ζ in the near neutral surface layer. Thomas and Foken (2002) found less scatter when using a fixed value $z^+ = 1\,\mathrm{m}$ instead of the measurement height; we verified that this was also the case for the provided data from the Bayelva EC system. The values of the constants for the slope and the intercept of the scaling factor in the neutral range proposed by Thomas and Foken (2002) are $c_{N1} = 0.21$ and $c_{N2} = 3.1$ respectively.

As briefly mentioned we revised the constants c_U, c_{N1} and c_{N2} to account for local effects at Bayelva. This was achieved in `MATLAB` by iteratively fitting the scaling factors in the proposed functional form, that is (2.64) and (2.65) with the c's as free parameters, to all the high quality measured block ITCs based on module output in the corresponding stability range via least squares regression. In such a way we arrive at revised ITC parametrizations by only considering blocks that were not hard or soft flagged by any of the previously discussed QC routines as such high quality data should satisfy the underlying assumptions of MOST. After each iteration we only considered blocks where the normalized absolute deviation between the measured ITC and the least squares fit was less than 0.3 for the least squares fits in subsequent iterations. We did this for 100 iterations in the two stability ranges considered. In both ranges the fits converged rapidly to the revised ITC parametrizations that we then employed in the ITC test in the `Fortran 90` module. In the near neutral range the result was $c_{N1} = 0.24$ and $c_{N2} = 2.9$ as the revised constants in (2.65), whereas in the unstable range the result was $c_U = 1.9$ as the revised constant in (2.64).

Having arrived at the revised parametrizations for the ITCs adapted to the Bayelva data we calculated the normalized absolute deviation, $\delta\mathrm{ITC}_{\sigma_w}$, from these parametrizations in each block, provided that it fell within either the given unstable or neutral stability range, following Foken and Wichura (1996)

$$\delta\mathrm{ITC}_{\sigma_w} = \left| \frac{\frac{\sigma_w}{u_*} - \left(\frac{\sigma_w}{u_*}\right)_{\mathcal{P}}}{\left(\frac{\sigma_w}{u_*}\right)_{\mathcal{P}}} \right| .$$

Subsequently the block is flagged for the deviation in the w-ITCs according to the admitedly somewhat arbitrary thresholds, based on long term experience, given in Mauder and Foken (2004) and Mauder and Foken (2011)

$$\mathrm{f}_{\mathrm{ITC}_{\sigma_w}} = \begin{cases} 2 & \text{if } \delta\mathrm{ITC}_{\sigma_w} > 0.75 \\ 1 & \text{if } 0.3 \leq \delta\mathrm{ITC}_{\sigma_w} \leq 0.75 \\ 0 & \text{otherwise} \end{cases} . \quad (2.66)$$

As with the other QC routines these flags were passed to the module to form a combined block quality flag and to track the result of the individual flagging procedues. The form of the proposed and revised ITC parametrizations with respect to the measured ITCs in both stability ranges is visualized in Figure 2.17, from which it is clear that the revision is considerable especially in the near neutral range (in the top panel). We have used the coefficient of variation, CV, calculated as the root mean square deviation of the revised ITC parametrizations from the measured ITCs normalized by the mean of the measured ITCs, to quantify scatter. This was done both for the w-ITCs shown (panel headers in Figure 2.17) and considered in the module, as well as for u and T_s-ITCs which we chose not to include in the module QC. Relatively speaking the CV, and hence the scatter, was quite small (< 0.2) for the w-ITCs in both stability ranges, hence their inclusion, when compared to u and T_s where the CV exceeded 0.5 in the unstable regime for the u-ITCs and 0.4 in the neutral regime for the T_s-ITCs. Furthermore, the correlation coefficient R between paramtrized and measured ITCs was considerably lower for the latter ITCs when compared to the w-ITCs where $R > 0.5$ (panel headers in Figure 2.17) in both the near neutral and unstable regimes. As a result if we were to employ these other ITCs as a QC we would end up discarding on the order of 10% of the high quality data that passed the remaining QC tests, compared to less than 1% for the w-ITCs.

The fact that there is a large scatter in the u-ITC parametrizations proposed by Thomas and Foken (2002) and references therein is not unexpected in the unstable regime. According to both Stull (1988) and Kaimal and Finnigan (1994), the variance of u in the surface layer under unstable stratification scales with the mixed layer height z_i, the height of the capping inversion (i.e. of the convective boundary layer), and not ζ as proposed in Thomas and Foken (2002). There is no way of determining z_i directly from EC measurements alone so it is not surprising that ITCs do not employ z_i as a scaling factor. As for the calculated T_s-ITCs (not shown) these reproduced the expected hyperbolic cotangent ζ dependence given by the form in Thomas and Foken (2002) and identified by Jocher et al. (2014) in a short field campaign in Ny Ålesund. Nonetheless, we found considerably more scatter particularly in the near neutral and stable ranges. In fact it was largely due to the fact that no specific parametrizations for ITCs are available in the stable regime, as acknowledged in Mauder and Foken (2004)[27], that Westermann

[27] Where the unstable parametrizations are used for lack of a better option in the stable regime.

et al. (2009) steered entirely clear of the use of ITCs in their QC procedure. In line with this we did not try to adopt ITCs parametrizations and flagging in the stable regime, finding considerable scatter in measured ITCs. Even so, at least in the neutral and unstable regimes the w-ITC parametrizations (green lines) displayed in Figure 2.17 show relatively little scatter and overall agreement with previous investigations (blue lines) and are as such useful indicators of the well developed turbulent conditions alluded to in Foken and Wichura (1996).

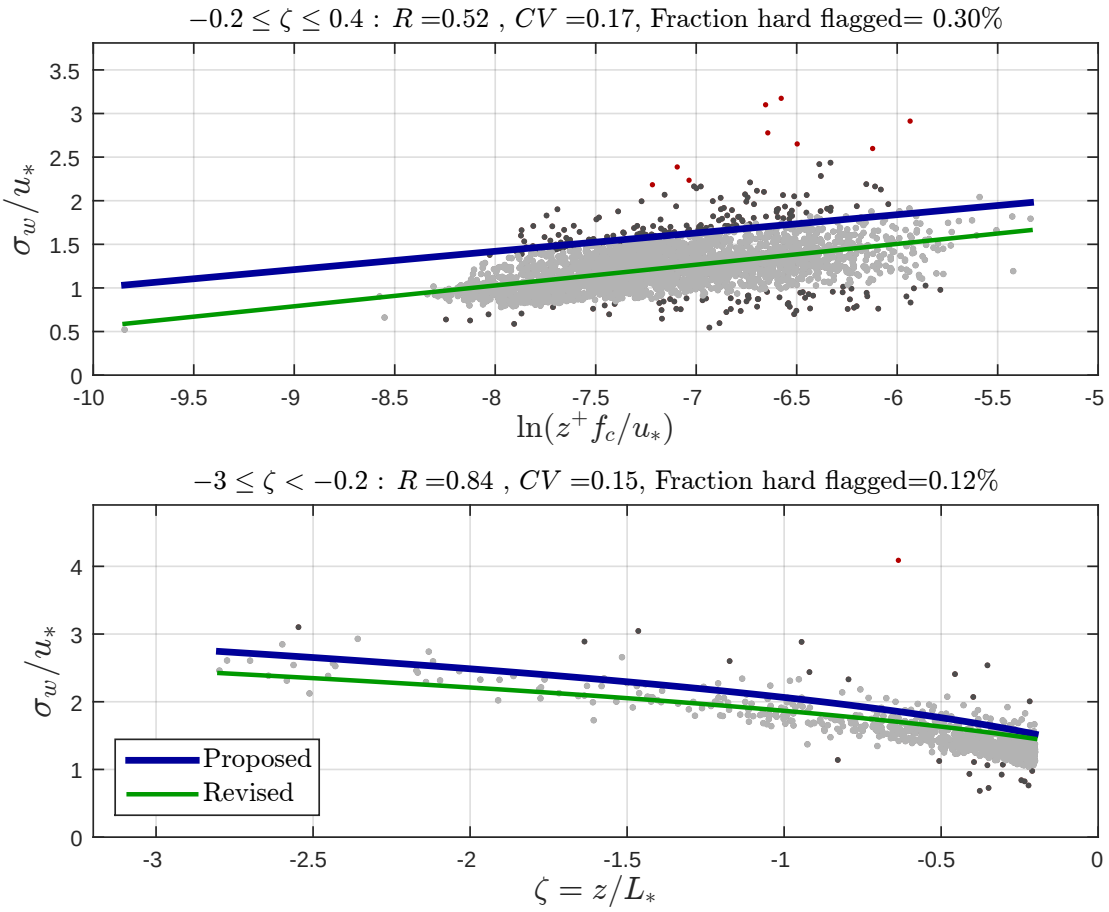


Figure 2.17: *The measured block ITCs σ_w/u_* (dots) as well as proposed (blue lines) and revised (green lines) parametrizations for the ITCs σ_w/u_* as functions of the scaling factors in the two stability ranges considered. Flags for the measured ITCs are color coded according to the flag in* (2.66) *with hard flags in red, soft flags in dark gray and high quality flags in light gray. Upper panel: The near neutral range ($-0.2 \leq \zeta \leq 0.4$) where the proposed parametrization of Thomas and Foken (2002) is given by* (2.65) *with $c_{N1} = 0.21$ and $c_{N2} = 3.1$ while the revised ITC is also given by* (2.65) *but with $c_{N1} = 0.24$ and $c_{N2} = 2.9$ as employed in the module. Lower panel: The unstable range ($-3 \leq \zeta < -0.2$) here the proposed parametrization of Panofsky* et al. *(1977) is given by* (2.64) *with $c_U = 3$ while the revised ITC is also given by* (2.64) *but with $c_U = 1.9$ as employed in the module. The value of the correlation coefficient and coefficient of variation between the revised and measured ITCs as well as the fraction of blocks under consideration that were hard flagged are displayed at the top of respective panels. Only blocks with high quality flags from remaining QC routines ($f = 0$) are displayed.*

2.6.6 Combined Block Quality Flag

Finally, all the quality flag categories are united into an overall combined block quality flag. Such a synthesis of the quality control procedures provides a simple tool that enables one to diagnose the general 'health' of a block of data and whether or not to include this block in subsequent analysis. Denoting the combined block quality flag as f_B we define it simply as the maximum value of all the quality flags for the given block, i.e.

$$\mathrm{f}_B = \max\left(\left[\mathrm{f}_{SK}, \mathrm{f}_F, \mathrm{f}_D, \mathrm{f}_V, \mathrm{f}_S, \mathrm{f}_{\mathrm{ITC}_{\sigma_w}}\right]\right) ,$$

where the skewness-kurtosis flag f_{SK} is set by (2.9), the faulty flag is set by (2.57), the flow distortion flag f_D is set by (2.6.2), the vertical velocity flag f_V is set by (2.59), the stationarity flag f_S is set by subjecting (2.62) to the thresholds in (2.61) and the vertical velocity integral turbulence characteristics flag is set by (2.66). The idea behind using the maximum value of all the block quality flag categories is that if any one flag indicates that the block is of poor quality (hard flagged) with re-

gards to either violation of the underlying assumptions in the point EC method, such as departures from statistical stationarity, *or* instrument related issues, such as flow distortion then the entire block should be ignored when estimating the surface exchange. Similarly if the block is not hard flagged by any of the categories but at least one flag is a soft flag (medium quality) then the block can still be included in the estimation of surface exchange but it should be excluded from investigations into the flux variance similarity (integral turbulence characteristics). To summarize, we reiterate that if a block is hard flagged by the combined block quality flag (i.e. $f_B = 2$) then all samples in the block are set to -1000 and discarded from further analysis.

2.6.7 Estimating Flux Sampling Uncertainty

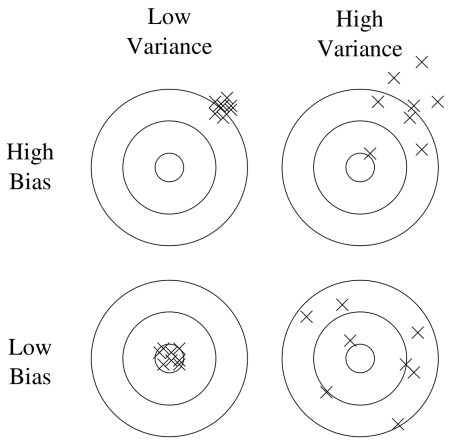


Figure 2.18: *The dartboard analogy. Imagine that we are playing darts and the* true *average is represented by the bullseye in the center of each dartboard. For each dartboard the crosses represent hits by thrown darts corresponding to our samples. In the corresponding texts for each board the bias is the systematic error and the (random) variance is the mean square random error. Figure adopted from Domingos (2012).*

In the final QC step we are interested in quantifying the uncertainty in our final block averaged flux estimates. The smaller (larger) the uncertainty relative to the estimated flux the more (less) confident we can be in the accuracy of our estimate (Aubinet et al., 2012). So, regardless of its magnitude relative to the estimated flux value, the flux sampling uncertainty is always a valuable diagnostic. As pointed out in both Finkelstein and Sims (2001) and Billesbach (2011) it is unfortunately not yet standard practice to include uncertainties when reporting fluxes produced through the EC method. For example no uncertainty procedure is included in the TK2 package (Mauder and Foken, 2004).

Uncertainties in our estimates, not just for fluxes but for any statistic, arise as a result of both systematic and random errors. Following the insightful discussion in Aubinet et al. (2012) (Chapter 7), we will distinguish between these two error sources for estimates within a given block of EC data. We may express some instantaneous EC estimate, obtained after processing raw data, ξ with respect to its (generally unknown) true value $\widehat{\xi}$ as

$$\xi = \widehat{\xi} + \delta_\xi + \epsilon_\xi \ , \tag{2.67}$$

where δ_ξ and ϵ_ξ are the systematic and random error respectively. From (2.67) the instantaneous error is the sum of these two error components, that is $\xi - \widehat{\xi} = \delta_\xi + \epsilon_\xi$. These are both depicted conceptually in Figure 2.18. In the context of a block average the systematic error is defined as the bias, the block averaged error, that is $\delta_\xi = \overline{\xi - \widehat{\xi}}$ (cf. Figure 2.18). Thereby it follows from (2.67) that the random error is the deviation in the instantaneous error from the bias, i.e. $\epsilon_\xi = \xi - \widehat{\xi} - \left(\overline{\xi - \widehat{\xi}}\right)$. So, remembering Reynolds averaging rules, it is clear that the block average random error is by definition zero. It is worth emphasizing for random error that, despite being zero on average, it influences other statistics such as variances and covariances signifcantly, while the systematic error does not. For example the relation between the true and estimated variance of ξ becomes

$$\overline{\widehat{\xi}'^2} = \overline{\xi'^2} - 2\,\overline{\xi'\epsilon_\xi} + \overline{\epsilon_\xi^2}\,, \tag{2.68}$$

where the term $\overline{\epsilon_\xi^2}$ is equivalent to the variance of the random error depicted in Figure 2.18.

With the previous discussion hopefully having provided an introduction to uncertainty, there is one outstanding issue. When employing the EC method we usually have no idea what the *true* value, of a flux or any other statistic, is. Measurements with higher accuracy (Finkelstein and Sims, 2001) are typically absent in the field and so calibration is not possible. With this in mind what are the sources of random and systematic error in the EC method? Adressing and describing all of these sources is beyond the scope of this work. As such, we are satisfied with presenting a summary following Aubinet et al. (2012) and Mauder et al. (2013). In the former references random error is attributed to the following three factors:

1. Turbulence being inherently stochastic ('random').
2. Noise attributed to the instruments.
3. A changing flux footprint.

Whereas systematic error is attributed to the following three factors:

1. Departures from the underlying assumptions in the EC method. For example a transient and/or horizontally heterogeneous surface layer.
2. Problems with instrument calibration.
3. Mistakes in EC data processing, in particular in the flux corrections and coordinate rotations.

Now the magnitude of the bias in our estimates is generally unknown. Nonetheless, we assume that the various steps in our EC processing module has sufficiently reduced the bias; in that this is their primary goal (Billesbach, 2011). So under this assumption we only have to contend with random error. Moreover, in line with Aubinet et al. (2012), we assume that the contribution towards random error of instrumental noise is negligible with respect to the stochastic error. In addition we assume that for the case of Bayelva a signifcant change in the flux footprint within an averaging period is unlikely in line with results of Westermann et al. (2009) based on the model of Schmid (1994). Note that these assumptions do not mean that we are overestimating the stochastic error, instead we are probably underestimating other error sources.

So what remains is to estimate the uncertainty in flux estimates associated with the stochasticity of turbulence (Mauder et al., 2013). Essentially this error is related to the inadequate sampling of the larger integral scale eddies that provide the greatest contribution to the true ensemble averaged flux (Lenschow et al., 1994). Consequently this random error is often refered to as sampling error (e.g. Finkelstein and Sims (2001)). The obvious solution to such an error would be to extend our averaging period to ensure adequate sampling, *however*, the catch is that the longer this period, the more likely we are to depart from statistical stationarity in our estimate (Lee et al., 2006). So, as the lesser of two evils, we are forced to contend with a certain degree of random error in our flux estimates; what remains is quantifying the uncertainty introduced.

A plethora of methods have been introduced to estimate the random, or sampling, error in flux estimates. Various examples are given in Lenschow et al. (1994) and Kaimal and Finnigan (1994). However, the problem with these methods is that they rely on estimating a parameter through relatively arbitrary means. This parameter is usually the integral time scale, discussed in Section 2.7.2. The issue with such a parameter is that its magnitude varies considerably with the method used in its estimation (see e.g. Finkelstein and Sims (2001) and Billesbach (2011)).

Consequently, we adopted the rigorous method of Finkelstein and Sims (2001), which requries no such arbitrary parameter (Billesbach, 2011), when estimating the flux sampling uncertainty as the variance of the flux (Appendix A.5). Such a calculation first demands a LDT of the sample time series in question because both the auto and crosscovariances are ill defined in the presence of a trend (cf. Finkelstein and Sims (2001) and Mauder et al. (2013)). Such a LDT, internal to this routine, should not distort the data considerably since we have already discarded blocks where the flux estimates depart significantly from statistical stationarity (Section 2.6.4). Subsequently we accounted for the effects of the propogation of uncertainty, following Billesbach (2011), in our particular EC processing module as outlined in Appendix A.6. As such we could diagnose the absolute uncertainty in our kinematic heat flux estimates using sensible and latent heat flux estimates. This was done via (A.46) for the kinematic sensible heat flux uncertainty, denoted σ_{F_T}, and (A.47) for the kinematic latent heat flux uncertainty, denoted $\sigma_{F_{\rho_v}}$, using (A.42) to estimate the variances of covariances. It is instructive to express the uncertanties relative to the magnitude of the flux. Due to the dynamic flux conversion factors canceling out, the dynamic relative flux uncertanties equal their kinematic counterparts such that for the sensible heat flux

$$\sigma_{Q_H}/|Q_H| = \sigma_{F_T}/|F_T|\,, \tag{2.69}$$

and for the latent heat flux

$$\sigma_{Q_E}/|Q_E| = \sigma_{F_{\rho_v}}/|F_{\rho_v}|\,. \tag{2.70}$$

2.7 Output

After having undergone all the processing steps that is: despiking, rotations, corrections and quality control the module produces a range of different arrays that are written to an output file. Each output file is in NetCDF format with the same name as the corresponding input file, but with a .cdf as opposed to .dat filename extension. The advantage of using the NetCDF format is that a good compromise is reached between file size, smaller than ASCII format but larger than binary, and the simplicity and speed of reading in individual arrays for visualization in an environment such as `MATLAB`. Herein we will briefly outline two of the three main categories of variables output by our module: 1. Block averaged statistics including fluxes and 2. Autostatistics. The final and third category, the turbulence (co)spectra and in particular how these are produced, has not yet been fully described and is thus outlined in a separate section (Section 2.8).

2.7.1 Block statistics

Our module outputs a wide range of block averaged statistics. Here we will simply list these unless undiscussed steps are used in their calculation. In line with typical EC practice (e.g. Lee et al. (2006)), the output block statistics are all based on the 30 minute non overlapping block averages[28]. To compliment the ensuing outputs a 3 dimensional array containing the timestamp t for the begining and end of each block is also produced in year, month, day, hour, minute, second format.

In that many symbols are previously defined we will present most of the block averaged outputs symbolically. All the ensuing symbols are output as 1 dimensional arrays. First in line are the output block averages of single variables these are: $\overline{u}$, $\overline{v}$, $\overline{w}$, $\overline{T}_s$, $\overline{T}_v$, $\overline{T}$, $\overline{\rho}_v$ and $\overline{q}$. In additon, the 6-hour pressure P and the diurnal measurement height z_m are also output as arrays with the same dimension and corresponding timestamps. Next in line are the variances: $\overline{u'^2}$, $\overline{v'^2}$, $\overline{w'^2}$, $\overline{T_s'^2}$ and $\overline{\rho_v'^2}$. The following covariances (turbulent fluxes) are also written to the output file: $\overline{u'w'}$, $\overline{v'w'}$, $\overline{T_s'w'}$ and $\overline{\rho_v'w'}$. All the block quality flags outlined in Section 2.6 as well as the flux fractional sampling errors (Section 2.6.7) and the correction factiors discussed in Section 2.5 are also returned as output by the module.

Additonaly after processing, in particular flux corrections, the following variables are calculated on a block by block basis. First off the friction velocity is given by (Stull, 1988)

$$u_* = \left(\overline{u'w'}^2 + \overline{v'w'}^2\right)^{1/4}$$

next in line is the block M-O stability parameter (Kaimal and Finnigan, 1994)

$$\zeta = z_m/L_* - \frac{z_m \kappa g \overline{T_s'w'}}{\overline{T}_s\, u_*^3}\,.$$

Finally the dynamic (energy) fluxes of sensible heat and latent heat (units Wm^{-2}), components in the SEB, are computeted. First off the block averaged sensible heat flux is computed through (cf. Fuehrer and Friehe (2002), Van Dijk et al. (2004), Mauder and Foken (2011) and Aubinet et al. (2012))

$$Q_H = \overline{\rho}\, c_p\, \overline{T'w'}\,, \tag{2.71}$$

where $c_p = c_{p,d}\,(1 + 0.84\,\overline{q})$ (Mauder and Foken, 2011) is the secific heat at constant pressure of moist air in which $c_{p,d} = 1004\,\mathrm{JK}^{-1}\mathrm{kg}^{-1}$ is the specific heat at constant pressure of dry air. Secondly the latent heat flux is computed via (Mauder and Foken, 2011)

$$Q_E = \lambda_v\, \overline{\rho_v w}\,, \tag{2.72}$$

where $\lambda_v = 2.501 \times 10^6 - 2360\left(\overline{T} - 273.15\right)$ is the latent heat of vaporization (Foken, 2008b) with units of Jkg^{-1} (with $\overline{T}$ in Kelvin). Note that the latent heat flux involves the total mass flux of water vapor whereas the specific heat flux only involves the turbulent flux. For an explanation of this apparent discrepancy we refer the reader to the extensive discussion Van Dijk et al. (2004). Both the second order dynamic fluxes in (2.71) and (2.72) look almost the same as their first order equivalents, i.e. (1.9) and (1.10). This is far from the case, remember that both the planar fitting and the numerous flux corrections are implicit in these second order fluxes, see Appendix A.6 for their full kinematic form.

2.7.2 Autostatistics

The next category in line for discussion are the autostatistics. These are output based on both an averaging time and a lag increment specified by the user of the module; in our runs we set these parameters to 30 minutes and $\Delta\tau_L = \Delta t$ (0.05 seconds) respectively. Before proceeding with calculating these autostatistics it is common practice (e.g. Kaimal and Finnigan (1994)) to apply a high pass filter such as a LDT to each statistical window. Such filters are needed in that these measures are highly sensitive to the presence of any low frequency trends in the data. Recall that as a result of Section 2.6 highly transient periods are already flagged and discarded so the application of a LDT is not a major intrusion but nonetheless needed to ensure that the autostatistics are well defined. In this section we will let ξ_i'' represent a generic linearly detrended variable with zero mean in a given segmement (averaging period) with I discrete entries (i.e. $i \in 1(1)I$).

We begin with the autocorrelation for a lag time τ_L, which when ξ'' is continuous, is given by (Stull, 1988)

$$R_{\xi''\xi''}(\tau_L) = \frac{\overline{\xi''(t)\xi''(t+\tau_L)}}{\overline{\xi''^2}}\,, \tag{2.73}$$

By definition for $\tau_L = 0$ the autocorrealtion is simply the correlation of $\xi''(t)$ with itself which will always be 1 (Stull, 1988). so for a discrete time series, as is the case for EC measurements, over the set averaging period for a given discreteme lag time $\tau_{L,j} = j\Delta\tau_L$ the

[28] The duration of the block average can be specified by the user of the module, but we will only discuss the traditional 30 minute block average.

autocorrelation is (cf. Section 2.5.1)

$$R_{\xi''\xi''}(\tau_{L,j}) = \frac{\frac{1}{I}\sum_{i=1}^{I-j}\xi_i''\xi_{i+j}''}{\overline{\xi''^2}}\,.$$

The autocorrelation is primarily used to calculate the (Eulerian) integral time scale of ξ'' which we denote $\mathcal{T}_{\xi''}$ strictly defined as (Wyngaard, 2010)

$$\mathcal{T}_{\xi''} = \int_0^\infty R_{\xi''\xi''}(\tau_L)d\tau_L\,.$$

Thereby $\mathcal{T}_{\xi''}$ represents the timescale overwhich ξ'' 'remembers' (is correlated with) itself (Andreas et al., 2008). Effectively we have to truncate the integration range somewhere in that our segments are not infinitely long. So, in practice $\mathcal{T}_{\xi''}$ is estimated from the discrete autocorrelation through a variety of means. Two common examples are 1) Setting $\mathcal{T}_{\xi''\xi''}$ as the e-folding time scale of the autocorrelation function (Kaimal and Finnigan, 1994) having assumed that $R_{\xi''\xi''}(\tau_L) \simeq e^{-\tau_L/\mathcal{T}_{\xi''}}$ and 2) By truncating the integral where $R_{\xi''\xi''}$ reaches a small near constant value (Finkelstein and Sims, 2001). Each method may yield a completely different $\mathcal{T}_{\xi''}$ estimate in that $R_{\xi''\xi''}$ does not always take such a simple exponential form (Finkelstein and Sims, 2001). This is somewhat problematic as the integral time scale is often used in to find an appropriate averaging time for flux estimates (Kaimal and Finnigan, 1994) and in diagnosing sampling errors in the flux estimates (Lenschow et al., 1994). Thus, by calculating the autocovariance we can in turn estimate the integral time scale and compare 'traditional' sampling error estimates (e.g. Lenschow et al. (1994)) with the flux sampling errors obtained using the method of Finkelstein and Sims (2001).

Next in line are the structure functions. Most generally the n-th order structure function in time of a continuous variable ξ'' for a time lag τ_L is given by (Mahrt, 1989)

$$D_{n,\xi''}(\tau_L) = \overline{[\xi''(t+\tau_L) - \xi''(t)]^n}$$

As such, we define the second order discrete structure function in time as (Mahrt, 1989)

$$D_{2,\xi''}(\tau_{L,j}) = \frac{1}{I}\sum_{i=1}^{I-j}\left(\xi_{i+j}'' - \xi_i''\right)^2$$

Note that all structure functions are identically zero for zero lag; so we do not need to calculate this value in that it is known beforehand and not helpful in the typical log-log representation used for structure functions (Stull, 1988). It is somewhat unusual to calculate structure functions in time, as opposed to space, however, by making use of Taylor's Hypothesis then the time lag is related to spatial separation r along the mean wind direction $\overline{u}$ through $\tau_L = r/\overline{u}$. So if turbulence is 'frozen' at a given scale the structure functions in time and space are merely scaled versions of one another. This is very useful in that Kolmogorov (1941) originally made the prediction of an inertial subrange based on the second order structure function in space. Thereby we should be able to identify inertial subranges as regions where a $\tau_L^{2/3}$ power law in time ($r^{2/3}$ in space) is satisfied.
Before closing the section we point out that all the discrete autostatistics are in practice only computed for $j << I$ that is $\tau_L \leq \tau_A/2$. This is done following a reccomendation in Stull (1988). The former points out that in shifting the segment by more than half the length of the segment itself the autostatistics at the given lag would no longer be statistically representative being based on too few samples. In the module the autostatistics are output for all LDT samples measured by the EC system that is $\xi'' = u''$, v'', w'', T_s'', ρ_v''. We will leave LDTs implicit, removing the clumsy ξ'' notation, when presenting and discussing the results.

2.8 Producing Turbulence (Co)Spectra

In the following we we will describe how estimates for various turbulence (co)spectra are produced in our `Fortran 90` module and what these estimates represent. The estimates are arrived at through a method that performs the following sequence of procedures:

- Conditioning of data via gap filling, applying a linear detrend (Gash and Culf, 1996) and tapering (Kaimal and Kristensen, 1991).
- Estimating forwards Discrete Fourier Transforms (DFTs) of the conditioned data using a Fast-Fourier Transfrom (FFT) algorithm via the DFFTPACK (Pumphrey and Swarztrauber, 1985) library.
- Calculating estimates for the resolved discrete (co)spectral intensities, densities and Ogives from the forwards DFTs (Stull, 1988).
- Dealiasing (Kaimal and Finnigan, 1994) the (co)spectral estimates using a suitable transfer function (Gobbi et al., 2006).
- Smoothing the (co)spectral estimates by applying a smoothing window (Konno and Ohmachi, 1998) corresponding to a logarithmically weighted mean.

Such a method could also readily be adapted to an implementation in e.g. `MATLAB`. As excellent reviews of the underlying theory and procedures are presented in both Stull (1988) and Kaimal and Finnigan (1994), on which our method is based, the ensuing description will

be kept as brief as possible with the overall aim of providing a means for the reproduction of our estimates. As usual, the first step in the process is to rotate the wind into the NESF (as described in Section 2.3.3). It is worth emphasizing already at this stage that all the ensuing conditioning of data is internal to the (co)spectral analysis routines and does not affect the already calculated block averaged quantities (e.g. fluxes). Herein our indexing will be in line with `Fortran 90` (and `Matlab`) array indexing such that all indices start at 1 as opposed to the zero indexing that as is used in many textbooks (e.g. Stull (1988)).

Before proceeding the resolution of the spectral analysis must be determined by the user by setting the duration of each segment (data window) to be analyzed. Through trial and error we found that the optimal segment duration was, upon accounting for subsequent *conditioning* (Stull, 1988) of the segments, 3 hours. This duration can of course be tuned as desired depending on what is being investigated. In our case we found that 3 hours was, in most cases, both sufficiently long so as to resolve the large energy containing eddies (Kaimal et al., 1989) and at the same time short enough to avoid departures from stationarity introduced by diurnal variation. In particular a choice of 3 hours allows for a study of what occurs at frequencies just beyond the corresponding standard block averaging period of 30 minutes. Now we may already allocate our discrete frequency array as $f_j = (j-1)\Delta f$ where $\Delta f = f_s/N_T$ is the frequency increment, $j \in 1(1)N_T$, $f_s = 20\,[\text{Hz}]$ is the sampling frequency of the EC system and $N_T = 2.16 \times 10^5$ is the total number of points in a 3 hour segment. The lowest eddy frequency, the first harmonic, we resolve is given by $f_2 = \Delta f = f_s/N_T = 9.26 \times 10^{-5}\,[\text{Hz}]$. As will be shown, at $f_1 = 0\,[\text{Hz}]$ estimates correspond to the mean values of the conditioned series (which are zero).

2.8.1 Conditioning

The sampled time series within each segment must also be conditioned for the subsequent Fourier analysis to be instructive. First off we only include segments where at most 10% of the data is faulty (spikes) to ensure that the estimates are as close to the true 3 hour (co)spectra as possible. So if any of the 30 minute blocks within the segment have been hard flagged, in which case the faulty portion would be $\geq 16\%$, the 3 hour segment is excluded from spectral analysis. Moreover, with each block in the segment having passed the QC (Section 2.6) we can be confident that the segments are near statistically stationary. As Fourier analysis does not allow for any gaps (e.g. discarded spikes) in the data we must also fill any gaps present. To do so we replace the faulty entries by values found via simple linear interpolation[29] of the two nearest valid data samples; with spikes constituting at most 10% of the segment this procedure is not as invasive as it may sound.

A LDT is performed on each of the gap filled time series in the segment. As in Section 2.2.3 we use the least squares method of Gash and Culf (1996) to find the linear trends and subsequently subtract the time dependent trend component from the corresponding entry in the time series. The purpose of the detrend is to high pass filter the segments by removing any unresolved low frequency components that would otherwise 'leak' into the low frequency end of the estimated spectra (Kaimal and Finnigan, 1994). In the absence of a detrend many of our estimates would be contaminated by so-called[30] 'red noise' (Stull, 1988) which manifests as a ficticious increase of the (co)variance contained in the low frequency end of the spectrum. Once more we emphasize that as a result of the QC (Section 2.6) nonstationary blocks where a large trend is typically present have been discarded from subsequent analysis. Hence the LDT procedure does not distort the original series considerably. Nonetheless, it is a necessary procedure in order to provide a faithful representation of the low frequency end of the turbulent (co)spectra. It is worth reemphasizing (see Section 2.2.3) that the mean of a LDT series is zero.

A final procedure in the conditioning of the sample series within each segment is needed before the FFT is applied. This procedure makes up for the leakage introduced by the sharp edges of the segment as a result of the implicit assumption in Fourier analysis that the segment under consideration is infinitely periodic (Stull, 1988). The LDT also helps in this regard, but spectral leakage may still occur from a residual jump (sharp edge) between when the segment ends and begins again periodically. The widely-used solution to the effect of finite sampling introduced by our choice of segment is to multiply by a window, a so-called taper, which reduces the sharp edges at the boundaries removing the corresponding discontinuity in the infinitely periodic extension of the segment (Kaimal and Kristensen, 1991). In Stull (1988) the so-called Bell-taper is presented as the only example. For a segment with $k \in 1(1)K$ evenly spaced discrete samples in time this Bell taper, $W_k^{(B)}$, is defined

$$W_k^{(B)} = \begin{cases} 1 & \text{if } \ 0.1K < k < 0.9K \\ \sin^2\left(5\pi \left\lceil \frac{k-1}{K-1} \right\rceil\right) & \text{otherwise} \end{cases}.$$

[29] If we have gaps at the end points of the segment we simply apply a linear extrapolation backwards (start of segment) or forwards (end of segment) in time.

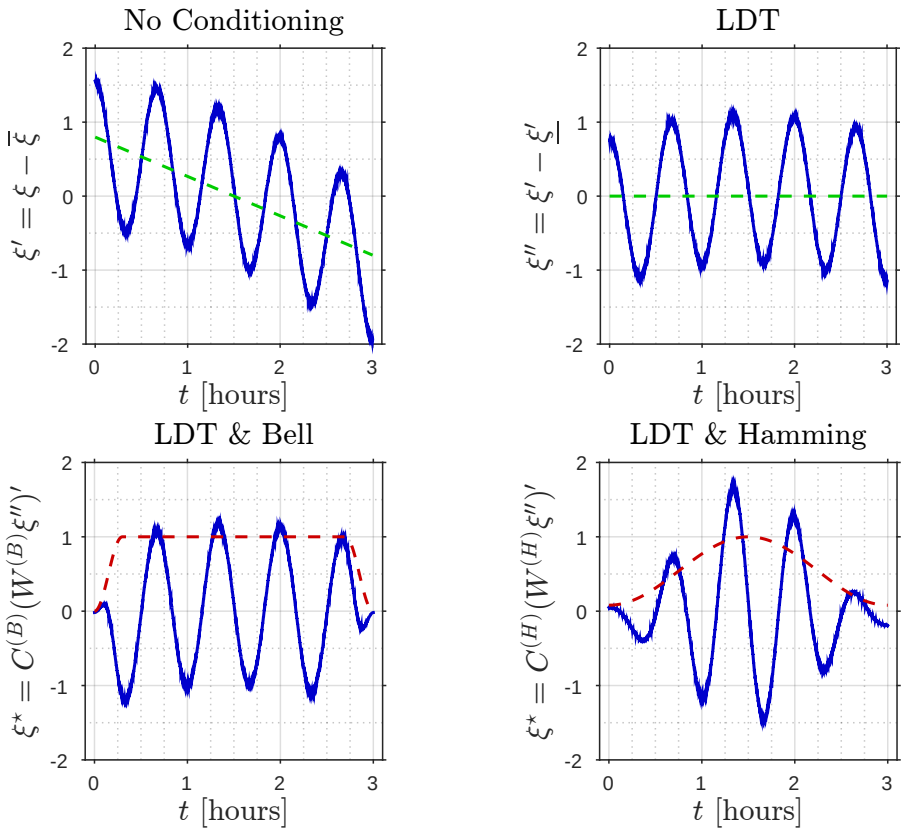
[30] An analogy to the visible portion of the electromagnetic spectrum where red light is found at the lower frequencies.

Kaimal and Kristensen (1991) tested a variety of commonly used tapers on spectra of short (order of one minute) segments and found that the Bell-taper did not perform well. A so-called Hamming window performed best with regards to eliminating the distortion of spectra and the former authors recommended its use even for spectra of longer segments as is our case. The distortion resulting from finite sampling was found to be particularly severe in the intertial subrange as predicted in Kaimal et al. (1989). As one of our goals is to identify the presence and extent of inertial subranges in our (co)spectral analysis it is prudent to follow the reccomendation of Kaimal and Kristensen (1991), which is reiterated in Kaimal and Finnigan (1994), and instead apply the Hamming window as our choice of tapering function. This Hamming window, $W_k^{(H)}$ is defined

$$W_k^{(H)} = 0.54 - 0.46\cos\left(2\pi\left[\frac{k-1}{K-1}\right]\right) .$$

Now the application of any taper will both damp the variance and introduce a residual mean in the LDT time series (Kaimal and Finnigan, 1994). As such we also need to apply a compensation factor such that the variance of the LDT time series is recovered, and remove any residual mean introduced.

Figure 2.19: *Steps in the conditioning of a synthetic 3 hour time series $\xi(t)$. Top left panel: Unconditioned deviation from the block average $\xi'(t)$ (blue sold line) with the linear trend component $\underline{\xi'}(t)$ (green dashed line). Top right panel: Deviation from the block average after applying a LDT $\xi''(t)$ (blue solid line) with the subsequent zero trend component $\underline{\xi''}(t)$ (green dashed line). Bottom panels: Hamming (right) and Bell (left) tapered LDT deviation from the block average $\xi^\star(t)$ (blue solid line) with the red dashed lines as the corresponding taper functions $W^{(B)}$ (left) and $W^{(H)}$ (right); both the conditioned series have been compensated for variance loss and the residual mean has been removed.*



With the above in mind the combined LDT and tapering operation for a given discrete and evenly spaced sample time series ξ_k with $k \in 1(1)K$ entries (in our case $K = N_T$) is applied as follows in the time domain

$$\xi_k^\star = CW_k\left(\xi_k''\right) - C\,\overline{W_k\xi_k''} = C\left(W_k\xi_k''\right)' ,$$

where $\xi_k^\star$ is the tapered and LDT sample, ξ_k'' is the LDT sample[31], C is the compensation factor and W_k is the taper. In the case of the Hamming window the compensation factor given by Kaimal and Kristensen (1991) is $C^{(H)} = \sqrt{2.52}$, while for the Bell taper we found $C^{(B)} = \sqrt{1.16}$ to be the appropriate choice based on sample data. We include the Bell taper since the choice of tape in our module (including no taper at all) is up to the user. However, due to the previous discussion we exclusively employ and recommend the Hamming

[31] As mentioned $\xi_k'' = \xi_k - \underline{\xi_k} = \xi_k' - \underline{\xi_k'}$ i.e. the LDT sample series is equal to the LDT deviation from the average of the sample series.

window when tapering time series prior to spectral analysis. The transformation of a 3 hour long synthetic time series as a result of both the LDT and the application of the tapers is shown in Figure 2.19. We verified that the variance was indeed conserved upon application of either of the two tapers and the corresponding compensation factor with $\sigma^2_{\xi^\star} = \sigma^2_{\xi''} = 0.51$ for the synthetic example in Figure 2.19.

2.8.2 DFTs

Now we are prepared to apply the Fourier transform (FT). Henceforth we will simply denote $\xi^\star$ as ξ leaving the conditioning (LDT, tapering, variance compensation and residual mean removal) of the series as implicit to avoid (even more) cumbersome notation. First we review the basic concept of the discrete Fourier transform (DFT); given a conditioned series in time[32] ξ_k with N_T entries that are evenly separated by the increment Δt, then the forwards DFT $\mathcal{F}_{\xi,j}$ at the discrete frequency $f_j = (j-1)f_s/N_T$ is given in by (e.g. Stull (1988))

$$\mathcal{F}_{\xi,j} = \frac{1}{N_T}\sum_{k=1}^{N_T} \xi_k \omega^{-(j-1)(k-1)}\,, \qquad (2.74)$$

where $\omega = \exp\{\mathrm{i}2\pi/N_T\}$ is the N_Tth root of unity and $\mathrm{i} = \sqrt{-1}$ is the imaginary unit. The inverse ('backwards') DFT has a similar definition (see e.g. Stull (1988)), but as we are not concerned with inverse transformations typically used in solving differential equations we will omitt a discussion of these here. From (2.74) it is readily seen that, with this definition, in the case of $j = 1$ ($f_j = 0$) the forwards DFT simply yields the mean of the ξ_k time series, which as a result of conditioning will always be zero. Moreover we note that, as opposed to ξ, $\mathcal{F}_\xi$ is complex.

Now having conditioned the processed EC data we compute the forwards DFT for multiple 3 hour segments time serie of the following samples $\xi = u, v, w, T_s, \rho_v$ all sampled at $f_s = 20\,[\mathrm{Hz}]$. As mentioned this is done via a FFT routine in DFFTPACK (Pumphrey and Swarztrauber, 1985), where the FFT is essentially a DFT that has been optimized in terms of speed taking advantage of binary computer architecture (Stull, 1988). We will not dwell on the details in the FFT as we are only concerned with the output; namely the estimated forwards DFT. What is worth mentioning is that the forwards DFT output by DFFTPACK is not normalized, so a subsequent division by N_T is required. Moreover, for efficiency the DFFTPACK simply transforms the input array containing ξ_k into an array containing $\mathcal{F}_{\xi,j}$ of the same length (N_T points) with only real entries overwriting the input array (an `INTENT(INOUT)` argument in `Fortran 90`)[33]. In the returned array the first entry is the mean entry (i.e. $\mathcal{F}_\xi(f_1)$), the second entry is the real component of the first harmonic ($\mathrm{Re}\{\mathcal{F}_\xi(f = f_2)\}$) the third entry is the imaginary component of the first harmonic ($\mathrm{Im}\{\mathcal{F}_\xi(f = f_2)\}$) and so on for the ensuing higher harmonics (see source code in Pumphrey and Swarztrauber (1985) for details and exceptions). For numerical convenience we mapped the entries of this array into a complex array containing the mean (strictly real) component along with the *paired* real and imaginary components of the respective harmonics. Note that both these arrays in effect only contain $N = N_T/2 + 1$ (as opposed to N_T) complex forwards DFFTs. As will be shown, this is not a problem.

2.8.3 (Co)Spectral estimates

Although it may not be immediately clear from (2.74), when ξ_k is real, as is our case, there is an inherent symmetry in $\mathcal{F}_{\xi,j}$. It turns out that for $2 \le j < N$ where $N = N_T/2 + 1$ the following holds equality holds (e.g. Stull (1988)):

$$\mathcal{F}_{\xi,j} = \mathcal{F}^*_{\xi,j^+}\,,$$

where

$$\mathcal{F}^*_{\xi,j^+} = \mathrm{Re}\left\{\mathcal{F}_{\xi,j^+}\right\} - \mathrm{i}\,\mathrm{Im}\left\{\mathcal{F}_{\xi,j^+}\right\}\,,$$

is the complex conjugate of F_{ξ,j^+}. The operators $\mathrm{Re}\{\}$ and $\mathrm{Im}\{\}$ pick out the real and imaginary components respectively. Here $j^+ = N_T + 2 - j$ such that $|N - j| = |N - j^+|$; or in terms of frequency $|f_N - f_j| = |f_N - f_{j^+}|$. So for any two frequency bands that are equidistant from, but on opposite sides of, the f_N band the forwards DFT at these bands are simply complex conjugates of one another. As a consequence

$$\mathcal{F}_{\xi,j}\mathcal{F}^*_{\xi,j} = \mathcal{F}^*_{\xi,j^+}\mathcal{F}_{\xi,j^+}\,,$$

that is the absolute values of the forwards DFT which corresponds to the spectral intensity[34], a measure of the contribution towards the variance, are identical at the two frequencies. As such, as shown in the right panel of Figure 2.20, when the spectral intensity is plotted against (linear) frequency it is symmetric about f_N.

[32] Although not directly applicable in the case of point EC data a DFT is equaly valid for a spatial distribution

[33] That is the time series stored in the array ξ_k is lost upon calling the forwards transform. So it is prudent to use a cloned version of ξ_k to avoid loosing ξ_k in the analysis steps.

[34] The same is true for the cospectral inensity soon to be defined.

Thereby no new information on the spectral intensity is contained in frequencies $f > f_N$ so effectively the highest frequency we can resolve is the Nyquist (folding) frequency $f_N = (N-1)f_s/N_T = f_s/2$. Inserting for $f_s = 1/\Delta t$ the Nyquist frequency can be stated alternatively as $f_N = 1/2\Delta t$; meaning that the fastest oscillations we can resolve via Fourier analysis have half a period per sampling interval Δt (Kaimal and Finnigan, 1994). Spectral intensities that include $f_{j^+} > f_N$ are refered to as unfolded and the reflection about f_N is always evident in such a representation. It is more instructive to consider the folded spectra, where the estimates at $f_{j^+} > f_N$ are folded back via doubling of the estimates at the corresponding f_j. In the folded spectral intensity the complete contribution to the variance at each frequency band $f_j \leq f_N$ is recovered. Note that in our case $N = N_T/2 + 1$ is an integer in that N_T is even so the frequency f_N is an entry in f_j. The estimate at the Nyquist frequency itself, where the fold occurs, is not subject to such a doubling. A visualization of a folded spectral intensity (right panel) and its unfolded counterpart (left panel) is shown in Figure 2.20 where the spectral intensity scale is logarithmic. In both cases the sum over all spectral intensities yields the variance, however, with the unfolded representation contributions to the variance are unphysical being located at unresolved frequencies.

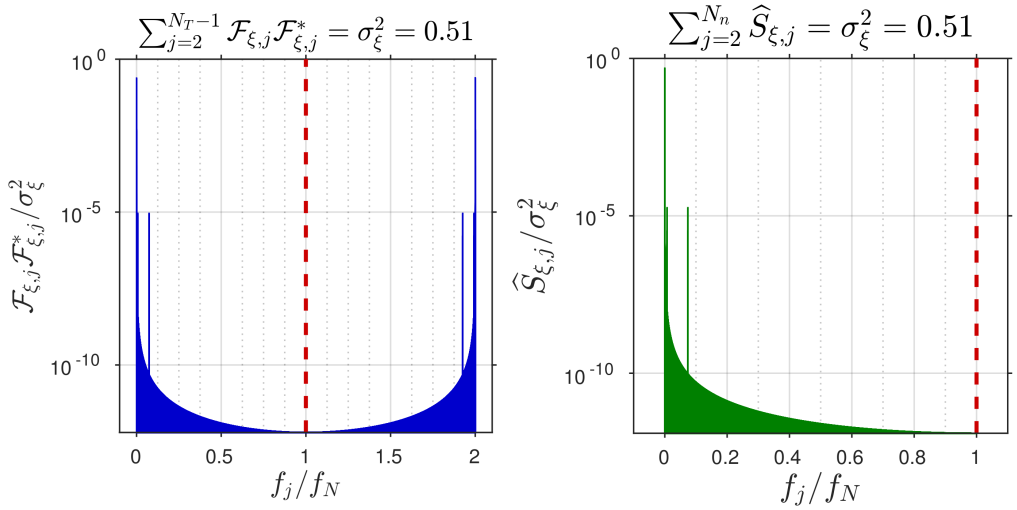


Figure 2.20: *Unfolded and folded spectral intensities in a* log-linear *representation. Left Panel: The unfolded discrete spectral intensity* $\mathcal{F}_{\xi,j}\mathcal{F}^*_{\xi,j}$ *normalized by the variance* σ_{ξ^2} *(blue bars) for* $N_T - 2$ *discrete frequency bins of width* Δf *centered on* $f_j = (j-1)f_s/(N_T)$ *with* $j \in 2(1)N_T - 1$ *normalized by the Nyquist frequency* f_N*. Right panel: The folded discrete spectral intensity* $\widehat{S}_{\xi,j}$ *(green bars), as given by equation* (2.75)*, normalized by the variance* σ^2_ξ *defined for* $N - 1 = N_T/2$ *discrete frequency bins of width* Δf *centered on* $f_j = (j-1)f_s/N$ *for* $j \in 2(1)N$*. The intensities in the right panel are* double *those at the correponding* f_j *in the left panel everywhere except at* f_N*. In both panels the position of this Nyquist (folding) frequency* $f_N = f_s/2$ *is marked by the red dashed line about which the unfolded discrete spectral intensity is symmetric. Also in this example* ξ *is the synthetic and conditioned sample time series displayed in the bottom right panel of Figure 2.19.*

Now as the mean (at $f_1 = 0$) contribution to the spectral intensity is always zero in that $\mathcal{F}_{\xi,1} = \overline{\xi} = 0$ we do not have to concern ourselves with this entry. So we will consequently ignore the mean frequences ($f_1 = 0$) and all entries $j > N$ (in accordance with the output of DFFTPACK). As such, to simplify the notation in the following we will henceforth let $N = N_T/2$ with $j \in 1(1)N$ and $f_j = jf_s/N_T$ with f_1 now representing the first harmonic and $f_N = f_s/2$ (the Nyquist frequency) representing the highest resolved frequency. From the forwards DFT the folded discrete spectral intensity, $\widehat{S}_{\xi,j}$, of the frequency band with width Δf centered about f_j is given by (Stull, 1988)

$$\widehat{S}_{\xi,j} = (2 - \delta_{jN})\, \mathcal{F}_{\xi,j}\mathcal{F}^*_{\xi,j}\,,$$

where δ_{jN} is 1 if $j = N$ and 0 otherwise while $\mathcal{F}^*_{\xi,j}$ is the complex conjugate of $\mathcal{F}_{\xi,j}$. Thereby, the above can be expressed equivalently in the following numerically convenient form

$$\widehat{S}_{\xi,j} = (2 - \delta_{jN})\, \mathrm{Re}\left\{\mathcal{F}_{\xi,j}\right\}^2 + (2 - \delta_{jN})\, \mathrm{Im}\left\{\mathcal{F}_{\xi,j}\right\}^2 . \qquad (2.75)$$

The spectral intensity for frequency band f_j is, as briefly mentioned, the contribution to the variance contained in the given frequency band (Stull, 1988). As such the variance of the conditioned sample time series is simply given by the sum

$$\overline{\xi'^2} = \sum_{j=1}^{N} \widehat{S}_{\xi,j}\,.$$

Similarly if we consider a complimentary conditioned sample time series χ_k, having computed $\mathcal{F}_{\chi,j}$, then the folded discrete cospectral intensity, $\widehat{Co}_{\xi\chi,j}$ of frequency band f_j is given by 198default

$$\widehat{Co}_{\xi\chi,j} = (2-\delta_{jN})\,\mathrm{Re}\left\{\mathcal{F}_{\xi,j}\mathcal{F}^*_{\chi,j}\right\}\,,$$

which analogously to (2.75) may be expressed equivalently in a form that is simple to implement numerically

$$\begin{aligned}\widehat{Co}_{\xi\chi,j} =& (2-\delta_{jN})\,\mathrm{Re}\left\{\mathcal{F}_{\xi,j}\right\}\mathrm{Re}\left\{\mathcal{F}_{\chi,j}\right\}\\ &+(2-\delta_{jN})\,\mathrm{Im}\left\{\mathcal{F}_{\xi,j}\right\}\mathrm{Im}\left\{\mathcal{F}_{\chi,j}\right\}\,. \quad (2.76)\end{aligned}$$

Not surprisingly the cospectral intensity for a given frequency band f_j is the contribution to the *co*variance contained in the given frequency band (Stull, 1988); as such the covariance of the two conditioned sample time series is

$$\overline{\xi'\chi'} = \sum_{j=1}^{N}\widehat{Co}_{\xi\chi,j}\,. \quad (2.77)$$

Now upon comparing (2.75) and (2.76) the spectral intensity is a special case of the more general cospectral intensity in that $\widehat{Co}_{\xi\xi,j} = \widehat{S}_{\xi,j}$, but as with variance and covariance we will keep the distinction (also in the notation) in line with convention in micrometeorological literature (e.g. Kaimal and Finnigan (1994)). Henceforth we will leave the term folded as implicit in that all of the subsequent discussion pretains only to folded (co)spectra.

A representation of (co)spectral intensity is in itself not very useful in that these represent (co)variance contributions for discrete frequency bands as opposed to the continuous (co)spectrum of turbulence (Taylor, 1938) to which the concepts of (co)spectral similarity (Kaimal et al. (1972) and Wyngaard and Coté (1972)) and the intertial subrange (Kolmogorov, 1941) can be properly attributed. Thereby following Stull (1988), noting that the spectral intensity of the band $f \in [f_j - \Delta f/2, f_j + \Delta f/2]$ is the total contribution of all eddies in this band towards the variance then the discrete spectral intensity $\widehat{S}_{\xi,j}$ is related to the continuous spectral *density* S_ξ through

$$\widehat{S}_{\xi,j} = \int_{f_j-\Delta f/2}^{f_j+\Delta f/2} S_\xi(f)\,df\,, \quad (2.78)$$

so we approximate the spectral *density* at f_j, $S_{\xi,j}$, as

$$S_{\xi,j} \simeq \frac{\widehat{S}_{\xi,j}}{\Delta f}\,. \quad (2.79)$$

Comparing (2.78) and (2.79) it is clear that the estimated spectral density $S_{\xi,j}$ is simply the average of continuous spectral density in the band $f \in [f_j - \Delta f/2, f_j + \Delta f/2]$. Moreover, we note that spectral density has units of variance per unit frequency as opposed to spectral intensity which has units of variance. Thereby, the continuous spectral density is related to the segment variance through

$$\overline{\xi'^2} = \int_{f_1}^{f_n} S_\xi(f)\,df\,. \quad (2.80)$$

The above reasoning also holds for the cospectral density, $Co_{\xi\chi}$, where we estimate the cospectral density at frequency f_j through

$$Co_{\xi\chi,j} \simeq \frac{\widehat{Co}_{\xi\chi,j}}{\Delta f}\,, \quad (2.81)$$

with the continuous cospectral density being related to the segment covariance through

$$\overline{\xi'\chi'} = \int_{f_1}^{f_n} Co_{\xi\chi}(f)\,df\,. \quad (2.82)$$

The cospectral density and its relation to the covariance leads naturaly to the Ogive (Desjardins et al. (1989) and Foken and Wichura (1996)), $Og_{\xi\chi}$, which is defined as follows for a given frequency $f_g \leq f_N$

$$Og_{\xi\chi}(f_g) = \int_{f_g}^{f_n} Co_{\xi\chi}(f)\,df\,. \quad (2.83)$$

Comparing (2.82) and (2.83) we note in particular that for $f_g = f_1$ then $Og_{\xi\chi}(f_1) = \overline{\xi'\chi'}$. In words the Ogive is a measure of the cumulative cospectral intensity in all resolved frequencies $f \geq f_g$. Recalling that the period of an eddy is the inverse of its frequency then $Og_{\xi\chi}(f_g)$ an equally be interpreted as the covariance contained in all periods $f^{-1} \leq f_g^{-1}$. Thus, the Ogive is typically used with $\chi = w$ to measure the convergence (or lackthereof) of a vertical turbulent flux within a given averaging period. For numerical convenience we compute the Ogive using the sum over cospectral intensities starting at f_g in that this is equivalent to the numerical integral over cospectral density estimates starting at f_g (cf. (2.77),(2.81),(2.82)). Moreover, for economy we do this by iteratively picking f_g as the nearest f_j to one of the frequencies $f_{C,i}$ that are evenly spaced on a logarithmic scale starting at f_1 and ending at f_N as given by (2.88), where $i \in 1(1)90$. That is for a given $f_{C,i}$ the discrete Ogive, $Og_{\xi\chi,i}$ is approximated by

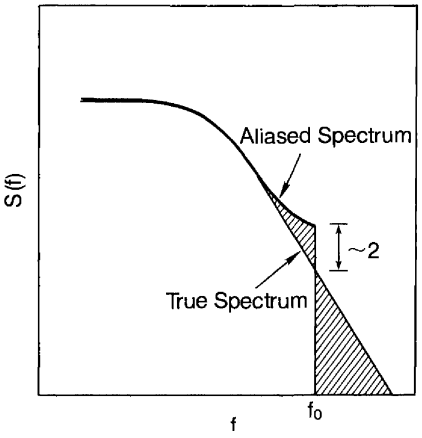
$$Og_{\xi\chi,i} = \sum_{j=l}^{N}\widehat{Co}_{\xi\chi,j}\,, \quad (2.84)$$

where, for a given $f_{C,i}$, the index l is the value of j that minimizes the expression $|f_j - f_{C,i}|$. The sum in (2.84) is carried out cumulatively for $i \in 1(1)90$ to yield the approximate Ogive distribution for the frequency array $f_{C,i}$. Note that with the ogive distribution being essentialy the cumulative sum over cospectral intensities we do not loose any of the covariance contributions with 90 (as opposed to N) entries since we retain the total bandwidth f_N. In particular it is obviously much cheaper to compute cumulative the sum in (2.84) for $i \in 1(1)90$ then for $i \in 1(1)N$ (since $N = 1.08 \times 10^5$).

Contrary to the cospectral density estimates, no smoothing or dealiasing is undertaken in the Ogive calculation since the aliased and unsmoothed cospectra contains all the covariance contained in the signal (Stull, 1988). The Ogive of the spectra, $Og_{\xi\xi,i}$, which corresponds the cumulative distribution of the spectral intensity, is also computed by replacing $\widehat{Co}_{\xi\chi,j}$ in (2.84) with $\widehat{S}_{\xi,j}$. In the module Ogives of the cospectra are computed for the conditioned samples $\xi = u, v, T_s, \rho_v$ with $\chi = w$ and Ogives for the spectra are computed for $\xi = u, v, w, T_s, \rho_v$.

2.8.4 Dealiasing

Figure 2.21: *A schemating representation of the effects of aliasing with $f_0 = f_N$ being the Nyquist frequency. The visualization is a log-log plot of the true spectral density (thin solid line) and the resolved and aliased spectral density (thick solid line) against natural frequency f. The variance contained in the true spectrum at $f > f_0$, hatched area to the right of the vertical line, is folded back in the resolved aliased spectrum falsely raising the spectral density at $f < f_0$ as indicated by the hatched area to the left of the veritcal line. Figure adopted from Kaimal and Finnigan (1994).*



Having computed all the Ogives we need to deal with what turns out to be mainly a cosmetic issue, namely alisasing. Aliasing is related to the previously described folding of the (co)spectra (Figure 2.20) and the fact that we expect some portion of the true (co)variance to exist in the unresolved frequencies $f > f_N$ (Stull, 1988) as shown by the thin solid line in Figure 2.21. The true (co)spectral density in these unresolved higher frequencies f_{j+} is folded back into corresponding lower frequencies f_j creating an aliased spectrum as shown by the thick solid line in Figure 2.21. So although the total (co)variance is resolved by the (co)spectrum, the estimated (co)spectral density in the resolved frequencies f_j is overestimated as an *unknown* portion of these estimates actually belongs to the corresponding unresolved frequencies f_{j+}. Note, however, that aliasing is not a problem if we know a priori that there is a well defined cut off frequency above which there is zero (or negligible) (co)spectral density and that this frequency is resolved (Kaimal and Finnigan, 1994). That is if it is less than or equal to the folding (Nyquist) frequency.

In our case such an ideal cutoff is nearly satisfied in that we are sampling at $f_s = 20\,\text{Hz}$ so our folding frequency $f_N = 10\,\text{Hz}$ is typically far into where we expect to find the inertial subrange. Recall that in the inertial subrange spectral density (cospectral density) is proportional to $f^{-5/3}$ ($f^{-7/3}$) and so a rapid drop off in the contribution towards the (co)variance is expected. Consequently, as noted in Kaimal and Finnigan (1994), the aliasing of unresolved (co)variance is rarely evident at frequencies lower than $f_N/2$. Nonetheless particularly in (co)spectra that are smoothed (using an averaging procedure) aliasing is usually evident near and at the Nyquist frequency where spectral density is typically raised by a factor of two as shown schematically in Figure 2.21. As such, the effects of aliasing often manifest as a sudden cesation of the inertial subrange as is clear from the red curves in Figure 2.22. It is worth emphasizing that aliasing would be a problem even if we did not fold our spectra; the (co)spectral density is always unresolved for $f > f_N$ and our unfolded (co)spectral estimates would just be the mirror image form of the folded (co)spectral estimates as shown in Figure 2.20. In fact in the unfolded representation, with falsely large estimates at $f > f_N$, we are even further from the true form of the (co)spectrum then in the folded case.

To deal with aliasing we simply apply a transfer function, effectively a dealiasing filter, to our (co)spectral density estimates that damps these at $f \lesssim f_N$. An example of a traditional dealiasing filter is given by the transfer function (Gobbi et al., 2006)

$$H_{r,j} = \frac{\left(1 - e^{-\Delta t/\tau_r}\right)^2}{1 + e^{2\Delta t/\tau_r} - 2e^{\Delta t/\tau_r}\cos\left(2\pi f_j \Delta t\right)}\,, \quad (2.85)$$

where τ_r is the time constant of the recursive filter. This is the transfer function corresponding to the classic approach of applying a low pass digital recursive filter in the time domain through to eliminate aliasing prior to Fourier analysis (e.g. Moore (1986))

$$\tilde{\xi}_k = \left(1 - \frac{\Delta t}{\tau_r}\right)\tilde{\xi}_{k-1} + \frac{\Delta t}{\tau_r}\xi_k \quad (2.86)$$

for $k \in 2(1)N_T$ where $\tilde{\xi}_k$ is the low pass filtered version of ξ_k. A frequent choice is τ_r satisfying $\Delta t/\tau_r = 0.8$ (Gobbi et al., 2006); this is a good compromise between not damping (co)spectral densities at frequencies unaffected by aliasing $f \leq 0.5 f_N$ (Kaimal and Finnigan, 1994) whilst still sufficiently damping estimates at frequencies that are significantly aliased ($f \lesssim f_N$). Although such a filter performs reasonably well, whether

it is applied in the time domain through (2.86) or directly in the spectral domain by explicitly using the transfer function in (2.85), it still damps the estimates at $f \lesssim f_N$ too much (see both Gobbi et al. (2006) and Figure 7.8 in Kaimal and Finnigan (1994)). As a consequence Gobbi et al. (2006) developed a dealisaing filter with all the desireable qualities of the recursive filter but *without* the excessive damping at $f \lesssim f_N$. Such a filter can be applied directly to the (co)spectral density estimates through application of the dealiasing transfer function (Gobbi et al., 2006)

$$H_{D,j} = \frac{\left[1 + \cos\left(\pi f_j \Delta t\right)\right]^2}{4} ,$$

through the operations

$$\tilde{Co}_{\xi\chi,j} = H_{D,j} Co_{\xi\chi,j} \ , \ \tilde{S}_{\xi,j} = H_{D,j} S_{\xi,j} \ , \quad (2.87)$$

for $j \in 1(1)N$. In our module the (co)spectral estimates are dealiased through (2.87) the result of which is shown (after smoothing) by the green curves in Figure 2.22. Note that after dealiasing the integral over resolved (co)spectral densities no longer yields the entire segment (co)variance having shifted a small portion of the (co)spectral intensity to the higher unresolved frequencies. This is why the Ogives are not calculated from t he dealiased (co)spectra. Nonteheless, after dealiasing the resolved (co)spectral densities are closer to the true (co)spectral densities for the same frequency range ($f \in [f_1, f_N]$) providing a more realistic representation of the high frequency extent of the inertial subrange. To summarize, the dealisaing procedure is mainly cosmetic in providing a more accurate representation of (co)spectra for $f \lesssim f_N$ where the true (co)spectral density is quite small being well into the intertial subrange (Figure 2.22).

2.8.5 Smoothing

The last step is to smoothe the spectral estimates before these are writen to the output file. The purpose of smoothing is two-fold in that not only is the result easier to interpret visually (Kaimal and Finnigan, 1994), unsmoothed spectra being inherently noisy (gray curves in Figure 2.22), but we also reduce the dimension of the output arrays by several orders of magnitude easing the burden on storeage capacity. As opposed to the commonly used splicing method proposed in e.g. Kaimal and Finnigan (1994), where spectra of high and low pass filtered segments are combined, the method we adopt is more direct and consequently alot simpler to implement. It is worth emphasizing this method, adopted from the seismology literature (Konno and Ohmachi, 1998), is a novel and convenient approach to spectral smoothing. To our knowledge in the field of micrometeorology we are the first to apply this particular smoothing method. First we define a frequency domain that is evenly spaced on a logarithmic scale with the same total bandwidth as our current frequency domain. We will denote these discrete frequencies as $f_{C,i}$ and subsequently define the $i \in 1(1)N_C$ entries (N.B. index i should not be confused with the imaginary unit i) as follows

$$f_{C,i} = f_1 10^{(i-1)\Delta_f} , \quad (2.88)$$

where

$$\Delta_f = \log_{10}\left(\frac{f_{C,i+1}}{f_{C,i}}\right) = \frac{1}{N_C - 1} \log_{10}\left(\frac{f_N}{f_1}\right) ,$$

is the logarithmically constant frequency increment. It is readily seen that $f_{C,1} = f_1$ and $f_{C,N_C} = f_N$ by definition; so we do not loose any resolution as the total bandwidth remains the same (up to the Nyquist frequency f_N). The logarithmic spacing also helps the visual interpretation as spectra are typically displayed in log-log plots allowing for easy identification of any power law behavior (Stull, 1988). Next we define a smoothing window, $W_{ij}^{(\mathrm{KO})}$, based on the work of Konno and Ohmachi (1998) as follows for $i \in 1(1)N_C$ and $j \in 1(1)N$

$$W_{ij}^{(\mathrm{KO})} = \begin{cases} 1 & \text{if } \beta_{ij} = 0 \\ \left[\sin\left(\beta_{ij}\right)/\beta_{ij}\right]^8 & \text{otherwise} \end{cases} , \quad (2.89)$$

where $\beta_{ij} = b \log_{10}\left(f_j / f_{C,i}\right)$ and b is a positive smoothing parameter producing smoother spectra for lower b values.We subsequently normalize this window through

$$\widetilde{W}_{ij}^{(\mathrm{KO})} = \frac{W_{ij}^{(\mathrm{KO})}}{\sum_{j=1}^{J} W_{ij}^{(\mathrm{KO})}} ,$$

then smoothe the spectrum $\tilde{S}_{\xi,}$ (or cospectrum $\tilde{Co}_{\xi w}$) via left multiplication by the matrix $\widetilde{W}_{ij}^{(\mathrm{KO})}$:

$$\widetilde{S}_{\xi,i} = \sum_{j=1}^{N} \widetilde{W}_{ij}^{(\mathrm{KO})} \tilde{S}_{\xi,j} .$$

Essentialy this smoothing operation, for a given $f_{C,i}$, corresponds to a weighted mean of spectral estimates at frequencies logarithmically centered on the frequency $f_{C,i}$; with increasing weight for unsmoothed spectral estimates at frequencies f_j close to $f_{C,i}$ on a logarithmic scale. Through trial and error we found $b = 10$, which corresponds to the actual total bandwidth, to be the best choice for the smoothing parameter when $N_C = 90$ central frequency bands $f_{C,i}$ are used. Note that in (2.89) we modified the exponent from its original value of 4 (as given in Konno and Ohmachi (1998)) to 8 in order to provide narrower lobes in the smoothing window, a logarithmically weighted mean, for a given $f_{C,i}$ accounting for the expected rapid dropoff in the intertial subrange such that estimates at frequencies over a

decade away have negligible influence on the smoothed mean estimate. For the direct result of smoothing, in the absence of a dealiasing filter, the reader is free to compare the gray and red curves for each panel in Figure 2.22.

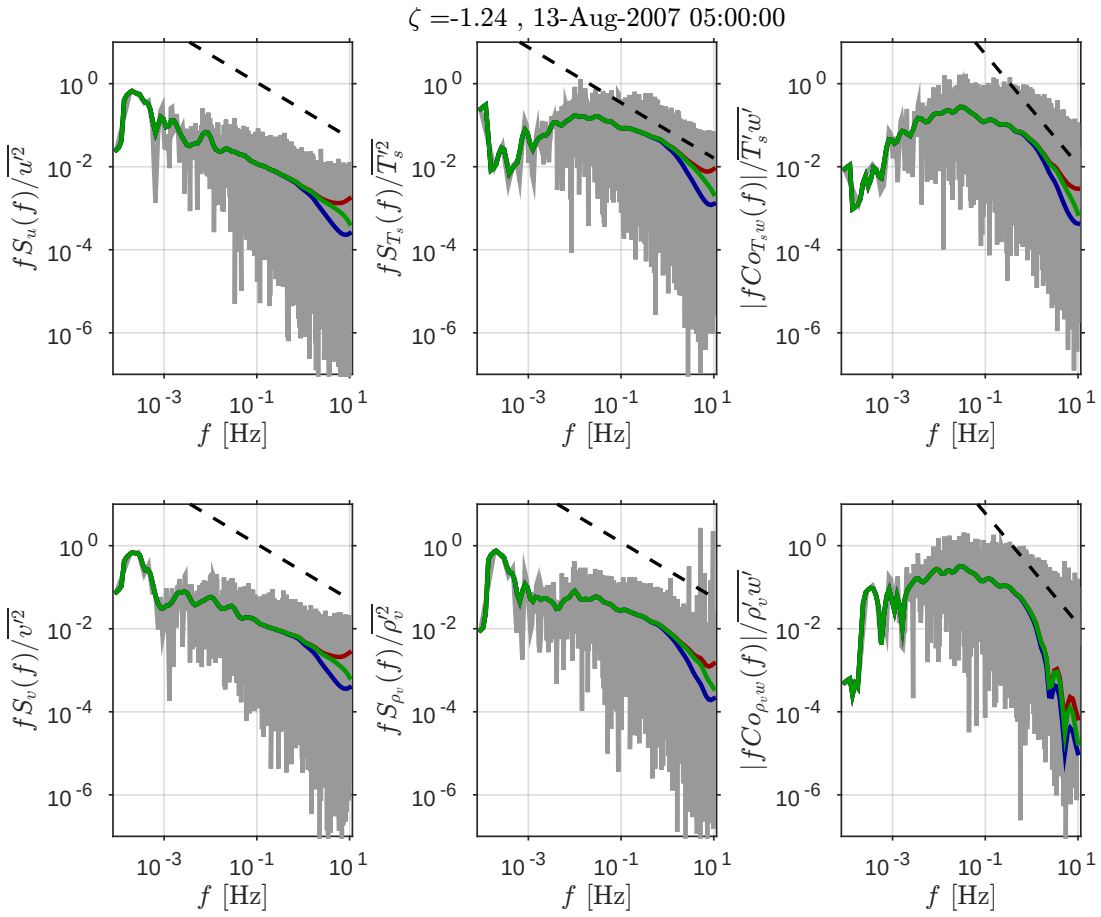


Figure 2.22: *Log-log representation of various normalized estimated cospectra and spectra as functions of natural frequency f [Hz] based on conditioned sampled time series in the 3 hour segment on the* 13.08.2007 *under unstable stratification. In each panel the grey lines are the unsmoothed and aliased estimates, the red lines are the smoothed aliased estimates, the blue lines the smoothed and dealiased (after Kaimal and Finnigan (1994)) estimates and the green lines are the smoothed and dealised (after Gobbi* et al. *(2006)) estimates as output by the module. The expected slope of the inertial subrange in each panel, $-2/3$ for normalized spectra and $-4/3$ for normalized cospectra, is represented by the dashed black line is included for orientation.*

Before finishing of the section it is worth emphasizing the (co) spectral estimates output by the module. We will drop the clumsy $\tilde{}$ notation leaving smoothing and dealising implicit wherever it is applied. The outputs are for the domain dictated by the logarithimically evenly spaced array $f_{C,i}$ with $i \in 1(1)90$ as given by (2.88). All the (co)spectral estimates are computed for conditioned overlapping 3 hour segments advanced one block average (30 minutes) at a time where only segments with $\leq 10\%$ spikes, along with associated timestamps, are included in the output. In total 3×10^3 different active 3 hour segments are output by the module based on the available raw data that pass the QC. The output Ogive distributions are $Og_{\rho_v w}, Og_{T_s w}, Og_{uw}, Og_{vw}$ based on the cumulative sum over cospectral intensities and $Og_{\rho_v \rho_v}, Og_{T_s T_s}, Og_{uu}, Og_{vv}, Og_{ww}$ based on the cumulative sum over spectral intensities. The output spectral densities are the dealiased and smoothed estimates S_u, S_v, S_w, S_{T_s} and S_{ρ_v}. Finally, the output cospectral densities are the dealised and smoothed estimates $Co_{uw}, Co_{vw}, Co_{T_s w}$ and $Co_{\rho_v w}$. Example spectra and cospectra produced by the module are shown by the green curves in Figure 2.22 along with their smoothed aliased counerparts (blue curves) and unsmoothed aliased counterparts (gray curves). Uiltimately (co)spectra allow us to identify important features of turbulence that are invisible from a time series, such as the scale of the eddies that contribute the most towards the (co)variance and the location and extent of inertial subranges (Kaimal and Finnigan, 1994). In addition through Ogive analysis (Section 3.4) it is possible to verify whether or not fluxes converge within the standard 30 minute averaging period (Foken et al., 2006).

Chapter 3

Results & Discussion

In the ensuing chapter we will present some of the results obtained after processing the Bayelva EC data through our module. The complete module output is not presented for two reasons. Firstly, around 100 different variable arrays are output by the module, and we simply do not have the space or time to discuss all of these. Secondly, we wish to focus on on as of yet undiscussed results and not restate the findings of Westermann et al. (2009) and Lüers et al. (2014). These novel results involve: 1) Block quality flags; 2) Flux uncertainties; 3) An intercomparison of the fluxes produced by our module and TK2; 4) Ogive analysis; 5) Validating model cospectra against observed cospectra. These results will be discussed in unison with their presentation as we find such an approach to be the most natural.

3.1 Quality Flags

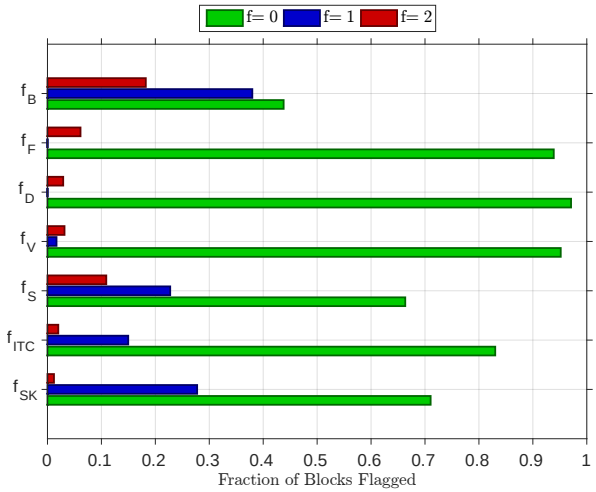


Figure 3.1: *Blocks flagged, as a fraction of* 7888 *which is the total number of 30-minute averaging blocks available, by the various flagging categories in the QC. Each category is displayed as a red-blue-green bar triplet. For each triplet there is one bar for the fraction of each integer quality flag: Highest quality flag ($f = 0$) fractions in green, medium quality ('soft') flag ($f = 1$) fractions in blue and finally poor quality ('hard') flag ($f = 2$) fractions in red. The flagging categories represented by each triplet are all defined in Section 2.6.*

Herein we will present and discuss the results of the extensive QC flagging procedure outlined in Section 2.6. The result of the entire procedure is summarized graphically in Figure 3.1. Recall that the combined (30 minute) block quality flag, f_B, takes the maximum value of all the individual flagging categories for a given block and is ultimately what we use to hard flag and discard blocks from further analysis. Thereby, it is encouraging that the sum of the category hard flagged fractions exceeds the combined block quality hard flag fraction by 0.08 indicating that there is considerable overlap in the category hard flags. This is not surprising, for example, both the w-ITC flag, being based on MOST, and the stationarity flag are likely to hard flag many of the same highly transient blocks.

It is also clear from Figure 3.1 that the stationarity flag is by far the strictest category operating across all stabilities (as opposed to the ITC flag) and hard flagging 11% of the blocks. For comparison the category with the closest hard flag fraction, the faulty flag, only flagged around half this number of blocks. Despite this the fraction of blocks hard flagged based on the relative nonstationarity test is, as previously discussed, in general agreement with other investigations at Bayelva (Westermann et al. (2009) and Lüers et al. (2014)).

On the other side of the scale is the skewness-kurtosis flag which is the least severe in terms of its hard flag fraction, and is not included in the TK2 (or TK3) QC procedures. Nonetheless, it was through a late implementation of this flag that we were able to systematically hard flag clearly unphysical fluxes (absolute sensible heat fluxes in excess of $600\mathrm{Wm}^{-2}$) that passed the remaining quality controls. Furthermore the flow distortion hard flag, occurring when the wind direction

is in the closed sector, occurred for $\simeq 3\%$ of the blocks for our batch of data which is in good agreement with the results of Westermann et al. (2009) based on a different period (March 2008-March 2009). The vertical velocity hard flag, blocks where $|\overline{w}| \geq 0.15$, also occurred for $\simeq 3\%$ of the blocks. Such an infrequent occurrence is to be expected after having successfully applied the planar fit of Wilczak et al. (2001). As with the skewness-kurtosis flag, the vertical velocity flag is crucial in isolating abnormal periods, in this case with strong advection, that are potentially not picked up by the remaining categories.

Finally the faulty flag, being based on the plausibility and MAD tests identifying spikes, hard flagged around 6% of the blocks. Most of these instances were flagged for implausibility due to the absolute humidity reported by the LI-7500 drifting to unphysical negative values. For the remainder the MAD test proved an invaluable tool for identifying isolated spikes amongst all the measured variables, allowing these to be ignored both in the block averaged statistics and the (co)spectral analysis. As for the portion of blocks flagged by the combined block quality flag: 19% were discarded having been hard flagged as poor quality, 38% received a soft flag as medium quality and the remaining 43% were of high quality. This is in general agreement with the flagging portions outlined in other EC investigations such as Nordbo et al. (2012) and Mauder et al. (2013).

3.2 Relative Flux Uncertainties

In the following we will present the result of the flux sampling uncertainty analysis for the highest quality ($f_B = 0$) sensible and latent heat fluxes produced by the module. Recall that both the fluxes and their relative uncertainties are 30 minute block averages. As outlined in Section 2.6.7, we arrive at the relative flux sampling uncertainties through (2.69) and (2.70) using a two step approach. The method of Finkelstein and Sims (2001) is used to estimate the variances of covariances (Appendix A.5) and subsequently we follow Billesbach (2011) by accounting for the propagation of uncertainty (Appendix A.6) through flux corrections. We re-emphasize that the uncertainty is considered to be solely due to the random stochastic error associated with the inadequate sampling of the flux within an averaging period discussed in Lenschow et al. (1994) and Mauder et al. (2013). So other sources of error, including systematic error, are not estimated.

It is also worthwhile discussing the following: 1) Are the relative uncertainties considerable? 2) Does the relative uncertainty depend on the magnitude of the flux itself? 3) Is there a dependence on stability? To questions 2) and 3) Finkelstein and Sims (2001) concluded that the answer was no, although the former acknowledged that only relatively large fluxes with magnitudes in excess of 10 Wm^{-2} were considered. We will see if our results support the conclusions of the former.

In Figure 3.2 the estimated block relative uncertainties for the respective fluxes are plotted as functions both of flux magnitude (upper panels) and stability (lower panels). Not included in the Figure are blocks with relative uncertainties of at least unity. For the latent/sensible heat flux such blocks occurred only $5\%/7\%$ of the time; the average of the absolute fluxes for these blocks was $2\,\mathrm{Wm}^{-2}/2\,\mathrm{Wm}^{-2}$ compared to $12\,\mathrm{Wm}^{-2}/19\,\mathrm{Wm}^{-2}$ for all blocks. Thereby, blocks with relative uncertainties in excess of unity are generally associated with low absolute flux values. Mahrt (2010) points out that in low flux conditions where only a few sporadic events typically contribute to the flux, the estimation of uncertainty based on the variance of the covariance is itself quite uncertain. So, uncertainties for low flux values should be interpreted with this in mind, even if estimating the uncertainty of the uncertainty is beyond the scope of this work. To avoid possible confusion by a statement such as 'the uncertainty of the uncertainty is large' we do not mean that the uncertainty itself has a large value, but rather that its estimated value is in itself uncertain. In the upper panels of Figure 3.2 this 'uncertainty of the uncertainty' is evident from the relatively large scatter, as shown by the binned standard deviations, in the relative uncertainty for low absolute fluxes. Conversely for the uncertainties in larger absolute fluxes the binned standard deviation is nearly cut in half. As opposed to Finkelstein and Sims (2001) we still include uncertainty estimates for absolute fluxes down to $1\,\mathrm{Wm}^{-2}$ in the hope that these estimates still provide reasonable approximations of the actual uncertainty.

Continuing we still focus our attention to the upper panels of Figure 3.2 when discussing the question of whether or not the relative uncertainties are dependent on the magnitude of the flux. As noted, Finkelstein and Sims (2001) found that the absolute uncertainty was proportional to the absolute flux implying that the relative uncertainty is independent of the absolute flux magnitude. It can be seen that for both fluxes the bin averaged relative uncertainties are near constant, on the order 0.1, for absolute fluxes in excess of $20\,\mathrm{Wm}^{-2}$. So, at least for the larger flux estimates our results are in agreement with those of Finkelstein and Sims (2001). If we instead start at the lowest absolute flux, then for

[1] A straight line with negative slope in a linear-logarithmic representation is the equivalent of logarithmic decay.

both latent and sensible heat flux the bin averaged relative uncertainty decays logarithmically[1] for increasing absolute flux values up to around $20\,\mathrm{Wm}^{-2}$. Thus, for low absolute fluxes the relative flux uncertainty is apparently not independent of the magnitude of the flux, although as mentioned the scatter is considerably greater.

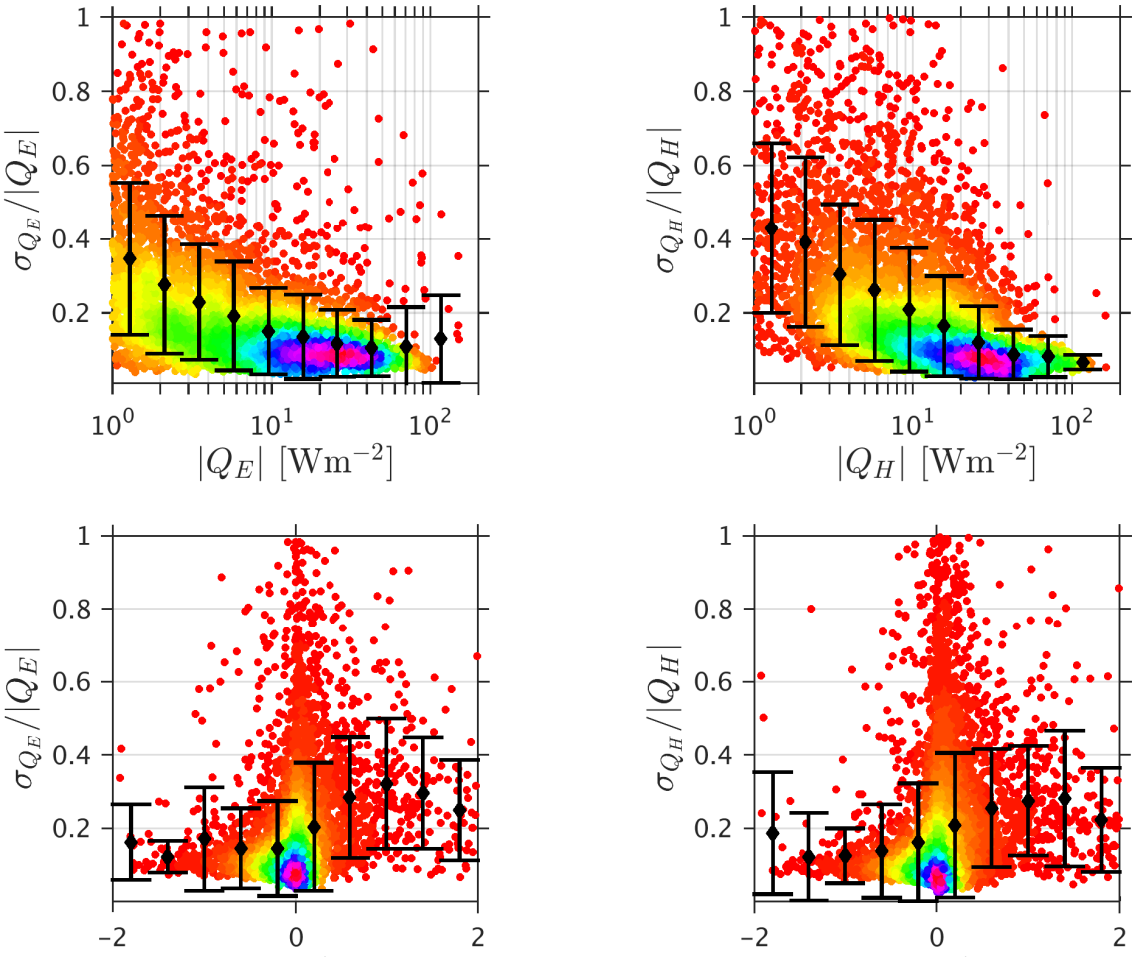


Figure 3.2: *Relative flux uncertainties for the latent heat flux ($\sigma_{Q_E}/|Q_E|$, left panels) and sensible heat flux ($\sigma_{Q_H}/|Q_H|$, right panels). Only fluxes with the high quality combined block flag ($f_B = 0$) are considered. The relative flux uncertainties are displayed as functions of the logarithmically scaled absolute value of the corresponding heat flux (top panels) and the M-O stability parameter ζ (bottom panels) on a linear scale. Pinks indicate the highest density of points whereas reds are the scarcest densities. The black 'error bars' show the mean (diamond) plus minus one standard deviation (horizontal bars) of relative flux uncertainty for 10 evenly spaced (logarithmically top panels, linearly bottom panels) bins of absolute flux (top panels) and ζ (bottom panels). For visualization purposes only relative flux uncertainties for absolute fluxes $\geq 1\,Wm^{-2}$ and stabilities $|\zeta| < 2$ are included.*

Next, we home in on the lower panels in Figure 3.2 that show the dependence of relative uncertainty on stability, ζ, for the respective fluxes. As in the upper panels there is considerable scatter in the uncertainty estimates. Using the binned standard deviations as a guide there does not seem to be a systematic dependence of scatter on stability, although the scatter appears to be somewhat smaller overall for unstable stratification when compared to neutral and stable stratification. Large scatter for neutral stratification is expected as here the magnitude of fluxes is typically small, by definition the perfectly neutral surface layer has zero buoyancy heat flux (Stull, 1988), and as discussed there is considerable uncertainty in the relative uncertainty estimates for low fluxes.

What is striking is that the relative uncertainty is not symmetric about $\zeta = 0$. On the unstable side in both panels starting at near neutral stratification and moving towards increasingly unstable stratification there appears to be, if we ignore the most unstable bin, a decay towards a constant value of the bin averaged relative uncertainty. Conversely on the stable side in both panels, starting at near neutral stratification and moving towards increasingly stable stratification, the bin averaged relative uncertainty initially increases and then levels

off and decreases again for $\zeta > 1$. The asymmetry is especially clear for $|\zeta| < 1$ in both panels where the bin averaged relative uncertainty increases as we transition from the unstable to the stable side reaching around double the moderately unstable value at moderately stable stratification. This is in contrast to the results of Finkelstein and Sims (2001) where no clear stability dependence was found. We are somewhat cautious in claiming that Figure 3.2 provides direct evidence of the relative flux uncertainty having a stability dependence due to the large scatter. Nonetheless, Finkelstein and Sims (2001) note that small nocturnal fluxes in stable stratification where the contributing eddies are sampled infrequently undoubtedly lead to a large relative flux uncertainty. Moreover the former point out that the exclusion of small fluxes in their analysis probably leads to a negative bias (underestimation) of their average relative flux uncertainty estimates which could explain why they found no stability dependence.

$\sigma_{Q_H}/|Q_H|$

Stability	Mean	Standard Deviation	CV	Number of Blocks	Fraction < 0.25
Neutral $\|\zeta\| \leq 0.1$	0.173	0.177	1.023	2542	0.80
Unstable $\zeta < -0.1$	0.138	0.125	0.906	1493	0.92
Stable $\zeta > 0.1$	0.272	0.190	0.700	1853	0.56
All	0.195	0.178	0.913	5888	0.76

$\sigma_{Q_E}/|Q_E|$

Stability	Mean	Standard Deviation	CV	Number of Blocks	Fraction < 0.25
Neutral $\|\zeta\| \leq 0.1$	0.168	0.155	0.923	2427	0.82
Unstable $\zeta < -0.1$	0.146	0.121	0.829	1457	0.90
Stable $\zeta > 0.1$	0.275	0.177	0.644	1219	0.54
All	0.187	0.160	0.856	5103	0.78

Table 3.1: *Simple statistics for the relative uncertainties of the sensible heat flux ($\sigma_{Q_H}/|Q_H|$, top panel) and the latent heat flux ($\sigma_{Q_E}/|Q_E|$, bottom panel) for four stability categories at Bayelva. In order (left to right) the columns represent the category mean, standard deviation, coefficient of variation (CV), number of blocks and the fraction of blocks within the category where the relative uncertainty is < 0.25. These statistics are based on the results presented in Figure 3.2.*

One question remains to be discussed, namely: 1) Are the relative flux uncertainties considerable? To answer this question in a simple manner we split the relative uncertainties into four stability categories and consider simple statistics for each of these. The result is presented in Table 3.2 from which it is evident that for the respective categories the relative uncertainties are quite similar for the sensible and latent heat flux. For both fluxes the mean relative uncertainty is largest in the stable category, smaller in the neutral category and smaller still in the unstable category at just over half the stable value. The same is true for the standard deviation in the relative uncertainty, although the difference between the stable and the unstable categories is not as large. Subsequently the same holds for the coefficient of variation (ratio of standard deviation to the mean) which is a measure of the relative scatter of the relative uncertainties in each category. The CV is large across the board being near or above unity. It is somewhat encouraging that the lowest, albeit still large, CV occurs for stable stratification; indicating that the stable uncertainties are less scattered than the other two categories. As for the number of blocks in each category, which are different for the two fluxes due to the exclusion of absolute fluxes that are less than unity, the neutral category is the most frequent, while there is a near even split between the number of blocks in the unstable and stable categories. Finally for the fraction of blocks where the relative uncertainty is below 0.25, for both fluxes this fraction is largest for unstable stratification followed by neutral stratification and finally stable stratification.

Overall we have that the relative flux uncertainties estimated for Bayelva are on average around 0.2 for both sensible (0.195) and latent heat flux (0.187). This is in line with (give or take 0.1) the averaged results of Finkelstein and Sims (2001) and Mauder et al. (2013) in which multiple EC sites were considered as well as Billesbach (2011) and Nordbo et al. (2012) both of which considered single EC sites. We find such an agreement encouraging in that we expect the long term average relative flux uncertainties to be around the same magnitude regardless of the location of an EC site. This is because the positive definite block flux uncertainty estimates related to stochastic error should 'average out' to a constant value over a long enough period (e.g. one year). As pointed out in Mauder et al. (2013), flux sampling uncertainty estimates are useful in that these provide modelers with quantitative estimates of the precision of observed parameters used in data assimilation. To our knowledge no analysis such as ours, in particular accounting for uncertainty propagation, has previ-

ously been undertaken for an Arctic EC site. Thereby we hope that the results in Table 3.1 can provide a useful reference to anyone wishing to assimilate EC data from similar sites into a model, particularly coupled land-atmosphere models. Moreover the uncertainty estimates could be useful when comparing the modeled and observed surface energy balance as is done in e.g. Aas et al. (2015). To finish off the section, whether or not the relative uncertainties are considerable will depend on the context in which the fluxes are used and the desired precision.

3.3 Module & TK2 Intercomparison

Primarily as a sanity check for our module output, we compared our flux estimates to those produced by the TK2 package (Mauder and Foken, 2004). The results of TK2 were available to us both for the period presented in Westermann et al. (2009), discussed in Section 1.3.3, as well as a pilot run with the same settings for the earlier period of 02.04.2007 to 14.09.2007 that we are studying. For the comparison it is of course the pilot run, with flux estimates that are synchronized with ours, that is relevant. As is typical for EC software intercomparisons, e.g. Mauder et al. (2008) and Fratini and Mauder (2014), we present the result in the form of a linear regression as depicted in Figure 3.3.

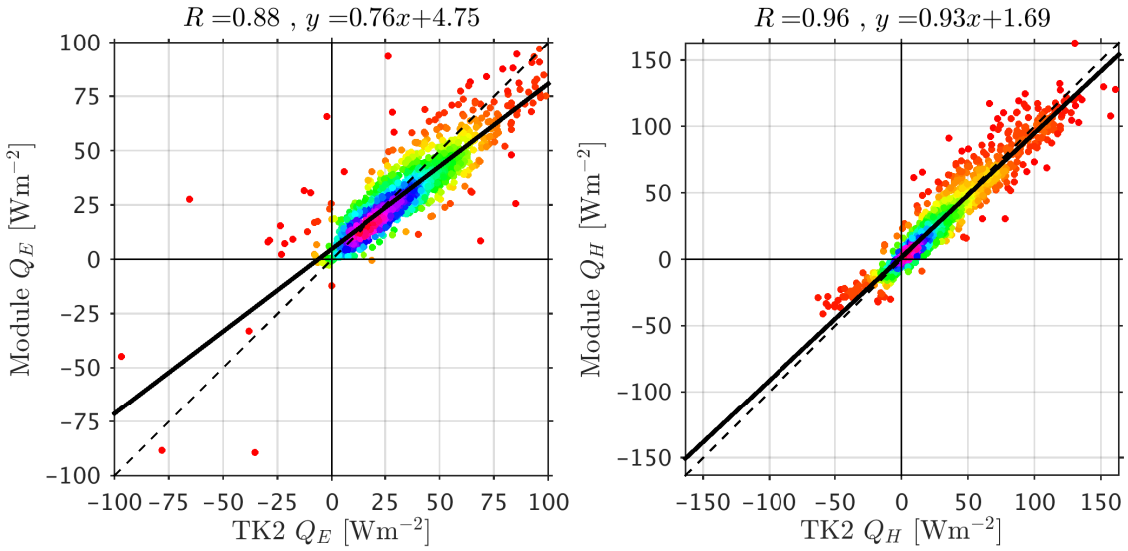


Figure 3.3: *Synchronous 30 minute block averaged flux estimates produced by TK2 (abscissa) and by our module (ordinate) in the period 02.04.2007 to 14.09.2007 at Bayelva. Individual flux estimates are displayed as colored points where pinks (reds) indicate the densest (sparsest) concentration of points. In each panel the dashed black line is the one to one line while the thick solid black line is the result of a linear regression, with the equation and linear correlation in the corresponding panel title. Only fluxes with the highest quality flag in both packages are included: 1315 latent heat fluxes (Q_E) in the left panel and 1567 sensible heat fluxes (Q_H) estimates in the right panel.*

We recall from Section 1.3.3 that the TK2 runs when combined with available energy (global radiation less the ground heat flux) measurements resulted in an average surface energy *imbalance* in excess of 50% and an energy balance ratio on the same order for hourly fluxes in the period March 2008 - March 2009. It is unlikely that the available energy estimates were significantly overestimated (Foken, 2008a), even with storage unaccounted for. We can make such a claim because even the diurnal average flux estimates still had a low (cf. Wilson et al. (2002)) energy balance ratio (74%). This corroborates the widely held view (e.g. Foken (2008a) and Leuning et al. (2012)) that the EC method tends to underestimate the magnitude of turbulent fluxes.

Why are we repeating ourselves? Well, the above discussion implies that our module would outperform TK2 in the case that our flux estimates are greater in magnitude. Clearly, from the linear regression in Figure 3.3 this is not the case, on average our module has somewhat lower absolute flux estimates than TK2. Even so, the estimates from TK2 and our module share a relatively high linear correlation, $R = 0.88$ and $R = 0.96$ for the latent and sensible heat flux, respectively, which is encouraging. Furthermore, our sensible heat flux estimates are on average quite close to the TK2 estimates, underestimating these by only 7%. In fact, the small difference in sensible heat flux estimates and high correlation are comparable to those found in other EC package intercomparissons for open path systems such as Fratini and Mauder (2014). For the latent heat flux, the agreement is not as good with our module

on average underestimating the TK2 estimates by 24%. With such a high linear correlation, albeit lower than for the sensible heat flux, this implies that the disagreement is fairly systematic. We will get back to possible explanations for this discrepancy. It was somewhat unfortunate that available energy measurements were not available for the majority of the period we considered in our study. Despite this, the performance of the module relative to TK2 provides an indication of how well we closed the surface energy balance. In that we underestimated the fluxes relative to TK2, the average energy in imbalance in our estimates is likely to also be in excess of 50%. Of course, it was not our goal to close the energy balance or reproduce TK2 estimated. Instead we wished to go beyond standard EC output and investigate the reasons for a lack of closure associated with deficiencies in the EC method under difficult conditions that can arise under stable stratification.

Processing Step	TK2	Module
Despiking	• Standard deviation test (Vickers and Mahrt, 1997).	• Median absolute deviation test (Mauder et al., 2013). • Skewness-Kurtosis test (Vickers and Mahrt, 1997).
Rotations	• Double rotation (Kaimal and Finnigan, 1994) ($\overline{w} = \overline{v} = 0$).	• Planar fit (Wilczak et al., 2001) ($< \overline{w} >= 0$). • Natural ensemble streamline frame ($\overline{v} = 0$, $\overline{w} \neq 0$).
Flux Corrections	• Integral cospectral attenuation (Moore, 1986).	• Analytical cospectral attenuation (Massman, 2000).
Quality Controls		• Integral turbulence characteristics test for w (Thomas and Foken, 2002). • Estimation of flux uncertainty (Billesbach, 2011).

Table 3.2: *The differences in the EC processing steps between the TK2 package (Mauder and Foken, 2004), as run by Westermann* et al. *(2009), and our module runs. Any step not listed was employed in both the TK2 runs and our module runs.*

To isolate the reason for the systematic underestimation of latent heat flux in our module when compared to TK2, it is instructive to contrast the two processing packages. All the differences between the two are listed in Table 3.2. Although there are quite a few, some of these have negligible effect on the flux. Most of those that do affect the latent heat flux should have the same effect on the sensible heat flux on which there is better agreement between the two packages. This leads us to the conclusion that the discrepancy in latent heat flux is due to the cospectral attenuation corrections. The reason being that this is the only processing step that is both performed differently in the two packages and corrects the latent and sensible heat flux separately using unique system transfer functions (cf. Section 2.5.2). Even though the attenuation correction factor for latent heat flux that we calculate through the method of Massman (2000) is larger than that for sensible heat flux (Figure 2.15), it is probable that we are still underestimating the attenuation of the latent heat flux. Visual evidence of the relatively strong high frequency cospectral attenuation suffered by the latent heat flux is provided in Figure 3.6, and it is unlikely that correction is large enough to systematically account for this. In TK2 the integral method of Moore (1986) is applied when correcting for attenuation, in which the lateral sensor separation is accounted for. Both our module and TK2 reconcile attenuation due to longitudinal separation, i.e. the component of the separation vector in the direction of the mean wind, through the maximum cross-correlation technique (Section 2.5.1). Therefore, it is possible that the attenuation resulting from lateral sensor separation, due to eddies moving across the mean wind direction, could play a more important role then we originally thought. Applying a lateral separation transfer function in the method of Massman (2000) is a possible solution, but unfortunately we did not have the time to implement and test its effect. Another reason that the method of Moore (1986) could provide a greater correction is that more realistic cospectral mod-

els are considered (see Mauder and Foken (2004)) than the simplified form proposed by Horst (1997) that we used in Section 2.5.2. We emphasize that despite our apparent underestimation of cospectral attenuation, the effect on relative flux uncertainties is negligible, while there is no effect at all on the observed cospectra (and Ogives) that are produced by our module. In particular, the produced cospectra are *not* in any way corrected for attenuation as it is the instantaneous values, as opposed to corrected fluxes, that are used in their calculation.

3.4 Ogive Analysis

In the following we will investigate whether or not our fluxes converge within the traditional 30 minute averaging period used in our module. This is the averaging period typically employed in EC campaigns (Lee et al., 2006); including the work of both Westermann et al. (2009) and Lüers et al. (2014). Such an investigation is possible through Ogive analysis first applied to EC studies by Desjardins et al. (1989). Our method for producing Ogives is outlined in Section 2.8. As previously discussed the Ogive, the cumulative cospectrum, is a measure of the cumulative covariance (flux) contained within frequencies (periods) greater than (lower than) the frequency (period) of the Ogive. Only the turbulent fluxes involved in the SEB are considered herein. Ogive analysis, despite traditionally not being a standard part of the EC method, has recently been applied successfully in a range of EC studies, e.g. Foken et al. (2006), Nordbo et al. (2012), Charuchittipan et al. (2014) and Sievers et al. (2015).

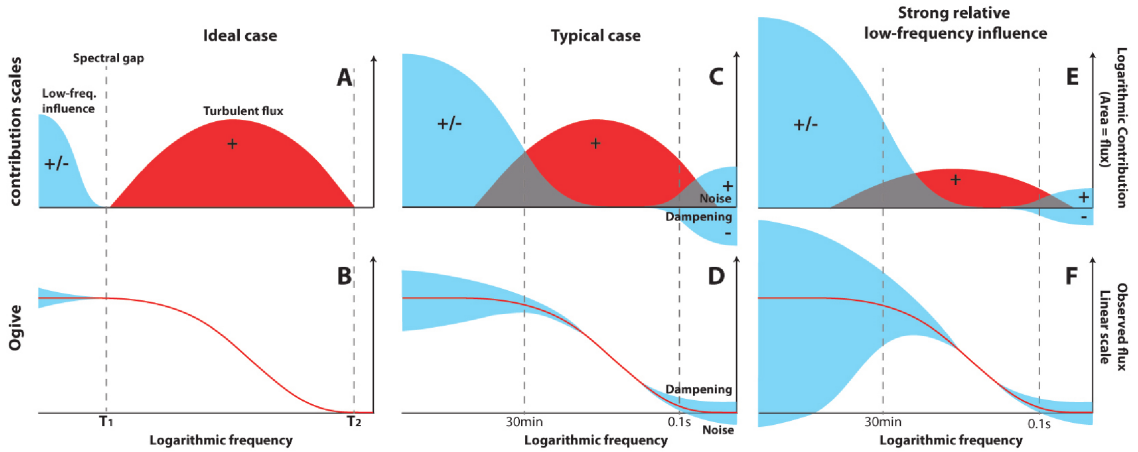


Figure 3.4: *Schematic relationship between the Ogive (cumulative cospectrum) as a function of frequency (lin-log) in the lower panels (B,D,F) and the cospectral density as a function of frequency (log-log) in the corresponding upper panels (A,C,E). Non turbulent low frequency flux contributions (e.g. drainage flows or gravity waves) and high frequency instrumental noise (vibrations) and dampening (attenuation) flux contributions are shown in blue while true turbulent flux contributions are shown in red. Left panels (A,B): The ideal case with a clear spectral gap separating surface layer turbulence from low frequency influence, with ideal instrumental response. Middle panels (C,D): Typical case with a less distinct cospectral gap and some instrumental problems. Right panels (E,F): Relatively strong non turbulent low frequency contribution and some instrumental problems. Note that the low frequency contribution can be of opposite sign to the turbulent flux contribution. In each panel dashed vertical lines correspond to a 30 minute averaging time (left line) and a Nyquist frequency of* 10 *Hz (right line). Figure adopted from (Sievers* et al.*, 2015).*

Ogive analysis allows us to verify if our fluxes converge to an asymptote within our averaging period (Charuchittipan et al., 2014). If the flux does converge at a frequency whose inverse is our averaging period (left panel in Figure 3.4) then this is evidence of a clearly defined cospectral gap (Sievers et al., 2015). The cospectral gap represents a minimum in the cospectrum that marks a clear separation between turbulent contributions to the flux from non turbulent low frequency (NTLF) contributions such as advection (Stull, 1988). Moreover, convergence implies not only that such a gap exists but that the high pass filtering operation of removing the block average successfully attenuates all low frequency influences beyond this gap and provides a faithful representation of the 'true' ensemble averaged turbulent flux (Wyngaard, 2010). Crucially, Ogive analysis is a powerful tool for identifying departures from the 'idealized' surface layer assumed by the EC method and required for estimating the SEB, NEE as well as providing universal forms for MOST (Section 1.3). As noted in Foken (2008a) the low frequency contributions that lead to an Ogive diverging at lower frequencies

(bottom right panel in Figure 3.4) may not only be due to non turbulent mechanisms such as drainage flows and gravity waves, but also unresolved slowly moving larger scale turbulence in the form of 'organized turbulent structures'. Thereby, the cospectral gap may not be as well defined as in the classical theory presented in e.g. Stull (1988); there is typically a considerable overlap between 'turbulent' and 'non-turbulent' contributions to the Ogive (Sievers et al., 2015). So, although the Ogive can tell us if our flux converges in an averaging period, if the Ogive does *not* converge there exists no accepted method for separating the unresolved low frequency turbulent and non turbulent contributions to the flux from one another (Sievers et al., 2015). Hence the Ogive is useful for identifying periods where the EC method works, but correcting the flux in the case that the Ogive does not converge may be far from trivial. Nonetheless, in the case that the Ogive clearly converges at lower frequencies then it is possible to correct the flux value by simply setting it to the Ogive (cumulative flux) at these low frequencies (Charuchittipan et al., 2014). Here our goal is not to correct the flux but to present a classification of the 2790 Ogives, for both sensible and latent heat flux, that we calculated for the Bayelva site. In particular we are interested in addressing the question of: How often do the fluxes converge within 30 minutes and is there a clear stability dependence on the fractional rate of Ogive convergence?

Before proceeding we define our normalized Ogives, denoted $\widehat{Og}_{\xi w}$, as the Ogive normalized by its maximum or minimum, whichever is greater in magnitude, i.e.

$$\widehat{Og}_{\xi w}(f) = \frac{Og_{\xi w}(f)}{\text{SAM}} \,. \tag{3.1}$$

Where the expression[2]

$$\text{SAM} = \text{sgn}\left\{\max(Og_{\xi w}) + \min(Og_{\xi w})\right\} \max\left(|Og_{\xi w}|\right) \,, \tag{3.2}$$

returns the sign of the Ogive at the frequency where the absolute value of the Ogive corresponds to its maximum absolute value. So, at the frequency where the Ogive peaks (in the absolute sense) the value of the normalized Ogive is unity. At the same time the normalized Ogive can take either sign away from its peak just like the Ogive itself. This is a more instructive (albeit cumbersome) definition for our purposes then that in Foken et al. (2006) where a single signed normalized Ogive is defined as the absolute Ogive normalized by its maximum. Such a definition will lead to confusion and missidentification of the Ogive cases to be described. For example (cf. Figure 3.4), in the presence of a SAM the mid frequency range combined with a relatively strong but oppositely signed NTLF the Ogive may take a maximum in the mid frequency range and then change sign and approach the same magnitude at the low frequency end. With our definition it is possible to identify such instances. Conversely, with the definition of Foken et al. (2006) these instances may unwittingly be classed as convergent Ogives even though there is relatively strong (and oppositely signed) NTLF influence. Sievers et al. (2015) point out that the occurrence of such 'reversal' cases can apparently not be explained by any existing theory. In particular these cases challenge the idea that the significant contributions towards the flux are always of the same sign regardless of eddy size.

As noted Secion 2.8 we verified that the values of all of our calculated 3 hour Ogives at the first harmonic (frequency whose inverse is 3 hours) matched the corresponding (conditioned) 3 hour segment covariances (or variances depending on the Ogive). Thereby, thanks to this control, we are fairly confident that no error was made in the Ogive calculation as the calculated value corresponds to that given by definition at the first harmonic (Stull, 1988). Subsequently we truncated our 3 hour calculated Ogives by removing frequencies corresponding to periods beyond 1 hour. We do so because these longer periods have few cycles within our 3 hour segments and are (apart from the first harmonic) too poorly resolved to be representative (Kaimal and Finnigan, 1994). Furthermore it is the Ogive value at 1 hour that we compared to that at our averaging period of 30 minutes to see if the Ogive converges. Hence the truncation is not problematic as we are only interested in frequencies $f^{-1} \leq 1\,\text{hour}$. So, when referring to Ogives, as well as their SAM, from here on it is understood that these only include periods up to 1 hour.

Next, in part following Foken et al. (2006), we split the behavior of the Ogive into one of four cases: 1) Convergent, 2) Extremum, 3) Divergent and 4) Reversed. We begin with fourth case, not considered in Foken et al. (2006), this is the reversed Ogive that we classify as any Ogive that satisfies

$$\widehat{Og}_{\xi w}(f) \leq -0.1 \,.$$

So a reversed Ogive is one for which there exists a frequency at which the Ogive value is in magnitude more than one tenth of the SAM (see (3.2)), but of the opposite sign. It is this relatively rare but important case that we use to identify segments with relatively strong, but oppositely signed, low frequency contributions (cf. Figure 3.4 with respect to the turbulent contribution (or vice versa). As mentioned without defining the normalized Ogive through (3.2) we would not be able to identify these cases and as a worst case risk them being falsely diagnosed as convergent Ogives. We use the term 'reversed' because these Ogives are classified as those that undergo a relatively large sign reversal. The threshold value of -0.1 as opposed to a value closer to 0 was chosen to avoid attributing Ogives where some high frequency instrument noise (due to e.g. vibrations) is present, unrelated to low frequency sign changes

[2] sgn is the signum function that returns $+1$ if its argument is positive, -1 if its argument is negative and 0 if its argument is zero.

characteristic of the 'reversal' case.

Now, the ensuing cases 1)-3) only apply to Ogives that have not been classified as case 4) (reversal). Case 1), the ideal case in Figure 3.4, is the convergent Ogive which we classify as any Ogive that satisfies

$$\widehat{Og}_{\xi w}\left(f^{-1} = 60\,\text{mins}\right) > 0.9\,,$$

and

$$\widehat{Og}_{\xi w}\left(f^{-1} = 30\,\text{mins}\right) > 0.9\,.$$

So in words for an Ogive to be convergent its value at both half an hour and one hour be within 10% of both its SAM. Case 2), shown by the lower blue Ogive curves in the bottom left panel of Figure 3.4, is the 'extremum' Ogive with a high frequency (low period) SAM that satisfies

$$\widehat{Og}_{\xi w}\left(f^{-1} = 60^{-1}\,\text{mins}\right) \leq 0.9\,.$$

Which is indicative of a relatively large SAM occurring at a higher frequencies. Finally case 3), shown by the upper blue Ogive curves in the bottom left panel of Figure 3.4, is the divergent Ogive which classify as an Ogive satisfies

$$\widehat{Og}_{\xi w}\left(f^{-1} = 60\,\text{mins}\right) > 0.9\,,$$

and

$$\widehat{Og}_{\xi w}(f^{-1} = 30\,\text{mins}) \leq 0.9\,.$$

Thus for the divergent Ogive the hourly value is greater than a factor 0.9 of the SAM (with the same *sign*) whereas the half hourly value is less than such a factor of the SAM. Indicating that beyond the half hourly time scale, which we are block averaging over, the flux continues to increase significantly in magnitude.

We performed the classification of the Ogive cases for a number of generalized stability categories: all stabilities (any ζ value), neutral stability ($|\zeta| \leq 0.1$), unstable stratification ($\zeta < -0.1$), and stable stratification ($\zeta > 0.1$). These stability categories are instructive in their simplicity, although they do not capture variations within a regime such as the transition from moderately to very stable stratification (Mahrt, 1999). In Figure 3.5 the Ogives of the sensible and latent heat fluxes, across all stabilities, are classified as case 1), 2) and 3) each shown in separate panels. Case 4) is not included owing to its relatively infrequent occurrence. Due to the large number of Ogives (gray curves) the shape of individual normalized Ogives is not clear in Figure 3.5, although the size of the gray area provides an indication of the overall spread within the Ogive cases. Moreover, the blue curves show the median, thus providing a visualization of the typical normalized Ogive shape for each case.

Since we have considered thousands of Ogives it is more instructive to present the result of the analysis in terms of case-wise fractions for each stability category. We provide such a presentation in Table 3.3 which is a convenient reference for the typical performance of a 30 minute averaging period across the stability categories at Bayelva. Furthermore, the results given in this Table are also applicable to other locations in the high Arctic, as well as similar tundra regions worldwide. Overall the classical 30 minute averaging period performs quite well for both the sensible heat flux and the latent heat flux with both Ogives converging $\sim 70\%$ of the time. The only other high Arctic Ogive analysis that we know of was conducted by Sievers et al. (2015) for a site on Greenland, however, the aim of their analysis was different from ours and no case fractions were given. So, to put our results in perspective we make due with one example campaign at lower latitude, namely LIFTASS-2003 (Mauder et al., 2006). Based on May-June data from a single site included in LIFTASS-2003 campaign Foken et al. (2006) calculated ~ 100 latent and sensible heat flux Ogives each based on a 4 hour segment and found that these converged 83% of the time. Foken et al. (2006) defined convergence as any normalized absolute Ogive that was within 0.9 of its absolute maximum at 150 minutes and at 30 minutes, as discussed this is similar but not the same as our definition. Charuchittipan et al. (2014) recently extended the analysis of Foken et al. (2006) by including 6 of the LIFTASS-2003 sites, each with different surface type, for the same May-June period. Considering only the unforested sites (being more comparable with Bayelva), with the same convergence threshold (0.9) and excluding the site investigated by Foken et al. (2006) then Charuchittipan et al. (2014) found minimum convergence fractions of 88% and 90% for the sensible and latent heat flux Ogives respectively. As such, the Ogive convergence fractions at Bayelva are considerably lower than those found for LIFTASS-2003.

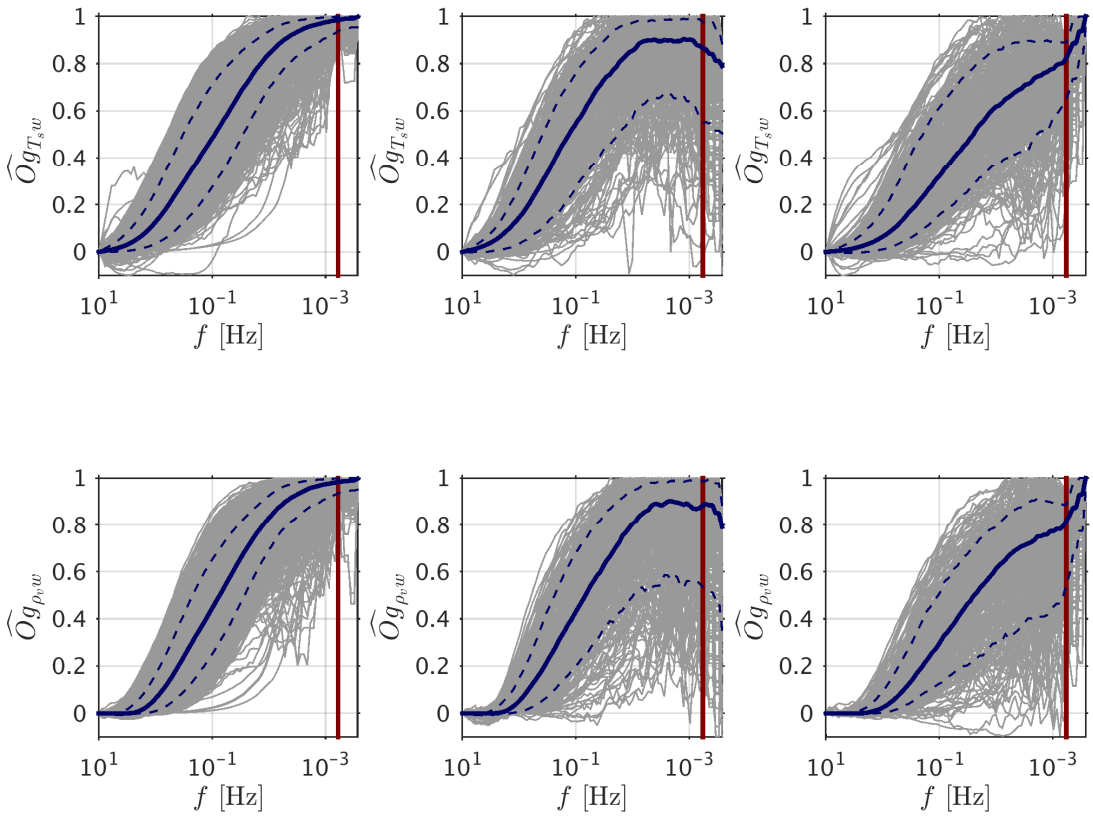


Figure 3.5: *A total of* 2790 *normalized Ogives (gray lines) of respectively sensible ($\widehat{Og}_{T_s w}$, top panels) and latent ($\widehat{Og}_{\rho_v w}$, bottom panels) heat flux as functions of natural frequency f (in Hz) where the frequency axis is reversed (from high to low) and logarithmic. For both the top and bottom panels: left, middle and right panel correspond to the convergent, extremum and divergent Ogive cases respectively. In each panel the case median normalized Ogive at each frequency is shown by the solid blue line, while the dashed blue lines depict the 10th (lower) and 90th (upper) percentiles. The vertical red solid line indicates the frequency whose inverse yields a period of 30 minutes and the inverse of the final (minimum) frequency yields a period of 1 hour.*

$Og_{T_s w}$				
Ogive Case	Neutral $\lvert\zeta\rvert \leq 0.1$	Unstable $\zeta < -0.1$	Stable $\zeta > 0.1$	All Stabilities
Case 1: Convergent	0.684	0.890	0.486	0.710
Case 2: Extremum	0.169	0.019	0.236	0.134
Case 3: Divergent	0.105	0.083	0.272	0.133
Case 4: Reversed	0.042	0.008	0.005	0.023
$Og_{\rho_v w}$				
Ogive Case	Neutral $\lvert\zeta\rvert \leq 0.1$	Unstable $\zeta < -0.1$	Stable $\zeta > 0.1$	All Stabilities
Case 1: Convergent	0.718	0.864	0.408	0.700
Case 2: Extremum	0.135	0.064	0.256	0.138
Case 3: Divergent	0.128	0.063	0.277	0.138
Case 4: Reversed	0.019	0.009	0.058	0.024

Table 3.3: *classification fractions of the Ogives (cumulative cospectra) for sensible ($Og_{T_s w}$, upper panel) and latent heat flux ($Og_{\rho_v w}$, lower panel) at Bayelva. Each row corresponds to one of three Ogive cases: convergent, extremum, divergent or reversed. Four stability categories are considered, one for each column. The numbers are the fraction of Ogives in a given category (column) that fall under the given classification (row). So the sum of the entries in a column is always* 1*. Each panel is based on a total of 2790 Ogives.*

Case stability fractions were not given in Foken et al. (2006) or Charuchittipan et al. (2014). As such we can only speculate as to the reasons, apart from the obvious difference in surface type, for the discrepancy between our convergence fractions and those of the former authors. Foken (2008a) noted that during the period they

considered the stability varied between a minimum of $\zeta = -0.5$ during the day up to a maximum of $\zeta = 0.5$ during the night. So, their Ogives likely include all the stability regimes we considered albeit no cases of very stable or very unstable stratification. Furthermore, as the period of LIFTASS-2003 considered by the former authors was during the early summer, which is the lightest part of the year for the sites considered, it is probable that unstable conditions occurred more frequently than stable ones. Our Bayelva data, on the other hand, contains a relatively large fraction of stably stratified Ogive segments. By inspection of Table 3.3 we note that for both the heat flux Ogives the convergent case was considerably less frequent only occurring 41% (for latent heat flux) and 49% (for sensible heat flux). Conversely during unstable stratification the Bayelva convergence fractions were high: 86% and 89% for, respectively, the latent and sensible heat flux Ogives. For neutral stratification the convergence fractions (70%) fell between the stable and unstable convergence fractions for both Ogives. As such, one possible reason for the discrepancy between our Ogive convergence fractions and those reported for LIFTASS-2003 is the fact that stable conditions, with low convergence fractions, were more prevalent in the Bayelva data.

In addressing why stable conditions have the lowest Ogive convergence fractions it is worth considering that stable cospectra peak at higher frequencies than their unstable and neutral counterparts (Kaimal et al., 1972). As mentioned in Section 2.5.2, the reason for the cospectral peaks shifting towards higher frequencies (lower periods) with increasingly stable stratification is, by definition of ζ, due to the increasing buoyant destruction of turbulence relative to the production from mean wind shear. In such a regime the large eddies present in unstable stratification can not exist (Wyngaard, 2010). Thereby, the cospectral gap also shifts towards higher frequencies. So, at least in theory, evaluating fluxes in stable conditions and ensuring convergence requires a *shorter* averaging period in that the integral time scale (inverse cospectral peak frequency) is much smaller than in unstable stratification (Kaimal and Finnigan, 1994). Moreover as noted in Sievers et al. (2015), the observed NTLF contributions towards the flux also increases. Recall that various mechanisms, outlined in Mahrt (1999), may lead to nonstationarity and horizontal heterogeneity of the flow in stable stratification. This departure from the idealized surface layer in turn expalins to increasing NTLF contributions towards the flux in stable stratification. Thus, by decreasing the averaging period we may also avoid some of these NTLF contributions. It turns out, however, that decreasing the averaging period in stable stratification is not generally a valid solution. This is because the significant decrease in averaging period needed to remove NTLF decreases the number of independent samples of turbulent flux contributions and leads to a large random error. In other words the averaging period would no longer be much greater than the integral time scale of the true turbulent cospectrum (Kaimal and Finnigan, 1994). In fact, as noted in Mahrt (2010) the estimates of uncertainty due to random error in cases with poorly sampled turbulent contributions are in themselves highly uncertain.

Finally we arrive at a major 'catch' with the EC method, noted in e.g. Foken and Wichura (1996) and Vickers and Mahrt (1997), namely that the averaging time must at the same time be: long enough to sufficiently sample the largest turbulent eddies and short enough to avoid nonstationary conditions that occur due to NTLF contributions (Sievers et al., 2015). So, if there is significant overlap in the cospectrum between the contributions from NTLFs and the large eddies we need to sample sufficiently, it is impossible to fulfill such a requirement on the averaging period. Note that this 'catch' can be problematic in any stability regime. Under unstable stratification the divergent Ogives are usually associated with instances where the largest eddies are not sufficiently sampled. In such cases extending the averaging period is not guaranteed to alleviate the problem as this may instead introduce significant nonstationarity in the form NTLF contributions. Conversely in stable stratification, as discussed, the problem is often that the NTLF contribution is very large relative to the turbulent contribution requiring a much shorter averaging time. If the averaging period is reduced accordingly, however, the true turbulent flux becomes inadequately sampled. In essence the problems with this 'catch' are related to the lack of a clearly defined spectral gap separating NTLF and turbulent contributions to the measured flux (Sievers et al., 2015). We conclude, based on convergence fractions in Table 3.3, that problems with this 'catch' can occur across all stabilities, but are most frequent in stable stratification.

An unanswered question, after having inspected Figure 3.5 and Table 3.3, is whether or not the cases of the Ogives of latent and sensible heat flux coincide (overlap) for the segments considered. Playing devils advocate, even if these two fluxes both converge 70% of the time, it could be that the *total* turbulent heat flux Ogive converged much less frequently leading to problems with (for example) SEB closure. With this in mind we also investigated whether or not there is a considerable coincidence of Ogive cases between the sensible and latent heat fluxes. For 58.75% of the 2790 total number of 3-hour Ogive segments both the sensible and latent heat fluxes converged. As such, c.f. Table 3.3, only for $\sim 12\%$ of the segments did the sensible heat flux ogive converge when the latent heat flux did not. So overall, when the latent heat flux Ogive converged then $\sim 82\%$ of the time the sensible heat flux Ogive case also converged. Similarly for 4.6%, 5.2% and 0.2% of the segments both the sensible Ogives were respectively of the case: extremum, divergent and reversed.

With the exception of the reversed case these percentages are relatively close to those given by the individual case fractions in the final column of Table 3.3. So more often than not, especially considering the convergent case is the most typical, the Ogive case for the sensible heat and latent heat flux Ogives in a given segment do coincide.

Why are emphasizing Ogive case coincidence? Well this presents evidence that, to a first order, scalar cospectral similarity (Wyngaard and Coté, 1972) where the shape of the properly nondimensionalized scalar flux cospectra collapse into a universal curve, is more often than not ($\sim 69\%$ of the time) satisfied at Bayelva. This in turn provides support for the use of the simple analytical cospectral attenuation model of Massman (2000) (and also the integral method of Moore (1986)). The model assumes that such a scalar cospectral similarity exists and that the universal curve is the same for the two scalar fluxes. Perhaps most importantly, the high fraction of coincidence of Ogive cases makes it straightforwards to not only identify problems related to low frequency influences on the surface energy balance but also how often these occur at Bayelva. The fact that both the Ogives converge 58% of the time implies that 42% of the time at least one of the Ogives does not converge. So, depending on the relative magnitude of the Ogive, this may have a considerable influence on the surface energy balance as well as nocturnal CO_2 NEE estimates and the applicability of MOST. This because the divergent/extremum/reversed Ogive provides evidence of low frequency contamination of the turbulent flux due to e.g. mesoscale circulations, drainage flows, gravity waves or transition periods (dawn/dusk) in the boundary layer (Mahrt, 2010) all of which make the underlying assumptions of the EC method invalid.

3.5 Cospectral Similarity

Here we pursue the claim, from Section 3.4, that the nondimensional scalar flux cospectra collapse into the same universal curves. We do so by calculating the nondimensional cospectra for the sensible and latent heat flux and dividing these into stability bins. Only the cospectra, truncated at inverse frequencies of 30 minutes, where the corresponding Ogives converged are considered as it is these that we expect to exhibit similar behavior. This corresponds to 1942 cospectra for the latent heat flux and 1938 cospectra for the sensible heat flux (cf. Table 3.3). Recall that the basic relation for cospectral similarity, proposed by Wyngaard and Coté (1972), was given by (2.32). Specifically, we need to express the nondimensional cospectra as functions of the nondimensional frequency $n = fz/\overline{u}$, where $\overline{u}$ is the segment average plane mean wind, to verify if (2.32) is satisfied and if the form is the same for both heat fluxes.

The result, in the form of an ensemble of nondimensional cosepctral densities and their dependence on nondimensional frequency, is shown through a log-log representation in Figure 3.6. We chose to consider two contrasting stability ranges to highlight differences, these were the moderate to strongly stable range $0.25 < \zeta < 2$ in the right panels and the moderate to strongly unstable case $-2 < \zeta < -0.25$ in the left panels. Multiple narrower stability ranges could have been chosen, yet our choice gets several points across in fewer panels. First of all, the agreement between the peaks of the ensemble median nondimensional cospectra (EMNC), the solid black lines in each panel, of the sensible and latent heat flux is remarkable in both stability ranges. For both the latent and sensible heat flux the EMNC peak at $n = 0.03$ in the unstable category, while the EMNC for the heat fluxes in the same order peak at $n = 0.25$ and $n = 0.29$ for the stable category. The neutral EMNC (not shown) for both fluxes were found to peak at the intermediate frequency of $n = 0.09$. So, the location of the EMNC peaks are more or less the same for the two fluxes in each stability range. Moreover, the claim of Kaimal and Finnigan (1994) that cospectra peak at higher frequencies the higher the stability is clearly demonstrated by both the stable EMNC peaking at a frequency that is nearly a decade higher than that for their unstable counterparts.

A clear difference, for both stabilities, is that the sensible heat flux nondimensional cospectra mainly follow the expected (e.g. (Kaimal and Finnigan, 1994)) $-4/3$ inertial subrange slope (in log-log representation) at high frequencies. Contrarily, most of the latent heat flux nondimensional cospectra drop off more steeply then expected at high frequency. The steeper drop-off provides evidence of more severe cospectral attenuation in the case of the latent heat flux; which is to be expected as w and ρ_v are not measured by collocated instruments (Section 2.5.1). As discussed in Section 3.3, the degree of latent heat flux high frequency cospectral attenuation is likely much larger than that which we correct for through the analytical approach of Massman (2000). At the same time, the high frequency cospectral attenuation in the heat flux appears to be quite negligible. If we disregard the high frequency attenuation the nondimensional cospectra are remarkably similar for the two fluxes. So it is likely that their unattenuated forms would be near identical as is assumed when correcting for cospectral attenuation either through the analytical method of Massman (2000) or the integral method of Moore (1986).

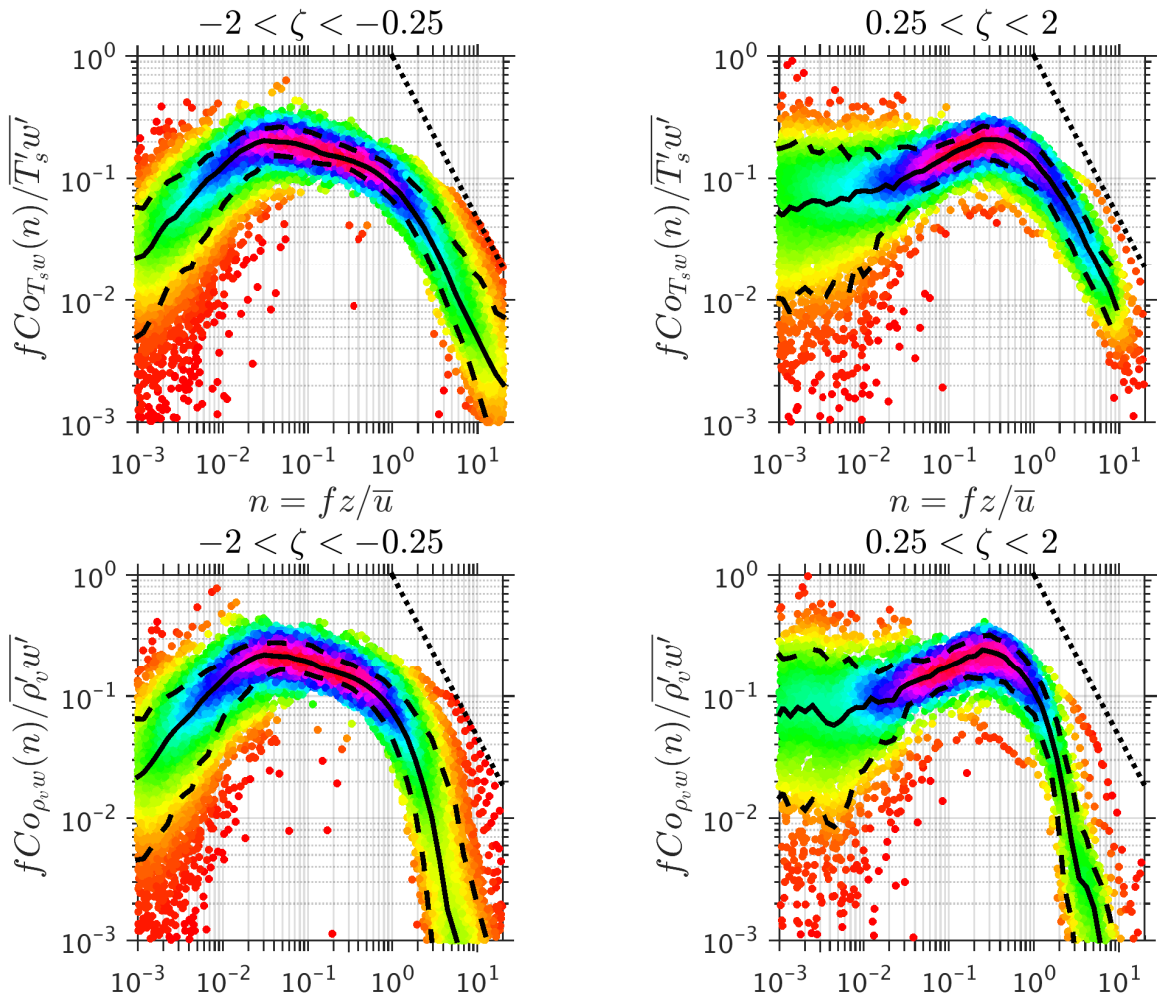


Figure 3.6: *Log-log representation of nondimensional 30 minute cospectra for the sensible heat flux ($fCo_{T_s w}/\overline{T_s' w'}$, upper panels) and latent heat flux ($fCo_{\rho_v w}/\overline{\rho_v' w'}$, lower panels) as functions of the nondimensional frequency $n = fz/\overline{u}$. Only cospectra where the corresponding Ogive (from which the 30 minute covariance used in normalization is obtained) converged are considered. In the left panels nondimensional cospectral densities in the stability range $-2 < \zeta < -0.25$ are presented based on 436 cospectra for both the sensible and latent heat flux. In the right panels nondimensional cospectral densities in the stability range $0.25 < \zeta < 2$ are presented based on 110 latent heat flux cospectra and 137 sensible heat flux cospectra. Pinks indicate the highest density of points whereas reds are the scarcest densities. in each panel the solid black curves represent the median nondimensional cospectrum, while the dashed blue lines depict the 10th (lower) and 90th (upper) percentiles. For orientation the expected $-4/3$ inertial subrange slope is indicated by the dotted black lines.*

There is a final feature that is worth highlighting in Figure 3.6 which circles back to the discussion in Section 3.4. This revolves around the difference in the low frequency behavior of both the flux EMNC in the two stability cases. The drop in the value of the EMNCs from the peak moving towards lower frequencies is much more pronounced in the unstable case. In fact, at the lowest frequencies the EMNC in the stable case has around the same value as the 90th percentile (upper dashed black line) in the unstable case, for both fluxes. Furthermore, the value of the 90th percentile of the stable nondimensional cospectra at the lowest frequencies are nearly equal to the stable and unstable EMNC peak values. This seemingly contradicts the notion of Massman (2000) that the cospectrum should be broader in the unstable case, and the stable EMNC bear little resemblance to the model cosepctra of Kaimal and Finnigan (1994) (cf. Figure 2.13). The apparent lack of agreement is a possible manifestation of significant NTLF influence on the stable cospectra, even in these cases where the corresponding Ogives converged. So, the lack of a strong low frequency drop off in the stable EMNC is likely not a feature of the turbulent cospectrum alone, but instead a manifestation of other influences on the measured cospectra such as gravity waves or drainage flows (Mahrt, 1999). Arguably, the right panels in Figure 3.6 epitomize the problem with resolving the turbulent cospectrum in the stable case, as it is not at all clear how one would go about disentangling the NTLF contributions from the turbulent low frequency contributions to the flux (Sievers et al., 2015).

Chapter 4

Summary & Conclusions

An extensive and transparent implementation of the EC method has been achieved in the form of the module, developed from scratch in `Fortran 90`, described in Chapter 2. All relevant procedures involved in modern EC data processing, as outlined in Aubinet et al. (2012) and Mauder et al. (2013), were included. A development from scratch was necessary in that a variety of new procedures, requiring access to the raw high frequency data streams, were introduced. These procedures, not contained in many of the standardized EC packages such as TK2 (Mauder and Foken, 2004), include the following:

1. The estimation of flux uncertainties using the variance of the covariance (Finkelstein and Sims, 2001) and subsequently accounting for the propogation of uncertainty through flux corrections (Billesbach, 2011).

2. Calculating second order structure functions (Mahrt, 1989) that allow for a simple identification of the inertial subrange and a diagnosis of dissipation (Högström, 1996).

3. Producing spectra, cospectra and Ogives of turbulent quantities (Kaimal and Finnigan, 1994). A new (to boundary layer meteorology) spectral smoothing window (Konno and Ohmachi, 1998) was applied for both cosmetic and practical (storage) reasons.

These are all output by the module, in addition to the standard output (e.g. fluxes). Furthermore, various novel approaches to traditional procedures (cf. Mauder and Foken (2004)) were introduced, these include:

1. Despiking using a median absolute deviation test (Mauder et al., 2013).

2. Planar fitting through unit vector operations (Lee et al., 2006) as opposed to rotation angles.

3. Using the analytical cospectral attenuation correction of Massman (2000) with revised time constants based on an exact scalar flux high frequency transfer function for the CSAT3 sonic anemometer presented by van Dijk (2002).

4. Adapting the 'SND' correction (Schotanus et al., 1983) such that all fluctuating quantities involved are measured directly through an open path EC system.

5. Revising the coefficients, proposed by Thomas and Foken (2002), in the flux-variance similarity relations (ITC) for vertical velocity by accounting for local effects at Bayelva.

Through such a customized development the goal was to construct a module that was tailor made for analyzing raw data from the Bayelva EC system. We achieved this by taking into account the specific conditions of the study site outlined in Westermann et al. (2009): instrument separation, flow distortion, cold temperatures, multiple phases of precipitation occuring, sloping terrain, extreme seasonal variations in the surface type, changes in the measurement height and, crucially, the prevalence of stable stratification during the Arctic night. Particularly the despiking routines and the quality control procedures needed to be tuned according to these local conditions. In addition, the novel procedures introduced allowed us to investigate problems that occur under stable stratification by comparing the behavior of module output in the stable regime to that in the neutral and unstable case.

The aim of our investigation was to assess to what degree the EC method can be relied on to accurately estimate surface exchange in the difficult conditions imposed by our high Arctic study site. As the northernmost member of a global network (FLUXNET) that includes hunderds of permanent flux tower sites across a vast array of terrestrial ecosystems (Baldocchi et al., 2001), our results have implications beyond the Arctic and the EC method. Bayelva, and more generally Svalbard, is an ideal 'laboratory' for studying problems with the EC method that occur in stable stratification. The reason is two fold. Firstly, we are spared from some overlapping complex features that exist in many of the lower latitude sites (e.g. forest canopies (Lee et al., 2006)). Secondly, the stable regime may persist longer during the Arctic night (Westermann et al.,

2009) than it does in a conventional diurnal cycle (Stull, 1988). Therefore, it is possible to study, practically in isolation, the effects of stable stratification on the EC method over extended periods. Nonetheless, the results extend to the typically nocturnal stable regime at lower latitudes in which various phenomena (see e.g. Aubinet (2008)) cause problems for many of the FLUXNET sites. In fact, a serious concern with FLUXNET data is that the EC methodology often fails in low flux stable conditions (Finnigan, 2008). Consequently, longterm budgets of energy, water and carbon exchange are invariably biased, even when gap filled, as most of the available flux estimates are from the unstable and neutral regimes (Aubinet et al., 2012). Even more alarming is the fact that the 'products' from flux towers are widely used by the modeling community: be it in the development of parametrizations (Högström, 1996), for validation (Aas et al., 2015) or to constrain model output through data assimilation (Leuning et al., 2012). So, if these products are accepted at face value, the error in these, both random and systematic, will invariably propogate through the desired development, validation or constraint into the model in question.

Despite the fact that the stable stratification may put the EC method in dire straits, a comprehensive set of tools (Mauder et al., 2013) are available to identify problematic periods. The underlying assumption of the EC method is that measurements are conducted in a surface layer that is simultaneously statistically stationary and horizontally homogeneous (Foken and Wichura, 1996). Thereby, the set of tools revolve around flagging periods where this assumption is clearly violated. We implemented these tools in a two-pronged approach. Initially the raw data was despiked as outlined in Section 2.2. Subsequently, after having undergone further processing, each 30 minute block of data was subjected to the extensive quality control procedure outlined in Section 2.6. In not being directly related to the identification of problems with the underlying assumption, the intermediate processing steps, rotations (Section 2.3) and flux corrections (Section 2.5), still play an important role in reducing the systematic error in flux estimates (Lee et al., 2006).

Overall, despite differences in the methods, we found satisfactory agreement for the high quality fluxes produced by our module and TK2 as shown through linear regression in Figure 3.3. The sensible heat flux showed particularly good agreement; on average our estimates were a factor 0.93 of those in TK2 with a correlation of $R = 0.96$, which is comparable to other EC software package intercomparisons for open path systems (Fratini and Mauder, 2014). For the latent heat flux the agreement was not as good; on average our estimates were a factor 0.76 of those in TK2 with a correlation of $R = 0.88$. Since the underestimation of latent heat flux is systematic (high correlation) it is likely due to a difference in the cospectral attenuation corrections used; for example we did not consider high frequency attenuation resulting from *lateral* sensor separation when implementing the analytical method of Massman (2000). Unfortunately, no radiation or ground heat flux measurements were available for the majority of the period we considered (see Section 2.1) so an evaluation of the SEB closure for our estimates was not possible. Nonetheless, as our estimates are somewhat in an agreement with those of TK2, it is probable that the energy imbalance in our estimates is comparable to that of TK2 shown in Figure 1.14. Thus we can assume that our energy imbalance is equally large, in excess of 50% for hourly fluxes. This comparison with TK2 merely provided a sanity check of the produced fluxes, we had to go beyond the standard TK2 output to investigate the reliability of the EC method under difficult conditions.

As for the identification of problems, one feature clearly stood out, namely departures from statistical stationarity. Of the individual quality flags considered in the QC (Section 2.6), this was by far the most frequently occuring hard flag as shown in Figure 3.1. In line with other investigations at Bayelva (Westermann et al. (2009) and Lüers et al. (2014)), 11% of the 7888 non-overlapping 30 minute blocks considered had to be discarded due to significant non-stationarity. Non-stationarity is symptomatic of a number of mechanisms, outlined in Andreas et al. (2008), although a strong contender at Bayelva is the weak intermittent turbulence disccused in Mahrt (1999) that may manifest in the stable boundary layer. As discussed in Section 3.1, despite not hard flagging as frequently, the remaining flag categories are still important in that each of these has the potential to identify problems that would otherwise be overlooked. Overall 43% of the blocks were flagged as being of high quality, 38% received a soft flag indicating medium quality and the remaining 19% were discarded being hard flagged as poor quality blocks. Discarding almost one fifth of the data may seem extreme, but it is in line with typical hard flag fractions (e.g. Mauder et al. (2013)). If such blocks were not discarded it would be detrimental to the estimated long term surface exchange budgets through the inclusion of clearly unphysical turbulent fluxes some of which have magnitudes in excess of $1000\,\mathrm{Wm}^{-2}$.

All the processing steps helped in reducing (or discarding) systematic errors in the estimated surface fluxes. These fluxes are also subject to considerable uncertainty due to the random error associated with insufficient sampling of integral scale (i.e. main flux containing) eddies within an averaging period (Finkelstein and Sims, 2001). It is not yet standard practice to include estimates of the flux uncertainties in EC packages such as TK2 (Billesbach, 2011). This is unfortunate as the flux uncertainty is an invaluable product for modelers with regards to, for example, the data assimilation of

observed fluxes (Mauder et al., 2013). As such, we calculated the relative uncertainty in our sensible and latent heat flux estimates as outlined in Section 2.6.7. For both fluxes the relative uncertainty for small (absolute) fluxes is found to be larger the lower the absolute flux value of the flux. Still, at larger flux magnitudes, i.e. in excess of $20\,\mathrm{Wm}^{-2}$, the relative uncertainty becomes nearly independent of the flux magnitude in agreement with Finkelstein and Sims (2001). As a function of stability the relative uncertainty in both fluxes nearly doubles in the passage from moderately unstable ($\zeta = -0.5$) to moderately stable stratification ($\zeta = 0.5$), which is at odds with the findings of Finkelstein and Sims (2001) where no clear stabiliy dependence was observed in the uncertainty. Overall, the mean relative uncertainty in the latent and sensible heat fluxes was found to be 18.7% and 19.5%, respectively. For stable stratification ($\zeta > 0.1$) the mean relative uncertainty was considerably higher, being on the order of 27% for both fluxes. All the averaged relative uncertanties shown in Table 3.1 are in general agreement with estimates from lower latitudes (e.g. Mauder et al. (2013)), but represents the only such analysis (to our knowledge) that has been carried out in the Arctic.

By means of the Ogive analysis presented in Section 3.4 we were able to verify how often the fluxes of sensible and latent heat converged within the traditional averaging period of thirty minutes. Flux convergence in turn indicates that we have resolved the entire turbulent cospectrum and captured a reliable estimate of the turbulent flux (Foken et al., 2006). The results of the analysis, in the form of Ogive classification, is summarized in Table 3.3. We found that overall the latent and sensible heat flux Ogives converged for respectively 70% and 71% of the segments considered, while their convergence coincided for 59% of the segments considered. Such a large degree of coincidence, combined with the results of Section 3.5, implies that the two nondimensional scalar cospectra are often nearly identical (cf. Figure 3.6), providing evidence of the scalar cospectral similarity (Wyngaard and Coté, 1972) that is assumed in cospectral attenuation corrections (e.g. Massman (2000)). Once again (to our knowledge) we are the first to apply such an analysis to an Arctic EC site. The convergence fractions were considerably lower than those found at lower latitudes, an indication of the problems caused by the relatively frequent occurence of stable stratification at Bayelva. In fact, the heat flux Ogives, in the same order, converged for only 41% and 48% of the stably stratified segments. We found, through Figure 3.6, that the reason for this lack of convergence in stable stratification can be attributed to an expected (Kaimal and Finnigan, 1994) shift of the cospectral peak towards higher frequencies in the stable regime compared to its unstable and neutral counterparts. That is to say, to resolve such a turbulent cospectrum while avoiding non-turbulent low frequency contributions, a much *shorter* averaging period is required. Shortening the averaging period the required amount, on the other hand, may lead to an inadequate number of samples of the cospectral peak which in turn leads to unacceptable levels of uncertainty. In conclusion, the traditional 30 minute averaging period is more often than not poorly suited for stable stratification. A standardization of novel methods, such as the numerical Ogive optimzation presented in Sievers et al. (2015), is needed if we are to resolve the lack of flux convergence in stable stratification. In so doing the bias introduced in longterm surface exchange budgets could be greatly reduced.

To anyone intending to make use of the EC method in any capacity we highly recommend making good use of the current state of the art in processing procedures given in e.g. Aubinet et al. (2012). The despiking and quality control procedures developed by Vickers and Mahrt (1997) and Foken and Wichura (1996) are mandatory. Respectively, these allow for the identification instrument related problems and violations of the underlying assumption that measurements are occuring within an idealized surface layer. As discussed in Section 1.3 three widely acknowledged problems outlined in Mahrt (1999), Finnigan (2008) and Foken (2008a) are symptomatic of frequent departures from this idealization, particularly under stable stratification (Mahrt, 1999). Moreover, regardless of the intended purpose of the EC investigation it should become standard practice to evaluate and report the uncertainty in flux estimates using one of the many methods presented in Aubinet et al. (2012). If possible we also recommend calculating cospectra, following e.g. Stull (1988) and Kaimal and Finnigan (1994), so as to perform Ogive analysis and investigate flux convergence within an averaging period (Foken et al., 2006).

Finishing off with a short outlook we would like to point out that in focusing on micrometeorology we have only scratched the surface of the wider field that is boundary layer meteorology. As an extension the current work it would be natural to compare some of the additional methods used to calculate flux uncertanties and estimate cospectral attenuation. Moreover, conducting a similar analysis for the flux of CO_2, also available from the Bayelva instrumentation, could confirm our suspicion that our findings in the Ogive analysis are the result of the physics of a stable boundary layer and apply to any scalar flux. In a wider context, running a high resolution Large Eddy Simulation in stable conditions for the Brøgger peninsula could yield insights into the flux footprint (Aubinet et al., 2012) and the three-dimensional structure of the local boundary layer (Westermann et al., 2009). Finally, working on a consistent way of assimilating flux measurements, *with* uncertainty, into models could go a long way towards replacing the current use of flux-profile relationships, at least where observations are available.

Appendix A

A.1 Indicial Notation

Following the tradition in the literature of boundary layer meteorology (e.g Monin and Obukhov (1954) and Stull (1988)) we make use of indicial notation[1]. Working within a purely Cartesian framework the transition from the more familiar vector notation to indicial notation is quite seemless. The benefit of applying indicial notation lies in the simplification of many vector operations common in vector calculus as well as in its (subjective) elegance.

To introduce the indicial framework we consider a familiar and intuitive starting point to be the wind vector $\mathbf{v}$. We may decompose this vector at a given point in space and time into three orthogonal components in our chosen Cartesian coordinate system, using a triplet of orthonormal unit vectors $\left[\widehat{\boldsymbol{\imath}}, \widehat{\boldsymbol{\jmath}}, \widehat{\boldsymbol{k}}\right]$ as follows

$$\mathbf{v} = u\widehat{\boldsymbol{\imath}} + v\widehat{\boldsymbol{\jmath}} + w\widehat{\boldsymbol{k}}\,.$$

From vector analysis the *magnitude* of this vector, $|\mathbf{v}|$, corresponds to the wind speed while the vector normalized by its magnitude is the wind *direction* $\widehat{\mathbf{v}} = \mathbf{v}/|\mathbf{v}|$. So an equivalent expression for the given wind vector would be $\mathbf{v} = |\mathbf{v}|\widehat{\mathbf{v}}$. Reminding us that a vector, as opposed to a scalar, is a physical quantity with both magnitue and direction.

Notice that if we were to set $\widehat{\boldsymbol{\imath}}_1 = \widehat{\boldsymbol{\imath}}, \widehat{\boldsymbol{\imath}}_2 = \widehat{\boldsymbol{\jmath}}, \widehat{\boldsymbol{\imath}}_3 = \widehat{\boldsymbol{k}}$ as the unit vectors of the orthogonal directions $x_1 = x, x_2 = y, x_3 = z$ and accordingly define $v_1 = u, v_2 = v, v_3 = w$ then the vector may be expressed compactly as a sum of components. Namely

$$\mathbf{v} = \sum_{i=1}^{3} v_i \widehat{\boldsymbol{\imath}}_i\,.$$

As specified in the above expression the index i runs over $i \in 1(1)3$ in that the vector is three-dimensional. Let us instead write the vector as

$$\mathbf{v} = v_i \widehat{\boldsymbol{\imath}}_i\,,$$

where we have applied the summation convention: if an index occurs twice in any one term in an expression it is always to be summed, unless the contrary is expressly stated. So we adopt an implicit summation over repated or 'dummy' indices appearing in any given term. Conversely indices that only appear once for any given term are refered to as 'free' indices in which case summation is not implied. When expressing vector equations unit vectors are typically omitted such that, as will be shown, an expression can be quickly decomposed to a given index by insertion of $i = 1, 2$ or 3 and hence to any direction $x_1 = x$, $y_1 = y$ or $z_1 = z$.

The number of free indices for a given term determines the *order* of the term. That is to say s (order zero) is a scalar quantity, a_i (order one) is a vector component, b_{ij} (order two) is a matrix entry while c_{ijk} is a third-order tensor element. As in linear algebra the order of the respective terms appearing as sums in an equation must always be equal. For example $a_i + a_i$ and $a_i + b_{ij}a_j$ are defined while $a_i + b_{ij}$ is not. To ensure clarity we require that the free indices appearing in respective terms must be the same. Furthermore with the aim of consistency we apply a convention of setting our first free index to i and dummy indices as the first available, i.e. not a free index, in the alphabetic sequence j, k, l, m, n.

Recalling the definition of the ∇ operator in Cartesian coordinates we translate this powerful tool into our new notation as follows

$$\nabla = \widehat{\boldsymbol{\imath}}\frac{\partial}{\partial x} + \widehat{\boldsymbol{\jmath}}\frac{\partial}{\partial y} + \widehat{\boldsymbol{k}}\frac{\partial}{\partial z} = \widehat{\boldsymbol{\imath}}_i\frac{\partial}{\partial x_i}\,.$$

It is worth emphasizing that as with all terms appearing in indicial notation the summation convention applies to ∇ as well, and of course that the standard rules of differentiation still hold. Summarizing the subsection we express different operations in both vector and indicial notation for comparisson using the wind vector

[1] Alternative names are Einstein notation, summation notation or suffix notation.

$$\frac{1}{2}\mathbf{v}\cdot\mathbf{v}=\frac{1}{2}v_jv_j\ ,\ \mathbf{vv}\ =v_iv_j\widehat{\boldsymbol{\imath}}_i\widehat{\boldsymbol{\imath}}_j\ ,\ \nabla k=\widehat{\boldsymbol{\imath}}_i\frac{\partial k}{\partial x_i}\ ,\ \nabla\cdot\mathbf{v}=\frac{\partial v_j}{\partial x_j}\ ,\ \nabla\mathbf{v}=\widehat{\boldsymbol{\imath}}_i\widehat{\boldsymbol{\imath}}_j\frac{\partial v_i}{\partial x_j}\,.$$

the terms from left to right express the kinetic energy (per unit mass) $k=\frac{1}{2}v_j^2$ as a dot product, the advective flux as an outer product, the gradient of kinetic energy, the divergence of the windfield and finally the gradient of the windfield (a dyadic product). In the same order these terms represent a scalar, a matrix, a vector, a scalar and finally a matrix. Note the inclusion of unit vectors for the matrix and vector quantities, without these the equalities would *not* hold as the indicial form would express a single element as i and j would be free. Nonetheless for ease of presentation these unit vectors are often made implicit.

A.2 Diagnostics

A derivation of the equations used in the sequential diagnosis (Section 2.4.3) is provided herein. Most of these manipulations are fairly simple with details available in numerous atmospheric physics (e.g. Wallace and Hobbs (2006)) and boundary layer meteorology (e.g. Stull (1988) and Foken (2008b)) textbooks. We will nonetheless take the time to elaborate whenever unusual steps are taken.

First of all recall that in our considerations of moist air we restrict our attention soley to an ideal gas mixture consisting of water vapor and dry air constitutents. The effect of trace gasses other than water vapor (e.g. CO_2 and O_3) on the density of the mixture is assumed to be negigible. Moreover the possible occurence of water in its liquid and solid phase is ignored both for simplicity (as in e.g. Fuehrer and Friehe (2002)) and due to the effects it has on the response of both the CSAT3 and the LI-7500 (Section 1.2.3). We anticipate that all instances where the presence of these other phases is significant will have been flagged and discarded in the despiking routine (Section 2.2). To start off we express the ideal gas law in two equivalent ways (e.g. Stull (1988))

$$P=\rho RT=\rho R_dT_v \tag{A.1}$$

where R is the specific gas constant for moist air and R_d the (known) specific gas constant for dry air; both with units $\mathrm{JK^{-1}kg^{-1}}$. From the above virtual temperature[2], T_v, is defined as the temperature required for a hypothetical dry air parcel to have the same density, ρ, as the given sample of moist air with temperature, T, at the same pressure P. It is a convenient quantity to employ for two reasons. Not only are we spared from having to keep track on the changes in the value of the specific gas constant R for moist air (Wallace and Hobbs, 2006). We also benefit from the fact that T_v is closely related to the sonic temperature, T_s, that we are measuring directly (Kaimal and Gaynor, 1991). For (A.1) to hold then based on some straightforwards manipulations involving partial pressures and densities (e.g. Wallace and Hobbs (2006)) the virtual temperature is related to absolute temperature through

$$T_v=\frac{T}{1-\frac{p_v}{P}\left(1-\mu^{-1}\right)}\simeq T\left[1+\frac{p_v}{P}\left(1-\mu^{-1}\right)\right]\ , \tag{A.2}$$

where p_v is the water vapor pressure[3] and $\mu^{-1}=M_v/M_d=0.622$ is the ratio of the molar mass of vapor ($M_v=0.01802\,[\mathrm{kg\ mol^{-1}}]$) to that of dry air ($M_d=0.02897\,[\mathrm{kg\ mol^{-1}}]$). For the approximation in (A.2) we have used a first order Taylor series expansion on $f(x)=(1-x)^{-1}$ 'forwards' from $x=0$ to simplify the relation for T_v making use of the fact that $0\leq p_v<<P$ so the forward step is small: $\Delta x=\frac{p_v}{P}\left(1-\mu^{-1}\right)<<1$. Note that the error in the approximation of $f(\Delta x)$ is on the order $\Delta x^2<<\Delta x<<1$ which is negiligible for our purposes and we will drop the '$\simeq$' henceforth. From (A.2) it is clear that the virtual temperature is always greater than the absolute temperature for moist air. As a corollary moist air is less dense than dry air (Stull, 1988) on account of having to isobarically increase the temperature of the hypothetical dry air parcel for its density to match that of the moist air parcel.

Next we consider an alternate relation for (dimensionless) specific humidity, q, following Wallace and Hobbs (2006) once more considering partial pressures, partial densities and employing the ideal gas law

$$q=\frac{\rho_v}{\rho_d+\rho_v}=\mu^{-1}\frac{p_v}{P\left[1-\frac{p_v}{P}\left(1-\mu^{-1}\right)\right]}\,. \tag{A.3}$$

In the above we used the following relation for the specific gas constant for water vapor $R_v=\mathcal{R}/M_v$ and similarly for that of dry air $R_d=\mathcal{R}/M_d$, where $\mathcal{R}$ is the universal gas constant (units $\mathrm{JK^{-1}mol^{-1}}$) when inserting for $\mu^{-1}=M_v/M_d=R_d/R_v$. Now using the same Taylor series expansion as in (A.2) then (A.3) is readily

[2] Here the subscript in T_v is short for 'virtual' not water vapor.

[3] Often given the symbol e which we circumvent here to avoid confusion with the turbulent kinetic energy.

simplified to

$$q = \mu^{-1}\frac{p_v}{P}\left[1 + \frac{p_v}{P}\left(1 - \mu^{-1}\right)\right] . \quad (A.4)$$

Now if we insert for $\delta = \mu^{-1}p_v/P << 1$ in (A.4) then using that $\mu - 1 \simeq \mu^{-1}$

$$q = \delta + (\mu - 1)\,\delta^2 \simeq \delta + \mu^{-1}\delta^2 , \quad (A.5)$$

then since $\delta >> \delta^2$ we have to a first order that

$$q \simeq \delta = \mu^{-1}\frac{p_v}{P} . \quad (A.6)$$

When we consider the Ny Ålesund temperature ranges (Figure 2.2) it turns out that the approxiation made is marginal. Taking a 'worst case' scenario (maximizing q) where we estimate, using Teten's equation as given in Foken (2008b), the saturation vapor pressure, p_v^*, at $T = 21°\mathrm{C}$ at a pressure of $P = 950\,\mathrm{hPa}$ we find that the relative error in the saturated specific humidity computed with (A.6) as opposed to (A.5) is only on the order of 1%. So as a good approximation we may use (A.6) in (A.2) and arrive at the following expression for T_v (as given directly in Kaimal and Gaynor (1991))

$$T_v \simeq T\left[1 + \mu q\left(1 - \mu^{-1}\right)\right] = T\left[1 + 0.61q\right] . \quad (A.7)$$

Furthermore we can use (A.6) to approximate the relation given in Kaimal and Businger (1963) for sonic temperature in terms of absolute temperature[4]

$$T_s = T\left(1 + 0.32\frac{p_v}{p}\right) \simeq T\,(1 + 0.51q) . \quad (A.8)$$

Subsequently we wish to relate sonic temperature to virtual temperature which would enable us to diagnose q directly through the ideal gas law. To do so we once more use a Taylor series expansion in (A.8), neglecting higher order terms using the same reasoning as previously, to arrive at the first order approximation

$$T = \frac{T_s}{1 + 0.51q} \simeq T_s\,(1 - 0.51q) . \quad (A.9)$$

Further we combine (A.9) with (A.7) to arrive at

$$T_v = T_s\left(1 + 0.1q - 0.31q^2\right) \simeq T_s\,(1 + 0.1q) . \quad (A.10)$$

where we have used $q << 1$ to justify dropping the q^2 term. Finally using (A.1) and (A.10) and inserting for $\rho = q\rho_v$ with some rearrangement we arrive at the following diagnostic equation for specific humidity where involving only known (measured) variables

$$q = \left(\frac{P}{R_d\rho_v T_s} - 0.1\right)^{-1} . \quad (A.11)$$

So (A.11) is the diagnostic equation employed for the estimation of q in the module from which ρ is estimated through the relation $\rho = \rho_v/q$. Subsequently (A.9) and (A.7) are used to estimate T and T_v respectively.

A.3 SND adaptation

No derivation of the SND corrections are given in either Schotanus et al. (1983) or Liu et al. (2001); instead the corrections are merely presented. So before proceeding with how we adapted the SND correction for the sensible heat flux we feel it is instructive to first present a quick derivation of the heat flux correction given in Schotanus et al. (1983) as many of the same procedures are used in the adaptation. To start off we note that the relation between instantaneous sonic and absolute temperature given by (A.8) is only valid if crosswind corrections have been applied (Liu et al., 2001). Thankfully this is the case for the CSAT3 instrument that we employ (see CSAT3 (2014)). Next applying a removal of the block average from T_s and using (A.8) we note that the fluctuating sonic temperature can be written as

$$T_s' = T_s - \overline{T} = T\,(1 + 0.51q) - \overline{T\,(1 + 0.51q)} . \quad (A.12)$$

So the block average of the product $T_s'w'$

$$\overline{T_s'w'} = \overline{\left[T\,(1 + 0.51q) - \overline{T\,(1 + 0.51q)}\right]w'} \quad (A.13)$$

where, once more noticing that multiple terms are zero upon Reynolds averaging ($\overline{\chi}\,\overline{\xi'} = 0$), we are left with

$$\overline{T_s'w'} = \overline{T\,(1 + 0.51q)\,w'} , \quad (A.14)$$

Reynolds averaging once again by first expanding and expressing quantities as the sum of their mean and fluctuating components yields

$$\overline{T_s'w'} = \overline{T'w'} + 0.51\overline{q'T'w'} + 0.51\overline{q}\,\overline{T'w'} + 0.51\overline{T}\,\overline{q'w'} . \quad (A.15)$$

Under the assumption that the scalar[5] fluctuations are small compared to the block averages (i.e. $\overline{q} >> |q'|$ and $\overline{T} >> |T'|$) combined with the fact that $|q'| << 1$ (*not* the case for $|T'|$) then the triple correlation term, $0.1\overline{q'T'w'}$, is negligible with respect to remaining terms. Moreover since $\overline{T}_s >> 1 >> \overline{q}$ we can safely assume that all but the first two remaining terms on the right hand side are negligible and we are left with the approximate relation

$$\overline{T_s'w'} \simeq \overline{T'w'} + 0.51\overline{q}\,\overline{T'w'} + 0.51\overline{T}\,\overline{q'w'} . \quad (A.16)$$

[4]The relation is only valid when crosswind has been corrected for (Schotanus et al., 1983); but as noted in Section 2.5.3 the crosswind correction is applied internally in the CSAT3 firmware.

[5]An assumption that does not hold for the vertical velocity since $\overline{w}$ is near zero whereas $|w'|$ can be very large in comparisson.

Which expressed in terms of the quantity we seek, $\overline{T'w'}$, reads

$$\overline{T'w'} \simeq \frac{\overline{w'T'_s} - 0.51\,\overline{T}\overline{q'w'}}{1 + 0.51\overline{q}}\,. \tag{A.17}$$

Finally noting that $\overline{q} << 1$ then as a good approximation

$$\overline{T'w'} \simeq \overline{w'T'_s} - 0.51\,\overline{T}\,\overline{q'w'}\,, \tag{A.18}$$

which is the form given directly in Schotanus et al. (1983) and Liu et al. (2001) when crosswind corrections are already applied.

Next we adapted (A.18) so as to be more tractable in an iterative procedure with two other flux corrections (spectral attenuation and WPL). To start we expressed the specific humidity as $q = \alpha\rho_v$ where $\alpha = \rho^{-1}$ is the instantaneous specific volume. Further we noted that α is approximately given through the ideal gas law by

$$\alpha = \frac{R_d T_v}{P} \simeq \frac{R_d T_s}{P}\,, \tag{A.19}$$

where we used that to a first order $T_v \simeq T_s$ based on (A.10) and the fact that $q << 1$. Next with the aid of mean removal and Reynolds averaging rules the fluctuating specific humidity can be exanded as

$$q' = (\alpha\rho_v)' = \alpha\rho_v - \overline{\alpha\rho_v} = \alpha\rho_v - \overline{\alpha}\,\overline{\rho_v} - \overline{\alpha'\rho'_v}\,. \tag{A.20}$$

As such the product block average of the product $q'w'$ in (A.18) can be expressed as

$$\overline{q'w'} = \overline{\left(\alpha\rho_v - \overline{\alpha}\,\overline{\rho_v} - \overline{\alpha'\rho'_v}\right)w'}\,, \tag{A.21}$$

which, when Reynolds averaging rules are invoked (i.e. $\overline{\overline{\alpha}\,\overline{\rho_v}\,w'} = \overline{\overline{\alpha'\rho'_v}\,w'} = 0$) ,becomes

$$\overline{q'w'} = \overline{\alpha\rho_v w'} = \overline{(\overline{\alpha} + \alpha')\,(\overline{\rho_v} + \rho'_v)\,w'}\,, \tag{A.22}$$

subsequently expanding the brackets and invoking Reynolds averaging rules once more we are left with

$$\overline{q'w'} = \overline{\alpha}\,\overline{\rho'_v w'} + \overline{\rho_v}\,\overline{\alpha' w'} + \overline{\alpha'\rho'_v w'}\,. \tag{A.23}$$

For the same reason as in (A.16) the magnitude of the triple correlation term, $\overline{\alpha'\rho'_v w'}$, is assumed to be negligible with respect to the other terms so to a first order

$$\overline{q'w'} \simeq \overline{\alpha}\,\overline{\rho'_v w'} + \overline{\rho_v}\,\overline{\alpha' w'}\,. \tag{A.24}$$

Subsequently we insert for the specific volume in (A.24) making use of the approximation in (A.19) where in our case the sonic temperature is the only fast responding variable whereby

$$\overline{q'w'} \simeq \frac{R_d}{P}\left(\overline{T_s}\,\overline{\rho'_v w'} + \overline{\rho_v}\,\overline{T'_s w'}\right)\,. \tag{A.25}$$

Finally by insertion of (A.25) into (A.18) we have our revised SND correction for sensible heat flux as employed in the module, namely

$$\overline{T'w'} \simeq \overline{T'_s w'}\left(1 - \frac{0.51 R_d \overline{\rho_v}\,\overline{T}}{P}\right) - \frac{0.51 R_d \overline{T_s}\,\overline{T}}{P}\,\overline{\rho'_v w'}\,, \tag{A.26}$$

where *all* the fluctuating quantities are measured directly by the EC system.

A.4 Origin of the WPL terms

Under the governing constraint of zero dry air mass flux proposed by Webb et al. (1980) the mean vertical velocity in a statistically stationary and horizontally homogeneous surface layer is given by

$$\overline{w} = -\frac{\overline{\rho'_d w'}}{\rho_d}\,. \tag{A.27}$$

To arrive at a tractable form of the above we need a measurable expression for the turbulent mass flux of dry air $\overline{\rho'_d w'}$. To start off we take the usual route (c.f. Webb et al. (1980),Fuehrer and Friehe (2002) and Lee and Massman (2011)) of seeking an expression for the dry air density ρ_d. Based on Daltons law of partial pressures $p = p_d + p_v = \rho_d R_d T + \rho_v R_v T$ (e.g. Wallace and Hobbs (2006)) upon rearrangement the dry air density may be expressed as

$$\rho_d = \frac{p}{R_d T} - \mu\rho_v\,, \tag{A.28}$$

where we now use p (as opposed to P) to emphasize that strictly speaking we are dealing with the instantaneous pressure. Next following Fuehrer and Friehe (2002) we note that

$$\frac{1}{T} = \frac{1}{\overline{T} + T'} = \frac{1}{\overline{T}}\left[\frac{1}{1 + \frac{T'}{\overline{T}}}\right]\,, \tag{A.29}$$

whereby upon Taylor series expanding the term in the brackets as in (A.9) we have to a first order that

$$\frac{1}{T} \simeq \frac{1}{\overline{T}}\left[1 - \frac{T'}{\overline{T}}\right]\,, \tag{A.30}$$

where the truncation error is marginal on accout of $1 >> |\frac{T'}{\overline{T}}|$ and we once more drop the '$\simeq$'. So inserting (A.30) into (A.28) and expanding the instantaneous pressure into the sum of its mean and fluctuating components it follows that

$$\rho_d = \frac{(\overline{p} + p')}{R_d\,\overline{T}} - \frac{(\overline{p} + p')\;T'}{R_d\,\overline{T}^2} - \mu\,(\overline{\rho_v} + \rho'_v)\,. \tag{A.31}$$

If we now consider that the turbulent vertical dry air

mass flux can be expressed equivalently as

$$\overline{\rho_d' w'} = \overline{(\rho_d + \overline{\rho_d})\, w'} = \overline{\rho_d w'}\,, \tag{A.32}$$

then upon inserting (A.31) into (A.32) with Reynolds averaging rules in mind as usual then

$$\overline{\rho_d' w'} = \frac{\overline{p' w'}}{R_d\, \overline{T}} - \frac{\overline{p}\, \overline{T' w'}}{R_d\, \overline{T}^2} - \frac{\overline{p' T' w'}}{R_d\, \overline{T}^2} - \mu\, \overline{\rho_v' w'}\,. \tag{A.33}$$

Now taking the average of (A.31) we note upon rearrangement that

$$\frac{\overline{p}}{R_d\, \overline{T}} = \overline{\rho_d} + \mu\, \overline{\rho_v}\,. \tag{A.34}$$

where we have ignored the $\frac{\overline{p' T'}}{R_d\, \overline{T}^2}$ term on account of $|p' T'| << \overline{p}\, \overline{T}$. Subsequently inserting (A.34) in (A.33) then

$$\overline{\rho_d' w'} = (\overline{\rho_d} + \mu\, \overline{\rho_v}) \left[\frac{\overline{p' w'}}{\overline{p}} - \frac{\overline{T' w'}}{\overline{T}} \right] - \mu \overline{\rho_v' w'}\,. \tag{A.35}$$

Subsequently inserting (A.35) in (A.27), ignoring the relatively small triple correlation term and following Webb et al. (1980) in defining the shorthand $\sigma = \overline{\rho_v}/\overline{\rho_d}$ yields the following expression for the mean vertical velocity

$$\overline{w} = (1 + \mu\sigma) \left[\frac{\overline{T' w'}}{\overline{T}} - \frac{\overline{p' w'}}{\overline{p}} \right] + \mu\sigma\, \frac{\overline{\rho_v' w'}}{\overline{\rho_v}}\,. \tag{A.36}$$

It is this expression for mean vertical velocity that is the source of the terms in the WPL correction presented in Section 2.5.4,

A.5 Estimating the Variance of a Covariance

In the following where our aim is to arrive at an estimate for the variance of a covariance we begin by highlighting some symmetries laid out in Finkelstein and Sims (2001). We consider block segments of two LDT variables ξ'' and χ''; the block averages of which are zero on account of the LDT. The symmetry present depends on which series is lagged/advanced with respect to the other by a constant time shift $\pm\tau_L$. For the crosscovariances the following holds within a block segment

$$\gamma_{\xi\chi}(\tau_L) = \gamma_{\chi\xi}(-\tau_L) = \overline{\xi''(t)\chi''(t+\tau_L)}\,, \tag{A.37}$$

and

$$\gamma_{\chi\xi}(\tau_L) = \gamma_{\xi\chi}(-\tau_L) = \overline{\chi''(t)\xi''(t+\tau_L)}\,. \tag{A.38}$$

It is worth emphasizing that by definition of a block average the points in time can not extend beyond the block itself. So these points, be it t, $t+\tau_L$ or $t-\tau_L$, always lie in the range $[t_m, t_m + \tau_A]$ where t_m is the timestamp for the start of the block and τ_A is the averaging time. Furthermore we will make use of the fact that the autocovariance is an even function:

$$\gamma_{\xi\xi}(\tau_L) = \gamma_{\xi\xi}(-\tau_L) = \overline{\xi''(t)\xi''(t+\tau_L)}\,, \tag{A.39}$$

for any choice of ξ''. As in Section 2.5.1 we will use the following shorthand $\gamma_{\xi\chi,j} = \gamma_{\xi\chi}(\tau_{L,j})$ for the crosscovariance with χ'' lagged in time by the discrete lag $\tau_{L,j} = j\Delta t$ with respect to ξ''. Analogously $\gamma_{\chi\xi,j} = \gamma_{\chi\xi}(\tau_{L,j})$ is the crosscovariance with ξ'' lagged in time by the discrete lag $\tau_{L,j} = j\Delta t$. Finally it is worth reminding the reader that for a given discrete lag $j\Delta t$ the crosscovariance is computed through

$$\gamma_{\xi\chi,j} = \frac{1}{I} \sum_{i=1}^{I-j} \xi_i'' \chi_{i+j}''\,,$$

in the case that χ'' is lagged in time by $j\Delta t$ with respect to ξ'', and through

$$\gamma_{\chi\xi,j} = \frac{1}{I} \sum_{i=1}^{I-j} \chi_i'' \xi_{i+j}''\,, \tag{A.40}$$

in the case that ξ'' is lagged in time by $j\Delta t$ with respect to χ''. Recall that in our case $I = 3.6 \times 10^4$ for the 30 minute block average sampled at 20 Hz. Thus it is worth bearing in mind that the act of computing the crosscovariance (or autocovariance) for a *single* lag involves summing over a large amount of terms.

Now the method of Finkelstein and Sims (2001) involves estimating the variance of a covariance, $\sigma^2_{\overline{\xi' w'}}$, based on the following mathematically rigorous expression from Fuller (1996). For the two discrete blocks segments, ξ'' and χ'', each with I entries the expression reads

$$\sigma^2_{\overline{\xi' w'}} \simeq \frac{1}{I} \sum_{j=-J}^{J} \left(\gamma_{\xi\xi,j} \gamma_{\chi\chi,j} + \gamma_{\xi\chi,j} \gamma_{\chi\xi,j} \right)\,. \tag{A.41}$$

Strictly speaking the summation in (A.41) should run from $-I$ to I but the expense of such a calculation is detrimental. Consider the following, there are: thousands of blocks that we need to apply (A.41) to, each of these 30 minute block consists of $I = 3.6 \times 10^4$ points in time *and* for each j calculating the auto/crosscovariances implies summing over $I - j$ terms. In fact with $J = I$ the expense would be on the order 10^{12} computations per 10^3 blocks. So in practice J is set to be several order of magnitudes lower than I on the assumption that the autocovariance and crosscovariances are negligbile at large lag. Based on our own sensitivity tests as well as those of Finkelstein and Sims (2001) and Billesbach (2011) it was found that increasing J

much beyond 400 has negligible effect on the variance of the covariance. Even with $J = 400$, however, (A.41) is still an expensive procedure (10^{10} computations for 10^3 blocks).

Here we can exploit the previously discussed symmetry in the crosscovariances and autocovariances. Using (A.37), (A.38) and (A.39) in (A.41) the variance of the covariance can be expressed equivalently as follows (with $J = 400$)

$$\sigma^2_{\overline{\xi' w'}} \simeq \frac{1}{I}\left[\overline{\xi'^2}\,\overline{\chi'^2} + \overline{\xi'\chi'}^2 + 2\sum_{j=1}^{J}\left(\gamma_{\xi\xi,j}\gamma_{\chi\chi,j} + \gamma_{\xi\chi,j}\gamma_{\chi\xi,j}\right)\right]. \tag{A.42}$$

Where the number computations are practically cut in half compared to (A.41) with the same choice of J. This because we no longer have to consider negative indices in the summation range. Note that in our presentation of (A.42) we have chosen to re-emphasize that for $j = 0$ the crosscovariances and the autocavariances correspond to the respective covariances and variances. It is (A.42) that we employ in the module when estimating the variance of the vertical turbulent fluxes of sonic heat and water vapor. These are then used to estimate the flux uncertainty of the sensible and latent heat flux in Section 2.6.7, where we account for the propogation of uncertanties through the flux corrections as outlined in Appendix A.6.

A.6 Propogation of Flux Uncertainty

The final sensible and latent heat flux estimates are produced after the raw data has undergone extensive processing, which includes multiple corrections. As such the uncertainty in the turbulent flux estimates can not simply be equated to those of the final dynamic flux estimates. Uncertainty will invaraibly propogate through the processing steps and in particular the flux corrections (Section 2.5). As noted in Billesbach (2011) accounting for the uncertainty propogation is *not* simply a matter of adding together uncertainty estimates.

In general denoting the vertical flux of some quantity ξ as F_ξ where the flux in question is typically a function of I independent parameters, X_i with $i \in 1(1)I$. Thereby we may write $F_\xi = F_\xi\left(X_1, \ldots, X_I\right)$ such that the uncertainty in the flux estimate is given by (Billesbach, 2011)

$$\sigma^2_{F_\xi} = \sum_{i=1}^{I}\left(\frac{\partial F_\xi}{\partial X_i}\right)^2 \sigma^2_{X_i} \tag{A.43}$$

in which $\sigma^2_{X_i}$ is the random uncertainty (variance) in the estimate of parameter X_i.

We will consider the uncertainty propogation for the corrected kinematic sensible heat flux and latent heat flux, which we will denote F_T and F_{ρ_v} respectively. That is $F_{\rho_v} = \left(\overline{\rho_v w}\right)_C$ and $F_T = \left(\overline{T' w'}\right)_C$ where the index C symbolized that the relevant sequence of flux corrections have been implemented. These corrected fluxes can be expressed with respect to the uncorrected (but planar fitted) turbulent fluxes:

$$F_T = CF_{a,H}\left(1 - \frac{0.51 R_d \overline{\rho}_v\,\overline{T}}{P}\right)\overline{T'_s w'} - 0.51 CF_s\,CF_{a,E}\,\frac{R_d\,\overline{T}_s\,\overline{T}}{P}\overline{\rho'_v w'}, \tag{A.44}$$

and

$$F_{\rho_v} = (1+\mu\sigma)\left[CF_s\,CF_{a,E}\left(1 - \frac{0.51 R_d\,\overline{T}_s\,\overline{\rho}_v}{P}\right)\overline{\rho'_v w'} + CF_{a,H}\,\frac{\overline{\rho_v}}{\overline{T}}\left(1 - \frac{0.51 R_d \overline{\rho}_v\,\overline{T}}{P}\right)\overline{T'_s w'}\right]. \tag{A.45}$$

In the above CF_s is the previously defined (Section 2.5.1) correction factor for sensor separation. Furthermore $CF_{a,H}$ and $CF_{a,E}$, in which subscript H is a shorthand for $\overline{T'_s w'}$ and subscript E is a shorthand for $\overline{\rho'_v w'}$, are the previously defined (Section 2.5.2) correction factors for cospectral attenuation of the vertical turbulent fluxes of sonic temperature and water vapor respectively.

When calculating the error propogation we assume that the only independent parameters for which the contribution to the uncertainty is signifcant are the covariances $\overline{\rho'_v w'}$ and $\overline{T'_s w'}$. There is undoubtedly some uncertainty in the estimation of the means as well, but we anticipate that terms involving the uncertainty of a mean will be small in comparisson. This because these uncertanties are weighted by small covariances as opposed to the relatively large means *and* the uncertainty in the

means is generally smaller than that of the covariance (Kaimal and Finnigan, 1994). Consider the example term, $G = \overline{T}\,\overline{T_s' w'}$, which is present in both (A.45) and (A.44), then from (A.43) total uncertainty in our G estimate would be $\sigma_G^2 = \overline{T_s' w'}^2\,\sigma_{\overline{T}}^2 + \overline{T}^2\,\sigma_{\overline{T_s' w'}}^2$. On account of $|\overline{T_s}| >> |\overline{T_s' w'}|$ and $\sigma_{\overline{T}} << \sigma_{\overline{T_s' w'}}$ then the uncertainty is to a very close approximation given by $\sigma_G^2 = \overline{T}^2\,\sigma_{\overline{T_s' w'}}^2$. For simplicity we also assume that there is no appreciable uncertainty in the correction factors (CF_a, CF_s). With the above considerations in mind then upon applying (A.43) to (A.44) and (A.45) the total uncertainties, now expressed as standard deviations with units of flux, in the kinematic sensible and latent heat flux are given by

$$\sigma_{F_T} = \left(\left[CF_{a,H} \left(1 - \frac{0.51 R_d \overline{\rho}_v \overline{T}}{P} \right) \right]^2 \sigma_{\overline{T_s' w'}}^2 + \left[0.51 CF_s\, CF_{a,E}\, \frac{R_d\, \overline{T}_s\, \overline{T}}{P} \right]^2 \sigma_{\overline{\rho_v' w'}}^2 \right)^{1/2}, \tag{A.46}$$

and

$$\sigma_{F_{\rho_v}} = (1 + \mu\sigma) \left(\left[CF_s\, CF_{a,E} \left(1 - \frac{0.51 R_d\, \overline{T}_s\, \overline{\rho}_v}{P} \right) \right]^2 \sigma_{\overline{\rho_v' w'}}^2 + \left[CF_{a,H}\, \frac{\overline{\rho_v}}{\overline{T}} \left(1 - \frac{0.51 R_d \overline{\rho}_v\, \overline{T}}{P} \right) \right]^2 \sigma_{\overline{T_s' w'}}^2 \right)^{1/2}. \tag{A.47}$$

Recall that $\sigma_{\overline{\rho_v' w'}}^2$ and $\sigma_{\overline{T_s' w'}}^2$ have already been computed using the method of Finkelstein and Sims (2001) through (A.42). So diagnosing (A.46) and (A.47) on a block-by-block basis in the module is relatively straightforwards. As pointed out in Billesbach (2011) although the expressions are seemingly both daunting and tedious, once these are coded up and checked the calculation is automatic. Moreover taking into account the propogation of uncertainty does not impose a heavy computational burden. In fact when compared to diagnosing the uncertainty of the uncorrected turbulent fluxes via the crosscovariance method (which is expensive) the additional computing power required by the propogation estimation is completely negligible.

Bibliography

Aas, K. S., Berntsen, T. K., Boike, J., Etzelmüller, B., Kristjánsson, J. E., Maturilli, M., Schuler, T. V., Stordal, F. and Westermann, S. (2015). A comparison between simulated and observed surface energy balance at the Svalbard archipelago. Journal of Applied Meteorology and Climatology (early online release) *(-)*.

Andreas, E. L., Geiger, C. A., Treviño, G. and Claffey, K. J. (2008). Identifying nonstationarity in turbulence series. Boundary-layer meteorology *127*, 37–56.

Arguez, A. and Vose, R. S. (2011). The definition of the standard WMO climate normal: The key to deriving alternative climate normals. Bulletin of the American Meteorological Society *92*, 699–704.

Aubinet, M. (2008). Eddy covariance CO2 flux measurements in nocturnal conditions: an analysis of the problem. Ecological Applications *18*, 1368–1378.

Aubinet, M., Vesala, T. and Papale, D. (2012). Eddy covariance: a practical guide to measurement and data analysis. Springer. Textbook.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R. et al. (2001). FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities. Bulletin of the American Meteorological Society *82*, 2415–2434.

Beare, R. J., Macvean, M. K., Holtslag, A. A., Cuxart, J., Esau, I., Golaz, J.-C., Jimenez, M. A., Khairoutdinov, M., Kosovic, B., Lewellen, D. et al. (2006). An intercomparison of large-eddy simulations of the stable boundary layer. Boundary-Layer Meteorology *118*, 247–272.

Billesbach, D. (2011). Estimating uncertainties in individual eddy covariance flux measurements: A comparison of methods and a proposed new method. Agricultural and Forest Meteorology *151*, 394–405.

Boike, J., Ippisch, O., Overduin, P. P., Hagedorn, B. and Roth, K. (2008). Water, heat and solute dynamics of a mud boil, Spitsbergen. Geomorphology *95*, 61–73.

Brezinski, C. and Tournès, D. (2014). The method of Cholesky for linear systems. In André-Louis Cholesky pp. 77–122. Springer.

Burba, G. (2013). Eddy Covariance Method for Scientific, Industrial, Agricultural and Regulatory Applications: A Field Book on Measuring Ecosystem Gas Exchange and Areal Emission Rates. LI-COR Biosciences.

Burba, G. G., McDermitt, D. K., Grelle, A., Anderson, D. J. and Xu, L. (2008). Addressing the influence of instrument surface heat exchange on the measurements of CO2 flux from open-path gas analyzers. Global Change Biology *14*, 1854–1876.

Businger, J. (1982). The fluxes of specific enthalpy, sensible heat and latent heat near the earth's surface. Journal of the Atmospheric Sciences *39*, 1889–1892.

Charuchittipan, D., Babel, W., Mauder, M., Leps, J.-P. and Foken, T. (2014). Extension of the averaging time in eddy-covariance measurements and its effect on the energy balance closure. Boundary-Layer Meteorology *152*, 303–327.

Clement, R. J. (2005). Mass and energy exchange of a plantation forest in Scotland using micrometeorological methods. PhD thesis, University of Edinburgh.

CSAT3 (2014). CSAT3 three dimensional sonic anemometer instruction manual. Campbell Scientific Inc.

Desjardins, R. L., MacPherson, J. I., Schuepp, P. H. and Karanja, F. (1989). An evaluation of aircraft flux measurements of CO_2, water vapor and sensible heat. Boundary Layer Meteorology *47*, 55–69.

Domingos, P. (2012). A few useful things to know about machine learning. Communications of the ACM *55*, 78–87.

eKlima (2014). eKlima: Free access to weather- and climate data from the Norwegian Meteorological Institute from historical data to real time observations. URL: `http://eklima.met.no`.

Esau, I., Argentini, S., Przybylak, R., Repina, I. and Sjöblom, A. (2012). Svalbard Meteorology. Advances in Meteorology *2012*.

Esau, I. and Repina, I. (2012). Wind climate in Kongsfjorden, Svalbard, and attribution of leading wind driving mechanisms through turbulence-resolving simulations. Advances in Meteorology *2012*.

Eugster, W., McFadden, J. P. and Chapin III, F. S. (1997). A comparative approach to regional variation in surface fluxes using mobile eddy correlation towers. Boundary-Layer Meteorology *85*, 293–307.

Finkelstein, P. L. and Sims, P. F. (2001). Sampling error in eddy correlation flux measurements. Journal of Geophysical Research: Atmospheres (1984–2012) *106*, 3503–3509.

Finnigan, J. (2004). A re-evaluation of long-term flux measurement techniques Part II: coordinate systems. Boundary-Layer Meteorology *113*, 1–41.

Finnigan, J. (2008). An introduction to flux measurements in difficult conditions. Ecological Applications *18*, 1340–1350.

Finnigan, J., Clement, R., Malhi, Y., Leuning, R. and Cleugh, H. (2003). A re-evaluation of long-term flux measurement techniques part I: averaging and coordinate rotation. Boundary-Layer Meteorology *107*, 1–48.

Foken, T. (2006). 50 years of the Monin–Obukhov similarity theory. Boundary-Layer Meteorology *119*, 431–447.

Foken, T. (2008a). The Energy Balance Closure Problem: An Overview. Ecological Applications *18*, 1351–1367.

Foken, T. (2008b). Micrometeorology. Springer. Textbook.

Foken, T., Aubinet, M., Finnigan, J. J., Leclerc, M. Y., Mauder, M. and Paw U, K. T. (2011). Results of a panel discussion about the energy balance closure correction for trace gases. Bulletin of the American Meteorological Society *92*, ES13–ES18.

Foken, T. and Wichura, B. (1996). Tools for quality assessment of surface-based flux measurements. Agricultural and Forest Meteorology *78*, 83–105.

Foken, T., Wimmer, F., Mauder, M., Thomas, C. and Liebethal, C. (2006). Some aspects of the energy balance closure problem. Atmospheric Chemistry and Physics *6*, 4395–4402.

Førland, E. J., Benestad, R., Hanssen-Bauer, I., Haugen, J. E. and Skaugen, T. E. (2012). Temperature and precipitation development at Svalbard 1900–2100. Advances in Meteorology *2011*.

Førland, E. J. and Hanssen-Bauer, I. (2000). Increased precipitation in the Norwegian Arctic: true or false? Climatic Change *46*, 485–509.

Fratini, G. and Mauder, M. (2014). Towards a consistent eddy-covariance processing: an intercomparison of EddyPro and TK3. Atmospheric Measurement Techniques *7*, 2273–2281.

Fuehrer, P. and Friehe, C. (2002). Flux corrections revisited. Boundary-Layer Meteorology *102*, 415–458.

Fuller, W. A. (1996). Introduction to Statistical Time Series, vol. 230,. John Wiley & Sons.

Gash, J. and Culf, A. (1996). Applying a linear detrend to eddy correlation data in realtime. Boundary-Layer Meteorology *79*, 301–306.

Gobbi, M., Chamecki, M. and Dias, N. (2006). Application of digital filtering for minimizing aliasing effects in atmospheric turbulent surface layer spectra. Water resources research *42*.

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

Gu, L., Massman, W. J., Leuning, R., Pallardy, S. G., Meyers, T., Hanson, P. J., Riggs, J. S., Hosman, K. P. and Yang, B. (2012). The fundamental equation of eddy covariance and its application in flux measurements. Agricultural and Forest Meteorology *152*, 135–148.

Hartmann, D., Klein Tank, A., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., Dentener, F., Dlugokencky, E., Easterling, D., Kaplan, A., Soden, B., Thorne, P., Wild, M. and Zhai, P. (2013). Observations: Atmosphere and Surface. In Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, (Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V. and Midgley, P., eds),. Cambridge University Press Cambridge, United Kingdom and New York, NY, USA.

Haugen, D., Kaimal, J. and Bradley, E. (1971). An experimental study of Reynolds stress and heat flux in the atmospheric surface layer. Quarterly Journal of the Royal Meteorological Society *97*, 168–180.

Högström, U. (1990). Analysis of turbulence structure in the surface layer with a modified similarity formulation for near neutral conditions. Journal of the atmospheric sciences *47*, 1949–1972.

Högström, U. (1996). Review of some basic characteristics of the atmospheric surface layer. In Boundary-Layer Meteorology 25th Anniversary Volume, 1970–1995 pp. 215–246. Springer.

Holtslag, A., Svensson, G., Baas, P., Basu, S., Beare, B., Beljaars, A., Bosveld, F., Cuxart, J., Lindvall, J., Steeneveld, G. et al. (2013). Stable atmospheric boundary layers and diurnal cycles: challenges for weather and climate models. Bulletin of the American Meteorological Society *94*, 1691–1706.

Horst, T. (1997). A simple formula for attenuation of eddy fluxes measured with first-order-response scalar sensors. Boundary-Layer Meteorology *82*, 219–233.

Jocher, G., Schulz, A., Ritter, C., Neuber, R., Dethloff, K. and Foken, T. (2014). The Sensible Heat Flux in the Course of the Year at Ny-Ålesund, Svalbard: Characteristics of Eddy Covariance Data and Corresponding Model Results. Advances in Meteorology *2015*.

Kaimal, J. and Businger, J. (1963). A continuous wave sonic anemometer-thermometer. Journal of Applied Meteorology *2*, 156–164.

Kaimal, J., Clifford, S. and Lataitis, R. (1989). Effect of finite sampling on atmospheric spectra. Boundary-Layer Meteorology *47*, 337–347.

Kaimal, J. and Gaynor, J. (1991). Another look at sonic thermometry. Boundary-Layer Meteorology *56*, 401–410.

Kaimal, J. and Kristensen, L. (1991). Time series tapering for short data samples. Boundary-Layer Meteorology *57*, 187–194.

Kaimal, J., Wyngaard, J. and Haugen, D. (1968). Deriving power spectra from a three-component sonic anemometer. Journal of Applied Meteorology *7*, 827–837.

Kaimal, J., Wyngaard, J., Izumi, Y. and Coté, O. (1972). Spectral characteristics of surface-layer turbulence. Quarterly Journal of the Royal Meteorological Society *98*, 563–589.

Kaimal, J. C. and Finnigan, J. J. (1994). Atmospheric boundary layer flows: their structure and measurement. Oxford University Press. Textbook.

Klipp, C. L. and Mahrt, L. (2004). Flux–gradient relationship, self-correlation and intermittency in the stable boundary layer. Quarterly Journal of the Royal Meteorological Society *130*, 2087–2103.

Kolmogorov, A. N. (1941). The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers. Dokl. Akad. Nauk SSSR *30*, 299–303.

Konno, K. and Ohmachi, T. (1998). Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor. Bulletin of the Seismological Society of America *88*, 228–241.

Kowalski, A. S. and Serrano-Ortiz, P. (2007). On the relationship between the eddy covariance, the turbulent flux, and surface exchange for a trace gas such as CO2. Boundary-layer meteorology *124*, 129–141.

Kristensen, L. and Fitzjarrald, D. R. (1984). The effect of line averaging on scalar flux measurements with a sonic anemometer near the surface. Journal of Atmospheric and Oceanic Technology *1*, 138–146.

Kristensen, L., Mann, J., Oncley, S. and Wyngaard, J. (1997). How close is close enough when measuring scalar fluxes with displaced sensors? Journal of Atmospheric and Oceanic Technology *14*, 814–821.

Lay, D. C. (2010). Linear Algebra and Its Applications. 4th edition, Pearson. Textbook.

Lee, X. (1998). On micrometeorological observations of surface-air exchange over tall vegetation. Agricultural and Forest Meteorology *91*, 39–49.

Lee, X., Massman, W. and Law, B. E. (2006). Handbook of micrometeorology: a guide for surface flux measurement and analysis, vol. 29,. Springer. Textbook.

Lee, X. and Massman, W. J. (2011). A perspective on thirty years of the Webb, Pearman and Leuning density corrections. Boundary-layer meteorology *139*, 37–59.

Lenschow, D., Mann, J. and Kristensen, L. (1994). How long is long enough when measuring fluxes and other turbulence statistics? Journal of Atmospheric and Oceanic Technology *11*, 661–673.

Leuning, R. (2007). The correct form of the Webb, Pearman and Leuning equation for eddy fluxes of trace gases in steady and non-steady state, horizontally homogeneous flows. Boundary-layer meteorology *123*, 263–267.

Leuning, R. and Judd, M. J. (1996). The relative merits of open-and closed-path analysers for measurement of eddy fluxes. Global Change Biology *2*, 241–253.

Leuning, R., Van Gorsel, E., Massman, W. J. and Isaac, P. R. (2012). Reflections on the surface energy imbalance problem. Agricultural and Forest Meteorology *156*, 65–74.

Li, M., Babel, W., Tanaka, K. and Foken, T. (2012). Note on the application of planar-fit rotation for non-omnidirectional sonic anemometers. Atmospheric Measurement Techniques Discussions *5*, 7323–7340.

LI-7500 (2001). LI-7500 CO_2/H_2O Analyzer instruction manual. LI-COR Biosciences.

LI-7500 (2005). LI-7500 Open Path CO_2/H_2O Analyzer Brochure. LI-COR Biosciences.

Liu, H., Peters, G. and Foken, T. (2001). New equations for sonic temperature variance and buoyancy heat flux with an omnidirectional sonic anemometer. Boundary-Layer Meteorology *100*, 459–468.

Lüers, J. and Bareiss, J. (2011). Direct near-surface measurements of sensible heat fluxes in the Arctic tundra applying eddy covariance and laser scintillometry—the Arctic Turbulence Experiment 2006 on Svalbard (ARCTEX-2006). Theoretical and applied climatology *105*, 387–402.

Lüers, J., Westermann, S., Piel, K. and Boike, J. (2014). Annual CO 2 budget and seasonal CO 2 exchange signals at a High Arctic permafrost site on Spitsbergen, Svalbard archipelago. Biogeosciences Discussions *11*, 1535–1559.

Mahrt, L. (1986). On the shallow motion approximations. Journal of Atmospheric Sciences *43*, 1036–1044.

Mahrt, L. (1989). Intermittency of atmospheric turbulence. Journal of the Atmospheric Sciences *46*, 79–95.

Mahrt, L. (1999). Stratified atmospheric boundary layers. Boundary-Layer Meteorology *90*, 375–396.

Mahrt, L. (2010). Computing turbulent fluxes near the surface: needed improvements. Agricultural and forest meteorology *150*, 501–509.

Massman, W. and Lee, X. (2002). Eddy covariance flux corrections and uncertainties in long-term studies of carbon and energy exchanges. Agricultural and Forest Meteorology *113*, 121–144.

Massman, W. J. (2000). A simple method for estimating frequency response corrections for eddy covariance systems. Agricultural and Forest Meteorology *104*, 185–198.

Massman, W. J. (2001). Reply to comment by Rannik on "A simple method for estimating frequency response corrections for eddy covariance systems". Agricultural and Forest Meteorology *107*, 247–251.

Mauder, M., Cuntz, M., Drüe, C., Graf, A., Schmid, H., Schmidt, M. and Steinbrecher, R. (2013). A strategy for quality and uncertainty assessment of long-term eddy-covariance measurements. Agricultural and Forest Meteorology *169*, 122–135.

Mauder, M. and Foken, T. (2004). Documentation and Instruction Manual of the Eddy Covariance Software Package TK2. Universitat Bäyreuth, Abteilung Mikrometeorologie: Arbeitsergebnisse 26.

Mauder, M. and Foken, T. (2011). Documentation and Instruction Manual of the Eddy Covariance Software Package TK3. Universitat Bäyreuth, Abteilung Mikrometeorologie: Arbeitsergebnisse 46.

Mauder, M., Foken, T., Clement, R., Elbers, J., Eugster, W., Grünwald, T., Heusinkveld, B. and Kolle, O. (2008). Quality control of CarboEurope flux data? Part 2: Inter-comparison of eddy-covariance software. Biogeosciences *5*, 451–462.

Mauder, M., Liebethal, C., Göckede, M., Leps, J.-P., Beyrich, F. and Foken, T. (2006). Processing and quality control of flux data during LITFASS-2003. Boundary-Layer Meteorology *121*, 67–88.

McMillen, R. T. (1988). An eddy correlation technique with extended applicability to non-simple terrain. Boundary-Layer Meteorology *43*, 231–245.

Monin, A. S. and Obukhov, A. M. (1954). Basic laws of turbulent mixing in the surface layer of the atmosphere. Contrib. Geophys. Inst. Acad. Sci. USSR *151*, 163–187.

Moore, C. (1986). Frequency response corrections for eddy correlation systems. Boundary-Layer Meteorology *37*, 17–35.

Nordbo, A., Järvi, L. and Vesala, T. (2012). Revised eddy covariance flux calculation methodologies–effect on urban energy balance. Tellus B *64*.

Nowak, A. and Hodson, A. (2013). Hydrological response of a High-Arctic catchment to changing climate over the past 35 years: a case study of Bayelva watershed, Svalbard. Polar Research *32*.

Obukhov, A. (1971). Turbulence in an atmosphere with a non-uniform temperature. Boundary-layer meteorology *2*, 7–29.

Ohmura, A., Gilgen, H., Hegner, H., Müller, G., Wild, M., Dutton, E. G., Forgan, B., Fröhlich, C., Philipona, R., Heimo, A. et al. (1998). Baseline Surface Radiation Network (BSRN/WCRP): New precision radiometry for climate research. Bulletin of the American Meteorological Society *79*, 2115–2136.

Oncley, S. P., Foken, T., Vogt, R., Kohsiek, W., DeBruin, H., Bernhofer, C., Christen, A., Van Gorsel, E., Grantz, D., Feigenwinter, C. et al. (2007). The energy balance experiment EBEX-2000. Part I: overview and energy balance. Boundary-Layer Meteorology *123*, 1–28.

Panofsky, H., Tennekes, H., Lenschow, D. H. and Wyngaard, J. (1977). The characteristics of turbulent velocity components in the surface layer under convective conditions. Boundary-Layer Meteorology *11*, 355–361.

Paw, U., Baldocchi, D., Meyers, T. and Wilson, K. (2000). Correction of eddy-covariance measurements incorporating both advective effects and density fluxes. Boundary-Layer Meteorology *97*, 487–511.

Peel, M. C., Finlayson, B. L. and McMahon, T. A. (2007). Updated world map of the Köppen-Geiger climate classification. Hydrology and earth system sciences discussions *4*, 439–473.

Priestley, C. (1954). Convection from a large horizontal surface. Australian Journal of Physics *7*, 176–201.

Pumphrey, H. and Swarztrauber, P. (1985). DFFTPACK V1.0, A Double precision clone by H. C. Pumphrey of: FFTPACK version 4 april 1985 by P.N. Swarztrauber. Available as `dp.tgz` at URL: `http://www.netlib.org/fftpack/`. Accessed most recently on the 20.04.2015.

Rudels, B., Björk, G., Nilsson, J., Winsor, P., Lake, I. and Nohr, C. (2005). The interaction between waters from the Arctic Ocean and the Nordic Seas north of Fram Strait and along the East Greenland Current: results from the Arctic Ocean-02 Oden expedition. Journal of Marine Systems *55*, 1–30.

Røed, L. P. (2013). Atmospheres and Oceans on Computers: Fundamentals. Compendium.

Schmid, H. (1994). Source areas for scalars and scalar fluxes. Boundary-Layer Meteorology *67*, 293–318.

Schotanus, P., Nieuwstadt, F. and De Bruin, H. (1983). Temperature measurement with a sonic anemometer and its application to heat and moisture fluxes. Boundary-Layer Meteorology *26*, 81–93.

Siebicke, L., Hunner, M. and Foken, T. (2012). Aspects of CO2 advection measurements. Theoretical and Applied Climatology *109*, 109–131.

Sievers, J., Papakyriakou, T., Larsen, S. E., Jammet, M., Rysgaard, S., Sejr, M. and Sørensen, L. (2015). Estimating surface fluxes using eddy covariance and numerical ogive optimization. Atmospheric Chemistry and Physics *15*, 2081–2103.

Smith, S. L., Burgess, M. M., Riseborough, D. and Mark Nixon, F. (2005). Recent trends from Canadian permafrost thermal monitoring network sites. Permafrost and Periglacial Processes *16*, 19–30.

Sodemann, H. and Foken, T. (2005). Special characteristics of the temperature structure near the surface. Theoretical and applied climatology *80*, 81–89.

SR50 (2007). SR50 Sonic Ranging Sensor. Campbell Scientific Inc.

Stephens, G. L., Li, J., Wild, M., Clayson, C. A., Loeb, N., Kato, S., L'Ecuyer, T., Stackhouse Jr, P. W., Lebsock, M. and Andrews, T. (2012). An update on Earth's energy balance in light of the latest global observations. Nature Geoscience *5*, 691–696.

Stull, R. B. (1988). An Introduction to Boundary Layer Meteorology. Kluwer Academic Publishers. Textbook.

Sun, J. (2007). Tilt corrections over complex terrain and their implication for CO2 transport. Boundary-layer meteorology *124*, 143–159.

Sun, J., Esbensen, S. and Mahrt, L. (1995). Estimation of Surface Heat Flux. Journal of the Atmospheric Sciences *52*, 3162–3171.

Taylor, G. I. (1938). The spectrum of turbulence. Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences *164*, 476–490.

Thomas, C. and Foken, T. (2002). Re-evaluation of integral turbulence characteristics and their parameterisations. In 15th Conference on Boundary Layer and Turbulence pp. 15–19, Proceedings.

Tillman, J. (1972). The indirect determination of stability, heat and momentum fluxes in the atmospheric boundary layer from simple scalar variables during dry unstable conditions. Journal of Applied Meteorology *11*, 783–792.

TopoSvalbard (2015). TopoSvalbard: Interactive topograpichal Svalbard map portal produced by the Norwegian Polar Institute. URL: `toposvalbard.npolar.no`.

van Dijk, A. (2002). Extension to 3D of "The effect of line averaging on scalar flux measurements with a sonic anemometer near the surface" by Kristensen and Fitzjarrald. Journal of atmospheric and oceanic technology *19*, 80–82.

Van Dijk, A., Moene, A., De Bruin, H. et al. (2004). The principles of surface flux physics: theory, practice and description of the ECPACK library. Meteorology and Air Quality Group, Wageningen University, Wageningen, The Netherlands.

Vickers, D. and Mahrt, L. (1997). Quality control and flux sampling problems for tower and aircraft data. Journal of Atmospheric and Oceanic Technology *14*, 512–526.

Wallace, J. M. and Hobbs, P. V. (2006). Atmospheric science: an introductory survey, vol. 92,. Academic press.

Webb, E. K., Pearman, G. I. and Leuning, R. (1980). Correction of flux measurements for density effects due to heat and water vapour transfer. Quarterly Journal of the Royal Meteorological Society *106*, 85–100.

Westermann, S. (2010). Sensitivity of permafrost. PhD thesis, University of Heidelberg.

Westermann, S., Lüers, J., Langer, M., Piel, K. and Boike, J. (2009). The annual surface energy budget of a high-arctic permafrost site on Svalbard, Norway. The Cryosphere Discussions *3*, 631–680.

Wilczak, J. M. and Bedard Jr, A. J. (2004). A new turbulence microbarometer and its evaluation using the budget of horizontal heat flux. Journal of Atmospheric and Oceanic Technology *21*, 1170–1181.

Wilczak, J. M., Oncley, S. P. and Stage, S. A. (2001). Sonic anemometer tilt correction algorithms. Boundary-Layer Meteorology *99*, 127–150.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C. et al. (2002). Energy balance closure at FLUXNET sites. Agricultural and Forest Meteorology *113*, 223–243.

Winther, J.-G., Godtliebsen, F., Gerland, S. and Isachsen, P. E. (2002). Surface albedo in Ny-Ålesund, Svalbard: variability and trends during 1981–1997. Global and Planetary Change *32*, 127–139.

WMO (2008). Chapter 5, Measurement of Surface Wind. In Guide to Meteorological Intstruments and Methods of Observation. World Meteorological Organization.

Wyngaard, J. and Coté, O. (1972). Cospectral similarity in the atmospheric surface layer. Quarterly Journal of the Royal Meteorological Society *98*, 590–603.

Wyngaard, J., Coté, O. and Izumi, Y. (1971). Local free convection, similarity, and the budgets of shear stress and heat flux. Journal of the Atmospheric Sciences *28*, 1171–1182.

Wyngaard, J. C. (2010). Turbulence in the Atmosphere. Cambridge University Press. Textbook.

Zhang, J., Lee, X., Song, G. and Han, S. (2011). Pressure correction to the long-term measurement of carbon dioxide flux. Agricultural and forest meteorology *151*, 70–77.