APPLICABILITY OF WEIGHT-SHIFT MICROLIGHT AIRCRAFT FOR MEASURING THE TURBULENT EXCHANGE ABOVE COMPLEX TERRAIN

A dissertation submitted to the

FACULTY OF BIOLOGY, CHEMISTRY AND GEOSCIENCES

OF THE UNIVERSITY OF BAYREUTH, GERMANY

to attain the academic degree of

DR. RER. NAT.

presented by

STEFAN METZGER

Diplom Geoökologe

born September 13, 1980

in Schweinfurt, Bavaria, Germany

Bayreuth, 2013-04-15

APPLICABILITY OF WEIGHT-SHIFT MICROLIGHT AIRCRAFT FOR MEASURING THE TURBULENT EXCHANGE ABOVE COMPLEX TERRAIN

Supervisors

Prof. Dr. Thomas Foken

Prof. Dr. Klaus Butterbach-Bahl

Die vorliegende Arbeit wurde in der Zeit von September 2007 bis Oktober 2012 an der Universität Bayreuth am Lehrstuhl für Hydrologie, Abteilung Mikrometeorologie, unter Betreuung von Prof. Dr. Thomas Foken und Prof. Dr. Klaus Butterbach-Bahl angefertigt.

Vollständiger Abdruck der von der Fakultät für Biologie, Chemie und Geowissenschaften der Universität Bayreuth genehmigten Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.).

Promotionsgesuch eingereicht am: 2012-10-08

Wissenschaftliches Kolloquium am: 2013-04-15

Prüfungsausschuss:

Prof. Dr. Thomas Foken (Erstgutachter)

Prof. Dr. Jens Bange (Zweitgutachter)

Prof. Dr. Bernd Huwe (Vorsitzender)

Prof. Dr. Michael Hauhs

Prof. Dr. Andreas Held

Dekan: Prof. Dr. Beate Lohnert

The present dissertation was conducted from September 2007 to October 2012, as external doctoral candidate of the University of Bayreuth, Department of Micrometeorology, Bayreuth, Germany.

The research aircraft and the majority of the infrastructure were provided by the Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research, Atmospheric Environmental Research, Garmisch-Partenkirchen, Germany.

The bulk of preparation, implementation and analysis of the experiments were performed at the Chinese Academy of Sciences, Institute of Atmospheric Physics, State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry, Beijing, China.

The doctoral thesis was completed while employed at the National Ecological Observatory Network, Fundamental Instrument Unit, Boulder, CO, U.S.A.

This doctoral thesis was performed under stipend funding by the German Academic Exchange Service, Helmholtz Association of German Research Centers, China Scholarship Council and the European Union under the Science and Technology Fellowship China.

The flight in Inner Mongolia was funded by the German Research Foundation, research group 536 “Matter fluxes in grasslands of Inner Mongolia as influenced by stocking rate”, and the National Natural Science Foundation of China (grant number 41021004).

The publications were funded by the German Research Foundation and Open Access Publishing Fund of the Karlsruhe Institute of Technology / Helmholtz Association.

This material is based upon work supported by the National Science Foundation under the grant DBI-0752017. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Contents

Acknowledgements

In 2007 I set out on a journey to combine my curiosity about the orient, family, and enthusiasm for atmospheric science… Now, about five years later I feel equally at home in Germany, China and the U.S., we are speaking Chinese at home, and I might soon be a doctor…

I want to use this opportunity to thank the wonderful persons who have made possible this development. From the very start Benjamin Wolf pointed me to Garmisch-Partenkirchen, and you always had a sympathetic ear – thank you! I am much obliged to my supervisors Thomas Foken and Klaus Butterbach-Bahl, for your trust in my abilities, your scientific and personal guidance, and the flexibility you provided me on this path. Special recognition goes to Wolfgang Junkermann, for your unfailing patience when introducing the aircraft to me, and your outstanding performance of the hair-rising flight patterns. My gratefulness to Xunhua Zheng for hosting me and providing indispensable infrastructure and negotiation skills. Further I am thankful to my co-authors, Frank Beyrich, Matthias Mauder, Frank Neidl, Klaus Schäfer, Hans-Peter Schmid, Baltasar Trancon y Widemann, and Sebastian Wieneke, for your pivotal stimuli and your contributions to the manuscripts. My appreciation to many other individuals who found innumerable ways to inspire and support this journey, including; Wolfgang Babel, Jens Bange, Frauke Barthold, Tobias Biermann, Benjamin Blank, Timothy Brown, Josef-Michael Burger, Yimin Cao, Dao Charuchittipan, Lifang Chen, Yong Chen, Michael Dannemann, Jia Deng, Francesco di Maio, Sofia Dos Santos Mendes, Elias, Leyla and Roland Felgenträger, Edwin Haas, Shenghui Han, Jordan Hixson, Natascha Kljun, Alexandra Lehmann, Jens-Peter Leps, Chunyan Liu, Ronghua Liu, Henry Loescher, Tiina Markkanen, Alan Page, Raül Ramos, Matthias Reiche, Jakob Reitberger, Charlotte Roehm, David Schaffrath, Hans-Eckhart Scheel, Andreas Schmidtler, Clemens Smolders, Christian Sperber, Rainer Steinbrecher, Jeffrey Taylor, Aline van den Kroonenberg, Ulrich Weisensee, Georg Wichmann, Hongkai Zhang, Yuandi Zhu, Rose Zuurbier.

My heartfelt thanks to my wife Liang Sun and to our families. Your steady support and unswerving faith were the building blocks on which I could develop my potential.

List of manuscripts

The dissertation is presented in cumulative form. It consists of three individual manuscripts. Two manuscripts have been published in the peer-reviewed journal Atmospheric Measurement Techniques, of which the first one also includes a 14 page supplement. A third manuscript has been published in the peer-reviewed journal Biogeosciences.

Published manuscripts

Metzger, S., Junkermann, W., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Corrigendum to "Measuring the 3-D wind vector with a weight-shift microlight aircraft" published in Atmos. Meas. Tech., 4, 1421–1444, 2011, Atmos. Meas. Tech., 4, 1515-1539, doi:10.5194/amt-4-1515-2011, 2011. Includes one supplement. ***Publisher's Note:*** *Due to mistakes on the publisher's side it was necessary to revise the original article Atmos. Meas. Tech., 4, 1421–1444, 2011 in this Corrigendum. Please use this Corrigendum as the main document.*

Metzger, S., Junkermann, W., Mauder, M., Beyrich, F., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Eddy-covariance flux measurements with a weight-shift microlight aircraft, Atmos. Meas. Tech., 5, 1699-1717, doi:10.5194/amt-5-1699-2012, 2012.

Metzger, S., Junkermann, W., Mauder, M., Butterbach-Bahl, K., Trancón y Widemann, B., Neidl, F., Schäfer, K., Wieneke, S., Zheng, X. H., Schmid, H. P., and Foken, T.: Spatially explicit regionalization of airborne flux measurements using environmental response functions, Biogeosciences, 10, 2193-2217, doi:10.5194/bg-10-2193-2013, 2013.

Publications not included in this thesis

Peer-reviewed

Sun, F., Ma, Y., Li, M., Ma, W., Tian, H., and Metzger, S.: Boundary layer effects above a Himalayan valley near Mount Everest, Geophys. Res. Lett., 34, L08808, doi:10.1029/2007gl029484, 2007.

Eigenmann, R., Metzger, S., and Foken, T.: Generation of free convection due to changes of the local circulation system, Atmos. Chem. Phys., 9, 8587-8600, doi:10.5194/acp-9-8587-2009, 2009.

Not peer-reviewed

Metzger, S., Ma, Y. M., Markkanen, T., Göckede, M., Li, M. S., and Foken, T.: Quality assessment of Tibetan Plateau eddy-covariance measurements utilizing footprint modeling, Advances in Earth Sciences, 21, 1260-1267, 2006.

Metzger, S., and Foken, T.: COPS experiment – Convective and orographically induced precipitation study, 1 June 2007 – 31 August 2007 – Documentation, Universität Bayreuth, Abteilung Mikrometeorologie, Bayreuth, Germany, 72 pp. ISSN:1614-8924, 2007.

Eigenmann, R., Metzger, S., and Foken, T.: Generation of free convection due to changes of the local circulation system, Atmos. Chem. Phys. Discuss., 9, 11367-11411, doi:10.5194/acpd-9-11367-2009, 2009.

Metzger, S., Junkermann, W., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Measuring the 3-D wind vector with a weight-shift microlight aircraft, Atmos. Meas. Tech. Discuss., 4, 1303-1370, doi:10.5194/amtd-4-1303-2011, 2011.

Broadwell, I., Jenkins, G., and Metzger, S.: Three perspectives: careers in China, Phys. World, 24, 42-43, 2011.

Metzger, S., Junkermann, W., Mauder, M., Beyrich, F., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Eddy-covariance flux measurements with a weight-shift microlight aircraft, Atmos. Meas. Tech. Discuss., 5, 2591-2643, doi:10.5194/amtd-5-2591-2012, 2012.

Summary

The possibility to reliably observe the exchange of heat and moisture between the land surface and the atmosphere is vital to our understanding of the regional and global cycling of energy and water. While ground-based flux measurements can be made continuously for long periods, they only represent a small landscape unit. On the other hand, aircraft-based measurements have the ability to directly measure the exchange over large areas. Especially over heterogeneous landscapes the spatio-temporal characteristics of both approaches complement each other. However, complex terrestrial ecosystems are sparsely investigated to date, in particular over topographically structured terrain. This can be attributed to; (i) limitations in the description of boundary layer processes over non-homogenous terrain, and (ii) a lack of applicable measurement platforms and techniques to study these processes. In pursue of a resolution strategy, this dissertation investigates the applicability of weight-shift microlight aircraft (WSMA) to gain new insights in the spatial variability of heat and moisture exchange over complex terrain.

WSMA are comparatively cheap in procurement and maintenance, and their unique structure provides exceptional transportability and climb rate. These structural features qualify the WSMA for terrain-following flight over complex and inaccessible terrain, but potentially influence measurements aboard the aircraft. In this dissertation a WSMA with a scientific payload enabling fast measurements of the 3D wind, temperature, water vapor concentration, position, and the radiative flux is used to;

(i) quantify the WSMA wind measurement uncertainty. A novel time-domain procedure is developed, which improves the accuracy of the WSMA wind measurement by 63% for the horizontal- and 72% for the vertical wind components. The resulting precisions are ± 0.09 m s^{-1} and ± 0.04 m s^{-1}, and the agreement with ground-based measurements is in the order of ± 0.4 m s^{-1} and ± 0.3 m s^{-1} (root mean square deviation), respectively.

(ii) quantify the WSMA eddy-covariance flux measurement uncertainty. From uncertainty propagation the smallest resolvable changes in friction velocity (0.02 m s^{-1}), and sensible- (5 W m^{-2}) and latent (3 W m^{-2}) heat flux are estimated. In comparison to tower measurements, the WSMA observes higher fluxes (17–21%). The differences are not statistically significant, and can be explained by the tower setup and non-propagating eddies.

(iii) spatially resolve and regionalize the heat and moisture exchange above a complex landscape. Wavelet decomposition of the turbulence data is used to yield a flux observation each 90 m along the flight path. For each flux observation the biophysical surface properties in the flux footprint are determined. An environmental response function between the flux observations and biophysical and meteorological drivers is then inferred using a machine learning technique. This function is used to produce regional maps of the heat and moisture exchange to an accuracy of ≤18% and a precision of ≤5% for individual land covers.

Hence this dissertation provides the necessary basis for using WSMA to investigate the mechanisms of turbulent exchange over heterogeneous and topographically structured terrain. Moreover, the developed algorithms are generally applicable to (i) partitioning flux uncertainty and environmental variability, (ii) extrapolating flux measurements, (iii) assessing the spatial representativeness of long-term tower flux measurements, and (iv) designing, constraining and evaluating flux algorithms for remote sensing and numerical modeling applications.

Zusammenfassung

Die experimentelle Untersuchung der Austauschströme von Wärme und Feuchte zwischen Landoberfläche und Atmosphäre trägt ausschlaggebend zu unserem Verständnis der regionalen und globalen Kreisläufe von Energie und Wasser bei. Bodengestütze Flussmessungen ermöglichen die Erfassungen langer Zeitreihen, repräsentieren aber lediglich kleine Landschaftsausschnitte. Auf der anderen Seite kann der Austausch über größeren Gebieten direkt aus flugzeuggestützte Messungen bestimmt werden. Insbesondere über heterogenem Gelände ergänzen sich diese raum-zeitlichen Eigenschaften beider Messverfahren. Die Wärme und Feuchteströme über komplexen terrestrischen Ökosystemen sind bis dato spärlich untersucht, insbesondere über topografisch geprägtem Gelände. Dies ist zurückzuführen auf; (i) die über heterogenem Gelände eingeschränkte Gültigkeit der vorhandenen Formulierungen von Austauschprozessen, und (ii) einen Mangel an geeigneten Messplattformen und -techniken zur weitergehenden Untersuchung dieser Prozesse. Die vorliegend Dissertation verfolgt eine mögliche Strategie zur Lösung dieser Zwangslage. Dazu wird die Eignung von schwerkraftgesteuerten Ultraleichtflugzeugen (WSMA) untersucht, neue Einsichten in die räumliche Variabilität von Wärme- und Feuchteaustausch über komplexem Gelände zu gewinnen.

WSMA sind vergleichsweise preiswert in Anschaffung und Unterhalt, und ihre charakteristische Bauweise resultiert in ausgezeichneter Transportfähigkeit und exzellenten Steigeigenschaften. Dies prädestiniert WSMA für dem Gelände folgende Flugmuster über komplexen und schwer zugänglichen Regionen. Zugleich muss aber ausgeschlossen werden, dass vorgenannte Eigenschaften die Messungen an Bord eines WSMA verfälschen. Im Rahmen dieser Dissertation wird ein mit Messinstrumenten zur schnellen Erfassung des 3D Windfeldes, der Lufttemperatur, des Wasserdampfgehalts, der Position und der Strahlungsflüsse ausgestattetes WSMA genutzt um:

(i) den Eingangsfehler in der WSMA Windmessung zu quantifizieren. Ein neuartiges Verfahren im Zeitbereich wird entwickelt, welches die Genauigkeit der WSMA Windmessung um 63% für den Horizontalwind, und um 72% für den Vertikalwind verbessert. Die resultierende Präzision ist $\pm 0.09\ m\ s^{-1}$ beziehungsweise $\pm 0.04\ m\ s^{-1}$, und die Übereinstimmung mit bodengestützen Messungen liegt bei $\pm 0.4\ m\ s^{-1}$ beziehungsweise $\pm 0.3\ m\ s^{-1}$ (Wurzel aus dem mittleren quadratischen Fehler).

(ii) den Eingangsfehler in der WSMA Eddy-Kovarianz Flussmessung zu quantifizieren. Anhand von Fehlerfortpflanzung wird die Auflösung der Schubspannungsgeschwindigkeit ($0.02\ m\ s^{-1}$), dem fühlbaren- ($5\ W\ m^{-2}$) und dem latenten Wärmestrom ($3\ W\ m^{-2}$) bestimmt. Im Vergleich zu Turmmessungen misst das WSMA um 17–21% höhere Flüsse. Die Unterschiede sind nicht statistisch signifikant, und können an Hand der Sensoreigenschaften am Turm und stationären Wirbeln erklärt werden.

(iii) den Wärme- und Feuchteaustausch über einer komplexen Landschaft räumlich aufzulösen und zu regionalisieren. Die Wavelet Zerlegung der Turbulenzdaten ermöglicht eine räumliche Auflösung der Flussmessung von 90 m entlang des Flugpfades. Für eine jede Messungen werden die biophysikalischen Oberflächeneigenschaften im jeweiligen Quellgebiet bestimmt. Mit Hilfe von stochastischen Entscheidungsbäumen wird eine Antwortfunktion zwischen den gemessenen Flüssen und ihren biophysikalischen und meteorologischen Treibern abgeleitet.

Diese Antwortfunktion wird genutzt um regionale Karten des Wärme- und Feuchteaustausches zu generieren. Hierbei liegt die Genauigkeit bei ≤18%, und die Präzision bei ≤5% für einzelne Landnutzungsklassen.

Die vorliegende Dissertation schafft die wissenschaftliche Basis um turbulente Austauschvorgänge über komplexem und topografisch geprägten Gelände mit WSMA untersuchen zu können. Des Weiteren sind die in der Dissertation entwickelten Algorithmen generell anwendbar um (i) Fehler und natürliche Variabilität in der Flussmessung zu partitionieren, (ii) Flussmessungen zu extrapolieren, (iii) die räumliche Repräsentativität von Langzeit-Flussmessungen zu bestimmen, und (iv) in der Fernerkundung und in numerischen Modellen genutzte Parametrisierungen für den turbulenten Austausch zu entwerfen, zu beschränken, und zu evaluieren.

1 Introduction

Climate change and its impact on natural resources is one of the most prominent global issues facing society, and has increasingly received attention over the past decades (e.g., IPCC, 1990, 1995, 2001, 2007). In a tremendous effort to improve our understanding of the underlying mechanisms, sensor networks which continuously monitor the environment are being established and operated from taxpayer money (e.g., Baldocchi et al., 2001; Hopkin, 2006; Yu et al., 2006). All of these observation platforms have in common the overarching goal of deducing appropriate mitigation strategies through enabling to forecast the effects of climate change. For this purpose the relationships between environmental drivers and responses are explored, e.g. through rigorous quantification of the pools and fluxes of energy, water, carbon and nitrogen.

1.1 The eddy-covariance method

The exchange (or flux) of heat, water and other air constituents between the earth's surface and the atmosphere is mainly governed by turbulent transport, both convective and mechanical. Buoyancy as well as shear stress results in a turbulent wind field for most of the day (e.g., Stull, 1988). The wind field mediates the fluxes between the earth's surface and the atmosphere, and its properties are used in the eddy-covariance (EC, a list of all notation can be found on p. 32 ff.) method to quantify their magnitude, direction, and timing (e.g., Aubinet et al., 2012). Moreover, the EC method combines the ability to directly and continuously measure the surface exchange with minimal intrusion. For these reasons, and despite several simplifications, the EC method has emerged as the *de facto* standard for continuous flux observations in many sensor networks (e.g., AmeriFlux, Fluxnet-Canada, Carboeurope).

The EC technique is based on mass conservation and makes use of the Reynolds decomposition (isolation of mean and fluctuating part) of relevant terms in the Navier-Stokes equation (e.g., Foken, 2008a; Stull, 1988). Among other simplifications, the Boussinesq-approximation (e.g., Foken et al., 2012) is particularly important because it permits flux measurements under convective conditions. This is achieved through only neglecting horizontal density gradients, but considering the vertical density gradient due to gravity (assumption of a local, incompressible atmosphere). Now the total flux F into or out of an ecosystem can be expressed as;

$$F = \underbrace{\int_0^z \frac{\partial \bar{s}}{\partial t}}_{\text{I}} + \underbrace{\int_0^z \frac{\partial \overline{u's'}}{\partial x}}_{\text{II}} + \underbrace{\int_0^z \frac{\partial \overline{v's'}}{\partial y}}_{\text{III}} + \underbrace{\int_0^z \frac{\partial \overline{w's'}}{\partial z}}_{\text{IV}} + \underbrace{\int_0^z \frac{\partial \bar{u}\bar{s}}{\partial x}}_{\text{V}} + \underbrace{\int_0^z \frac{\partial \bar{v}\bar{s}}{\partial y}}_{\text{VI}} + \underbrace{\int_0^z \frac{\partial \bar{w}\bar{s}}{\partial z}}_{\text{VII}}, \tag{1}$$

with overbars indicating an average and primes denoting turbulent fluctuations (i.e., deviations from the average). Here s is a scalar quantity such as temperature or water vapor concentration; u, v and w

are along-, cross-, and vertical wind speeds with respect to the Cartesian coordinates x, y, and z; t is time, and z is the measurement height.

Term I in Eq. (1) represents the tendency of s in the vertical column below the sensor, i.e. storage. Terms II–IV represent the turbulent flux divergence, and terms V–VII represent advection through the layer between the surface and sensor. For low or sparse canopies it is assumed that this layer is well-mixed, and that steady state conditions prevail during the averaging period ($\partial/\partial t = 0$), which cancels term I from Eq. (1). Assuming horizontal homogeneity ($\partial/\partial x = 0$, $\partial/\partial y = 0$) implies equal horizontal inflow and outflow at opposite faces of a control volume (e.g., Finnigan et al., 2003), and cancels horizontal flux divergence (terms II–III) and advection (terms V–VI) from Eq. (1). Furthermore, neglecting large-scale subsidence or convection ($\overline{w} \equiv 0$) cancels term VII from Eq. (1), and hence only term IV remains. Using dimensional analysis and similarity numbers it can be shown that the eddy covariance $\overline{w's'}$ is constant with height in the atmospheric surface layer to within approximately 10% (e.g., Foken, 2008a). When all above conditions hold true, the total flux is then equal to the EC measurement;

$$F = \overline{w's'}. \tag{2}$$

1.2 Effects of complex terrain

As long as the one-dimensional transport Eq. (2) is valid, a single ground-based EC measurement can represent the spatially averaged flux to or from an underlying ecosystem. In reality however, every ecosystem is inhomogeneous to some extent, and a graduation from less to more complex surfaces is much more applicable. Here and in the following, complex refers not only to the structure, function and spatio-temporal distribution of the surface sources and sinks, but also to surface topography. Consequently, real-world measurement conditions principally invalidate the assumption of one-dimensional transport in various forms and degrees (e.g., Finnigan, 2008; Foken et al., 2011; Leuning et al., 2012; Mahrt, 2010; Mauder et al., 2007b). Aside from sensor limitations (e.g., Dellwik et al., 2010; Kochendorfer et al., 2012), it is this oversimplification that is held responsible for the systematic energy imbalance that is frequently observed across EC sensor networks (e.g., Foken et al., 2011; Leuning et al., 2012; Wilson et al., 2002). The lack of agreement between the sum of the turbulent heat fluxes and the available energy also challenges the validity of EC flux measurements of climate effective trace gases (Ruppert et al., 2006).

Kaminski et al. (2012) find that EC flux measurements efficiently constrain modeling approaches for relatively homogeneous situations, but are not robust against unknown complexity. Failure to explicitly consider spatial representativeness increases the uncertainty in the observed spatio-temporal flux structure of EC sensor networks. This effect is also known as location bias (Schmid and Lloyd, 1999). Such bias can be introduced by preferential measurement locations, e.g., on hill tops rather than in valleys (Finnigan, 2008), or in dryer rather than in moister areas (Desjardins et al., 1997). Also, the source area of an EC flux measurement varies in space and time, as a function of atmospheric stability and the wind field. Over complex terrain, the spatial representativeness of an EC measurement thus varies from diurnal over seasonal to inter-annual time scales (Göckede et al.,

2008). Biases are introduced when the source area changes systematically over surfaces with distinctly different biophysical properties and source/sink behavior (e.g., Chen et al., 2012).

1.3 Motivation

Data from EC sensor networks are being used to parameterize, constrain and evaluate land surface schemes in climate models (e.g., Williams et al., 2009). In this way biases in EC flux observations propagate directly into the resultant data products (Leuning et al., 2012). Improved characterization of the uncertainties in the EC observations would significantly foster our ability to understand, model, forecast and eventually mitigate environmental change (Jung et al., 2011). The motivation of this dissertation is to provide an airborne measurement platform which enables studying some of the mechanisms that lead to systematically biased flux estimates from EC sensor networks.

Despite most EC flux observations are performed at ground-based sites in the time domain, much of turbulence theory is conceptually formulated in the spatial domain (e.g., Kaimal and Finnigan, 1994; Lenschow et al., 1980; Mahrt, 2010; Panofsky and Dutton, 1984). As compared to time domain measurements, EC flux measurements in the spatial domain do not rely on the passive transport of atmospheric eddies with the mean wind (Taylor, 1915). E.g., the speed of a research aircraft is typically much faster than the wind speed. Through active propulsion in space, fast response airborne measurements more completely capture the true spatial variation of the turbulent exchange over heterogeneous surfaces (e.g., Kustas et al., 2006). Thus, airborne soundings cover a comparatively larger range in the state space of environmental drivers and responses. If sufficient measurement accuracy is warranted, this results in a high signal to noise ratio in the airborne observations (Lenschow and Sun, 2007).

However, spatial sampling does not *per se* obsolete the assumptions in the one-dimensional transport Eq. (2). Yet, compared to their ground-based pendants, aircraft EC measurements tend to better fulfill some of the assumptions that justify cancellation of the respective terms from Eq. (1). Because of their short duration, change in the storage term I is likely small (steady state conditions). It is hypothesized that, during unstable stratification, surface heterogeneity induces non-propagating eddies (NPE, term VII), which in turn cause horizontal compensatory flows (terms II–III, V–VI, e.g., Foken et al., 2011; Mahrt, 2010). These NPEs are explicitly resolved in a spatial sample, but not in a temporal one. Moreover, each of terms II–III, V–VII is more likely approaching zero for an instantaneous spatial average than for a stationary temporal average (Steinfeld et al., 2007). Lastly, the effect of vertical flux divergence (term IV) on the airborne EC measurement is negligible, as long as the soundings are performed close to the surface (e.g., Isaac et al., 2004). If higher flight altitudes are required, term IV can also be approximated from convective boundary layer (CBL) scaling (e.g., Deardorff, 1974), stacked flight patterns (e.g., Betts et al., 1990), or from a conservation approach (Bange et al., 2006).

Above properties make airborne flux data particularly valuable for the analysis of atmospheric transport mechanisms (e.g., Hiyama et al., 2007; Mauder et al., 2007a; Strunin et al., 2004), and facilitates the derivation of scaling laws (e.g., Lenschow and Stankov, 1986; Mahrt, 2000). This motivates using aircraft observations of the surface-atmosphere exchange also for separating uncertainties in EC flux measurements due to;

(i) sampling errors;

(ii) direct effects of surface heterogeneity, such as location bias through variable source/sink behavior and surface roughness, and;

(iii) indirect effects of surface heterogeneity, such as transport by NPEs.

However, to date manned platforms, such as fixed wing aircraft and helicopters, are expensive to operate or not applicable in settings such as remote areas beyond the range of an airfield. Unmanned aerial vehicles on the other hand provide mobility, yet do not allow a comprehensive sensor package due to payload restrictions (e.g., overview in Dias et al., 2012; Egger et al., 2002; Hobbs et al., 2002; Martin et al., 2011; Thomas et al., 2012). In order to enable the relation of airborne EC flux measurements to surface properties, aircraft are bound to fly at low and constant altitude above ground. To follow topographically structured terrain, an aircraft must therefore possess a high ratio of climb rate to airspeed, which only few airborne platforms provide. This dissertation intends to make accessible an alternative airborne measurement platform which combines comprehensive, terrain-following observations over complex terrain with a minimal demand in cost, transport and infrastructure. After successfully applying a weight-shift microlight aircraft (WSMA) to aerosol and radiation transfer studies (e.g., Junkermann, 2001; Junkermann, 2005), it was hypothesized that WSMA can also fulfill above requirements. The possibility of low-level, terrain-following EC flux measurements with WSMA is explored in this dissertation, which can provide a viable addition to existing airborne measurement platforms. Such development could aid overcoming the lack of applicable measurement platforms and techniques to study thus far poorly understood CBL processes over complex terrain.

1.4 Objectives of this thesis

The overarching goal of this dissertation is to evaluate the suitability of WSMA for gaining new insights in the spatial variability of heat and moisture exchange above complex terrestrial surfaces. Three concrete objectives are assigned to this goal;

(i) assessing the suitability of WSMA based 3D wind vector measurement for EC applications;

(ii) quantifying the uncertainty in WSMA based measurements of turbulence statistics and EC flux, and;

(iii) demonstrating the usefulness of WSMA based EC flux measurements to spatially resolve the land-atmosphere exchange above complex and not readily accessible terrain.

These objectives are addressed by three individual publications as well as supplementary materials presented in Appendices B–E of this thesis.

Metzger et al. (2011, Appendix B) address objective (i) by developing, establishing and evaluating a generalized calibration strategy for the wind measurement from airborne platforms. This strategy consists of an expanded algorithmic description of the wind measurement (Supplement to Metzger et al., 2011, Appendix C), in combination with a specific sequence of flight patterns used for calibration. The principal aim of this paper is to enable accurate wind measurements even during low-level, terrain-following flight, such as required for EC soundings over complex terrain. Hence, the

possibility to dynamically offset the effect of pilot input on the wind measurement is sought. It is shown that this can be realized through consideration of the lift coefficient, i.e. the ratio of aircraft gravitational to propulsion forces, in the wind computation. All sources of uncertainty are quantitatively propagated through the algorithmic description, and the WSMA wind measurement is evaluated against independent reference measurements. The findings emphasize that WSMA are capable of accurately measuring the 3D wind vector. Hence the necessary basis is provided for the study of precision and spectral quality of the wind measurement, which is prerequisite for reliable EC flux measurements.

Metzger et al. (2012, Appendix D) address objective (ii) by, first of all, quantifying the precision of the wind measurement from WSMA, the lynchpin of flux calculations from aircraft. From here, the smallest resolvable changes in friction velocity, and sensible- (H) and latent (LE) heat flux are estimated. Secondly, measurements of wind, temperature, humidity and respective fluxes are compared between independent ground-based reference installations and the WSMA. Some differences are observed, which could be related to inconsistencies in the ground-based installations, and to their different abilities to capture flux contributions from NPEs. These findings encourage the use of WSMA as a low cost and highly versatile flux measurement platform. In combination with its high transportability and unique flight characteristics, the WSMA is thus well suited to study land-atmosphere interactions above complex terrain.

The objective (iii) is addressed in Metzger et al. (2013, Appendix E). In this paper, low-level EC measurements from WSMA are used to characterize the exchange of heat and moisture over a complex landscape in Inner Mongolia, P.R. China. From CBL scaling it is found that the vertical flux gradients below the terrain-following flights satisfy the surface layer definition. Consequently the WSMA based EC measurements can be interpreted as surface fluxes. Wavelet decomposition of the WSMA turbulence data is then combined with footprint modeling and a machine learning technique to infer environmental response functions. From these response functions high resolution maps of the heat and moisture exchange over the entire Xilin River Catchment are extrapolated. These maps are then summarized for each land cover type, providing information on the individual source strength and spatial variability. This study emphasizes the potential of WSMA based EC flux measurements above complex terrain to (a) provide high-resolution inventories of the land-atmosphere exchange, (b) advancing the location bias treatment of ground-based flux measurements from diagnostic assessment (e.g., Chen et al., 2011) to prognostic transfer functions, and consequently (c) separating the effects of sampling errors and surface heterogeneity in long-term sensor networks.

The algorithms developed in the course of this thesis are intended for future use beyond the data collected and analyzed here. For this purpose all algorithms are implemented in a software package in the R programming language (R Development Core Team, 2012). This software package is currently available upon request, and will be released to The Comprehensive R Archive Network (http://cran.r-project.org/) in the near future.

2 Experiments and data

The results presented in this thesis are based on three extensive flight campaigns with the WSMA. In Sect. 2.1 the WSMA and its scientific equipment are briefly introduced. The publications presented in Appendices B and C are based on data that were collected during two flight campaigns in Germany (Sects. 2.1, 2.3). These campaigns were conducted with the sole purpose to explore the possibility of wind and EC flux measurements with the WSMA. The dataset used for the publication in Appendix D was collected during a third flight campaign in China (Sect. 2.4). This campaign was performed as integral part of the Sino-German research collaboration MAGIM (DFG research group 536). All flight campaigns were initiated by K. Butterbach-Bahl, I developed the measurement strategies, and W. Junkermann performed the majority of the research flights.

2.1 The weight-shift microlight aircraft

According to the safety and regulatory standards of the European Civil Aviation Conference, microlight aircraft are defined as aircraft with a maximum stall speed of 65 km h^{-1} and a take-off mass of no more than 450 kg. Figure 1 shows the weight-shift microlight research aircraft D–MIFU. It consists of two distinct parts, the wing and the trike (the unit hung below the wing, containing pilot, engine and the majority of the scientific equipment). The weight-shift control system is enabled by the pilot's direct application of pitching or rolling moments to the wing via the basebar. Counterbalance is provided by the mass of the trike unit suspended below the wing. Simple procedures for certification of installations on an open aircraft allow a wide spectrum of applications as well as the flexible installation of scientific equipment. At an operational airspeed of ≈100 km h^{-1}, D–MIFU can carry a maximum of 80 kg scientific payload from 15 m a.g.l. (above ground level) to 4000 m a.s.l. (above sea level). The full performance characteristics can be found in Junkermann (2001), and additional technical detail is provided in Metzger et al. (2011, Appendix B).

The structure of a WSMA differs from common fixed-wing aircraft, which provides exceptional transportability and climb rate, and qualifies it for applications in complex and inaccessible terrain. However, these structural features might also expose the trike-based wind measurement to variable distortion in the flow field around the wing. Firstly, the trike, i.e. the turbulence measurement platform, is free to rotate in pitch and roll against the wing. Secondly, the wing deforms aeroelastically with aircraft trim, which changes its aerodynamic properties. These characteristics complicate the accurate and precise measurement of the true 3D wind vector, and consequently the dependable measurement of EC fluxes.

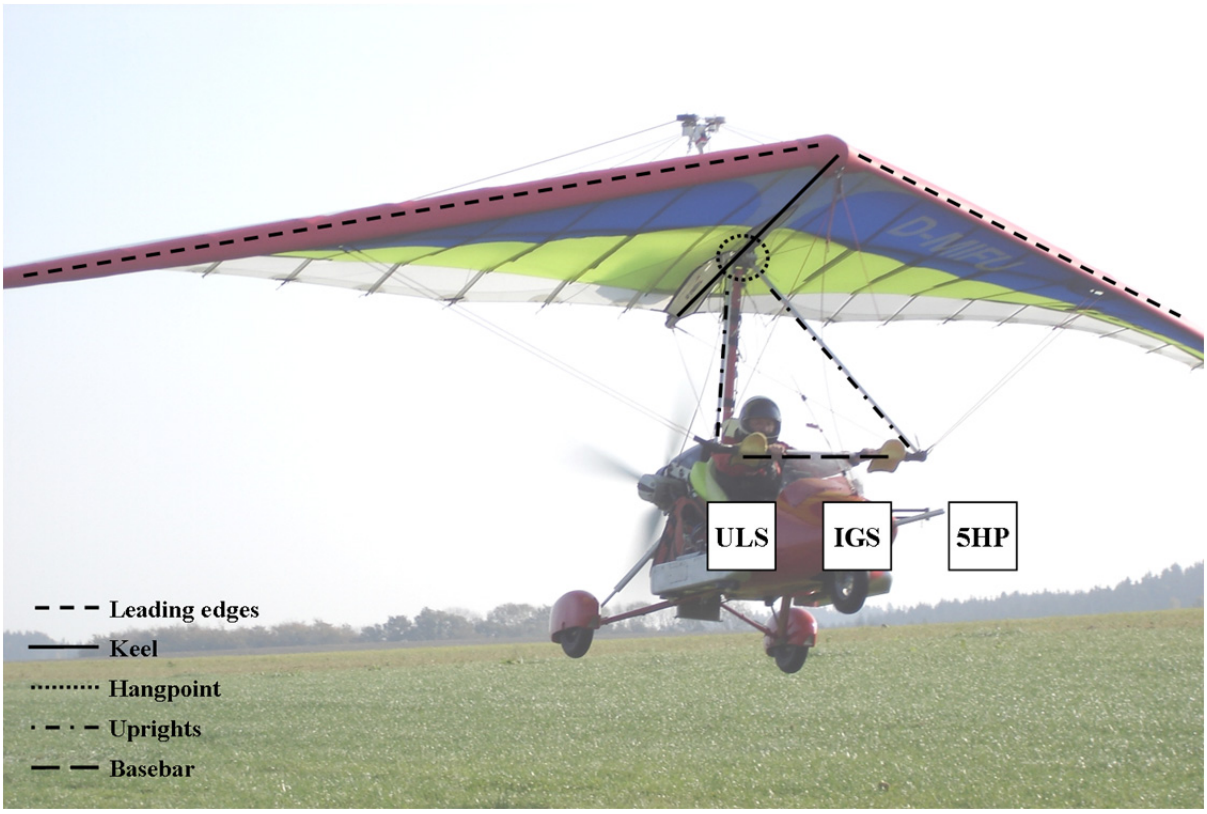


Fig. 1. Weight-shift microlight research aircraft D-MIFU, aircraft structural features are highlighted by dash-dotted lines. Sensor locations of wind-measuring five hole probe (5HP), inertial measurement and global positioning system (IGS, inside aircraft nose) and universal laser altitude sensor (ULS, below pilot seat) are indicated. Figure modified after Metzger et al. (2011, Appendix B).

The WSMA is equipped with fast response instruments that enable capturing the turbulent scales of atmospheric motion. A detailed description of the installation points, models and characteristics of the deployed sensors and data acquisition is given in Metzger et al. (2011, Appendix B). In short, most variables are sampled at 100 Hz and are block-averaged and stored at 10 Hz, yielding a horizontal resolution of approximately 2.5 m. To conduct fast wind measurements, the WSMA is outfitted with a combination of inertial measurement and global positioning system (IGS, RT3102, Oxford Technical Solutions, Upper Heyford, England), and a five-hole pressure probe (5HP, in-house development). The principle is to resolve the meteorological wind vector from the vector difference of the aircraft's inertial velocity (captured by the IGS) and the wind vector relative to the aircraft (captured by the 5HP, Supplement to Metzger et al., 2011, Appendix C). Additional acceleration measurements (ADXL330, Analog Devices, Inc., Norwood, U.S.A.) were installed in the hang point of trike and wing, and in the 5HP. These enable investigating potential influences on the wind measurement from WSMA engine or propeller resonance, or from the natural frequencies of the trike and the wing. On the underside of the 5HP, air temperature is measured with a 50 µm thermocouple (CHAL-002, OMEGA Engineering, Inc. Stamford, U.S.A.). An open path infrared gas analyzer (IRGA, OP2, ADC Bioscientific, Great Amwell, UK) to the port side of the 5HP is used to measure

the concentration of water vapor. The instrument response of both the thermocouple and the IRGA is 50 Hz. In addition, a slow (2 Hz instrument response) humidity reference from a TP3 dew point mirror (Meteolabor AG, Wetzikon, Switzerland) is stored at ≤0.1 Hz. A third-order Savitzky-Golay complementary filter (Chen et al., 2004) is used to correct drift of the IRGA due to pressure changes with altitude. This filter bases the humidity fluctuations measured by the IRGA on the slow dew point mirror reference. A filter window size of 13.9 s or ≈350 m maximizes the integral over the humidity power spectrum, and is used to correct the measurements. Additional slow measurements (≤0.1 Hz) of surface temperature (CT infrared thermometer, Optris GmbH, Berlin, Germany) and downwelling short wave radiation (LI–200 SZ, LI–COR Inc., Lincoln, Nebraska) are used in this thesis.

2.2 Bavaria, Germany

The first flight campaign took place from 19 June to 11 July 2008 over Lake Starnberg, Germany (47.9°N, 11.3°E). The lake is located in the foreland of the German Alps, which is a slightly rolling landscape (600–800m a.s.l.) and mainly consists of grassland with patches of forest. The campaign was designed to acquire the necessary datasets for developing and calibrating a WSMA wind computation algorithm which enables to dynamically offset the effects of pilot input. For this purpose 31 flights of 8 specific patterns were performed in the free atmosphere above Lake Starnberg. Of these, 10 ideal realizations of these patterns are used in Metzger et al. (2011, Appendix B). Aside from the WSMA, no additional scientific equipment was deployed. This campaign was funded and implemented solely under the Karlsruhe Institute of Technology.

2.3 Brandenburg, Germany

An evaluation of the WSMA wind and EC flux measurements against ground-based reference measurements was carried out during the second flight campaign between 14 and 21 October 2008. This experiment was performed around the boundary layer field site Falkenberg (52.2°N, 14.1°E) of the German Meteorological Service (DWD), Richard-Aßmann Observatory, Lindenberg, Germany. This field site lies in the basically flat North German Plain, and the terrain height varies between 40 m and 130 m a.s.l. within an area of 20×20 km^2. The land cover is dominated by agriculture and forests, interspersed by equal amounts of lakes, meadows and settlements. This campaign was funded by the Karlsruhe Institute of Technology, and implemented in close collaboration with the DWD.

To evaluate the accuracy of the WSMA wind measurement, Metzger et al. (2011, Appendix B) use data from an instrumented 99 m tower and from sonic detection and ranging (SODAR, PCS-2000/64, Metek GmbH, Elmshorn, Germany). The 99 m tower provided cup measurements of wind speed at four levels (40, 60, 80, and 98 m a.g.l., Wind Sensor Classic, Adolf Thies GmbH, Göttingen, Germany). The wind direction was measured with vanes at heights of 40 and 98 m a.g.l. (Wind Direction Sensor Classic, Adolf Thies GmbH, Göttingen, Germany), and a static pressure reference was provided at 1 m a.g.l. (PTB220A, Vaisala Oy, Helsinki, Finland, 10 min averages). Tower profile and pressure data were averaged and stored in 10 min intervals. Sonic anemometers (USA-1, Metek GmbH, Elmshorn, Germany) provided wind vector measurements at a rate of 20 Hz at 50 and 90 m a.g.l. The SODAR wind vector profiles (15 min averages) reached, at increments of 20 m, from 40 to 240 m a.g.l. 17 cross-shaped flight patterns with legs of 3 km length were performed with their center

between the 99 m tower and the SODAR. The flights were carried out at the approximate sounding levels of the 99 m tower and the SODAR (50, 100, 150, 200 and 250 m a.g.l.). This enables the direct comparison of the wind components between the WSMA and the ground-based measurements.

Moreover, Metzger et al. (2012, Appendix D) compare turbulence statistics, EC fluxes and power spectra between the WSMA and ground-based measurements. For this purpose additional data from installations on the 99 m tower, an EC surface flux measurement, a large-aperture scintillometer (LAS), a wind profiler, and from radio soundings are used. In combination with the sonic anemometers, open path IRGAs (LI-7500, LI-COR Biosciences, Lincoln, U.S.A.) at 50 m and 90 m a.g.l. enable the comparison of turbulence statistics, EC fluxes and power spectra between WSMA and tower. The 99 m tower was further equipped with profile measurements of temperature (HMP-45, Vaisala Oy, Helsinki, Finland) and humidity (Frankenberger Psychrometer, Theodor Friedrichs GmbH, Hamburg, Germany) at 40, 60, 80, and 98 m a.g.l (10 min averages). The profiles are interpolated to the heights of the EC installations and are used to compare average temperature and humidity measurements between WSMA and tower. The two lowest levels of the cross-shaped flight patterns were performed at the approximate height of the tower EC installations, and provided 36 individual flight legs for this comparison. Identical instrumentation as on the tower was used for an additional EC surface flux measurement upwind (south) of the tower base, at 2.4 m a.g.l. Also 10 min area-averaged surface sensible heat fluxes were derived from the LAS. Furthermore, hourly estimates of the CBL depth were derived from SODAR and wind profiler data, and from six-hourly routine radio soundings performed by the DWD. The surface sensible heat fluxes measured at the 2.4 m EC and the LAS are used in conjunction with the CBL depths to approximate vertical flux profiles. On two days, simultaneous tower and WSMA measurements of the sensible heat flux are compared to these flux profiles.

2.4 Inner Mongolia, P.R. China

The third flight campaign was performed from 23 June to 4 August 2009 over the remote steppe of the Mongolian Plateau. The hilly investigation area south of the provincial capital Xilinhot, Inner Mongolia, China (43.6°N, 116.7°E, 1000–1400 m a.s.l.) is covered by semi-arid grassland, intersected by a dune belt. The land cover in the investigation area varies distinctly in space and time (Ketzer et al., 2008; Schaffrath et al., 2011). On the small scale landscape position effects soil moisture availability. On larger scales management practices such as haymaking, grazing or irrigated agriculture are suspected to cause a vast heterogeneity in the heat and moisture fluxes between surface and atmosphere. As integral part of the Sino-German research collaboration MAGIM (Matter fluxes in grasslands of Inner Mongolia as influenced by stocking rate), the goal of this flight campaign was to support long-term ground-based measurements through gaining new insights in the spatial distribution of fluxes on the regional scale. In addition to the WSMA measurements, Metzger et al. (2013, Appendix E) deployed a ceilometer (LD40 – Vaisala Oy, Helsinki, Finland) for this purpose. The ceilometer provided 10 min vertical profiles of the atmospheric laser radiation backscatter intensity. The depth of the convective boundary layer (CBL) was inferred from this data using the maximum gradient method (Emeis et al., 2008) in combination with semi-daily radiosonde ascends in nearby Xilin Hot (World Meteorological Organization station 54102, http://weather.uwyo.edu/upperair/sounding.html). The CBL depth is used to characterize the

horizontal mixing between surface and flight level and to determine the source area of the WSMA based EC flux measurements. EC measurements with the WSMA were performed along 14 individual straight line flight patterns, which followed the terrain at ≈50 m a.g.l. Two flight lines perpendicular to the prevailing wind direction were chosen on a daily basis, and repeated until a minimum of 40 km of data was acquired. This resulted in a total of 35 WSMA soundings during the flight campaign, of which Metzger et al. (2013, Appendix E) selected 12 flights for analysis with stationarity of the solar irradiance being the selection criteria. This flight campaign was funded by the German Research Foundation, and was implemented in collaboration between the Karlsruhe Institute of Technology and the Chinese Academy of Sciences.

3 Results

In the following, the results of the development and application of WSMA based EC flux measurement are outlined and related to the corresponding publications in Appendices B–E. This development process starts with the wind measurement (Sect. 3.1), continues to the turbulent flux measurement (Sect. 3.2) and concludes with a sample application of the WSMA based EC flux measurement over complex and previously inaccessible terrain (Sect. 3.3).

3.1 Wind measurement

The EC technique relies upon the precise measurement of fluctuations of atmospheric quantities, based on negligible sensor drift throughout an averaging period. Measured from aircraft, the determination of the wind vector requires a sequence of thermodynamic and trigonometric equations (Supplement to Metzger et al., 2011, Appendix C). These equations propagate various sources of error, and are consequently the lynchpin for EC flux measurements from aircraft. An abundance of methods exists for calibrating these algorithms to the individual aerodynamic properties of a wide range of fixed-wing aircraft (e.g., Tjernström and Friehe, 1991; van den Kroonenberg et al., 2008; Williams and Marcotte, 2000). However, none of the existing methods is capable of treating WSMA specific effects on the wind measurement, originating from trike rotation and wing deformation. Hence, Metzger et al. (2011, Appendix B) advance the existing wind calibration procedures with focus on (i) creating a formal calibration framework, (ii) retaining applicability to fixed-wing aircraft, and (iii) offsetting WSMA specific effects on the wind measurement. The result is a bottom-up calibration procedure in three stages;

(i) basic calibration of temperature and pressure sensors;

(ii) general in-flight calibration of flow angle measurements, and;

(iii) WSMA specific in-flight calibration of flow angle measurements.

These three stages are composed of a sequence of seven calibration steps A–G, which is shown in Fig. 2. Each step targets one or more variables in the wind vector computation, utilizes experimental data under defined environmental conditions, and results in incrementally refined system performance. Additional information on the flight maneuvers used for this purpose as well as variable definitions is provided in Metzger et al. (2011, Appendix B).

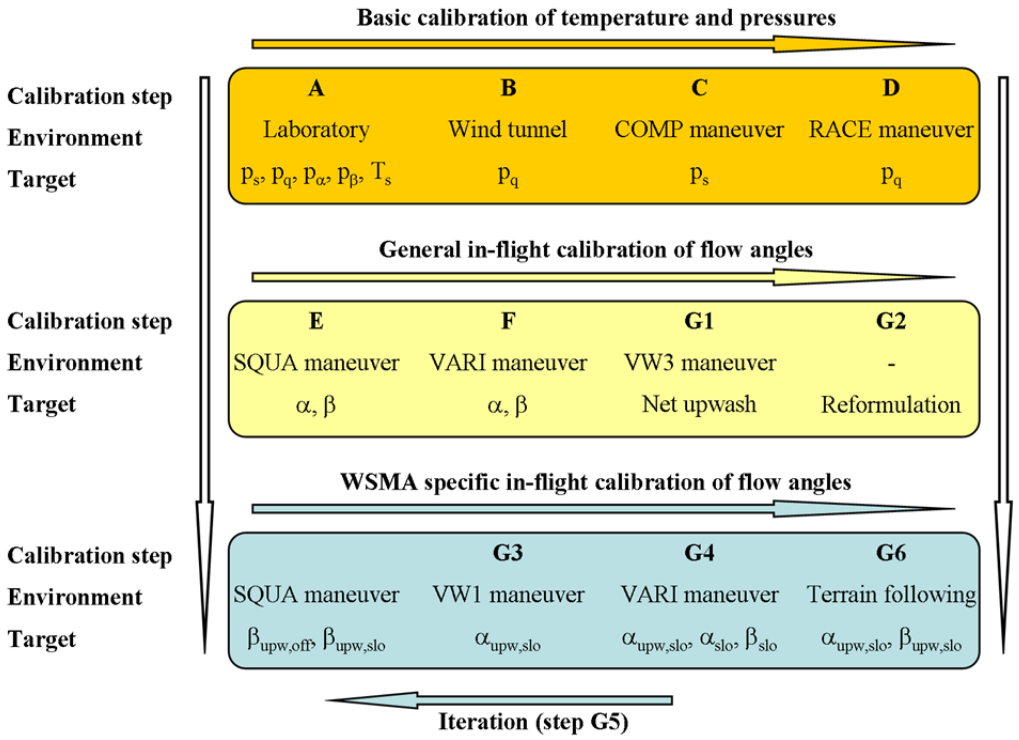


Fig. 2. Flow chart of the wind calibration framework. The calibration steps A–G are carried out in a sequence from left to right, top to bottom. Each step results in an incrementally refined system. The iterative step G5 (blue background) comprises the three flight maneuvers SQUA, VW1 and VARI. Within G5 the SQUA maneuver is not associated with an individual calibration step. Figure from Metzger et al. (2011, Appendix B).

The individual calibration steps A–F are well documented in the literature (e.g., Garman et al., 2006; Kalogiros and Wang, 2002a; van den Kroonenberg et al., 2008). Here, these steps are combined to a scalable wind calibration framework, which is sufficiently universal to apply to a wide range of research aircraft beyond the WSMA. This calibration framework is expanded by an additional step G. In this last step, the ability to overcome the effects of variable aircraft lift on the wind measurement is advanced from diagnostic observation (Garman et al., 2008) to prognostic treatment. For this purpose the lift coefficient is used, defined as (Metzger et al., 2011, Appendix B);

$$CL = \frac{2}{\rho (v_{tas})^2} \frac{L}{S}, \tag{3}$$

with air density ρ, true airspeed v_{tas}, and the wing loading $\frac{L}{S}$, where the lift L equals the aircrafts sum of forces perpendicular to the airstream, and S is the wing's surface area. To disentangle the comprehensive sequence of assessment and treatment, step G is further divided into six sub-steps G1–G6.

In the first step G1, the net effect of changing aircraft trim (v_{tas}) and wing loading on the wind measurement is investigated during pilot-forced altitude oscillations (flight pattern VW3 in the free

atmosphere above Lake Starnberg, Sect. 2.2). As a consequence of lift generation air rises in front of the wing, which is defined as upwash. From Fig. 3 it is apparent that the upwash modeled after a parameterization for fixed wing aircraft (Crawford et al., 1996) is proportional to *CL*. As opposed to this parameterization, the observed net upwash at the 5HP location is smaller by one order of magnitude and more variable, as well as phase inverted with *CL*. These findings are confirmed with the level acceleration – deceleration flight VW1 with a long period (180 s) and negligible vertical aircraft velocity (Fig. 3). Considering a change from high v_{tas} (low *CL*) to low v_{tas} (high *CL*) during level flight actually a number of effects contribute to the observed net upwash: (i) increase of upwash production from the wing according to the upwash parameterization for fixed wing aircraft, (ii) decrease of wing circulation effective at the 5HP through larger distance and opening angle between 5HP and wing, and (iii) decrease of propeller induced upwash. The latter effects counteract the wing induced upwash. In addition (iv) the shape of the aeroelastic wing, as well as (v) the flow around the trike change. In place of a static parameterization, Garman et al. (2008) propose to correct for upwash by considering the wing loading, which carries information on the aircraft's vertical acceleration. In the following, *CL* is used instead of wing loading, which has the advantage that additional information on the aircraft's trim, i.e. information on above effects (ii)–(v), is included. Over eight independent flights of patterns VW1, VW2 and VW3 the observed net upwash is correlated with *CL* (−0.53±0.16), change in v_{tas} (0.57±0.16), and wing pitch (−0.50±0.20). I.e. the application of control forces leads to a simultaneous change in both, the wind field and the lift coefficient. In the following this relationship will be used to treat the effect of pilot input on the wind measurement.

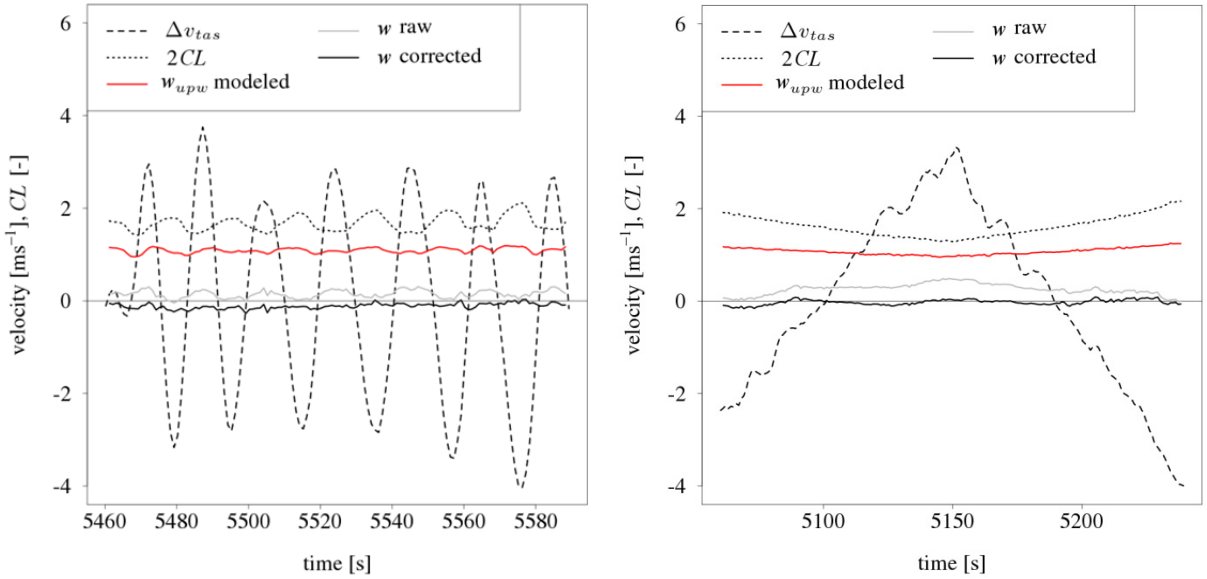


Fig. 3. Forced oscillation pattern (VW3, left) and level acceleration–deceleration pattern (VW1, right) on 25 June 2008. For improved legibility the average is subtracted from true airspeed (Δv_{tas}) and lift coefficient is inflated by the factor two (2 *CL*). Displayed is the vertical wind (w) before (raw) and after (corrected) correction for dependence on *CL*. For comparison the modeled upwash (w_{upw}) after Crawford et al. (1996) is presented. Figure modified after Metzger et al. (2011, Appendix B).

In step G2, a correction for the net upwash at the measurement location is proposed. Crawford et al. (1996) and Kalogiros and Wang (2002b) have shown that the upwash parameterization for fixed-wing aircraft can be reformulated as a function of *CL* in the 5HP measured airstream attack angle (α_m) of the airstream. On this basis a linear model in α is proposed to treat variable upwash for the WSMA wind measurement (Metzger et al., 2011, Appendix B);

$$\alpha_\infty = \alpha_m - \alpha_{upw},$$
$$\alpha_\infty = \alpha_m - \left(\alpha_{upw,off} + \alpha_{upw,slo}\, CL\right), \tag{4}$$

with α_∞ the (desired) free airstream attack angle and α_{upw} an additive attack angle provoked by the upwash with $\alpha_{upw,off}$ and $\alpha_{upw,slo}$ being its constant part and slope with *CL*, respectively. As outlined in step G1, the complex interaction of wing upwash and aeroelasticity, distance and opening angle with the 5HP, propeller slipstream and flow around the trike is collectively correlated in *CL*. This offers the possibility of a dynamic treatment of the net flow distortion in one single explanatory variable. The objective of this correction is to reposition the mean vertical wind under the effects of terrain-following flight, i.e. to correct for offset drift.

In step G3 the coefficients in Eq. (4) are calibrated using specific flight patterns in the free atmosphere. From a simple least-squares regression procedure $\alpha_{upw,off}$=0.031 rad and $\alpha_{upw,slo}$ =−0.027 rad are determined. That is, α_m would be overestimated by ≈1.7° if the WSMA could fly at zero lift. The effect decreases with slower flight at a rate of ≈−1.7° per *CL*. For the VW3 flight (Fig. 3), the application of this procedure improved the decorrelation of *w* with v_{tas} from 0.79 to −0.11, as well as the decorrelation with wing pitch from −0.78 to 0.17. Also, changes in aircraft trim during the VW1 flight did not alter the offset in *w* anymore (Fig. 3).

In step G4 the calibration of $\alpha_{upw,off}$=0.039 rad is updated from flights in the atmospheric surface layer in combination with a variance optimization criterion (Kalogiros and Wang, 2002a; Khelif et al., 1999; Tjernström and Friehe, 1991). Such procedure provides a much tighter constraint compared to the flights in the free atmosphere used in step G3. While contamination of the latter, e.g. by large scale subsidence, will not influence the determination of $\alpha_{upw,slo}$, it might have biased the initial $\alpha_{upw,off}$.

In step G5 sidewash in the sideslip angle β is calibrated in analogy to Eq. (3). According to Eq. (A11) in Supplement to Metzger et al. (2011, Appendix C) cross dependence occurs between the calibrations in α and β. This problem is solved by iterating the optimality criteria for β, $\alpha_{upw,slo}$ and $\alpha_{upw,off}$ in sequence. Compared to the upwash parameterization, sidewash is found to be modest ($\beta_{upw,off}$=−0.004 rad) and less sensitive regarding *CL* ($\beta_{upw,slo}$=−0.010 rad). Leaving dynamic considerations aside (i.e., *CL*=0), the magnitude of the sideslip angle correction is one order of magnitude lower than the attack angle correction. For v_{tas}=30 m s^{-1} this affects the wind measurement to approximately −0.1 m s^{-1} and 1.2 m s^{-1}, respectively. The transverse distortions increase and the vertical distortions decrease at a ratio of ≈1:3, when considering interactions with propeller and wing (i.e., *CL*>0). A summary of all calibration coefficients can be found in Metzger et al. (2011, Table 6, Appendix B).

In step G6 the uncertainty of the proposed correction is assessed during terrain-following flight. For this purpose, the combined regression errors in $\alpha_{upw,off}$ and $\alpha_{upw,slo}$ are numerically propagated through the entire wind computation for a terrain-following flight. The resulting uncertainty in the vertical wind measurement is within 0.1 m s^{-1} root mean square deviation (RMSD) for the mean and 1% for the variance. This compares to the magnitude of the correction, which is in the order of 0.5 m s^{-1} for the mean and 3% for the variance, respectively. In order to adjust aircraft altitude, the pilot anticipates the terrain contours at a scale of kilometers. At typical airspeeds this corresponds to an adjustment of power setting and wing pitch by the pilot at frequencies <0.1 Hz. To avoid artifacts from turbulent atmospheric motions at higher frequencies, the present correction is only applied to frequencies <0.1 Hz. This is achieved by calculating Eq. (4) through a third order Savitzky-Golay complementary filter (e.g., Chen et al., 2004). The treatment leads to a decrease in the vertical wind variance in the order of −3%. This is expected because the impact of the low frequency pilot actions on the wind measurement is removed. However, low frequency atmospheric motions, such as NPEs, overlap with pilot actions in frequency space. In Metzger et al. (2011, Sect. 4.1.7, Appendix B) the effect of low frequency atmospheric motions on the correction is estimated using a numerical procedure. It is found that the maximum deviation from an undisturbed measurement is within ±3%, or sub centimeter, and that the variance of the vertical wind changes by <0.01%. On this basis it is concluded that, (i) the proposed correction Eq. (4) can be applied to the entire frequency range <0.1 Hz without introducing significant artifacts to the wind measurement, and (ii) the resulting wind measurement is independent of slow aircraft rising and sinking maneuvers, as well as changes in aircraft trim and lift. Hence the WSMA wind measurement fulfills the basic requirements to be used for EC measurements during terrain-following flight.

Finally, Metzger et al. (2011, Appendix B) use three independent lines of analysis to quantify the overall system uncertainty; (i) uncertainty propagation through the wind computation equations, (ii) in-flight testing and (iii) comparison of the measured wind vector with ground-based measurements. In the following, only the results of 36 tower–aircraft comparison flights at the boundary layer field site Falkenberg (Sect. 2.3) are presented. The comparability of WSMA and ground based wind measurement is quantified by calculating RMSD and offset (BIAS) for an equal number of flights at 24 m s^{-1} and 27 m s^{-1} true airspeed. The impact of calibration steps C–G on these measures is displayed in Fig. 4. The measurement of the horizontal wind components is mainly improved (14%, relative to the initial uncertainty) by the in-flight dynamic pressure correction (step D). After the wind square analysis (step E) the measurement is no further improved nor deteriorated. Yet the vertical wind measurement receives its greatest improvement (31% RMSD) during steps F–G, i.e. variance optimization and the consideration of the lift coefficient. During these steps BIAS and ΔBIAS, i.e. the dependence of BIAS on true airspeed, are reduced. Subsequently, three outliers are removed from the WSMA measurements, on grounds of turbulence and wake effects generated at a forest edge. The final RMSD and BIAS between WSMA and ground based measurements amount to 0.39 m s^{-1} and −0.11 m s^{-1} for the horizontal-, and 0.27 m s^{-1} and −0.10 m s^{-1} for the vertical wind components, respectively.

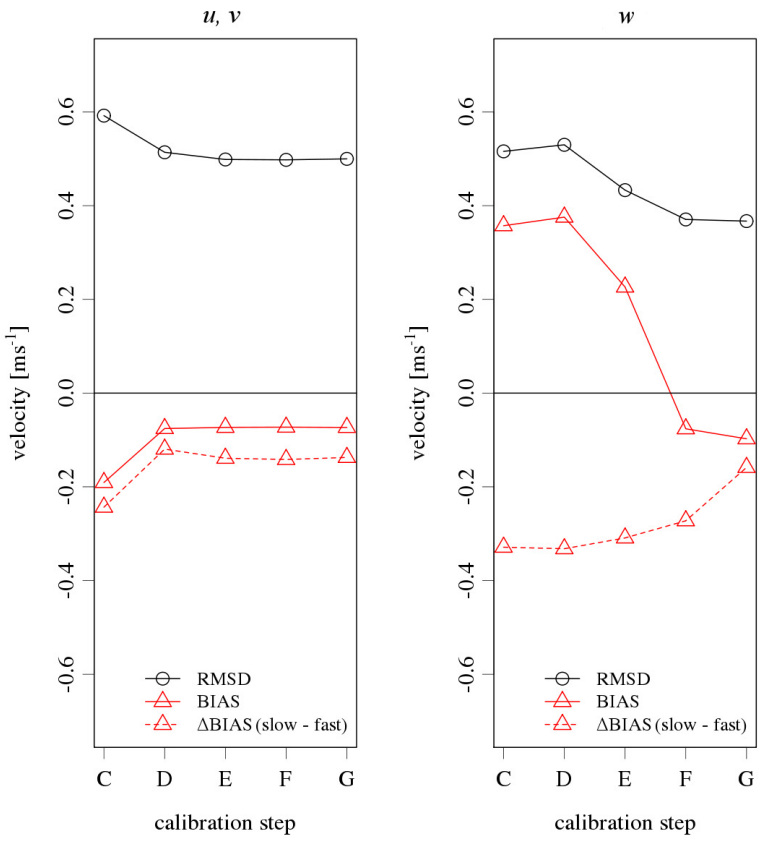


Fig. 4. Influence of the calibration steps C–G on RMSD and BIAS between weight-shift microlight aircraft and all simultaneous ground-based measurements of the horizontal (u, v) and the vertical (w) wind components. ΔBIAS indicates the difference in BIAS between measurements at 27 m s^{-1} and 24 m s^{-1} true airspeed. Figure modified after (Metzger et al., 2011, Appendix B).

3.2 Turbulent flux measurement

Contributions to the EC flux measurement originate from turbulent atmospheric motions on a variety of wavelengths and amplitudes. In order to reliably estimate the total flux, the fluctuations of the vertical wind and the scalars must be measured with high accuracy and precision. Furthermore, the instrumentation and data acquisition must possess a suitable frequency response and sampling rate. In the case of airborne measurements, the carrier can additionally influence the spectral quality of the measurement. Hence, Metzger et al. (2012, Appendix D) evaluate the performance of the WSMA flux measurement through;

(i) assessing measurement precision and accuracy;

(ii) comparing spectral properties, averages, deviations and fluxes between WSMA and tower-based EC measurements;

(iii) using footprint modeling to study the measurements' spatial context, and;

(iv) bringing to attention the effect of larger-scale atmospheric motions on the results by means of comparison to a LAS.

(i) The measurement precisions in Metzger et al. (2012, Table 1, Appendix D) are superimposed over the WSMA turbulence raw data of the 36 tower–aircraft comparison flights at the boundary layer field site Falkenberg (Sect. 2.3). Both original and manipulated datasets are processed through the entire wind computation, and the deviations in the wind components are summarized (σ_u=0.07 m s^{-1}, σ_v=0.09 m s^{-1}, σ_w=0.04 m s^{-1}). Drawing on Lenschow and Sun (2007), we assume that a minimum signal-to-noise ratio of 5:1 is required to measure the wind fluctuations with sufficient precision for EC applications. Thus, standard deviations of 0.45 m s^{-1} and 0.20 m s^{-1} are reliably resolved in the horizontal and vertical wind components, respectively. For all 36 flights, the original and manipulated datasets are further propagated through the EC computation, now also considering the precisions of the fast temperature and humidity measurement. The result is an estimate of the least resolvable change in the measured flux (σ_{u*}=0.003 m s^{-1}, σ_H=0.9 W m^{-2}, and σ_E=0.5 W m^{-2}). Using the above signal-to noise-ratio of 5:1, changes in friction velocity, sensible- and latent heat of 0.02 m s^{-1}, 5 W m^{-2}, and 3 W m^{-2}, respectively, are reliably resolved. The above repeatability does not consider the environmental changes (temperature, humidity, pressure etc.) which are experienced by the sensors measuring aboard a moving aircraft. Changes in the environment likely lead to sensor drift, increasingly deteriorating the measurement with flight duration. Hence, Metzger et al. (2012, Appendix D) assess whether the measurement accuracy warrants the resolution of horizontal CBL structures up to the mesoscale (10–100 km). Using the procedure of Lenschow and Sun (2007), it is confirmed that the wind measurement accuracy is equal or better the required signal level for this purpose. Also, it is seen that the wind measurement accuracy is limited by the 5HP dynamic and differential pressure measurements used to infer the flow angles α, β. Accuracy in the scalar measurements along a flight leg is constrained by the drifts of the fast thermocouple and the dew point mirror. Using the same 1: 5 signal-to-noise criteria as above, temperature- and humidity fields differing >1.3 K or >5% mixing ratio, respectively, can be reliably distinguished over a 100 km flight leg.

(ii) Various motions of the WSMA can potentially disturb the wind measurement, and consequently the EC flux measurement. Metzger et al. (2012, Appendix D) examine potential resonance from the engine or propeller of the WSMA, as well as the natural frequencies of trike and wing. Transverse and especially vertical to the WSMA body, accelerations measured in the 5HP agree well with measurements from the IGS up to a frequency of 2–3 Hz (Fig. 5). Beyond that the acceleration measurements at the 5HP continue to follow the pattern of IGS accelerations, but are slightly enhanced. This is expected, since the 5HP has a longer lever (≈0.5 m) with respect to the center of rotation, i.e. the hangpoint of WSMA wing and trike. Consequently the acceleration amplitudes are higher at the 5HP, which is also accounted in the lever arm correction of the wind measurement (Supplement to Metzger et al., 2011, Eq. A12, Appendix C). Two more spectral peaks are found at 30 and 45 Hz. These peaks are likely to be associated with harmonics from the engine and propeller, rotating at ≈100 Hz and ≈30 Hz, respectively. However, the −3 dB point (20 Hz) of the 5HP's low-pass filter is lower than the Nyquist frequency (50 Hz) of its data acquisition. This prevents aliasing of these spectral peaks into the final 10 Hz wind vector observations. For the acceleration component longitudinal to the WSMA, the 5HP pattern is enhanced compared to the IGS. This is surprising,

because it is the axis of plug- and socket connection between IGS and 5HP, i.e. the axis with the least margin for resonance. No remnants of this scale discrepancy are evident in the wind measurements, in particular between 1–5 Hz (Metzger et al., 2012, Fig. 3, Appendix D). This leads to the conclusions that (i) the wind vector computation correctly accounts for the displacement of 5HP and IGS, and (ii) the cause of enhanced 5HP acceleration measurements (especially longitudinal to the body) lies in the fixture of the acceleration sensor in the 5HP, rather than in the mounting of the 5HP against the IGS. The spectral behavior of the wing acceleration measurements is different to those of the trike. It displays a distinct peak around 0.7 Hz, which is only present in the transverse component of the trike measurements. It can be understood as the wing's natural frequency, i.e. its inertia.

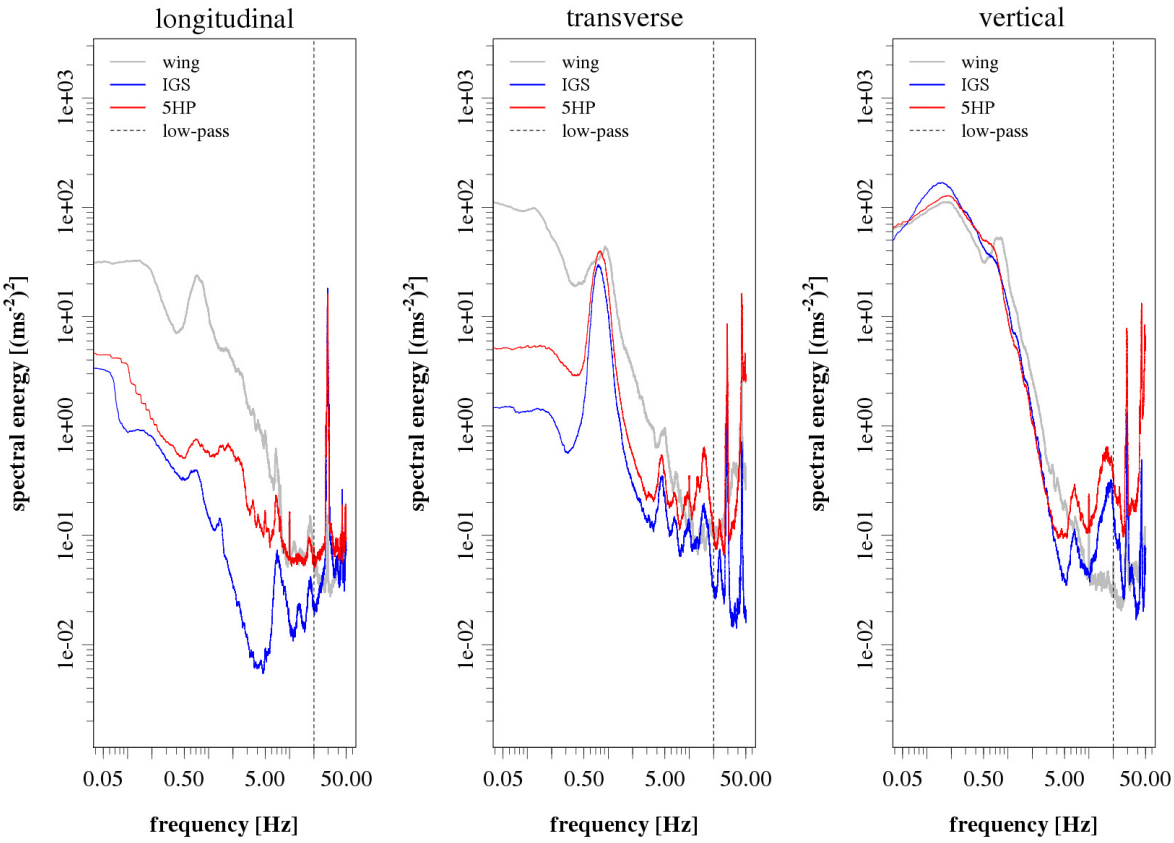


Fig. 5. Smoothed power spectra of acceleration measurements in the WSMA trike coordinate system. The dashed vertical line indicates the −3 dB frequency (20 Hz) of the Butterworth low-pass filter in the wind vector data acquisition system. Figure modified after Metzger et al. (2012, Appendix D).

Subsequently, Metzger et al. (2012, Appendix D) quantify the impact of WSMA spectral properties on the wind- and EC flux measurement. For this purpose a fast Fourier transformation is applied to the 36 tower and WSMA fast response data series during the comparison flights. Each individual transform is normalized to a sum of unity. To reduce scatter, the normalized transforms of all tower and WSMA measurements, respectively, are then binned into frequency bands and ensemble-

averaged. Here, only the results for the cospectra are discussed. In Fig. 6 ensemble cospectra (*Co*) are presented as function of the normalized frequency $n=f\cdot z/\bar{U}$, with f being the sampling frequency, z being the measurement height, and $\bar{U}$ being the horizontal wind speed for the tower and the true airspeed for the WSMA, respectively. Also shown is the reference cospectrum of Massman and Clement (2004), with the spectral maximum at n=0.1 for unstable stratification (Kaimal and Finnigan, 1994, 24 out of 36 flights). The momentum flux at the tower exhibits large scatter in the individual cospectra at both installation heights, and the ensemble cospectrum is not calculated. The scatter might result from the wind direction dependent correlation of the horizontal- and vertical wind components at the USA–1 sonic anemometers (e.g., Mauder et al., 2007c). All analyzed ensemble cospectra approximately follow the reference cospectrum, and for the heat fluxes the peak of the tower cospectra coincide with n=0.1 of the reference cospectrum. The cospectral peaks of the WSMA measurements are marginally shifted towards higher frequencies around n=0.2. Metzger et al. (2012, Sect. 3.3.3, Appendix D) associate increased variance in the WSMA wind components with spectral artifacts resulting from the treatment of the net flow distortion in the time-, but not in the frequency domain (Metzger et al., 2011, Appendix B). In order to quantify the impact on the WSMA flux measurement, all individual cospectra are compared between the WSMA and the reference cospectrum in the range of the wing's natural frequency ($0.4\leq f\leq 2$ Hz). To account for the influence of stratification, the peaks of the reference cospectra are calculated using the forms of Kaimal et al. (1972). Relative to the entire frequency range the spectral artifacts lead to a systematic deviation in the fluxes of momentum, sensible- and latent heat of 3±6%, −1±6% and 1±3% (median differences), respectively.

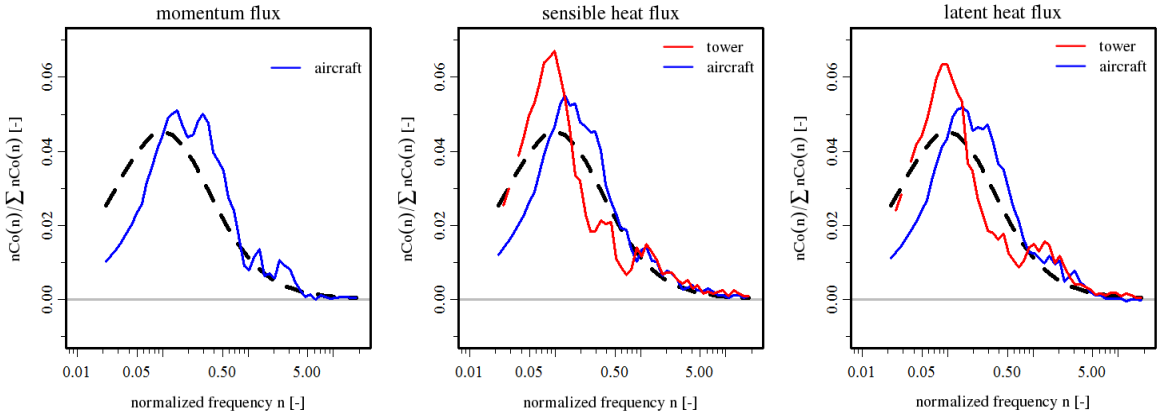


Fig. 6. Average cospectra of all measurements between tower and WSMA. Also shown is the reference cospectrum of Massman and Clement (2004, dashed line). Figure from Metzger et al. (2012, Appendix D).

The WSMA wind and flux measurements are corrected for this spectral inconsistency before comparing them to ground based measurements. The appropriate correction factors are estimated from the comparison of measured spectra and cospectra to modeled ones (Metzger et al., 2012, Sects. 3.3.3, 3.3.4, Appendix D). A maximum likelihood functional relationship (MLFR, Ripley and

Thompson, 1987), which considers the random error in the data, is used to compare tower and WSMA measurements. For this purpose the statistical random error σ_{ran} is calculated after Lenschow and Stankov (1986) and Lenschow et al. (1994), and consolidated after Mahrt (1998) for the ensemble of all measurements (σ_{ens}). Averages of along-wind component u, temperature T and absolute humidity a agree very well between tower and WSMA (Table 1). The MLFRs for standard deviations (2–34%) and fluxes (17–21%) indicate higher estimates of the airborne measurements compared to the tower. Considering the 99.5% confidence intervals the observed differences are not significant, with exception of the temperature standard deviations. It must be noted that the magnitude of the flow distortion correction for the USA–1 sonic anemometers at the tower alone is in the order of the observed differences between the tower and the WSMA. Moreover, the amplitude resolution test by Vickers and Mahrt (1997) would reject 28 out of 36 tower sonic temperature data sets. The problem is related to the insufficient sonic temperature resolution (0.01 K) of the USA–1, which appears as additional spectral energy in the form of high frequency white noise. Consequently the USA–1 measurements cannot be regarded as reliable reference for the temperature standard deviations, and the applied spectral correction factors and resulting MLFRs must be interpreted with caution. The

Table 1. Results of the maximum likelihood functional relationships between tower and WSMA measurements. Shown are the MLFR slope and its standard error Slope±σ, weighted coefficient of determination R^2, residual standard error σ_{res}, the average statistical random error σ_{ran} and the ensemble random error σ_{ens}. Table from Metzger et al. (2012, Appendix D).

Variable	Slope±σ	R^2	σ_{res}	σ_{ran}	σ_{ens}
		Averages			
u	0.99±0.02	1.00	9%	15%	3%
T	1.00±0.00	1.00	0%	0%	0%
a	0.99±0.00	1.00	1%	2%	0%
		Standard deviations			
u	1.15±0.05	0.99	21%	9%	2%
v	1.02±0.05	0.98	20%	8%	1%
w	1.10±0.03	0.99	9%	5%	1%
T	1.34±0.07	0.98	17%	7%	1%
a	1.17±0.08	0.98	23%	9%	1%
		Fluxes			
u_*	1.21±0.07	0.98	13%	25%	5%
H	1.17±0.08	0.98	10%	29%	8%
LE	1.17±0.10	0.96	25%	34%	7%

problem is less pronounced for the sensible heat flux. The white noise in the USA–1 sonic temperature measurement does not affect the measurement of, or the correlation with, the vertical wind measurement. The result is a modest underestimation of −3% of the tower sensible heat flux due to reduced coherence of sonic temperature and vertical wind at high frequencies.

(iii) Subsequently Metzger et al. (2012, Appendix D) investigate whether the differences between WSMA and tower measurements can be related to their different spatial representativeness. For this purpose a cross-wind distributed footprint parameterization is used together with Corine Land Cover 2006 data (Version 13, European Environment Agency, 2010, 100 m horizontal resolution). The spatial context of the platforms agrees well when considering the average footprint contributions over all tower-WSMA comparison measurements. For both platforms most of the footprint covers arable land (95–97%). Contributions from the remaining land covers are sub-percent except for forest (2–3%), and meadows do not contribute at all. Taking a closer look at the individual, simultaneous measurements the source areas can however differ considerably. Over all simultaneous measurements the actual overlap ranges from 12–68% of the footprint weights, with a median of 35±17%. However varying overlap did not systematically alter the differences in the flux measurements between tower and WSMA ($R^2 \leq 0.07$). Hence, spatial representativeness cannot explain the remaining differences between WSMA and tower measurements.

(iv) Lastly, Metzger et al. (2012, Appendix D) assess the potential impact of principal differences of spatial averaging (LAS, WSMA) and temporal averaging (tower EC) on the measured fluxes. For this purpose simultaneous measurements of the sensible heat flux on two days are inter-compared between the measurement platforms using boundary layer scaling. Any H measured by tower EC and extrapolated to flight altitude is lower by 25–40% compared to the LAS. At the same time H measured by the WSMA is ≤25% lower compared to the LAS, but 15–25% higher compared to the tower EC. The footprints of all measurements is dominated by >90% contributions from arable land. Consequently differing source areas of the measurements do not qualify as potential reason for the observed differences. Foken (2008b) and Mahrt (2010) suggest that the energy balance non-closure frequently observed from tower EC measurements is connected to the interaction of terrain heterogeneity and turbulent scales. Following their hypothesis, the tower EC cannot adequately capture flux contributions from NPEs due to its inability of spatial sampling. On the other hand a LAS captures NPEs up to the dimension of its path length, with increasing sensitivity towards the center of its optical path (Foken et al., 2010). Also airborne EC is capable of spatial sampling and captures some of the associated flux, depending on the horizontal extent of the NPEs and the flight path. The presence of NPEs in the study area has been shown (Steinfeld et al., 2007; Uhlenbrock et al., 2004), and can thus be considered a potential explanation for the deviation of the tower EC results from the WSMA, and even more so from the LAS results.

Metzger et al. (2012, Appendix D) show that turbulence measurements from WSMA can be achieved with sufficient precision and accuracy to enable EC flux calculation. Differences in the order of 15–25% remain between the fluxes measured by the ground based instruments and the WSMA. However, the 99.5% confidence intervals in the comparison between tower and WSMA indicate that the differences are insignificant and the WSMA flux measurement is unbiased. The precision of the WSMA flux measurement can be quantified to ≤10% (1σ MLFR slope error).

3.3 Spatial resolution and extrapolation of turbulent fluxes

The overarching goal of airborne eddy-covariance flux measurements is to bridge the gap between observations and data assimilation approaches on different spatial scales. Metzger et al. (2013, Appendix E) develop a procedure that aids this purpose by ‘mining‘ the information content of EC flux observations. Through accomplishment of four subsequent steps this LTFM procedure extracts quantitative response functions with environmental drivers;

(i) low level EC flux flights;

(ii) time-frequency analysis of the flux observations;

(iii) composition of continuous biophysical surface properties in the flux footprint, and;

(iv) environmental response function from non-parametric machine learning techniques.

In developing the LTFM procedure, the objective of Metzger et al. (2013, Appendix E) is to spatially explicitly characterize the exchange of sensible- and latent heat over the heterogeneous steppe landscape of the Xilin River Catchment (XRC), Inner Mongolia, P.R. China (Sect. 2.4).

(i) The LTFM procedure requires the relation of the airborne measured fluxes to land cover properties. To enable this requirement an aircraft is bound to measure close to the surface, where characteristic fluxes from different land covers are not yet fully homogenized (or blended, Mason, 1988; Wood and Mason, 1991). Moreover the flux must be measured at a constant altitude above ground, so as to avoid artificial flux contributions through altitude fluctuations along vertical gradients (Vickers and Mahrt, 1997). However investigation areas are seldom ideally flat, and topography can vary significantly across a domain. To safely follow terrain contours at a low and constant altitude above ground, the aircraft must possess a low ratio of true airspeed to climb rate. The WSMA fulfills this requirement, and terrain-following EC flux measurements with the WSMA were conducted at 50 m a.g.l. in the XRC. From boundary layer scaling it is found that the vertical flux gradients below the flight level satisfy the surface layer definition (constant within |5–10%|). Hence measured H and E can be interpreted as surface fluxes.

(ii) Metzger et al. (2013, Appendix E) use the wavelet cross-scalogram technique to enable a high spatial discretization of turbulent flux measurements, without neglecting flux contributions from long wavelengths. For each individual 10 Hz observation the wavelet cross-scalogram represents the measured turbulent flux as contributions from different atmospheric transport scales. Theoretically, this enables a spatial resolution of ≈2.5 m of the WSMA flux measurement. However, for an individual sample the random error is excessively large, but decreases inversely proportional to the square root of the sample size (e.g., Lenschow and Stankov, 1986). In search of a suitable sample size, a compromise must be found between random error (high resolution) and smearing (low resolution) of the resulting flux estimates. Metzger et al. (2013, Appendix E) find that the typical length scale of surface heterogeneity is in the order of 1000 m. Thus, a flight path length of 1000 m is a physically meaningful window for the computation of turbulence statistics and fluxes. Hence, while retaining a spatial discretization of 90 m, the wavelet cross-scalogram is integrated over centered subintervals of 1000 m length. Compared to an integration length of 90 m this results in a decrease of the ensemble random error (Mahrt, 1998) of ≈70% for a flight line of 20 km length. This procedure

yields a high number of flux observations along a flight line, which provides previously unachievable resolution and coverage of the environmental state space. The downside of the low-frequency support of wavelets is that edge effects due to the finite overall data set increase with scale. Torrence and Compo (1998) define the cone of influence as the boundary where the power of edge-related artifacts is damped by a factor of e^{-2}. In general, more certain flux contributions below the cone of influence include transport scales up to ≈⅓ of the flight length. For the flights in the XRC the less certain flux contributions above the cone of influence are small (−15% to −4% median differences for all flights). These contributions are included in the analysis, to ensure inclusion of identical transport scales close to the start or end, and at the center of a dataset, and to preserve the global covariance.

(iii) The spatial variation of temperature and precipitation in the XRC follows altitudinal and latitudinal trends (Auerswald et al., 2009; Wittmer et al., 2010). To resolve the effective state of biophysical surface properties over time, Metzger et al. (2013, Appendix E) use Moderate Resolution Imaging Spectroradiometer (MODIS) data. Here, 8-day composites of the land surface temperature (*LST*, MOD11A2.5, 1 km resolution), and 16-day composites of the enhanced vegetation index (*EVI,* MOD13Q1, MYD13Q1, 250 m resolution) are used. The *LST* and *EVI* datasets are bi-linearly interpolated to the 90 m resolution of an existing land cover classification, and linearly interpolated in time to yield an individual map for each flight day. Turbulence statistics for each 1000 m subinterval over a wavelet scalogram is used to evaluate the 2D footprint model described in Metzger et al. (2012, Appendix D). An individual evaluation is carried out for each overflown cell of the land cover, *LST* and *EVI* grids (i.e., every 90 m along the flight path). With the overflown grid cell as base point, and the footprint weights w_{xy} $\left(\sum w_{xy} = 1\right)$ for each grid cell with position x, y, relative to the base point, the footprint composition is calculated;

$$LST = \sum_x \sum_y w_{xy} \cdot LST_{xy}, \tag{5}$$

$$EVI = \sum_x \sum_y w_{xy} \cdot EVI_{xy}, \tag{6}$$

with the land surface temperature and enhanced vegetation index for each grid cell, LST_{xy} and EVI_{xy}, respectively. For graphical representation the footprint weights of all evaluations along a flight line are superimposed and normalized to a sum of unity. In Fig. 7 *LST* and *EVI* generally follow the land cover patterns, e.g. lower temperature and higher greenness for irrigated agriculture and marshland. However, it is also evident that the static land cover classification cannot reflect the current surface conditions. E.g., the marshland in the north-western quadrant appears dried-out (high *LST* and low *EVI*), while the steppe area in the north-eastern quadrant shows large variations in *LST*. I.e., the land cover classification represents the long-term effects of vegetation, climate, soil and topography. However, biophysical surface properties also vary significantly within land cover classes. This is likely as a function of geomorphological properties such as aspect, slope and soil type, but also due to the large variability of convective rainfall events across the study area (e.g., Schaffrath et al., 2011).

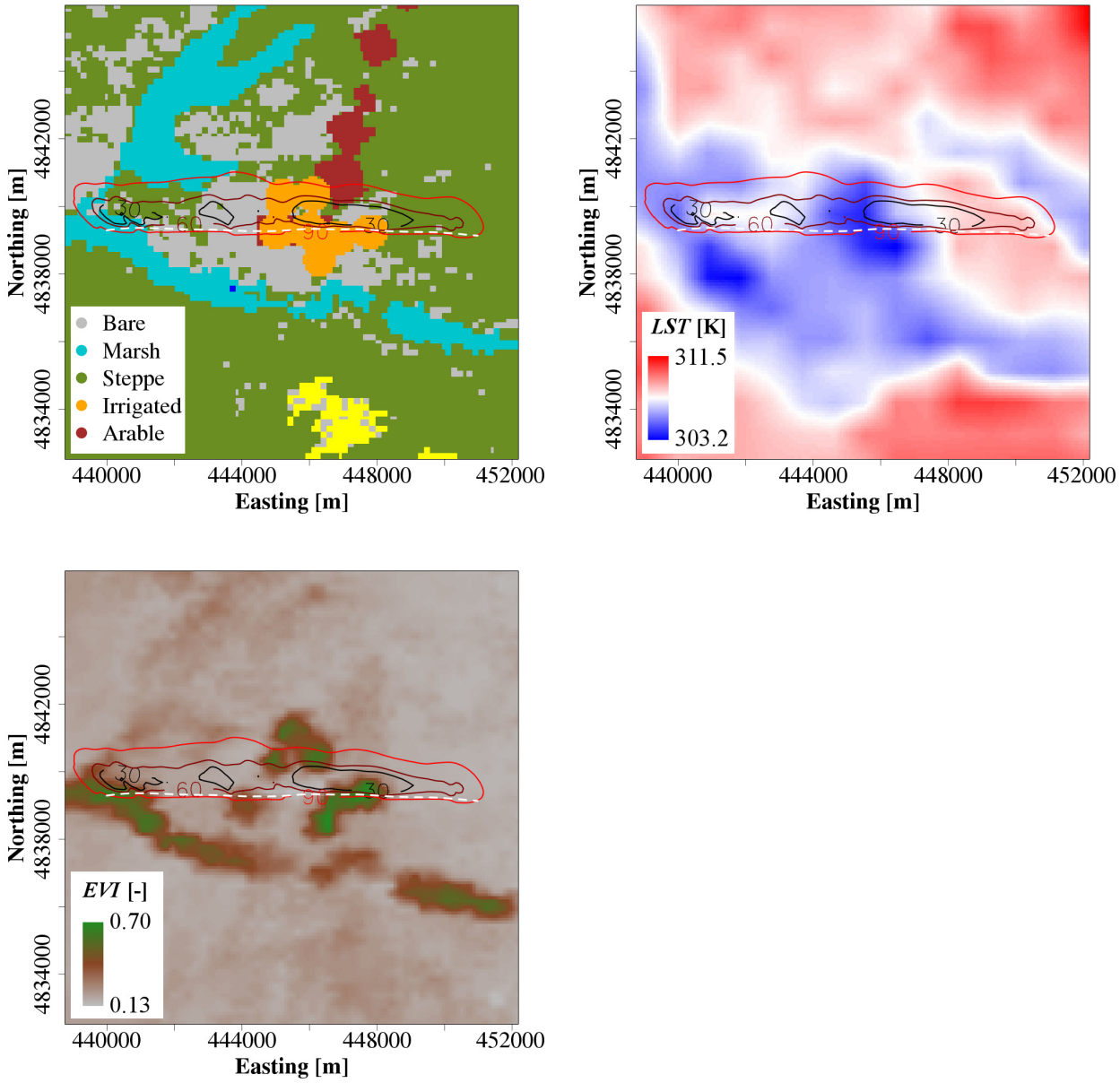


Fig. 7. Flight along pattern O12 on 8 July 2009, 12:16–12:24 Chinese standard time (UTC+8 h, white dashed line). The composite flux footprint along the flight line (30%, 60%, 90% contour lines) is superimposed over maps of land cover (upper left panel), land surface temperature (LST, upper right panel), and enhanced vegetation index (EVI, lower left panel). The land cover color codes are abbreviated for bare soil (Bare), marshland (Marsh), generic steppe (Steppe), irrigated agriculture (Irrigated), and rainfed agriculture (Arable). Figure from Metzger et al. (2013, Appendix E).

(iv) Environmental response functions (ERF, Desjardins et al., 1994) are an approach to utilize quantitative information about the EC measurement's spatial context. The general idea is to establish a functional relationship between spatially or temporally resolved flux observations (responses) and

corresponding environmental drivers. Metzger et al. (2013, Appendix E) base the development of a catchment-specific ERF on the works of Chen et al. (1999), Hutjes et al. (2010) and Ogunjemiyo et al. (2003). The LTFM procedure advances these approaches;

(a) Thus far, a suitable number of flux observations was obtained by either shortening the time-domain EC averaging interval (Chen et al., 1999; Ogunjemiyo et al., 2003), or by stratifying repeated observations along the same flight line on different days (Hutjes et al., 2010). The inherent drawbacks are the neglect of either long wavelength contributions to the flux measurement, or inter-day variability of ecosystem drivers. Both are overcome using the wavelet cross-scalogram technique;

(b) Previously, the development of ERFs has solely focused on drivers in the footprint of the flux observations, namely discrete land cover classifications. This procedure ignores within-class variability across a catchment, e.g. along climatic or altitudinal gradients, which is overcome by using continuous variables such as *LST* and *EVI* instead. Also, substituting discrete with continuous variables enables the use of more advanced scaling algorithms. In addition, the present approach considers meteorological drivers such as downwelling shortwave radiation ($S\downarrow$), mixing ratio (*MR*), and potential temperature (θ). This avoids the need to stratify or pre-select data, and enables constructing a single ERF that is valid for the observation period and, within in range of the measured variables, across a catchment of interest.

(c) Hitherto, ERFs were determined as the inverse of a linear mixing matrix, using either numerical (Chen et al., 1999) or regression methods (Hutjes et al., 2010; Ogunjemiyo et al., 2003). Such procedure assumes a linear relationship between drivers and responses, which is subject of on-going discussion and research (e.g., Raupach and Finnigan, 1995). Instead, the present approach uses boosted regression trees (BRT), a non-parametric machine learning technique, to establish an ERF between drivers and responses. In contrast to parametric approaches, BRT does not assume a predetermined form of the response, but constructs an ERF according to information in the data. BRT can fit complex nonlinear relationships, automatically handle interactions between drivers, and provide predictive performance that is superior to most traditional modeling methods (e.g., Hu et al., 2010).

Metzger et al. (2013, Appendix E) apply BRTs to N=8446 observations during 35 flights in the XRC (Sect. 2.4). The purpose is to extract the relationships between *H*, *LE* and land cover (*LST*, *EVI*) and meteorological ($S\downarrow$, *MR*, and θ) variables. In Fig. 8 the BRTs for *H* are summarized as partial dependence plots. These show the effect of each individual variable on the response after subtraction of the offset (161 W m^{-2}), and after accounting for the average effects of all other variables in the model. The partial dependence plots in Fig. 8 are sorted in order of the relative importance of the response variables (Friedman, 2001). The most important responses of *H* are non-linear (*LST*, θ), followed by linear responses ($S\downarrow$, *MR*, and *EVI*). With the exception of *MR* and *EVI* the individual responses are positive in sign.

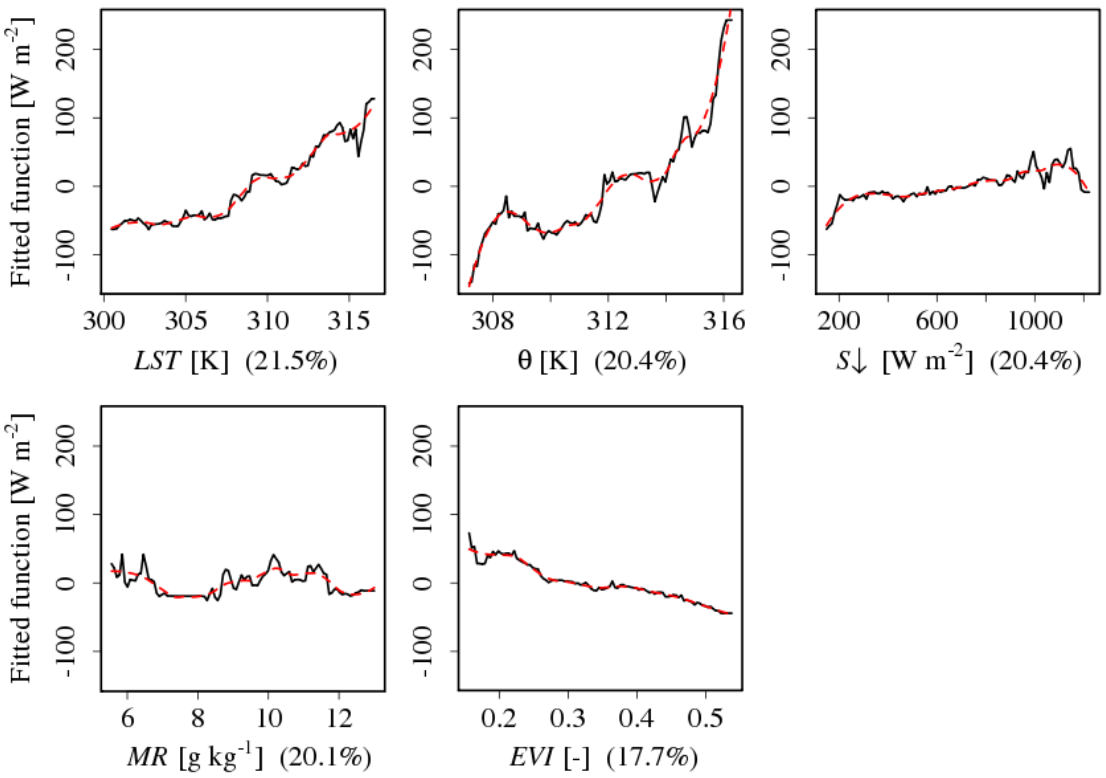


Fig. 8. Boosted regression tree partial response plots of *H* for all five state variables in order of their relative importance (in braces). The fitted function (black) shows the variable response of the BRT over the range of one individual state variable, while the remaining state variables are held at an average, constant value. The red dashed line is a smoothed representation of the fitted function (locally weighted polynomial regression). Figure from Metzger et al. (2013, Appendix E).

To evaluate the performance of the BRT models, MLFRs are established between BRT fitted values for *H* and *LE* and the observed fluxes. For both, *H* and *LE* the agreement between the BRT fitted values and the observed fluxes is excellent, with approximately unity slope and ≤1% median absolute deviation in the residuals. For the duration of each flight pattern, the trained BRT models are used to extrapolate *H* and *LE* across the XRC. For this purpose the median meteorological state variables during each flight pattern as well as topical grids of MODIS *LST* and *EVI* data are used. Grid cells that exceed the state space of the BRT training data set are excluded from extrapolation. Fig. 9 shows the resulting flux grids for three different days, with a spatial coverage of ≥92%. Despite the land cover classification was never used during the extrapolation process, several landscape units are clearly recognizable in the flux maps. For instance both, the Xilin river valley to the west, as well as the mountainous headwater area to the east display low sensible- and high latent heat fluxes. On the contrary, the non-vegetated basin on the northern tip shows consistently low evapotranspiration.

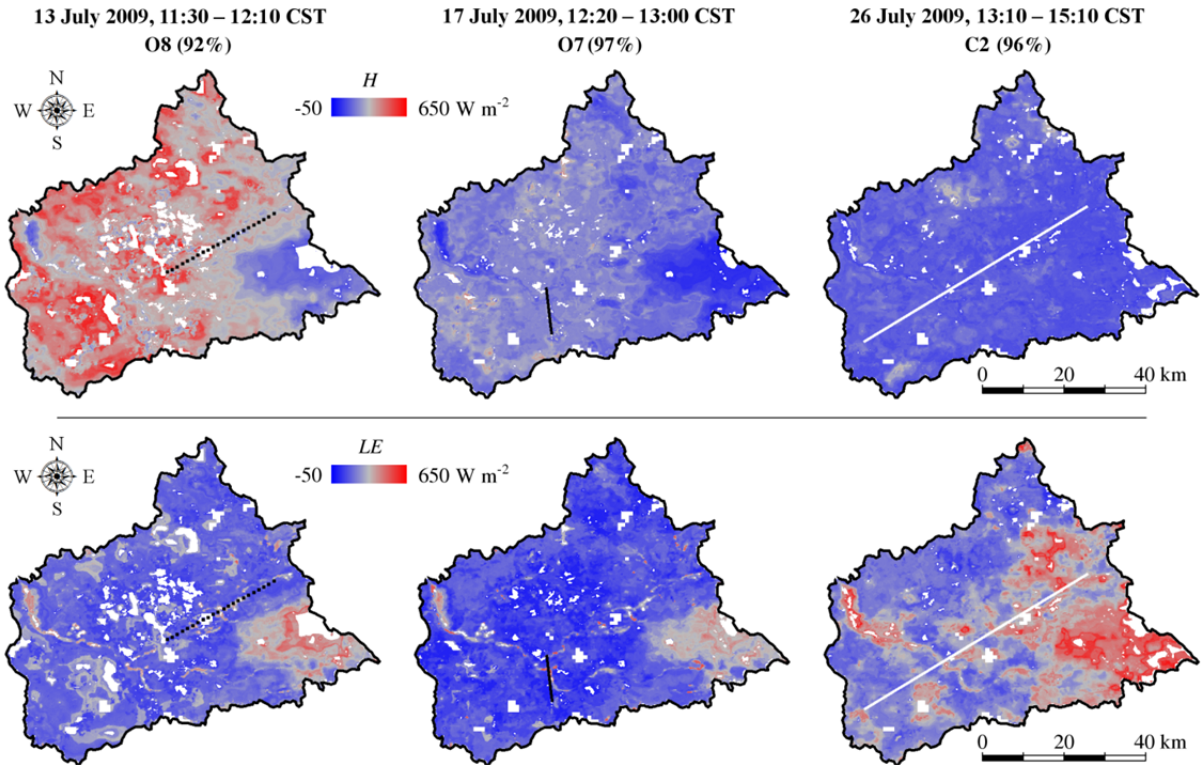


Fig. 9. Maps of LTFM predicted fluxes of sensible heat (*H*, top) and latent heat (*LE*, bottom) on 13, 17 and 26 July 2009 (left to right). Percentages in braces after the flight identifier (O8, O7, C2) indicate the spatial coverage of the prediction throughout the catchment. Meteorological state variables from the superimposed flight lines are used in the respective LTFM prediction. Figure taken from Metzger et al. (2013, Appendix E).

To assess the reliability of the LTFM method for spatially resolving and extrapolating turbulent fluxes, Metzger et al. (2013, Appendix E) establish an uncertainty budget (Table 2);

(a) Metzger et al. (2012, Appendix D) have shown that turbulent flux measurements with the WSMA platform and instrumentation are unbiased, and precise to ≤10%;

(b) For a single flux measurement the systematic and random components of the uncertainty due to the limited sampling size of turbulent eddies are <1±57% for *H* and <1±121% for *LE*, respectively;

(c) Small median residuals of 0±5% for *H* and 0±6% for *LE* between fitted and observed values emphasize that the BRT fitting technique is unbiased;

(d) However, potential biases in the LTFM procedure can result from using non-linear BRT response functions for prediction. E.g., deviations from the true value of *LST* of similar magnitude but opposite sign do not cancel out in the predictions. Metzger et al. (2013, Appendix E) quantify the resulting bias for three different test cases (Table 2: Spatio-temporal analysis, BRT response function, BRT state variables). The systematic differences

do not exceed 18%, and in all cases the 99.9% confidence intervals include unity slope. Hence, an accuracy of ≤20% can be assigned to the LTFM method.

Table 2. Median systematic ± random uncertainty terms for a single flux observation or grid cell throughout the LTFM procedure. Table from Metzger et al. (2013, Appendix E).

Source	*H*	*LE*
Instrumentation and Hardware	0±8%	0±10%
Turbulence sampling	0±57%	0±121%
Spatio-temporal analysis	2±40%	4±47%
BRT residuals	0±5%	0±6%
BRT response function	11±69%	18±77%
BRT state variables	13±77%	14±75%

Assuming normal distribution and independence, the random parts of all uncertainty terms in Table 2 can be combined to their Gaussian sum. Then, the ensemble random uncertainty quantifies the level of confidence that is expected from aggregating multiple observations or grid cells (e.g., Mahrt, 1998). The resulting ensemble random uncertainty for land cover specific flux estimates throughout the XRC ranges from <1% for steppe to 5% for settlements and irrigated agriculture Metzger et al. (2013, Appendix E).

4 Conclusions

The overarching goal of this dissertation is to evaluate the suitability of WSMA for gaining new insights in the spatial variability of the heat and moisture exchange above complex terrestrial surfaces. In pursue of this goal three objectives are addressed, and corresponding infrastructure and post-processing/analysis techniques are developed. With the WSMA a measurement platform is provided that enables (i) characterizing transport processes in the CBL, and (ii) quantifying the surface-air exchange over topographically structured and previously inaccessible terrain. Lastly, (iii) the LTFM analysis procedure aids the resolution and extrapolation of the surface-air exchange over non-homogenous landscapes. These techniques are developed using the example of terrain-following EC measurements with WSMA, but are general enough to be used with other airborne or ground-based EC measurement systems.

(i) Metzger et al. (2011, Appendix B) assess the suitability of WSMA based 3D wind vector measurement for EC applications. It is shown that carefully computed wind vector measurements using a WSMA are not inferior to those from other airborne platforms. A novel calibration approach is developed, which minimizes the influence of flow distortion from fuselage, propeller and wing in the dynamic pressure and flow angle computations. For flights including rising and sinking of the aircraft, such as during terrain-following EC applications, three independent lines of analysis yield comparable uncertainty. This convergence underpins the integrity of sensing elements and wind model description, and enables to quantify the overall operational uncertainty (RMSD) to 0.4 m s^{-1} for the horizontal and 0.3 m s^{-1} for the vertical wind components. These findings emphasize that the 3D wind vector can be measured reliably from a highly transportable and low-cost weight-shift microlight aircraft. Hence the necessary basis is provided for the study of precision and spectral quality of the wind measurement, which is prerequisite for reliable EC flux measurements.

(ii) Metzger et al. (2012, Appendix D) quantify the uncertainty in WSMA based measurements of turbulence statistics and EC flux. It is shown that turbulence measurements from a WSMA can be achieved with sufficient precision to enable EC flux calculation. Furthermore a coordinated setup of tall tower, LAS and WSMA measurements avoids typical platform inter-comparison errors due to averaging intervals and vertical flux divergence (e.g., Betts et al., 1990). Differences in the order of 15–25% remain between the fluxes measured by the ground based instruments and the WSMA. However, the 99.5% confidence intervals of the MLFRs between tower and WSMA include unity slope. Consequently the observed differences can be considered insignificant, and the precision of the WSMA flux measurement is quantified to ≤10% (1σ slope error). The energy transfer by NPEs above heterogeneous terrain is discussed as potential reason for the observed differences between spatially and temporally averaging flux measurements. Hence, the WSMA is a suitable tool to promote the on-going research of surface-atmosphere interactions over heterogeneous landscapes. The flux measurement is sufficiently accurate to cover the required length of the flight transects (10–100 km). Moreover, the WSMA's low ratio of true airspeed to climb rate is well suited for terrain-following flight over complex terrain. All of the above features are beneficial for the study of yet poorly understood exchange mechanisms between the earth's surface and the atmosphere.

This further substantiates the versatility of the WSMA as a low cost and widely applicable environmental research aircraft.

(iii) Metzger et al. (2013, Appendix E) demonstrate the usefulness of WSMA based EC flux measurements above complex and not readily accessible terrain. The WSMA is applied to EC flux measurements over the remote steppe of the Mongolian Plateau. This data basis is used to develop the LTFM analysis procedure, which 'mines' the information content of EC flux observations and extracts quantitative relationships with environmental drivers. The improvements of the LTFM procedure over existing methods are;

(a) The use of a WSMA with low airspeed and high climb rate enables low level flights at constant height even above topographically structured terrain;

(b) Wavelet decomposition of the turbulence data yields unprecedented spatial resolution of the flux observations. However, due to edge effects flux observations close to the start or end of a dataset can contain spectral artifacts. Alternative techniques such as empirical mode decomposition (Barnhart et al., 2012) might help to further improve the results.

(c) An 'offline' footprint parameterization considering 3D dispersion is suitable to map the differences in surface properties encountered by a flux aircraft. However, when adapting LTFM e.g. to ground-based measurements, the range of surface properties is likely to be significantly reduced. In order to improve the signal to noise level it might become important to also consider the local flow field, especially when measuring at greater heights. E.g. closure models with terrain-following coordinates (Hsieh and Katul, 2009; Sogachev and Lloyd, 2004) or 'online' Lagrangian dispersion modeling (Markkanen et al., 2010; Matross et al., 2011; Wang and Rotach, 2010) could aid this purpose.

(d) In spite of a static and discrete land cover classification, the LTFM method uses spatio-temporally continuous and topical information of biophysical surface properties. Only the continuous nature of MODIS land surface data enabled the use of the BRT machine learning technique. In this combination the climatic and altitudinal gradients throughout a heterogeneous steppe landscape are successfully reproduced. In the interest of further advancing LTFM, it is desirable to also consider the uncertainty in the observations during BRT fitting, and to explore alternative machine learning techniques such as support vector machines (e.g., Yang et al., 2007).

Current methods to spatially resolve surface fluxes are mainly focused on remote sensing algorithms (e.g., Fan et al., 2007) and process-based land surface models (e.g., Vetter et al., 2012). These procedures often demand far-reaching assumptions, such as the closure of the energy and water balances (e.g., Anderson et al., 2012), or are challenging with respect to the required data basis (e.g., Kaminski et al., 2012; Ziehn et al., 2011). In contrast, the LTFM approach allows inferring high-resolution surface flux maps directly from observational data with minimal and quantifiable assumptions.

The WSMA flux measurement and the LTFM procedure enable studying some of the mechanisms that can lead to systematically biased flux estimates from EC sensor networks. E.g., flux uncertainty and environmental variability can be partitioned through isolating and quantifying relevant land-

atmosphere exchange processes. Also, direct and indirect effects of surface heterogeneity, such as location bias and transport by NPEs, can be quantified and adjusted. Thereby the WSMA flux measurement provides novel possibilities to explore the sparsely investigated mechanisms that yield the regional land-atmosphere exchange. In combination with ground-based measurements, such information can significantly contribute to a spatio-temporal continuum of observations to foster our understanding of earth system dynamics. Analogously applying LTFM to long-term ground-based EC measurements could aid, e.g., constraining local to regional water budgets, distinguishing anthropogenic and natural sources/sinks in urban environments, and substantiating process-studies. Moreover, LTFM can be used to directly extrapolate flux measurements to the catchment scale, or to design, constrain and evaluate flux algorithms for remote sensing and numerical modeling applications. Hence, the infrastructure and methods developed in the course of this dissertation support and complement the rigorous quantification of the pools and fluxes of energy, water, carbon and nitrogen from long-term sensor networks. In consequence this might contribute to improving our understanding of climate change and its impact on natural resources, and to deduce appropriate mitigation strategies.

Notation

Abbreviations

3D	Three-dimensional
5HP	Five hole probe
a.g.l.	Above ground level
a.s.l.	Above sea level
Arable	Rainfed agriculture
Bare	Bare soil
BIAS	Bias (offset)
BRT	Boosted regression trees
CBL	Convective boundary layer
COMP	Wind calibration flight maneuver, see Metzger et al. (2011, Appendix B)
D–MIFU	Name of research aircraft
DFG	German Research Foundation
DWD	German Meteorological Service
EC	Eddy covariance
ERF	Environmental response function
IGS	Inertial measurement and global positioning system
IPCC	Intergovernmental Panel on Climate Change
IRGA	Infrared gas analyzer
Irrigated	Irrigated agriculture
LAS	Large-aperture scintillometer
LTFM	Low level flights, Time-frequency-, Footprint-, and Machine learning analyses
MAGIM	Matter fluxes in grasslands of Inner Mongolia as influenced by stocking rate
Marsh	Marshland
MLFR	Maximum likelihood functional relationship
MODIS	Moderate Resolution Imaging Spectroradiometer
NPE	Non-propagating eddies
RACE	Wind calibration flight maneuver, see Metzger et al. (2011, Appendix B)
RMSD	Root mean square deviation
SODAR	Sonic detection and ranging
SQUA	Wind calibration flight maneuver, see Metzger et al. (2011, Appendix B)
Steppe	Generic steppe

Abbreviations (continued)

ULS	Universal laser altitude sensor
UTC	Coordinated universal time
VARI	Wind calibration flight maneuver, see Metzger et al. (2011, Appendix B)
VW1–VW3	Wind calibration flight maneuver, see Metzger et al. (2011, Appendix B)
WSMA	Weight-shift microlight aircraft
XRC	Xilin River Catchment

Subscripts

∞	Free airstream
ens	Ensemble error
H	Sensible heat flux
m	Measured
LE	Latent heat flux
off	Offset
ran	Random error
res	Residual standard error
slo	Slope
upw	Upwash
u	Along-wind component
u_*	Friction velocity
v	Cross-wind component
w	Vertical wind component
x	Longitudinal coordinate
y	Latitudinal coordinate

Functions

Overbars denote averages, and primes denote deviations from the average.

Δ	Finite difference
σ	Standard deviation
Co	Cospectrum

Parameters and variables

α	Attack angle	[rad]
β	Sideslip angle	[rad]
ρ	Air density	[kg m^{-3}]
θ	Potential temperature	[K]
a	Absolute humidity	[g m^{-3}]
CL	Lift coefficient	[-]
e	Euler's number (≈2.71828)	[-]
EVI	Enhanced vegetation index	[-]
f	Measurement frequency	[Hz]
H	Sensible heat flux	[W m^{-2}]
L	Lift	[N]
LE	Latent heat flux	[W m^{-2}]
LST	Land surface temperature	[K]
MR	Mixing ratio	[g kg^{-1}]
n	Normalized frequency	[-]
N	Sample size	[-]
R^2	Weighted coefficient of determination	[-]
s	Scalar quantity (wildcard)	[-]
S	Surface area of wing	[m^2]
$S\downarrow$	Downwelling shortwave radiation	[W m^{-2}]
u_*	Friction velocity	[m s^{-1}]
t	Time	[s]
T	Air temperature	[K]
u	Along-wind speed	[m s^{-1}]
$\bar{U}$	Horizontal wind speed for the tower and the true airspeed for the WSMA	[m s^{-1}]
v	Cross-wind speed	[m s^{-1}]
v_{tas}	True airspeed	[m s^{-1}]
w_{upw}	Upwash velocity	[m s^{-1}]
w	Footprint weight	[-]
w	Vertical wind speed	[m s^{-1}]
x, y, z	Cartesian coordinates	[-]
z	Measurement height	[m]

References

Anderson, M. C., Kustas, W. P., Alfieri, J. G., Gao, F., Hain, C., Prueger, J. H., Evett, S., Colaizzi, P., Howell, T., and Chávez, J. L.: Mapping daily evapotranspiration at Landsat spatial scales during the BEAREX'08 field campaign, Adv. Water Res., 50, 162-177, doi:10.1016/j.advwatres.2012.06.005, 2012.

Aubinet, M., Vesala, T., and Papale, D., (Eds.): Eddy covariance: A practical guide to measurement and data analysis, Springer, Dordrecht, Heidelberg, London, New York, 438 pp., 2012.

Auerswald, K., Wittmer, M. H. O. M., Männel, T. T., Bai, Y. F., Schäufele, R., and Schnyder, H.: Large regional-scale variation in C3/C4 distribution pattern of Inner Mongolia steppe is revealed by grazer wool carbon isotope composition, Biogeosciences, 6, 795-805, doi:10.5194/bg-6-795-2009, 2009.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., U, K., Pilegaard, K., Schmid, H., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, Bull. Am. Meteorol. Soc., 82, 2415-2434, doi:10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2, 2001.

Bange, J., Zittel, P., Spiess, T., Uhlenbrock, J., and Beyrich, F.: A new method for the determination of area-averaged turbulent surface fluxes from low-level flights using inverse models, Boundary Layer Meteorol., 119, 527-561, doi:10.1007/s10546-005-9040-6, 2006.

Barnhart, B. L., Eichinger, W. E., and Prueger, J. H.: Introducing an Ogive method for discontinuous data, Agric. For. Meteorol., 162–163, 58-62, doi:10.1016/j.agrformet.2012.04.003, 2012.

Betts, A. K., Desjardins, R. L., Macpherson, J. I., and Kelly, R. D.: Boundary-layer heat and moisture budgets from FIFE, Boundary Layer Meteorol., 50, 109-138, doi: 10.1007/BF00120520, 1990.

Chen, B., Coops, N. C., Fu, D., Margolis, H. A., Amiro, B. D., Barr, A. G., Black, T. A., Arain, M. A., Bourque, C. P. A., Flanagan, L. B., Lafleur, P. M., McCaughey, J. H., and Wofsy, S. C.: Assessing eddy-covariance flux tower location bias across the Fluxnet-Canada Research Network based on remote sensing and footprint modelling, Agric. For. Meteorol., 151, 87-100, 10.1016/j.agrformet.2010.09.005, 2011.

Chen, B., Coops, N. C., Fu, D., Margolis, H. A., Amiro, B. D., Black, T. A., Arain, M. A., Barr, A. G., Bourque, C. P. A., Flanagan, L. B., Lafleur, P. M., McCaughey, J. H., and Wofsy, S. C.: Characterizing spatial representativeness of flux tower eddy-covariance measurements across the Canadian Carbon Program Network using remote sensing and footprint analysis, Remote Sens. Environ., 124, 742-755, doi:10.1016/j.rse.2012.06.007, 2012.

Chen, J., Jonsson, P., Tamura, M., Gu, Z. H., Matsushita, B., and Eklundh, L.: A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter, Remote Sens. Environ., 91, 332-344, doi:10.1016/j.rse.2004.03.014, 2004.

Chen, J. M., Leblanc, S. G., Cihlar, J., Desjardins, R. L., and MacPherson, J. I.: Extending aircraft- and tower-based CO_2 flux measurements to a boreal region using a Landsat thematic mapper land cover map, J. Geophys. Res. Atmos., 104, 16859-16877, doi:10.1029/1999JD900129, 1999.

Crawford, T., Dobosy, R., and Dumas, E.: Aircraft wind measurement considering lift-induced upwash, Boundary Layer Meteorol., 80, 79-94, doi:10.1007/BF00119012, 1996.

Deardorff, J. W.: Three-dimensional numerical study of turbulence in an entraining mixed layer, Boundary Layer Meteorol., 7, 199-226, doi:10.1007/bf00227913, 1974.

Dellwik, E., Mann, J., and Larsen, K. S.: Flow tilt angles near forest edges – Part 1: Sonic anemometry, Biogeosciences, 7, 1745-1757, doi:10.5194/bg-7-1745-2010, 2010.

Desjardins, R. L., MacPherson, J. I., Schuepp, P. H., and Hayhoe, H. N.: Airborne flux measurements of CO_2, sensible, and latent heat over the Hudson Bay lowland, J. Geophys. Res. Atmos., 99, 1551-1561, doi:10.1029/93JD01296, 1994.

Desjardins, R. L., MacPherson, J. I., Mahrt, L., Schuepp, P., Pattey, E., Neumann, H., Baldocchi, D., Wofsy, S., Fitzjarrald, D., McCaughey, H., and Joiner, D. W.: Scaling up flux measurements for the boreal forest using aircraft-tower combinations, J. Geophys. Res. Atmos., 102, 29125-29133, doi:10.1029/97JD00278, 1997.

Dias, N., Gonçalves, J., Freire, L., Hasegawa, T., and Malheiros, A.: Obtaining potential virtual temperature profiles, entrainment fluxes, and spectra from mini unmanned aerial vehicle data, Boundary Layer Meteorol., 145, 93-111, doi:10.1007/s10546-011-9693-2, 2012.

Egger, J., Bajrachaya, S., Heinrich, R., Kolb, P., Lammlein, S., Mech, M., Reuder, J., Schaper, W., Shakya, P., Schween, J., and Wendt, H.: Diurnal winds in the Himalayan Kali Gandaki valley. Part 3: Remotely piloted aircraft soundings, Monthly Weather Review, 130, 2042-2058, doi:10.1175/1520-0493(2002)130<2042:DWITHK>2.0.CO;2, 2002.

Emeis, S., Schäfer, K., and Münkel, C.: Surface-based remote sensing of the mixing-layer height - a review, Meteorol. Z., 17, 621-630, doi:10.1127/0941-2948/2008/0312, 2008.

Fan, L., Liu, S., Bernhofer, C., Liu, H., and Berger, F. H.: Regional land surface energy fluxes by satellite remote sensing in the Upper Xilin River Watershed (Inner Mongolia, China), Theor. Appl. Climatol., 88, 231-245, doi:10.1007/s00704-006-0241-9, 2007.

Finnigan, J.: An introduction to flux measurements in difficult conditions, Ecological Applications, 18, 1340-1350, doi:10.1890/07-2105.1, 2008.

Finnigan, J. J., Clement, R., Malhi, Y., Leuning, R., and Cleugh, H. A.: A re-evaluation of long-term flux measurement techniques. Part 1: Averaging and coordinate rotation, Boundary Layer Meteorol., 107, 1-48, doi:10.1023/A:1021554900225, 2003.

Foken, T.: Micrometeorology, Springer, Berlin, Heidelberg, 306 pp., 2008a.

Foken, T.: The energy balance closure problem: An overview, Ecological Applications, 18, 1351-1367, doi:10.1890/06-0922.1, 2008b.

Foken, T., Mauder, M., Liebethal, C., Wimmer, F., Beyrich, F., Leps, J.-P., Raasch, S., DeBruin, H. A. R., Meijninger, W. M. L., and Bange, J.: Energy balance closure for the LITFASS-2003 experiment, Theor. Appl. Climatol., 101, 149-160, doi:10.1007/s00704-009-0216-8, 2010.

Foken, T., Aubinet, M., Finnigan, J. J., Leclerc, M. Y., Mauder, M., and Paw U, K. T.: Results of a panel discussion about the energy balance closure correction for trace gases, Bull. Am. Meteorol. Soc., 92, ES13-ES18, doi:10.1175/2011bams3130.1, 2011.

Foken, T., Aubinet, M., and Leuning, R.: The eddy-covariance method, in: Eddy covariance: A practical guide to measurement and data analysis, edited by: Aubinet, M., Vesala, T., and Papale, D., Springer, Dordrecht, Heidelberg, London, New York, 1-19, 2012.

Friedman, J. H.: Greedy function approximation: A gradient boosting machine, Ann Stat, 29, 1189-1232, doi:10.1214/aos/1013203451, 2001.

Garman, K. E., Hill, K. A., Wyss, P., Carlsen, M., Zimmerman, J. R., Stirm, B. H., Carney, T. Q., Santini, R., and Shepson, P. B.: An airborne and wind tunnel evaluation of a wind turbulence measurement system for aircraft-based flux measurements, J. Atmos. Oceanic Technol., 23, 1696-1708, doi:10.1175/JTECH1940.1, 2006.

Garman, K. E., Wyss, P., Carlsen, M., Zimmerman, J. R., Stirm, B. H., Carney, T. Q., Santini, R., and Shepson, P. B.: The contribution of variability of lift-induced upwash to the uncertainty in vertical winds determined from an aircraft platform, Boundary Layer Meteorol., 126, 461-476, doi:10.1007/s10546-007-9237-y, 2008.

Göckede, M., Foken, T., Aubinet, M., Aurela, M., Banza, J., Bernhofer, C., Bonnefond, J. M., Brunet, Y., Carrara, A., Clement, R., Dellwik, E., Elbers, J., Eugster, W., Fuhrer, J., Granier, A., Gruenwald, T., Heinesch, B., Janssens, I. A., Knohl, A., Koeble, R., Laurila, T., Longdoz, B., Manca, G., Marek, M., Markkanen, T., Mateus, J., Matteucci, G., Mauder, M., Migliavacca, M., Minerbi, S., Moncrieff, J., Montagnani, L., Moors, E., Ourcival, J. M., Papale, D., Pereira, J., Pilegaard, K., Pita, G., Rambal, S., Rebmann, C., Rodrigues, A., Rotenberg, E., Sanz, M. J., Sedlak, P., Seufert, G., Siebicke, L., Soussana, J. F., Valentini, R., Vesala, T., Verbeeck, H., and Yakir, D.: Quality control of CarboEurope flux data - Part 1: Coupling footprint analyses with flux data quality assessment to evaluate sites in forest ecosystems, Biogeosciences, 5, 433-450, doi:10.5194/bg-5-433-2008, 2008.

Hiyama, T., Strunin, M. A., Tanaka, H., and Ohta, T.: The development of local circulations around the Lena River and their effect on tower-observed energy imbalance, Hydrol. Processes, 21, 2038-2048, doi:10.1002/hyp.6705, 2007.

Hobbs, S., Dyer, D., Courault, D., Olioso, A., Lagouarde, J. P., Kerr, Y., McAneney, J., and Bonnefond, J.: Surface layer profiles of air temperature and humidity measured from unmanned aircraft, Agronomie, 22, 635-640, doi:10.1051/agro:2002050, 2002.

Hopkin, M.: Ecology: Spying on nature, Nature, 444, 420-421, doi:10.1038/444420a, 2006.

Hsieh, C.-I., and Katul, G.: The Lagrangian stochastic model for estimating footprint and water vapor fluxes over inhomogeneous surfaces, Int. J. Biometeorol., 53, 87-100, doi:10.1007/s00484-008-0193-0, 2009.

Hu, F. S., Higuera, P. E., Walsh, J. E., Chapman, W. L., Duffy, P. A., Brubaker, L. B., and Chipman, M. L.: Tundra burning in Alaska: Linkages to climatic change and sea ice retreat, J. Geophys. Res., 115, G04002, doi:10.1029/2009jg001270, 2010.

Hutjes, R. W. A., Vellinga, O. S., Gioli, B., and Miglietta, F.: Dis-aggregation of airborne flux measurements using footprint analysis, Agric. For. Meteorol., 150, 966-983, doi:10.1016/j.agrformet.2010.03.004, 2010.

IPCC: Climate change 1990 - Scientific assessment of climate change, contribution of working group 1 to the first assessment report of the Intergovernmental Panel on Climate Change, Intergovernmental Panel on Climate Change, Cambridge, United Kingdom, 410 pp., 1990.

IPCC: Climate change 1995 - The science of climate change, contribution of working group 1 to the second assessment report of the Intergovernmental Panel on Climate Change, Intergovernmental Panel on Climate Change, Cambridge, United Kingdom, 588 pp., 1995.

IPCC: Climate change 2001 - The scientific basis, contribution of working group 1 to the third assessment report of the Intergovernmental Panel on Climate Change, Intergovernmental Panel on Climate Change, Cambridge, United Kingdom, 881 pp., 2001.

IPCC: Climate change 2007 - The physical science basis, contribution of working group 1 to the fourth assessment report of the Intergovernmental Panel on Climate Change, Intergovernmental Panel on Climate Change, Cambridge, United Kingdom, 996 pp., 2007.

Isaac, P. R., McAneney, J., Leuning, R., and Hacker, J. M.: Comparison of aircraft and ground-based flux measurements during OASIS95, Boundary Layer Meteorol., 110, 39-67, doi:10.1023/A:1026002301152, 2004.

Jung, M., Reichstein, M., Margolis, H. A., Cescatti, A., Richardson, A. D., Arain, M. A., Arneth, A., Bernhofer, C., Bonal, D., Chen, J. Q., Gianelle, D., Gobron, N., Kiely, G., Kutsch, W., Lasslop, G., Law, B. E., Lindroth, A., Merbold, L., Montagnani, L., Moors, E. J., Papale, D., Sottocornola, M., Vaccari, F., and Williams, C.: Global patterns of land-atmosphere fluxes of carbon dioxide, latent heat, and sensible heat derived from eddy covariance, satellite, and meteorological observations, J. Geophys. Res. Biogeosci., 116, G00J07, doi:10.1029/2010jg001566, 2011.

Junkermann, W.: An ultralight aircraft as platform for research in the lower troposphere: System performance and first results from radiation transfer studies in stratiform aerosol layers and broken cloud conditions, J. Atmos. Oceanic Technol., 18, 934-946, doi:10.1175/1520-0426(2001)018<0934:AUAAPF>2.0.CO;2, 2001.

Junkermann, W.: The actinic UV-radiation budget during the ESCOMPTE campaign 2001: Results of airborne measurements with the microlight research aircraft D-MIFU, Atmos. Res., 74, 461-475, doi:10.1016/j.atmosres.2004.06.009, 2005.

Kaimal, J. C., Wyngaard, J. C., Izumi, Y., and Coté, O. R.: Spectral characteristics of surface-layer turbulence, Q. J. R. Meteorolog. Soc., 98, 563-589, doi:10.1002/qj.49709841707, 1972.

Kaimal, J. C., and Finnigan, J. J.: Atmospheric boundary layer flows: Their structure and measurement, Oxford University Press, New York, 289 pp., 1994.

Kalogiros, J. A., and Wang, Q.: Calibration of a radome-differential GPS system on a Twin Otter research aircraft for turbulence measurements, J. Atmos. Oceanic Technol., 19, 159-171, doi:10.1175/1520-0426(2002)019%3C0159:COARDG%3E2.0.CO;2, 2002a.

Kalogiros, J. A., and Wang, Q.: Aerodynamic effects on wind turbulence measurements with research aircraft, J. Atmos. Oceanic Technol., 19, 1567-1576, doi:10.1175/1520-0426(2002)019<1567:AEOWTM>2.0.CO;2, 2002b.

Kaminski, T., Rayner, P. J., Voßbeck, M., Scholze, M., and Koffi, E.: Observing the continental-scale carbon balance: Assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design, Atmos. Chem. Phys., 12, 7867-7879, doi:10.5194/acp-12-7867-2012, 2012.

Ketzer, B., Liu, H., and Bernhofer, C.: Surface characteristics of grasslands in Inner Mongolia as detected by micrometeorological measurements, Int. J. Biometeorol., 52, 563-574, doi:10.1007/s00484-008-0148-5, 2008.

Khelif, D., Burns, S. P., and Friehe, C. A.: Improved wind measurements on research aircraft, J. Atmos. Oceanic Technol., 16, 860-875, doi:10.1175/1520-0426(1999)016<0860:IWMORA>2.0.CO;2, 1999.

Kochendorfer, J., Meyers, T. P., Frank, J., Massman, W. J., and Heuer, M. W.: How well can we measure the vertical wind speed? Implications for fluxes of energy and mass, Boundary Layer Meteorol., in press, doi:10.1007/s10546-012-9738-1, 2012.

Kustas, W. P., Anderson, M. C., French, A. N., and Vickers, D.: Using a remote sensing field experiment to investigate flux-footprint relations and flux sampling distributions for tower and aircraft-based observations, Adv. Water Res., 29, 355-368, doi:10.1016/j.advwatres.2005.05.003, 2006.

Lenschow, D., and Sun, J.: The spectral composition of fluxes and variances over land and sea out to the mesoscale, Boundary Layer Meteorol., 125, 63-84, doi:10.1007/s10546-007-9191-8, 2007.

Lenschow, D. H., Wyngaard, J. C., and Pennell, W. T.: Mean-field and second-moment budgets in a baroclinic, convective boundary layer, Journal of the Atmospheric Sciences, 37, 1313-1326, doi:10.1175/1520-0469(1980)037<1313:mfasmb>2.0.co;2, 1980.

Lenschow, D. H., and Stankov, B. B.: Length scales in the convective boundary layer, Journal Of The Atmospheric Sciences, 43, 1198-1209, doi:10.1175/1520-0469(1986)043<1198:LSITCB>2.0.CO;2, 1986.

Lenschow, D. H., Mann, J., and Kristensen, L.: How long is long enough when measuring fluxes and other turbulence statistics?, J. Atmos. Oceanic Technol., 11, 661-673, doi:10.1175/1520-0426(1994)011<0661:HLILEW>2.0.CO;2, 1994.

Leuning, R., van Gorsel, E., Massman, W. J., and Isaac, P. R.: Reflections on the surface energy imbalance problem, Agric. For. Meteorol., 156, 65-74, doi:10.1016/j.agrformet.2011.12.002, 2012.

Mahrt, L.: Flux sampling errors for aircraft and towers, J. Atmos. Oceanic Technol., 15, 416-429, doi:10.1175/1520-0426(1998)015<0416:FSEFAA>2.0.CO;2, 1998.

Mahrt, L.: Surface heterogeneity and vertical structure of the boundary layer, Boundary Layer Meteorol., 96, 33-62, doi:10.1023/a:1002482332477, 2000.

Mahrt, L.: Computing turbulent fluxes near the surface: Needed improvements, Agric. For. Meteorol., 150, 501-509, doi:10.1016/j.agrformet.2010.01.015, 2010.

Markkanen, T., Steinfeld, G., Kljun, N., Raasch, S., and Foken, T.: A numerical case study on footprint model performance under inhomogeneous flow conditions, Meteorol. Z., 19, 539-547, doi:10.1127/0941-2948/2010/0488, 2010.

Martin, S., Bange, J., and Beyrich, F.: Meteorological profiling of the lower troposphere using the research UAV "M^{2}AV Carolo", Atmos. Meas. Tech., 4, 705-716, doi:10.5194/amt-4-705-2011, 2011.

Mason, P. J.: The formation of areally-averaged roughness lengths, Q. J. R. Meteorolog. Soc., 114, 399-420, doi:10.1002/qj.49711448007, 1988.

Massman, W. J., and Clement, R.: Uncertainty in eddy covariance flux estimates resulting from spectral attenuation, in: Handbook of micrometeorology: A guide for surface flux measurement and analysis, 1 ed., edited by: Lee, X., Law, B., and Massman, W., Springer, Dordrecht, 67-100, 2004.

Matross, D. M., Andrews, A., Pathmathevan, M., Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Lee, J. T., Zhao, C., Bakwin, P. S., Munger, J. W., and Hollinger, D. Y.: Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data, Tellus B, 58, 344-358, doi:10.1111/j.1600-0889.2006.00206.x, 2011.

Mauder, M., Desjardins, R. L., and MacPherson, I.: Scale analysis of airborne flux measurements over heterogeneous terrain in a boreal ecosystem, J. Geophys. Res. Atmos., 112, D13112, doi:10.1029/2006JD008133, 2007a.

Mauder, M., Jegede, O. O., Okogbue, E. C., Wimmer, F., and Foken, T.: Surface energy balance measurements at a tropical site in West Africa during the transition from dry to wet season, Theor. Appl. Climatol., 89, 171-183, doi:10.1007/s00704-006-0252-6, 2007b.

Mauder, M., Oncley, S., Vogt, R., Weidinger, T., Ribeiro, L., Bernhofer, C., Foken, T., Kohsiek, W., De Bruin, H., and Liu, H.: The energy balance experiment EBEX-2000. Part II: Intercomparison of eddy-covariance sensors and post-field data processing methods, Boundary Layer Meteorol., 123, 29-54, doi:10.1007/s10546-006-9139-4, 2007c.

Metzger, S., Junkermann, W., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Corrigendum to "Measuring the 3-D wind vector with a weight-shift microlight aircraft" published in Atmos. Meas. Tech., 4, 1421–1444, 2011, Atmos. Meas. Tech., 4, 1515-1539, doi:10.5194/amt-4-1515-2011, 2011.

Metzger, S., Junkermann, W., Mauder, M., Beyrich, F., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Eddy-covariance flux measurements with a weight-shift microlight aircraft, Atmos. Meas. Tech., 5, 1699-1717, doi:10.5194/amt-5-1699-2012, 2012.

Metzger, S., Junkermann, W., Mauder, M., Butterbach-Bahl, K., Trancón y Widemann, B., Neidl, F., Schäfer, K., Wieneke, S., Zheng, X. H., Schmid, H. P., and Foken, T.: Spatially explicit regionalization of airborne flux measurements using environmental response functions, Biogeosciences, 10, 2193-2217, doi:10.5194/bg-10-2193-2013, 2013.

Ogunjemiyo, S. O., Kaharabata, S. K., Schuepp, P. H., MacPherson, I. J., Desjardins, R. L., and Roberts, D. A.: Methods of estimating CO_2, latent heat and sensible heat fluxes from estimates of land cover fractions in the flux footprint, Agric. For. Meteorol., 117, 125-144, doi:10.1016/S0168-1923(03)00061-3, 2003.

Panofsky, H. A., and Dutton, J. A.: Atmospheric turbulence: Models and methods for engineering applications, John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore, 397 pp., 1984.

R Development Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2012.

Raupach, M. R., and Finnigan, J. J.: Scale issues in boundary-layer meteorology: Surface energy balances in heterogeneous terrain, Hydrol. Processes, 9, 589-612, doi:10.1002/hyp.3360090509, 1995.

Ripley, B. D., and Thompson, M.: Regression techniques for the detection of analytical bias, Analyst, 112, 377-383, doi:10.1039/an9871200377, 1987.

Ruppert, J., Thomas, C., and Foken, T.: Scalar similarity for relaxed eddy accumulation methods, Boundary Layer Meteorol., 120, 39-63, doi:10.1007/s10546-005-9043-3, 2006.

Schaffrath, D., Barthold, F., and Bernhofer, C.: Spatiotemporal variability of grassland vegetation cover in a catchment in Inner Mongolia, China, derived from MODIS data products, Plant and Soil, 340, 181-198, doi:10.1007/s11104-010-0465-4, 2011.

Schmid, H. P., and Lloyd, C. R.: Spatial representativeness and the location bias of flux footprints over inhomogeneous areas, Agric. For. Meteorol., 93, 195-209, doi:10.1016/s0168-1923(98)00119-1, 1999.

Sogachev, A., and Lloyd, J.: Using a one-and-a-half order closure model of the atmospheric boundary layer for surface flux footprint estimation, Boundary Layer Meteorol., 112, 467-502, doi:10.1023/B:BOUN.0000030664.52282.ee, 2004.

Steinfeld, G., Letzel, M. O., Raasch, S., Kanda, M., and Inagaki, A.: Spatial representativeness of single tower measurements and the imbalance problem with eddy-covariance fluxes: Results of a large-eddy simulation study, Boundary Layer Meteorol., 123, 77-98, doi:10.1007/s10546-006-9133-x, 2007.

Strunin, M. A., Hiyama, T., Asanuma, J., and Ohata, T.: Aircraft observations of the development of thermal internal boundary layers and scaling of the convective boundary layer over non-homogeneous land surfaces, Boundary Layer Meteorol., 111, 491-522, doi:10.1023/B:BOUN.0000016542.72958.e9, 2004.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Kluwer Academic Publishers, Dordrecht, 670 pp., 1988.

Taylor, G. I.: Eddy motion in the atmosphere, Philosophical Transactions of the Royal Society of London, A 215, 1-26, doi:10.1098/rsta.1915.0001, 1915.

Thomas, R. M., Lehmann, K., Nguyen, H., Jackson, D. L., Wolfe, D., and Ramanathan, V.: Measurement of turbulent water vapor fluxes using a lightweight unmanned aerial vehicle system, Atmos. Meas. Tech., 5, 243-257, doi:10.5194/amt-5-243-2012, 2012.

Tjernström, M., and Friehe, C. A.: Analysis of a radome air-motion system on a twin-jet aircraft for boundary-layer research, J. Atmos. Oceanic Technol., 8, 19-40, doi:10.1175/1520-0426(1991)008%3C0019:AOARAM%3E2.0.CO;2, 1991.

Torrence, C., and Compo, G. P.: A practical guide to wavelet analysis, Bull. Am. Meteorol. Soc., 79, 61-78, doi:10.1175/1520-0477(1998)079<0061:apgtwa>2.0.co;2, 1998.

Uhlenbrock, J., Raasch, S., Hennemuth, B., Zittel, P., and Meijninger, W. M. L.: Effects of land surface heterogeneities on the boundary structure and turbulence during LITFASS-2003: Large-eddy simulations in comparison with turbulence measurements, 16th AMS Symposium on Boundary Layers and Turbulence, Portland, U.S.A., 8-13 August, 2004.

van den Kroonenberg, A., Martin, T., Buschmann, M., Bange, J., and Vorsmann, P.: Measuring the wind vector using the autonomous mini aerial vehicle M^2AV, J. Atmos. Oceanic Technol., 25, 1969-1982, doi:10.1175/2008JTECHA1114.1, 2008.

Vetter, S., Schaffrath, D., and Bernhofer, C.: Spatial simulation of evapotranspiration of semi-arid Inner Mongolian grassland based on MODIS and eddy covariance data, Environmental Earth Sciences, 65, 1567-1574, doi:10.1007/s12665-011-1187-5, 2012.

Vickers, D., and Mahrt, L.: Quality control and flux sampling problems for tower and aircraft data, J. Atmos. Oceanic Technol., 14, 512-526, doi:10.1175/1520-0426(1997)014<0512:QCAFSP>2.0.CO;2, 1997.

Wang, W., and Rotach, M. W.: Flux footprints over an undulating surface, Boundary Layer Meteorol., 136, 325-340, doi:10.1007/s10546-010-9498-8, 2010.

Williams, A., and Marcotte, D.: Wind measurements on a maneuvering twin-engine turboprop aircraft accounting for flow distortion, J. Atmos. Oceanic Technol., 17, 795-810, doi:10.1175/1520-0426(2000)017%3C0795:WMOAMT%3E2.0.CO;2, 2000.

Williams, M., Richardson, A. D., Reichstein, M., Stoy, P. C., Peylin, P., Verbeeck, H., Carvalhais, N., Jung, M., Hollinger, D. Y., Kattge, J., Leuning, R., Luo, Y., Tomelleri, E., Trudinger, C. M., and Wang, Y. P.: Improving land surface models with FLUXNET data, Biogeosciences, 6, 1341-1359, doi:10.5194/bg-6-1341-2009, 2009.

Wilson, K., Goldstein, A., Falge, E., Aubinet, M., Baldocchi, D., Berbigier, P., Bernhofer, C., Ceulemans, R., Dolman, H., Field, C., Grelle, A., Ibrom, A., Law, B. E., Kowalski, A., Meyers, T., Moncrieff, J., Monson, R., Oechel, W., Tenhunen, J., Valentini, R., and Verma, S.: Energy balance closure at FLUXNET sites, Agric. For. Meteorol., 113, 223-243, doi:10.1016/s0168-1923(02)00109-0, 2002.

Wittmer, M. H. O. M., Auerswald, K., Bai, Y., Schäufele, R., and Schnyder, H.: Changes in the abundance of C3/C4 species of Inner Mongolia grassland: Evidence from isotopic composition of soil and vegetation, Global Change Biol., 16, 605-616, doi:10.1111/j.1365-2486.2009.02033.x, 2010.

Wood, N., and Mason, P.: The influence of static stability on the effective roughness lengths for momentum and heat transfer, Q. J. R. Meteorolog. Soc., 117, 1025-1056, doi:10.1002/qj.49711750108, 1991.

Yang, F., Ichii, K., White, M. A., Hashimoto, H., Michaelis, A. R., Votava, P., Zhu, A. X., Huete, A., Running, S. W., and Nemani, R. R.: Developing a continental-scale measure of gross primary production by combining MODIS and AmeriFlux data through support vector machine approach, Remote Sens. Environ., 110, 109-122, doi:10.1016/j.rse.2007.02.016, 2007.

Yu, G.-R., Wen, X.-F., Sun, X.-M., Tanner, B. D., Lee, X., and Chen, J.-Y.: Overview of ChinaFLUX and evaluation of its eddy covariance measurement, Agric. For. Meteorol., 137, 125-137, doi:10.1016/j.agrformet.2006.02.011, 2006.

Ziehn, T., Knorr, W., and Scholze, M.: Investigating spatial differentiation of model parameters in a carbon cycle data assimilation system, Global Biogeochem. Cycles, 25, GB2021, doi:10.1029/2010gb003886, 2011.

List of Appendices

Appendix A: Individual contributions to the joint publications

This cumulative dissertation consists of publications that were composed in close cooperation with other researchers. This section specifies the individual contributions to the manuscripts in Appendices B–D.

Appendix B: Metzger et al. (2011)

Metzger, S., Junkermann, W., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Corrigendum to "Measuring the 3-D wind vector with a weight-shift microlight aircraft" published in Atmos. Meas. Tech., 4, 1421–1444, 2011, Atmos. Meas. Tech., 4, 1515-1539, doi:10.5194/amt-4-1515-2011, 2011.

Publisher's Note: *Due to mistakes on the publisher's side it was necessary to revise the original article Atmos. Meas. Tech., 4, 1421–1444, 2011 in this Corrigendum. Please use this Corrigendum as the main document.*

I alone developed, established and verified the seven-step wind calibration procedure incl. the net-upwash correction, and the end-to-end uncertainty budget as presented in the paper. It was me who calibrated the sensors and analyzed the wind tunnel measurements. The installation of the acceleration sensors in the five hole probe and the wing hangpoint, as well as the design of the measurement campaigns arose from my ideas. I implemented the described algorithms in a software tool and also wrote the entire text of the manuscript.

W. Junkermann installed and maintained all scientific equipment on the research aircraft. He also maintained and certified the aircraft itself, obtained flight permissions, performed the research flights, and edited the manuscript.

In many discussions K. Butterbach-Bahl, H. P. Schmid and T. Foken shared their vast experience on experiment design and analysis. They also supervised the manuscript composition, and contributed by editorial work and in discussing its contents. The general idea of using low-pass filtering in the wind calculation originates from T. Foken.

Appendix C: Supplement to Metzger et al. (2011)

This supplement was published together with the main article Metzger et al. (2011, Appendix B) and is available at http://www.atmos-meas-tech.net/4/1421/2011/amt-4-1421-2011.html.

The mathematical description of the wind measurement system and its uncertainty analysis was compiled by me alone. I also implemented the described algorithms in a software tool and drafted the entire text of the supplement.

At the recommendation of K. Butterbach-Bahl and T. Foken this supplement was introduced in order to extract predominantly algorithmic descriptions from the main article.

Appendix D: Metzger et al. (2012)

Metzger, S., Junkermann, W., Mauder, M., Beyrich, F., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Eddy-covariance flux measurements with a weight-shift microlight aircraft, Atmos. Meas. Tech., 5, 1699-1717, doi:10.5194/amt-5-1699-2012, 2012.

I initiated and designed the experiment and its evaluation strategy as presented in the article. It was me who conceptualized and performed the aircraft- and tower eddy-covariance data processing, incl. median-based de-spiking, complementary filtering (inspiration by J. Bange), spectral transforms and -references. The uncertainty budget on basis of the maximum-likelihood functional relationship is based on my ideas. Furthermore I validated and applied a footprint parameterization, implemented the described algorithms in a software tool, and wrote the entire text of the manuscript.

W. Junkermann installed and maintained all scientific equipment on the research aircraft. He also maintained and certified the aircraft itself, obtained flight permissions, performed the research flights, and edited the manuscript.

M. Mauder contributed to the manuscript through sharing his thoughts in many discussions and editorial work. He called attention to the amplitude resolution problem of the Metek USA–1 sonic anemometer, and suggested the final order in which the results are presented.

F. Beyrich (with the assistance of J. P. Leps and U. Weisensee) performed the ground-based reference measurements. In addition F. Beyrich determined the scintillometer fluxes as well as the CBL depth, and contributed his expertise in the regional land-atmosphere exchange to the manuscript editing process.

K. Butterbach-Bahl contributed to the interpretation of the presented results by suggesting the use of confidence intervals. He edited and substantially clarified the manuscript through encouraging the use of Appendices.

H. P. Schmid raised awareness to consider the uncertainties in both, dependent- and independent variables during regression analysis. This resulted in using the maximum-likelihood functional relationship in this study. Also, he shared his conceptual idea for a footprint parameterization considering 3D dispersion, which was formally communicated via K. Heidbach and M. Mauder.

By sharing his encouraging and critical thoughts T. Foken guided me through the process of composing the manuscript. He emphasized the use of spectral analysis and called attention to the influence of the Metek USA–1 anemometer geometry on the correlation between the horizontal- and vertical wind components. Furthermore he encouraged and contributed to the footprint validation by sharing the results of a reference Lagrangian dispersion model (calculated by T. Markkanen, and communicated through W. Babel and D. Charuchittipan).

Appendix E: Metzger et al. (2013)

Metzger, S., Junkermann, W., Mauder, M., Butterbach-Bahl, K., Trancón y Widemann, B., Neidl, F., Schäfer, K., Wieneke, S., Zheng, X. H., Schmid, H. P., and Foken, T.: Spatially explicit regionalization of airborne flux measurements using environmental response functions, Biogeosciences, 10, 2193-2217, doi:10.5194/bg-10-2193-2013, 2013.

I alone developed the experiment strategy and designed the flight patterns. I contributed to the implementation of the experiment on a daily basis through flight planning and coordination, data analyses and archiving, ground-based observations, and on-site negotiations with the local bureaucracy. It was my idea to; (i) draw on environmental response functions (ERF) for the purpose of flux regionalization, and to advance the concept through; (ii) considering 3D dispersion source area modeling, (iii) scaling the ERF by consideration of meteorological and land surface drivers, (iv) using current and continuous biophysical surface properties instead of a static and discrete land cover classification, (v) mining the data by use of non-parametric machine learning methods, and (vi) designing and rigorously quantifying an end-to-end uncertainty budget. I alone performed the entire data analysis, and compiled a software package that includes all algorithms used in this study. The idea to use a particular machine learning method, namely boosted regression trees, was inspired by H. Loescher and T. Stohlgren. J. Taylor shared his encouraging and critical thoughts which helped structuring the uncertainty budget. It was me who came up with a strategy to condense and discuss the results, and composed the vast majority of the manuscript.

W. Junkermann installed and maintained all scientific equipment on the research aircraft. He also maintained and certified the aircraft itself, performed the majority of research flights, and edited the manuscript. During the second part of the flight campaign W. Junkermann and A. Schmidtler shared the pilot duties among them.

During many discussions M. Mauder significantly contributed to the soundness of this study through (i) vetting the flight strategy, (ii) his idea to combine the ERF with wavelet analysis, (iii) his insistence to investigate the mixing regime, and (iv) editorial work.

It was K. Butterbach-Bahl's hypothesis, that airborne eddy-covariance flux measurements are a promising tool to gain new insights in the spatial variability of heat and moisture exchange across the Xilin River Catchment. On basis of this hypothesis he secured the funding that enabled us applying the research aircraft in Inner Mongolia, P.R. China. Also, it was him who raised awareness of the climatic gradients in the Xilin River Catchment, and participated during parts of the measurement campaign. He actively edited the manuscript, which streamlined its structure and aided the ecological interpretation of the results.

B. Trancon y Widemann contributed the wavelet algorithms used in this study in form of a software package, he drafted the chapter on wavelet analysis, and he reviewed the manuscript.

F. Neidl programmed and continuously maintained a new data acquisition system on the research aircraft.

K. Schäfer provided the ceilometer for deployment in the Xilin River Catchment, he derived the convective boundary layer depth from its data, and reviewed the manuscript.

S. Wieneke prepared the MODIS data for ingestion into the ERF algorithm, he performed the update of the land cover classification and drafted the corresponding paragraph in the manuscript.

X. H. Zheng was hosting this project by providing indispensable infrastructure in Beijing and in Inner Mongolia together with C. Liu, and by obtaining the required flight permissions together with S. Han.

H. P. Schmid participated in an early-stage site inspection during which we discussed scaling concepts and potential measurement strategies.

T. Foken encouraged the composition of this manuscript from an early stage, and guided me through its process. During several review cycles his critical and explicit comments helped to substantialize, focus and clarify the manuscript.

Appendix B: Metzger et al. (2011)

Atmos. Meas. Tech., 4, 1515–1539, 2011
www.atmos-meas-tech.net/4/1515/2011/
doi:10.5194/amt-4-1515-2011

Atmospheric Measurement Techniques

Corrigendum to

"Measuring the 3-D wind vector with a weight-shift microlight aircraft" published in Atmos. Meas. Tech., 4, 1421–1444, 2011

S. Metzger[1,2], **W. Junkermann**[1], **K. Butterbach-Bahl**[1], **H. P. Schmid**[1], **and T. Foken**[3]

[1]Karlsruhe Institute of Technology, Institute for Meteorology and Climate Research, Garmisch-Partenkirchen, Germany
[2]Chinese Academy of Sciences, Institute of Atmospheric Physics, Beijing, China
[3]University of Bayreuth, Department of Micrometeorology, Bayreuth, Germany

Received: 20 January 2011 – Published in Atmos. Meas. Tech. Discuss.: 28 February 2011
Revised: 29 June 2011 – Accepted: 7 July 2011 – Published: 26 July 2011

Publisher's Note: *Due to mistakes on publisher's side it was necessary to revise the original article Atmos. Meas. Tech., 4, 1421–1444, 2011 in this Corrigendum. Please use this Corrigendum as the main document.*

Abstract. This study investigates whether the 3-D wind vector can be measured reliably from a highly transportable and low-cost weight-shift microlight aircraft. We draw up a transferable procedure to accommodate flow distortion originating from the aircraft body and -wing. This procedure consists of the analysis of aircraft dynamics and seven successive calibration steps. For our aircraft the horizontal wind components receive their greatest single amendment (14 %, relative to the initial uncertainty) from the correction of flow distortion magnitude in the dynamic pressure computation. Conversely the vertical wind component is most of all improved (31 %) by subsequent steps considering the 3-D flow distortion distribution in the flow angle computations. Therein the influences of the aircraft's trim (53 %), as well as changes in the aircraft lift (16 %) are considered by using the measured lift coefficient as explanatory variable. Three independent lines of analysis are used to evaluate the quality of the wind measurement: (a) A wind tunnel study in combination with the propagation of sensor uncertainties defines the systems input uncertainty to $\approx$0.6 m s^{-1} at the extremes of a 95 % confidence interval. (b) During severe vertical flight manoeuvres the deviation range of the vertical wind component does not exceed 0.3 m s^{-1}. (c) The comparison with ground based wind measurements yields an overall operational uncertainty (root mean square error) of $\approx$0.4 m s^{-1} for the horizontal and $\approx$0.3 m s^{-1} for the vertical wind components. No conclusive dependence of the uncertainty on the wind magnitude ($<$8 m s^{-1}) or true airspeed (ranging from 23–30 m s^{-1}) is found. Hence our analysis provides the necessary basis to study the wind measurement precision and spectral quality, which is prerequisite for reliable Eddy-Covariance flux measurements.

Correspondence to: W. Junkermann
(wolfgang.junkermann@kit.edu)

1 Introduction

In environmental science, spatial representativeness of measurements is a general problem. The limited coverage of ground based measurements requires strategies to better understand spatial patterns (e.g., Baldocchi et al., 2001; Beyrich et al., 2006). Here airborne measurements are capable of supplementing and extrapolating ground based information (e.g., Lenschow, 1986; Desjardins et al., 1997; Mauder et al., 2008). However, to date manned platforms, such as fixed-wing aircraft (FWA, a summary of all notation is provided in Supplement C, see Sect. 3) and helicopters, are expensive to operate. Furthermore, their application is often not possible in settings such as remote areas beyond the range of an airfield. Here small size unmanned aerial vehicles are of use. These allow the measurement of a limited range of variables, such as temperature, humidity and wind vector (e.g., Egger et al., 2002; Hobbs et al., 2002; van den Kroonenberg et al., 2008). However due to payload constraints, they do not allow a comprehensive sensor package. A weight-shift microlight aircraft (WSMA) may provide a low-cost and easily transportable alternative, which also places a minimal demand on infrastructure in the

measurement location. After successfully applying a WSMA to aerosol and radiation transfer studies (e.g., Junkermann, 2001, 2005), the possibility of 3-D wind vector measurement from WSMA shall be explored. The underlying motivation is to work towards Eddy-Covariance (EC) flux measurements in the atmospheric boundary layer (ABL).

The determination of the 3-D wind vector from an airborne, i.e. moving platform, requires a high degree of sophistication. Specially designed probes enable the measurement of the 3-D turbulent wind field with respect to the aircraft (e.g. Brown et al., 1983; Crawford and Dobosy, 1992). At the same time the aircraft's movement with respect to the earth must be captured (e.g. Lenschow, 1986; Kalogiros and Wang, 2002a). A total of 15 measured quantities are involved in the computation of the 3-D wind vector (Supplement A), and consequently a similar number of potential uncertainty sources need to be considered. Furthermore, flow distortion by the aircraft itself can affect the measurement (e.g. Crawford et al., 1996; Kalogiros and Wang, 2002b; Garman et al., 2008). This complexity led to a number of quantitative uncertainty assessments of the wind measurement from aircraft, of which a few shall be mentioned here. While the carriers are commonly FWA, they cover a wide range, from single-engined light aircraft (e.g. Crawford and Dobosy, 1992) to twin-engined business jet (Tjernström and Friehe, 1991, e.g.) and quad-engined utility aircraft (e.g. Khelif et al., 1999). A similar variety of methodologies is used for the individual proof-of-concept. Widespread are uncertainty propagation of sensor uncertainties (e.g. Tjernström and Friehe, 1991; Crawford and Dobosy, 1992; Garman et al., 2006) and the analysis of specific flight manoeuvres (e.g. Tjernström and Friehe, 1991; Williams and Marcotte, 2000; Kalogiros and Wang, 2002a). Probably due to the higher infrastructural demand, wind tunnel studies (e.g. Garman et al., 2006), comparison to ground based measurements (e.g. Tjernström and Friehe, 1991) and aircraft intercomparisons (e.g. Khelif et al., 1999) are less common. Often statistical measures are used to express uncertainty, such as repeatability (e.g. $0.03\,\mathrm{m\,s^{-1}}$, Garman et al., 2006), deviation range (e.g. 0.4–$0.6\,\mathrm{m\,s^{-1}}$, Williams and Marcotte, 2000), median differences (e.g. $0.1 \pm 0.4\,\mathrm{m\,s^{-1}}$, Khelif et al., 1999), or root mean square error (e.g. $\geq 0.1\,\mathrm{m\,s^{-1}}$ at $\leq 2\,\mathrm{m\,s^{-1}}$ deviation range, Kalogiros and Wang, 2002a).

The EC technique (e.g., Kaimal and Finnigan, 1994) relies upon the precise measurement of atmospheric fluctuations, including the fluctuations of the vertical wind. Measured from aircraft, the determination of the wind vector requires a sequence of thermodynamic and trigonometric equations (Supplement A). These ultimately define the wind component's frame of reference. Yet, owing to its flexible wing- and aircraft architecture, the dynamics and flow distortion of the WSMA are likely more complex than those of FWA. Therefore the use of well established wind vector algorithms for FWA requires adaptation and correction. Consequently this study first and foremost investigates the feasibility and reliability of the wind measurement from WSMA. Based on these findings the measurement precision will be addressed in a successive study. The WSMA's overall measurement uncertainty was quantified by one standard deviation (σ) for sensor uncertainties provided by the manufacturers (combined effects of temperature dependence, gain error, nonlinearity), and one root mean square error (RMSE) for uncertainties from comparison experiments (including the uncertainty of the external reference, where applicable). Due to their analogous role in variance statistics, σ and RMSE are both referred to with one σ for convenience.

At the outset of this study the use of computational fluid dynamics was envisaged. The assessment was that, with great effort, such model could describe the mean aerodynamic properties of the WSMA. The challenge of the wind measurement from WSMA however comprises from the sources of variability. Consequently preference was given to a detailed experimental study.

After introducing the WSMA and outlining its physical properties, the sensor package for this study is presented. Following the analysis of the aircraft's dynamics, a toolbox is derived for the calibration of the 3-D wind vector measurement and the assessment of its uncertainty. It consists of a wind tunnel study, uncertainty propagation and in-flight manoeuvres. The toolbox is used to customize a wind vector algorithm for use with the WSMA. To evaluate this procedure, the final calibration is applied to measurements in the ABL. Wind measurements from the WSMA are compared to simultaneous ground based measurements from sonic detection and ranging (SODAR) and tall tower sonic- and cup anemometer and vane measurements. Based on three independent lines of analysis the overall uncertainty of the WSMA wind measurement is determined.

2 The weight-shift microlight aircraft

According to the safety and regulatory standards of the European Civil Aviation Conference, microlight aircraft are defined as aircraft with a maximum stall speed of $65\,\mathrm{km\,h^{-1}}$ and a take-off mass of no more than 450 kg. Figure 1 shows the weight-shift microlight research aircraft D-MIFU. It consists of two distinct parts, the wing and the trike (the unit hung below the wing, containing pilot, engine and the majority of the scientific equipment). The weight-shift control system is enabled by the pilot's direct application of pitching or rolling moments to the wing via the basebar. Counterbalance is provided by the mass of the trike unit suspended below the wing. Simple procedures for certification of installations on an open aircraft allow a wide spectrum of applications as well as flexible installation of scientific equipment. At an operational airspeed of $\approx 100\,\mathrm{km\,h^{-1}}$ D-MIFU can carry a maximum of 80 kg scientific payload from 15 m a.g.l. (above ground) to 4000 m a.s.l. (above sea level).

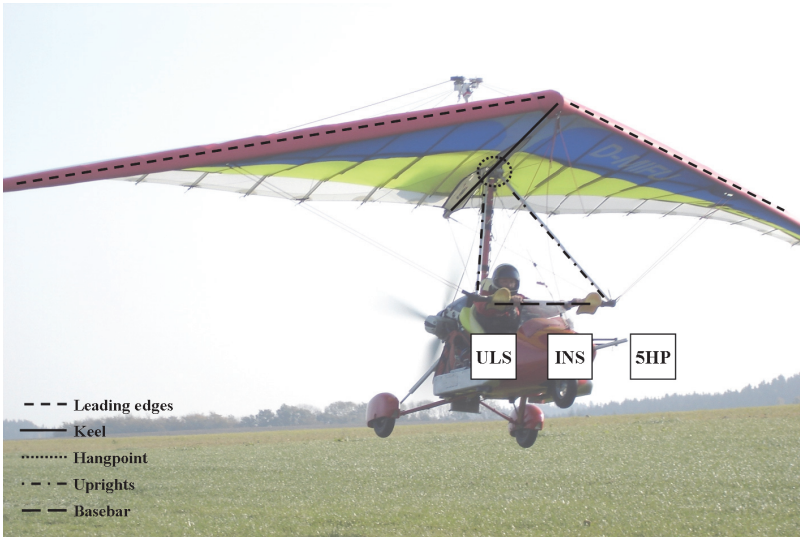


Fig. 1. Weight-shift microlight research aircraft D-MIFU, aircraft structural features are highlighted by dash-dotted lines. Sensor locations of the five hole probe (5HP), inertial navigation system (INS, inside aircraft nose) and universal laser sensor (ULS, below pilot seat) are indicated. For details on the respective installations see Sect. 2.2 and Table 2. Figure 3 details the layout of the five hole probe.

The full performance characteristics can be found in Junkermann (2001).

D-MIFU consists of a KISS 450 cambered wing by Air Creation, France, and the ENDURO-1150 trike manufactured by Ultraleichtflug Schmidtler, Germany. Owing to its aeroelasticity, the tailless delta wing is termed a flex-wing, contributing ≈15 % to the aircraft weight. The primary parts of the wing structure are the leading edges joined at the nose to the keel tube, which runs the root length of the wing (Fig. 1). Stretched over upper and lower surface is a high strength polyester sail. At a span of 9.8 m and keel length of 2.1 m, the wing provides a surface (S) of 15.1 m^2. It is put under considerable internal loads during rigging, it's form and rigidity being ensured by cross-tubes, rods and a wiring system. The basebar in front of the pilot seat is linked to the keel via two uprights and tensioned flying wires. It provides transmission of pitch and roll forces and is the primary flight control (Gratton, 2001). In the hangpoint on the wing keel the trike is attached to the wing. Since the trike is free to rotate in pitch and roll without hindrance, there is no pendular stability. In this regard the relationship of trike to wing is similar to the relationship of a trailing bomb to its carrier (e.g. HELIPOD, Bange and Roth, 1999). However trike and wing are fixed in their longitudinal axis, i.e. in the heading direction. The trike does not contribute significantly to the WSMA's lift, but represents a large portion of weight (≈85 %), drag, and provides all thrust through a 73 kW pusher engine-propeller combination. Flight stability in three axes is based on the offset of torques appearing at different locations on the wing (Cook, 1994). Torques result from wing aerodynamical effects, which sum nearest to neutral (slight nose-down torque for cambered wings) in one point along the wing's chord line, termed the wing's centre of pressure (Fig. 2). The centre of gravity, as far as the wing is concerned, is located in the hangpoint. The net aerodynamical torque is offset by a longitudinal lever arm between the centres of pressure and by a longitudinal -gravity, determining the aircraft's trim speed (the airspeed at which the aircraft will fly steadily without pilot input). Moreover increasing airspeed will result in an aeroelastical flattening of the wing, which is in contrast to FWA. This in turn can alter the balance of torsional loads and with it the circulation about the wing (Cook and Spottiswoode, 2006).

2.1 Physical properties

The need to adapt wind calibration procedures designed for fixed-wing aircraft is mainly caused by two structural features of the WSMA. The trike, i.e. the turbulence measurement platform, is mobile for pitching and rolling movements below the wing. Therefore the trike-based flow- and attitude angles must be measured with high resolution, precision and accuracy. Moreover, wing aerodynamics depends on its aeroelasticity with airspeed, and varying flow distortion in front of the wing must be considered. The effects of these WSMA features are not necessarily independent of each other, and may have a different impact on the wind measurement depending on the aircraft dynamics at a particular time. Therefore the WSMA was equipped with motion sensors. On the trike these were placed in the fuselage (Inertial

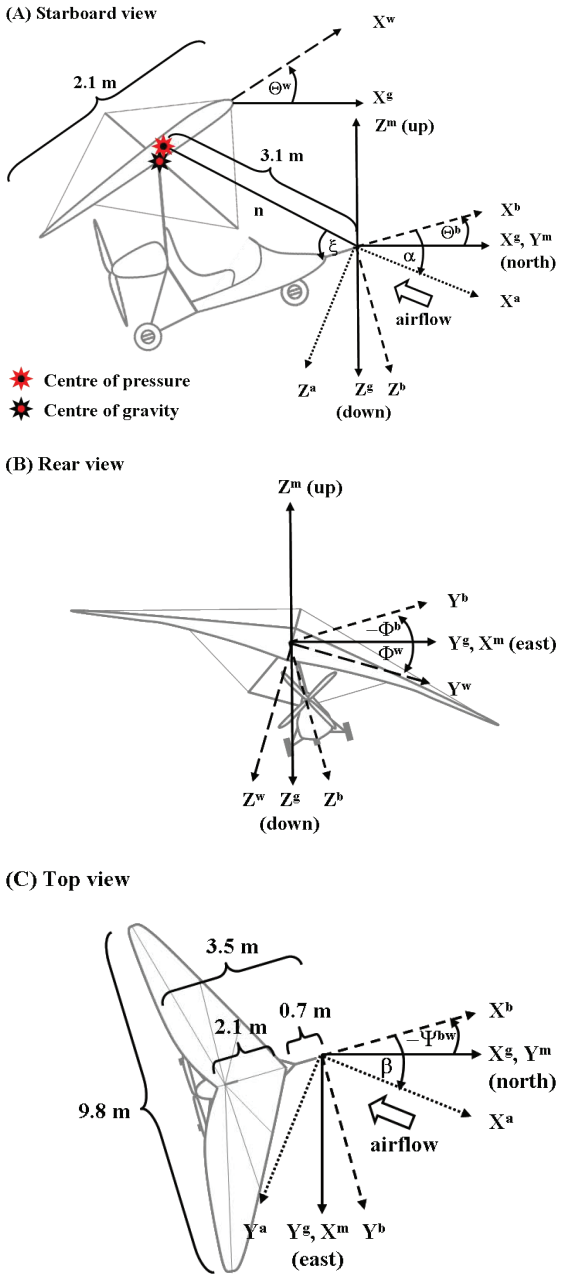


Fig. 2. Geometrical features of the weight-shift microlight aircraft and coordinate systems with axes X, Y, and Z used to compute the wind vector. The superscripts a, b, g, m and w represent, respectively, the aerodynamic-, body-, geodetic-, meteorological and wing coordinate systems (Supplement A). **(A)** Starboard view: Angle of attack (α), pitch angle (Θ), normalized radius (n), wing upwash direction (ξ), centre of gravity and centre of pressure. **(B)** Rear view: Roll angle (Φ); **(C)** Top view: Sideslip angle (β) and true heading (Ψ).

Navigation System, INS) and the wind measuring pressure probe (3-D acceleration), extending ≈0.7 m and ≈3.5 m forward from fuselage and aft-mounted propeller, respectively (Figs. 1 and 2). Further, the wing was equipped with motion sensors in the hangpoint (3-D acceleration) and atop the wing (3-D attitude). The INS is the most reliable motion sensor (Table 2), since it integrates the complementary characteristics of global positioning system (unbiased) and inertial measurement (precise). Position and velocity are calculated from inertial measurements of 3-D acceleration and 3-D angular rate, and matched with data from two global positioning units using a Kalman filter. The INS outputs 3-D vectors of position, attitude, velocity, angular rates and acceleration.

Airborne wind measurements are susceptible to distortion, since the aircraft itself is (a) a flow barrier and (b) must produce lift to remain airborne (Wyngaard, 1981; Cooper and Rogers, 1991). The aircraft's propeller, trike, and wing can be sources of flow distortion. Only little distortion from trike structural features is expected transverse to the pressure probe: the trike body is symmetric on its port and starboard side, and the pressure probe, propeller and pilot are centred on its longitudinal axis (Figs. 1 and 2). In contrast the body is asymmetric on its upside and underside, and the propeller location is 0.8 m higher than the pressure probe. This suggests symmetric flows in transverse, and asymmetric flows in longitudinal and vertical directions. All of which are expected to carry continuously through the pressure probe location, since the probe is rigidly fixed to the trike. This however is not the case for distortion from the WSMA wing. While the wind measurement encounters lift-induced upwash from the wing (Crawford et al., 1996; Garman et al., 2008), the trike, and with it the pressure probe, has rotational freedom in pitch and roll towards the WSMA wing. In the following we will outline the dependences of upwash generation from the wing. The amount of lift (L) generated by the wing equals the aircraft's sum of forces perpendicular to the airstream:

$$L = m\, a^{\mathrm{g},z}, \tag{1}$$

with the aircraft mass (m) and the vertical acceleration ($a^{\mathrm{g},z}$) in the geodetic coordinate system (GCS, superscript g, positive northward, eastward and downward) at the wing's centre of gravity (measured at, or dislocated to the hangpoint). For simplicity the acceleration perpendicular to the airstream was approximated by the vertical acceleration in the GCS. The maximum deviation during severe vertical manoeuvring (only used for evaluation) does not exceed ±1 %. Also the aircraft control forces applied by the pilot meet in the hangpoint. During ABL measurements these are primarily changes in power setting and wing pitch to adjust the aircraft altitude. For level, unaccelerated flight, lift essentially equals the aircraft's weight force, but is opposite in sign. The loading factor (LF) during vertically accelerated flight is then $\mathrm{LF} = \frac{L}{mg}$, the ratio of lift- to weight force with $g = 9.81\,\mathrm{m\,s^{-2}}$. Normalizing L for the airstream's dynamic pressure (p_q) and

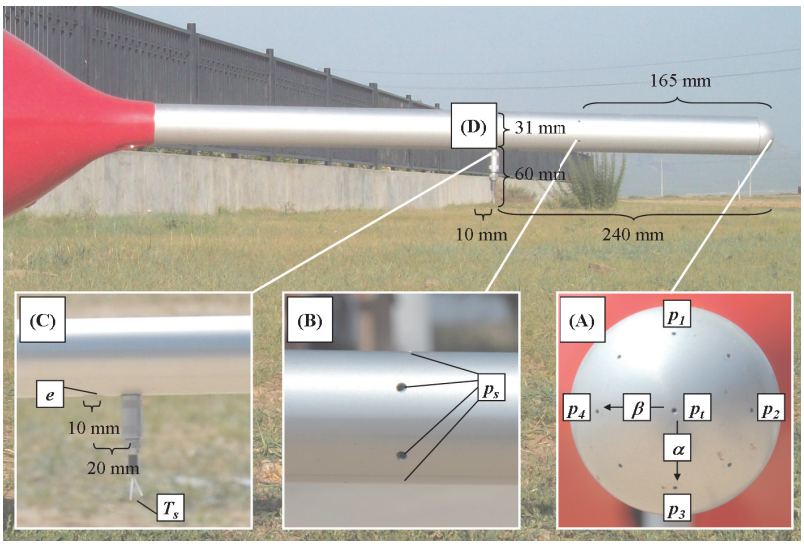


Fig. 3. Layout of the five hole probe, with letters indicating sensor locations. **(A)** The half sphere tip of the five hole probe, with ports for total- and differential pressure measurements. **(B)** Ports for static pressure measurement downstream of the half sphere. **(C)** Thermocouple and port for the capacitive humidity measurement. **(D)** Location of five hole probe 3-D acceleration sensor. Additional information is given in Sect. 2.2.

the wing's surface area (S) yields the unit-free lift coefficient (CL):

$$\mathrm{CL} = \frac{1}{p_q} \frac{L}{S} = \frac{2}{\rho v_{\mathrm{tas}}^2} \frac{L}{S}, \tag{2}$$

with wing loading ($\frac{L}{S}$). Moreover p_{q} in Eq. (2) can be substituted by air density (ρ) and true airspeed (v_{tas}). In CL the wing's ability to generate lift is determined to be approximately linear with wing pitch. As a consequence of lift generation air rises in front of the wing, which is defined as upwash. Crawford et al. (1996) provide the following parametrization to calculate the upwash velocity (v_{up}^w) for FWA:

$$v_{\mathrm{up}}^w = \frac{1}{\pi^2 n} v_{\mathrm{tas}} \mathrm{CL} = \frac{1}{\pi^2 n} \frac{v_{\mathrm{tas}}}{p_q} \frac{L}{S}, \text{ with } \frac{\delta \frac{v_{\mathrm{tas}}}{p_q}}{\delta v_{\mathrm{tas}}} \approx -0.3\ \mathrm{hPa}^{-1}. \tag{3}$$

Here v_{up}^w is defined as the tangent on a circle with normalized radius n. Thereby n is the separation distance from the wing's centre of pressure to the position of the pressure probe, normalized by the effective wing chord (Fig. 2). The wing upwash direction ξ is then enclosed by n and the trike body axis X_{b}. Since the wing is free to rotate in pitch and roll, v_{up}^w carries the orientation of the wing coordinate system (WCS, superscript w, positive forward, starboard, and downward). In Eq. (3) v_{up}^w varies inversely with n. Furthermore v_{up}^w can be expressed either directly proportional to v_{tas} and CL, or directly proportional to relative airspeed ($\frac{v_{\mathrm{tas}}}{p_{\mathrm{q}}}$) and $\frac{L}{S}$. Based on the functional relation between lift and upwash generation a treatment for the wind measurement from WSMA is derived in Sect. 4.1.

2.2 Instrumentation and data processing

Wind measurement by airborne systems is challenging. High resolution sensors are needed to determine the attitude, position, and velocity of the aircraft relative to the earth, as well as the airflow in front of the fuselage. The instrumentation involved in the wind measurement and data acquisition, including the respective manufacturers, is summarized in Table 1. A more detailed description of sensor characteristics and uncertainties is provided in Table 2, while respective locations are displayed in Figs. 1 and 3.

The principle is to resolve the meteorological wind vector from the vector difference of the aircraft's inertial velocity (recorded by the inertial navigation system) and the wind vector relative to the aircraft. To determine the latter, the aircraft was outfitted with a specially designed lightweight five hole half sphere pressure probe (5HP, e.g. Crawford and Dobosy, 1992; Leise and Masters, 1993). Figure 3a shows the half sphere tip of the 5HP, with a total pressure (p_{t}) port at its centre. Two additional pressure ports on each, the vertical (p_1, p_3) and the horizontal axis (p_2, p_4), surround the central port at an angle of $\tau = 45°$. These differential

Table 1. Overview of sensors and electronic instrumentation used for the wind measurement.

Component	Model	Manufacturer	Address
Butterworth low pass filter	AF40-4BU TP	E.S.F. electronic	Göttingen, Germany
Electronic compass module	TCM2-20	PNI Sensor Corporation	Santa Rosa, USA
Humidity sensor	SHT75	Sensirion AG	Staefa, Switzerland
Industrial computer	PR-Z32-EA-ST	Diamond Systems Corporation	Newark, USA
Inertial navigation system	RT3102	Oxford Technical Solutions	Upper Heyford, England
Differential pressure sensor	PCLA12X5D	Sensortechnics GmbH	Puchheim, Germany
Statistic pressure sensor	SP82AL	Capto As.	Horten, Norway
Thermocouple	CHAL-002	OMEGA Engineering, Inc.	Stamford, USA
Three-axis accelerometer	ADXL330	Analog Devices, Inc.	Norwood, USA
Universal laser sensor	ULS (Second edition)	Laser Technology, Inc.	Centennial, USA
Operating system	Minix 2.0	Andrew Stuart Tanenbaum	Amsterdam, Netherlands

pressure readings are used to determine attack angle (α) and sideslip angle (β), respectively, arrows indicate the direction of positive angular measurement. Polyetherketone tubings of $\leq$80 mm length and 1 mm inner diameter are used to connect these ports of 1.5 mm diameter to their respective pressure transducers. Additional (unnumbered) pressure ports at 45° increments are not used in this study. Six pressure ports are located downstream of the half sphere (Fig. 3b). These are ring-compensated around the circumference of the five hole probe for flow angle independent static pressure (p_s) measurement. Figure 3c shows the freely suspended 50 µm type K thermocouple for fast temperature (T_s) measurement and the 10 mm port for a capacitive humidity measurement (e). Time constants of thermocouple and humidity sensor are $<$0.02 s and $<$5 s at $v_{tas} = 27\,\mathrm{m\,s^{-1}}$, respectively. Humidity readings are solely used to provide the air density correction (Eq. A10) for the v_{tas} computation. At a typical true airspeed of $28\,\mathrm{m\,s^{-1}}$ only about 30 % and 15 % of the dynamic- and differential pressure transducers' range is exploited, respectively. This however enables the 5HP to be used also on faster aircraft such as motorized gliders, e.g. for inter-comparison measurements. Plug- and-socket connectors with locating pins insure a repeatable location of the 5HP with respect to the INS within $<$0.1°. The whole installation weights in at 350 g.

100 Hz temperature and pressure signals pass through hardware (analogue) four-pole Butterworth filters with 20 Hz cut-off frequency to filter high-frequency noise. Filter slope and frequency were chosen to allow miniaturization and comply with the system's 15 Hz bottleneck filter frequency of the infra-red gas analyser for EC flux calculation (not used in this study). The filter leads to a phase shift in the signal of $\approx$20 ms, and the amplitude of a 10 Hz sine signal is reduced by $<$1 %. The INS data are stored in a standalone system at a rate of $100\,\mathrm{s^{-1}}$. Remaining data streams for the wind computation are stored centrally at a rate of $10\,\mathrm{s^{-1}}$ by an in-house developed data acquisition system (embedded Institute for Meteorology and Climate Research data acquisition system, EIDAS). EIDAS is based on a ruggedized industrial computer and a real-time UNIX-like operating system. 5 V analogue signals at $\geq$10 Hz pass through a multiplexer and A/D converter at a resolution of 16 bits. For oversampled variables (100 Hz) the resulting signal is block averaged.

The INS has a latency time for internal calculations of $\approx$4 ms. Yet INS and EIDAS data streams have to be merged to calculate the ambient wind, and later turbulent fluxes. Therefore the resulting time lag between INS and 5HP of $\approx$16 ms has to be considered. The appropriate time shift of one to two 100 Hz increments is determined via lagged correlation. During post-processing the 100 Hz INS data set is then shifted by this increment before block averaging to 10 Hz. A spike test revealed $\approx$7 % missing values in the wing attitude data, which were filled via linear interpolation. To enable angular averaging or interpolation, heading angles were transformed from polar to Cartesian coordinates.

3 Wind vector

Approaches to compute the wind vector from fixed-wing aircraft are often similar in principle, though differ considerably in detail (e.g., Tjernström and Friehe, 1991; Williams and Marcotte, 2000; van den Kroonenberg et al., 2008). Therefore we provide a supplement to this study at http://www.bayceer.uni-bayreuth.de. Supplement A details the specific implementation that was found suitable for the wind measurement with our weight-shift microlight aircraft. A model to propagate uncertainty through the wind vector equations is provided in Supplement B. Relevant notation and abbreviations are listed in Supplement C.

The system's calibration was arranged bottom-up, i.e. from single instrument to collective application. The procedure starts with the laboratory calibration of the individual sensors, continues with the characterization of flow around the 5HP, and concludes with the treatment of WSMA specific effects on the wind measurement. Finally three

Table 2. List of measured variables, sensor characteristics, signal processing and data acquisition. Individual sensor locations are described in Sect. 2.2 and displayed in Figs. 1 and 3. Resolution refers to the smallest change registered by the data acquisition (DAQ) units. σ is the overall sensor uncertainty provided by the manufacturer in form of one standard deviation. Signal rates are displayed for sampling, filtering and storing (Signal SFS). Data acquisition takes place in two forms, standalone (SA) and on the central DAQ unit EIDAS. For non SA devices signal forwarding via A/D converter, recommended standard 232 (RS232) or serial peripheral interface (SPI) is indicated (Interface DAQ).

Quantity	Variable	Sensor	Range	Resolution	σ	Signal SFS [s^{-1}]			Interface DAQ	
Airframe motion										
Latitude/longitude		RT3102	$\pm 89.9°/\pm 180°$	6×10^{-15}°	1.1 m	100		100		SA
Altitude sea level		RT3102	$<18\,000$ m	0.001 m	2.7 m	100		100		SA
Altitude ground level		ULS	0.15–500 m	0.001 m	0.04 m	10		10	RS232	EIDAS
Heading, body	Ψ^b	RT3102	0–360°	0.00006°	0.1°	100		100		SA
Heading, wing	Ψ^w	TCM2-20	0–360°	0.1°	0.5°	16		10	RS232	EIDAS
Pitch/roll, body	Θ^b/Φ^b	RT3102	$\pm 90°/\pm 180°$	0.00006°	0.06°	100		100		SA
Pitch/roll, wing	Θ^w/Φ^w	TCM2-20	±20°	0.1°	0.2°	16		10	RS232	EIDAS
3-D velocity, body	v_{gs}^m	RT3102	0–515 m s^{-1}	0.0001 m s^{-1}	0.02 m s^{-1}	100		100		SA
3-D ang. rat., body	Ω^b	RT3102	$\pm 100°$ s^{-1}	0.0006° s^{-1}	0.01° s^{-1}	100		100		SA
3-D accel., body	a^b	RT3102	±10 g	0.00001 g	0.001 g	100		100		SA
3-D accel., wing		ADXL330	±3 g	0.0003 g	0.01 g	100		100	A/D	EIDAS
3-D accel., 5HP		ADXL330	±3 g	0.0003 g	0.01 g	100		100	A/D	EIDAS
Relative air motion										
Static pressure	$p_{s,A}$	SP82AL	0–1000 hPa	0.02 hPa	0.1 hPa	100	20	10	A/D	EIDAS
Dynamic pressure	$p_{q,A}$	PCLA12X5D	±12.5 hPa	0.0005 hPa	0.06 hPa	100	20	10	A/D	EIDAS
Attack pressure	p_α	PCLA12X5D	±12.5 hPa	0.0005 hPa	0.06 hPa	100	20	10	A/D	EIDAS
Sideslip pressure	p_β	PCLA12X5D	±12.5 hPa	0.0005 hPa	0.06 hPa	100	20	10	A/D	EIDAS
Fast temp.	T_s	CHAL-002	−20–60 °C	0.0015 K	0.5 K	100	20	10	A/D	EIDAS
Humidity, 5HP	e	SHT75	0–70 hPa	0.07 hPa	0.3 hPa	10		10	SPI	EIDAS

independent lines of analysis are used to quantify the overall system uncertainty: (a) uncertainty propagation through respective equations, (b) in-flight testing and (c) comparison of the measured wind vector with ground based measurements.

3.1 Wind tunnel study

Prior to in-flight use, the five hole probe was tested in an open wind tunnel at the Technical University of Munich, Germany, Institute for Fluid Mechanics. Objectives were to (a) confirm the applicability of transformation Eqs. (A5)–(A7) and (b) determine the 5HP's uncertainty in the operational range of the WSMA. The 5HP was mounted on D-MIFU's nosecap and measuring occurred at airflow velocities ranging from 20 to 32 m s^{-1} (equivalent to 2–6 hPa wind tunnel dynamic pressure). The dynamic pressure at the design stagnation point (i.e. the wind tunnel angles of attack $\tilde{\alpha} = 0°$ and sideslip $\tilde{\beta} = 0°$) was measured at airflow velocity increments of 1 m s^{-1}. At increments of 2 m s^{-1} a total of 570 permutations of 10 predefined angles $\tilde{\alpha}$ and $\tilde{\beta}$, each ranging from 0° to +20°, were measured. In addition one-dimensional symmetry tests were performed for six predefined angles $\tilde{\alpha}$ and $\tilde{\beta}$ ranging from −20° to +20° at an airflow velocity of 30 m s^{-1}. For the WSMA operational true airspeed of 28 m s^{-1} (or 4.5 hPa dynamic pressure during flight) the uncertainty of the wind tunnel airflow velocity was 0.7 % or $\sigma = 0.03$ hPa dynamic pressure. The airflow angles were varied by a calibration robot, the uncertainty in the wind tunnel angles was $\sigma_{\tilde{\alpha},\tilde{\beta}} < 0.1°$ (equal to the alignment repeatability between 5HP and INS). The wind tunnel angles $\tilde{\alpha}$, $\tilde{\beta}$ are related to the airflow angles α and β used for the wind calculation (Boiffier, 1998):

$$\begin{aligned} \alpha &= \tilde{\alpha}, \\ \beta &= \arctan\left(\frac{\tan\tilde{\beta}}{\cos\tilde{\alpha}}\right). \end{aligned} \tag{4}$$

The wind vector calculated from airborne measurements is very sensitive to uncertainties in its input variables. Calibration in laboratory and assessment in wind tunnel yield the basic sensor setup. However the effect of sensor and alignment uncertainties on the wind vector is not straightforward, and involves numerous trigonometric functions (Supplement A). To make the influence of individual measured quantities on the wind vector transparent, linear uncertainty propagation models were used (Supplement B). The intention is to investigate the wind measurement's uncertainty constraint by sensor setup and wind model description under controlled

Table 3. Flight campaign summary for locations Lake Starnberg (ST), Lindenberg (LI), and Xilinhot (XI). Anticyclonic and cyclonic conditions are indicated by a and c, respectively. For the flight patterns racetrack (RACE), wind square (SQUA), variance optimization (VARI), vertical wind specific flights (VW1–VW3) and the comparison to ground based measurements (COMP) the number of available datasets for each date is given together with respective track length (km) in parenthesis. Additional information is given in Sect. 3.2.

Date	19 Jun 2008	24 Jun 2008	25 Jun 2008	11 Jul 2008	15 Oct 2008	16 Oct 2008	18 Oct 2008	20 Oct 2008	21 Oct 2008	31 Jul 2009
Location	ST	ST	ST	ST	LI	LI	LI	LI	LI	XI
CYC 950 hPa	a	a	c	c	a	c	a	a	c	a
CYC 500 hPa	a	a	a	a	a	c	a	a	a	a
p [hPa]	1019	1021	1020	1015	1017	1008	1020	1018	1012	1010
T_{max} [°C]	22.4	21.6	27.7	27.8	14.8	14.3	13.1	16.5	21.7	31.1
Cloud cover	5/8	4/8	4/8	4/8	8/8	8/8	5/8	4/8	4/8	7/8
RACE	2 (10)	4 (10)	4 (10)				4 (10)			
SQUA		5 (12)	1 (12)							
VARI								6 (20)	4 (20)	2 (80)
VW1			1 (4)							
VW2		1 (11)	1 (11)							
VW3			1 (9)							
COMP					6 (12)	5 (12)	6 (12)			

boundary conditions. Because of flow distortion effects (Sect. 2.1) the boundary conditions during flight however are less well known and might be significantly different from the laboratory. Therefore a methodology for in-flight calibration and evaluation was derived. It consists of a WSMA specific calibration model and -flight patterns.

3.2 Flight campaigns

These patterns were carried out during three flight campaigns at different sites, each with its characteristic landscape and meteorological forcing:

3.2.1 Lake Starnberg, Germany

The first flight campaign took place from 19 June to 11 July 2008 over Lake Starnberg (47.9° N, 11.3° E). The lake is located in the foreland of the German Alps, that is a slightly rolling landscape (600–800 m a.s.l.) and mainly consists of grassland with patches of forest. The campaign focused on early morning soundings in the free atmosphere above Lake Starnberg.

3.2.2 Lindenberg, Germany

In a second campaign from 14–21 October 2008 comparison flights were carried out at the boundary layer measurement field of the German Meteorological Service, Richard-Aßmann-Observatory, near Lindenberg (52.2° N, 14.1° E). The area lies in the flat North German Plain (40–100 m a.s.l.), where land-use in the vicinity is dominated by an equal amount of agriculture and forests, interspersed by lakes. Flights in the atmospheric boundary layer were conducted under near-neutral stratification (stability parameter $|\frac{z}{L}| \leq 0.2$).

3.2.3 Xilinhot, China

To extend the operational range, an additional dataset under conditions approaching free convection ($\frac{z}{L} \ll -0.2$) was included in this study: From 23 June to 4 August 2009 an Eddy-Covariance flux campaign was performed over the steppe of the Mongolian Plateau. The hilly investigation area south of the provincial capital Xilinhot, Inner Mongolia, China (43.6° N, 116.7° E, 1000–1400 m a.s.l.) is covered by semi-arid grassland, intersected by a dune belt.

A summary of all flights as well as an overview of the synoptic weather conditions is provided in Table 3. Synoptic wind direction and cyclonality (CYC) were retrieved from the objective weather type data base of the German Meteorological Service (Bissolli and Dittmann, 2001). The XI flight on 31 July 2009 was supplemented with publicly available data from the US National Centre for Environmental Prediction. Prevailing wind direction throughout all flight days was south-west. Sea level pressure (p), 2 m a.g.l. maximum temperature (T_{max}) and cloud coverage are 24 h observations of the closest national meteorological service station on the respective day.

3.3 Flight patterns

In the following, the strategies of the individual flight patterns at these three sites are categorized in five classes and briefly outlined. The first four of them serve to isolate independent parameters for the flow distortion correction,

while the last one is used to compare aircraft to ground based measurements. The patterns are used for the actual calibration and evaluation of the wind measurement in Sect. 4.

3.3.1 Racetrack pattern

The first type of flight pattern consists of two legs parallel to the mean wind direction at constant altitude (one pair), one upstream leg (subscript +) and one downstream leg (subscript −). The legs are suitably aligned with the mean wind when having opposite tracks for identical aircraft settings. For any racetrack pair flown at constant true airspeed (v_{tas}), the (assumed homogeneous and stationary) mean wind ($\overline{\boldsymbol{v}^{\mathrm{m}}}$) cancels out (Leise and Masters, 1993; Williams and Marcotte, 2000):

$$\begin{aligned}\overline{|\boldsymbol{v}^{\mathrm{m}}_{\mathrm{gs}}|} &= \frac{1}{2}\left(\overline{|\boldsymbol{v}^{\mathrm{m}}_{\mathrm{gs},+}|} + \overline{|\boldsymbol{v}^{\mathrm{m}}_{\mathrm{gs},-}|}\right)\\ &= \frac{1}{2}\left(\left(\overline{v_{\mathrm{tas},+}} + \overline{|\boldsymbol{v}^{\mathrm{m}}|}\right) + \left(\overline{v_{\mathrm{tas},-}} - \overline{|\boldsymbol{v}^{\mathrm{m}}|}\right)\right)\\ &= \overline{v_{\mathrm{tas}}}. \end{aligned} \quad (5)$$

In this way the INS measured ground speed ($\overline{|\boldsymbol{v}^{\mathrm{m}}_{\mathrm{gs}}|}$) can be used to minimize the difference $|\overline{|\boldsymbol{v}^{\mathrm{m}}_{\mathrm{gs}}|} - \overline{v_{\mathrm{tas}}}|$ by iteratively adjusting dynamic pressure in Eq. (A8). This yields an inverse reference for dynamic pressure, which is solely based on INS data. Since the temperature and static pressure sensitivities of Eq. (A8) are two orders of magnitude lower than that of the dynamic pressure (Table 5), the inverse reference can now be used to adjust the 5HP measured dynamic pressure to in-flight conditions. A total of 14 racetrack pairs at airspeeds ranging from 21 to 32 m s^{-1} were conducted in the calm and steady atmosphere above the ABL (Table 3).

3.3.2 Wind square pattern

The second type of flight pattern consists of four legs flown at constant altitude and constant v_{tas} in the cardinal directions (north (N), east (E), south (S), west (W)). Assuming that the flights were carried out in a homogeneous and stationary wind field, the measured horizontal wind components (v^{m}_u, v^{m}_v) should be independent of aircraft heading, i.e. constant at each side of the wind square. With it a potential offset in β can be determined: The offset in β is changed iteratively, until the standard deviation of v^{m}_u and v^{m}_v throughout a wind square is minimized. For flights above the ABL, in addition the vertical wind component can be expected to be negligible. A potential offset in α can be determined in a similar fashion to β, however, under the constraint of minimizing the absolute value of the vertical wind component (v^{m}_w). The wind square pattern further allows to estimate the uncertainties of v_{tas} and β: Since the flight legs are aligned in the cardinal directions, along-track wind components (v^{m}_u (N, S), v^{m}_v (E, W)) are predominantly sensitive to errors in v_{tas}. Cross-track wind components (v^{m}_v (N, S), v^{m}_u (E, W)) are predominantly sensitive to errors in β. Thus, errors in v_{tas} and β can be estimated as:

$$\begin{aligned}\sigma_{uv,\mathrm{tas}} &= \sqrt{\frac{1}{2}\left(\left(\overline{v^{\mathrm{m}}_u}(\mathrm{N}) - \overline{v^{\mathrm{m}}_u}(\mathrm{S})\right)^2 + \left(\overline{v^{\mathrm{m}}_v}(\mathrm{E}) - \overline{v^{\mathrm{m}}_v}(\mathrm{W})\right)^2\right)}\\ \sigma_{uv,\beta} &= \sqrt{\frac{1}{2}\left(\left(\overline{v^{\mathrm{m}}_v}(\mathrm{N}) - \overline{v^{\mathrm{m}}_v}(\mathrm{S})\right)^2 + \left(\overline{v^{\mathrm{m}}_u}(\mathrm{E}) - \overline{v^{\mathrm{m}}_u}(\mathrm{W})\right)^2\right)}. \end{aligned} \quad (6)$$

Six wind squares were flown above the ABL at airspeeds from 23 to 29 m s^{-1} (Table 3).

3.3.3 Variance optimization pattern

The third type of flight pattern is a straight and level ABL sounding, intended for EC flux measurement. The assumption made here is that errors in the flow angles increase the wind variance. In contrast to the previous two patterns, this method does not imply homogeneity or stationarity. It can therefore be applied even in the presence of thermal turbulence, i.e. in the convective ABL (Tjernström and Friehe, 1991; Khelif et al., 1999; Kalogiros and Wang, 2002a). Offsets and slopes for α and β were computed to minimize (a) the sum of the wind components variances plus (b) the absolute value of the mean vertical wind. Here it is expected that $\overline{v^{\mathrm{m}}_w}$ approaches zero for a sufficiently high number of datasets above approximately level terrain. Twelve straight and level ABL soundings (or 360 km of flight data, Table 3) at airspeeds from 24 to 28 m s^{-1} between 50 and 160 m above ground were used for this variance optimization.

3.3.4 Vertical wind specific patterns

The fourth type of flight pattern specifically addresses errors in v^{m}_w, the wind component crucial for EC flux applications. Based on Lenschow (1986) straight-flight calibration patterns were performed above the ABL. These are intended to assess and minimize the possible influence of aircraft (in our case WSMA) lift and trim on v^{m}_w. At airspeeds ranging from 21 to 32 m s^{-1} five vertical wind (VW) specific flights, divided into three sub-patterns, were utilized in this study (Table 3):

- VW1 – (Level acceleration – deceleration): whilst the engine's power setting was gradually varied, the wing pitch (and with it lift coefficient) was adjusted to maintain flight altitude. With this pattern the influence of aircraft trim on v^{m}_w can be determined.
- VW2 – (Smooth oscillation): starting from level flight the power setting was slowly varied, while the wing pitch was adjusted to maintain constant v_{tas}. In consequence, the aircraft ascended and descended about the mean height, while CL remained approximately unchanged. VW2 was used to assess the influence of wing pitch and aircraft vertical velocity on v^{m}_w.

– VW3 – (Forced oscillation): starting from level flight the wing pitch was forcibly alternated. The aircraft ascended and descended around the mean height, while power setting remained unchanged. In response aircraft accelerations and velocities, and with it the airflow around the aircraft, changed. VW3 was used to assess the integral influence of vertically accelerated flight on v_w^{m}, as e.g. during terrain following flights in the ABL (see Sect. 4.1, Step G6).

3.3.5 Comparison to ground based reference measurements

The fifth and last type of flight pattern is a series of comparison measurements between WSMA and ground based measurements. These were carried out at the boundary layer measurement field of the German Meteorological Service, Richard-Aßmann-Observatory, near Lindenberg. The lower part of the ABL was probed by a 99-m tower and a SODAR with their base at 73 m a.s.l. The 99-m tower provided cup measurements (10 min averages) of wind speed at four levels (40, 60, 80, and 98 m a.g.l.), the wind direction was measured with vanes at heights of 40 and 98 m a.g.l. (10 min averages). Sonic anemometers mounted at the tower provided turbulent wind vector measurements at 50 and 90 m a.g.l. The SODAR wind vector profiles (15 min averages) reached, at increments of 20 m, from 40 to 240 m a.g.l. In addition a reference for static pressure was provided at 1 m a.g.l. 17 cross-shaped patterns (van den Kroonenberg et al., 2008), with flight legs of 3 km centred between tower and SODAR, were performed at 24 and 27 m s^{-1} airspeed (Table 3). The flights were carried out at the approximate sounding levels of tower and SODAR (50, 100, 150, 200 and 250 m a.g.l.). This allows a direct comparison of WSMA and ground based measured wind components. Aircraft and sonic wind measurements were filtered using the stationarity test for wind measurements by Foken and Wichura (1996). SODAR, cup and vane data were stratified for the best quality rating assigned by the German Meteorological Service. Simultaneous wind data of WSMA and ground based measurements were accepted for comparison only if they agreed to within ± 20 m height above ground (which equals $\approx 2\,\sigma$ of variations in WSMA altitude). This data screening resulted in a total of 20 data couples (between WSMA and cups/vanes, sonics and SODAR) for v_{uv}^{m}, and 19 data couples for v_w^{m}. Compared to cups/vanes, sonics and SODAR, the WSMA soundings were on average higher above ground by 0.1 ± 5.5, 8.7 ± 5.6, and 0.5 ± 5.3 m, respectively.

4 Application to weight-shift microlight aircraft

To understand operational requirements for setup and calibration of the wind vector measurement, aircraft attitude and dynamics were assessed for a straight and level boundary layer flight (Table 3, variance optimization flight on 31 July 2009). Histograms of aircraft properties were calculated from $\approx 3 \times 10^4$ data points sampled ≈ 50 m a.g.l. (Fig. 4). Variations in true airspeed and aircraft vertical movement were resulting from aircraft manoeuvres to follow the terrain contours as well as thermal turbulence (labile stratification, stability parameter $\frac{z}{L} \approx -0.9$). Attitude angles (Θ^{b}, Φ^{b}) indicate constant upward pitching and anticlockwise roll of the trike, respectively. Pitching as well as rolling increase in magnitude with v_{tas}, i.e. power setting of the engine. The pitching moment can be understood as the dynamic balance with v_{tas} between propeller thrust and the drag difference between the trike (low) and the wing (high). This is confirmed by an estimate of the attack angle (α), which shows fewer variation due to alignment with the streamlines, though alike Θ^{b} increases with v_{tas} ($\approx 0.4°$ per m s^{-1}). The rolling moment can be understood as counterbalance of the clockwise rotating propeller torque. In addition side-slipping of the trike over its port side was detected from an estimate of the sideslip angle (β), increasing at a rate of $\approx -0.6°$ per m s^{-1} with v_{tas}. The operational range in α and β estimates were found $\approx |15°|$, averaging to $6.0 \pm 1.8°$ and $-5.5 \pm 3.2°$, respectively (Fig. 4). Following the lift Eq. (2), wing pitch decreases with v_{tas}. That is, with increasing v_{tas} the noses of wing and trike approach each other. Wing roll does not display dependence on v_{tas}, i.e. no counter reaction on propeller torque or trike roll. The wing loading factor (LF) was found to vary within a range of $\sigma \approx 0.1$ g (Fig. 4), from which the upwash variation in front of the wing can be assessed.

Using five hole probe measured v_{tas} in Eq. (3) the upwash velocity (v_{up}^{w}) at 5HP location was determined to 1.52 ± 0.19 m s^{-1}. D-MIFU is travelling at low airspeed and has a small relative separation (n) between wing and 5HP. Both factors lead to an increase in v_{up}^{w}. Various research aircraft have been assessed with regard to upwash generation (Crawford et al., 1996), compared to which D-MIFU ranges mid-table. This can be ascribed to the low wing loading, which is a fraction of those of fixed-wing aircraft, and decreases v_{up}^{w}. Wing loading, and with it v_{up}^{w}, are directly proportional to vertical acceleration and aircraft mass in Eq. (1). Hence $\sigma \approx 10\,\%$ variation in LF (Fig. 4) accounts for most of the variance in v_{up}^{w}. In addition aircraft mass can vary during the flight due to fuel consumption ($\pm 4\,\%$) and among measurements due to weight differences of pilots ($\pm 2\,\%$). Due to the trike's rotational freedom, upwash about the wing's centre of pressure can partially translate into along- and sidewash (longitudinal and transverse to the trike body, respectively) at the 5HP location in the trike body coordinate system (BCS). Mean aerodynamic chord theory yields the centre of pressure's position of the wing within 0.2 m or $<10\,\%$ chord length of the centre of gravity. Assuming the centres of pressure and gravity to coincide, the pitch difference between wing and trike can be neglected, and v_{up}^{w} is easily transformed

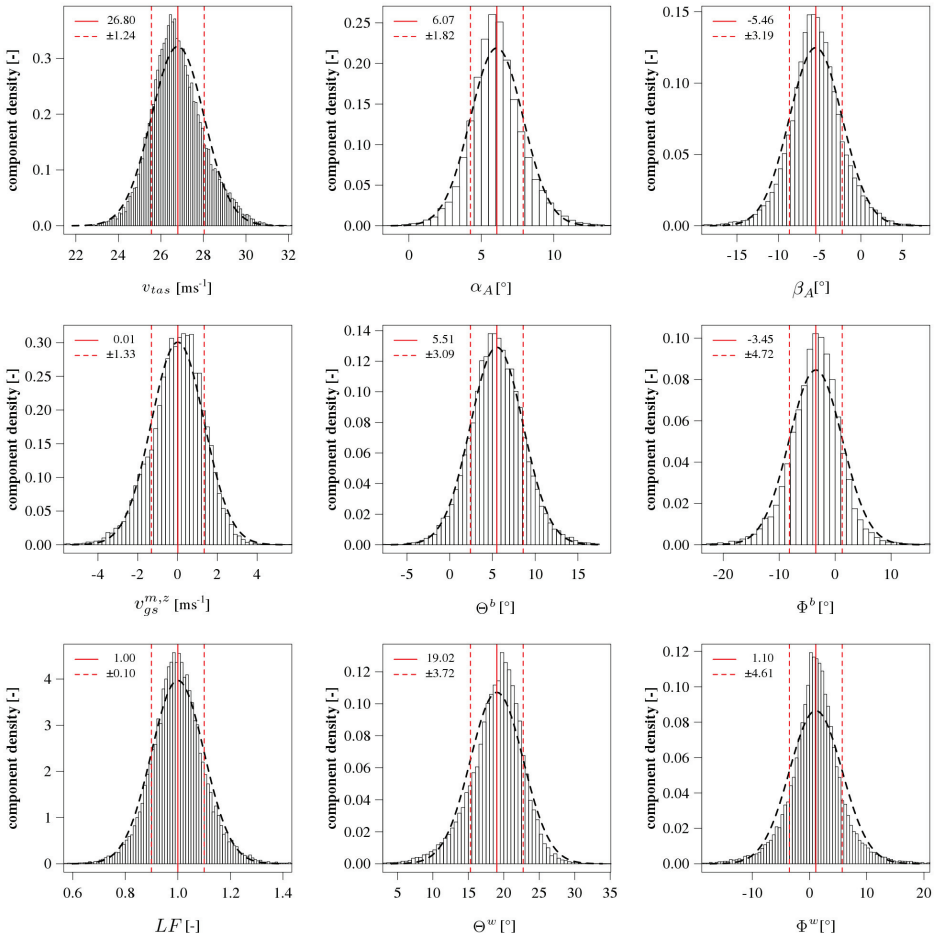


Fig. 4. Histograms of aircraft properties derived for the flight on 31 July 2009 (Table 3). Component density is scaled so that the histograms have a total area of one. Red vertical lines indicate distribution average (solid) and standard deviation (dashed). The black dashed bell curve displays a reference normal distribution: True airspeed (v_{tas}), attack angle (α_{A}), sideslip angle (β_{A}), aircraft vertical velocity ($v_{\mathrm{gs}}^{\mathrm{m},z}$), trike pitch- ($\Theta^{\mathrm{b}}$) and roll ($\Phi^{\mathrm{b}}$) angles, loading factor (LF, the ratio of lift- to weight force), as well as wing pitch- (Θ^{w}) and roll (Φ^{w}) angles.

into the BCS: the transformation Eq. (A13) was carried out about zero heading difference, the wing upwash direction ($\xi = -41.9 \pm 0.3°$), and the roll difference between wing and trike. Wing upwash net effect at the 5HP location was then directed forward, right and upward with $1.01 \pm 0.13\,\mathrm{m\,s^{-1}}$, $0.12 \pm 0.13\,\mathrm{m\,s^{-1}}$, and $-1.12 \pm 0.14\,\mathrm{m\,s^{-1}}$ in trike body coordinates (Fig. 5).

4.1 Wind measurement calibration

The sensitivity of the wind model description was analysed by linear uncertainty propagation models (Supplement B). The first model in Eq. (B1) permits to express the sensitivity of the wind computation as a function of attitude angles, flow angles and true airspeed. It was carried out for two reference flight states at $v_{\mathrm{tas}} = 27\,\mathrm{m\,s^{-1}}$. In State 1 attitude and flow angles were assumed small (1°), as it would be typical for calm atmospheric conditions. This allows for the small-angle approximation in Eq. (B1), resulting in uncertainties for the wind components ($\Delta v_{uvw}^{\mathrm{m}}$) as a function of the heading angle (Ψ). In State 2 attitude (10°) and flow angles (−15°) were approximately increased to their 95 % confidence intervals during soundings in the convective ABL (Fig. 4). Consequently the full form of Eq. (B1) must be used for State 2. It allows to calculate the maximum uncertainty in the wind components ($|\Delta v_{uvw}^{\mathrm{m}}|$) over all Ψ, as well as to compare these between both flight states. Both states were inferred uncertainties of 1° and $0.5\,\mathrm{m\,s^{-1}}$ for

Table 4. Input uncertainty (IU) from the linear uncertainty propagation model Eq. (B1). For the sensitivity analysis the model was forced with two different reference states, State 1 with small and State 2 with enhanced flow (α, β) and attitude (Θ^b, Φ^b, Ψ^b,) angles. Both states were inferred similar uncertainty quantities Δf_i in α, β, Θ^b, Φ^b, Ψ^b, and true airspeed (v_{tas}). After calibration Step B the reference State 2 was used for the uncertainty propagation: the actual uncertainties in (a) the flow computation (α, β and v_{tas}, Table 5), and (b) the sensor alignment (Θ^b, Φ^b, Ψ^b) were inferred. Additional information is given in Sect. 4.1.

Variable	α	β	Θ^b	Φ^b	Ψ^b	v_{tas}	IU
Model forcing							
State 1	1°	1°	1°	1°	0...360°	27 m s^{-1}	
State 2	−15°	−15°	10°	10°	0...360°	27 m s^{-1}	
$\Delta f_{i,\text{sensitivity}}$	1°	1°	1°	1°	1°	0.5 m s^{-1}	
$\Delta f_{i,\text{propagation}}$	0.76°	0.76°	0.1°	0.1°	0.1°	0.34 m s^{-1}	
Results State 1 – sensitivity							
Δv_u^m [m s^{-1}]	<0.01	0.47cosΨ	<0.01	<0.01	0.47cosΨ	0.50	1.08
Δv_v^m [m s^{-1}]	<0.01	−0.47sinΨ	<0.01	<0.01	−0.47sinΨ	0.50	1.08
Δv_w^m [m s^{-1}]	0.47	<0.01	−0.47	<0.01	<0.01	<0.01	0.95
$\lvert\Delta v_{uv}^m\rvert$ [m s^{-1}]	0.01	0.47	<0.01	0.01	0.47	0.5	1.08
$\lvert\Delta v_w^m\rvert$ [m s^{-1}]	0.47	0.01	0.48	0.01	0.00	<0.01	0.97
Results State 2 – sensitivity							
$\lvert\Delta v_{uv}^m\rvert$ [m s^{-1}]	0.21	0.47	0.21	0.14	0.42	0.45	1.34
$\lvert\Delta v_w^m\rvert$ [m s^{-1}]	0.41	0.05	0.32	0.14	0.00	0.22	1.14
Results State 2 – propagation							
$\lvert\Delta v_v^m\rvert$ [m s^{-1}]	0.16	0.36	0.02	0.01	0.04	0.30	0.64
$\lvert\Delta v_w^m\rvert$ [m s^{-1}]	0.31	0.04	0.03	0.01	0.00	0.15	0.55

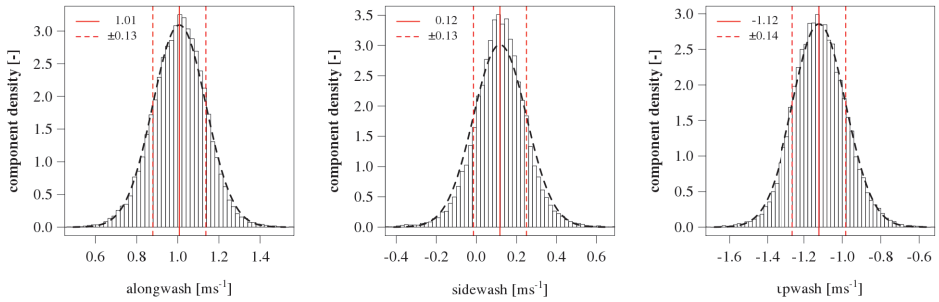


Fig. 5. Histograms of wing-generated alongwash, sidewash and upwash at the five hole probe location. Results are calculated from wing properties in Eqs. (1)–(3) and then rotated from wing- into trike body coordinates (Fig. 2) using Eq. (A13). Presented is the same dataset and in the same manner as in Fig. 4.

angular- and v_{tas} measurements, respectively. From State 1 it can be seen that the major uncertainty in the horizontal wind components (v_{uv}^m) originates from v_{tas}, sideslip angle (β) and heading angle (Ψ), where β and Ψ carry similar sign and sensitivity (Table 4). On the contrary, the vertical wind component (v_w^m) is similarly sensitive to attack angle (α) and pitch angle (Θ), yet with reversed sign. As compared to State 1, in State 2 the absolute uncertainties in the horizontal ($|\Delta v_{uv}^m|$) and vertical ($|\Delta v_w^m|$) wind components are increased by 24 % and 18 %, respectively. The increase however does not originate from the most sensitive terms, but from formerly negligible terms such as trike roll (Φ^b).

Table 5. Uncertainty of variables entering the wind vector computation Eq. (A1): Static pressure (p_s), dynamic pressure as used in the computation of flow angles ($p_{q,A}$) and the true airspeed ($p_{q,B}$), differential pressures (p_α, p_β), static temperature (T_s) and water vapour pressure (e). Sources of uncertainty (σ) are subscripted as follows: manufacturer provided sensor uncertainty (SEN), calibration in laboratory (LAB), wind tunnel (TUN), and wind model description (MOD). The 0.05 K and 0.36 K uncertainties for radiation and ram rise errors in T_s were accounted in σ_{mod}. Additional information is given in Sect. 4.1.

Variable		Input					Sensitivities			Output		
	Unit x	σ_{SEN} $[x]$	σ_{LAB} $[x]$	σ_{TUN} $[x]$	σ_{MOD} $[x]$	$\sigma_{GAU,i}$ $[x]$	S_α $[°/x]$	S_β $[°/x]$	$S_{v_{tas}}$ $[m\,s^{-2}/x]$	σ_α [°]	σ_β [°]	$\sigma_{v_{tas}}$ $[m\,s^{-1}]$
Propagated												
p_s	[hPa]	0.10	0.58			0.59			−0.01			0.01
$p_{q,A}$	[hPa]	0.06	0.03			0.07	4.26	3.60		0.29	0.24	
$p_{q,B}$	[hPa]	0.06	0.03	0.04		0.08			3.85			0.30
p_α	[hPa]	0.06	0.03			0.07	4.57	0.97		0.31	0.07	
p_β	[hPa]	0.06	0.03			0.07	0.48	4.74		0.03	0.32	
T_s	[K]	0.26	0.40		0.36	0.60			0.04			0.02
e	[hPa]	0.30				0.30			0.01			0.01
$\sum$										0.62	0.62	0.34
Non-propagated												
α	[°]			0.44		0.44	1.00			0.44		–
β	[°]			0.43		0.43		1.00			0.43	–
$\sigma_{gau,o}$										0.76	0.76	0.34

The latter now account for up to 50 % of $|\Delta v_{uv}^m|$ and 37 % of $|\Delta v_w^m|$. In Table 5 similar sensitivity analyses were carried out for α in Eq. (A5), β in Eq. (A6) and the thermodynamic derivation of v_{tas} in Eq. (A8). The input uncertainties were Gaussian summarized ($\sigma_{gau,i}$) and propagated into output uncertainties (σ) of attack angle (α), sideslip angle (β) and true airspeed (v_{tas}), using the sensitivities (S) in their respective computations Eqs. (A5), (A6), and (A8). Propagated output uncertainties were summed up in analogy to Eq. (B1) before Gaussian summarizing them with the non-propagated uncertainties for α and β wind tunnel measurements to the final output uncertainties ($\sigma_{gau,o}$). Also here $v_{tas} = 27\,m\,s^{-1}$ was assumed as reference state, parametrized as 3.7 hPa dynamic pressure (p_q), 21 °C static temperature, 850 hPa static pressure, and 9.5 hPa water vapour pressure. Derived sensitivities indicate a dominant dependence of α and β on their respective differential pressure measurement, as well as on p_q. In case of v_{tas} sensitivity on the p_q measurement clearly prevails. This procedure allows to separate, and consequently further concentrate on, the variables most sensitive to the wind vector calculation. For v_w^m, the central wind component in the Eddy-Covariance flux technique, the variables to focus calibration effort on are α, Θ and p_q. Likewise correct readings of β, p_q and Ψ are of greatest importance for the calculation of v_{uv}^m.

Due to the same adiabatic heating effect (ram rise) as in Eq. (A9), the temperature measured by the thermocouple might be slightly higher than the static temperature intrinsic to the air. At the same time the measured temperature is smaller than the total temperature at the stagnation point on the tip of the 5HP, since the air at the thermocouple is not brought to rest. Even at peak $v_{tas} = 30\,m\,s^{-1}$ of the WSMA the ram rise of 0.4 K does not surpass the overall uncertainty of the thermocouple (Table 2). As a practical advantage of the slow flying WSMA therefore no fractional "recovery factor" correction as known from faster fixed-wing aircraft needs to be introduced (Trenkle and Reinhardt, 1973). Using above sensitivity analysis the associated uncertainty amounts to $0.02\,m\,s^{-1}$ in v_{tas}. According to the parametrizations (5) and (7) in Foken (1979) the error caused by solar radiation intermittently incident at the unshielded thermocouple was estimated to be <0.05 K. Since no radiation shielding was applied, both temperature errors were included in the uncertainty propagation (Table 5).

The actual calibration sequence was organized in seven steps (Fig. 6), resulting in an incrementally refined system. To reduce scatter and facilitate convergence of the iterative processes the 10 Hz aircraft data were block averaged to 1 Hz for Steps D–G:

Step A – Laboratory: Initial calibration of all A/D devices.
Step B – Wind tunnel: Assessment of attack- (α) and sideslip angle (β) and first correction of dynamic pressure (p_q).

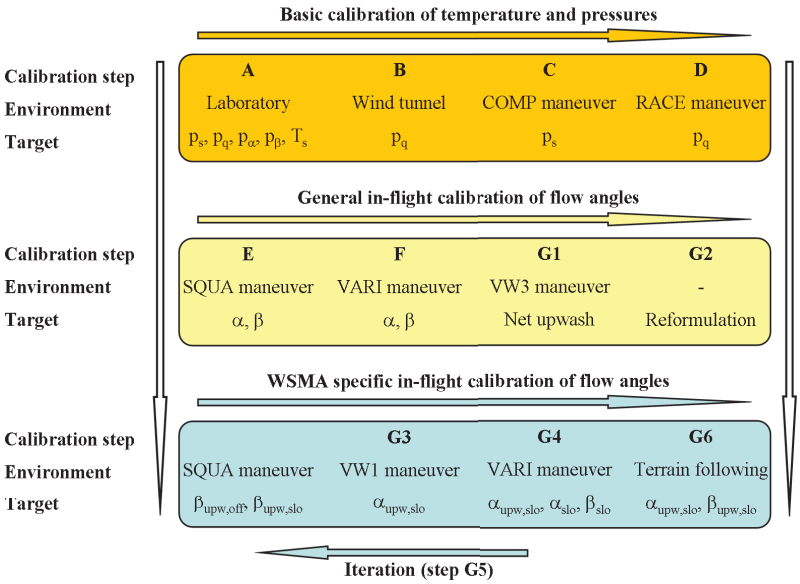


Fig. 6. Flow chart of the calibration process. The calibration Steps A–G are carried out in a sequence from left to right, top to bottom. Each step results in an incrementally refined system. The iterative Step G5 (blue background) comprises the three flight manoeuvres SQUA, VW1 and VARI. Within G5 the SQUA manoeuvre is not associated with an individual calibration step.

Step C – Tower fly-bys: Adjustment of static pressure (p_s).
Step D – Racetracks: Second p_q correction.
Step E – Wind squares: First estimate of α and β correction.
Step F – Variance optimization: Second estimate of α and β correction.
Step G – Vertical wind treatment: Relation of measured net flow distortion to lift coefficient, iterative optimization with Steps E–F.

4.1.1 Step A – Laboratory

Calibration coefficients from laboratory and all successive steps are summarized in Table 6. Residuals are propagated together with sensor uncertainties provided by the manufacturers. The resulting uncertainties are summarized in Table 5.

4.1.2 Step B – Wind tunnel

Since the wind tunnel was too small for the complete aircraft, the setup was reduced to the five hole probe and the aircraft's nose-cap. Therefore the actual flow distortion during flight was not included in this step. For angles of attack (α) and sideslip (β) within $\pm 17.5°$ the first-order approximations Eqs. (A5)–(A6) were most effective for deriving flow angles from our miniaturized 5HP. Root mean square error (RMSE) and bias (BIAS) amounted to 0.441°, 0.144° and 0.428°, 0.047° for α and β, respectively, with a Pearson Coefficient of determination $R^2 > 0.99$. Residuals did not scale with true airspeed, but resulted from incomplete removal of α and β cross dependence (Fig. 7). The probe design was working less reliably with the exact solutions for flow angle determination (e.g. Eq. 7 in Crawford and Dobosy, 1992). We speculate that this behaviour arises from the amplified pressure drops in the attack and sideslip differential pressures ($p_{\alpha,\beta}$) at elevated angles. In contrast to their 1.5 mm pressure ports, the dynamic pressure ($p_{q,A}$, subscript upper-case letters A–G indicating calibration stage) is measured against a direction-independent static pressure port (Fig. 3). While allowing for slightly more scatter due to angular cross-dependency, the approximate Eqs. (A5) and (A6) compensate the different pressure drops in the quotient $\frac{p_{\alpha,\beta}}{p_{q,A}}$. On the other hand use of a calibration polynomial as suggested by Bohn and Simon (1975) has the advantage that it does not assume rotational symmetry. A fit of the calibration polynomial yielded high precision, however did not prove robust for in-flight use and was discarded. For dynamic pressure ($p_{q,A}$), offset (0.22 hPa) and slope (1.05) were corrected from zero working angle ($\arccos(\cos\alpha\cos\beta)$) measurements. Applying the pressure drop correction Eq. (A7) thereafter reduced the scatter significantly, in particular for elevated working angles (Fig. 7). Below 20° working angle (≈15° flow angle) $p_{q,B}$ was slightly overestimated, above this a loss of only ≈ -0.1 hPa remained. RMSE and BIAS amounted to 0.042 and 0.012 hPa, respectively, with $R^2 = 0.999$. The wind tunnel study proofs the applicability of the spherical model Eqs. (A5)–(A7) to determine flow angles and dynamic pressure from our 5HP: the deviations are well within the

Table 6. Coefficients for static pressure (p_s), dynamic pressure (p_q), differential pressures (p_α, p_β), static temperature (T_s), and flow angle measurements (α, β) during calibration Steps A–G. Respective environments are laboratory (LAB), wind tunnel (TUN), comparison to ground based measurements (COMP), racetrack (RACE), wind square (SQUA), variance optimization (VARI) and vertical wind (VW) specific flight patterns. Coefficients are distinguished in offset (off) and slopes (slo), where applicable with lift coefficient in the upwash corrections (upw). Cross-calibration is referred to with the calibration steps in parentheses. Coefficients in parentheses were only used for intermediate calculations.

Variable	Coefficient	A. LAB	B. TUN	C. COMP	D. RACE	E. SQUA	F. VARI	G. VW
p_s	off [hPa]	−1.220		−2.26				
	slo [−]	225.170						
p_q	off [hPa]	−13.895	0.216					
	slo [−]	6.068	1.049		1.085			
p_α	off [hPa]	−13.706						
	slo [−]	6.088						
p_β	off [hPa]	−13.704						
	slo [−]	6.060						
T_s	off [K]	−33.821						
	slo [−]	9.762						
α	upw,off [rad]					(0.005)	(0.017)	0.039 (F)
	upw,slo [rad]							−0.027
	slo [−]							1(F)
β	upw,off [rad]					(−0.012)	(−0.014)	−0.004 (E)
	upw,slo (rad)							−0.010 (E)
	slo [−]							1(F)

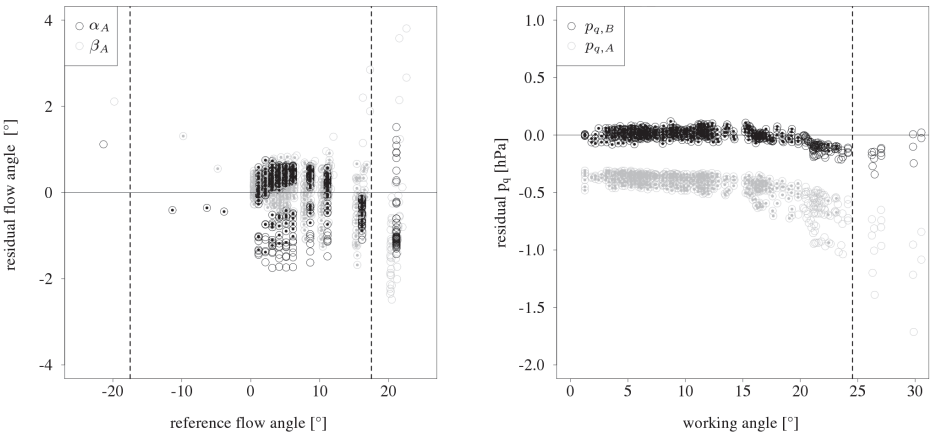


Fig. 7. Wind tunnel evaluation results: open circles represent the residuals for all combinations of flow angles (α_A, β_A, left) and dynamic pressure before ($p_{q,A}$) and after ($p_{q,B}$) wind tunnel correction (right). Full circles indicate subsets that lie in the (extended) operational flow angle range of ±17.5°. These subsets are used for the uncertainty assessment. Dashed vertical lines indicate the corresponding thresholds of flow angle and working angle (acos(cosα cosβ)), respectively.

effects of sensor accuracies, i.e. 0.6° and 0.06 hPa, respectively (Table 5). Consequently in-flight tests with similar scope were omitted.

4.1.3 Step C – Tower fly-bys

A wing induces lift by generating lower pressure atop and higher pressure below the airfoil. Since the five hole probe is measuring at a position being located below the wing, the static pressure (p_s) measurement is potentially biased. An offset adjustment was estimated to -2.26 ± 0.43 hPa from comparison with tower based measurements (Table 6). No dependence of the adjustment on true airspeed, i.e. propeller slipstream or lift coefficient, could be detected. This can most probably be attributed to the small v_{tas} range of the WSMA. The dynamic pressure (p_q) is measured using

a differential pressure sensor between the static- and total pressure ports (Fig. 3). These ports are located only 165 mm apart from each other, and are therefore subject to the same position error. Consequently position error cancels out in the p_q measurement.

4.1.4 Step D – Racetracks

For racetrack and wind square flights, inhomogeneous flight legs were discarded using the stationarity test for wind measurements by Foken and Wichura (1996). Respective optimality criteria Eqs. (5)–(6) were applied to 1 Hz block averages of the remaining legs. The dynamic pressure inverse reference from racetracks suggests an offset (0.213 hPa) and slope (1.085) correction. Without considering additional dependences, the fit for different power settings is well determined with 0.115 hPa residual standard error and $R^2 = 0.974$. We have seen that the wing upwash (v_{up}^w) in front of the wing of the WSMA is effective forward, right and upward (Fig. 5), while the propeller slipstream is directed backward and upward at the 5HP location in body coordinate system. As net effect we find that the magnitude of dynamic pressure ($p_{q,B}$) measured at the 5HP tip, and with it the calculated true airspeed, is reduced. The slope correction from racetracks was used to account for this loss in $p_{q,B}$. The suggested offset was considered as inversion residue of atmospheric inhomogeneities during the racetrack manoeuvres, and consequently discarded. Also an analogous correction for the static pressure measurement has been discarded: at a flight altitude of ≈1000 m a.s.l. the maximum correction at 6 hPa dynamic pressure (0.51 hPa) corresponds to ≤10 m difference in altitude. This is in the order of the aircraft altitude fluctuations. With this the accepted error is in the same order as the uncertainty of the static pressure offset from tower fly-bys.

4.1.5 Step E – Wind squares

Over all wind square flights the optimality criteria for horizontal and vertical wind components were averaged. Offsets for α (0.005 rad) and β (−0.012 rad) were iteratively adjusted to minimize this single measure.

4.1.6 Step F – Variance optimization

From the variance optimization method a second set of offsets for α (0.017 ± 0.003 rad) and β (-0.014 ± 0.001 rad) was found. The optimality criteria were applied to each leg individually and the offsets determined were averaged. The estimates differ from those for the wind squares by 0.6° for α and by 0.1° for β. While the deviation for β lies within the installation repeatability, the deviation for α corresponds to ≈0.3 m s^{-1} uncertainty in the vertical wind (Table 4). The wing's upwash in Eq. (3), and its variation due to different aircraft trim was considered as one potential reason for this deviation: While flying level with similar power setting, flights in denser air in the atmospheric boundary layer (e.g. variance optimization flights) require a smaller lift coefficient, i.e. less wing pitch, than flights in the less dense air in the free atmosphere (e.g. wind square flights). That is CL in Eq. (2) is inversely proportional to air density. For terrain following flights in the ABL vertical acceleration due to pilot input is likely to additionally alter the wing loading, and with it CL.

4.1.7 Step G – Vertical wind treatment

Among all the wind components the vertical wind measurement is of prevailing importance to reliably compute Eddy-Covariance fluxes. Correspondingly its assessment and treatment is the centrepiece of this calibration procedure. To disentangle the comprehensive sequence of assessment and treatment, Step G is further divided into six sub-steps:

Step G1 – Net effect of aircraft trim and wing loading.
Step G2 – Reformulation of the upwash correction.
Step G3 – Parametrization of aircraft trim and wing loading effects.
Step G4 – Parametrization of offsets.
Step G5 – Iterative treatment of cross dependences.
Step G6 – Application to terrain following flight.

Step G1 – Net effect of aircraft trim and wing loading

The net effect of changing aircraft trim and wing loading was investigated with the forced oscillation (VW3) flight pattern. During the flight on 25 June 2008 the wing pitching angle was modified by ±5°. This results in a climb angle amplitude >10° (compared to <5° during terrain following flight) and a maximum vertical velocity of |4| m s^{-1} (Fig. 8). It is evident that the modelled upwash (v_{up}^w) is proportional to the lift coefficient, as defined in Eq. (3). Assuming a constant vertical wind, not necessarily but likely approaching zero above the ABL, measured variations in v_w^m are referred to as "observed net upwash". As opposed to the parametrization by Crawford et al. (1996) for fixed-wing aircraft, the observed net upwash at the five hole probe location is smaller by one order of magnitude but more variable, as well as phase inverted with CL. These findings are confirmed with the level acceleration – deceleration flight VW1 with a long period (180 s) and negligible vertical velocity (Fig. 8). With it a potential phase difference between airspeed and wing loading during the VW3 flight can be ruled out as well explanation for the antagonistic relationship between CL and the observed upwash. Introducing a lower proportionality factor to Eq. (3) could account for the particular properties of the WSMA wing. This would reduce the magnitude of the modelled upwash, but could not explain the higher variability as well as the phase inversion of the observed upwash. In contrast to FWA, the WSMA's wing-tip and trike nose approach each other with increasing airspeed (Sect. 4). The wing's centre of pressure is within <10 % chord length of the centre of gravity. Through this distance, wing pitching by −5° would result in a decrease

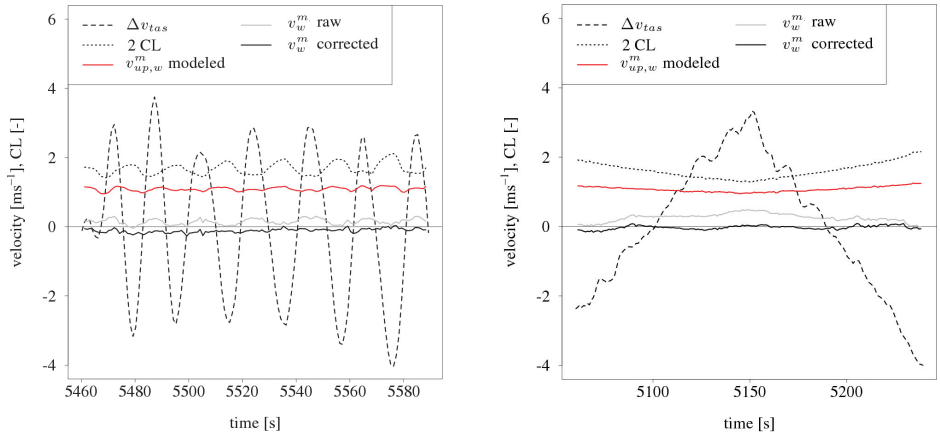


Fig. 8. Forced oscillation pattern (VW3, left) and level acceleration–deceleration pattern (VW1, right) on 25 June 2008. For improved legibility the average is subtracted from true airspeed (Δv_{tas}) and lift coefficient is inflated by the factor two (2 CL). Displayed is the vertical wind (v^{m}_w) before (raw) and after (corrected) correction for dependence on the lift coefficient. For comparison the modelled upwash ($v^{\text{m}}_{\text{up},w}$) is presented, which was computed using Eq. (3) and decomposed and rotated from wing- into meteorological coordinates using Eq. (A13).

of the normalized distance between centre of pressure and 5HP (n), by $\approx -1\,\%$. Though modelled upwash inversely varies with n in Eq. (3), the approach of wing and trike alone can not explain the upwash phase inversion. Considering a change from high true airspeed (low lift coefficient) to low true airspeed (high lift coefficient) during level flight actually a number of effects contribute to the observed net upwash: (a) increase of upwash production from the wing according to Eq. (3), (b) decrease of wing circulation effective at the 5HP through larger distance and opening angle between 5HP and wing, and (c) decrease of propeller induced upwash. The latter effects counteract the wing induced upwash. In addition (d) the shape of the aeroelastic wing, as well as (e) the flow around the trike change. Therefore the net upwash of a WSMA can neither be parametrized nor corrected with the Crawford et al. (1996) wing upwash model alone. Garman et al. (2008) on the other hand proposed to correct for upwash by considering the actual wing loading factor (LF), which carries information on the aircraft's vertical acceleration. In contrast to the study of Garman et al. (2008), WSMA weight, fuel level as well as dynamic pressure (p_{q}) are known. Therefore CL can be directly determined and used instead of LF. This has the advantage that information on the aircraft's trim, i.e. information on above effects (b)–(e), is included: as formulated in Eq. (2), p_{q} carries information on v_{tas} at given air density. Over eight independent flights of patterns VW1, VW2 and VW3 the observed net upwash is correlated with CL (-0.53 ± 0.16), change in v_{tas} (0.57 ± 0.16), and wing pitch (-0.50 ± 0.20). I.e. the application of control forces leads to a simultaneous change in both, the wind field and the lift coefficient. In the following section we will use this relationship to treat the effect of pilot input on the wind measurement.

Step G2 – Reformulation of the upwash correction

Crawford et al. (1996) and Kalogiros and Wang (2002b) have shown that the upwash Eq. (3) can be reformulated as a function of CL in the 5HP measured attack angle (α). Yet, as opposed to FWA, the WSMA is defined in two different coordinate systems, those of the wing (upwash) and the trike (5HP measurement, Fig. 2). Therefore an upwash correction in α would not explicitly consider the mobility of the trike in the wing circulation. As shown above only minor uncertainty would be introduced for pitching movements, though rolling movements and their possible influence would be left out. Consequently wind measurements during turning manoeuvres would not be covered, which however are not the subject of this study. In return correcting the upwash in α yields several advantages compared to explicitly modelling and subsequently subtracting the upwash: one explanatory variable is sufficient to explain the upwash variability effectively incident at the 5HP. With it a potential phase shift between variables measured in the wing and the trike body coordinate systems, as well as additional coordinate transformations are omitted. Therefore the net upwash variability was treated for straight and level flight (such as during EC soundings) using a linear model in α:

$$\begin{aligned}\alpha_\infty &= \alpha_{\text{A}} - \alpha_{\text{upw}} \\ &= \alpha_{\text{A}} - \left(\alpha_{\text{upw,off}} + \alpha_{\text{upw,slo}}\text{CL}\right),\end{aligned} \tag{7}$$

with α_∞ the (desired) free air stream angle of attack, α_{A} being the 5HP derived attack angle, and α_{upw} an additive attack angle provoked by the upwash with $\alpha_{\text{upw,off}}$ and $\alpha_{\text{upw,slo}}$ being its constant part and sensitivity on CL, respectively. As outlined above the complex interaction of wing upwash and aeroelasticity, distance and opening angle with the 5HP,

propeller slipstream and flow around the trike is collectively correlated in CL. This offers the possibility of a dynamic treatment of the net flow distortion in one single explanatory variable. The purpose of this correction is to reposition the mean vertical wind under the effects of terrain following flight in the ABL, i.e. correct for its offset drift.

Step G3 – Parametrization of aircraft trim and wing loading effects

For vertical wind specific flights (VW) above the ABL, α in Eq. (A11) was changed iteratively until yielding a vertical wind (v_w^{m}) of zero. Subtracting this inverse reference of α_∞ from α_{A} gives us an estimate of α_{upw}. To reduce scatter, α_{upw} was averaged after binning over increments of 0.01 CL. From this binned and averaged data $\alpha_{\mathrm{upw,off}}$ and $\alpha_{\mathrm{upw,slo}}$ were obtained with a linear fit (Fig. 9). Scatter for the level acceleration–deceleration (VW1) flight and the forced oscillation (VW3) flight (both on 25 June 2008) is significantly reduced by implementing the binning procedure. Before binning, the VW1 flight shows a slight hysteresis, probably due to the accelerating- and decelerating legs. Non-binned values of the VW3 flight are considerably more scattered than for VW1. This can be attributed to the rising and sinking process of the aircraft and changing flow regimes about the wing during load change at the turning points. Fitted coefficients differed slightly between the two flights. The analysis was continued with the coefficients of the better determined VW1 flight ($R^2 = 0.85$), which amount to $\alpha_{\mathrm{up,off}} = 0.031$ rad and $\alpha_{\mathrm{upw,slo}} = -0.027$ rad. That is α_{A} would be overestimated by $\approx 1.7°$ if the WSMA could fly at zero lift. The effect decreases with slower flight at a rate of $\approx -1.7°$ per CL. The correction reduces the vertical wind fluctuations for systematic deviations resulting from varying wing trim (53 %, relative to the bias-adjusted overall fluctuation) and wing loading (16 %) for above named VW1 and VW3 flights, respectively. For the VW3 flight (Fig. 8) the decorrelation of v_w^{m} with v_{tas} improves from 0.79 to -0.11, and the decorrelation with wing pitch improves from -0.78 to 0.17. Assuming zero vertical wind, RMSE and BIAS slightly improved from 0.17 and 0.15 m s^{-1} to 0.13 and -0.11 m s^{-1}, respectively. Lenschow (1986) proposed a 10 % criteria for the effect of the aircraft's vertical velocity ($v_{\mathrm{gs}}^{\mathrm{m},z}$) on v_w^{m}. It is employed as an operational limit by the Research Aviation Facility of the US National Centre for Atmospheric Research (NCAR, Tjernström and Friehe, 1991). Using the upwash correction this measure was improved from 3.8 % to 2.7 % (σ). A slight trend in v_w^{m} remains. The correction was also applied to two smooth oscillation (VW2) patterns. The flight on 24 June 2008 was conducted in less calm air and two different power settings were applied (Fig. 10). The correction changed overall RMSE and BIAS from 0.26 and 0.13 m s^{-1} to 0.25 and -0.13 m s^{-1}, respectively. That is the quality measures did not indicate significant improvement, but the vertical wind BIAS was inverted. However after correction the change in power settings (4800–5000 s: $v_{\mathrm{tas}} = 26$ m s^{-1}, 5200–5400 s: $v_{\mathrm{tas}} = 28$ m s^{-1}) did not alter the offset in v_w^{m} anymore (correlation of v_w^{m} with v_{tas} decreased from 0.42 to 0.21). The dependence on vertical movement decreased only slightly from 14.7 % to 13.5 % (σ), however correlation of v_w^{m} with $v_{\mathrm{gs}}^{\mathrm{m},z}$ is <0.02. Due to the less calm atmosphere σ might not be representative for their cross dependence in this case. The VW2 flight on 25 June 2008 was again conducted in calm air at 28 m s^{-1} true airspeed (Fig. 10). Here our correction leads to a change in RMSE and BIAS from 0.22 and 0.20 m s^{-1} to 0.09 and -0.02 m s^{-1}. After correction the dependence on vertical aircraft movement increased slightly from 7.7 % to 8.3 % (σ), which still well agrees with the limit used by NCAR.

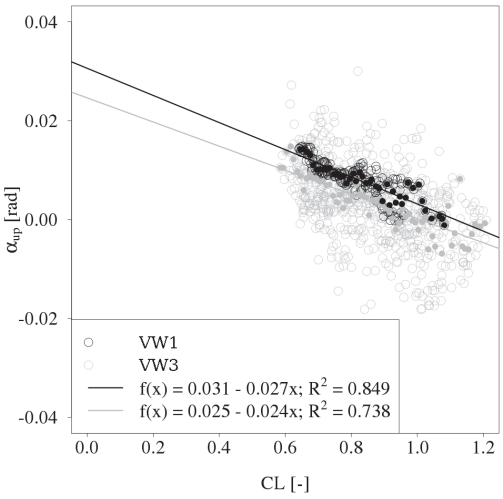


Fig. 9. Upwash attack angle (α_{up}) as function of the lift coefficient (CL) for two vertical wind specific flight patterns on 25 June 2008 (VW1, VW3). α_{up} is the difference of attack angle as measured by the five hole probe and the inverse reference of the free airstream attack angle (Sect. 4.1 Step G3). Open circles depict the entire 1 Hz dataset. Full circles are averages after binning over increments of 0.01 CL.

Step G4 – Parametrization of offsets

We have learned from the VW3 pattern (Fig. 8), that calculation of v_w^{m} was improved for flights which include changes in aircraft trim and lift. From the VW2 pattern we have seen that v_w^{m} is independent of slow aircraft rising and sinking manoeuvres, and the decorrelation of v_w^{m} with v_{tas} was improved (Fig. 10). Hence the WSMA wind measurement fulfils the requirements to be used during terrain following flights in the ABL. Yet after applying the correction, BIAS in v_w^{m} was negative, ranging from -0.13 to -0.02 m s^{-1}. Assuming

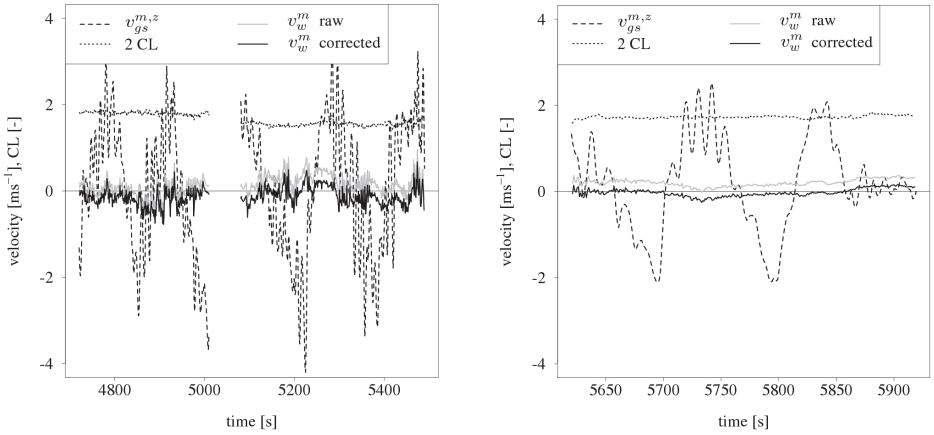


Fig. 10. Smooth oscillation flights (VW2) on 24 June 2008 (left) and 25 June 2008 (right). In addition to the variables explained in Fig. 8 the vertical aircraft velocity ($v_{\mathrm{gs}}^{\mathrm{m},z}$) is shown. Additional information is given in Sect. 4.1 Step G3.

independence of v_w^{m} from v_{tas}, the detected BIAS depends on $\alpha_{\mathrm{up,off}}$ in Eq. (7). Both, $\alpha_{\mathrm{up,off}}$ and $\alpha_{\mathrm{upw,slo}}$ were determined using the VW1 flight on 25 June 2008 during ambiguous cyclonality atop and below measurement altitude (Table 3). In Fig. 9 the determination of $\alpha_{\mathrm{upw,slo}}$ depends on the change of CL, while the offset $\alpha_{\mathrm{up,off}}$ depends on the ambient vertical wind. During the inverse reference procedure v_w^{m} was forced to zero while, e.g. in an anticyclone, subsidence occurs. In such a situation $\alpha_{\mathrm{up,off}}$ would be underestimated. During the VW flights on 24 and 25 June 2008, cyclonality and BIAS in v_w^{m} both changed. While $\alpha_{\mathrm{upw,slo}}$ is insensitive, no constant $\alpha_{\mathrm{up,off}}$ could be determined from the VW flights. At this point the variance optimization flights in the ABL are of importance. Assuming constant ABL height (approximately fulfilled for noontime EC soundings) the second optimality criteria states that due to mass conservation $\overline{v_w^{\mathrm{m}}}$ approaches zero for a sufficiently high number of datasets. With it $\alpha_{\mathrm{up,off}}$ was determined directly from ABL flights. Using the first variance optimization optimality criteria, i.e. the minimization of the wind variance, also α and β slopes were tested.

Step G5 – Iterative treatment of cross dependences

An approach similar to Eq. (7), the explanation of upwash in α, was used to explain sidewash in β:

$$\beta_\infty = \beta_{\mathrm{A}} - \left(\beta_{\mathrm{upw,off}} + \beta_{\mathrm{upw,slo}}\mathrm{CL}\right), \tag{8}$$

using the calibration criteria of the wind square flights for parametrization. According to Eq. (A11) cross dependence occurs between the parametrizations in α and β. This problem was solved by iterating the optimality criteria for wind square, vertical wind, and variance optimization flights in sequence. The order of this sequence, i.e. first optimizing for the horizontal wind components (v_{uv}^{m}), then for the vertical wind component (v_w^{m}), was chosen due to their different order of magnitude and importance for EC application. Spurious contamination with v_w^{m} would change v_{uv}^{m} only by a fraction. The other way around however would result in considerably higher contamination in v_w^{m}. The final calibration coefficients are summarized in Table 6.

Compared to the upwash parametrization, sidewash was found to be modest ($\beta_{\mathrm{upw,off}} = -0.004$ rad) and less sensitive regarding CL ($\beta_{\mathrm{upw,slo}} = -0.010$ rad, Table 6). This is in line with the initial attempt to resolve the circulation around the wing and the trike movement explicitly (Fig. 5). The findings also confirm our initial hypothesis that flow transverse to the pressure probe requires less correction than in the vertical direction (Sect. 2.1): leaving dynamic considerations aside (i.e. lift coefficient is zero), the magnitude of the sideslip angle correction is one order of magnitude lower than the attack angle correction (0.039 rad, offsets in Table 6). For a true airspeed of 30 m s^{-1} this affects the wind measurement to approximately -0.1 m s^{-1} and 1.2 m s^{-1}, respectively. The transverse distortions increase and the vertical distortions decrease at a ratio of ≈1:3, when considering interactions with propeller and wing (i.e. non-zero lift coefficient, slopes in Table 6).

Step G6 – Application to terrain following flight

The uncertainty of the correction during terrain following flight can be assessed from the regression errors, e.g. in the upwash attack angle for the vertical wind. In Step G3 the level acceleration-deceleration flight was used to calibrate the regression slope. This slope did not differ significantly compared to the forced oscillation flight, with -0.027 ± 0.002 and -0.024 ± 0.002, respectively. In Step G4 we used 12 ABL flights or 360 km of flight data to

parametrize the regression offset (0.039 ± 0.003). The combined errors in slope and intercept were applied to the terrain following flight on 31 July 2009. The resulting uncertainty in the vertical wind measurement is within $0.1\,\mathrm{m\,s^{-1}}$ (RMSE) for the mean and 1 % for the variance. This compares to the magnitude of the correction, which is in the order of $0.5\,\mathrm{m\,s^{-1}}$ for the mean and 3 % for the variance, respectively.

Many variables in the ABL scale with distance from the exchange surface (e.g. Mahrt, 2000; Mahrt et al., 2001). Interpretable results can be achieved by flying at approximately constant altitude above ground (e.g. Betts et al., 1990; Vickers and Mahrt, 1997). Yet many terrestrial surfaces are not ideally flat. During terrain following measurements we focus on mostly horizontal flight tracks between 40 and 80 km extend. Typical altitude gradients during such flight patterns are 100 m vertical on 10 km horizontal, and rarely reach $\pm 5°$ climb angle. In order to adjust aircraft altitude, the pilot anticipates the terrain contours at a scale of kilometres. At typical airspeeds this corresponds to an adjustment of power setting and wing pitch through the pilot at frequencies <0.1 Hz. In addition to the low frequency control forces, also external forces due to atmospheric turbulence and mesoscale motions meet in the hangpoint. An increase in the vertical wind variance in the order of 1 % would result when applying the correction for low frequency pilot actions to the entire frequency spectrum. Consequently the correction is only applied to frequencies <0.1 Hz. This is achieved by calculating Eqs. (7), (8) through a third order Savitzky-Golay complementary filter (e.g. Chen et al., 2004). The treatment leads to a decrease in the vertical wind variance in the order of -3 %. This is expected since the impact of the low frequency pilot actions on the wind measurement is removed.

Low frequency atmospheric motions, such as turbulent organized structures, overlap with pilot actions in frequency space. The effect of low frequency atmospheric motions on the correction can be estimated with a simple example. In a large-scale downdraft of velocity w_d an aircraft of mass m and airspeed v_tas has to produce the total lift $L = m a^{g,z}(1 + w_\mathrm{d}/v_\mathrm{tas})$, see Eq. (1). As compared to zero vertical wind conditions the lift, and with it the lift coefficient in Eq. (2), is changed by the ratio of $w_\mathrm{d}/v_\mathrm{tas}$. For the flight on 31 July 2009 a ratio of ± 1 % is equivalent to $w_\mathrm{d} = \pm 0.27\,\mathrm{m\,s^{-1}}$, a typical value for turbulent organized structures in the ABL (e.g., Steinfeld et al., 2007). To test the influence on the correction a sinusoidal signal with amplitude 1 % and frequency 0.01 Hz was added to the measured lift coefficient. The maximum deviation from the undisturbed measurement is within ± 3 %, or sub centimetre. The variance of the vertical wind is changed by <0.01 %. We conclude that the correction can be applied to the entire frequency range ≤ 0.1 Hz without introducing significant uncertainty to the wind measurement.

4.2 Wind measurement evaluation

After completing all calibration steps, the wind measurement with the WSMA was evaluated. The evaluation was carried out in three lines of analysis, (a) uncertainty propagation, (b) wind square flights, and (c) comparison to ground based wind measurements. For a true airspeed of $27\,\mathrm{m\,s^{-1}}$ the propagation of uncertainties in sensors (flow angle differential pressures, dynamic- and static pressures, static temperature, and water vapour pressure), their basic calibration and wind model description yield an uncertainty (σ) of 0.76°, 0.76°, and $0.34\,\mathrm{m\,s^{-1}}$ in attack angle (α), sideslip angle (β) and true airspeed, respectively (Table 5). Feeding the input uncertainty Eq. (B1) with these quantities extends the uncertainty propagation to the wind components (Table 4). The input error is formulated worst case, and parametrized for the 95 % confidence intervals of the attitude and flow angles. In addition the uncertainty of the inertial navigation system ($0.02\,\mathrm{m\,s^{-1}}$) was considered in the wind vector Eq. (A1). This allows to estimate the maximum potential uncertainty by sensor setup and wind model description. The results for the maximum overall uncertainty bounds are 0.66 and $0.57\,\mathrm{m\,s^{-1}}$ for the horizontal (v_{uv}^m) and vertical (v_w^m) wind components, respectively.

Figure 11 shows the results of all wind square flights. For wind velocities $>2\,\mathrm{m\,s^{-1}}$ v_{uv}^m determined for individual legs deviate less than 10 % from the average for the entire square. The residuals did not scale with the average wind velocity, to a greater degree they are likely to result from an incomplete removal of wind field inhomogeneities over the 12 km long flight paths. Therefore a horizontal wind velocity of $2\,\mathrm{m\,s^{-1}}$ can not be considered as a detection limit for wind measurements from WSMA. Also no systematic deviation for aircraft orientation could be detected. However v_w^m shows a slight sensitivity of -0.05 on v_tas ($R^2 = 0.46$). Using the cardinal direction evaluation criteria Eq. (6), RMSE in α_∞, β_∞ and $|v_\mathrm{tas}^\mathrm{m}|$ were computed to 0.31, 0.33 and $0.26\,\mathrm{m\,s^{-1}}$, respectively. These compare well to the results from the uncertainty propagation (Tables 4 and 5), which amount to 0.31, 0.36 and $0.34\,\mathrm{m\,s^{-1}}$ for α_A, β_A and v_tas, respectively.

Figure 12 shows a qualitative comparison of WSMA and ground based wind measurements for the flight on 15 October 2008. The vertical profile shows an equal number of flights at 24 and $27\,\mathrm{m\,s^{-1}}$ true airspeed. Despite one outlier in v_v^m and v_w^m at 120 m a.g.l., no distinct differences in average wind velocities between ground based measurements and WSMA are apparent. The comparability of WSMA and ground based wind measurement was further quantified by calculating RMSE and BIAS for all measurements accepted for the comparison (Table 3). The impact of calibration Steps C–G on these measures is displayed in Fig. 13. The measurement of the horizontal wind components (v_{uv}^m) was mainly improved (14 %, relative to the initial uncertainty) by means of the in-flight dynamic pressure correction (Step D). After the wind square analysis (Step E) the measurement

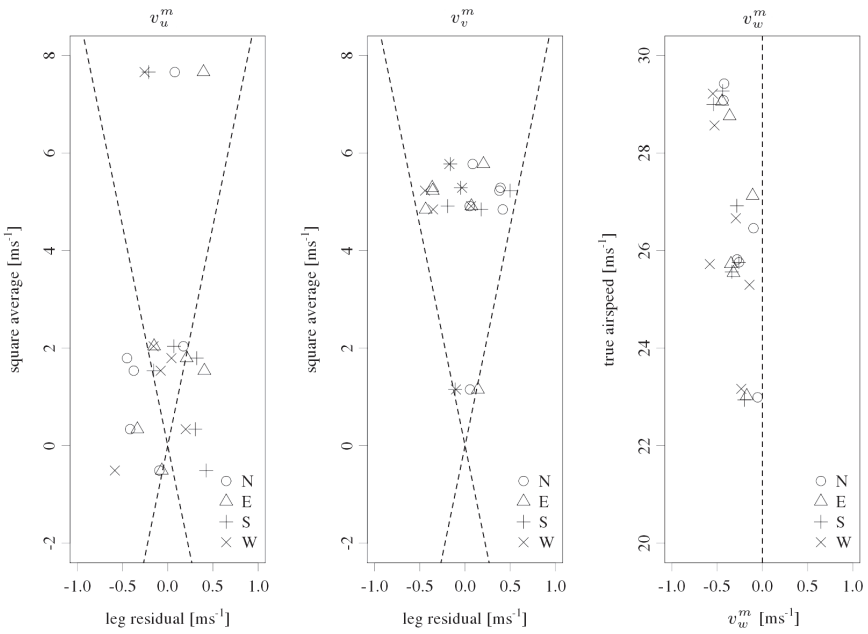


Fig. 11. Results from the wind square flights. For the horizontal wind components (v_{uv}^{m}) the x-axis displays the residuals (leg average–square average), while the y-axis shows the wind magnitude. In contrast the vertical wind component (v_w^{m}) is plotted against the true airspeed. Flight legs are depicted with different symbols according to their position in the square pattern. Dashed lines indicate a 10 % criteria for v_{uv}^{m}, and the zero line for v_w^{m}.

was not further improved nor deteriorated. Yet the vertical wind measurement (v_w^{m}) receives its greatest improvement (31 %) during Steps F–G, i.e. variance optimization and vertical wind specific patterns: During these steps BIAS and dBIAS, i.e. its dependence on v_{tas}, were reduced. In contrast to the findings from the wind square analysis, with a sensitivity of ≈+0.05 a slight positive dependence of all wind components on v_{tas} remained. Considering all data couples between WSMA and ground based measurements, RMSE and BIAS amount to 0.50 and $-0.07\,\mathrm{m\,s^{-1}}$ for v_{uv}^{m} and 0.37 and $-0.10\,\mathrm{m\,s^{-1}}$ for v_w^{m}, respectively. In addition to the above mentioned outlier, two more suspects were identified for the flight on 18 October 2008, again concurrent for v_v^{m} and v_w^{m}. A possible explanation is the increased land surface heterogeneity sensed by the aircraft while travelling through the wind field. On the northern and western limbs of the aircraft cross pattern, forest patches of ≥200 m edge length interrupt the flat arable land immediately upwind. Therefore WSMA measurements can include turbulence and wake effects generated at the forest edges. In contrast tower measurements are not subject to comparable roughness changes until ≈2 km in upwind direction. Omitting the three outliers from the statistics, RMSE and BIAS between WSMA and ground based measurements improve to 0.39 and $-0.11\,\mathrm{m\,s^{-1}}$ for v_{uv}^{m} and 0.27 and $-0.10\,\mathrm{m\,s^{-1}}$ for v_w^{m}, respectively.

4.3 Discussion

Distortions of the wind measurement originating from the interactions of the aeroelastic wing, propeller and trike structural features were successfully corrected for conditions approximating straight, terrain following flight. Yet the treatments integral to Eqs. (A7), (7) and (8) leave room for improvement: Compared to ground based measurements the aircraft underestimated the wind components $\approx -0.1\,\mathrm{m\,s^{-1}}$. A possible reason could be the discarded offset during the dynamic pressure (p_{q}) in-flight calibration (Sect. 4.1, Step D). Rather forcing the linear fit to zero would slightly enhance the slope of p_{q} and with it compliance to the aircraft's inertial speed.

During the wind square and comparison flights contradictory sensitivities (regression slope −0.05 versus +0.05) of the wind components on the true airspeed were found. For the variability in v_{tas} during a thermally turbulent flight in the atmospheric boundary layer ($\sigma = 1.24\,\mathrm{m\,s^{-1}}$, Fig. 4) this corresponds to $\pm 0.06\,\mathrm{m\,s^{-1}}$ deviation in the wind components. Since this deviation is one order of magnitude lower than the system's input uncertainty, it was not further treated.

The lift coefficient is used as sole explanatory variable for the observed net upwash in the linear calibration models Eqs. (7) and (8). This treats the influence of aircraft trim (i.e. dynamic pressure) and lift (i.e. loading factor) on

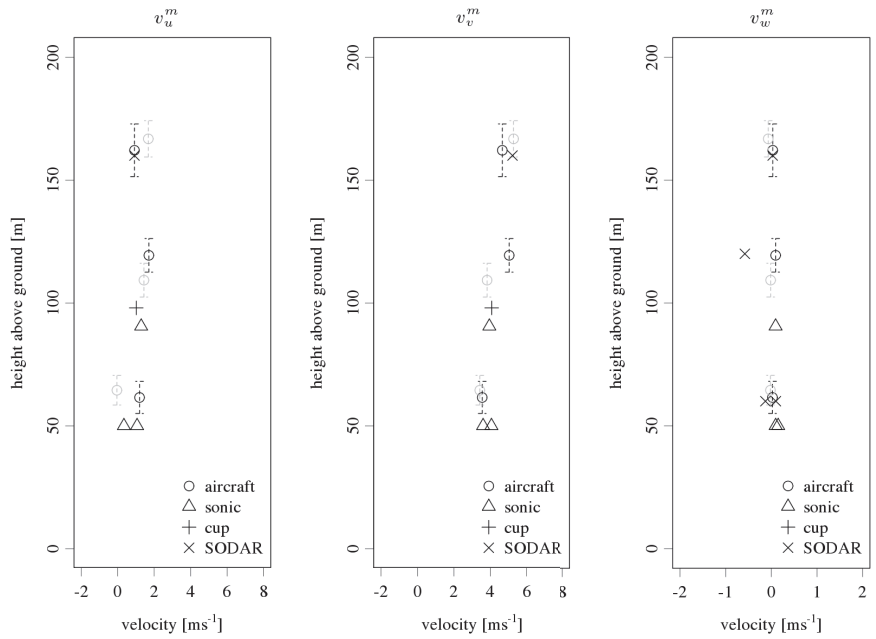


Fig. 12. Vertical profiles for horizontal (v_{uv}^{m}) and vertical (v_w^{m}) wind components of simultaneous ground based and weight-shift microlight aircraft measurements on 15 October 2008, 14:50–16:00 CET. Different symbols indicate the different wind sensors. Black circles represent aircraft measurements at 24 m s^{-1} true airspeed, while grey circles represent measurements at 27 m s^{-1} true airspeed. Vertical error bars indicate one standard deviation of the aircraft altitude.

the wind measurement with similar sensitivity. The study by Visbal and Shang (1989) however shows that the flow field response of airfoils to pitch oscillations depends on the excitation frequency. With the Fourier method proposed by Kalogiros and Wang (2002b) the frequency dependence of the wing induced upwash can be modelled for FWA. The distinct difference from time domain methods is an amplified (≈20 %) upwash correction in the inertial subrange of atmospheric turbulence compared to lower frequencies. Due to little contributions of the inertial subrange, the effect on the eddy flux measurement at flight altitude (≤4 %) is however relatively small. At the same time a transformation from the wing to the trike coordinate system would be required, carrying a potentially variable phase difference. Moreover the interactions with propeller and trike, resulting in the net flow distortion, remain untreated. Isolating these interactions would require considerably more in-flight data and analytical effort. In return such procedure could address forenamed dependence of the wind components on v_{tas} and additionally allow for superior wind measurements during turning manoeuvres.

5 Conclusions

We have shown that carefully computed wind vector measurements using a weight-shift microlight aircraft are not inferior to those from other airborne platforms. A 10 % limit of contamination of the wind components by the aircraft movement, as used by the US National Centre for Atmospheric Research, was fulfilled even during severe vertical manoeuvring. For flights including rising and sinking of the aircraft, such as during terrain following Eddy-Covariance applications, three independent lines of analysis yield comparable uncertainty. This convergence is remarkable and emphasizes the integrity of sensing elements and wind model description. The procedure further enables to quantify the overall operational uncertainty (root mean square error) to 0.4 m s^{-1} for the horizontal and 0.3 m s^{-1} for the vertical wind components.

Independent consideration of trike movement and wing circulation according to the fixed-wing aircraft theory was not successful. Instead flow distortion of fuselage, propeller and wing were minimized by an approach integrated in the dynamic pressure and flow angle computations. The magnitude of distortion was treated as slope correction in the dynamic pressure computation. The distortion's distribution in components longitudinal, transverse and vertical to the wind measurement was subsequently parametrized in the attack- and sideslip angle computations. The lift coefficient

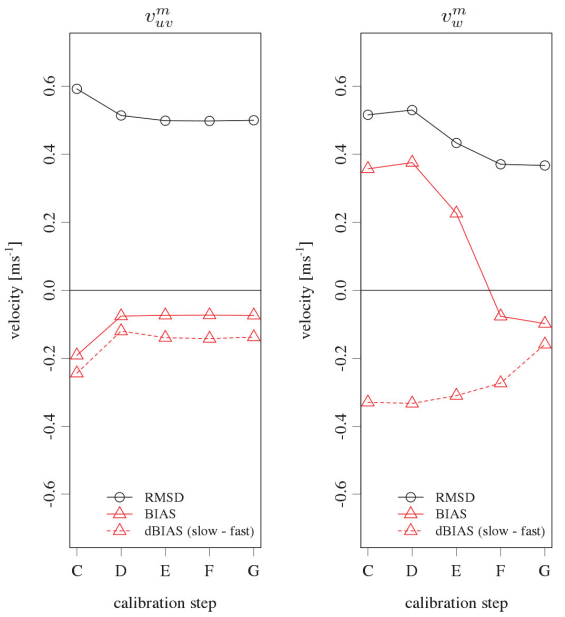


Fig. 13. Influence of the calibration Steps C–G on root mean square error (RMSE) and bias (BIAS) between weight-shift microlight aircraft and all simultaneous ground based measurements of the horizontal (v_{uv}^{m}) and the vertical (v_w^{m}) wind components. dBIAS indicates the difference in BIAS between measurements at 27 and 24 m s^{-1} true airspeed.

was successfully used as sole variable explaining the upwash distribution, containing in it the effects of aircraft trim and lift. After the treatment an inconclusive dependence of the vertical wind measurement on the aircraft's true airspeed remained. In-flight tests relate this dependence to an uncertainty of 0.06 m s^{-1} in the vertical wind measurement. As compared to ground based measurements the final wind components were marginally underestimated by the aircraft (≈ -0.1 m s^{-1}).

Our findings emphasize that the 3-D wind vector can be measured reliably from a highly transportable and low-cost weight-shift microlight aircraft. Hence the necessary basis is provided for the study of precision and spectral quality of the wind measurement, which is prerequisite for reliable Eddy-Covariance flux measurements. This brings the weight-shift microlight aircraft platform an important step closer towards a fullfeatured environmental research aircraft.

Supplementary material related to this article is available online at: http://www.atmos-meas-tech.net/4/1515/2011/amt-4-1515-2011-supplement.pdf.

Acknowledgements. Special recognition has to be given to Josef-Michael Burger, Matthias Mauder, Frank Neidl, Rainer Steinbrecher and Rose Zuurbier at the Karlsruhe Institute of Technology, Institute for Meteorology and Climate Research. Burger and Zuurbier carried out the bulk of the wind-tunnel measurements, whereas Mauder contributed his micrometeorological advise. Neidl programmed and continuously maintained the data acquisition system, while Steinbrecher initiated wind- and flux measurements with the weight-shift microlight aircraft in the first place. Our gratefulness to Xunhua Zheng and her work-group at the Chinese Academy of Sciences, Institute of Atmospheric Physics, who was hosting our project and providing indispensable infrastructure. We are much obliged to Frank Beyrich of the German Meteorological Service, Richard-Aßmann-Observatory, who permitted us to carry out the evaluation flights and provided us with the corresponding SODAR- and tower data for this study. Our thanks to Jens Bange and Aline van den Kroonenberg of the Technical University of Braunschweig, Institute of Aerospace Systems (now Eberhard Karls University of Tübingen, Institute for Geoscience), for their tireless advise. Stipend funding by the German Academic Exchange Service, Helmholtz Association of German Research Centrers, China Scholarship Council and the European Union under the Science and Technology Fellowship China is acknowledged. The flight in Inner Mongolia was funded by the German Research Foundation, research group 536 "Matter fluxes in grasslands of Inner Mongolia as influenced by stocking rate". The publication was funded by the German Research Foundation and Open Access Publishing Fund of the Karlsruhe Institute of Technology. We thank John Kalogiros and an anonymous reviewer for their valuable comments and detailed feedback.

Edited by: J.-P. Pommereau

References

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W. U. K., Pilegaard, K., Schmid, H., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapour, and energy flux densities, B. Am. Meteorol. Soc., 82, 2415–2434, 2001.

Bange, J. and Roth, R.: Helicopter-borne flux measurements in the nocturnal boundary layer over land – a case study, Bound.-Lay. Meteorol., 92, 295–325, 1999.

Betts, A. K., Desjardins, R. L., Macpherson, J. I., and Kelly, R. D.: Boundary-layer heat and moisture budgets from FIFE, Bound.-Lay. Meteorol., 50, 109–138, 1990.

Beyrich, F., Leps, J. P., Mauder, M., Bange, J., Foken, T., Huneke, S., Lohse, H., Ludi, A., Meijninger, W. M. L., Mironov, D., Weisensee, U., and Zittel, P.: Area-averaged surface fluxes over the LITFASS region based on Eddy-Covariance measurements, Bound.-Lay. Meteorol., 121, 33–65, 2006.

Bissolli, P. and Dittmann, E.: The objective weather type classification of the German Weather Service and its possibilities of application to environmental and meteorological investigations, Meteorol. Z., 10, 253–260, 2001.

Bohn, D. and Simon, H.: Mehrparametrige Approximation der Eichräume und Eichflächen von Unterschall- bzw. Überschall-5-Loch-Sonden, Archiv für Technisches Messen und Meßtechnische Praxis, 470, 31–37, 1975.

Boiffier, J.-L.: The Dynamics of Flight: The Equations, Wiley, 1998.

Brown, E. N., Friehe, C. A., and Lenschow, D. H.: The use of pressure-fluctuations on the nose of an aircraft for measuring air motion, J. Clim. Appl. Meteorol., 22, 171–180, 1983.

Chen, J., Jonsson, P., Tamura, M., Gu, Z. H., Matsushita, B., and Eklundh, L.: A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter, Remote Sens. Environ., 91, 332–344, 2004.

Cook, M. V.: The theory of the longitudinal static stability of the hang-glider, Aeronaut. J., 98, 292–304, 1994.

Cook, M. V. and Spottiswoode, M.: Modelling the flight dynamics of the hang glider, Aeronaut. J., 110, 1–20, 2006.

Cooper, W. A. and Rogers, D.: Effects of air-flow trajectories around aircraft on measurements of scalar fluxes, J. Atmos. Ocean. Tech., 8, 66–77, 1991.

Crawford, T. L. and Dobosy, R. J.: A sensitive fast-response probe to measure turbulence and heat-flux from any airplane, Bound.-Lay. Meteorol., 59, 257–278, 1992.

Crawford, T., Dobosy, R., and Dumas, E.: Aircraft wind measurement considering lift-induced upwash, Bound.-Lay. Meteorol., 80, 79–94, 1996.

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

Egger, J., Bajrachaya, S., Heinrich, R., Kolb, P., Lammlein, S., Mech, M., Reuder, J., Schaper, W., Shakya, P., Schween, J., and Wendt, H.: Diurnal winds in the Himalayan Kali Gandaki valley, Part III: Remotely piloted aircraft soundings, Mon. Weather Rev., 130, 2042–2058, 2002.

Foken, T.: Temperaturmessungen mit dünnen Platindrähten, Z. Meteorol., 29, 299–307, 1979.

Foken, T. and Wichura, B.: Tools for quality assessment of surface-based flux measurements, Agr. Forest Meteorol., 78, 83–105, 1996.

Garman, K. E., Hill, K. A., Wyss, P., Carlsen, M., Zimmerman, J. R., Stirm, B. H., Carney, T. Q., Santini, R., and Shepson, P. B.: An airborne and wind tunnel evaluation of a wind turbulence measurement system for aircraft-based flux measurements, J. Atmos. Ocean. Tech., 23, 1696–1708, 2006.

Garman, K. E., Wyss, P., Carlsen, M., Zimmerman, J. R., Stirm, B. H., Carney, T. Q., Santini, R., and Shepson, P. B.: The contribution of variability of lift-induced upwash to the uncertainty in vertical winds determined from an aircraft platform, Bound.-Lay. Meteorol., 126, 461–476, 2008.

Gratton, G. B.: The weightshift-controlled microlight aeroplane, Proceedings of the Institution of Mechanical Engineers Part G, J. Aerospace Eng., 215, 147–154, 2001.

Hobbs, S., Dyer, D., Courault, D., Olioso, A., Lagouarde, J. P., Kerr, Y., McAneney, J., and Bonnefond, J.: Surface layer profiles of air temperature and humidity measured from unmanned aircraft, Agronomie, 22, 635–640, 2002.

Junkermann, W.: An ultralight aircraft as platform for research in the lower troposphere: system performance and first results from radiation transfer studies in stratiform aerosol layers and broken cloud conditions, J. Atmos. Ocean. Tech., 18, 934–946, 2001.

Junkermann, W.: The actinic UV-radiation budget during the ESCOMPTE campaign 2001: results of airborne measurements with the microlight research aircraft D-MIFU, Atmos. Res., 74, 461–475, 2005.

Kaimal, J. and Finnigan, J.: Atmospheric Boundary Layer Flows: Their Structure and Measurement, Oxford University Press, 1994.

Kalogiros, J. A. and Wang, Q.: Calibration of a radome-differential GPS system on a twin otter research aircraft for turbulence measurements, J. Atmos. Ocean. Tech., 19, 159–171, 2002a.

Kalogiros, J. A. and Wang, Q.: Aerodynamic effects on wind turbulence measurements with research aircraft, J. Atmos. Ocean. Tech., 19, 1567–1576, 2002b.

Khelif, D., Burns, S. P., and Friehe, C. A.: Improved wind measurements on research aircraft, J. Atmos. Ocean. Tech., 16, 860–875, 1999.

Leise, J. A. and Masters, J. M.: Wind measurement from aircraft, Tech. rep., United States Department of Commerce, National Oceanic and Atmospheric Administration, Aircraft Operations Center, MacDill Air Force Base, 1993.

Lenschow, D. H.: Aircraft measurements in the boundary layer, in: Probing the Atmospheric Boundary Layer, American Meteorological Society, Boston, Massachusetts, USA, 39–55, 1986.

Mahrt, L.: Surface Heterogeneity and Vertical Structure of the Boundary Layer, Bound.-Lay. Meteorol., 96, 33–62, 2000.

Mahrt, L., Vickers, D., and Sun, J.: Spatial variations of surface moisture flux from aircraft data, Adv. Water Res., 24, 1133–1141, 2001.

Mauder, M., Desjardins, R. L., and MacPherson, I.: Creating surface flux maps from airborne measurements: application to the Mackenzie area GEWEX study MAGS 1999, Bound.-Lay. Meteorol., 129, 431–450, 2008.

Steinfeld, G., Letzel, M. O., Raasch, S., Kanda, M., and Inagaki, A: Spatial representativeness of single tower measurements and the imbalance problem with eddy-covariance fluxes: results of a large-eddy simulation study, Bound.-Lay. Meteorol., 123, 77–98, 2007.

Tjernström, M. and Friehe, C. A.: Analysis of a radome air-motion system on a twin-jet aircraft for boundary-layer research, J. Atmos. Ocean. Tech., 8, 19–40, 1991.

Trenkle, F. and Reinhardt, M.: In-flight temperature measurements, vol. 2 of Advisory Group for Aerospace Research and Development: Flight Test Instrumentation Series, North Atlantic Treaty Organization, Neuilly-sur-Seine, France, 1973.

van den Kroonenberg, A., Martin, T., Buschmann, M., Bange, J., and Vorsmann, P.: Measuring the wind vector using the autonomous mini aerial vehicle M2AV, J. Atmos. Ocean. Tech., 25, 1969–1982, 2008.

Vickers, D. and Mahrt, L.: Quality control and flux sampling problems for tower and aircraft data, J. Atmos. Ocean. Tech., 14, 512–526, 1997.

Visbal, M. R. and Shang, J. S.: Investigation of the flow structure around a rapidly pitching airfoil, AIAA J., 27, 1044–1051, 1989.

Williams, A. and Marcotte, D.: Wind measurements on a maneuvering twin-engine turboprop aircraft accounting for flow distortion, J. Atmos. Ocean. Tech., 17, 795–810, 2000.

Wyngaard, J. C.: The effects of probe-induced flow distortion on atmospheric-turbulence measurements, J. Appl. Meteorol., 20, 784–794, 1981.

Appendix C: Supplement to Metzger et al. (2011)

Supplement to

Metzger, S., Junkermann, W., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.:
Measuring the 3-D wind vector with a weight-shift microlight aircraft

It is the intention of this supplement to make transparent the procedures in the main article, and to guide the reader through the relevant calculations. Supplement A provides the formulary necessary to compute the wind vector from a weight-shift microlight aircraft. A model to propagate uncertainty through the wind vector equations is provided in Supplement B1. Relevant notation and abbreviations are listed in Supplement C. References to literature are given at the end of the document.

Supplement A Wind measurement transformation equations

The wind measurement from aircraft requires several coordinate systems, as well as angles to transform between them (Fig. 2). We define the wind vector $\boldsymbol{v}^{\mathrm{m}} = (v_{\mathrm{u}}^{\mathrm{m}}, v_{\mathrm{v}}^{\mathrm{m}}, v_{\mathrm{w}}^{\mathrm{m}})$ in the standard meteorological coordinate system (MCS, superscript m, positive eastward, northward, and upward). Then $\boldsymbol{v}^{\mathrm{m}}$ can be calculated from navigation, flow and attitude measurements: In the MCS $\boldsymbol{v}^{\mathrm{m}}$ is expressed as the vector difference between the aircraft's ground speed vector ($\boldsymbol{v}_{\mathrm{gs}}^{\mathrm{m}}$), directly measured by the inertial navigation system (INS), and the true airspeed vector ($\boldsymbol{v}_{\mathrm{tas}^{\mathrm{m}}}$), essentially measured by the five hole probe (5HP, Williams and Marcotte, 2000):

$$\begin{aligned}\boldsymbol{v}^{\mathrm{m}} &= \boldsymbol{v}_{\mathrm{gs}}^{\mathrm{m}} - \boldsymbol{v}_{\mathrm{tas}}^{\mathrm{m}} \\ &= \boldsymbol{v}_{\mathrm{gs}}^{\mathrm{m}} - \mathbf{M}^{\mathrm{bm}} \times \left(\mathbf{M}^{\mathrm{ab}}(-v_{\mathrm{tas}}) + \boldsymbol{v}_{\mathrm{lev}}^{\mathrm{b}} \right). \end{aligned} \tag{A1}$$

Yet the quantity directly measured by the 5HP is the true airspeed scalar v_{tas}. The second, decomposed form of the wind vector Eq. (A1) indicates that several calculation steps are necessary to arrive at the desired vector quantity $\boldsymbol{v}_{\mathrm{tas}}^{\mathrm{m}}$.

In the following we will walk through these successive steps, starting with the 5HP measurements. From the ports of the 5HP (Fig. 3) three differential pressures were measured:

$$p_{\mathrm{q,A}} = p_{\mathrm{t}} - p_{\mathrm{s}}, \tag{A2}$$

$$p_{\alpha} = p_3 - p_1, \text{ and} \tag{A3}$$

$$p_{\beta} = p_4 - p_2. \tag{A4}$$

Measured dynamic pressure $p_{\mathrm{q,A}}$ (subscript upper-case letters A–G indicate calibration stage), and attack- and sideslip differential pressures p_{α}, p_{β} were used to calculate the airflow angles (Williams and Marcotte, 2000):

$$\alpha_{\mathrm{A}} = \frac{2}{9\sin(2\tau)} \frac{p_{\alpha}}{p_{\mathrm{q,A}}}, \text{ and} \tag{A5}$$

$$\beta_{\mathrm{A}} = \frac{2}{9\sin(2\tau)} \frac{p_{\beta}}{p_{\mathrm{q,A}}}. \tag{A6}$$

Here $\tau = 45^{\circ}$ is the angle between the central port p_t and the other ports p_1 through p_4 on the 5HP half sphere. Defining the normalization factor $D = \sqrt{1 + \tan^2\alpha_{\mathrm{A}} + \tan^2\beta_{\mathrm{A}}}$ the measured dynamic pressure $p_{\mathrm{q,A}}$ can be corrected for the pressure drop occurring at elevated airflow angles:

$$p_{\mathrm{q,B}} = p_{\mathrm{q,A}} \left(\frac{9 - 5D^2}{4D^2} \right)^{-1}. \tag{A7}$$

Now we can derive v_{tas} from the thermodynamic measurements of the 5HP: Due to stagnation at the tip of the 5HP ambient air is heated from its intrinsic temperature (T_{s}) to total temperature (T_{t}). Assuming adiabatic heating, Bernoulli's equation

$$\begin{aligned} v_{\mathrm{tas}}^2 &= 2c_{\mathrm{p,h}}(T_{\mathrm{t}} - T_{\mathrm{s}}) \\ &= 2c_{\mathrm{p,h}}T_{\mathrm{s}} \left[\left(\frac{p_{\mathrm{s}}}{p_{\mathrm{s}} + p_{\mathrm{q}}} \right)^{-\kappa} - 1 \right], \end{aligned} \tag{A8}$$

gives v_{tas} as a function of the temperature difference (Leise and Masters, 1993). Since T_{t} can not be measured directly, it is substituted in Eq. (A8) by the adiabatic process (ram rise)

$$T_{\mathrm{t}} = T_{\mathrm{s}} \left(\frac{p_{\mathrm{s}}}{p_{\mathrm{s}} + p_{\mathrm{q}}} \right)^{-\kappa}, \tag{A9}$$

with the Poisson number $\kappa = 1 - \frac{c_{\mathrm{v,h}}}{c_{\mathrm{p,h}}}$. Furthermore the wind measurement should be independent of air humidity (subscript h). Therefore the specific heats under constant pressure (subscript p) $c_{\mathrm{p,h}}$ or constant volume (subscript v) $c_{\mathrm{v,h}}$ of moist air have to be derived from the specific heat constants for dry air (subscript d) and water vapour (subscript w), $c_{\mathrm{p,d}} = 1005\,\mathrm{J\,kg^{-1}\,K^{-1}}$, $c_{\mathrm{p,w}} = 1846\,\mathrm{J\,kg^{-1}\,K^{-1}}$, $c_{\mathrm{v,d}} = 718\,\mathrm{J\,kg^{-1}\,K^{-1}}$, and $c_{\mathrm{v,w}} = 1384\,\mathrm{J\,kg^{-1}\,K^{-1}}$ (Khelif et al., 1999):

$$\begin{aligned}
c_{\mathrm{p,h}} &= c_{\mathrm{p,d}} \left[1 + q \left(\frac{c_{\mathrm{p,w}}}{c_{\mathrm{p,d}}} - 1 \right) \right], \\
c_{\mathrm{v,h}} &= c_{\mathrm{v,d}} \left[1 + q \left(\frac{c_{\mathrm{v,w}}}{c_{\mathrm{v,d}}} - 1 \right) \right], \text{ with specific humidity being} \\
q &= \varepsilon \frac{e}{p_{\mathrm{s}} + e(\varepsilon - 1)},
\end{aligned} \tag{A10}$$

where $\varepsilon = 0.622$ is the ratio of molecular weight of water vapour to that of dry air, and e is the 5HP measured water vapour pressure.

Once derived, the scalar quantity v_{tas} has to be transformed into a vector quantity. This can be achieved by defining the aerodynamic coordinate system (ACS, superscript a, positive forward, starboard, and downward), which has its origin at the 5HP tip. In this coordinate system the true airspeed vector has the components $\boldsymbol{v}_{\mathrm{tas}}^{\mathrm{a}} = (-v_{\mathrm{tas}}, 0, 0)$. Since the ACS is aligned with the streamlines its orientation however varies in time. Therefore $\boldsymbol{v}_{\mathrm{tas}}^{\mathrm{a}}$ is transformed into a fixed coordinate system, that is the trike body coordinate system (BCS, superscript b, positive forward, starboard, and downward) with its origin in the INS. This is accomplished by successive rotations about the vertical axis $\mathrm{Z^a}$ and the transverse axis $\mathrm{Y^a}$. Following Lenschow (1986) the

rotations can be summarized in the operator

$$\mathbf{M}^{\mathrm{ab}} = D^{-1} \begin{pmatrix} 1 \\ \tan\beta \\ \tan\alpha \end{pmatrix}, \tag{A11}$$

with the 5HP derived airflow angles of attack α and sideslip β, and the normalization factor D as derived in Eqs. (A5)–(A7). Since $\boldsymbol{v}_{\mathrm{tas}^{\mathrm{a}}}$ carries all its information in the first vector component, it is sufficient to apply this transformation to $-v_{\mathrm{tas}}$ in the wind vector Eq. (A1).

Now the wind vector is known in the orientation of the BCS, yet with its origin still at the 5HP tip as initially defined in the ACS. To allow for the vector difference as required in the wind Eq. (A1) we have to account for the displacement of ACS origin (5HP tip) relative to the BCS origin (INS). This is done by considering the lever arm correction vector (Williams and Marcotte, 2000):

$$\boldsymbol{v}^{\mathrm{b}}_{\mathrm{lev}} = \begin{pmatrix} \Omega^{\mathrm{b}}_{\Phi} \\ \Omega^{\mathrm{b}}_{\Theta} \\ \Omega^{\mathrm{b}}_{\Psi} \end{pmatrix} \times \begin{pmatrix} x^{\mathrm{b}} \\ y^{\mathrm{b}} \\ z^{\mathrm{b}} \end{pmatrix}, \tag{A12}$$

with INS measured body rates $\Omega^{\mathrm{b}}_{\Phi}$, $\Omega^{\mathrm{b}}_{\Theta}$, $\Omega^{\mathrm{b}}_{\Psi}$ about the X^{b}, Y^{b}, Z^{b} axes, and the displacement of the 5HP with respect to the INS along these axes, $x^{\mathrm{b}} = -0.73\,\mathrm{m}$, $y^{\mathrm{b}} = -0.01\,\mathrm{m}$, and $z^{\mathrm{b}} = 0\,\mathrm{m}$. The vector sum $\mathbf{M}^{\mathrm{ab}}(-v_{\mathrm{tas}}) + \boldsymbol{v}^{\mathrm{b}}_{\mathrm{lev}}$ in the wind Eq. (A1) then describes the true airspeed vector in the BCS.

A last step remains to obtain $\boldsymbol{v}^{\mathrm{m}}_{\mathrm{tas}}$ for use in the wind Eq. (A1), that is the transformation of the true airspeed vector from the BCS into the MCS. This is achieved by a first transformation into the geodetic coordinate system (GCS, superscript g, positive northward, eastward and downward) via successive rotations about the X^{b}, Y^{b}, Z^{b} axes (Lenschow, 1986). From there the wind vector is permutated into the MCS (positive eastward, northward and upward). The

transformations can be summarized in the operator

$$\mathbf{M}^{\mathrm{bm}} = \begin{pmatrix} M_{11}^{\mathrm{bm}} & M_{12}^{\mathrm{bm}} & M_{13}^{\mathrm{bm}} \\ M_{21}^{\mathrm{bm}} & M_{22}^{\mathrm{bm}} & M_{23}^{\mathrm{bm}} \\ M_{31}^{\mathrm{bm}} & M_{32}^{\mathrm{bm}} & M_{33}^{\mathrm{bm}} \end{pmatrix}, \text{ with} \tag{A13}$$

$$\begin{aligned}
M_{11}^{\mathrm{bm}} &= \sin\Psi^{\mathrm{b}}\cos\Theta^{\mathrm{b}}, \\
M_{12}^{\mathrm{bm}} &= \cos\Psi^{\mathrm{b}}\cos\Phi^{\mathrm{b}} + \sin\Psi^{\mathrm{b}}\sin\Phi^{\mathrm{b}}\sin\Theta^{\mathrm{b}}, \\
M_{13}^{\mathrm{bm}} &= \sin\Psi^{\mathrm{b}}\sin\Theta^{\mathrm{b}}\cos\Phi^{\mathrm{b}} - \cos\Psi^{\mathrm{b}}\sin\Phi^{\mathrm{b}}, \\
M_{21}^{\mathrm{bm}} &= \cos\Psi^{\mathrm{b}}\cos\Theta^{\mathrm{b}}, \\
M_{22}^{\mathrm{bm}} &= \cos\Psi^{\mathrm{b}}\sin\Theta^{\mathrm{b}}\sin\Phi^{\mathrm{b}} - \sin\Psi^{\mathrm{b}}\cos\Phi^{\mathrm{b}}, \\
M_{23}^{\mathrm{bm}} &= \sin\Psi^{\mathrm{b}}\sin\Phi^{\mathrm{b}} + \cos\Psi^{\mathrm{b}}\sin\Theta^{\mathrm{b}}\cos\Phi^{\mathrm{b}}, \\
M_{31}^{\mathrm{bm}} &= \sin\Theta^{\mathrm{b}}, \\
M_{32}^{\mathrm{bm}} &= -\cos\Theta^{\mathrm{b}}\sin\Phi^{\mathrm{b}}, \\
M_{33}^{\mathrm{bm}} &= -\cos\Theta^{\mathrm{b}}\cos\Phi^{\mathrm{b}},
\end{aligned}$$

where Φ^{b}, Θ^{b}, and Ψ^{b} are the INS measured attitude angles roll, pitch and heading, respectively. Finally the movement of the BCS with respect to the MCS is described by $\boldsymbol{v}_{\mathrm{gs}}^{\mathrm{m}}$ in the wind vector Eq. (A1).

Supplement B Uncertainty quantification

B1 Uncertainty propagation

In Eq. (A1) the wind vector is the difference between the aircraft's ground speed vector ($\boldsymbol{v}_{\mathrm{gs}}^{\mathrm{m}}$) and the true airspeed vector ($\boldsymbol{v}_{\mathrm{tas}}^{\mathrm{m}}$). The measurement of $\boldsymbol{v}_{\mathrm{gs}}^{\mathrm{m}}$ is readily provided by the inertial navigation system, together with the related uncertainty (Table 2). Uncertainty propagation is however required for $\boldsymbol{v}_{\mathrm{tas}}^{\mathrm{m}}$, since 12 measured quantities are merged during its calculation. The magnitude of the lever arm correction Eq. (A12), and with it possible uncertainty from this source, is two orders lower than $\boldsymbol{v}_{\mathrm{tas}}^{\mathrm{m}}$. It can therefore be neglected in the uncertainty propagation, which leaves nine measured quantities. By preprocessing Eqs. (A5)–(A10) these are further condensed to three measured quantities and three derived variables (see next paragraph for respective uncertainty propagation). Modified after Vörsmann (1985) the input uncertainty of the $\boldsymbol{v}_{\mathrm{tas}}^{\mathrm{m}}$ measurement can then be calculated from a linearised uncertainty propagation model in the vector components $v_{\mathrm{tas,c}}^{\mathrm{m}}$ (c = u, v, or w):

$$\Delta v_{\mathrm{tas,c}}^{\mathrm{m}} = \sum_{i=1}^{i=6} \left| \frac{\delta v_{\mathrm{tas,c}}^{\mathrm{m}}}{\delta f_i}\, \sigma(f_i) \right|, \tag{B1}$$

with $\frac{\delta v_{\mathrm{tas,c}}^{\mathrm{m}}}{\delta f_i}$ being the partial derivatives of Eqs. (A11) and (A13) inserted into the wind vector Eq. (A1). Thereby the input uncertainty of $\boldsymbol{v}_{\mathrm{tas}}^{\mathrm{m}}$ can be expressed as function of the (assumed independent) input variables (f_i), with $\sigma(f_i)$ being their respective uncertainty. Here f_i are three quantities directly measured by the INS (i.e. pitch- (Θ^{b}), roll- (Φ^{b}) and heading- (Ψ^{b}) angles) and three variables derived from five hole probe measurements (i.e. attack angle (α), sideslip angle (β) and true airspeed scalar (v_{tas})). Such a procedure is conservative, since it assumes uncertainty interference, but not cancellation. It yields the maximum possible uncertainty triggered by the combined effects of σf_i. The derivatives were further simplified by small-angle approximation. This simplification allows to express the input uncertainty with sign and sensitivity as a function of Ψ^{b}, whereas the full form yields the maximum absolute input uncertainty for different flight states.

In analogy uncertainty propagation models were formulated for the three derived variables α in Eq. (A5), β in Eq. (A6) and v_{tas} in Eq. (A8). These permit to express the actual uncertainties originating from the six remaining directly measured quantities, i.e. both flow angle differential pressures, dynamic- and static pressures, static temperature, and water vapour pressure.

With this setup the overall uncertainty at each stage of the wind calculation procedure can be evaluated through Gaussian uncertainty propagation (e.g., Taylor, 1997):

$$\sigma_{\mathrm{gau}} = \sqrt{\sum_{i=1}^{N} \sigma_i^2}, \tag{B2}$$

with N being the number of (assumed linear and independent) uncertainty terms contributing to the stage investigated.

B2 Uncertainty measures

For applications in the atmospheric boundary layer the comparison to a reference standard can yield an integral measure of confidence under varying conditions (e.g., Vogt and Thomas, 1995; Mauder et al., 2006). Therefore this study employs two basic bivariate criteria for the comparison of wind components. These are the root mean square deviation (RMSE) and bias (BIAS) between sample and reference (ISO, 1993):

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A_i - R_i)^2}, \tag{B3}$$

$$\mathrm{BIAS} = \frac{1}{N}\sum_{i=1}^{N}(A_i - R_i), \tag{B4}$$

with N being the number of data couples R_i and A_i, R_i being the ith reference observation and A_i the ith observation by aircraft sensors, sampled simultaneously. RMSE is also called comparability and is a measure of overall uncertainty. BIAS is the systematic difference between

the mean of the measurements and the reference. These criteria were not normalized, since no consistent dependence on the wind magnitude or the aircraft's true airspeed was found.

Supplement C Notation

Scalars and vector components are displayed in italics, vectors are displayed in bold italics, and matrices are displayed in bold roman typeface, respectively. Where applicable coordinate systems and respective axes are indicated by superscripts, whereas subscripts are used as specifiers.

C1 Operators

$[\mathbf{M}]$	Transformation matrix
$[\delta]$	Differential operator
$[\Delta]$	Difference operator

C2 Parameters and variables

a	Acceleration
A	Aircraft measurement
BIAS	Bias
CL	Lift coefficient
c_{p}	Specific heat at constant pressure
c_{v}	Specific heat at constant volume
D	Derived term containing airflow angles
f	Place-holder for input variables
g	Gravitational acceleration
i	Continuous index
L	Lift
LF	Loading factor
$\frac{L}{S}$	Wing loading
m	Mass
n	Normalized centre of pressure – 5HP separation distance
N	Sample size
p	Pressure

q	Specific humidity
R	Reference measurement
RMSE	Root mean square error
S	Wing surface area
T	Temperature
v	Velocity scalar or vector component
$\boldsymbol{v}$	Velocity vector
x, y, z	Distances on respective coordinate axes
$\frac{z}{L}$	Stability parameter
α	Angle of attack
β	Angle of sideslip
ε	Ratio of molecular masses
Θ	Pitch
κ	Poisson number
Φ	Roll
ξ	Wing upwash direction
π	Perimeter constant
ρ	Air density
σ	Standard deviation, RMSE
τ	Angle between central and surrounding ports on half-sphere
Ψ	Heading
Ω	Body rate

C3 Subscripts – superscripts

1–4	Pressure ports
∞	Free airstream
$+, -$	Into wind, with wind
$\tilde{}$	Wind tunnel

a	Aerodynamic coordinate system, positive forward, starboard, and downward
A–G	Calibration steps
b	Body coordinate system, positive forward, starboard and downward
d	Dry air
g	Geodetic coordinate system, positive northward, eastward and downward
gau	Gaussian uncertainty propagation
gs	Ground speed
h	Humid air
lev	Lever arm
m	Meteorological coordinate system, positive eastward, northward and upward
off	Offset
q	Dynamic-
r	Inverse reference
s	Static-
slo	Slope
t	Total-
tas	True airspeed
u, v, w	Wind components in x, y, z directions
up	Upwash
w	Water vapour; Wing coordinate system, positive forward, starboard and downward
x, y, z	Standard Cartesian coordinate axes
α	Angle of attack
β	Angle of sideslip

C4 Abbreviations

5HP	Five hole probe
ABL	Atmospheric boundary layer
ACS	Aerodynamic coordinate system, positive forward, starboard, and downward
a.g.l.	Above ground level
a.s.l.	Above sea level
BCS	Body coordinate system, positive forward, starboard and downward
D-MIFU	Name of aircraft
DAQ	Data acquisition
E	East
EC	Eddy covariance
EIDAS	Embedded Institute for Meteorology and Climate Research data acquisition system
FWA	Fixed-wing aircraft
GCS	Geodetic coordinate system, positive northward, eastward and downward
INS	Inertial navigation system
IU	Input uncertainty
LI	Lindenberg
MCS	Meteorological coordinate system, positive eastward, northward and upward
N	North
S	South
ST	Lake Starnberg
ULS	Universal laser sensor
VW1–VW3	Vertical wind specific flight patterns
W	West

WCS	Wing coordinate system, positive forward, starboard and downward
WSMA	Weight-shift microlight aircraft
XI	Xilinhot

References

ISO: Statistics – vocabulary and symbols – Part 1: Probability and general statistical terms, vol. ISO 3534-1, International Organization for Standardization, Geneva, Switzerland, 1993.

Khelif, D., Burns, S. P., and Friehe, C. A.: Improved wind measurements on research aircraft, J. Atmos. Ocean. Tech., 16, 860–875, 1999.

Leise, J. A. and Masters, J. M.: Wind measurement from aircraft, Tech. rep., United States Department of Commerce, National Oceanic and Atmospheric Administration, Aircraft Operations Center, MacDill Air Force Base, 1993.

Lenschow, D. H.: Aircraft measurements in the boundary layer, in: Probing the Atmospheric Boundary Layer, American Meteorological Society, Boston, Massachusetts, USA, 39–55, 1986.

Mauder, M., Liebethal, C., Göckede, M., Leps, J. P., Beyrich, F., and Foken, T.: Processing and quality control of flux data during LITFASS-2003, Bound.-Lay. Meteorol., 121, 67–88, 2006.

Taylor, J. R.: An Introduction to Error Analysis: the Study of Uncertainties in Physical Measurements, University Science Books, Mill Valley, California, USA, 2nd edn., 1997.

Vogt, S. and Thomas, P.: Sodar – a useful remote sounder to measure wind and turbulence, J. Wind Eng. Ind. Aerod., 54, 163–172, 1995.

Vörsmann, P.: Ein Beitrag zur Bordautonomen Windmessung, Ph.D. thesis, Technische Universität Braunschweig, 1985.

Williams, A. and Marcotte, D.: Wind measurements on a maneuvering twin-engine turboprop aircraft accounting for flow distortion, J. Atmos. Ocean. Tech., 17, 795–810, 2000.

Appendix D: Metzger et al. (2012)

Atmos. Meas. Tech., 5, 1699–1717, 2012
www.atmos-meas-tech.net/5/1699/2012/
doi:10.5194/amt-5-1699-2012

Atmospheric Measurement Techniques

Eddy-covariance flux measurements with a weight-shift microlight aircraft

S. Metzger[1,2,3], **W. Junkermann**[1], **M. Mauder**[1], **F. Beyrich**[4], **K. Butterbach-Bahl**[1], **H. P. Schmid**[1], **and T. Foken**[5]

[1]Karlsruhe Institute of Technology, Institute for Meteorology and Climate Research, Atmospheric Environmental Research, Garmisch-Partenkirchen, Germany
[2]Chinese Academy of Sciences, Institute of Atmospheric Physics, State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry, Beijing, China
[3]National Ecological Observatory Network, Fundamental Instrument Unit, Boulder, USA
[4]German Meteorological Service, Richard-Aßmann-Observatory, Lindenberg, Germany
[5]University of Bayreuth, Department of Micrometeorology, Bayreuth, Germany

Correspondence to: W. Junkermann (wolfgang.junkermann@kit.edu)

Received: 18 February 2012 – Published in Atmos. Meas. Tech. Discuss.: 30 March 2012
Revised: 17 June 2012 – Accepted: 18 June 2012 – Published: 19 July 2012

Abstract. The objective of this study is to assess the feasibility and quality of eddy-covariance flux measurements from a weight-shift microlight aircraft (WSMA). Firstly, we investigate the precision of the wind measurement ($\sigma_{u,v} \leq 0.09\,\mathrm{m\,s^{-1}}$, $\sigma_w = 0.04\,\mathrm{m\,s^{-1}}$), the lynchpin of flux calculations from aircraft. From here, the smallest resolvable changes in friction velocity ($0.02\,\mathrm{m\,s^{-1}}$), and sensible- ($5\,\mathrm{W\,m^{-2}}$) and latent ($3\,\mathrm{W\,m^{-2}}$) heat flux are estimated. Secondly, a seven-day flight campaign was performed near Lindenberg (Germany). Here we compare measurements of wind, temperature, humidity and respective fluxes between a tall tower and the WSMA. The maximum likelihood functional relationship (MLFR) between tower and WSMA measurements considers the random error in the data, and shows very good agreement of the scalar averages. The MLFRs for standard deviations (SDs, 2–34 %) and fluxes (17–21 %) indicate higher estimates of the airborne measurements compared to the tower. Considering the 99.5 % confidence intervals, the observed differences are not significant, with exception of the temperature SD. The comparison with a large-aperture scintillometer reveals lower sensible heat flux estimates at both tower (−40 to −25 %) and WSMA (−25–0 %). We relate the observed differences to (i) inconsistencies in the temperature and wind measurement at the tower and (ii) the measurement platforms' differing abilities to capture contributions from non-propagating eddies. These findings encourage the use of WSMA as a low cost and highly versatile flux measurement platform.

1 Introduction

Energy and matter fluxes between the Earth's surface and the atmosphere can be determined using the eddy-covariance (EC) method. This method is based on the Reynolds decomposition of the Navier-Stokes equation, and it assumes steady state conditions and horizontal homogeneity (e.g. Kaimal and Finnigan, 1994). Nevertheless, the EC method is frequently used in complex terrain, for which applicability is subject of on-going research (e.g. Foken et al., 2010; Göckede et al., 2008). In particular, it is assumed that the mean vertical wind approaches zero for a sufficiently long averaging interval. This requirement is more likely fulfilled by spatial than by temporal measurements, because spatial measurements enable registering atmospheric motions on larger scales (e.g. Mahrt, 2010). Under conditions of negligible advection and horizontal flux divergence, the total vertical flux is then inferred from the covariance between the vertical wind and the scalar of interest (e.g. temperature, humidity).

Ground-based measurements of turbulent fluxes are of local character and are therefore not necessarily representative of their greater surroundings, especially in complex terrain (e.g. Desjardins et al., 1997; Isaac et al., 2004b; Mahrt, 2010). The spatial gap between in-situ observations, satellite observations and modelled data needs to be considered as one plausible explanation for their frequently observed mismatch (e.g. Kanda et al., 2004; Lu et al., 2005). Here

process studies with airborne platforms provide a valuable link to understand and bridge scale discrepancies (e.g. Bange et al., 2002; Davis et al., 1992; Hiyama et al., 2007; Isaac et al., 2004a). At the same time, fixed-wing aircraft and helicopters are expensive to operate or not applicable in settings such as remote areas beyond the range of an airfield. Unmanned aerial vehicles on the other hand provide mobility, yet do not allow a comprehensive sensor package due to payload restrictions (e.g. Egger et al., 2002; Hobbs et al., 2002; Martin et al., 2011; Thomas et al., 2012). Here the weight-shift microlight aircraft (WSMA) can provide an alternative at low cost-, transport- and infrastructural demand. After successfully applying a WSMA to aerosol and radiation transfer studies (Junkermann, 2001, 2005), Metzger et al. (2011) showed that carefully computed wind measurements from WSMA are not inferior to those from other airborne platforms. On this basis, the feasibility of EC flux measurements from WSMA in the atmospheric boundary layer (ABL) is explored in this study. The overarching perspective is to work towards an airborne platform that allows characterising complex terrain in remote areas, including the measurement of regional turbulent fluxes.

Contributions to the EC flux measurement originate from turbulent atmospheric motions on a variety of wavelengths and amplitudes. In order to reliably estimate the total flux, the fluctuations of the vertical wind and the scalars must be measured with high accuracy and precision. Furthermore, the instrumentation and data acquisition must possess a suitable frequency response and sampling rate. In the case of airborne measurements, the carrier can additionally influence the spectral quality of the measurement. Therefore, the present study commences with (i) an assessment of the measurement errors. To evaluate the system performance, we (ii) compare spectral properties, averages, deviations and fluxes between WSMA and tower-based EC measurements. The analysis continues with (iii) a study of the measurements' spatial context, which is inferred from footprint modelling. The (iv) comparison to a large-aperture scintillometer (LAS) brings to attention the effect of larger-scale atmospheric motions on the results and completes the study.

2 Materials and methods

2.1 The weight-shift microlight aircraft

The structure of a WSMA differs from common fixed-wing aircraft: it consists of two distinct parts, the wing and the trike, which hangs below the wing and contains the pilot, engine and the majority of the scientific equipment. This particular structure provides the WSMA with exceptional transportability and climb rate, which qualifies it for applications in complex and inaccessible terrain. A detailed description of the physical properties of the WSMA used in this study as well as characteristics and manufacturers of sensors and data acquisition is given in Metzger et al. (2011). In short, most variables are sampled at 100 Hz and are block-averaged and stored at 10 Hz, yielding a horizontal resolution of approximately 2.5 m. To conduct fast wind measurements, the WSMA is outfitted with a combination of global positioning system and inertial measurement unit (GPS/IMU), and a five-hole pressure probe (5HP). The principle is to resolve the meteorological wind vector from the vector difference of the aircraft's inertial velocity (captured by the GPS/IMU) and the wind vector relative to the aircraft (captured by the 5HP). The structural features of the WSMA also influence the wind measurement: (i) the wing deforms aeroelastically with aircraft trim, and (ii) the trike is free to rotate in pitch and roll against the wing. Metzger et al. (2011) present a time domain procedure which treats the impact of the WSMA's structural features as well as pilot input on the wind measurement. The remaining maximum deviation of the vertical wind component is $0.15\,\mathrm{m\,s^{-1}}$ during severe vertical manoeuvres. At typical airspeeds between 23–30 $\mathrm{m\,s^{-1}}$, simultaneous wind measurements from WSMA and ground-based instrumentation agree within $0.3\,\mathrm{m\,s^{-1}}$ for the vertical and within $0.4\,\mathrm{m\,s^{-1}}$ for the horizontal components (root mean square error). The present study investigates the potential influence of resonance from the WSMA's engine or propeller, or from the natural frequencies of trike and wing, on the wind measurement. For this purpose, acceleration measurements in the hang point of trike and wing, in the global positioning system/inertial measurement unit and in the five-hole probe, are used. The acceleration measurement in the hang point is transformed to the trike coordinate system. A 100 Hz dataset consisting of $\approx 3 \times 10^5$ data points sampled during a level long-distance flight on 31 July 2009 (Metzger et al., 2011, Table 3) is used for the assessment.

Air temperature is measured with a 50 µm thermocouple. The temperature error introduced by intermittent solar radiation at the unshielded thermocouple is < 0.05 K at nominal true airspeed (Metzger et al., 2011). An OP2 infrared gas analyser (IRGA, ADC Bioscientific, Great Amwell, UK) is used to measure the concentration of water vapour. The instrument response of both the thermocouple and the IRGA is 50 Hz. In addition, a slow (2 Hz instrument response) humidity reference from a TP3 dew point mirror (Meteolabor AG, Wetzikon, Switzerland) is stored at ≥ 0.1 Hz. Vertical profile flights revealed a dependence of the IRGA measurements on flight altitude. This dependence was related to a malfunctioning temperature compensation of the light source as well as air permeability of the light chamber. From measurements of calibration gases in a climate chamber, the temperature compensation is updated in post-processing. Similar measurements were conducted in a pressure chamber to determine the time constant of the light chamber permeability (≈ 60 s or 1500 m of horizontal flight). In order to correct the permeability effect, a third-order Savitzky-Golay complementary filter (Chen et al., 2004) is used. The complementary filter corrects for the IRGA's drift by basing the humidity

fluctuations measured by the IRGA on the slow dew point mirror reference. A window size of 13.9 s or ≈ 350 m maximises the integral over the humidity power spectrum, and is used to correct the measurements.

In the present study, also the influence of measurement precision on the eddy-covariance flux results is investigated. For this purpose, we follow Garman et al. (2006) and define measurement precision as $1\,\sigma$ repeatability. The precision of all variables entering the EC flux calculation is presented in Table 1. In the case of the GPS/IMU, precision originates from Kalman filter outputs, and in the case of the 5HP, it is calculated from laboratory and wind tunnel measurements (Metzger et al., 2011).

Table 1. Measurement precision of global positioning system/inertial measurement unit (GPS/IMU), five-hole probe (5HP), thermocouple, and infrared gas analyzer (IRGA).

Quantity	Precision
Heading (GPS/IMU)	0.1°
Pitch, Roll (GPS/IMU)	0.04°
3-D velocity (GPS/IMU)	0.02 m s^{-1}
3-D angular rate (GPS/IMU)	0.01° s^{-1}
3-D acceleration (GPS/IMU)	0.01 m s^{-2}
Attack angle (5HP)	0.08°
Sideslip angle (5HP)	0.18°
True airspeed (5HP)	0.05 m s^{-1}
Temperature (thermocouple)	0.04 K
Humidity (IRGA)	0.005 g m^{-3}

2.2 Field campaign

A comparison between the airborne WSMA and ground-based measurements was carried out during a flight campaign between 14 and 21 October 2008. This experiment was performed around the boundary layer field site Falkenberg (52.2° N, 14.1° E) of the German Meteorological Service (DWD), Richard-Aßmann Observatory, Lindenberg, Germany. This field site lies in the basically flat North German Plain, and the terrain height varies between 40 m and 130 m above sea level (a.s.l.) within an area of $20 \times 20\,\mathrm{km}^2$. To characterize surface heterogeneity, we use the Corine Land Cover 2006 data with a horizontal resolution of 100 m (Version 13, European Environment Agency, 2010). The arable land was harvested before the study period, and, consequently, the surface properties differed mainly between but not within landscape units. We thus regrouped the 28 Corine Land Cover fractions in the study area into five landscape units, thereby reducing unnecessary scatter (Figs. 7 and 9). The resulting representation of the landscape around the Falkenberg site ($20 \times 20\,\mathrm{km}^2$) is dominated by agriculture (47 %) and forests (38 %), interspersed by equal amounts (5 %) of lakes, meadows and settlements.

A full characterisation of the Falkenberg site and its instrumentation is presented by Beyrich and Adam (2007). Data from an instrumented 99 m tower are used for the comparison of the WSMA measurements. Sonic anemometers (USA-1 – Metek GmbH, Elmshorn, Germany) as well as open path IRGAs (LI-7500 – LI-COR Biosciences, Lincoln, USA) were installed at 50 m and 90 m above ground level (a.g.l.). These instruments sampled the wind vector, sonic temperature and humidity at a rate of 20 Hz, enabling EC flux computation. For the USA-1, the manufacturer's 2-D flow distortion correction was operationally applied to the wind vector measurement. These data are used for the comparison of the average wind as well as variances, covariances and power spectra between the WSMA and the tower. The tower was further equipped with profile measurements of temperature (HMP-45 – Vaisala Oy, Helsinki, Finland) and humidity (Frankenberger Psychrometer – Theodor Friedrichs GmbH, Hamburg, Germany) at 40, 60, 80, and 98 m a.g.l. The profiles are interpolated to the heights of the EC installations and are used to compare average temperature and humidity between tower and WSMA. A static pressure measurement (PTB220A – Vaisala Oy, Helsinki, Finland) at 74 m a.s.l. is extrapolated to the heights of the tower EC installations using the hypsometric equation. It is used for the conversion of the tower EC fluxes from kinematic units to units of energy. Tower profile and pressure data were averaged and stored in 10 min intervals.

Identical instrumentation as on the tower was used for an additional EC surface flux measurement upwind (south) of the tower base, at 2.4 m a.g.l. The half-hourly sensible heat flux was determined from this measurement as an operational product of the DWD. Global radiation was measured at 2 m using a CM24 pyranometer/albedometer (Kipp and Zonen, Delft, The Netherlands) and stored as 10 min averages. Also 10 min area-averaged surface sensible heat fluxes were derived from a large-aperture scintillometer. At an effective beam height of 43 m a.g.l., the near-infrared LAS runs along a path length of 4.7 km (Fig. 9). The LAS was developed and built by the Meteorology and Air Quality Group of the Wageningen University; technical details are presented in Meijninger et al. (2006). Furthermore, hourly estimates of the ABL depth were derived from sonic detection and ranging and wind profiler data, and from six-hourly routine radio soundings performed by the DWD. The surface sensible heat fluxes measured at the 2.4 m EC and the LAS are used in conjunction with the ABL depths to approximate vertical flux profiles. Simultaneous WSMA measurements of the sensible heat flux are compared to these flux profiles.

In the course of the flight campaign, the atmospheric conditions changed from very weak to strong turbulent mixing. The cloud cover (09:00 to 15:00 UTC) decreased from 8/8 to 4/8, and the maximum available global radiation increased from 280 W m^{-2} at the beginning to 460 W m^{-2} at the end of the campaign. Also the wind speed increased from 2 m s^{-1} to 10 m s^{-1} at the 50 m tower level, with the wind direction

changing from west to south. The ranges of the surface sensible and the latent heat fluxes were 0–100 W m^{-2} and 0–200 W m^{-2}, respectively. The sensible heat flux slightly increased in the course of the campaign, while the latent heat flux remained approximately comparable throughout the flight days. Also the maximum ABL depth increased from 250 m to 1150 m in the course of the campaign. The atmospheric stratification was neutral to unstable, with the median of the stability parameter $z/L = -0.18 \pm 0.21$ from WSMA and -0.14 ± 0.29 from tower measurements.

2.3 Data processing

Eddy-covariance data were post-processed analogously for the 99 m tower and for the WSMA turbulence measurements. The software package TK3 (Mauder and Foken, 2011) was used to process the tower EC data, applying the raw data treatments and flux corrections as put forward in Foken et al. (2012). (i) The raw data were screened for spikes using the algorithm of Hojstrup (1993). Visual inspection revealed that neighbouring spikes in the IRGA data were not detected by the algorithm. The original algorithm uses average and standard deviation criteria with low break down points for small sample sizes (Rousseeuw and Verboven, 2002). After substituting the criteria with the median and the median absolute deviation, the spikes were efficiently removed. (ii) The time delay due to separation between the vertical wind measurement and adjacent sensors was determined and corrected by maximizing their lagged correlation. (iii) To correct for potential misalignment, the USA-1 wind measurement was rotated into the streamline coordinate system using the planar-fit method by Wilczak et al. (2001). (iv) The temperature variance as well as the sensible heat flux was calculated using the crosswind correction by Liu et al. (2001). (v) The formulations by Webb et al. (1980) were used to correct the latent heat flux for density fluctuations.

To handle the WSMA data, an analysis package with similar processing steps was developed in GNU R version 2.13 (R Development Core Team, 2011), which is available upon request. Several forenamed corrections are not applicable to the WSMA measurement and were omitted: (iii) the aircraft vertical wind is already defined in geodetic normal, which is perpendicular to the spatial average of the streamlines, and (iv) the air temperature is directly measured by the thermocouple. The sensitivity of the WSMA measured fluxes on the remaining corrections was tested for a flight in the convective boundary layer ($z/L = -0.8$) at 50 m a.g.l. No spikes were present in this dataset, and consequently the spike elimination (i) had no influence on the results. The corrections for time delay (ii), high frequency spectral loss (Moore, 1986) and density fluctuations (v) only affected the latent heat flux. The median differences between applying and neglecting these corrections were in the order of 5 %, 1 % and 20 %, respectively, which is in agreement with the findings of Mauder and Foken (2006). In the following, the correction for high frequency loss due to sensor separation is not applied because its influence is negligible at measuring heights ≥ 50 m. The fluxes computed from both software packages were compared using regression analysis and showed perfect agreement with unity slope. Consequently, comparability is ensured when calculating fluxes from tower and WSMA platforms with their respective software packages.

2.4 Evaluation strategy

2.4.1 Propagation of sensor errors

The eddy-covariance technique relies upon the precise measurement of fluctuations of atmospheric quantities, based on negligible sensor drift throughout an averaging period. Our intention is to evaluate whether sensor precision and drift facilitate the use of the weight-shift microlight aircraft as turbulence measurement platform. Measured from aircraft, the determination of the wind vector requires a sequence of thermodynamic and trigonometric equations (e.g. Metzger et al., 2011). These equations propagate various sources of error, and are consequently the lynchpin for EC flux measurements from aircraft. Here we propagate known sensor precisions to atmospheric quantities, which yields the minimum resolvable change in the associated fluxes. Thereafter, the maximum achievable averaging period for the flux calculation is determined as a function of sensor drift.

2.4.2 Spectral properties of the aircraft

As opposed to ground-based measurements, the weight-shift microlight aircraft is subject to several simultaneous motions, such as locomotion, engine and propeller rotation. Our intention is to assess if and to what extent the WSMA's motions influence the wind measurement. For this purpose, a spectral analysis was carried out by fast Fourier transformation of acceleration measurements in the hang point of trike and wing, in the global positioning system/inertial measurement unit and in the five-hole probe. We present power spectra representative for these structural parts of the WSMA and interpret the spectral behaviour in the context of the wind measurement.

2.4.3 Comparison between tower and aircraft measurements

Measurements with the weight-shift microlight aircraft were conducted along a cross-shaped pattern within 1.5 km horizontal distance of the tall tower from 15 to 18 October 2008 (Fig. 7). A total of 36 flights of 3 km length or ≈ 120 s duration are compared to simultaneous tower measurements. The WSMA was travelling at two different airspeeds, 24 m s^{-1} and 27 m s^{-1}, and was flying within 0.5 ± 5.3 m altitude of the corresponding installations on the tower. The WSMA temperature and densities were transformed to potential quantities at the respective tower height. The objective of the

comparison is to assess the quality of the WSMA measurement, with focus on the EC flux. The objective is not to quantify the actual exchange between surface and atmosphere. The comparatively short flight legs are therefore a compromise of sample size (≈ 1200 data points from WSMA) and vicinity of the measurement platforms.

Differing spectral contributions can lead to a systematic bias in the flux estimates between the platforms. To ensure equal contributions from the long wave part of the spectrum, we constrain the averaging periods to the same normalized frequency. The tower averaging period $\tau_{tow} = \tau_{air} \cdot v_{tas}/|\boldsymbol{uvw}|$ then results from the flight duration for one leg τ_{air} and the ratio of airspeed v_{tas} to the module of the wind vector $|\boldsymbol{uvw}|$. For $\tau_{air} \approx 120$ s and the ratio $v_{tas}/|\boldsymbol{uvw}| \approx 5$, the appropriate tower averaging period is $\tau_{tow} \approx 600$ s or 10 min. Using this averaging period, the tower results were calculated at increments of 1 min. The WSMA was more frequently travelling upwind (180 ± 720 m median difference) than downwind of the tower. In a window of ± 10 increments, the tower result was chosen that minimized the scatter (root mean square error) between all flux measurements of both platforms. This allows taking into account advection between the platforms, as well as potential timing differences of the data acquisition systems. Best agreement was reached for a shift of 2 ± 6 increments, corresponding to an upwind distance of 600 ± 1800 m for an air mass travelling at 5 m s^{-1}.

In order to detect systematic differences between tower and WSMA measurements, a regression-like analysis was applied to all 36 flights. Simple least-squares regression is strictly applicable only when one measurement is without error (Lindley, 1947). This however is not the case for the measurements in our study, which are subject to uncertainties such as random statistical error. Instead, we use maximum-likelihood fitting of a functional relationship (MLFR, Ripley and Thompson, 1987). This method assigns a weight to each data couple in the relationship, which is inversely proportional to its error variances. In our case, the squared random statistical errors in the tower and WSMA measurements are used, which appreciates reliable data and depreciates uncertain data couples. These errors are inferred from the integral length scales of the WSMA measurements (Appendix A), and define an inner and an outer scale of confidence in the comparison. The errors in the MLFR coefficients are determined from a jackknife estimator (Quenouille, 1956; Tukey, 1958). Since the regression intercepts were not significant, the relationships were forced through the origin, and confidence intervals were determined from the slope error. The coefficient of determination R^2 was calculated in analogy to weighted least-squares regression (Kvalseth, 1985; Willett and Singer, 1988). It is the proportion of variation in weighted Y that can be accounted for by weighted X. Finally, the residual standard error is determined using Eq. (A4).

2.4.4 Spatial analysis

We use footprint modelling in order to assess the spatial context of measurements. For this purpose, the along-wind footprint parameterization of Kljun et al. (2004) was combined with a suitable crosswind distribution (Appendix B). The resulting model is computationally fast, considers 3-D dispersion and is applicable beyond the atmospheric surface layer. We compare the overlap of the tower and WSMA footprints as well as the contribution of different land covers to the measurements.

2.4.5 Comparison with large-aperture scintillometer

In addition to tower- and aircraft-based eddy-covariance measurements, we also include sensible heat flux estimates from the large-aperture scintillometer in the comparison. On seasonal average, LAS measures 10–20 % higher values of the sensible heat flux compared to tower EC measurements (e.g. Liu et al., 2011; Meijninger et al., 2006). In particular above heterogeneous terrain, the capture of elevated, non-propagating eddies (NPE) by the LAS, but not by the tower EC, is discussed as a potential reason for the observed differences (Foken et al., 2010). In this respect, the spatially averaged EC measurement from WSMA is similar to the LAS. Consequently, the objective of the comparison with LAS is to aid the interpretation of systematic differences between the spatially (aircraft) and temporally (tower) averaged EC measurements. On 20 and 21 October 2008, two flights of 4.7 km length or $\approx$150 s duration were conducted $\approx$1 km to the east of, and parallel to the LAS measuring path (Fig. 9). The duration of the WSMA flight translates to $\approx$750 s or 12.5 min averaging interval at the 50 m and 90 m levels of the tower. Nevertheless, the tower flux measurements are averaged over 10 min, identical to Sect. 2.4.3. For longer averaging intervals, the time series became increasingly instationary on 21 October 2008, and the flux magnitude decreased. WSMA temperature and density measurements were transformed to potential quantities at the mean flight altitude, i.e. 108 m and 119 m a.g.l. on 20 and 21 October 2008, respectively. Using boundary layer scaling, the results from LAS and WSMA are compared to simultaneous tower EC measurements at 2.4 m, 50 m and 90 m a.g.l., which are located at the LAS transmitter site.

3 Results

3.1 Propagation of sensor errors

In the first part of this section, we assess the WSMA's measurement precision and its impact on EC flux measurements over short flight legs. In the second part, we evaluate the maximum flux averaging period facilitated by the drift (accuracy) of the sensors.

3.1.1 Measurement precision and least resolvable flux

The measurement precisions in Table 1 were superimposed over the turbulence raw data of the 36 tower–aircraft comparison flights ($N = 37\,000$). Both original and manipulated datasets were processed through the entire wind computation, and the deviations in the wind components were compared ($\sigma_u = 0.07\,\mathrm{m\,s^{-1}}$, $\sigma_v = 0.09\,\mathrm{m\,s^{-1}}$, $\sigma_w = 0.04\,\mathrm{m\,s^{-1}}$). Drawing on Lenschow and Sun (2007), we assume that a minimum signal-to-noise ratio of 5 : 1 is required to measure the wind fluctuations with sufficient precision for EC applications. Thus, standard deviations of $\leq 0.45\,\mathrm{m\,s^{-1}}$ and $0.20\,\mathrm{m\,s^{-1}}$ are reliably resolved in the horizontal and vertical wind components, respectively. For all 36 flights, the original and manipulated datasets were further propagated through the EC algorithm, also considering the precisions of the fast temperature and humidity measurement (Table 1). The result is an estimate of the least resolvable change in the measured flux ($\sigma_{u^*} = 0.003\,\mathrm{m\,s^{-1}}$, $\sigma_H = 0.9\,\mathrm{W\,m^{-2}}$, and $\sigma_E = 0.5\,\mathrm{W\,m^{-2}}$). Using the above signal-to noise-ratio of 5 : 1, changes in friction velocity, sensible- and latent heat of $0.02\,\mathrm{m\,s^{-1}}$, $5\,\mathrm{W\,m^{-2}}$, and $3\,\mathrm{W\,m^{-2}}$, respectively, are reliably resolved.

3.1.2 Measurement accuracy and maximum averaging interval

The above repeatability does not consider the environmental changes (temperature, humidity, pressure …) which are experienced by the sensors measuring aboard a moving aircraft. Changes in the environment likely lead to sensor drift, increasingly deteriorating the measurement with flight duration. In the following, we assess whether the measurement accuracy warrants the resolution of horizontal ABL structures up to the mesoscale (10–100 km). We start with the vertical wind measurement, because its signal levels and large-scale variability are low compared to the horizontal wind components or the scalars. We use the methods of Lenschow and Sun (2007), and first estimate the required signal level:

$$\frac{\partial w}{\partial t} < 0.2\sqrt{2}\,\sigma_w 2\pi k v_{\mathrm{tas}} = 1.4\times10^{-4}\,\mathrm{m\,s^{-2}} \qquad (1)$$

from the mesoscale variability of the vertical wind $\sigma_w = 0.1\,\mathrm{m\,s^{-1}}$, corresponding wavenumber $k = 2.8\times10^{-5}\,\mathrm{m^{-1}}$, and true airspeed $v_{\mathrm{tas}} = 28\,\mathrm{m\,s^{-1}}$. The required signal level is compared to the accuracy of the vertical wind measurement, using Eq. (5) from Lenschow and Sun (2007):

$$\frac{\partial w}{\partial t} \cong \underset{\mathrm{I}}{\Theta\frac{\partial v_{\mathrm{tas}}}{\partial t}} + \underset{\mathrm{II}}{v_{\mathrm{tas}}\frac{\partial \Theta}{\partial t}} + \underset{\mathrm{III}}{v_{\mathrm{tas}}\frac{\partial w_{\mathrm{AIR}}}{\partial t}} = 1.1\times10^{-4}\,\mathrm{m\,s^{-2}}, \qquad (2)$$

with $\Theta = \alpha - \theta$, the (radians) angles of attack α, and pitch θ, and the aircraft vertical velocity w_{AIR}. Here we apply the combined accuracies of the sensors and the wind model description, $\partial v_{\mathrm{tas}} = 0.34\,\mathrm{m\,s^{-1}}$, $\partial\Theta = 1.1\times10^{-2}$, and $\partial w_{\mathrm{AIR}} = 0.02\,\mathrm{m\,s^{-1}}$ (Metzger et al., 2011) over $\partial t = 1\,\mathrm{h}$ duration of a 100 km flight leg. With Θ rarely exceeding ±0.17 radians, terms I, II and III in Eq. (2) equate to $1.7\times10^{-5}\,\mathrm{m\,s^{-2}}$, $8.4\times10^{-5}\,\mathrm{m\,s^{-2}}$, and $0.6\times10^{-5}\,\mathrm{m\,s^{-2}}$, respectively. It can be seen that the overall performance is limited by the accuracy of Θ in the second term. This accuracy is dominated by the dynamic and differential pressure measurements used to infer α.

Analogously, the signal level required for the horizontal wind components ($1.8\times10^{-3}\,\mathrm{m\,s^{-2}}$) and their measurement accuracy ($\leq 1.9\times10^{-4}\,\mathrm{m\,s^{-2}}$) are calculated (Lenschow and Sun, 2007). Again, the dynamic and differential pressure measurements used to infer true airspeed and sideslip angle are the weakest link.

Accuracy in the scalar measurements along a flight leg is constrained by the drifts of the fast thermocouple ($7.2\times10^{-5}\,\mathrm{K\,s^{-1}}$) and the dew point mirror ($2.8\,\mathrm{ppm\,s^{-1}}$). Using the same 1 : 5 signal-to-noise criteria as in Eq. (1), temperature- and humidity fields differing > 1.3 K or $> 5\,\%$ mixing ratio, respectively, can be reliably distinguished throughout a 100 km flight leg.

3.2 Spectral properties of the aircraft

Various motions of the weight-shift microlight aircraft can potentially disturb the wind measurement. For comparison with the wind power spectra, we assess potential resonance from the engine or propeller of the WSMA, as well as the natural frequencies of trike and wing. Transverse and especially vertical to the WSMA body, accelerations of the five-hole probe agree well with measurements from the global positioning system/inertial measurement unit up to a frequency of 2–3 Hz (Fig. 1). At higher frequencies, the acceleration measurements at the 5HP continue to follow the pattern of GPS/IMU accelerations, though are slightly enhanced. This is expected, since the 5HP has a longer lever (≈ 0.5 m) with respect to the centre of rotation, i.e. the hang point of wing and trike. Consequently, the acceleration amplitudes are higher at the 5HP, which is accounted for in the lever arm correction of the wind measurement. The spectral peaks at 30 and 45 Hz are likely to be associated with harmonics from the engine and propeller, rotating at ≈ 100 Hz and ≈ 30 Hz, respectively. Because the -3 dB point (20 Hz) of the 5HP's low-pass filter is lower than the data acquisition's Nyquist frequency (50 Hz), aliasing of the wind measurement is however not a problem. For the acceleration component longitudinal to the WSMA, the 5HP pattern is enhanced compared to the GPS/IMU. This is surprising, since it is the axis of plug- and socket connection between GPS/IMU and 5HP, i.e. the axis with the least margin for resonance. We speculate that the reason for the observed difference lies in the fixture of the acceleration sensor in the 5HP, rather than in the mounting of the 5HP against the GPS/IMU. This could also partially explain the slightly enhanced energy in the 5HP transverse and vertical acceleration spectra. The spectral

behaviour of the wing acceleration measurements is different to those of the trike. It displays a distinct peak around 0.7 Hz, which is only present in the transverse component of the trike measurements. It can be understood as the wing's natural frequency, i.e. its inertia. In Sect. 3.3 the spectral characteristics of the wind measurement are related to these properties.

3.3 Comparison between tower and aircraft measurements

Here we compare average quantities, standard deviations (SD) as well as turbulent fluxes between tower and weight-shift microlight aircraft measurements. The objective is to reveal systematic differences between the platforms and identify their causes, such as instrument- or platform-related problems.

3.3.1 Statistical error

To unveil systematic differences between the platforms, we have to take into account the random statistical error of the measurements (Sect. 2.4.3). For this purpose, the integral length scales of scalars and fluxes were computed from WSMA measurements using Eq. (A1). The length scales for each flight were then used to calculate the average- and the ensemble random errors σ_{ran} and σ_{ens}, respectively, using Eqs. (A2)–(A5). In Table 2, the errors are summarized for each variable in the comparison between tower and WSMA. The average random errors are low for the measurement of averages ($<$ 1–15 %) and standard deviations (5–9 %), and higher for the friction velocity and the heat fluxes (25–34 %). Likewise, σ_{ens} increases from averages and SDs to fluxes. The comparison of the entire dataset between tower and WSMA is associated with an overall uncertainty of ≤ 3 % for the averages, ≤ 2 % for SDs and ≤ 8 % for the fluxes. In the following sections, these errors are used to derive the maximum likelihood functional relationships between tower and WSMA measurements.

3.3.2 Averages

The averages of vertical and along-wind components, temperature and absolute humidity were compared between tower and airborne measurements. The vertical wind measurements agree within several centimetres and show similar variability (Fig. 2). Taking into account the natural scatter in the data, the average vertical wind is not significantly different from zero for both platforms. This confirms our assumption that the aircraft vertical wind (in geodetic normal) is perpendicular to the spatial average of the streamlines. For the along-wind component, temperature and humidity measurements, the maximum likelihood relationship between tower and WSMA was calculated (Fig. 2). The error bars correspond to the random statistical error in the measurements, and the weight of each point in the relationship is represented by the size of the circles. Also shown are the slope $f(x)$, the residual standard error σ_{res} and the weighted coefficient of determination R^2 of the MLFR. Along-wind, temperature and humidity measurements agree well between the two platforms. The slopes are not significantly different from unity, and the MLFR explains ≥ 99 % of the variance in the dataset. In case of temperature and humidity, the 99.5 % confidence intervals are actually too narrow to be distinguished from the MLFR line. In order to assess the integrity of the MLFR, the results can be compared to the average- and the ensemble random errors σ_{ran} and σ_{ens}, respectively. The error in the residuals σ_{res} measures the random scatter not accounted for by the MLFR, and should be lower than the average random error σ_{ran}. In the presented data, σ_{res} is below σ_{ran} (Table 2), indicating that the significance of the MLFR exceeds the average random error. Likewise, we expect more confidence in the MLFR slope when the ensemble random error σ_{ens} is small. Here the errors in the slopes and σ_{ens} are equally small, emphasizing the close relationships.

Table 2. Results of the maximum likelihood functional relationships between tower and WSMA measurements. Shown are the MLFR slope and its standard error Slope $\pm\,\sigma$, weighted coefficient of determination R^2, residual standard error σ_{res}, the average random statistical error σ_{ran} and the ensemble random error σ_{ens}.

Variable	Slope $\pm\,\sigma$	R^2	σ_{res}	σ_{ran}	σ_{ens}
Averages					
u	0.99 ± 0.02	1.00	9 %	15 %	3 %
T	1.00 ± 0.00	1.00	0 %	0 %	0 %
a	0.99 ± 0.00	1.00	1 %	2 %	0 %
Standard deviations					
u	1.15 ± 0.05	0.99	21 %	9 %	2 %
v	1.02 ± 0.05	0.98	20 %	8 %	1 %
w	1.10 ± 0.03	0.99	9 %	5 %	1 %
T	1.34 ± 0.07	0.98	17 %	7 %	1 %
a	1.17 ± 0.08	0.98	23 %	9 %	1 %
Fluxes					
u_*	1.21 ± 0.07	0.98	13 %	25 %	5 %
H	1.17 ± 0.08	0.98	10 %	29 %	8 %
LE	1.17 ± 0.10	0.96	25 %	34 %	7 %

3.3.3 Spectra and standard deviations

Before investigating covariances, the spectra and standard deviations of the individual variables are compared between the tower and WSMA measurements. In order to assess the spectral quality, the 36 tower and WSMA data series used for the MLFRs were transferred into frequency domain using fast Fourier transformation. Each individual transform was normalized to a sum of unity. To reduce scatter, the normalized transforms of all tower and WSMA measurements, respectively, were then binned into frequency bands and ensemble-averaged. Figure 3 shows the average power spectra of vertical and horizontal wind, temperature and humidity for the tower and WSMA measurements. The power spectra

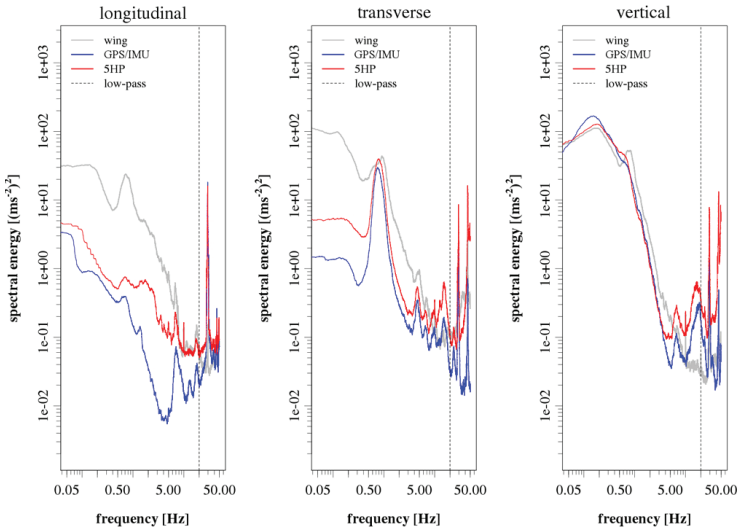


Fig. 1. Smoothed power spectra of acceleration measurements in the WSMA trike coordinate system. The dashed vertical line indicates the −3 dB frequency (20 Hz) of the Butterworth low-pass filter in the wind vector data acquisition system.

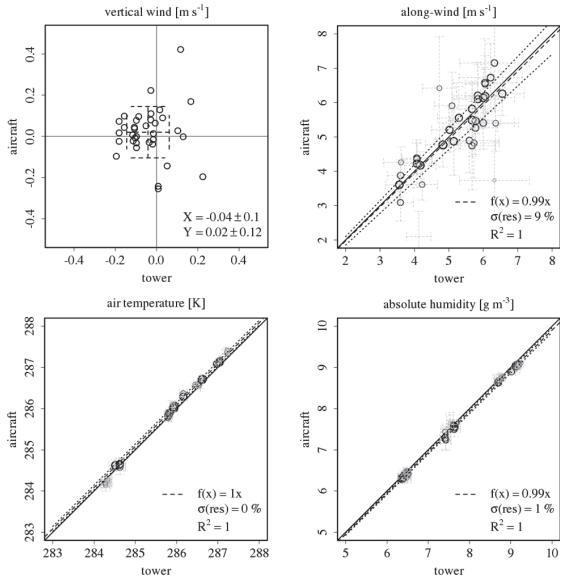


Fig. 2. Comparison between tower and WSMA averages of vertical- and along-wind components, temperature and absolute humidity. Average and standard deviation are given for the vertical wind. For along-wind, temperature and humidity, the solid, dashed and dotted lines are the 1 : 1 line, the maximum likelihood functional relationship, and the 99.5 % confidence interval, respectively. Additional information is given in the text.

are presented as function of observation frequency to enable the association with the physical properties of the WSMA. Due to fewer samples per dataset, the scatter in the WSMA spectra is higher at low frequencies. We use the streamwise component of the horizontal wind, i.e. in the direction of the mean wind and in the direction of the mean aircraft heading for tower and WSMA, respectively. The sonic temperature spectrum at the tower is compared to the air temperature spectrum at the WSMA. Kolmogorov (1941) defined the $f^{-5/3}$ law of isotropic turbulence for the inertial sub-range of atmospheric turbulence ($\approx$ 0.05–5 Hz). All variables, with exception of the WSMA vertical wind and the tower sonic temperature, follow the $f^{-5/3}$ law well (Fig. 3). The vertical wind and to a lesser degree also the streamwise wind component of the WSMA show a spectral peak between 0.4–2 Hz, coinciding with the wing's natural frequency (Sects. 3.2 and 4). Compared to the $f^{-5/3}$ law and relative to the entire frequency range, the standard deviations in the wind components are overestimated by 8 ± 3 % and 2 ± 1 %, respectively (median differences). In the WSMA transverse direction, the wind SD is overestimated by 5 ± 4 % at lower frequencies from 0.15–0.4 Hz (not shown). On the other hand, the power spectrum of the tower sonic temperature does not follow the $f^{-5/3}$ law very well. For frequencies above 0.3 Hz, the SD is overestimated by 29 ± 19 %. The pattern is consistent for both sonic anemometers at different measurement heights.

The SD of each individual measurement was corrected for spectral artifacts before case-by-case comparison between WSMA and tower. The MLFRs show that the WSMA measures 10 % and 15 % higher SDs in the vertical and along-wind components, respectively, compared to the tower

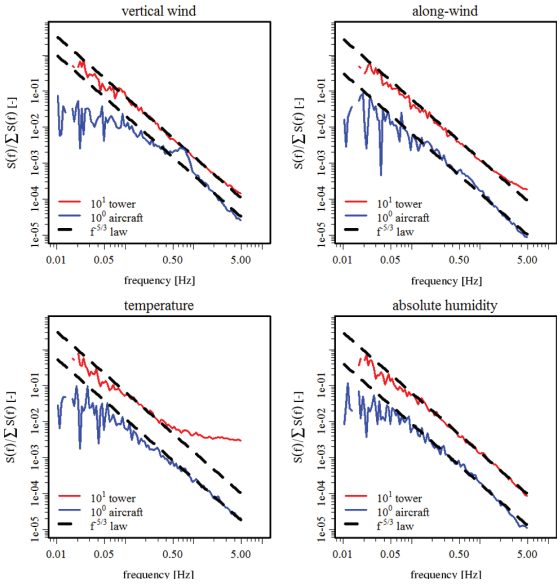


Fig. 3. Average power spectra of all measurements between tower and WSMA. To improve legibility, the tower data are offset by one order of magnitude. Also shown is the $f^{-5/3}$ law of isotropic turbulence (dashed line). Additional information is given in the text.

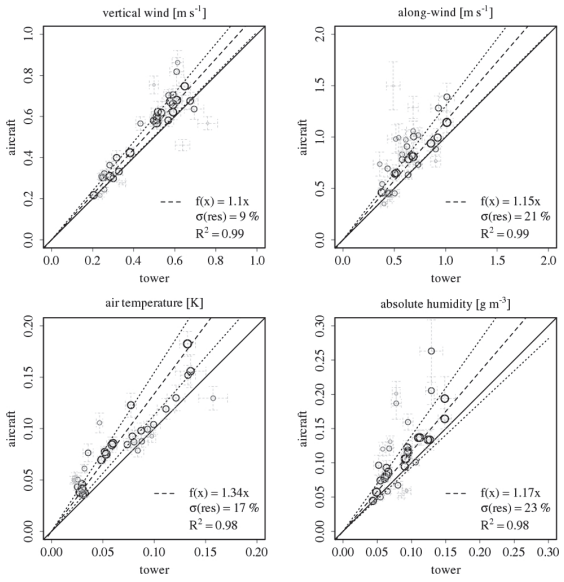


Fig. 4. Comparison between tower and WSMA standard deviations. The results are displayed in the same way as Fig. 2. Additional information is given in the text.

(Fig. 4). To a lesser degree, this behaviour is also found in the SD of the cross-wind component (2 %, not shown). Also for temperature (34 %) and humidity (17 %), the WSMA measures higher SDs compared to the tower. However, the difference is significant only for the temperature measurement, while the 99.5 % confidence intervals for all other measurements approximately include unity slope. All MLFRs explain ≥ 98 % of the variance in the data. The ensemble random errors are smaller than the slope errors, but the residual standard errors exceed their respective average random statistical error (Table 2). This can partially be attributed to very small denominators in the calculation of σ_{res}, but it is also a result of the natural scatter in the data. This indicates that σ_{ran} and σ_{ens} are potentially overoptimistic estimates of the spatial variability in the scalar fields.

3.3.4 Cospectra and fluxes

The correlation of horizontal and vertical wind components was compared between tower (-0.31 ± 0.14) and WSMA (-0.33 ± 0.13) measurements. Both values are close to one another and within the characteristic range from -0.15 to -0.35 (e.g. Foken et al., 2004). However, increased correlation due to spectral properties of the WSMA would result in systematically biased EC fluxes. In the following, we quantify to what extent the spectral properties of the WSMA contaminate the measured fluxes, and we correct the resulting bias. For this purpose, cospectra of the fluxes of momentum, sensible heat and latent heat were calculated in analogy to the power spectra in Sect. 3.3.3. In Fig. 5, ensemble cospectra are presented as function of the normalized frequency $n = f \cdot z/\bar{U}$, with $\bar{U}$ being the horizontal wind speed for the tower and the true airspeed for the WSMA, respectively. This facilitates the summarization of measurements at different heights above ground and the comparison between different platforms (Desjardins et al., 1989). Also shown is the reference cospectrum of Massman and Clement (2004), with the spectral maximum at $n = 0.1$ for unstable stratification (Kaimal and Finnigan, 1994; 24 out of 36 flights). The cospectra approximately follow the reference cospectrum, with exception of the momentum flux at the tower (not shown). The latter exhibits large scatter in the individual cospectra at both installation heights, which was not reduced by increasing the length of the dataset up to 30 min. Consequently, no ensemble cospectrum was calculated for the momentum flux at the tower. For the heat fluxes, the peak of the tower cospectra coincide with $n = 0.1$ of the reference cospectrum. The cospectral peaks of the WSMA measurements are marginally shifted towards higher frequencies around $n = 0.2$. A slight bimodality is attributable to combining measurements under slightly varying stratification, and is not present in the cospectra of single flights (peaking between $0.05 < n < 0.25$). Nevertheless, the shape of the reference cospectrum is generally better resembled by the WSMA measurements. In Sect. 3.3.3, we related increased variance in the WSMA wind components to spectral artifacts originating from the wing. In order to quantify the impact on

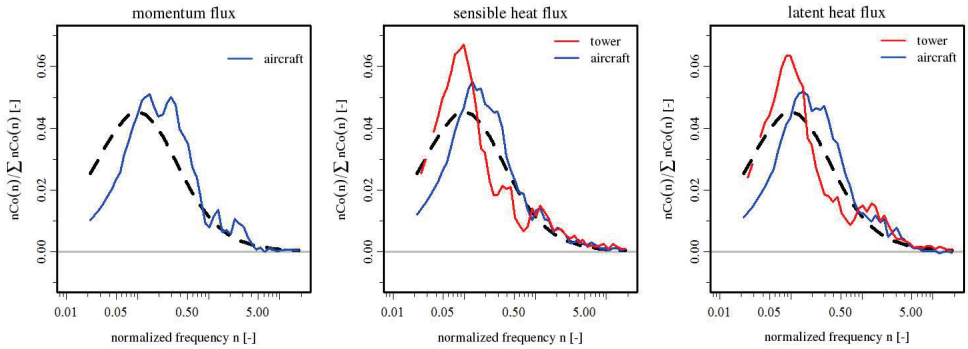


Fig. 5. Average cospectra of all measurements between tower and WSMA. Also shown is the reference cospectrum of Massman and Clement (2004, dashed line). Additional information is given in the text.

the WSMA flux measurement, all individual cospectra are compared between the WSMA and the reference cospectrum in the frequency range $f = 0.4$–2 Hz of the spectral artifacts. To account for the influence of stratification, the peaks of the reference cospectra were calculated using the forms of Kaimal et al. (1972). Relative to the entire frequency range, the spectral artifacts lead to a systematic deviation in the fluxes of momentum, sensible- and latent heat of 3 ± 6 %, -1 ± 6 % and 1 ± 3 % (median differences), respectively. Similarly, the inadequate frequency response of the tower sonic temperature measurements results in an underestimation of the sensible heat flux of -3 ± 5 % for frequencies > 0.3 Hz.

The covariances of the individual measurements were corrected for spectral artifacts before continuing the comparison. Finally, fluxes of momentum, sensible and latent heat are compared between EC measurements from tower and WSMA using the maximum likelihood functional relationship (Fig. 6). The WSMA estimates are 17–21 % higher compared to the tower, and the 99.5 % confidence intervals enclose unity slope. Alike the SDs, the ensemble random errors of the heat fluxes are less than or equal to the slope errors (Table 2). This indicates a sufficient sample size supporting the MLFR results. The explained variance is high (≥ 96 %) for all observed fluxes. All residual standard errors are ≤ 25 %, and significantly lower than the respective average random statistical error (Table 2).

Our analysis revealed that differences in the measurements between tower and weight-shift microlight aircraft partially originate from spectral artifacts. After factoring in these effects, the SDs at the WSMA differ from the tower by 2–15 % for the wind components, 34 % for temperature and 17 % for the humidity measurement. Likewise, the EC fluxes from WSMA remain in excess of 17–21 % compared to the tower.

3.4 Spatial analysis

In the following, we investigate whether the remaining differences between WSMA and tower measurements can be related to their spatial representativeness. For this purpose, we use a cross-wind distributed footprint parameterization (Appendix B) together with the Corine 2006 Land Cover raster. The contributions from all raster cells are cumulated with distance from the measurements, resulting in footprint effect level rings (Fig. 7). When considering the average footprint contributions over all tower-WSMA comparison measurements, the spatial context of the platforms appears to agree quite well. For both platforms, most of the footprint covers arable land (95–97 %). Contributions from the remaining land covers are sub-percent except for forest (2–3 %), and meadows do not contribute at all.

Taking a closer look at the individual, simultaneous measurements, the source areas can however differ considerably (Fig. 7). The degree of overlap is less related to flight altitude than to flight direction. The footprints often agree better for flights in along-wind direction (Fig. 7a, c). Yet this is not a general rule, because the flight paths are horizontally displaced from the tower (Fig. 7b). For flights in cross-wind direction, the footprint fans out, which additionally reduces the overlap (Fig. 7d). Over all measurements, the overlap ranges from 12–68 % of the footprint weights, with a median of 35 ± 17 %. However, varying overlap did not systematically alter the differences in the flux measurements between tower and WSMA ($R^2 \leq 0.07$).

3.5 Comparison with large-aperture scintillometer

In order to assess principal differences of spatial averaging (large-aperture scintillometer, weight-shift microlight aircraft) and temporal averaging (tower eddy covariance), we intercompare simultaneous measurements of the sensible heat flux. In Fig. 8 (abscissa), the fluxes from all platforms are normalized by the reference sensible heat flux derived

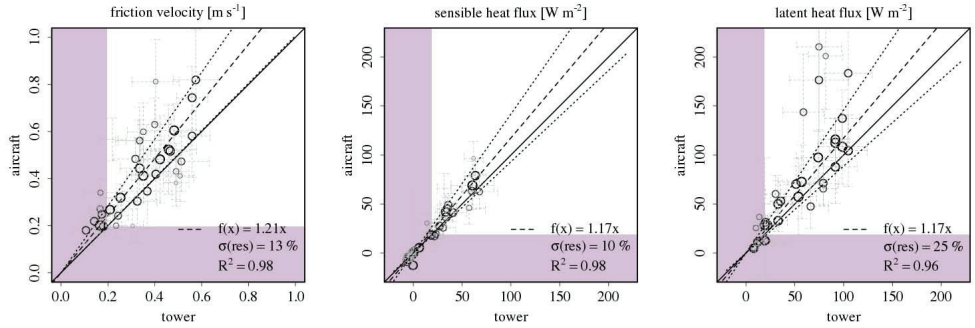


Fig. 6. Comparison between tower and WSMA eddy-covariance fluxes: friction velocity, sensible heat and latent heat. The results are displayed in the same way as Fig. 2. The data points in the shaded areas close to zero are omitted in the calculation of the normalized random errors (see Appendix A). Additional information is given in the text.

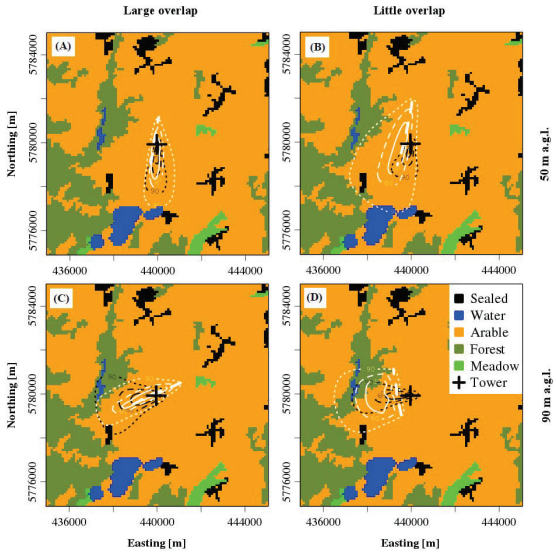


Fig. 7. Footprint effect levels of four simultaneous tower and WSMA measurements against the background of the Corine Land Cover raster (10 × 10 km). The flight path (white dashed line) and the footprint effect levels are shown. Solid, dashed and dotted lines represent a cumulative footprint of 30 %, 60 % and 90 %, respectively. Flights **(A)** and **(B)** on 15 October 2008, 14:51–14:52 UTC and 14:09–14:12 UTC at 50 m a.g.l. Flights **(C)** and **(D)** on 16 October 2008, 11:36–11:38 UTC and 11:09–11:10 UTC at 90 m a.g.l.

from the LAS, 55 W m^{-2} and 45 W m^{-2} on 20 and 21 October 2008, respectively. Similarly, the measuring heights (Fig. 8, ordinate) are normalized by the boundary layer depth, 450 m and 600 m on 20 and 21 October 2008, respectively. Under conditions of forced convection, H can be assumed to diminish below the entrainment zone around 0.8 ABL depth (Deardorff, 1974; Sorbjan, 2006). This allows us to approximate the vertical gradient of H, for which the measurements of LAS and 2.4 m EC are used as surface flux reference. The tower measurements at 50 m and 90 m approximately follow the vertical flux gradient derived from the EC surface flux measurements. However, it is evident that any H measured by tower EC and extrapolated to flight altitude is lower by 25–40 % compared to the LAS. At the same time, H measured by the WSMA is $\leq$ 25 % lower compared to the LAS, but 15–25 % higher compared to the tower measurements.

Differing source areas of the measurements (Fig. 9) qualify as a potential reason for the observed differences in the sensible heat flux. In absence of a better estimate, the LAS footprint was derived from the turbulence statistics at the 50 m tower EC measurement, and weighted along the LAS path (Meijninger et al., 2002; Wang et al., 1978). While southerlies prevailed on 20 October 2008 (Fig. 9a), southwesterly winds were observed on 21 October 2008 (Fig. 9b). Despite weaker winds on 20 October 2008, the footprint extent of the tower measurements is longer compared to 21 October 2008. This can be explained by different surface roughness upwind of the measurements and corresponding differences in the mechanical generation of turbulence ($0.2 < u_* < 0.5$). The roughness lengths are $\approx 10^{-4}$ m (water upwind) and $\approx 10^{-2}$ m (forest upwind) on 20 and 21 October 2008, respectively.

The overlap of the source areas of the tower and WSMA measurements with the LAS measurement was < 10 % on 20 October 2008. On 21 October 2008, however, the overlap of WSMA and LAS increased to 37 %, while remaining < 10 % between tower and LAS measurements. Despite the increasing overlap, WSMA and LAS fluxes did not agree as well as on the previous day. The footprints of all measurements were dominated by contributions from arable land, which generally were > 90 %. Only the 90 m tower EC measurement on 20 October 2008 had significant contributions from water bodies (13 %) and forest (6 %). Nevertheless, this measurement encloses the tower vertical flux gradient well within its random error (Fig. 8).

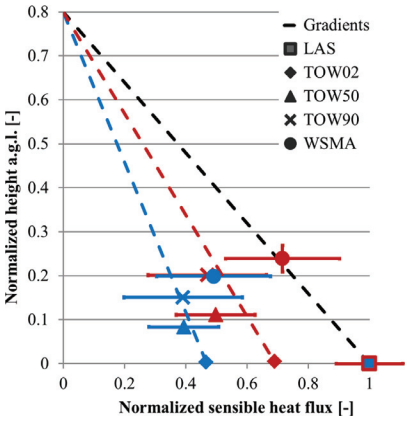


Fig. 8. Vertical profile of the sensible heat flux during two flights: 20 October 2008, 10:53–10:55 UTC (red) and 21 October 2008, 09:56–09:59 UTC (blue). Heat fluxes are compared between large-aperture scintillometer (LAS), tower eddy covariance at 2.4 m, 50 m and 90 m (TOW02, TOW50, TOW90) and the weight-shift microlight aircraft (WSMA). Where available, error bars show the random statistical error and the altitude standard deviation. Additional information is given in the text.

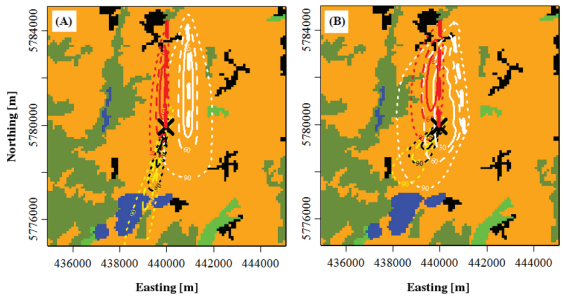


Fig. 9. Footprint effect levels of the measurements from Fig. 8 on 20 October 2008 **(A)** and 21 October 2008 **(B)**, presented similarly to Fig. 7. In addition, the weighted footprint along the LAS path is shown (red), and the footprints from the tower measurements at 50 m (black) and 90 m (yellow) are distinguished.

4 Discussion

The propagation of sensor errors enables defining the minimum change in an atmospheric quantity that can be reliably resolved by the WSMA measurements of respective variables. For the wind measurement, this coincides with the lower margin of the standard deviations observed at both tower and WSMA (Fig. 4). We thus conclude that the precision of the wind measurement warrants eddy-covariance flux measurements under unstable to slightly stable stratifications. The precision of the vertical wind is better by a factor of two compared to the horizontal wind components. This can be traced back to the better precision of the attack- and pitch ($\leq 0.08°$) angles compared to the sideslip- and heading ($\leq 0.18°$) angles. From the assessment of sensor accuracies, we found that the wind and scalar measurements facilitate the signal levels required for the resolution of mesoscale ABL structures. This enables extending averaging intervals and spectral analyses up to a scale of tens of kilometres.

The focus of this study is on the comparison of turbulence statistics between weight-shift microlight aircraft and tower measurements. A potential source of uncertainty is the flow distortion correction of the USA-1 sonic anemometers used at the tall tower. Two different corrections are provided by the manufacturer, of which the "milder" 2-D version was used in the operational setup of the DWD. A post processing comparison for the presented tower data shows that the 3-D version would lead to a systematic increase in the wind SDs by $\approx 35\,\%$, and in the fluxes by $\approx 15\,\%$. From instrument comparison, the 2-D correction seems to be more appropriate. However, it must be noted that the magnitude of this correction alone is in the order of the differences in the wind SDs and the fluxes observed between the tower and the WSMA. Longer averaging intervals at the tower or detrending of the WSMA data did not change the general behaviour, but increased the scatter in the comparison. Moreover, vertical flux divergence can be ruled out as potential error source, since the measurements were conducted at approximately the same altitude above ground. Also, altitude fluctuations by the WSMA were accounted for by using potential quantities of temperature and densities at the tower pressure level. In the following, we consequently focus on the effects of spectral artifacts and surface heterogeneity.

During the investigation of the WSMA spectral properties, a scale discrepancy was detected between accelerations acting at the five-hole probe and their accounting in the global positioning system/inertial measurement unit, i.e. the wind computation (Fig. 1). No remnants of this scale discrepancy are evident in the wind measurements, in particular between 1–5 Hz. This leads to the conclusions that (i) the wind vector computation correctly accounts for the displacement of 5HP and GPS/IMU, and (ii) the cause of enhanced 5HP acceleration measurements (especially longitudinal to the body) lies in the fixture of the acceleration sensor in the 5HP, rather than in the mounting of the 5HP against the GPS/IMU. The spectral peak in the WSMA vertical and streamwise wind components (Fig. 3) coincides with the wing's natural frequency around 0.7 Hz (Sect. 3.2). A less pronounced peak between 0.15–0.4 Hz in the transverse wind component coincides only with a peak in vertical accelerations (Fig. 1). Both spectral features can potentially be associated with the treatment of wing upwash in the time domain, but not in the frequency domain (Metzger et al., 2011). The WSMA wind and flux measurements were corrected for this spectral inconsistency before comparing them to ground-based measurements. The appropriate correction factors were estimated

from the comparison of measured spectra and cospectra to modelled ones.

Because of large scatter in the momentum flux cospectra at the tower, no ensemble was calculated. We speculate that the scatter originates from the wind direction-dependent correlation of the horizontal and vertical wind components at the USA-1 sonic anemometers (e.g. Mauder et al., 2007b). The erroneous sonic temperature spectrum indicates problems with the measurement of the temperature SD and the sensible heat flux at the tower. The amplitude resolution test by Vickers and Mahrt (1997) would reject 28 out of 36 tower sonic temperature data sets. The problem was related to the insufficient sonic temperature resolution (0.01 K) of the USA-1, which appears as superficial spectral energy in the form of high-frequency white noise. Spectral correction factors analogous to the WSMA measurements were used in an attempt to correct the systematic overestimation of the temperature SD. Such a procedure changes the maximum likelihood functional relationship between tower and aircraft temperature SD from −9 % underestimation to 34 % overestimation of the WSMA measurement. At the same time, the residual standard error in the MLFR increases from 9 % to 17 %. The scattering can be suppressed by rejecting data points with weak temperature SD (< 0.05 K), resulting in a reduced overestimation of 15 % of the WSMA measurement. Consequently, the USA-1 measurements cannot be regarded as reliable reference for the temperature SD, and the spectral correction factors must be interpreted with caution. The problem is less pronounced for the sensible heat flux. The white noise in the USA-1 sonic temperature measurement does not affect the measurement of, or the correlation with, the vertical wind measurement. The result is a modest underestimation of −3 % of the tower sensible heat flux due to reduced coherence of sonic temperature and vertical wind at high frequencies. The humidity measurements agree well between the platforms in the time and in the frequency domain. Several outliers in the WSMA latent heat flux measurement coincide with WSMA flights west of the tall tower, which are closer to a forest edge. Increased mechanical turbulence downwind of the forest edge is a potential explanation for increased turbulent fluxes. This finding however does not hold for the sensible heat flux and the friction velocity, which is in contradiction to scalar similarity.

As a potential source for the observed differences, we assess the spatial context of the tower and aircraft measurements. The footprint results illustrate that the land cover contributions are very similar for both platforms. Provided the land cover data are a suitable proxy for the land-atmosphere exchange, differences between tower and WSMA measurements cannot be attributed to different land cover contributions alone. However, the footprint analysis also reveals that the source areas only share a fractional overlap. Consequently, the observed differences can potentially originate from the platforms' principally different sampling strategies. Foken (2008) and Mahrt (2010) conclude that the energy balance non-closure frequently observed from tower EC measurements is connected to the interaction of terrain heterogeneity and turbulent scales. For this purpose, the transfer of heat between surface and atmosphere is considered separately for smaller, random eddies and for larger, non-propagating eddies. Thereof, the transfer by the small, random eddies is measured by the tower EC. However, an additional transfer component is suspected to occur at significant surface heterogeneities, leading to the generation of non-uniformly distributed NPEs. Based on intensive measurement campaigns and modelling efforts, the presence of NPEs in the study area has been shown (Uhlenbrock et al., 2004). At 100 m measuring height, the resulting average horizontal imbalance of the sensible heat flux is in the order of 4–19 % (Steinfeld et al., 2007). In the same study area, Beyrich et al. (2006) found comparable differences between surface flux estimates from aircraft and tower of 11 % for the sensible heat and 23 % for the latent heat. Furthermore, a tendency was shown to close the energy balance on the regional scale with spatially averaging methods (Mauder et al., 2007a). This suggests that the tower EC cannot adequately capture all flux contributions due to its inability of spatial sampling. On the other hand, a large-aperture scintillometer captures NPEs up to the dimension of its path length, with increasing sensitivity towards the centre of the path (Foken et al., 2010). Also airborne EC is capable of spatial sampling and captures some of the associated flux, depending on the horizontal extent of the NPEs and the flight path. Therefore, the presence of NPEs in the study area can be considered a potential explanation for the deviation of the tower EC results from the WSMA, and even more so from the LAS results. In order to further adjust results from tower and airborne EC measurements, the raw data could be high-pass filtered, which restricts low-frequency flux contribution to an identical threshold (e.g. Thomas et al., 2012).

This study combines several principles and methods for the purpose of quantifying: (i) the suitability of airborne instrumentation for EC measurements, (ii) the complex feedback of WSMA motions on the EC measurement, and (iii) the direct intercomparison of measurement platforms with differing spatial representativeness. The applied techniques are not restricted to use with the WSMA, but are general enough to be used for the development and assessment of other airborne platforms.

5 Conclusions

We have shown that turbulence measurements from a weight-shift microlight aircraft can be achieved with sufficient precision to enable eddy-covariance flux calculation. Furthermore, a coordinated setup of tall tower, large-aperture scintillometer and weight-shift microlight aircraft measurements avoids typical errors due to averaging intervals and vertical flux divergence (e.g. Betts et al., 1990). Differences

on the order of 15–25 % remain between the fluxes measured by the ground-based instruments and the WSMA. At that, the LAS generally measured the highest flux magnitude, followed by the WSMA and the tower. However, the 99.5 % confidence intervals of the maximum likelihood functional relationships between tower and WSMA include unity slope. Consequently, the observed differences can be considered insignificant, and the accuracy of the WSMA flux measurement is quantified to $\leq 10\,\%$ ($1\,\sigma$ slope error).

Nevertheless, several potential reasons for the disagreement of the results between the measurement platforms are investigated. (i) The WSMA wind and flux measurement is subject to spectral artifacts originating from its wing, and the results are corrected prior to the comparison. (ii) The flow distortion correction and the temperature resolution at the tower sonic anemometer measurement alone can explain the full magnitude of the disagreement. (iii) A footprint analysis allows excluding differences in the surface areas as the primary reason for the disagreement. (iv) Principal differences between spatially and temporally averaging flux measurements may also explain the observed differences. In particular, energy transfer by non-propagating eddies above heterogeneous terrain is discussed as a potential reason.

We conclude that the WSMA is a suitable tool to promote the on-going research of surface-atmosphere interactions in heterogeneous landscapes. The flux measurement is sufficiently accurate to cover the required flight transect length (10–100 km). Moreover, the WSMA's low ratio of true airspeed to climb rate is well suited for terrain-following flight over complex terrain. Using e.g. wavelet analysis, the regional turbulent exchange measured from the aircraft can be located in space and in spectral scale (Mauder et al., 2007a; Strunin and Hiyama, 2004). All of the above features are beneficial for the study of yet poorly understood exchange mechanisms between the Earth's surface and the atmosphere. This further substantiates the versatility of the WSMA as a low cost and widely applicable environmental research aircraft.

Appendix A

Integral length scales and statistical error

The integral length scale λ can be interpreted as the typical size of the most energy-transporting eddies. It is calculated by integration of the autocorrelation function from zero lag to the first crossing with zero at lag r_0 (Bange et al., 2002; Lenschow and Stankov, 1986):

$$\lambda = \int_0^{r_0} \frac{\langle f'(x)\, f'(x+r)\rangle}{\langle f'(x)^2\rangle} \mathrm{d}r \,. \tag{A1}$$

Here f represents a turbulent quantity c (scalars or wind components), but also combinations of these with the vertical wind $f(x) = w'(x) \cdot c'(x)$ (Bange, 2007). Hence, the integral scales of the turbulent fluxes were directly calculated from the data of the weight-shift microlight aircraft. The transformation into the integral time scale τ of simultaneous tower measurements is carried out by division of λ with the mean horizontal wind speed at the tower. This assumes that Taylor's hypothesis of frozen turbulence is valid (Taylor, 1915). The random statistical error of the sample average $\bar{f}$ is simply the square root of its variance $\sqrt{\overline{(f')^2}}$. However, the errors in the variance V of f, or the covariance F of its combinations, are functions of the integral scales λ, τ. The random statistical errors σ_V and σ_F were determined after Lenschow and Stankov (1986); Lenschow et al. (1994):

$$\sigma_V^2 = 2 \cdot V^2 \cdot \frac{\lambda}{L} \,, \tag{A2}$$

$$\sigma_F^2 = 2 \cdot F^2 \cdot \frac{\lambda}{L} \cdot \frac{1 + r_{wc}^2}{r_{wc}^2} \,, \tag{A3}$$

with the averaging length L and the correlation coefficient between vertical wind and the turbulent quantity r_{wc}. It is assumed that $\lambda \ll L$, and that w and c are Gaussian distributed. The momentum flux consists of two orthogonal components, $\overline{u'w'}$ and $\overline{v'w'}$, with the wind components u, v and w. The calculation of its random error is obtained from Gaussian error reproduction of the errors in its components (Bange et al., 2002). The individual random errors of the WSMA and the tower measurement, σ_{air} and σ_{tow}, respectively, were summarized for each variable over all flight legs:

$$\sigma_{\mathrm{ran}} = \frac{\sum_{\mathrm{air}}^{n} \sigma_{\mathrm{air}} + \sum_{\mathrm{tow}}^{n} \sigma_{\mathrm{tow}}}{2n - 1} \,, \tag{A4}$$

resulting in the average random error in the data couples σ_{ran} with the sample size n. The ensemble random error σ_{ens} considers the reduction of the random error with the sample size (Mahrt, 1998):

$$\sigma_{\mathrm{ens}} = \frac{\sigma_{\mathrm{ran}}}{\sqrt{n}} \,, \tag{A5}$$

with zero expected value σ_{ens} and the standard deviation σ_{ran} of the population. While σ_{ran} is a measure for the average dispersion of the data couples, σ_{ens} quantifies the level of confidence we can expect from comparing the entire dataset between the two platforms. To use Eqs. (A4)–(A5) for data obtained during different flight days, we use normalized error estimates. Yet, the normalized errors are excessively large when the denominator, i.e. the measurement quantity, approaches zero. For turbulent fluxes, this is usually the case under stable conditions, where e.g. intermittent turbulence can violate the assumptions in the integral length scales. Consequently, we constrain the calculation of σ_{ran} and σ_{ens} for the fluxes to values of $u_* > 0.2\,\mathrm{m\,s^{-1}}$ and H, $E > 20\,\mathrm{W\,m^{-2}}$, resulting in $N = 28$, 15 and 24 samples, respectively.

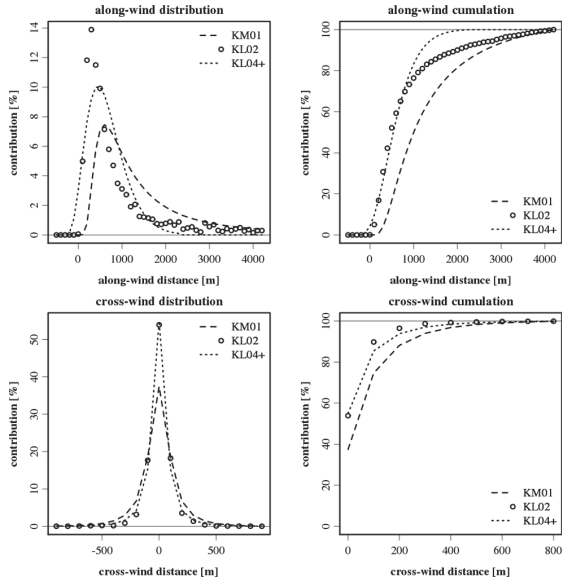


Fig. B1. Cross-wind and along-wind integrated distributions of the footprints for case 2, $z = 100$ m from Markkanen et al. (2009). Upper and lower panels display longitudinal and cross-sections, respectively. The footprint weight distributions are shown on the left side, and the cumulative distributions are shown on the right side.

Appendix B

Footprint modelling

The footprint- or source weight function quantifies the spatial contributions to each measurement (Schmid, 2002; Vesala et al., 2008). Analytical footprint models are often limited, e.g. regarding stability regimes or measurement heights (e.g. Kormann and Meixner, 2001, subsequently referred to with KM01). Lagrangian footprint models overcome these limitations and additionally consider 3-D dispersion, but are computationally expensive. The footprint model of Kljun et al. (2004, KL04) is a parameterization of the backward Lagrangian model of Kljun et al. (2002, KL02) in the range $-200 \leq z/L \leq 1$, $u_* \geq 0.2\,\mathrm{m\,s^{-1}}$, and $1\,\mathrm{m} \leq z \leq z_i$, with the boundary layer depth z_i. Thus, it combines little computational effort with broad applicability. The parameterization depends upon friction velocity u_*, measurement height z, standard deviation of the vertical wind σ_w and the aerodynamic roughness length z_0, of which u_*, z and σ_w are measured directly. The roughness length is inferred using the logarithmic wind profile with the integrated universal function for momentum exchange after Businger et al. (1971) in the form of Högström (1988). The KL04 is a cross-wind integrated footprint model, i.e. it does not resolve the distribution perpendicular to the main wind direction. In order to account for cross-wind dispersion, the KL04 was combined with a Gaussian cross-wind distribution function (Kljun et al., 2012). The combination of KL04 with this cross-wind distribution results in a computationally fast footprint parameterization which considers 3-D dispersion and is not constrained to applications in the surface layer. In the following, we refer to this model as KL04+.

Table B1. Median performance of the footprint parameterizations KL04+ and KM01 compared to the reference Lagrangian model of Kljun et al. (2002). Uncertainty measures NMSE, MAD and r are explained in the text.

Direction	Along-wind		Cross-wind	
Model	KL04+	KM01	KL04+	KM01
NMSE	0.34 ± 0.05	0.95 ± 0.48	0.34 ± 0.05	0.33 ± 0.36
MAD [%]	0.43 ± 0.10	0.63 ± 0.28	0.08 ± 0.05	0.12 ± 0.14
r	0.91 ± 0.00	0.66 ± 0.12	0.99 ± 0.01	0.99 ± 0.01

To evaluate model performance, KM01 and KL04+ were compared to the reference Lagrangian model KL02. For this purpose, four existing realizations of the KL02 model were used from Markkanen et al. (2009, Table 1, case 1 ($L = -32$ m, $u_* = 0.27\,\mathrm{m\,s^{-1}}$) and case 2 ($L = -76.6$ m, $u_* = 0.295\,\mathrm{m\,s^{-1}}$), $z = 50$ m and 100 m). Markkanen et al. (2009) adopted case 1 from Leclerc et al. (1997), and referenced the results to large eddy simulations (Raasch and Schröter, 2001). The above parameter sets do not include σ_v and σ_w, which were derived using the integral turbulence characteristics proposed by Lumley and Panofsky (1964) and Panofsky et al. (1977), respectively. The computed footprint weights were summarized for each cell of a grid with 100 m horizontal spacing and subsequently integrated over cross-wind and along-wind direction, respectively (Fig. B1). In both directions, KL04+ assigns more weight to the close range compared to KM01 (Fig. B1 left panels). This is consistent throughout the range of the tested atmospheric conditions (not shown). In along-wind direction, KL04+ reproduces KL02 very well until the cumulative distribution accounts for approximately 80 % of the footprint (Fig. B1 upper right panel). Contributions from below the measurement location due to along-wind dispersion are slightly pronounced by KL04+, and neglected by KM01. The cross-wind distributions of both KL04+ and KM01 agree reasonably well with KL02.

To quantify the model comparison, normalized mean square error (NMSE, Hanna and Paine, 1989), median absolute deviation (MAD, Rousseeuw and Verboven, 2002) and Pearson's coefficient of correlation (r) are used. The NMSE is based on variance statistics and is thus sensitive to the few largest deviations in the dataset. In contrast, the MAD is the middle value of the error distribution, and is more sensitive to the error frequency. We use NMSE and MAD to assess the model performance around the peak and the tail of the footprint, respectively. The correlation coefficient provides information on the degree of similarity between the models' distributions. Table B1 summarizes the results of the model

comparison. In both along-wind and cross-wind directions, the similarity between KL04+ and KL02 is as good or better than between KM01 and KL02 (r). In along-wind direction, KL04+ is considerably closer to KL02 in the close range (NMSE) and the far range (MAD) compared to KM01. A similar deviation between KM01 and KL02 has been reported by Kljun et al. (2003). In cross-wind direction, both KL04+ and KM01 agree equally well with KL02. For all metrics, the standard errors over the four parameter sets are smaller for the KL04+ model, pointing out its reliability.

Acknowledgements. Recognition has to be given to Frank Neidl and Rainer Steinbrecher at the Karlsruhe Institute of Technology, Institute for Meteorology and Climate Research. Neidl programmed and continuously maintained the data acquisition system, while Steinbrecher initiated wind- and flux measurements with the weight-shift microlight aircraft in the first place. Thanks go to Ulrich Weisensee and Jens-Peter Leps at the German Meteorological Service, Richard-Aßmann-Observatory Lindenberg for performing the tower turbulence measurements. We are obliged to Xunhua Zheng and her workgroup at the Chinese Academy of Sciences, Institute of Atmospheric Physics, which was hosting our project and providing indispensable infrastructure. Our thanks to Jens Bange at the Eberhard Karls University of Tübingen, Institute for Geoscience, for his advice regarding the flight campaign design. The authors wish to thank Henry Loescher at the National Ecological Observatory Network, Fundamental Instrument Unit for his continued support. Stipend funding by the German Academic Exchange Service, Helmholtz Association of German Research Centres, China Scholarship Council and the European Union under the Science and Technology Fellowship China is acknowledged. The National Ecological Observatory Network is a project sponsored by the National Science Foundation and managed under cooperative agreement by NEON, Inc. This material is based upon work supported by the National Science Foundation under the grant DBI-0752017. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Edited by: S. Malinowski

References

Bange, J.: Airborne Measurement of Turbulent Energy Exchange Between the Earth Surface and the Atmosphere, Faculty of Engineering, Braunschweig University of Technology, Braunschweig, Germany, 2007.

Bange, J., Beyrich, F., and Engelbart, D. A. M.: Airborne measurements of turbulent fluxes during LITFASS-98: Comparison with ground measurements and remote sensing in a case study, Theor. Appl. Climatol., 73, 35–51, doi:10.1007/s00704-002-0692-6, 2002.

Betts, A. K., Desjardins, R. L., Macpherson, J. I., and Kelly, R. D.: Boundary-layer heat and moisture budgets from FIFE, Bound.-Lay. Meteorol., 50, 109–138, doi:10.1007/BF00120520, 1990.

Beyrich, F. and Adam, W. K.: Site and data report for the Lindenberg reference site in CEOP – phase I, Berichte des Deutschen Wetterdienstes, 55 pp., 2007.

Beyrich, F., Leps, J. P., Mauder, M., Bange, J., Foken, T., Huneke, S., Lohse, H., Ludi, A., Meijninger, W. M. L., Mironov, D., Weisensee, U., and Zittel, P.: Area-averaged surface fluxes over the LITFASS region based on eddy-covariance measurements, Bound.-Lay. Meteorol., 121, 33–65, doi:10.1007/s10546-006-9052-x, 2006.

Businger, J. A., Wyngaard, J. C., Izumi, Y., and Bradley, E. F.: Flux-profile relationships in the atmospheric surface layer, J. Atmos. Sci., 28, 181–189, doi:10.1175/1520-0469(1971)028<0181:FPRITA>2.0.CO;2, 1971.

Chen, J., Jonsson, P., Tamura, M., Gu, Z. H., Matsushita, B., and Eklundh, L.: A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky-Golay filter, Remote Sens. Environ., 91, 332–344, doi:10.1016/j.rse.2004.03.014, 2004.

Davis, F. W., Schimel, D. S., Friedl, M. A., Michaelsen, J. C., Kittel, T. G. F., Dubayah, R., and Dozier, J.: Covariance of biophysical data with digital topographic and land use maps over the FIFE site, J. Geophys. Res.-Atmos., 97, 19009–19021, doi:10.1029/92jd01345, 1992.

Deardorff, J. W.: Three-dimensional numerical study of turbulence in an entraining mixed layer, Bound.-Lay. Meteorol., 7, 199–226, doi:10.1007/bf00227913, 1974.

Desjardins, R. L., MacPherson, J. I., Schuepp, P. H., and Karanja, F.: An evaluation of aircraft flux measurements of CO_2, water vapor and sensible heat, Bound.-Lay. Meteorol., 47, 55–69, doi:10.1007/BF00122322, 1989.

Desjardins, R. L., MacPherson, J. I., Mahrt, L., Schuepp, P., Pattey, E., Neumann, H., Baldocchi, D., Wofsy, S., Fitzjarrald, D., McCaughey, H., and Joiner, D. W.: Scaling up flux measurements for the boreal forest using aircraft-tower combinations, J. Geophys. Res.-Atmos., 102, 29125–29133, doi:10.1029/97JD00278, 1997.

Egger, J., Bajrachaya, S., Heinrich, R., Kolb, P., Lammlein, S., Mech, M., Reuder, J., Schaper, W., Shakya, P., Schween, J., and Wendt, H.: Diurnal winds in the Himalayan Kali Gandaki valley. Part III: Remotely piloted aircraft soundings, Month. Weather Rev., 130, 2042–2058, doi:10.1175/1520-0493(2002)130<2042:DWITHK>2.0.CO;2, 2002.

Foken, T.: The energy balance closure problem: an overview, Ecol. Appl., 18, 1351–1367, doi:10.1890/06-0922.1, 2008.

Foken, T., Göckede, M., Mauder, M., Mahrt, L., Amiro, B., and Munger, W.: Post-field Data Quality Conirol, in: Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis, 1st Edn., edited by: Lee, X., Law, B., and Massman, W., Springer Netherlands, Dordrecht, The Netherlands, 264 pp., 2004.

Foken, T., Mauder, M., Liebethal, C., Wimmer, F., Beyrich, F., Leps, J.-P., Raasch, S., DeBruin, H. A. R., Meijninger, W. M. L., and Bange, J.: Energy balance closure for the LITFASS-2003 experiment, Theor. Appl. Climatol., 101, 149–160, doi:10.1007/s00704-009-0216-8, 2010.

Foken, T., Leuning, R., Oncley, S. P., Mauder, M., and Aubinet, M.: Corrections and data quality, in: Eddy Covariance: A Practical Guide to Measurement and Data Analysis, edited by: Aubinet, M., Vesala, T., and Papale, D., Springer, Dordrecht, Heidel-

berg, London, New York, 85–131, 2012.

Garman, K. E., Hill, K. A., Wyss, P., Carlsen, M., Zimmerman, J. R., Stirm, B. H., Carney, T. Q., Santini, R., and Shepson, P. B.: An airborne and wind tunnel evaluation of a wind turbulence measurement system for aircraft-based flux measurements, J. Atmos. Ocean. Tech., 23, 1696–1708, doi:10.1175/JTECH1940.1, 2006.

Göckede, M., Foken, T., Aubinet, M., Aurela, M., Banza, J., Bernhofer, C., Bonnefond, J. M., Brunet, Y., Carrara, A., Clement, R., Dellwik, E., Elbers, J., Eugster, W., Fuhrer, J., Granier, A., Grünwald, T., Heinesch, B., Janssens, I. A., Knohl, A., Koeble, R., Laurila, T., Longdoz, B., Manca, G., Marek, M., Markkanen, T., Mateus, J., Matteucci, G., Mauder, M., Migliavacca, M., Minerbi, S., Moncrieff, J., Montagnani, L., Moors, E., Ourcival, J.-M., Papale, D., Pereira, J., Pilegaard, K., Pita, G., Rambal, S., Rebmann, C., Rodrigues, A., Rotenberg, E., Sanz, M. J., Sedlak, P., Seufert, G., Siebicke, L., Soussana, J. F., Valentini, R., Vesala, T., Verbeeck, H., and Yakir, D.: Quality control of CarboEurope flux data – Part 1: Coupling footprint analyses with flux data quality assessment to evaluate sites in forest ecosystems, Biogeosciences, 5, 433–450, doi:10.5194/bg-5-433-2008, 2008.

Hanna, S. R. and Paine, R. J.: Hybrid Plume Dispersion Model (HPDM) development and evaluation, J. Appl. Meteorol., 28, 206–224, doi:10.1175/1520-0450(1989)028<0206:HPDMDA>2.0.CO;2, 1989.

Hiyama, T., Strunin, M. A., Tanaka, H., and Ohta, T.: The development of local circulations around the Lena River and their effect on tower-observed energy imbalance, Hydrol. Process., 21, 2038–2048, doi:10.1002/hyp.6705, 2007.

Hobbs, S., Dyer, D., Courault, D., Olioso, A., Lagouarde, J. P., Kerr, Y., McAneney, J., and Bonnefond, J.: Surface layer profiles of air temperature and humidity measured from unmanned aircraft, Agronomie, 22, 635–640, doi:10.1051/agro:2002050, 2002.

Högström, U.: Non-dimensional wind and temperature profiles in the atmospheric surface layer: A re-evaluation, Bound.-Lay. Meteorol., 42, 55–78, doi:10.1007/BF00119875, 1988.

Hojstrup, J.: A statistical data screening procedure, Meas. Sci. Technol., 4, 153–157, doi:10.1088/0957-0233/4/2/003, 1993.

Isaac, P. R., Leuning, R., Hacker, J. M., Cleugh, H. A., Coppin, P. A., Denmead, O. T., and Raupach, M. R.: Estimation of regional evapotranspiration by combining aircraft and ground-based measurements, Bound.-Lay. Meteorol., 110, 69–98, doi:10.1023/A:1026054317990, 2004a.

Isaac, P. R., McAneney, J., Leuning, R., and Hacker, J. M.: Comparison of aircraft and ground-based flux measurements during OASIS95, Bound.-Lay. Meteorol., 110, 39–67, doi:10.1023/A:1026002301152, 2004b.

Junkermann, W.: An ultralight aircraft as platform for research in the lower troposphere: System performance and first results from radiation transfer studies in stratiform aerosol layers and broken cloud conditions, J. Atmos. Ocean. Tech., 18, 934–946, doi:10.1175/1520-0426(2001)018<0934:AUAAPF>2.0.CO;2, 2001.

Junkermann, W.: The actinic UV-radiation budget during the ESCOMPTE campaign 2001: results of airborne measurements with the microlight research aircraft D-MIFU, Atmos. Res., 74, 461–475, doi:10.1016/j.atmosres.2004.06.009, 2005.

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

Kaimal, J. C., Wyngaard, J. C., Izumi, Y., and Coté, O. R.: Spectral characteristics of surface-layer turbulence, Q. J. Roy. Meteorol. Soc., 98, 563–589, doi:10.1002/qj.49709841707, 1972.

Kanda, M., Inagaki, A., Letzel, M. O., Raasch, S., and Watanabe, T.: LES study of the energy imbalance problem with Eddy covariance fluxes, Bound.-Lay. Meteorol., 110, 381–404, doi:10.1023/B:BOUN.0000007225.45548.7a, 2004.

Kljun, N., Rotach, M. W., and Schmid, H. P.: A three-dimensional backward lagrangian footprint model for a wide range of boundary-layer stratifications, Bound.-Lay. Meteorol., 103, 205–226, doi:10.1023/A:1014556300021, 2002.

Kljun, N., Kormann, R., Rotach, M. W., and Meixer, F. X.: Comparison of the Langrangian footprint model LPDM-B with an analytical footprint model, Bound.-Lay. Meteorol., 106, 349–355, doi:10.1023/A:1021141223386, 2003.

Kljun, N., Calanca, P., Rotach, M. W., and Schmid, H. P.: A simple parameterisation for flux footprint predictions, Bound.-Lay. Meteorol., 112, 503–523, doi:10.1023/B:BOUN.0000030653.71031.96, 2004.

Kljun, N., Rotach, M. W., and Schmid, H. P.: A Simple Parameterisation for Two-dimensional Flux Footprints, in preparation, 2012.

Kolmogorov, A. N.: Lokalnaja struktura turbulentnosti v neschtschimaemoi schidkosti pri otschen bolschich tschislach Reynoldsa (The local structure of turbulence in incompressible viscous fluid for very large Reynolds numbers), Proceedings of the USSR Academy of Sciences, 30, 299–303, 1941.

Kormann, R. and Meixner, F. X.: An analytical footprint model for non-neutral stratification, Bound.-Lay. Meteorol., 99, 207–224, doi:10.1023/A:1018991015119, 2001.

Kvalseth, T. O.: Cautionary note about R^2, Am. Stat., 39, 279–285, 1985.

Leclerc, M. Y., Shen, S., and Lamb, B.: Observations and large-eddy simulation modeling of footprints in the lower convective boundary layer, J. Geophys. Res.-Atmos., 102, 9323–9334, doi:10.1029/96JD03984, 1997.

Lenschow, D. H. and Stankov, B. B.: Length Scales in the Convective Boundary Layer, J. Atmos. Sci., 43, 1198–1209, doi:10.1175/1520-0469(1986)043<1198:LSITCB>2.0.CO;2, 1986.

Lenschow, D. and Sun, J.: The spectral composition of fluxes and variances over land and sea out to the mesoscale, Bound.-Lay. Meteorol., 125, 63–84, doi:10.1007/s10546-007-9191-8, 2007.

Lenschow, D. H., Mann, J., and Kristensen, L.: How long is long enough when measuring fluxes and other turbulence statistics?, J. Atmos. Ocean. Tech., 11, 661–673, doi:10.1175/1520-0426(1994)011<0661:HLILEW>2.0.CO;2, 1994.

Lindley, D. V.: Regression lines and the linear functional relationship, Supplement, J. R. Stat. Soc., 9, 218–244, 1947.

Liu, H. P., Peters, G., and Foken, T.: New equations for sonic temperature variance and buoyancy heat flux with an omnidirectional sonic anemometer, Bound.-Lay. Meteorol., 100, 459–468, doi:10.1023/A:1019207031397, 2001.

Liu, S. M., Xu, Z. W., Wang, W. Z., Jia, Z. Z., Zhu, M. J., Bai, J., and Wang, J. M.: A comparison of eddy-covariance and large aperture scintillometer measurements with respect to the energy bal-

ance closure problem, Hydrol. Earth Syst. Sci., 15, 1291–1306, doi:10.5194/hess-15-1291-2011, 2011.

Lu, L., Denning, A. S., da Silva-Dias, M. A., da Silva-Dias, P., Longo, M., Freitas, S. R., and Saatchi, S.: Mesoscale circulations and atmospheric CO_2 variations in the Tapajos Region, Para, Brazil, J. Geophys. Res.-Atmos., 110, 1–17, doi:10.1029/2004JD005757, 2005.

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

Mahrt, L.: Flux sampling errors for aircraft and towers, J. Atmos. Ocean. Tech., 15, 416–429, doi:10.1175/1520-0426(1998)015<0416:FSEFAA>2.0.CO;2, 1998.

Mahrt, L.: Computing turbulent fluxes near the surface: needed improvements, Agr. Forest Meteorol., 150, 501–509, doi:10.1016/j.agrformet.2010.01.015, 2010.

Markkanen, T., Steinfeld, G., Kljun, N., Raasch, S., and Foken, T.: Comparison of conventional Lagrangian stochastic footprint models against LES driven footprint estimates, Atmos. Chem. Phys., 9, 5575–5586, doi:10.5194/acp-9-5575-2009, 2009.

Martin, S., Bange, J., and Beyrich, F.: Meteorological profiling of the lower troposphere using the research UAV "M^2AV Carolo", Atmos. Meas. Tech., 4, 705–716, doi:10.5194/amt-4-705-2011, 2011.

Massman, W. J. and Clement, R.: Uncertainty in Eddy Covariance flux estimates resulting from spectral attenuation, in: Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis, 1st Edn., edited by: Lee, X., Law, B., and Massman, W., Springer Netherlands, Dordrecht, The Netherlands, 264 pp., 2004.

Mauder, M. and Foken, T.: Impact of post-field data processing on eddy covariance flux estimates and energy balance closure, Meteorol. Z., 15, 597–609, doi:10.1127/0941-2948/2006/0167, 2006.

Mauder, M. and Foken, T.: Documentation and Instruction Manual of the Eddy-Covariance Software Package TK3, Universitätsbibliothek Bayreuth, Universität Bayreuth, Abteilung Mikrometeorologie, Bayreuth, Germany, 2011.

Mauder, M., Desjardins, R. L., and MacPherson, I.: Scale analysis of airborne flux measurements over heterogeneous terrain in a boreal ecosystem, J. Geophys. Res.-Atmos., 112, D13112, doi:10.1029/2006JD008133, 2007a.

Mauder, M., Oncley, S. P., Vogt, R., Weidinger, T., Ribeiro, L., Bernhofer, C., Foken, T., Kohsiek, W., De Bruin, H. A. R., and Liu, H.: The energy balance experiment EBEX-2000. Part II: Intercomparison of eddy-covariance sensors and post-field data processing methods, Bound.-Lay. Meteorol., 123, 29–54, 2007b.

Meijninger, W. M. L., Hartogensis, O. K., Kohsiek, W., Hoedjes, J. C. B., Zuurbier, R. M., and De Bruin, H. A. R.: Determination of area-averaged sensible heat fluxes with a large aperture scintillometer over a heterogeneous surface – Flevoland field experiment, Bound.-Lay. Meteorol., 105, 37–62, doi:10.1023/a:1019647732027, 2002.

Meijninger, W. M. L., Beyrich, F., Lüdi, A., Kohsiek, W., and De Bruin, H. A. R.: Scintillometer-based turbulent fluxes of sensible and latent heat over a heterogeneous land surface – a contribution to LITFASS-2003, Bound.-Lay. Meteorol., 121, 89–110, doi:10.1007/s10546-005-9022-8, 2006.

Metzger, S., Junkermann, W., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Measuring the 3-D wind vector with a weight-shift microlight aircraft, Atmos. Meas. Tech., 4, 1421–1444, doi:10.5194/amt-4-1421-2011, 2011.

Moore, C. J.: Frequency-response corrections for Eddy-correlation systems, Bound.-Lay. Meteorol., 37, 17–35, doi:10.1007/BF00122754, 1986.

Panofsky, H., Tennekes, H., Lenschow, D., and Wyngaard, J.: The characteristics of turbulent velocity components in the surface layer under convective conditions, Bound.-Lay. Meteorol., 11, 355–361, doi:10.1007/bf02186086, 1977.

Quenouille, M. H.: Notes on bias in estimation, Biometrika, 43, 353–360, doi:10.1093/biomet/43.3-4.353, 1956.

R Development Core Team: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2011.

Raasch, S. and Schröter, M.: PALM – A large-eddy simulation model performing on massively parallel computers, Meteorol. Z., 10, 363–372, doi:10.1127/0941-2948/2001/0010-0363, 2001.

Ripley, B. D. and Thompson, M.: Regression techniques for the detection of analytical bias, Analyst, 112, 377–383, doi:10.1039/an9871200377, 1987.

Rousseeuw, P. J. and Verboven, S.: Robust estimation in very small samples, Comput. Stat. Data An., 40, 741–758, doi:10.1016/S0167-9473(02)00078-6, 2002.

Schmid, H. P.: Footprint modeling for vegetation atmosphere exchange studies: a review and perspective, Agr. Forest Meteorol., 113, 159–183, doi:10.1016/S0168-1923(02)00107-7, 2002.

Sorbjan, Z.: Statistics of scalar fields in the atmospheric boundary layer based on large-Eddy simulations. Part II: Forced convection, Bound.-Lay. Meteorol., 119, 57–79, doi:10.1007/s10546-005-9014-8, 2006.

Steinfeld, G., Letzel, M. O., Raasch, S., Kanda, M., and Inagaki, A.: Spatial representativeness of single tower measurements and the imbalance problem with eddy-covariance fluxes: results of a large-eddy simulation study, Bound.-Lay. Meteorol., 123, 77–98, doi:10.1007/s10546-006-9133-x, 2007.

Strunin, M. A. and Hiyama, T.: Applying wavelet transforms to analyse aircraft-measured turbulence and turbulent fluxes in the atmospheric boundary layer over Eastern Siberia, Hydrol. Process., 18, 3081–3098, doi:10.1002/hyp.5750, 2004.

Taylor, G. I.: Eddy motion in the atmosphere, Philos. T. R. Soc. Lond., 215, 1–26, doi:10.1098/rsta.1915.0001, 1915.

Thomas, R. M., Lehmann, K., Nguyen, H., Jackson, D. L., Wolfe, D., and Ramanathan, V.: Measurement of turbulent water vapor fluxes using a lightweight unmanned aerial vehicle system, Atmos. Meas. Tech., 5, 243–257, doi:10.5194/amt-5-243-2012, 2012.

Tukey, J. W.: Bias and confidence in not-quite large samples, Ann. Math. Stat., 29, p. 614, 1958.

Uhlenbrock, J., Raasch, S., Hennemuth, B., Zittel, P., and Meijninger, W. M. L.: Effects of land surface heterogeneities on the boundary structure and turbulence during LITFASS-2003: large-eddy simulations in comparison with turbulence measurements, 16th AMS Symposium on Boundary Layers and Turbulence, Portland, USA, Abstract No. 9.3, 8–13 August, 2004.

Vesala, T., Kljun, N., Rannik, U., Rinne, J., Sogachev, A., Markkanen, T., Sabelfeld, K., Foken, T., and Leclerc, M. Y.: Flux and concentration footprint modelling: State of the art, Environ. Pollut., 152, 653–666, doi:10.1016/j.envpol.2007.06.070, 2008.

Vickers, D. and Mahrt, L.: Quality control and flux sampling problems for tower and aircraft data, J. At-

mos. Ocean. Tech., 14, 512–526, doi:10.1175/1520-0426(1997)014<0512:QCAFSP>2.0.CO;2, 1997.

Wang, T. I., Ochs, G. R., and Clifford, S. F.: A saturation-resistant optical scintillometer to measure C_n^2, J. Opt. Soc. Am., 68, 334–338, doi:10.1364/JOSA.68.000334, 1978.

Webb, E. K., Pearman, G. I., and Leuning, R.: Correction of flux measurements for density effects due to heat and water vapour transfer, Q. J. Roy. Meteorol. Soc., 106, 85–100, doi:10.1002/qj.49710644707, 1980.

Wilczak, J. M., Oncley, S. P., and Stage, S. A.: Sonic anemometer tilt correction algorithms, Bound.-Lay. Meteorol., 99, 127–150, doi:10.1023/A:1018966204465, 2001.

Willett, J. B. and Singer, J. D.: Another cautionary note about R^2: its use in weighted least-squares regression analysis, Am. Stat., 42, 236–238, 1988.

Appendix E: Metzger et al. (2013)

Biogeosciences, 10, 2193–2217, 2013
www.biogeosciences.net/10/2193/2013/
doi:10.5194/bg-10-2193-2013

Biogeosciences Open Access

Spatially explicit regionalization of airborne flux measurements using environmental response functions

S. Metzger[1,2,*,**], W. Junkermann[1], M. Mauder[1], K. Butterbach-Bahl[1], B. Trancón y Widemann[3,6,***], F. Neidl[1], K. Schäfer[1], S. Wieneke[4], X. H. Zheng[2], H. P. Schmid[1], and T. Foken[5,6]

[1]Karlsruhe Institute of Technology, Institute for Meteorology and Climate Research, Atmospheric Environmental Research Division, Garmisch-Partenkirchen, Germany
[2]Chinese Academy of Sciences, Institute of Atmospheric Physics, State Key Laboratory of Atmospheric Boundary Layer Physics and Atmospheric Chemistry, Beijing, China
[3]University of Bayreuth, Chair of Ecological Modelling, Bayreuth, Germany
[4]University of Cologne, Institute of Geophysics and Meteorology, Cologne, Germany
[5]University of Bayreuth, Department of Micrometeorology, Bayreuth, Germany
[6]Member of Bayreuth Centre of Ecology and Ecosystem Research, Bayreuth, Germany
[*]now at: National Ecological Observatory Network, Fundamental Instrument Unit, Boulder, USA
[**]now at: University of Colorado, Institute of Arctic and Alpine Research, Boulder, USA
[***]now at: Technical University of Ilmenau, Programming Languages and Compilers, Ilmenau, Germany

Correspondence to: S. Metzger (smetzger@neoninc.org)

Received: 1 October 2012 – Published in Biogeosciences Discuss.: 14 November 2012
Revised: 8 February 2013 – Accepted: 9 February 2013 – Published: 3 April 2013

Abstract. The goal of this study is to characterize the sensible (H) and latent (LE) heat exchange for different land covers in the heterogeneous steppe landscape of the Xilin River catchment, Inner Mongolia, China. Eddy-covariance flux measurements at 50–100 m above ground were conducted in July 2009 using a weight-shift microlight aircraft. Wavelet decomposition of the turbulence data enables a spatial discretization of 90 m of the flux measurements. For a total of 8446 flux observations during 12 flights, MODIS land surface temperature (LST) and enhanced vegetation index (EVI) in each flux footprint are determined. Boosted regression trees are then used to infer an environmental response function (ERF) between all flux observations (H, LE) and biophysical (LST, EVI) and meteorological drivers. Numerical tests show that ERF predictions covering the entire Xilin River catchment ($\approx 3670\,\mathrm{km}^2$) are accurate to $\leq 18\,\%$ ($1\,\sigma$). The predictions are then summarized for each land cover type, providing individual estimates of source strength ($36\,\mathrm{W\,m^{-2}} < H < 364\,\mathrm{W\,m^{-2}}$, $46\,\mathrm{W\,m^{-2}} < \mathrm{LE} < 425\,\mathrm{W\,m^{-2}}$) and spatial variability ($11\,\mathrm{W\,m^{-2}} < \sigma_H < 169\,\mathrm{W\,m^{-2}}$, $14\,\mathrm{W\,m^{-2}} < \sigma_{\mathrm{LE}} < 152\,\mathrm{W\,m^{-2}}$) to a precision of $\leq 5\,\%$. Lastly, ERF predictions of land cover specific Bowen ratios are compared between subsequent flights at different locations in the Xilin River catchment. Agreement of the land cover specific Bowen ratios to within $12 \pm 9\,\%$ emphasizes the robustness of the presented approach. This study indicates the potential of ERFs for (i) extending airborne flux measurements to the catchment scale, (ii) assessing the spatial representativeness of long-term tower flux measurements, and (iii) designing, constraining and evaluating flux algorithms for remote sensing and numerical modelling applications.

1 Introduction

Measurements of the exchange of heat and moisture between the land surface and the atmosphere are critical to our understanding of the role of terrestrial ecosystems in the global climate system. Ground-based eddy-covariance (EC, a summary of all notation is provided in Appendix A) measurements are suited to continuously monitor selected sites for long periods and enable the integration in time (Baldocchi et al., 2001). However, these results might only represent

Fig. 1. Location of the Xilin River catchment in the Inner Mongolia Autonomous Region, China (modified after Steffens et al., 2008).

small areas around the immediate measurement locations (e.g. Kaharabata et al., 1997; Schuepp et al., 1992). On the other hand aircraft-based measurements can provide flux information at regional scales (e.g. Desjardins et al., 1995) but are restricted to short periods of time. Thus the temporal and spatial characteristics of ground-based and airborne measurements complement each other (Gioli et al., 2004; Mauder et al., 2007). It is desirable to integrate both approaches in an effort to provide suitable datasets for the design, constraint, and evaluation of mass and energy exchange models at site as well as at regional scales (Chen et al., 1999; Desjardins et al., 1997). In the following we briefly review the requirements for spatial scaling of airborne EC measurements, and the applicability of airborne EC measurements over complex terrain.

Aggregation approaches enable estimating the exchange over entire landscapes, provided fluxes for characteristic land cover features or domains are known (Beyrich et al., 2006). Flight path segmentation can be a useful tool to directly relate airborne EC measurements to landscape units (e.g. Desjardins et al., 1994; Vellinga et al., 2010). It is also possible to functionally relate these measurements to land cover properties, which then reflect the effects of vegetation, climate, soil and topography on the flux strength. For example, Kirby et al. (2008) propose a method for discerning individual fluxes in a heterogeneous landscape based on subsets of "pure" flux fragments. Another approach is to utilize quantitative information about the EC measurement's spatial context, on which basis environmental response functions (ERFs, Desjardins et al., 1994) can be derived. The general idea of ERFs is to establish a relationship between spatially or temporally resolved flux observations (responses) and corresponding environmental drivers. Hence ERFs are a quantitative mechanism to extract relationships from, and to condense the information content in a dataset. If sufficiently accurate, the extracted relationships can then be used, e.g. to bridge observational scales or to adjust the spatial representativeness of ground-based flux measurements. In addition, current methods to spatially resolve surface fluxes are mainly focused on remote sensing algorithms (e.g. Fan et al., 2007) and process-based land surface models (e.g. Vetter et al., 2012). These procedures often demand far-reaching assumptions, such as the closure of the energy and water balances (e.g. Anderson et al., 2012), or are challenging with respect to the required data basis (e.g. Kaminski et al., 2012; Ziehn et al., 2011). In contrast, accurate ERFs enable inferring high-resolution surface flux maps directly from observational data with minimal, quantifiable assumptions. However, ERFs cannot provide insights, e.g. into ecosystem pools. Consequently, ERFs might be suitable for complementing data assimilation and remote sensing approaches, e.g. through contributing to the design, constraint and evaluation of flux algorithms.

The forenamed applications require the relation of the airborne measured fluxes to land cover properties. To enable this requirement an aircraft is bound to measure close to the surface, where characteristic fluxes from different land covers are not yet fully homogenized (or blended, Mason, 1988; Wood and Mason, 1991). Moreover, the flux must be measured at a constant altitude above ground, so as to avoid artificial flux contributions through altitude fluctuations along vertical gradients (Vickers and Mahrt, 1997). However investigation areas are seldom ideally flat, and topography can vary significantly throughout a domain. To safely follow terrain contours at a low and constant altitude above ground, the aircraft must possess a low ratio of true airspeed to climb rate. Only a few airborne platforms fulfil this requirement (e.g. Bange et al., 2006; Gioli et al., 2004; Thomas et al., 2012), with the weight-shift microlight aircraft (WSMA) being one of them (Metzger et al., 2011, 2012). In forenamed studies we describe a WSMA that enables airborne EC flux measurements in remote settings at reasonable cost and minimal infrastructural demand. The objectives of the present study are to investigate the possibilities of (i) deriving meaningful EC fluxes from WSMA measurements over complex terrain, and (ii) scaling the results to a domain of interest.

We applied the WSMA over the undulating steppe of the Xilin River catchment (XRC), Inner Mongolia, China (Fig. 1). On 21 days in the summer of 2009, flights along line transects were conducted at 50–100 m a.g.l. From boundary layer scaling it is found that the vertical flux gradients below the flight level satisfy the surface layer definition (constant within 5–10 %). Hence measured sensible (H) and latent heat flux (LE) can be interpreted as surface fluxes (Sect. 3.1). Because of its climate and management practices typical for semiarid grasslands of China (Butterbach-Bahl et al., 2011), intensive ecological research commenced in the XRC in the late 1970s (Jiang, 1985). Besides Tibet, Inner Mongolia is

China's most important province for grassland-based livestock production, and desertification due to overgrazing is a major problem for extensive areas (Meurer and Jiang, 2001). In addition, the land cover in the investigation area varies distinctly in space and time (Ketzer et al., 2008; Schaffrath et al., 2011). We postulate that airborne EC flux measurement is a promising tool to gain new insights into the spatial variability of heat and moisture exchange across the XRC.

In the present paper we firstly introduce the WSMA and the on-board measurements, and give an overview of the climate and physical composition of the study area (Sects. 2.1, 2.2). We then describe a measurement strategy (Sect. 2.3) which is linked to a novel data processing approach (Sect. 2.4). A wavelet transformation allows us to resolve fluxes above each overflown cell of a 90 m land cover raster without neglecting flux contributions on much larger scales. In combination with footprint modelling and a non-parametric machine learning technique, an ERF is computed between the airborne flux observations and meteorological and land surface drivers (Fig. 5 provides an overview of the data flow). In Sect. 3 we present the results of this functional relationship and evaluate its potential to explain the spatial distribution of the heat and moisture exchange along the flight lines. We interpret the results in the context of blending scales and statistical errors, and discuss the uncertainty associated with using the ERF to predict fluxes to freely selectable domains in the XRC (Sect. 3.2.4). Lastly, we give an outlook on potential applications of WSMA flux measurements and future improvements of the presented methodology (Sect. 4).

2 Materials and methods

2.1 The weight-shift microlight aircraft

The structure of a WSMA differs from common fixed-wing aircraft: it consists of two distinct parts, the wing and the trike, which hangs below the wing and contains pilot, engine and the majority of the scientific equipment. This particular structure provides the WSMA with exceptional transportability and climb rate, which qualifies it for applications in inaccessible and topographically structured terrain. A detailed description of the physical properties of the WSMA used in this study as well as characteristics and manufacturers of sensors and data acquisition is given in Metzger et al. (2011, 2012). In short, most variables are sampled at 100 Hz and are block-averaged and stored at 10 Hz, yielding a horizontal resolution of approximately 2.5 m. In this study, we use the 10 Hz measurements of the 3-D wind speed ($\sigma_w = 0.04\,m\,s^{-1}$ precision), temperature ($\sigma = 0.04\,K$), humidity ($\sigma = 0.005\,g\,m^{-3}$), and the height a.g.l. ($\sigma = 0.04\,m$). From error propagation it was found that changes in friction velocity (u_*), H and LE of $0.02\,m\,s^{-1}$, $5\,W\,m^{-2}$, and $3\,W\,m^{-2}$, respectively, can be reliably distinguished (Metzger et al., 2012). In addition, we use for this study slow measurements (≤ 0.1 Hz) of humidity (TP3 dew point mirror, Meteolabor AG, Wetzikon, Switzerland), surface temperature (CT infrared thermometer, Optris GmbH, Berlin, Germany), and down-welling shortwave radiation (LI-200 SZ, LI-COR Inc., Lincoln, Nebraska, 400–1100 nm, within an error of 5 % equal to pyranometer measurements, 300–3000 nm).

2.2 Study area

Airborne EC flux measurements were performed in the XRC from 23 June to 4 August 2009. The hilly investigation area lies south of the provincial capital Xilinhot, Inner Mongolia, China (43.1–43.9° N, 116.0–117.2° E; 1000–1500 m a.s.l., Fig. 2). The XRC covers an area of $\approx 3670\,km^2$ and is characterized by temperate continental monsoon climate, with cold and dry winters and warm and wet summers. From data of the years 1982–2005 at the Inner Mongolia Grassland Ecosystem Research Station (IMGERS, 43.63° N, 116.70° E; 1187 m a.s.l., Fig. 2), the monthly mean air temperature ranges from −21 °C in January to +19 °C in July, with an annual mean of +1 °C (Liu et al., 2008). Variability in total annual precipitation is high (166–507 mm) with a mean annual sum of 335 mm. Typically, 60–80 % of the rainfall occurs from June to August (Chen, 1988). June 2009 (57 mm) and August 2009 (60 mm) were in the usual range, but July 2009 (35 mm) only received half of the long-term average rainfall. Detailed information on the meteorological conditions during the flight campaign is provided in Appendix B. Chestnut soils are the main zonal soil types, with a land cover dominated by *Stipa grandis*, *S. krylovii*, *Artemisia* and *Leymus chinensis* steppe. Throughout the XRC the abundance of C4 species in the steppe composition is relatively homogeneous (15–25 %, Auerswald et al., 2009). The growing season usually lasts from the end of May to late September (Liang et al., 2001).

Land cover in the XRC had been classified on the basis of a Landsat 7 Thematic Mapper image of 17 August, 2005 (Wiesmeier et al., 2011). In recent years however the development of settlements sprawled, and irrigated agriculture is gaining popularity (Qi et al., 2007). The Bowen ratio (Bo) of the latter is distinctly different from the land cover classes that already exist in the classification of Wiesmeier et al. (2011). This land cover classification was thus updated and extended by visual reclassification of Advanced Spaceborne Thermal and Reflection Radiometer (ASTER) images of 7 and 28 April 2009. The result is a land cover map with a resolution of 90 m, which is dominated by generic steppe (71 % coverage), intersected by a dune belt (10 %, Fig. 2). The coverage of bare soil, mountain meadow, marshland and rainfed agriculture is each ≈ 5 %, and the coverage of water bodies, settlements and irrigated agriculture is sub-per cent. In the context of this study the land cover

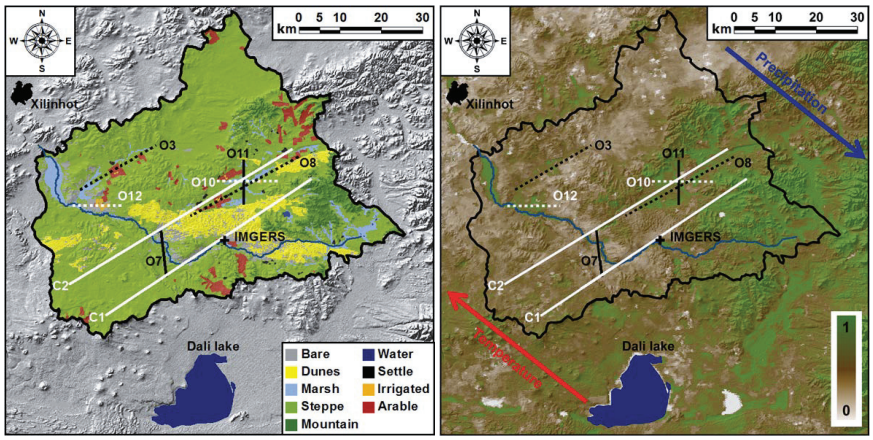


Fig. 2. Maps of the Xilin River catchment (black boundary), with the IMGERS research station and pairs of flight lines. Left: land cover classification (modified after Wiesmeier et al., 2011) over a digital elevation model (Shuttle Radar Topography Mission, Tile 60_04, data version 4.1, Jarvis et al., 2008). The colour codes are abbreviated for bare soil (Bare), marshland (Marsh), generic steppe (Steppe), mountain meadow (Mountain), settlements (Settle), irrigated agriculture (Irrigated), and rainfed agriculture (Arable). Right: MODIS-enhanced vegetation index of 20 July 2009 with a colour bar ranging from $0 < \mathrm{EVI} < 1$.

classification represents the longer-term effects of vegetation, climate, soil and topography.

The spatial variation of temperature and precipitation in the XRC follows altitudinal and latitudinal trends (Auerswald et al., 2009; Wittmer et al., 2010). To resolve the effective state of biophysical surface properties over time, we use Moderate Resolution Imaging Spectroradiometer (MODIS) data. We chose 8-day composites of the daytime land surface temperature (LST, MOD11A2.5, 1 km resolution), and 16-day composites of the enhanced vegetation index (EVI, MOD13Q1, MYD13Q1, 250 m resolution) for this purpose. Due to an 8-day overlap of the EVI data products by the MODIS Terra and Aqua missions, LST and EVI datasets could both be acquired for 4, 12, 20, 28 July, and 5 August 2009. The LST and EVI datasets were bi-linearly interpolated to the 90 m resolution of the land cover classification, and linearly interpolated in time to yield an individual map for each flight day. The spatial gradients in temperature and precipitation throughout the XRC are clearly reproduced by the greenness of the vegetation (Fig. 2). The spatio-temporal resolution of the MODIS data enables assessing the actual state of the biophysical conditions at the land surface. LST and EVI vary significantly not only throughout the study period, but also between the different land cover types (Fig. 3). All land cover types follow a similar temporal trend, with LST and EVI peaking mid-July and end-July, respectively. While open water is the coolest surface, LST and greenness increase from mountain meadow over marshland to irrigated agriculture, which are likely strong sources of evapotranspiration. The reverse relationship (increasing LST and decreasing EVI) is found for settlements, rainfed agriculture, dunes, steppe and bare soil, which are likely strong sources of H.

A ceilometer (LD40 – Vaisala, Helsinki, Finland) was deployed at IMGERS and provided vertical profiles of the atmospheric laser radiation backscatter intensity (Mini light detection and ranging, originally applied for the detection of the cloud base height). The depth of the convective boundary layer (CBL) was inferred from 10 min means of these data, in combination with semi-daily radiosonde ascends in nearby Xilinhot (World Meteorological Organization station 54102, http://weather.uwyo.edu/upperair/sounding.html). For this purpose the maximum gradient method is used, which enables the detection of up to five lifted inversions (Emeis et al., 2008; Helmis et al., 2012; Münkel and Roininen, 2010). It is assumed that the aerosol number concentration, size distribution, shape and chemical composition (refractive index, absorption) adapt rapidly to the CBL structure. If there was more than one maximum or layer detected, the lowest one is taken as the CBL depth. The CBL depth is used in Sect. 2.4.1 for the calculation of atmospheric length scales, and in Sect. 2.4.3 for the source area calculation of the airborne flux measurement. In addition, the cloud cover during the flight periods was monitored by ground personnel at IMGERS.

2.3 Measurement strategy

Advancing into more complex terrain, a flight strategy needs to be derived that considers (i) pilot safety, (ii) vertical flux gradients, (iii) orographically induced effects on radiative transfer and turbulence generation, (iv) statistical errors, and (v) the land cover distribution. Such a flight strategy was derived for the XRC study region using the geographic information system ArcMap 9.2 (Environmental

Systems Research Institute, Redlands, CA, USA). (i) A minimum flight level of 50 m above ground was found to provide the pilot with sufficient clearance for safe flight even under buoyancy-driven turbulence. (ii) Measuring at a constant pressure level would make a conversion of the measured temperature and densities into potential quantities (Sect. 2.4) less important. However, over tilted or undulating terrain the aircraft would partially travel along the vertical flux gradients, thus spuriously contaminating the measured flux signal. A correction for the vertical flux gradients over complex terrain is not as straightforward and well conditioned as the conversion into potential quantities. In order for the fluxes to remain interpretable, the aircraft should thus measure at approximately constant height above terrain; i.e. the flight paths should follow the terrain contours. (iii) We use a digital elevation model (Shuttle Radar Topography Mission, Tile 60_04, data version 4.1, Jarvis et al., 2008) to calculate slope angles. In an effort to avoid immediate orographically induced effects on radiative transfer and turbulence generation, all locations within 500 m radius around slopes exceeding 6° were masked. This radius approximately equals five times the standard deviation (SD) of the terrain elevation. (iv) To reduce the statistical errors, the flight path should be long and perpendicular to the mean wind direction. Thus we aligned straight flight lines along four wind axes in the areas which were not masked in the previous step. (v) Of the flight lines that were frequently perpendicular to the mean wind direction, those that best represented the land cover distribution in the XRC were covered on multiple flight days. This strategy results in a terrain-following flight patterns, with typical altitude gradients of 100 m vertical on 10 km horizontal. The climb angle of the aircraft rarely exceeds ±5°, and on average the height above ground is constant to within 12 m (Table 1). Each flight line was repeated until a minimum of 40 km of data were acquired.

The aircraft was operated from IMGERS (Fig. 2). For the present study we use data from six days in July 2009 (Table 1), which were selected according to the availability of auxiliary datasets and homogeneity of the down-welling shortwave radiation $S\downarrow$ along the flight tracks (Table B1). On each day measurements were carried out along pairs of approximately parallel flight lines at a nominal airspeed of 27 m s^{-1}, with eight individual flight lines in total. Each pair is located across or along the humidity and temperature gradients in the XRC (Fig. 2). This strategy provides two independent datasets for each flight day, and covers the fundamental climatic gradients in this area.

The land cover type most frequently observed below all flight lines is steppe (Table 1). Flight lines with significant surface coverage of marshland or irrigation agriculture tend to be greener (higher EVI) compared to flight lines with significant coverage of dunes or bare soil. In the following, we investigate whether spatially resolved land cover information can be used as predictor for H and LE measured along the flight lines. At this we hypothesize that LST and EVI are representative proxies for heat and moisture sources on the surface, respectively (e.g. Glenn et al., 2008; Lyons and Halldin, 2004; Nagler et al., 2007). Because aircraft measurements cover a broad state space, the resulting observations are particularly suited to infer ERFs.

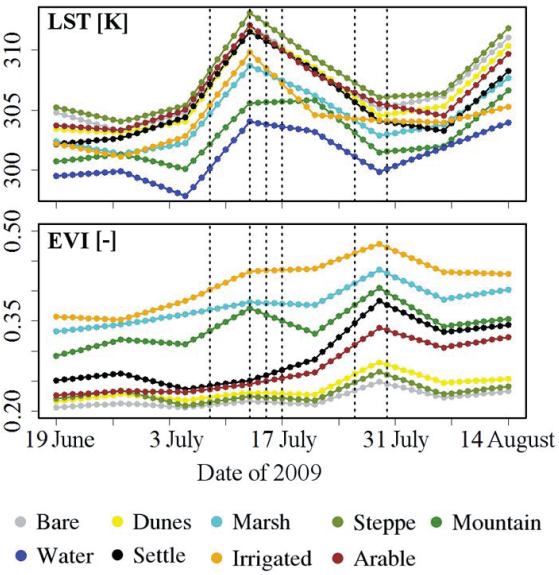


Fig. 3. Change of land surface temperature (top) and enhanced vegetation index (bottom) for each land cover class throughout the study period. The land cover colour code and corresponding abbreviations are identical with Fig. 2. Vertical dashed lines indicate the flight dates. The land cover "Water" is not present for the EVI, because water absorbs strongly in the near infrared, leading to negative EVI values that are not indicative of vegetation greenness.

2.4 Data processing

An analysis package for the processing of the airborne EC data was developed in GNU R version 2.13 (R Development Core Team, 2012). The analysis package is described in detail in Metzger et al. (2012) and is available upon request. Relevant processing steps are: (i) the raw data are screened for spikes; (ii) humidity from fast response and slow reference sensors are merged using a complementary filter; (iii) the WSMA temperature and densities are transformed to potential quantities at the mean flight altitude (pressure level) of each flight line; (iv) the time delay due to separation between the vertical wind measurement and the temperature and humidity measurements is corrected by maximising their lagged correlation; (v) the WPL correction according to Webb et al. (1980) is used to correct LE for density fluctuations; (vi) correction for spectral artefacts (-1 ± 1 %, -2 ± 1 %, and -6 ± 2 % for the SDs in the wind components along, transverse and vertical to the aircraft coordinate

Table 1. Summary of the WSMA flights selected for analysis and related surface conditions. Shown are date, Chinese standard time (CST = coordinated universal time + 8), flight identifier (ID), length of each flight line l, repetitions rep, cumulated precipitation in a 10-day trailing window P, most frequently occurring land cover classes LC1–LC3, and the enhanced vegetation index EVI immediately below the flight lines. A legend with colour codes for LC and EVI is provided at the bottom. The LC colour code and corresponding abbreviations are identical with Fig. 2.

Date	Time CST	ID	l [km]	rep	P [mm]	LC1	LC2	LC3	EVI
8 July 2009	10:20–10:50	O10	15	3	1.4	54%	35%	6%	35 ± 13%
	12:00–12:50	O12	13	6		32%	30%	21%	35 ± 18%
13 July 2009	11:30–12:10	O8	30	2	5.0	59%	15%	11%	32 ± 10%
	12:40–13:10	O3	21	2		70%	17%	13%	25 ± 6%
15 July 2009	11:30–12:20	O11	11	6	5.0	48%	42%	10%	33 ± 9%
	12:30–13:00	O7	11	4		79%	16%	4%	21 ± 6%
17 July 2009	11:00–11:30	O11	11	4	5.2	54%	40%	4%	36 ± 12%
	12:20–13:00	O7	11	5		82%	10%	5%	21 ± 8%
26 July 2009	12:50–15:30	C1	60	2	13.4	51%	24%	20%	25 ± 7%
	13:10–15:10	C2	63	2		73%	10%	9%	27 ± 9%
30 July 2009	11:00–13:30	C1	60	2	14.3	52%	24%	18%	25 ± 7%
	11:10–13:20	C2	63	2		74%	10%	8%	26 ± 8%

LC	Arable	Bare	Dunes	Irrigated	Marsh	Steppe
EVI	> 20%–25%		> 25%–30%	> 30%–35%		> 35%

system, respectively) and cospectral artefacts (-4 ± 4 %, 0 ± 1 %, and 0 ± 2 % for u_*, H and LE, respectively) after Metzger et al. (2012); and (vii) calculation of random and systematic statistical errors after Lenschow and Stankov (1986) and Lenschow et al. (1994).

2.4.1 Horizontal mixing between surface and flight level

Horizontal mixing between the surface and the flight level results in the spatial integration of fluxes above heterogeneous terrain, a process also referred to as "blending" (e.g. Mason, 1988). Working toward an ERF between surface properties (driver) and flux measurement (response), we will test three hypotheses related to horizontal mixing. At flight level (i) the turbulence is in approximate equilibrium with the land surface in the flux footprint; (ii) the measured turbulence statistics are representative of the mechanical setting upwind; (iii) changes in the turbulent flux can be resolved at the horizontal scale of surface heterogeneities.

Several analytical formulations have been developed to characterize a mixing regime (e.g. Mahrt, 2000; Raupach and Finnigan, 1995; Wood and Mason, 1991). Such formulations are usually based on the comparison of characteristic length scales for surface heterogeneity and CBL mixing. Here we use the autocorrelation function to estimate the typical horizontal scale of surface heterogeneity L_H. For this purpose we integrate the autocorrelation function of the WSMA-measured surface temperature T_s at distance d along a flight line from zero lag to the first crossing with zero at lag c_0;

$$L_H = \int_0^{c_0} \frac{\overline{T_s'(d) T_s'(d+c)}}{\overline{T_s'(d)^2}} \mathrm{d}c, \quad (1)$$

where overbars denote the mean along a flight line, and primes denote the deviations from this mean. L_H can be interpreted as the spatial coherence of surface features along the flight path (e.g. Strunin et al., 2004). In the following we assume that L_H is isotropic within the flux footprint, i.e. also representative perpendicular to the flight path.

In order to further characterize the mixing regime, L_H can be compared to atmospheric length scales. These length scales inter-relate the transport strengths in the horizontal and vertical directions, and correspond to the along wind distance after which the air mass below a reference level is approximately homogenized. Formulations for atmospheric length scales mainly differ in their use of (i) the measures of transport strengths, and (ii) the vertical or horizontal reference scale. Raupach and Finnigan (1995) proposed a length scale L_R (now also referred to as Raupach length), which characterizes the mixing regime throughout the entire CBL;

$$L_R = 0.8 z_i \frac{\overline{u}}{\overline{w_*}}, \quad (2)$$

with CBL depth z_i, average (bulk) horizontal transport velocity u, and convective velocity w_*. Because u cannot be measured directly, it is substituted in Eq. (2) and in the following investigations with the measured horizontal wind speed at flight altitude. The influence of surface heterogeneities with spatial scales L_H that are small compared to L_R is confined below the CBL top. In this case the concept of a "blending height" within the CBL arises. The blending height corresponds to a vertical level at which the turbulent flow field over heterogeneous terrain approaches equilibrium with the local vertical gradient. If the blending height is confined within the surface layer, Monin–Obukhov similarity can be applied above the blending height (Mahrt, 2000). Wood and Mason (1991) define the thermal blending height for unstable stratification;

$$z_{TB1} = L_H \frac{\overline{w'\theta'_{0,v}}}{\overline{u}\,\overline{\theta_{0,v}}} \cdot 323, \tag{3}$$

with the buoyancy heat flux from the virtual potential temperature $\theta_{0,v}$, and the horizontal wind speed u at the blending height. This thermal blending height can be rearranged as thermal blending length;

$$L_{TB1} = \overline{z}\frac{\overline{u}\,\overline{\theta_{0,v}}}{\overline{w'\theta'_{0,v}}} \cdot 3.1 \cdot 10^{-3}, \tag{4}$$

now representing the smallest scale of surface heterogeneity that significantly influences the turbulent flow at flight level z above ground. An improved version of the thermal blending length was proposed by Mahrt (2000);

$$L_{TB2} = L_R \frac{\overline{\theta_{0,v}}}{\sigma_{T_s}} \cdot 4.3 \cdot 10^{-3}, \tag{5}$$

which considers the SD of T_s as a measure for the amplitude of surface heterogeneity. The numeral factors in Eqs. (3)–(5) were estimated from observations by Mahrt (2000). In Sect. 3.1 the results of above formulations are used to test the initial hypotheses related to horizontal mixing.

2.4.2 Wavelet cross-scalogram

The differentiation of land cover types is at odds with the classical time-domain EC method, which assumes homogeneous terrain. However, Parseval's theorem implies that the covariance of signals may be studied not only in the time domain, but equivalently in the frequency domain. There, we have the wavelet transform family of methods at our disposal, which are particularly suited for the spectral analysis of non-stationary signals.

A wavelet is a signal that is localized in both time and frequency. Different localizations of the same basic shape (daughter wavelets) are constructed as a function of time t by defining;

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) = \frac{1}{\sqrt{|a|}}\psi(q), \tag{6}$$

where ψ is a suitable mother wavelet, a a scale parameter (in frequency domain), b a location parameter (in time domain), and $q = (t-b)/a$ a dimensionless coordinate (in time–frequency space). The convolution $\int x(t)\psi_{a,b}(t)\,dt$ of a signal x with a daughter wavelet $\psi_{a,b}$ yields a wavelet coefficient $W_x(a, b)$, a wavelet transform being a collection of such coefficients. We follow the procedure of Torrence and Compo (1998), using the continuous wavelet transform approximation for discrete input. The chosen mother wavelet is the Morlet wavelet $\psi(q) = \pi^{-1/4}e^{i\omega_0 q}e^{-q^2/2}$ with the frequency parameter $\omega_0 = 6$. The relevant parameters are spaced exponentially in frequency and linearly in time, respectively: $a_j = a_0 2^{j\delta j}$ for $j = 0, \ldots, J$ and $b_n = n\delta t$ for $n = 0, \ldots, N-1$, with length of the dataset N, initial scale parameter $a_0(\delta t, \psi)$ and number of scale increments $J(a_0, \delta j, \delta t, N, \psi)$. a_0 and J are chosen such that the extreme wavelet scales match the period of the Nyquist frequency, here 0.2 s, and the duration of the dataset, respectively. The unit of increment in the time domain, δt, is given by the sampling period of the time series, here 0.1 s. The unit of increment in frequency domain, δj, can be set to different values, with smaller values increasing both resolution and redundancy. For the present data, the results of the wavelet analyses were insensitive to the choice of δj in the range $0.0625 < \delta j < 0.25$. Hence we follow the example of Torrence and Compo (1998) and use $\delta j = 0.125$. The wavelet scalogram of a signal x is defined as the matrix of $|W_x(a, b)|^2$ for all admissible a, b. Likewise, the wavelet cross-scalogram of two signals x, y is the matrix of $W_x(a, b)W_y(a, b)^*$, where * denotes the complex conjugate. The global covariance of x and y can be estimated by weighted averaging;

$$\mathrm{cov}_{a,b} = \frac{\delta j \delta t}{C_\delta N}\sum_{j=0}^{J}\sum_{n=0}^{N-1}\frac{W_x(a_j, b_n)W_y(a_j, b_n)^*}{a_j}, \tag{7}$$

where C_δ is a reconstruction factor specific to each mother wavelet, here 0.776 for the Morlet wavelet. The covariance can also be estimated locally for a subinterval of either j or n. This is a useful feature for dealing with changes in land cover: the continuous wavelet transform is highly redundant, with high correlation between adjacent low-frequency coefficients. Therefore, the covariance for a subinterval in time can be estimated without neglecting low-frequency, large-scale contributions. The downside of the large low frequency support is that edge effects due to the finite overall dataset increase with scale. Torrence and Compo (1998) define the cone of influence (COI) as the boundary where the power of edge-related artefacts is damped by a factor of e^{-2}. Integration over all scales yields results close to the time-domain EC method (here, -7 to -3 % median differences), but also includes less reliable estimates above the COI (e.g. Strunin and Hiyama, 2004). Considering only scales below the COI rejects those less reliable estimates (e.g. Mauder et al., 2007). However, because part of the scale range is excluded, such a

procedure also systematically increases the discrepancy between wavelet and time-domain EC methods (here, −22 to −7 % median differences). Moreover, the COI tapers toward the centre of the dataset (Fig. 4). Different scales would be included when estimating the covariance for subintervals at different positions along the flight path. Hence, for the localization of flux contributions in space we (i) integrate over all scales of a subinterval, and (ii) use a correction factor for each individual flight path to compensate the difference between wavelet and time-domain EC methods. Such a procedure is suited for the derivation of ERFs, because (i) it enables localization in space, (ii) considers contributions to the local flux from scales that are larger than the subinterval, (iii) is not biased with respect to the global time-domain covariance, and (iv) uncertainty arising from edge-effects is propagated in the ERF and included in the final uncertainty metric (Sect. 2.5.1).

Computations were performed with a continuous wavelet transform package written in GNU R (R Development Core Team, 2012), partially based on the published code of Torrence and Compo (1998) available from http://atoc.colorado.edu/research/wavelets. u_*, H and LE (and analogously the SDs of the wind components) are calculated for overlapping subintervals of 1000 m length. The subintervals are centred above each cell of the land cover/LST/EVI grids that was overflown by the WSMA, principally yielding one flux observation every 90 m. The resulting sample size for all 12 flights in Table 1 is $N = 8446$.

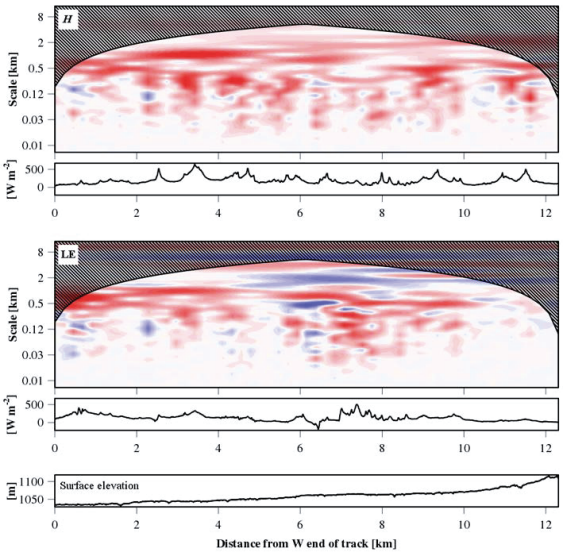


Fig. 4. Wavelet cross-scalograms for the sensible heat flux (top panel) and the latent heat flux (centre panel) along flight pattern O12 on 8 July 2009, 12:16–12:24 CST. The colour palette changes from blue (downward fluxes) over white (neutral) to red (upward fluxes). The shaded areas identify the cone of influence. Below each cross-scalogram the integrated flux over all scales is shown for each overflown 90 m cell of the land cover grid. The surface elevation along the flight pattern is displayed in the bottom panel.

2.4.3 Footprint modelling

The footprint- or source weight function quantifies the spatial contributions to each flux observation (Schmid, 2002; Vesala et al., 2008). For this purpose we use the footprint model of Kljun et al. (2004, KL04), which is a parameterisation of the backward Lagrangian model of Kljun et al. (2002) in the range $-200 \leq z/L \leq 1$, $u_* \geq 0.2\,\mathrm{m\,s^{-1}}$, and $1\,\mathrm{m} \leq z \leq z_i$. The parameterisation depends upon u_*, measurement height z, SD of the vertical wind σ_w and the aerodynamic roughness length z_0, of which u_*, z and σ_w are measured directly by the WSMA. The roughness length is inferred using the logarithmic wind profile with the integrated universal function for momentum exchange after Businger et al. (1971) in the form of Högström (1988). The KL04 is a cross-wind integrated footprint model; i.e. it does not resolve the distribution perpendicular to the main wind direction. In order to account for cross-wind dispersion, the KL04 was combined with a Gaussian cross-wind distribution function (Kljun et al., 2013, in the following referred to as KL04+). In addition to above variables, the SD of the crosswind from WSMA measurements and the depth of the CBL z_i from ceilometer measurements (Sect. 2.2) are used. This results in a computationally fast footprint parameterisation which considers 3-D dispersion and is not constrained to applications in the surface layer. Metzger et al. (2012) evaluated KL04+ against a backward Lagrangian reference footprint model, and good agreement was found for all considered cases.

Turbulence statistics for a 1000-m-long subinterval over the wavelet scalograms (Sect. 2.4.2) are used to evaluate the KL04+. One evaluation is carried out for each overflown cell of the land cover, LST and EVI grids (i.e. every 90 m along the flight path). With the overflown grid cell as base point, the footprint weights $w_{xy}(\sum w_{xy} = 1)$ are calculated for each grid cell with position x, y, relative to the base point. From here the footprint composition is calculated;

$$\mathrm{LST} = \sum_x \sum_y w_{xy} \cdot \mathrm{LST}_{xy}, \quad (8)$$

$$\mathrm{EVI} = \sum_x \sum_y w_{xy} \cdot \mathrm{EVI}_{xy}, \quad (9)$$

with the land surface temperature and enhanced vegetation index for each grid cell, LST_{xy} and EVI_{xy}, respectively. For graphical representation all evaluations of KL04+ along a flight line are superimposed and normalized to a sum of unity. Additional information and references regarding footprint calculations along line transects can be found in Hutjes et al. (2010) and Meijninger et al. (2006).

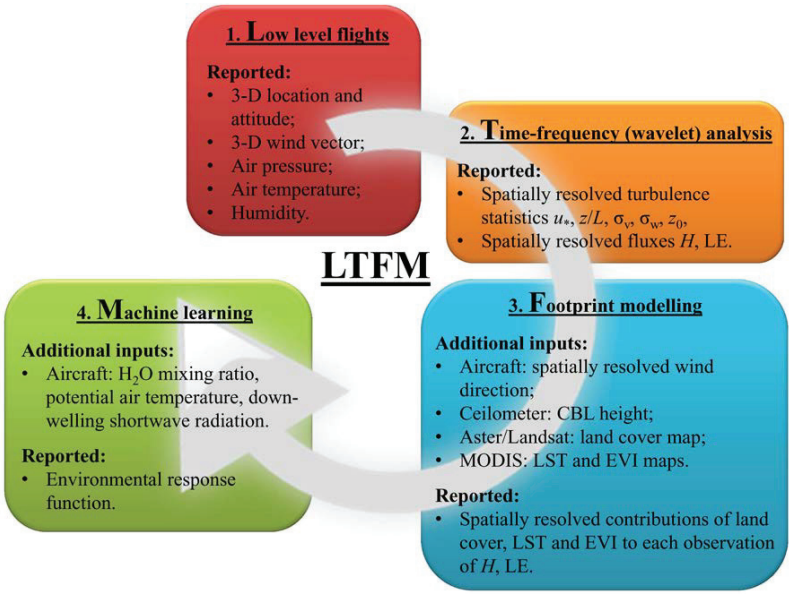


Fig. 5. Flow chart showing how input and reported data streams are processed along the four principal steps of the LTFM method. Additional detail is provided in Sects. 2.4.4 and 4, and a summary of all notation can be found in Appendix A.

2.4.4 Environmental response function

We base the development of a catchment-specific ERF on the works of Chen et al. (1999), Hutjes et al. (2010) and Ogunjemiyo et al. (2003). The general idea is to establish a functional relationship between spatially or temporally resolved flux observations (responses) and corresponding environmental drivers. Figure 5 provides an overview of the novel approach to ERF presented in the following.

Thus far, a suitable number of flux observations was obtained by either shortening the time-domain EC averaging interval (Chen et al., 1999; Ogunjemiyo et al., 2003), or by stratifying repeated observation along the same flight line on different days (Hutjes et al., 2010). The inherent drawbacks are the neglect of either long wavelength contributions to the flux measurement, or inter-day variability of ecosystem drivers. Both are overcome using the wavelet cross-scalogram technique (Sect. 2.4.2).

Previously, the development of ERFs has solely focused on drivers in the footprint of the flux observations, namely discrete land cover classifications. This procedure ignores within-class variability across a catchment, e.g. along climatic or altitudinal gradients, which can be overcome by using continuous variables such as LST and EVI instead. In addition, the present approach considers the meteorological drivers $S\downarrow$, mixing ratio (MR), and potential temperature (θ). This avoids the need for stratifying or pre-selecting data, and enables constructing a single ERF that is valid for the entire observation period and, within the range of the measured variables, throughout a catchment of interest.

Hitherto, ERFs were determined as the inverse of a linear mixing matrix, using either numerical (Chen et al., 1999) or regression methods (Hutjes et al., 2010; Ogunjemiyo et al., 2003). Such a procedure assumes a linear relationship between drivers and responses, which is subject of on-going discussion and research (e.g. Raupach and Finnigan, 1995). Instead, the present approach uses boosted regression trees (BRTs), a non-parametric machine learning technique, to establish an ERF between drivers and responses. In contrast to parametric approaches, BRTs do not assume a predetermined form of the response, but construct an ERF according to the information in the data. It is for this reason that not the absolute values of the land surface and meteorological drivers are important, but rather their spatial variability and coherence. In case of the land surface drivers for example, the only assumption made here is that the spatial patterns of LST and EVI approximate the spatial patterns of source strength in H and LE (e.g. Holmes, 1970; Oke, 1987). This is a much weaker assumption than a mechanistic link, and adds power to the method. BRTs can fit complex nonlinear relationships, automatically handle interactions between drivers, and provide predictive performance that is superior to most traditional modelling methods (e.g. Hu et al., 2010). Here we use the BRT work package by Elith et al. (2008), which builds upon the GBM library by Ridgeway (2012). To identify the optimal choice of parameters and variables for the BRTs, a sensitivity analysis was conducted

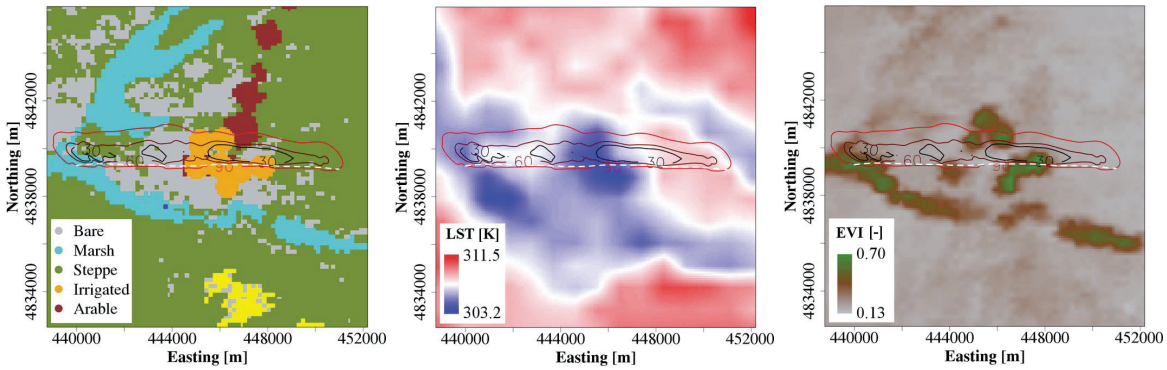


Fig. 6. Flight along pattern O12 on 8 July 2009, 12:16–12:24 CST (white dashed line). The composite flux footprint along the flight line (30 %, 60 %, 90 % contour lines) is superimposed over maps of land cover (left panel), land surface temperature (LST, centre panel), and enhanced vegetation index (EVI, right panel). The land cover colour code and corresponding abbreviations are identical with Fig. 2.

using the cross-validation (CV) procedure described in Elith et al. (2008). During cross validation all available data are divided into 10 random combinations of training (90 %) and evaluation (10 %) fractions, which allows assessing and optimising model performance. The parameter settings that minimized predictive deviance for the present dataset were found to be: absolute (Laplace) error structure, bag fraction (0.7), tree complexity (5), learning rate (0.1), and number of trees (10^4). The initial set of variables also included time of the day, MODIS albedo, atmospheric pressure, land cover, z, z_i, u, u_*, z_0, virtual potential temperature, as well as elevation, topographic wetness index, aspect, and slope of the footprint modelled source area. We use the variable dropping algorithm by Elith et al. (2008) to reach a compromise between predictive deviance and model parsimony. This algorithm (i) fits a BRT model, (ii) performs a 10-fold CV, (iii) drops the least important predictor (determined from the improvement to the model and the number of splits, Friedman, 2001), and (iv) repeats this sequence until a stopping criterion is reached. The mean CV deviance can be used to decide how many variables can be removed without significantly affecting predictive performance. Here, we set an upper threshold of 30 W m^{-2} for the mean CV deviance, which equals $\leq 1/2$ the random sampling error in the flux observations (Table 4). The dropping of variables first stopped for LE at 29.2 W m^{-2} mean CV deviance, yielding a set of the five most important predictors (LST, EVI, $S\downarrow$, MR, and θ). For H the same predictor set yields a mean CV deviance of only 22.6 W m^{-2}. Remarkably, atmospheric pressure, z, and z_i were no significant predictors for the observed fluxes. This indicates that the chosen flight/analysis strategy effectively minimizes cross-contamination of the flux observations by vertical flux/pressure gradients. Analogously the algorithm dropped elevation, aspect, and slope of the footprint modelled source area as predictors. This shows that slope-induced effects on radiative transfer or turbulence generation do not significantly impact the flux observations. Consequently, the final BRT model is fitting an ERF to H and LE as function of only the five most important predictors. This ERF is then used to predict H and LE throughout the XRC, as a function of LST and EVI for each grid cell, and the median $S\downarrow$, MR, and θ for the duration of a flight.

In the following we will use the term LTFM to refer to the overall procedure consisting of Low level flights, Time–frequency-, Footprint-, and Machine learning analyses (Fig. 5).

2.5 Uncertainty

Throughout the present study, we use the median and the median absolute deviation as preferred measure of location and scale, respectively (Croux and Rousseeuw, 1992; Rousseeuw and Verboven, 2002). All resulting uncertainty estimates are representative of one standard deviation. For the purpose of detecting systematic differences between observations and predictions, we use the maximum-likelihood fitting of a functional relationship (MLFR, Ripley and Thompson, 1987). This method assigns a weight to each data couple in the relationship, which is inversely proportional to its error variances. In our case, the squared random flux errors in the observations, and the residuals in the BRT cross-validation ensemble are used. This appreciates reliable data and depreciates uncertain data couples. The errors in the MLFR coefficients are determined from a jackknife estimator (Quenouille, 1956; Tukey, 1958). If the regression intercepts were not significant, the relationships were forced through the origin, and confidence intervals were determined from the slope error. The coefficient of determination R^2 was calculated in analogy to weighted least-squares regression (Kvalseth, 1985; Willett and Singer, 1988). It is the proportion of variation in the weighted dependent variable that can be accounted for by the weighted independent variable.

Uncertainty in the LTFM up-scaling procedure originates from different sources during measurement and data analysis. Part of these uncertainty terms exhibit random characteristics; i.e. they tend to cease with sample size. Another part however will systematically bias the results, independent of sample size. An uncertainty budget for the random and systematic uncertainties in the LTFM procedure will consist of uncertainty terms for (i) instrumentation and hardware, (ii) turbulence sampling, (iii) spatio-temporal analysis, (iv) BRT residuals, (v) BRT response function, and (vi) BRT state variables. While uncertainty terms (i), (ii), and (iv) can be quantified with readily available procedures (Sects. 2.4, 3.3), in teh following we describe several techniques to assess terms (iii), (v) and (vi).

2.5.1 Spatio-temporal analysis in heterogeneous terrain

Under the umbrella of spatio-temporal analysis, we quantify in the following the uncertainty contribution from wavelet analysis, footprint modelling, and the assumption of linear mixing. The fluxes derived from the wavelet cross-scalogram were adjusted to match the leg-averaged fluxes from time-series EC, which avoids bias between both techniques. Also areas above the wavelet cross-scalogram COI were used in the flux calculation to ensure including all scales of turbulent transport along the entire transect. However, values above the COI are potentially distorted due to edge effects, in particular close to the beginning and the end of each transect. These artefacts propagate in the resulting variances and fluxes, and consequently into the footprint estimates. Additional spatial uncertainty terms result from the use of an "offline" footprint model that does not consider the actual flow field, as well as from the MODIS EVI and LST data. The use of BRTs does not expect a linear response between the state variables and the flux signal. However, LTFM still assumes the linear mixing of the flux signal with respect to the contributing surface patches with different biophysical properties and source strengths.

To quantify the error inherent in the above analysis steps, we compare maps of LTFM predicted fluxes to airborne flux observations. (i) The BRT is trained with all available observations ($N = 8466$). (ii) Using the median state variables along each flight leg ($N = 42$), the BRTs response function is used to predict a similar number of flux maps (Fig. 11). (iii) The LTFM footprints are superimposed over these flux maps. (iv) For each flux observation, a predicted flux is calculated as the footprint-weighted average of all contributing cells, and (v) predictions and observations are compared.

2.5.2 Response function

BRTs are a non-parametric machine learning technique in which a response function is constructed according to the coherencies in the training data. As a direct consequence the predictive performance of BRTs depends on how complete the combinations of state variables in the evaluation data are represented in the training data. Here we assess the susceptibility of the BRT response function and predictive performance to missing state variable combinations in the training data. For this purpose, 12 incomplete training datasets are created, each of which omitting a different flight out of the total of 12 flights in Table 1. For each incomplete training dataset, (i) the BRT is trained, (ii) the resulting response function is used with the state variables along the omitted flight for prediction, and (iii) predictions and observations are compared.

2.5.3 State variables

Here, we consider the uncertainty resulting from disregarding part of the natural variability in the state variables that are used for spatially and temporally explicit BRT predictions. For this purpose we quantify the disregarded parts of the natural variability in each state variable and propagate it through the full BRT model. While explicit in time, the meteorological variables measured by the aircraft do not cover the entire catchment. We estimate a measure of spatial variability from all subsequent pairs of flights that are located in different areas of the catchment (Table 1, Fig. 2). The median differences throughout the catchment for $S\downarrow$ ($-6 \pm 12\,\mathrm{W\,m^{-2}}$), θ ($-1.1 \pm 1.1\,\mathrm{K}$), and MR ($-0.5 \pm 0.3\,\mathrm{g\,kg^{-1}}$) are not significant (Wilcoxon rank-sum test, $p \geq 0.18$). On the contrary, MODIS EVI and LST are explicit in space, but not continuous in time. The 8-day trends from one scene to the next are accounted for in the BRT procedure through temporal interpolation between the MODIS scenes (Sect. 2.2). However, processes of shorter duration, such as frequent events of small-scale convective precipitation, go unaccounted. Hence we estimate a measure of the natural variability between two MODIS scenes. For this purpose we calculate the median change of all grid cells between all subsequent MODIS scenes, amounting to 0.01 ± 0.05 for EVI and $-0.5 \pm 6.2\,\mathrm{K}$ for LST. The random part of EVI and LST natural variability by far exceeds the MODIS data product uncertainty of ≈ 0.015 (Xiang et al., 2003) and $\approx 1\,\mathrm{K}$ (Wan and Li, 2008), respectively. Hence MODIS data product uncertainty was not considered separately.

The correlation matrix between the state variables was calculated using all 8446 aircraft observations. A variance-covariance matrix was calculated from this correlation matrix and the random part of the state variables' natural variability. Preserving the variance-covariance relationship, 1000 samples were drawn from a multivariate normal distribution with zero mean. These represent 1000 combinations of co-existing natural variability in the state space of the BRT model. The propagation through the BRT model was performed individually for each combination by (i) superimposing the estimated natural variability over the measured state variables of all 8446 observations, (ii) performing a BRT prediction, and (iii) comparing the results to the undisturbed predictions.

3 Results and discussion

In the first part of this section, we assess the surface–atmosphere mixing regimes. From there wavelet analysis, footprint modelling and BRTs are used to infer ERFs between land surface properties and the flux measurements. Lastly, uncertainties in the LTFM up-scaling procedure are analysed and discussed.

3.1 Horizontal mixing between surface and flight level

On spatial average energy conservation requires that the vertical profiles of H and LE approach their respective entrainment flux at the top of the CBL (e.g. Deardorff, 1974; Sorbjan, 2006). The linear vertical flux gradient of H throughout the CBL was calculated ($-0.21\,\mathrm{W\,m^{-2}\,m^{-1}}$ – $-0.06\,\mathrm{W\,m^{-2}\,m^{-1}}$), assuming that H ceases at the statically stable entrainment zone around 0.8 CBL. However, the entrainment flux of E is unknown. Hence we cannot estimate the vertical flux gradient of E, but assume a comparable order of magnitude as for H. The resulting effect of the vertical flux gradient below the flight level is $-5 \pm 2\,\%$ of H, which falls well within the surface layer definition (e.g. Raupach and Finnigan, 1995; Stull, 1988, flux constant within $5-10\,\%$). Thus, it is feasible to assume that H and E measured at flight level are representative of surface fluxes.

The characteristic length scale of surface heterogeneity is on the order of several hundred to thousand meters, with an average of $L_\mathrm{H} = 1012 \pm 715\,\mathrm{m}$ (Table 2). Along identical flight paths L_H is comparable between days with different meteorological settings (e.g. 15 and 17 July 2009, 26 and 30 July 2009). This confirms the usefulness of the surface temperature measurement as a proxy for surface heterogeneity. Only the longer flight paths C1 and C2 cross the dune belt in the centre of the catchment (Fig. 2). The dune belt is the largest continuous land cover after steppe, and consequently the autocorrelation function of T_s estimates large values of L_H (1458–2615 m). During all flights, L_H was small compared to the Raupach length ($L_\mathrm{R} = 1532$–$5214\,\mathrm{m}$), and thus the influence of the surface heterogeneity is confined within the CBL ($z_i = 1100$–$2500\,\mathrm{m}$). Here we use the thermal blending height ($z_{\mathrm{TB1}} = 40 \pm 29\,\mathrm{m}$) as an estimate for the vertical level where quasi-equilibrium of the turbulent exchange between land surface and atmosphere is reached. At all times the flight level ($z = 48$–$102\,\mathrm{m}$) is above z_{TB1} and below $\approx 10\,\%$ of the CBL depth, a common estimate for the depth of the atmospheric surface layer (e.g. Raupach and Finnigan, 1995; Stull, 1988). Hence it is feasible to assume that the turbulence measurement at flight level is representative for the land surface in the flux footprint. The interpretation of the flux observations might be more complicated for measurement heights below the thermal blending height (limited spatial representativeness) or above the surface layer (vertical flux gradient). The blending length formulations L_{TB1} (1660 ± 723) and L_{TB2} (957 ± 441) are used to assess the minimum size of surface heterogeneity that significantly influences the flow at flight level. Here we use $255\,\mathrm{m} < L_{\mathrm{TB2}} < 1852\,\mathrm{m}$ as a guideline, because L_{TB2} is also representative of the magnitude of surface heterogeneity. The native resolution of the EVI data (230 m) and the land cover data (90 m) is equal to or better than L_{TB2}, and thus sufficient to reproduce the variability of the land cover. In comparison, the native resolution of the LST data (1000 m) is coarse, potentially leading to an attenuation of the ERFs.

Using the wavelet cross-scalogram, long wavelength contributions to the flux do not constrain the spatial resolution of the flux computation along the flight path. Nevertheless, the random flux error is inversely proportional to the square root of the averaging length (e.g. Lenschow and Stankov, 1986), and propagates directly into the computation of the ERFs. Hence, we consider a trade-off between random error (high resolution) and smearing (low resolution) of the resulting flux estimates. The upwind distance (perpendicular to the WSMA flight path) where 80 % of the flux contributions are included in the footprint, $L_{80\%} = 1171 \pm 314\,\mathrm{m}$, is comparable in magnitude to L_H. Thus, a flight path length of similar extent (1000 m) is a physically meaningful window for the computation of turbulence statistics and fluxes, because (i) changes in the turbulent flux (response) are resolved at the same spatial scale as the characteristic surface heterogeneity (driver); (ii) the turbulence statistics used for footprint calculations are representative on the same spatial scale as the upwind extent; and (iii) the random error for each flux estimate decreases by $\approx 70\,\%$ compared to a window length of 90 m.

The (aerodynamic) roughness length is usually below 1 m, with exception of the low wind speed situation on 17 July 2009, pattern O11, and the higher flight levels ($z \geq 97\,\mathrm{m}$) on 26 July 2009 (Table 2).

3.2 Flux un-mixing

The presentation of the flux un-mixing results follows the sequence of the LTFM analysis steps. In Sect. 3.2.1 the spatially resolved flux observations from the wavelet cross-scalogram are illustrated. Subsequently, footprint modelling is used to infer the biophysical surface properties in the source area of each flux observation (Sect. 3.2.2). In Sect. 3.2.3 the ERFs between flux observations and meteorological and land surface drivers are established. These ERFs are then used to predict the surface fluxes throughout the XRC, which are finally summarized for different land covers (Sect. 3.2.4).

3.2.1 Spatially resolved flux measurement

Here and in the following we use a flight along pattern O12 for illustration, which follows a shallow elevation gradient (Fig. 4 bottom panel). This flight pattern is particularly suitable for this purpose because of its marked land cover changes over a relatively short distance. The wavelet

Table 2. Mean length scales ±SD between repetitions during the WSMA flights selected for analysis. Shown are CBL depth z_i, aerodynamic roughness length z_0, flight altitude z, thermal blending height z_{TB1}, length scale of surface heterogeneity L_H, Raupach length L_R, the thermal blending lengths L_{TB1} and L_{TB2}, and the upwind distance from the WSMA $L_{80\%}$, where 80 % of the flux contributions are included in the flux footprint.

Date	Time (CST)	ID	z_i (m)	z_0 (m)	z (m)	z_{TB1} (m)	L_H (m)	L_R (m)	L_{TB1} (m)	L_{TB2} (m)	$L_{80\%}$ (m)
8 Jul 2009	10:20–10:50	O10	1100	0.21 ± 0.13	59 ± 4	45 ± 7	802 ± 192	1532 ± 105	1046 ± 171	255 ± 5	931 ± 52
	12:00–12:50	O12	1800	0.07 ± 0.05	72 ± 6	24 ± 2	700 ± 58	4093 ± 164	2160 ± 263	628 ± 39	1574 ± 147
13 Jul 2009	11:30–12:10	O8	1900	0.04 ± 0.05	51 ± 0	41 ± 4	1055 ± 5	4770 ± 3	1330 ± 122	999 ± 12	1440 ± 127
	12:40–13:10	O3	2100	0.05 ± 0.07	51 ± 2	38 ± 5	858 ± 95	4292 ± 236	1134 ± 45	1108 ± 220	1305 ± 64
15 Jul 2009	11:30–12:20	O11	2200	0.06 ± 0.04	55 ± 4	11 ± 2	370 ± 54	5214 ± 508	1950 ± 223	1368 ± 304	1470 ± 168
	12:30–13:00	O7	2100	0.26 ± 0.19	57 ± 7	17 ± 6	298 ± 76	3507 ± 82	984 ± 117	1207 ± 132	1081 ± 128
17 Jul 2009	11:00–11:30	O11	1400	1.13 ± 0.81	48 ± 1	22 ± 0	366 ± 37	1589 ± 65	788 ± 74	440 ± 28	720 ± 74
	12:20–13:00	O7	1400	0.05 ± 0.06	52 ± 2	19 ± 8	507 ± 219	3136 ± 169	1381 ± 54	950 ± 159	1296 ± 150
26 Jul 2009	12:50–15:30	C1	2500	1.75 ± 1.91	97 ± 4	85 ± 68	1458 ± 913	2626 ± 284	1974 ± 562	752 ± 196	991 ± 331
	13:10–15:10	C2	2500	2.30 ± 2.12	102 ± 5	83 ± 51	1459 ± 632	2430 ± 382	1921 ± 325	916 ± 310	945 ± 191
30 Jul 2009	11:00–13:30	C1	1600	0.48 ± 0.34	56 ± 3	46 ± 21	1653 ± 161	3097 ± 340	2403 ± 1438	1015 ± 22	1042 ± 100
	11:10–13:20	C2	1600	0.11 ± 0.01	54 ± 0	51 ± 13	2615 ± 84	3798 ± 735	2853 ± 601	1852 ± 51	1206 ± 0

cross-scalogram allows a high spatial discretization of turbulent flux measurements. At the same time it includes flux contributions from wavelength that are significantly longer than the 1000 m subinterval for each flux observation (Fig. 4). The resulting high number of flux observations along a flight line leads to previously unachievable resolution and coverage of the state space. Spatially coherent flux contributions are detected on transport scales (eddy sizes) of 500–2000 m, that is of similar size as the extent of homogeneous surface patches L_H. Strong local flux contributions are confined to scales < 500 m, and approximately decay within the lower threshold of the observed blending lengths L_{TB1} and L_{TB2}. This confirms a close coupling between atmospheric turbulence structures with surface patchiness, and consolidates the interpretation of the length-scale approach. The less certain flux contributions above the wavelet COI are small (-15 to -4 % median differences for all flights). In the present example the COI is confined to relatively small scales (≤ 4 km), which is a direct result of the comparatively short flight. In general, more certain flux contributions below the COI include transport scales up to $\approx 1/3$ of the flight length, and can reach ≈ 16 km for flight patterns C1 and C2. However, this also implies that the maximum considered transport scale differs between the flight patterns, just as it would be the case for the time-domain EC method. The wavelet cross-scalogram reveals strong turbulent transport in the second and fourth quarter of the flight for H, and in the first and third quarter for LE (Fig. 4). When integrated over all transport scales for each overflown 90 m cell of the land cover grid, these patterns correspond to strong upward fluxes.

3.2.2 Land cover

In Sect. 3.2.1, turbulence statistics and fluxes were integrated for each overflown 90 m cell of the land cover grid. In the following we expand the integration window to overlapping subintervals of 1000 m length, while retaining a spatial discretization of 90 m. Such a procedure significantly reduces the random sampling error (Sect. 3.1), though at the cost of decreasing the number of resulting observations by one window size ($dN \approx 10$). The resulting turbulence statistics are used to calculate the source area of each individual flux observation along the flight line, which are superimposed over the land cover grids. Figure 6 shows that in general LST and EVI follow the land cover patterns, e.g. lower temperature and higher greenness for irrigated agriculture and marshland. However, it is also evident that the static land cover classification cannot reflect the current surface conditions. For example, the marshland in the north-western quadrant appears dried-out (high LST and low EVI), while the steppe area in the north-eastern quadrant shows large variations in LST. Hence, biophysical surface properties also vary significantly within the land cover classes. This is likely a function of geomorphological properties such as aspect, slope and soil type, but also due to the large variability of convective rainfall events across the study area (e.g. Schaffrath et al., 2011).

Following superimposition of the footprints over the land cover data, the spatial contributions of different surface properties to each flux observation can be quantified (Fig. 7). It is evident that measured Bo changes in correspondence with the dominating land cover, i.e. low Bo for marshland and irrigated agriculture, and high Bo for bare soil and steppe. LST and EVI are stratified between the land covers, although in different sequence compared to the regional average (Fig. 3). The variability of LST and EVI within the land cover classes is equal to or larger than the between-class variability, in particular for marshland, irrigated and rainfed agriculture. While LST and EVI behave inversely for all natural land covers ($-0.78 < r < -0.10$), the contrary is true for irrigated ($r = 0.92$) and rainfed ($r = 0.30$) agriculture.

The latter finding appears counter-intuitive, but can be explained by tillage farming in the low-level plains with crops that are not adapted to the semiarid climate, such as potatoes. The albedo of these densely vegetated crops can be lower compared to the sparsely vegetated steppe land cover ($\alpha \approx 0.2$, Ketzer et al., 2008), resulting in higher foliage temperatures. Only two natural land covers, marshland and mountain meadow, exhibit similarly high EVI values as the field crops (Figs. 2, 3). Nevertheless, the LST of these land covers is comparatively low. In case of the marshland this can be explained by water-saturated soils with high heat capacity. Conversely, lower temperatures in accordance with the adiabatic temperature gradient are expected for the mountain meadows at higher altitudes.

In Fig. 8 H and LE observations along the flight line are shown together with the LST and EVI in the respective source area. Because of the 1000 m integration window over the wavelet cross-scalogram, the results appear smoother compared to Fig. 4, where a 90 m integration window is used. It is apparent that H and LE both systematically change with LST ($r_H = 0.64$, $r_{LE} = -0.84$) and EVI ($r_H = -0.62$, $r_{LE} = 0.73$). However, peaks in H (3 km and 9 km in Fig. 8) and in LE (0 km and 7 km) do not manifest when their respective land surface drivers in the footprint are maximal. Instead they seem to follow a trade-off function between LST and EVI.

3.2.3 Environmental response functions

Thus far our findings indicate that the interactions between land surface and atmosphere are multi-facetted and potentially non-linear. Hence we use LST and EVI as topical, spatio-temporal proxies for the source strength of H and LE, rather than using the land cover classification directly. In comparison to earlier flux un-mixing studies (Chen et al., 1999; Hutjes et al., 2010; Ogunjemiyo et al., 2003), this has the benefit of (i) providing individual source strength representations for the effects of surface moisture and temperature, and (ii) representing the land surface by continuous (LST, EVI) rather than discrete variables (land cover classes), thus enabling the use of more advanced scaling algorithms.

Here, we use BRTs to extract the relationships between all ($N = 8446$) flux observations and land cover (LST, EVI) and meteorological ($S\downarrow$, MR, and θ) variables. While BRTs are capable of reproducing complex interactions through multilayered branching, the fitted function can be summarized, e.g. as partial dependence plots (Fig. 9). These show the effect of each individual variable on the response after (i) subtraction of the offset ($H_0 = 161\,W\,m^{-2}$, $LE_0 = 176\,W\,m^{-2}$), and (ii) accounting for the average effects of all other variables in the model. The partial dependence plots in Fig. 9 are sorted in order of the relative importance of the response variables (Friedman, 2001). The most important responses of H are non-linear (LST, θ), followed by linear responses ($S\downarrow$, MR, and EVI). With the exception of MR and EVI, the individual responses are positive in sign. The order of the responses for LE is partially different (MR, LST, θ, $S\downarrow$, and EVI), and only the responses on $S\downarrow$ and EVI are approximately linear (not shown). With exception of MR (concave, maximal response around $10\,g\,kg^{-1}$) and LST (convex, minimal response around $\approx 310\,K$), the signs of the responses for LE are positive. It appears surprising that H and LE are only weakly related to $S\downarrow$. This can be explained by using only noontime flights in the present study, where $S\downarrow$ mainly fulfils the purpose of accounting for varying cloud/radiation conditions between different measurement days. In addition, during individual flights $S\downarrow$ was usually constant to within $\leq 10\,\%$ (Table B1). However, when using ERFs to reproduce a diurnal cycle, a much larger dependence of H and LE on $S\downarrow$ would be expected.

In Fig. 10 MLFRs are established between BRT fitted values for H and LE and the observed fluxes ($N = 8446$). Here we use the BRT cross-validation residuals and the random sampling errors in the observations to determine the MLFR weights of each data point. Uncertainty terms (i), (iii), (v) and (vi) (Sect. 2.5) cannot be quantified individually for each observation. Hence these terms are not considered here, but in the final uncertainty budget (Tables 3 and 4). For both H and LE the agreement between the BRT fitted values and the observed fluxes is excellent. Contrary to our initial anticipation, the ERFs are not attenuated by the relatively coarse MODIS LST resolution, as indicated by approximately zero MLFR offset and unity slope. The median absolute deviation in the residuals is small ($\leq 1\,\%$). However, several outliers are found for moderate to high fluxes of H ($N = 41$) and LE ($N = 133$), for which the BRTs underestimate the observed value by $-150\,W\,m^{-2}$ or more. The majority of these cases occur during the flights O8 on 13 July 2009 and C1 on 26 July 2009, respectively. On both dates the outliers concur with highly intermittent solar irradiance ($200 < S\downarrow < 1200\,W\,m^{-2}$) along a short section of the flight paths. For instance an intermittent cloud cover can disrupt the functional relation between the irradiance (driver) and the flux (response) observations, because, (i) at a flight level of 50–100 m a.g.l., the aircraft irradiance measurement does not represent $S\downarrow$ in the source areas of H and LE, and (ii) the plant physiological response can vary substantially on spatio-temporal scales that are small compared to atmospheric transport processes between the land surface and the aircraft.

Our choice of land surface and meteorological drivers appears to work well for describing the noontime surface–atmosphere exchange of heat and water vapour over a moisture-limited landscape. However, it is important to note that appropriately describing exchange processes over longer periods of time, for different landscapes or scalars might require finding an entirely different set of predictors.

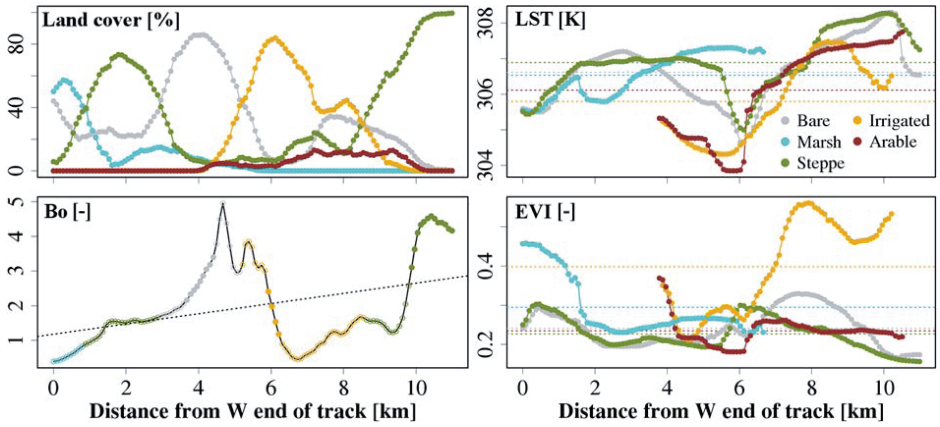


Fig. 7. Biophysical surface properties in the footprint of each observation ($N = 124$) along the flight pattern O12 on 8 July 2009, 12:16–12:24 CST, summarized by land cover. Shown are (clockwise from top right panel) land surface temperature, enhanced vegetation index, Bowen ratio, and the land cover fraction in the footprint. The dashed lines are land cover averages for LST and EVI, and the spatial trend for Bo. The land cover colour code and corresponding abbreviations are identical with Fig. 2.

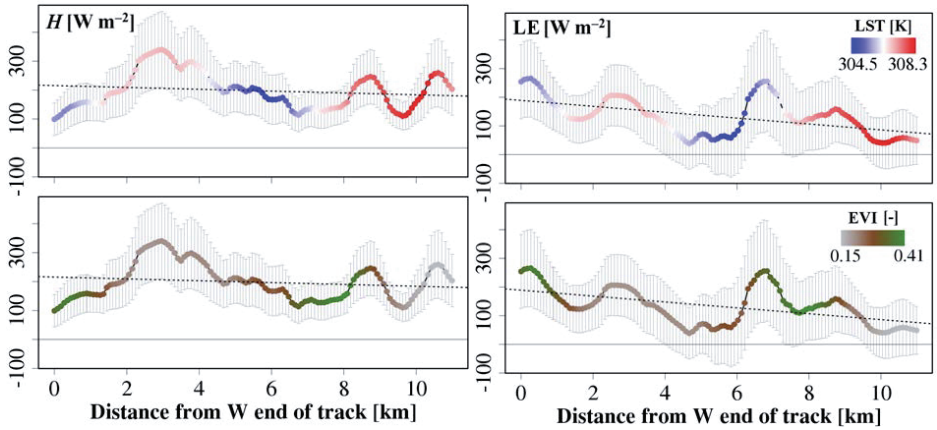


Fig. 8. Sensible heat flux (left panels) and latent heat flux (right panels) along the flight pattern O12 on 8 July 2009, 12:16–12:24 CST. Also shown is the random sampling error (error bars) for each observation ($N = 124$), and the spatial trend (dashed line). The top and bottom panels show the land surface temperature and the enhanced vegetation index in the footprint of each observation, respectively.

3.2.4 Extrapolation and summarization

For the duration of each flight pattern, the trained BRT models are used to extrapolate H and LE throughout the XRC. For this purpose the median meteorological state variables during each flight pattern as well as topical grids of MODIS LST and EVI data are used. Grid cells that exceed the state space of the BRT training dataset ($N = 8446$) are excluded from extrapolation. Fig. 11 shows the resulting flux grids for three different days, with a spatial coverage of $\geq 92\,\%$. Because of the identical state space ranges for BRT training and prediction, also the ranges of the extrapolated turbulent fluxes are within limits of the observations. Despite that the land cover classification was never used during the extrapolation process, several landscape units are clearly recognizable in the flux maps. For instance bot, the Xilin River valley and the mountainous headwater area to the east display low H and LE. On the contrary, the non-vegetated basin on the northern tip shows consistently low evapotranspiration.

For a given meteorological boundary condition (MR, θ, $S\downarrow$), the heat fluxes within several hours of solar zenith can be expressed as a function of LST and EVI (Fig. 9). In turn, these biophysical surface properties are characteristic within a land cover class (Fig. 3). Here, we aggregate all grid cells of the flux maps according to land cover class, resulting in sample distributions of H and LE. This allows

Table 3. Median land cover specific flux estimates of H and LE from the LTFM procedure over all flight patterns ± median spatial variability within the respective land cover. Also shown are the corresponding median ensemble random uncertainties $\sigma_{ens}(H)$, σ_{ens}(LE) and land cover specific sample size N.

Land cover	H ($W m^{-2}$)	LE ($W m^{-2}$)	$\sigma_{ens}(H)$	σ_{ens}(LE)	N
Bare soil	193 ± 32	136 ± 38	1 %	1 %	22049
Sand dunes	188 ± 37	144 ± 55	1 %	1 %	43424
Marshland	125 ± 55	230 ± 59	1 %	1 %	20722
Steppe	202 ± 40	138 ± 46	< 1 %	< 1 %	321956
Mountain meadow	114 ± 47	260 ± 69	1 %	1 %	25175
Settlements	172 ± 40	155 ± 46	3 %	5 %	1404
Rainfed agriculture	183 ± 35	147 ± 40	1 %	1 %	17024
Irrigated agriculture	116 ± 32	224 ± 41	5 %	5 %	1068

Table 4. Median systematic- and random uncertainty terms (in parentheses) for a single flux observation or grid cell throughout the LTFM procedure.

Source	H	LE
Instrumentation and hardware	0 % (8 %)	0 % (7 %)
Turbulence sampling	0 % (57 %)	0 % (121 %)
Spatio-temporal analysis	2 % (40 %)	4 % (47 %)
BRT residuals	0 % (5 %)	0 % (6 %)
BRT response function	11 % (69 %)	18 % (77 %)
BRT state variables	13 % (77 %)	14 % (75 %)

a formal transition from a mosaic- to a tile representation of H and LE over the XRC for the duration of each flight pattern (Mengelkamp et al., 2006). These sample distributions then enable the analysis of land cover specific source strengths ($36 W m^{-2} < H < 364 W m^{-2}$, $46 W m^{-2} <$ LE $< 425 W m^{-2}$), as well as the spatial variability within a land cover ($11 W m^{-2} < \sigma_H < 169 W m^{-2}$, $14 W m^{-2} < \sigma_{LE} < 152 W m^{-2}$). Table 3 gives an overview of the median land cover specific H and LE over all flight patterns, and their median spatial variability. These results fall well within the range of summertime ensemble average fluxes during solar noon observed by ground-based EC measurements over different land covers in this region ($100 W m^{-2} < H < 310 W m^{-2}$ and $100 W m^{-2} <$ LE $< 480 W m^{-2}$; Gao et al., 2009; Hao et al., 2007; Hao et al., 2008; Shao et al., 2008). In comparison, the flight-line average heat fluxes are in the range of $71 W m^{-2} < H < 310 W m^{-2}$ and $46 W m^{-2} <$ LE $< 300 W m^{-2}$ (Table B2).

However, the magnitudes of H and LE are not only functions of land cover, but also proportional to the available energy. The available energy changes within, but in particular between flight days. To alleviate this effect and to enable the comparison between different flights, we calculate the Bowen ratio Bo $= H/$LE between the sample distributions. Despite differences in the meteorological drivers (MR, θ, $S\downarrow$), the median land cover specific Bo agrees well between subsequent flight patterns on all measurement days (Fig. 12).

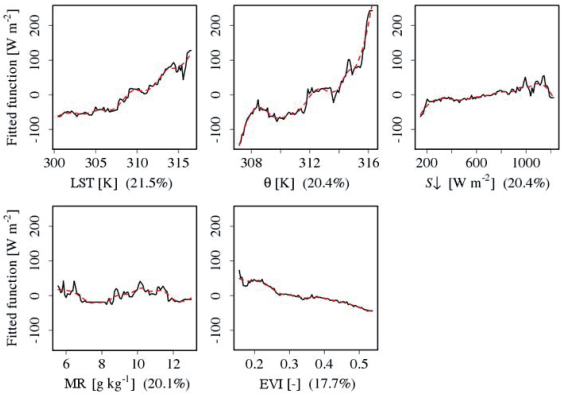


Fig. 9. Boosted regression tree partial response plots of H for all five state variables in order of their relative importance (in braces). The fitted function (black) shows the variable response of the BRT over the range of one individual state variable, while the remaining state variables are held at an average, constant value. The red dashed line is a smoothed representation of the fitted function (locally weighted polynomial regression).

During the afternoon flights, 12 ± 9 % higher Bo values are observed compared to the morning flights, as expected from a land surface that desiccates in the course of the day. Nevertheless, the 99.9 % confidence interval includes unity slope. Hence, for several hours within solar zenith Bo does not change significantly, and can be interpreted as a characteristic land surface property. On this basis we summarize the regional flux estimates for the duration of the flight campaign as time series of land cover specific Bo ratios (Fig. 12). The order of Bo between the land covers follows the order of the land cover specific EVI approximately inversely, while the temporal pattern follows the pattern of the land cover specific LST (Fig. 3). High Bo values until mid-campaign indicate that the land surface dries out. This trend is reversed toward the end of the campaign, when the approach of humid air masses leads to considerable precipitation. The median daily

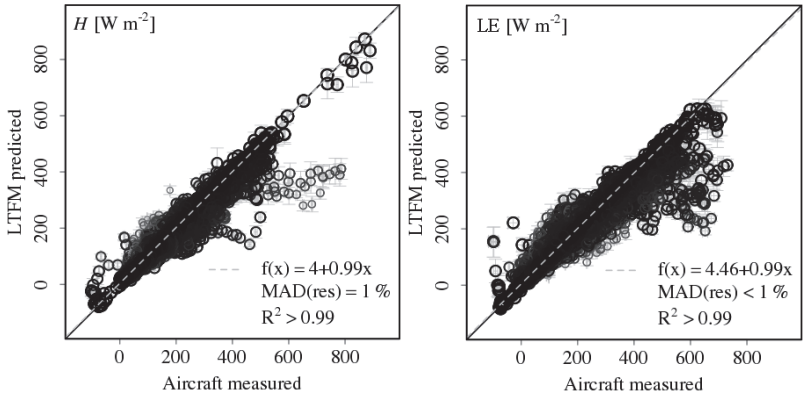


Fig. 10. Maximum likelihood functional relationships between $N = 8446$ aircraft observation and LTFM predictions of sensible heat flux (left) and latent heat flux (right). The weight of each data point in the relationship is represented by the size of the circles. The error bars show the cross-validation residuals for the LTFM predictions, and the ensemble random sampling error for the aircraft measurement. The 99.9 % confidence intervals are too narrow to be displayed properly.

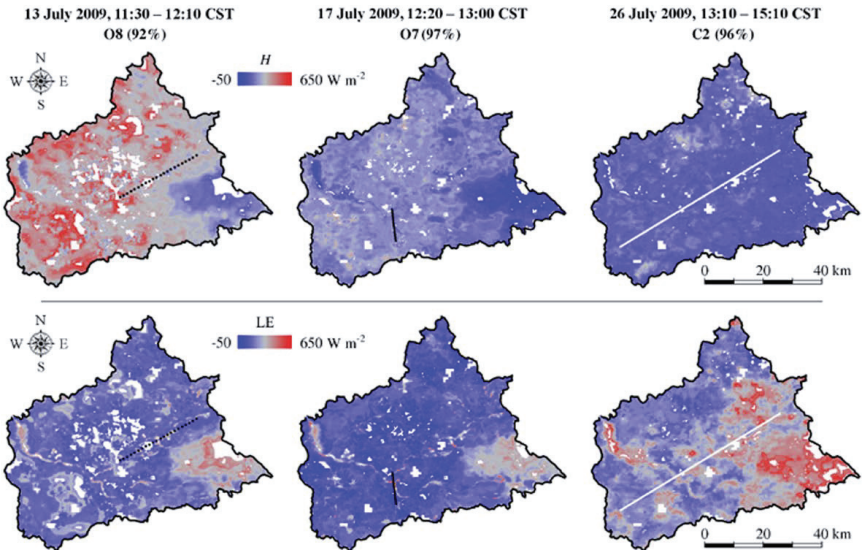


Fig. 11. Maps of the LTFM predicted fluxes of sensible heat (H, top) and latent heat (LE, bottom) on 13, 17 and 26 July 2009 (left to right). The colour gradient from blue over grey to red represents values that are lower, equal to, or greater than the average of the values, respectively (see legend). Percentages in braces after the flight ID indicate the spatial coverage of the prediction throughout the catchment. Meteorological state variables from the superimposed flight lines are used in the respective LTFM prediction (illustration identical with Fig. 2).

natural variability of Bo within the land covers ranges from 48 % (rainfed agriculture) to 79 % (marshland). Water absorbs strongly in the near infrared, leading to negative EVI values that are not indicative of vegetation greenness. Hence EVI values for water surfaces are discarded, and the land cover "water" cannot be modelled by the present ERFs.

3.3 Uncertainty

Metzger et al. (2012) have shown that turbulent flux measurements with the WSMA platform and instrumentation are unbiased, and precise to within 8 %. Uncertainty due to the limited sampling size of turbulent eddies is estimated using the methods of Lenschow and Stankov (1986) and Lenschow et al. (1994). Details on the implementation can be found in Metzger et al. (2012). For a single flux measurement, the

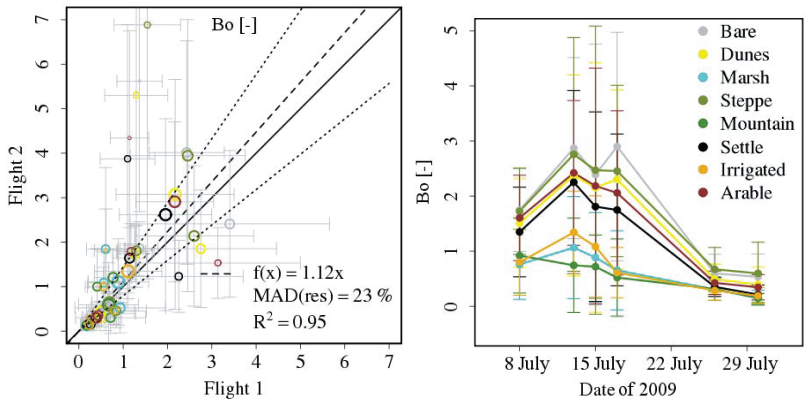


Fig. 12. Left: MLFR of Bowen ratio between the first and the second flight pattern on every measurement day. The weight of each data point in the relationship is represented by the size of the circles. Right: time series of Bo for different land covers throughout the measurement campaign. In both images the error bars represent the Gaussian sum of the natural variability in each land cover class and the ensemble random error in the LTFM procedure. The land cover colour code and corresponding abbreviations are identical with Fig. 2.

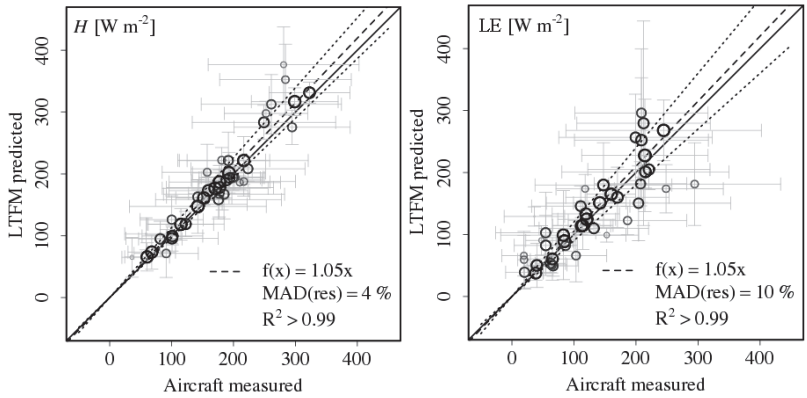


Fig. 13. MLFRs of median observed and predicted fluxes along 42 flight lines. The error bars correspond to the variability of the fluxes along the flight line, and the weight of each data point in the relationship is represented by the size of the circles.

systematic (and random) components of this sampling uncertainty range from $< 1\,\%$ (57 %) for H to $< 1\,\%$ (121 %) for LE. Table 4 summarizes above uncertainty sources, as well as additional sources which are discussed in the following.

In order to assess the uncertainty arising from the spatio-temporal analyses (Sect. 2.5.1), we compare the median observed and predicted fluxes along all flight legs (Fig. 13). The LTFM predictions slightly overestimate the observed fluxes ($H = 5\,\%$, LE $= 5\,\%$), but in both cases the 99.9 % confidence intervals include unity slope. The median differences of d$H = 2\,\%$ (40 %), and dLE $= 4\,\%$ (47 %) agree marginally more closely. Moreover, the median residuals between fitted and observed values emphasize that the BRT fitting technique is unbiased (Table 4).

Subsequently, we assess the predictive performance of the BRT response function in light of missing state variable combinations in the training data. For this purpose one flight at a time was omitted from the training data, and the incompletely trained BRT model was used to predict the missing data. The resulting median differences amount to 11 % (69 %, $N = 7311$) for H and 18 % (77 %, $N = 7265$) for LE. During prediction, cases where one or more state variables exceed their respective range during training were excluded. As a consequence the sample size is $\approx 14\,\%$ smaller than the total number of observations ($N = 8466$).

Lastly, we consider the uncertainty resulting from disregarding part of the spatio-temporal variability in the state variables during BRT predictions. For this purpose we quantify the disregarded parts of the natural variability, and propagate it through the full BRT model. The resulting median differences amount to 13 % (77 %) and 14 % (75 %) for H and LE, respectively, and are dominated by the effect

of LST natural variability ($r = 0.81$, and -0.69). Because the response of the BRT predictions on LST is non-linear (Fig. 9), deviations of similar magnitude but opposite sign in LST do not cancel out in the predictions. This can lead to a systematic overestimation as a function of the specific state variable combination in each prediction, and is hence dependent on the catchment composition. However, in all test cases the 99.9 % confidence intervals between observed and predicted fluxes include unity slope. Hence we go without introducing a non-linearity response factor, but assign an accuracy of ≤ 20 % to the LTFM method.

Assuming normal distribution and independence, the random parts of all uncertainty terms (Table 4, in parentheses) can be combined to their Gaussian sum. Then, the ensemble random uncertainty σ_{ens} considers the reduction of the random uncertainty with sample size (e.g. Mahrt, 1998);

$$\sigma_{\text{ens}} = \frac{\sigma_{\text{ran}}}{\sqrt{N}}, \tag{10}$$

with zero expected value σ_{ens} and the SD σ_{ran} of the population with size N. While σ_{ran} is a measure for the average dispersion of a single observation or grid cell, σ_{ens} quantifies the level of confidence we can expect from aggregating multiple observations or grid cells. The resulting ensemble random uncertainty for land cover specific flux estimates throughout the XRC ranges from < 1 % for steppe to 5 % for settlements and irrigated agriculture (Table 3).

4 Conclusions

The overarching goal of airborne EC flux measurements is to bridge the gap between observations and data assimilation approaches on different spatial scales. This study develops the LTFM procedure to characterize the exchange of sensible and latent heat for different land covers in a heterogeneous steppe landscape. The procedure "mines" the information content of EC flux observations and extracts quantitative relationships with environmental drivers. In the process LTFM maximises objectivity and data use efficiency – all available observations are considered. The subsequent steps of LTFM are (1) low level EC flux flights, (2) time–frequency analysis of the flux observations, (3) source area modelling of continuous biophysical surface properties, and (4) inferring ERFs from non-parametric machine learning.

(i) The use of a weight-shift microlight aircraft with low airspeed and high climb rate enables low level flights at constant height even above topographically structured terrain. Masking out slopes during flight planning effectively minimizes cross-contamination of the flux observations by slope-induced effects on radiative transfer or turbulence generation. This reduces the degrees of freedom in explaining the observed flux responses, albeit potentially at the expense of oversimplifying surface–air exchange processes.

(ii) Wavelet decomposition of the turbulence data yields unprecedented spatial resolution of the flux observations. However, due to edge effects flux observations close to the start or end of a dataset can contain spectral artefacts. Using alternative techniques such as empirical mode decomposition (Barnhart et al., 2012a, b) or structure-parameter methods (Van Kesteren et al., 2013) might help to further improve the results.

(iii) An "offline" footprint parameterization considering 3-D dispersion is suitable to map the differences in surface properties encountered by a flux measuring aircraft. However, when adapting LTFM e.g. to ground-based measurements, the range of surface properties is likely to shrink significantly. In order to improve the decreased signal-to-noise ratio, it might become important to also consider the local flow field, especially when measuring at greater heights. For example, closure models with terrain-following coordinates (Hsieh and Katul, 2009; Sogachev and Lloyd, 2004) or "online" Lagrangian dispersion modelling (Markkanen et al., 2010; Matross et al., 2011; Wang and Rotach, 2010; Weil et al., 2012) could be useful for such a purpose.

(iv) Instead of a static and discrete land cover classification, the LTFM method uses spatio-temporally continuous and topical information of biophysical surface properties. Only the continuous nature of MODIS land surface data enabled the use of the BRT machine learning technique. In this combination the climatic and altitudinal gradients throughout the XRC are successfully reproduced. In the interest of further advancing LTFM, it is desirable to also consider the uncertainty in the observations during machine learning, and to explore alternative machine learning techniques such as support vector machines (e.g. Yang et al., 2007).

The ERFs resulting from LTFM can aid bridging observational scales, e.g. by isolating and quantifying relevant land–atmosphere exchange processes, estimating land cover specific emission factors, extending flux measurements to the catchment scale, assessing the spatial representativeness of EC flux measurements, etc. Analogously applying LTFM to ground-based EC measurements could aid, e.g. advancing the treatment of location bias from diagnostic assessment (e.g. Chen et al., 2012) to prognostic transfer functions, constraining local to regional water budgets, distinguishing anthropogenic and natural sources/sinks in urban environments and substantiating process-studies.

Appendix A

Notation

A1 Abbreviations

3-D	Three-dimensional
a.g.l.	Above ground level
a.s.l.	Above sea level
Arable	Rainfed agriculture
ASTER	Advanced Spaceborne Thermal and Reflection Radiometer
Bare	Bare soil
BRT	Boosted regression tree
CBL	Convective boundary layer
COI	Cone of influence
CST	Chinese standard time (CST = coordinated universal time + 8)
CV	Cross-validation
EC	Eddy covariance
ERF	Environmental response function
IMGERS	Inner Mongolia Grassland Ecosystem Research Station
Irrigated	Irrigated agriculture
KL04	Footprint parameterisation of Kljun et al. (2004)
KL04+	Footprint parameterisation of Kljun et al. (2004) with superimposed cross-wind dispersion function
LTFM	Low level flights, time–frequency-, footprint-, and machine learning analyses as influenced by stocking rate
Marsh	Marshland
MLFR	Maximum likelihood functional relationship
MODIS	Moderate Resolution Imaging Spectroradiometer
Mountain	Mountain meadow
SD	Standard deviation
Settle	Settlements
Steppe	Generic steppe
WSMA	Weight-shift microlight aircraft
XRC	Xilin River catchment

A2 Functions

Overbars denote the mean along a flight line, and primes denote the deviations from this mean.

*	Complex conjugate
cov	Covariance
d	Difference
σ	Standard deviation
ψ	Mother wavelet

A3 Parameters and variables

α	Albedo (–)
a	Wavelet scale parameter (s)
a_0	Initial wavelet scale parameter (s)
b	Wavelet location parameter (s)
Bo	Bowen ratio (–)
c	Lag of autocorrelation function (m)
CC	Cloud cover (–)
C_δ	Wavelet reconstruction factor (–)
d	Distance along a flight line (m)
DIR	Wind direction (°)
δj	Wavelet frequency increment (–)
δt	Wavelet time increment (s)
e	Euler's number ≈ 2.71828 (–)
EVI	Enhanced vegetation index (–)
H	Sensible heat flux ($\mathrm{W\,m^{-2}}$)
i	Imaginary unit $i^2 = -1$ (–)
ID	Flight identifier (–)
j	Running index (–)
J	Number of wavelet scale increments (–)
l	Length of flight line (km)
L	Monin–Obukhov length (m)
$L_{80\%}$	Upwind distance where 80 % of the flux contributions are included in the footprint (m)
LC	Land cover class coverage (%)
LE	Latent heat flux ($\mathrm{W\,m^{-2}}$)
LST	Land surface temperature (K)
L_H	Horizontal scale of surface heterogeneity (m)
L_R	Raupach length (m)
L_TB1	Thermal blending length (m)
L_TB2	Improved thermal blending length (m)
MR	Mixing ratio ($\mathrm{g\,kg^{-1}}$)
n	Running index (–)
N	Sample size (–)
p	Probability of test statistic (–)
P	Cumulated precipitation in a 10-day trailing window (mm)
q	Dimensionless wavelet coordinate (–)
r	Pearson correlation coefficient (–)
rep	Repetitions (–)
res	residuals (Depending on variable)
$S\downarrow$	Down-welling shortwave radiation ($\mathrm{W\,m^{-2}}$)
t	Time (s)
T_s	Surface temperature (K)
θ	Potential temperature (K)
$\theta_{0,\,v}$	Virtual potential temperature (K)
u	Horizontal wind speed ($\mathrm{m\,s^{-1}}$)
u_*	Friction velocity ($\mathrm{m\,s^{-1}}$)
w	Footprint weight (–)
W	Wavelet coefficient (–)
w_*	Convective velocity ($\mathrm{m\,s^{-1}}$)
x	Wildcard for a signal (–)
y	Wildcard for a signal (–)
z	Measurement height (m)
z_0	Aerodynamic roughness length (m)
ω_0	Wavelet frequency parameter (–)
z_i	Convective boundary layer depth (m)
z_TB1	Thermal blending height (m)

Table B1. Mean meteorological conditions ±SD between repetitions during the WSMA flights selected for analysis. Shown are cloud cover CC, shortwave down-welling radiation $S\downarrow$, mixing ratio MR, horizontal wind speed u, wind direction DIR, virtual potential temperature θ_v, surface temperature T_S, and the SD of the surface temperature σ_{TS}.

Date	Time (CST)	ID	CC	$S\downarrow$ ($W m^{-2}$)	MR ($g kg^{-1}$)	u ($m s^{-1}$)	DIR (°)	θ_v (K)	T_S (°C)	σ_{TS} (K)
8 Jul 2009	10:20–10:50	O10	7/8	842 ± 55	6.4 ± 0.0	3.2 ± 0.3	221 ± 12	312.0 ± 0.4	40.8 ± 1.2	8.1 ± 0.4
	12:00–12:50	O12	7/8	773 ± 53	5.9 ± 0.2	6.5 ± 0.3	320 ± 2	313.7 ± 0.3	39.8 ± 1.0	8.8 ± 0.4
13 Jul 2009	11:30–12:10	O8	4/8	810 ± 16	9.3 ± 0.1	8.3 ± 0.4	291 ± 6	309.7 ± 0.1	41.9 ± 0.3	6.4 ± 0.1
	12:40–13:10	O3	4/8	838 ± 6	8.6 ± 0.1	6.9 ± 0.6	297 ± 6	311.2 ± 0.3	45.8 ± 0.2	5.3 ± 0.8
15 Jul 2009	11:30–12:20	O11	7/8	796 ± 72	7.1 ± 0.1	7.0 ± 0.8	253 ± 9	315.3 ± 0.2	42.1 ± 1.1	5.3 ± 0.7
	12:30–13:00	O7	7/8	843 ± 56	6.8 ± 0.0	5.8 ± 0.2	255 ± 8	316.7 ± 0.2	50.8 ± 0.7	4.0 ± 0.5
17 Jul 2009	11:00–11:30	O11	7/8	589 ± 39	9.4 ± 0.1	2.7 ± 0.2	102 ± 5	309.3 ± 0.2	35.9 ± 0.5	4.8 ± 0.3
	12:20–13:00	O7	7/8	682 ± 122	11.2 ± 0.2	5.9 ± 0.4	144 ± 4	310.4 ± 0.1	40.3 ± 2.5	4.5 ± 0.9
26 Jul 2009	12:50–15:30	C1	7/8	668 ± 46	9.6 ± 0.3	2.9 ± 0.1	174 ± 5	312.8 ± 0.5	36.4 ± 1.2	4.8 ± 0.7
	13:10–15:10	C2	7/8	747 ± 67	9.1 ± 0.1	2.7 ± 0.3	178 ± 23	313.0 ± 0.4	36.5 ± 0.5	3.7 ± 0.7
30 Jul 2009	11:00–13:30	C1	7/8	715 ± 82	11.6 ± 0.3	4.3 ± 0.7	159 ± 15	311.8 ± 0.9	34.6 ± 5.1	4.1 ± 0.5
	11:10–13:20	C2	7/8	567 ± 11	11.6 ± 0.0	4.9 ± 0.9	154 ± 9	311.3 ± 1.2	32.8 ± 1.0	2.7 ± 0.5

Table B2. Mean turbulence statistics ±SD between repetitions during the WSMA flights selected for analysis. Shown are friction velocity u_*, sensible heat flux H, latent heat flux LE, Monin–Obukhov length L, SD of vertical wind σ_w, and convective velocity w_*.

Date	Time (CST)	ID	u_* ($m s^{-1}$)	H ($W m^{-2}$)	LE ($W m^{-2}$)	L (m)	σ_w ($m s^{-1}$)	w_* ($m s^{-1}$)
8 Jul 2009	10:20–10:50	O10	0.31 ± 0.03	154 ± 20	194 ± 45	−14 ± 3	0.88 ± 0.03	1.82 ± 0.08
	12:00–12:50	O12	0.46 ± 0.07	199 ± 14	110 ± 36	−38 ± 16	0.97 ± 0.04	2.29 ± 0.05
13 Jul 2009	11:30–12:10	O8	0.52 ± 0.09	290 ± 47	196 ± 43	−40 ± 24	0.99 ± 0.05	2.65 ± 0.13
	12:40–13:10	O3	0.42 ± 0.12	288 ± 18	86 ± 51	−23 ± 19	0.95 ± 0.01	2.70 ± 0.07
15 Jul 2009	11:30–12:20	O11	0.47 ± 0.07	176 ± 17	138 ± 30	−46 ± 23	0.83 ± 0.05	2.36 ± 0.07
	12:30–13:00	O7	0.52 ± 0.12	310 ± 28	65 ± 12	−38 ± 19	1.07 ± 0.11	2.77 ± 0.08
17 Jul 2009	11:00–11:30	O11	0.39 ± 0.07	156 ± 26	46 ± 67	−30 ± 15	0.79 ± 0.04	1.93 ± 0.09
	12:20–13:00	O7	0.41 ± 0.07	206 ± 18	48 ± 45	−27 ± 13	0.83 ± 0.06	2.11 ± 0.07
26 Jul 2009	12:50–15:30	C1	0.33 ± 0.09	120 ± 33	300 ± 76	−20 ± 10	0.94 ± 0.11	2.25 ± 0.21
	13:10–15:10	C2	0.37 ± 0.05	117 ± 11	227 ± 2	−29 ± 8	0.94 ± 0.04	2.22 ± 0.06
30 Jul 2009	11:00–13:30	C1	0.43 ± 0.13	107 ± 76	194 ± 27	−57 ± 14	0.75 ± 0.12	1.78 ± 0.41
	11:10–13:20	C2	0.38 ± 0.05	71 ± 4	223 ± 11	−50 ± 19	0.68 ± 0.04	1.65 ± 0.02

A4 Subscripts

In general, subscripts follow the parameter and variable definitions in Appendix A3. Instances with differing use of subscripts are defined in the following.

ens	Ensemble
ran	Random
v	Cross-wind component
w	Vertical wind component
x	Longitudinal coordinate
y	Latitudinal coordinate

Appendix B

Meteorological conditions

The midday flights are usually accompanied by a thin layer of cirrus clouds, interspersed with local convective cumuli, resulting in a cloud cover between 4/8 and 7/8 (Table B1). The down-welling shortwave radiation decreases over the duration of the campaign, with minima on 17 and 30 July 2009. These minima coincide with the advection of comparatively moist air, as evident from the higher mixing ratios. $S\downarrow$ also correlates with the precipitation history (Table 1, $r = -0.40$). The wind speed at flight level decays from up to 8.3 $m s^{-1}$ at the beginning down to 2.7 $m s^{-1}$ towards the last quarter of the flight campaign. All wind sectors with the exception

of northerlies occur. Both the virtual potential air temperature and the surface temperature peak during the middle of the flight campaign. As a result of several convective precipitation events, the mixing ratio increases over the flight campaign, accompanied by a dampening of the surface temperature variability.

The ranges of the flight line average turbulent fluxes are $0.3 < u_* < 0.5\,\mathrm{m\,s^{-1}}$, $71\,\mathrm{W\,m^{-2}} < H < 310\,\mathrm{W\,m^{-2}}$ and $46\,\mathrm{W\,m^{-2}} < \mathrm{LE} < 300\,\mathrm{W\,m^{-2}}$ (Table B2). The friction velocity peaks during flights under high wind speeds. While H dominates the heat exchange during the middle of the campaign, LE peaks at the beginning and end of the campaign. The Bowen ratio throughout the campaign correlates ($r = -0.67$) with precipitation history (Table 1), i.e. the moisture available for evapotranspiration. Moreover, H clearly correlates with $S\downarrow$ ($r = 0.68$), while no such relationship was found for LE ($r = 0.02$). The atmospheric stratification was unstable throughout all flights (Monin–Obukhov length $L = -34 \pm 20\,\mathrm{m}$), with corresponding high values of the SD of the vertical wind $\sigma_w = 0.88 \pm 0.11\,\mathrm{m\,s^{-1}}$ and convective velocity $w_* = 2.21 \pm 0.39\,\mathrm{m\,s^{-1}}$.

Acknowledgements. We wish to thank Henry Loescher and Jeffrey Taylor at the National Ecological Observatory Network, Fundamental Instrument Unit, for their continued support and statistical advice. The National Ecological Observatory Network is a project sponsored by the National Science Foundation and managed under cooperative agreement by NEON, Inc. This material is based upon work supported by the National Science Foundation under the grant DBI-0752017. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. Thanks go to Matthias Reiche at the Leibniz Centre for Agricultural Landscape Research, Institute of Soil Landscape Research, who provided the ASTER imagery used in this study. Recognition has to be given to Benjamin Blank, formerly at Justus-Liebig-University Giessen, Institute of Landscape Ecology and Resources Management, who prepared the overview Fig. 1. The MODIS data were obtained through the online Data Pool at the NASA Land Processes Distributed Active Archive Center, USGS/Earth Resources Observation and Science Centre (https://lpdaac.usgs.gov/get_data). The flight campaign in Inner Mongolia was funded by the German Research Foundation, research group FOR 536 MAGIM, "Matter fluxes in grasslands of Inner Mongolia as influenced by stocking rate", and the National Natural Science Foundation of China, grant number 41021004. Stipend funding by the German Academic Exchange Service, Helmholtz Association of German Research Centres, China Scholarship Council and the European Union under the Science and Technology Fellowship China is acknowledged. We thank Ankur Desai, Raymond Desjardins and Paul Stoy for their valuable comments in the review process.

The service charges for this open access publication have been covered by a Research Centre of the Helmholtz Association.

Edited by: P. Stoy

References

Anderson, M. C., Kustas, W. P., Alfieri, J. G., Gao, F., Hain, C., Prueger, J. H., Evett, S., Colaizzi, P., Howell, T., and Chávez, J. L.: Mapping daily evapotranspiration at Landsat spatial scales during the BEAREX'08 field campaign, Adv. Water Res., 50, 162–177, doi:10.1016/j.advwatres.2012.06.005, 2012.

Auerswald, K., Wittmer, M. H. O. M., Männel, T. T., Bai, Y. F., Schäufele, R., and Schnyder, H.: Large regional-scale variation in C3/C4 distribution pattern of Inner Mongolia steppe is revealed by grazer wool carbon isotope composition, Biogeosciences, 6, 795–805, doi:10.5194/bg-6-795-2009, 2009.

Baldocchi, D., Falge, E., Gu, L., Olson, R., Hollinger, D., Running, S., Anthoni, P., Bernhofer, C., Davis, K., Evans, R., Fuentes, J., Goldstein, A., Katul, G., Law, B., Lee, X., Malhi, Y., Meyers, T., Munger, W., Oechel, W., U., K., Pilegaard, K., Schmid, H., Valentini, R., Verma, S., Vesala, T., Wilson, K., and Wofsy, S.: FLUXNET: A new tool to study the temporal and spatial variability of ecosystem-scale carbon dioxide, water vapor, and energy flux densities, Bull. Am. Meteorol. Soc., 82, 2415–2434, doi:10.1175/1520-0477(2001)082<2415:FANTTS>2.3.CO;2, 2001.

Bange, J., Spiess, T., Herold, M., Beyrich, F., and Hennemuth, B.: Turbulent fluxes from Helipod flights above quasi-homogeneous patches within the LITFASS area, Bound.-Lay. Meteorol., 121, 127–151, doi:10.1007/s10546-006-9106-0, 2006.

Barnhart, B. L., Eichinger, W. E., and Prueger, J. H.: Introducing an Ogive method for discontinuous data, Agr. Forest. Meteorol., 162–163, 58–62, doi:10.1016/j.agrformet.2012.04.003, 2012a.

Barnhart, B. L., Eichinger, W. E., and Prueger, J. H.: A new eddy-covariance method using empirical mode decomposition, Bound.-Lay. Meteorol., 145, 369–382, doi:10.1007/s10546-012-9741-6, 2012b.

Beyrich, F., Leps, J. P., Mauder, M., Bange, J., Foken, T., Huneke, S., Lohse, H., Ludi, A., Meijninger, W. M. L., Mironov, D., Weisensee, U., and Zittel, P.: Area-averaged surface fluxes over the LITFASS region based on eddy-covariance measurements, Bound.-Lay. Meteorol., 121, 33–65, doi:10.1007/s10546-006-9052-x, 2006.

Businger, J. A., Wyngaard, J. C., Izumi, Y., and Bradley, E. F.: Flux-profile relationships in the atmospheric surface layer, J. Atmos. Sci., 28, 181–189, doi:10.1175/1520-0469(1971)028<0181:FPRITA>2.0.CO;2, 1971.

Butterbach-Bahl, K., Kögel-Knabner, I., and Han, X.: Steppe ecosystems and climate and land-use changes – vulnerability, feedbacks and possibilities for adaptation, Plant Soil, 340, 1–6, doi:10.1007/s11104-010-0651-4, 2011.

Chen, B., Coops, N. C., Fu, D., Margolis, H. A., Amiro, B. D., Black, T. A., Arain, M. A., Barr, A. G., Bourque, C. P. A., Flanagan, L. B., Lafleur, P. M., McCaughey, J. H., and Wofsy, S. C.: Characterizing spatial representativeness of flux tower eddy-covariance measurements across the Canadian Carbon Program Network using remote sensing and footprint analysis, Remote

Sens. Environ., 124, 742–755, doi:10.1016/j.rse.2012.06.007, 2012.

Chen, J. M., Leblanc, S. G., Cihlar, J., Desjardins, R. L., and MacPherson, J. I.: Extending aircraft- and tower-based CO_2 flux measurements to a boreal region using a Landsat thematic mapper land cover map, J. Geophys. Res. Atmos., 104, 16859–16877, doi:10.1029/1999JD900129, 1999.

Chen, Z. Z.: Topography and climate of the Xilin river basin, Research on grassland ecosystem, 3, 13–22, 1988.

Croux, C. and Rousseeuw, P. J.: Time-efficient algorithms for two highly robust estimators of scale, Computation. Stat., 1, 411–428, 1992.

Deardorff, J. W.: Three-dimensional numerical study of turbulence in an entraining mixed layer, Bound.-Lay. Meteorol., 7, 199–226, doi:10.1007/bf00227913, 1974.

Desjardins, R. L., MacPherson, J. I., Schuepp, P. H., and Hayhoe, H. N.: Airborne flux measurements of CO_2, sensible, and latent heat over the Hudson Bay lowland, J. Geophys. Res. Atmos., 99, 1551–1561, doi:10.1029/93JD01296, 1994.

Desjardins, R. L., Macpherson, J. I., Neumann, H., Den Hartog, G., and Schuepp, P. H.: Flux estimates of latent and sensible heat, carbon dioxide, and ozone using an aircraft-tower combination, Atmos. Environ., 29, 3147–3158, doi:10.1016/1352-2310(95)00007-L, 1995.

Desjardins, R. L., MacPherson, J. I., Mahrt, L., Schuepp, P., Pattey, E., Neumann, H., Baldocchi, D., Wofsy, S., Fitzjarrald, D., McCaughey, H., and Joiner, D. W.: Scaling up flux measurements for the boreal forest using aircraft-tower combinations, J. Geophys. Res. Atmos., 102, 29125–29133, doi:10.1029/97JD00278, 1997.

Elith, J., Leathwick, J. R., and Hastie, T.: A working guide to boosted regression trees, Journal of Animal Ecology, 77, 802–813, doi:10.1111/j.1365-2656.2008.01390.x, 2008.

Emeis, S., Schäfer, K., and Münkel, C.: Surface-based remote sensing of the mixing-layer height – a review, Meteorol. Z., 17, 621–630, doi:10.1127/0941-2948/2008/0312, 2008.

Fan, L., Liu, S., Bernhofer, C., Liu, H., and Berger, F. H.: Regional land surface energy fluxes by satellite remote sensing in the Upper Xilin River Watershed (Inner Mongolia, China), Theor. Appl. Climatol., 88, 231–245, doi:10.1007/s00704-006-0241-9, 2007.

Friedman, J. H.: Greedy function approximation: A gradient boosting machine, Ann. Stat., 29, 1189–1232, doi:10.1214/aos/1013203451, 2001.

Gao, Z., Lenschow, D. H., He, Z., Zhou, M.: Seasonal and diurnal variations in moisture, heat and CO_2 fluxes over a typical steppe prairie in Inner Mongolia, China, Hydrol. Earth Syst. Sci., 13, 987–998, doi:10.5194/hess-13-987-2009, 2009.

Gioli, B., Miglietta, F., De Martino, B., Hutjes, R. W. A., Dolman, H. A. J., Lindroth, A., Schumacher, M., Sanz, M. J., Manca, G., Peressotti, A., and Dumas, E. J.: Comparison between tower and aircraft-based eddy covariance fluxes in five European regions, Agr. Forest. Meteorol., 127, 1–16, doi:10.1016/j.agrformet.2004.08.004, 2004.

Glenn, E. P., Huete, A. R., Nagler, P. L., and Nelson, S. G.: Relationship between remotely-sensed vegetation Indices, canopy attributes and plant physiological processes: What vegetation indices can and cannot tell us about the landscape, Sensors, 8, 2136–2160, doi:10.3390/s8042136, 2008.

Hao, Y., Wang, Y., Huang, X., Cui, X., Zhou, X., Wang, S., Niu, H., and Jiang, G.: Seasonal and interannual variation in water vapor and energy exchange over a typical steppe in Inner Mongolia, China, Agric. For. Meteorol., 146, 57–69, doi:10.1016/j.agrformet.2007.05.005, 2007.

Hao, Y., Wang, Y., Mei, X., Huang, X., Cui, X., Zhou, X., and Niu, H.: CO_2, H_2O and energy exchange of an Inner Mongolia steppe ecosystem during a dry and wet year, Acta Oecologica, 33, 133–143, doi:10.1016/j.actao.2007.07.002, 2008.

Helmis, C., Sgouros, G., Tombrou, M., Schäfer, K., Münkel, C., Bossioli, E., and Dandou, A.: A comparative study and evaluation of mixing-height estimation based on SODAR-RASS, ceilometer data and numerical model simulations, Bound.-Lay. Meteorol., 145, 507–526, doi:10.1007/s10546-012-9743-4, 2012.

Högström, U.: Non-dimensional wind and temperature profiles in the atmospheric surface layer: A re-evaluation, Bound.-Lay. Meteorol., 42, 55–78, doi:10.1007/BF00119875, 1988.

Holmes, R. M.: Meso-scale effects of agriculture and a large prairie lake on the atmospheric boundary layer, Agron. J., 62, 546–549, doi:10.2134/agronj1970.00021962006200040037x, 1970.

Hsieh, C.-I. and Katul, G.: The Lagrangian stochastic model for estimating footprint and water vapor fluxes over inhomogeneous surfaces, Int. J. Biometeorol., 53, 87–100, doi:10.1007/s00484-008-0193-0, 2009.

Hu, F. S., Higuera, P. E., Walsh, J. E., Chapman, W. L., Duffy, P. A., Brubaker, L. B., and Chipman, M. L.: Tundra burning in Alaska: Linkages to climatic change and sea ice retreat, J. Geophys. Res., 115, G04002, doi:10.1029/2009jg001270, 2010.

Hutjes, R. W. A., Vellinga, O. S., Gioli, B., and Miglietta, F.: Dis-aggregation of airborne flux measurements using footprint analysis, Agr. Forest. Meteorol., 150, 966–983, doi:10.1016/j.agrformet.2010.03.004, 2010.

Jiang, S.: An introduction to the Inner Mongolia Grassland Ecosystem Research Station, Research on grassland ecosystem, 1, 1–11, 1985.

Kaharabata, S. K., Schuepp, P. H., Ogunjemiyo, S., Shen, S., Leclerc, M. Y., Desjardins, R. L., and MacPherson, J. I.: Footprint considerations in BOREAS, J. Geophys. Res. Atmos., 102, 29113–29124, doi:10.1029/97JD02559, 1997.

Kaminski, T., Rayner, P. J., Voßbeck, M., Scholze, M., and Koffi, E.: Observing the continental-scale carbon balance: Assessment of sampling complementarity and redundancy in a terrestrial assimilation system by means of quantitative network design, Atmos. Chem. Phys., 12, 7867–7879, doi:10.5194/acp-12-7867-2012, 2012.

Ketzer, B., Liu, H., and Bernhofer, C.: Surface characteristics of grasslands in Inner Mongolia as detected by micrometeorological measurements, Int. J. Biometeorol., 52, 563–574, doi:10.1007/s00484-008-0148-5, 2008.

Kirby, S., Dobosy, R., Williamson, D., and Dumas, E.: An aircraft-based data analysis method for discerning individual fluxes in a heterogeneous agricultural landscape, Agr. Forest. Meteorol., 148, 481–489, doi:10.1016/j.agrformet.2007.10.011, 2008.

Kljun, N., Rotach, M. W., and Schmid, H. P.: A three-dimensional backward lagrangian footprint model for a wide range of boundary-layer stratifications, Bound.-Lay. Meteorol., 103, 205–226, doi:10.1023/A:1014556300021, 2002.

Kljun, N., Calanca, P., Rotach, M. W., and Schmid, H. P.: A simple parameterisation for flux footprint predictions, Bound.-Lay. Meteorol., 112, 503–523, 2004.

Kljun, N., Rotach, M. W., and Schmid, H. P.: A simple parameterisation for two-dimensional flux footprints, in preparation, 2013.

Kvalseth, T. O.: Cautionary note about R^2, Am. Stat., 39, 279–285, 1985.

Lenschow, D. H. and Stankov, B. B.: Length scales in the convective boundary layer, J. Atmos. Sci., 43, 1198–1209, doi:10.1175/1520-0469(1986)043<1198:LSITCB>2.0.CO;2, 1986.

Lenschow, D. H., Mann, J., and Kristensen, L.: How long is long enough when measuring fluxes and other turbulence statistics?, J. Atmos. Oceanic Technol., 11, 661–673, doi:10.1175/1520-0426(1994)011<0661:HLILEW>2.0.CO;2, 1994.

Liang, E., Shao, X., Hu, Y., and Lin, J.: Dendroclimatic evaluation of climate-growth relationships of Meyer spruce (*Picea meyeri*) on a sandy substrate in semi-arid grassland, North China, Trees-Struct. Funct., 15, 230–235, doi:10.1007/s004680100097, 2001.

Liu, C., Holst, J., Brüggemann, N., Butterbach-Bahl, K., Yao, Z., Han, S., Han, X., and Zheng, X.: Effects of irrigation on nitrous oxide, methane and carbon dioxide fluxes in an Inner Mongolian steppe, Adv. Atmos. Sci., 25, 748–756, doi:10.1007/s00376-008-0748-3, 2008.

Lyons, T. J. and Halldin, S.: Surface heterogeneity and the spatial variation of fluxes, Agr. Forest. Meteorol., 121, 153–165, doi:10.1016/j.agrformet.2003.08.031, 2004.

Mahrt, L.: Flux sampling errors for aircraft and towers, J. Atmos. Oceanic Technol., 15, 416–429, doi:10.1175/1520-0426(1998)015<0416:FSEFAA>2.0.CO;2, 1998.

Mahrt, L.: Surface heterogeneity and vertical structure of the boundary layer, Bound.-Lay. Meteorol., 96, 33–62, doi:10.1023/a:1002482332477, 2000.

Markkanen, T., Steinfeld, G., Kljun, N., Raasch, S., and Foken, T.: A numerical case study on footprint model performance under inhomogeneous flow conditions, Meteorol. Z., 19, 539–547, doi:10.1127/0941-2948/2010/0488, 2010.

Mason, P. J.: The formation of areally-averaged roughness lengths, Q. J. R. Meteorolog. Soc., 114, 399–420, doi:10.1002/qj.49711448007, 1988.

Matross, D. M., Andrews, A., Pathmathevan, M., Gerbig, C., Lin, J. C., Wofsy, S. C., Daube, B. C., Gottlieb, E. W., Chow, V. Y., Lee, J. T., Zhao, C., Bakwin, P. S., Munger, J. W., and Hollinger, D. Y.: Estimating regional carbon exchange in New England and Quebec by combining atmospheric, ground-based and satellite data, Tellus B, 58, 344–358, doi:10.1111/j.1600-0889.2006.00206.x, 2011.

Mauder, M., Desjardins, R. L., and MacPherson, I.: Scale analysis of airborne flux measurements over heterogeneous terrain in a boreal ecosystem, J. Geophys. Res. Atmos., 112, D13112, doi:10.1029/2006JD008133, 2007.

Meijninger, W. M. L., Beyrich, F., Lüdi, A., Kohsiek, W., and De Bruin, H. A. R.: Scintillometer-based turbulent fluxes of sensible and latent heat over a heterogeneous land surface – a contribution to LITFASS-2003, Bound.-Lay. Meteorol., 121, 89–110, doi:10.1007/s10546-005-9022-8, 2006.

Mengelkamp, H. T., Beyrich, F., Heinemann, G., Ament, F., Bange, J., Berger, F., Bosenberg, J., Foken, T., Hennemuth, B., Heret, C., Huneke, S., Johnsen, K. P., Kerschgens, M., Kohsiek, W., Leps, J. P., Liebethal, C., Lohse, H., Mauder, M., Meijninger, W., Raasch, S., Simmer, C., Spiess, T., Tittebrand, A., Uhlenbrock, J., and Zittel, R.: Evaporation over a heterogeneous land surface – the EVA-GRIPS project, Bull. Am. Meteorol. Soc., 87, 775–786, doi:10.1175/BAMS-87-6-775, 2006.

Metzger, S., Junkermann, W., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Measuring the 3-D wind vector with a weight-shift microlight aircraft, Atmos. Meas. Tech., 4, 1421–1444, doi:10.5194/amt-4-1421-2011, 2011.

Metzger, S., Junkermann, W., Mauder, M., Beyrich, F., Butterbach-Bahl, K., Schmid, H. P., and Foken, T.: Eddy-covariance flux measurements with a weight-shift microlight aircraft, Atmos. Meas. Tech., 5, 1699–1717, doi:10.5194/amt-5-1699-2012, 2012.

Meurer, M. and Jiang, Y.: Die Steppen Nordchinas und ihre Belastung durch weide- und landwirtschaftliche Landnutzung, Geographische Rundschau, 53, 48–52, 2001.

Münkel, C. and Roininen, R.: Automatic monitoring of boundary layer structures and depth with ceilometer using a novel robust algorithm, International Symposium for the Advancement of Boundary Layer Remote Sensing, Paris, France, 2010.

Nagler, P. L., Glenn, E. P., Kim, H., Emmerich, W., Scott, R. L., Huxman, T. E., and Huete, A. R.: Relationship between evapotranspiration and precipitation pulses in a semi-arid rangeland estimated by moisture flux towers and MODIS vegetation indices, J. Arid. Environ., 70, 443–462, doi:10.1016/j.jaridenv.2006.12.026, 2007.

Ogunjemiyo, S. O., Kaharabata, S. K., Schuepp, P. H., MacPherson, I. J., Desjardins, R. L., and Roberts, D. A.: Methods of estimating CO_2, latent heat and sensible heat fluxes from estimates of land cover fractions in the flux footprint, Agr. Forest. Meteorol., 117, 125–144, doi:10.1016/S0168-1923(03)00061-3, 2003.

Oke, T. R.: Boundary layer climates, Taylor & Francis, London, 435 pp., 1987.

Qi, Y.-C., Dong, Y.-S., Liu, J.-Y., Domroes, M., Geng, Y.-B., Liu, L.-X., Liu, X.-R., and Xiao, H. Y.: Effect of the conversion of grassland to spring wheat field on the CO_2 emission characteristics in Inner Mongolia, China, Soil Tillage Res., 94, 310–320, doi:10.1016/j.still.2006.08.008, 2007.

Quenouille, M. H.: Notes on bias in estimation, Biometrika, 43, 353–360, doi:10.1093/biomet/43.3-4.353, 1956.

R Development Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, 2012.

Raupach, M. R. and Finnigan, J. J.: Scale issues in boundary-layer meteorology: Surface energy balances in heterogeneous terrain, Hydrol. Processes, 9, 589–612, doi:10.1002/hyp.3360090509, 1995.

Ridgeway, G.: gbm V1.6-3.2 – Generalized boosted regression models, available under: http://cran.r-project.org/web/packages/gbm/, 2012.

Ripley, B. D. and Thompson, M.: Regression techniques for the detection of analytical bias, Analyst, 112, 377–383, doi:10.1039/an9871200377, 1987.

Rousseeuw, P. J. and Verboven, S.: Robust estimation in very small samples, Computation. Stat. Data An., 40, 741–758, doi:10.1016/S0167-9473(02)00078-6, 2002.

Schaffrath, D., Barthold, F., and Bernhofer, C.: Spatiotemporal variability of grassland vegetation cover in a catchment in Inner Mongolia, China, derived from MODIS data products, Plant Soil, 340, 181–198, doi:10.1007/s11104-010-0465-4, 2011.

Schmid, H. P.: Footprint modeling for vegetation atmosphere exchange studies: A review and perspective, Agr. Forest. Meteorol., 113, 159–183, doi:10.1016/S0168-1923(02)00107-7, 2002.

Schuepp, P. H., MacPherson, J. I., and Desjardins, R. L.: Adjustment of footprint correction for airborne flux mapping over the FIFE site, J. Geophys. Res. Atmos., 97, 18455–18466, doi:10.1029/92JD00884, 1992.

Shao, C., Chen, J., Li, L., Xu, W., Chen, S., Gwen, T., Xu, J., and Zhang, W.: Spatial variability in soil heat flux at three Inner Mongolia steppe ecosystems, Agric. For. Meteorol., 148, 1433–1443, doi:10.1016/j.agrformet.2008.04.008,2008.

Sogachev, A. and Lloyd, J.: Using a one-and-a-half order closure model of the atmospheric boundary layer for surface flux footprint estimation, Bound.-Lay. Meteorol., 112, 467–502, 2004.

Sorbjan, Z.: Statistics of scalar fields in the atmospheric boundary layer based on large-eddy simulations. Part 2: Forced convection, Bound.-Lay. Meteorol., 119, 57–79, doi:10.1007/s10546-005-9014-8, 2006.

Steffens, M., Kölbl, A., Totsche, K. U., and Kögel-Knabner, I.: Grazing effects on soil chemical and physical properties in a semiarid steppe of Inner Mongolia (China), Geoderma, 143, 63–72, doi:10.1016/j.geoderma.2007.09.004, 2008.

Strunin, M. A. and Hiyama, T.: Applying wavelet transforms to analyse aircraft-measured turbulence and turbulent fluxes in the atmospheric boundary layer over Eastern Siberia, Hydrol. Process., 18, 3081–3098, doi:10.1002/hyp.5750, 2004.

Strunin, M. A., Hiyama, T., Asanuma, J., and Ohata, T.: Aircraft observations of the development of thermal internal boundary layers and scaling of the convective boundary layer over non-homogeneous land surfaces, Bound.-Lay. Meteorol., 111, 491–522, 2004.

Stull, R. B.: An Introduction to Boundary Layer Meteorology, Kluwer Academic Publishers, Dordrecht, 670 pp., 1988.

Thomas, R. M., Lehmann, K., Nguyen, H., Jackson, D. L., Wolfe, D., and Ramanathan, V.: Measurement of turbulent water vapor fluxes using a lightweight unmanned aerial vehicle system, Atmos. Meas. Tech., 5, 243–257, doi:10.5194/amt-5-243-2012, 2012.

Torrence, C. and Compo, G. P.: A practical guide to wavelet analysis, Bull. Am. Meteorol. Soc., 79, 61–78, doi:10.1175/1520-0477(1998)079<0061:apgtwa>2.0.co;2, 1998.

Tukey, J. W.: Bias and confidence in not-quite large samples, Ann. Math. Stat., 29, 614, 1958.

Van Kesteren, B., Hartogensis, O. K., van Dinther, D., Moene, A. F., and De Bruin, H. A. R.: Measuring H_2O and CO_2 fluxes at field scales with scintillometry: Part 1 – Introduction and validation of four methods, Agric. For. Meteorol., in press, doi:10.1016/j.agrformet.2012.09.013, 2013.

Vellinga, O. S., Gioli, B., Elbers, J. A., Holtslag, A. A. M., Kabat, P., and Hutjes, R. W. A.: Regional carbon dioxide and energy fluxes from airborne observations using flight-path segmentation based on landscape characteristics, Biogeosciences, 7, 1307–1321, doi:10.5194/bg-7-1307-2010, 2010.

Vesala, T., Kljun, N., Rannik, U., Rinne, J., Sogachev, A., Markkanen, T., Sabelfeld, K., Foken, T., and Leclerc, M. Y.: Flux and concentration footprint modelling: State of the art, Environ. Pollut., 152, 653–666, doi:10.1016/j.envpol.2007.06.070, 2008.

Vetter, S., Schaffrath, D., and Bernhofer, C.: Spatial simulation of evapotranspiration of semi-arid Inner Mongolian grassland based on MODIS and eddy covariance data, Environ. Earth Sci., 65, 1567–1574, doi:10.1007/s12665-011-1187-5, 2012.

Vickers, D. and Mahrt, L.: Quality control and flux sampling problems for tower and aircraft data, J. Atmos. Oceanic Technol., 14, 512–526, doi:10.1175/1520-0426(1997)014<0512:QCAFSP>2.0.CO;2, 1997.

Wan, Z. and Li, Z. L.: Radiance-based validation of the V5 MODIS land-surface temperature product, Int. J. Remote Sens., 29, 5373–5395, doi:10.1080/01431160802036565, 2008.

Wang, W. and Rotach, M. W.: Flux footprints over an undulating surface, Bound.-Lay. Meteorol., 136, 325–340, doi:10.1007/s10546-010-9498-8, 2010.

Webb, E. K., Pearman, G. I., and Leuning, R.: Correction of flux measurements for density effects due to heat and water vapour transfer, Q. J. R. Meteorolog. Soc., 106, 85–100, doi:10.1002/qj.49710644707, 1980.

Weil, J., Sullivan, P., Patton, E., and Moeng, C.-H.: Statistical variability of dispersion in the convective boundary layer: Ensembles of simulations and observations, Bound.-Lay. Meteorol., 145, 185–210, doi:10.1007/s10546-012-9704-y, 2012.

Wiesmeier, M., Barthold, F., Blank, B., and Kögel-Knabner, I.: Digital mapping of soil organic matter stocks using random forest modeling in a semi-arid steppe ecosystem, Plant Soil, 340, 7–24, doi:10.1007/s11104-010-0425-z, 2011.

Willett, J. B. and Singer, J. D.: Another cautionary note about R^2: Its use in weighted least-squares regression analysis, Am. Stat., 42, 236–238, 1988.

Wittmer, M. H. O. M., Auerswald, K., Bai, Y., Schäufele, R., and Schnyder, H.: Changes in the abundance of C3/C4 species of Inner Mongolia grassland: Evidence from isotopic composition of soil and vegetation, Global Change Biol., 16, 605–616, doi:10.1111/j.1365-2486.2009.02033.x, 2010.

Wood, N. and Mason, P.: The influence of static stability on the effective roughness lengths for momentum and heat transfer, Q. J. R. Meteorolog. Soc., 117, 1025–1056, doi:10.1002/qj.49711750108, 1991.

Xiang, G., Huete, A. R., and Didan, K.: Multisensor comparisons and validation of MODIS vegetation indices at the semiarid Jornada experimental range, IEEE Trans. Geosci. Remote Sens., 41, 2368–2381, doi:10.1109/tgrs.2003.813840, 2003.

Yang, F., Ichii, K., White, M. A., Hashimoto, H., Michaelis, A. R., Votava, P., Zhu, A. X., Huete, A., Running, S. W., and Nemani, R. R.: Developing a continental-scale measure of gross primary production by combining MODIS and AmeriFlux data through support vector machine approach, Remote Sens. Environ., 110, 109–122, doi:10.1016/j.rse.2007.02.016, 2007.

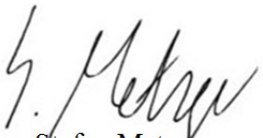
Ziehn, T., Knorr, W., and Scholze, M.: Investigating spatial differentiation of model parameters in a carbon cycle data assimilation system, Global Biogeochem. Cy., 25, GB2021, doi:10.1029/2010gb003886, 2011.

Erklärung

Hiermit erkläre ich, dass ich die Arbeit selbständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe.

Ferner erkläre ich, dass ich anderweitig mit oder ohne Erfolg nicht versucht habe, diese Dissertation einzureichen. Ich habe keine gleichartige Doktorprüfung an einer anderen Hochschule endgültig nicht bestanden.

Bayreuth, 2013-04-15



Stefan Metzger