Calibration and Quality Assurance of Flux Observations from a Small Research Aircraft*

OLAF S. VELLINGA

Earth System Science Group, Wageningen University and Research Centre, Wageningen, Netherlands

RONALD J. DOBOSY AND EDWARD J. DUMAS

Air Resources Laboratory, National Oceanic and Atmospheric Administration, Oak Ridge, Tennessee

BENIAMINO GIOLI

Institute of Biometeorology, National Research Council, Firenze, Italy

JAN A. ELBERS

Alterra, Wageningen University and Research Centre, Wageningen, Netherlands

RONALD W. A. HUTJES

Earth System Science Group, Wageningen University and Research Centre, Wageningen, Netherlands

(Manuscript received 26 August 2011, in final form 21 September 2012)

ABSTRACT

Small environmental research aircraft (ERA) are becoming more common for detailed studies of air–surface interactions. The Sky Arrow 650 ERA, used by multiple groups, is designed to minimize the complexity of high-precision airborne turbulent wind measurement. Its relative wind probe, of a nine-port design, is furthermore used with several other airplanes. This paper gives an overview of 1) calibration of the model that converts the probe's raw measurements to meteorological quantities; 2) quality control and assurance (QC–QA) in postprocessing of these quantities to compute fluxes; and 3) sensitivity of fluxes to errors in calibration parameters. The model, an adapted version of standard models of potential flow and aerodynamic upwash, is calibrated using an integrated method to derive a globally optimum set of parameters from in-flight maneuvers. Methods of QC–QA from the tower flux community are adopted for use with airborne flux data to provide more objective selection criteria for large datasets. Last, measurements taken from a standard operational flight are used to show fluxes to be most sensitive to calibration parameters that directly affect the vertical wind component. In another test with the same data, varying all calibration parameters simultaneously by $\pm 10\%$ of their optimum values, the model computes a response in the fluxes smaller than 10%, though a larger response may occur if only a subset of parameters is perturbed. A MATLAB toolbox has been developed that facilitates the procedures presented here.

* Supplemental information related to this paper is available at the Journals Online website: <http://dx.doi.org/10.1175/JTECH-D-11-00138.s1>.

Corresponding author address: O. S. Vellinga, Earth System Science Group, Wageningen University and Research Centre, P.O. Box 47, 6700 AA Wageningen, Netherlands.
E-mail: olaf.vellinga@wur.nl

1. Introduction

The land surface with its different land use types affects atmospheric processes at scales from local to global. The land surface and the climate system are related through a number of bidirectional feedbacks, involving surface fluxes of sensible and latent energy, greenhouse gases and momentum, planetary boundary layer dynamics, clouds, and free-atmosphere characteristics (van

FIG. 1. The Sky Arrow 650 TCNS (equivalent to the ERA model) with a mobile flux platform (see instrument labeling) on board. The PH-WUR is owned and operated by Wageningen University and Research Centre, the Netherlands.

Heerwaarden et al. 2009). Both numerical models and observations are used to study the behavior of these fluxes and their surface and atmospheric controls. To assess surface fluxes at regional scales, often vertical concentration profiles have been used in budgeting studies (Dolman et al. 1997), accounting for atmospheric transport as computed with forward (Sarrat et al. 2009) and inverse (Ciais et al. 2010) modeling techniques.

Small environmental research aircraft (SERA; Crawford et al. 2001) are increasingly being used to provide direct observations of surface fluxes with improved space and time resolution (e.g., Dolman et al. 2009), because they have some significant benefits over larger ones, such as the Twin Otter 300 or the Lockheed WP-3D Orion. Their smaller size and slower airspeed make them far less intrusive at low altitudes. They can often fly in the surface layer well below 60 m, hence measuring air–surface exchange that more closely resembles fluxes measured at tower sites. This assures minimum vertical flux divergence and smaller, better defined footprint areas.

Arguably, one of the more popular SERA is the Sky Arrow 650 environmental research aircraft (ERA; Fig. 1) with currently eight in operation by different research institutes. Its small dimensions and single-engine pusher make the airplane especially suited for airborne flux measurement. The ERA version of this airplane was designed to minimize the complexities of this type of measurement, enabling the use of a theoretical model with a minimum number of calibration parameters to be determined in flight or in a wind tunnel. Its turbulence probe, called Best Aircraft Turbulence (BAT) probe (Hacker and Crawford 1999), has both relative flow sensors and the probe's

motion sensors are collocated. Because of the airplane's design, it can be mounted on the nose tip well forward of the propeller and wing. This configuration is assumed to minimize sources of uncertainty caused by sidewash, large separation of relative velocity sensors from airplane motion sensors, and short distance from probe to wing.

The BAT probe developed by Air Resources Laboratory (ARL) of the National Oceanic and Atmospheric Administration (NOAA, United States) in collaboration with Airborne Research Australia (ARA, Australia) has become a widely used wind measurement system for SERA. It has a pressure hemisphere with nine pressure ports instead of the typical five holes (e.g., Rosemount-858AJ28; Brown et al. 1983), containing solid-state differential and absolute pressure transducers and three orthogonal accelerometers. It also holds a fast temperature probe. The hemisphere is mounted on a cylinder- or cone-shaped supporting tube that holds its electronics and optional additional instruments. The probe is usually mounted on the tip of the nose (e.g., Sky Arrow) or on a pod under one of the wings (e.g., HK36 ECO Dimona). Typically, together with a gas analyzer on or near the BAT probe, an inertial navigation system (INS)–GPS system, a second group of accelerometers at the center of gravity (CG) of the aircraft, auxiliary instrumentation (e.g., radiation sensors), and data acquisition hardware and software, it constitutes the “mobile flux platform” (MFP) system. The MFP is a proven concept (Dumas et al. 2001; Gioli et al. 2004) that gives the aircraft the ability to measure airborne fluxes of carbon dioxide ($f\text{CO}_2$), sensible heat (H), latent heat (λE), and momentum by the eddy covariance technique.

The MFP system is a tightly coupled and heavily interdependent set of instruments. These provide input to a potential flow model that converts them to meteorological quantities. The aircraft distorts the flow and pressure fields around the instruments as a function of airspeed and aircraft attitude. To account for these effects (in all but the most extreme maneuvers) along with any remaining misalignment and tilt of the instruments, the whole system needs to be calibrated in dedicated flight maneuvers. Descriptions of these flights and their corresponding procedures to determine the parameters can be found in peer-reviewed literature (e.g., Lenschow 1986; Bögel and Baumann 1991; Williams and Marcotte 2000; Garman et al. 2008), and in gray literature (e.g., Eckman et al. 1999; Isaac 2004; Hall et al. 2006). But perhaps the most comprehensive and elaborate description is found in the work by Leise and Masters (1993). The Sky Arrow's particular configuration features, that is, pusher propulsion and nose-mounted BAT probe, are all intended to simplify airborne flux measurement. This paper presents experience gained in calibrating its MFP system.

Quality control and quality assurance (QC–QA) are not only important for interpreting airborne flux data but also provide objective selection criteria for large datasets. Few publications about QC–QA for airborne flux data can be found. Some discuss the implemented method (e.g., Vickers and Mahrt 1997; Hall et al. 2006), others analyze typical errors or interpretation difficulties in airborne flux measurements (e.g., Mann and Lenschow 1994; Mahrt 1998). Unlike for flux data from ground stations [e.g., Aubinet et al. (2000), or guidelines of the AmeriFlux organization found at <http://public.ornl.gov/ameriflux>], a widely accepted standard methodology about QC–QA for airborne flux data is presently not available. However, SERA are being used more closely with tower-mounted instruments in detailed studies of air–surface interaction. This makes adoption of a method of QC–QA for airborne flux data similar to the tower flux community more attractive.

Hence, the objective of this paper is to provide a threefold overview of calibration and quality control for airborne flux data obtained from the Sky Arrow and the BAT probe. First, we present a procedure of calibration, which includes adapted versions of the flow model and the upwash model. Second, we present a procedure of postprocessing with a main focus on QC–QA adopted from the tower flux community. Third, we assess the sensitivity of the fluxes of carbon dioxide, momentum, and energy to variations of the wind model's parameters about their optimum values, determined by our calibrations.

2. Theory

a. Computing wind components

The eddy covariance technique requires time series of both the transported scalar quantity and the transporting turbulent wind, each measured at sufficient frequency to resolve the eddies. From an airplane, the turbulent wind is measured in a fast-moving accelerated reference frame having 6 degrees of freedom, but wind is defined in the quasi-inertial reference frame of the earth. Besides the true airspeed vector $\hat{\mathbf{U}}_a$ at the probe, the airplane's ground speed vector $\mathbf{U}_p$ and rotation vector $\hat{\boldsymbol{\omega}}_p$ must, therefore, be measured at sufficient frequency to resolve the airplane's motion. Typically, the same sample rate (e.g., 50 Hz) suffices both for eddy covariance and for the airplane's motions. This will be called the full sample rate (FSR). The turbulent wind $\mathbf{U}(t)$ is then the sum of the airflow at the probe and the probe's motion with respect to the earth:

$$\mathbf{U}(t) = \mathbf{G}(t)[\hat{\mathbf{U}}_a(t) + \hat{\boldsymbol{\omega}}_p(t) \times \hat{\mathbf{r}}] + \mathbf{U}_p(t). \quad (1)$$

The unadorned symbols in Eq. (1) are in earth coordinates (x to east, y to north, and z opposite gravity). The airplane's coordinates, denoted by $(\cdot)$, form a right-hand orthogonal system with $\hat{x}$ directed forward, $\hat{y}$ toward the port wing, and $\hat{z}$ “upward.” The $\mathbf{G}$ is the transformation matrix. The cross product in Eq. (1) can be ignored if the displacement $\hat{\mathbf{r}}$ between the INS–GPS's measurement ($\hat{\boldsymbol{\omega}}_p$, $\mathbf{U}_p$) and the probe's measurement ($\hat{\mathbf{U}}_a$) is negligible. This has been an advantage of the BAT probe, which collocates GPS and accelerometers with the wind sensors (Crawford and Dobosy 1992).

The rotation matrix $\mathbf{G}$ from airplane to earth coordinates is normally implemented in a sequence of two-dimensional rotations: roll ϕ , pitch θ , and heading ψ . These attitude angles are determined at the FSR by the INS–GPS system obtained by blending, as velocity vector. Offset corrections (ε_ϕ , ε_θ , ε_ψ) are introduced here to correct for possible misalignment of the INS–GPS instrument to the BAT probe's relative wind sensors. The generally small values of these correction constants are determined via dedicated flight maneuvers.

Wind vector $\hat{\mathbf{U}}_a(t)$ comes from sampling the distribution of pressure over the BAT probe's hemisphere at the FSR to resolve the complex turbulent motions in the airplane's reference frame. The motion $\mathbf{U}_p(t)$ of the probe itself is a blend. Low-frequency velocity (up to about 10 Hz) is directly reported in the earth's coordinates by the GPS system. Higher-frequency velocities are obtained by integrating acceleration measurements. The latter is derived in the airplane's coordinates and then rotated at the

FSR by $\mathbf{G}(t)$ before blending. The resulting $\mathbf{U}_p(t)$ and $\hat{\mathbf{U}}_a(t)$ are combined as in Eq. (1), assuming $\hat{\mathbf{r}} = 0$.

The BAT probe derives $\hat{\mathbf{U}}_a$ from three differential pressures and one absolute pressure (see Fig. 2 in Garman et al. 2006): $\delta p_z = p_2 - p_4$ in the $\hat{x} - \hat{z}$ plane, $\delta p_y = p_3 - p_1$ in the $\hat{x} - \hat{y}$ plane, and $\delta p_x = p_0 - \bar{p}_r$ between the central pressure port, p_0 , and the absolute reference pressure, $\bar{p}_r$. For the BAT probe, $\bar{p}_r$ is a hydraulically derived average ($\bar{\cdot}$) from the four diagonal ports, approximately the ambient atmospheric pressure.

The accuracy of the hydraulic average has been established in multiple studies by in-flight calibrations, as described below, and by comparison with measurements from towers and other aircraft. Recent wind tunnel tests (Dobosy et al. 2013) confirmed these field results and revealed other aspects of the BAT probe's operation.

The derivation uses a potential flow model described in depth by Leise and Masters (1993) to provide the magnitude and direction of the oncoming flow. The direction of $\hat{\mathbf{U}}_a$ is the angular displacement of the stagnation point from the BAT probe's central port. It has two components: α , the angle of attack ($\hat{x}, \hat{z}$); and β , the angle of sideslip ($\hat{x}, \hat{y}$).

Leise and Masters (1993) provide several variations of the equations used to determine these angles, depending on information available for input. The BAT probe, which provides its own absolute reference pressure $\bar{p}_r$, uses a further variation (Eckman 1999):

$$\begin{aligned}\tan \alpha &= 2 \frac{F_z}{Q}, \\ \tan \beta &= 2 \frac{F_y}{Q}, \\ Q &= 1 + \sqrt{1 + 2(F_y^2 + F_z^2)},\end{aligned}\quad (2)$$

where auxiliary quantities

$$\begin{aligned}F_y &= \frac{\sin^2 \xi_r}{4 \sin \xi \cos \xi} \frac{\delta p_y}{\delta p_x \varepsilon_q} \equiv C_\beta \frac{\delta p_y}{\delta p_x \varepsilon_q}, \\ F_z &= \frac{\sin^2 \xi_r}{4 \sin \xi \cos \xi} \frac{\delta p_z}{\delta p_x \varepsilon_q} \equiv C_\alpha \frac{\delta p_z}{\delta p_x \varepsilon_q},\end{aligned}\quad (3)$$

are defined to simplify the notation. Here, ξ is the angle between p_0 and each of the pressure ports in the $(\hat{x}, \hat{z})$ plane or in the $(\hat{x}, \hat{y})$ plane, and ξ_r , which is the angle between p_0 and each of the ports accounting for $\bar{p}_r$. The ε_q is an empirical correction factor for δp_x .

All outer pressure ports of the BAT probe are located at the same angle of 45° from the central pressure port ($\xi = \xi_r = 45^\circ$). The first terms at the right-hand side of Eq. (3) correspond then to a theoretical value of 0.250.

They are replaced by coefficients C_α and C_β as shown by the formulas after the equivalent sign, to account for departures from the ideal potential flow model, and need to be determined by dedicated flight maneuvers. As a result, the constants represented by Eq. (2) differ from Eckman (1999). Correction factor ε_q , ideally unity, accounts for any small remaining imperfections in the pressure hemisphere, especially the placement and configuration of the ports. It is speed dependent, determined by a dedicated flight maneuver.

Next, the dynamic pressure (q) and the ambient pressure (p_s) are determined:

$$q = \frac{2}{k_p} \frac{\delta p_x \varepsilon_q}{\sin^2 \xi_r} \left[\frac{1 + \tan^2 \alpha + \tan^2 \beta}{2 - \tan^2 \alpha - \tan^2 \beta} \right], \quad (4)$$

$$p_s = \bar{p}_r - q \left\{ k_p \left[\frac{2 + \sin^2 \xi (\tan^2 \alpha + \tan^2 \beta - 2)}{2(1 + \tan^2 \alpha + \tan^2 \beta)} - 1 \right] + 1 \right\}, \quad (5)$$

where $k_p = 2.07$ and is an empirical constant reported by Traub and Rediniotis (2003), accounting for the effects of the cylinder on the hemisphere. For a standard potential flow model of a sphere, this constant is $3/4$. Note that in Eq. (4), δp_x is also corrected by ε_q .

From the dynamic and ambient pressure, one can compute the Mach number (M) as

$$M^2 = \frac{2}{\gamma - 1} \left[\left(1 + \frac{q}{p_s} \right)^{(\gamma-1)/\gamma} - 1 \right], \quad (6)$$

where $\gamma = 1.4$ is the ratio of the specific heats at constant pressure and constant volume. This provides the ambient temperature T , given the measured temperature T_{obs} , contaminated by adiabatic heating as oncoming air is compressed around the temperature probe (e.g., John 1984, chapter 17):

$$T = T_{\text{obs}} \left(1 + \frac{M^2 R_m}{2 c_{\text{vm}}} \varepsilon_r \right)^{-1}. \quad (7)$$

Here, R_m is the gas constant for moist air and c_{vm} is the associated specific heat at constant volume. Again, there is an empirical adjustment for nonideal behavior. This is the nondimensional temperature recovery factor (ε_r), ideally unity. At the Sky Arrow's low Mach number ($M = 0.1$), the compression effects are small, which makes its calibration less significant, as will be noted later.

The ambient temperature determines the speed of sound, from which the true airspeed (U_a) can be calculated using M and R_m :

TABLE 1. Variables and constants to determine C_α , C_u , and α_0 according to our version of the upwash model of Crawford et al. (1996).

Symbol	Description	Dependency	Value	Unit
A	Aspect ratio wing	b^2/S	7.123	—
b	Wing span	Constant	9.680	m
B	Offset parameter upwash model	Fig. 2b	1.41*	°
c	Length wing's chord	Constant	1.400	m
C_L	Coefficient of lift	$L/(q S)$	Varies	—
C_u	Sensitivity coefficient lift-induced upwash	Eq. (13)	0.191*	—
C_{ux}	Forewash-accounting C_u	Eq. (14)	Varies	°
g_0	Standard gravity	Constant	9.806 65	m s ⁻²
L	Total lift	$m g_0$	6.30×10^3	N
m	Mass aircraft, including actual payload	Constant	642	kg
n	Nondimensional distance	x_{wp}/c	2.414	—
q	Dynamic pressure	Eq. (4)	Varies	hPa
S	Wing area	Constant	13.155	m ²
U_a	True airspeed	Eq. (8)	Varies	m s ⁻¹
w_p	Probe's vertical velocity (earth coordinates)	Observed	Varies	m s ⁻¹
x_{wp}	Distance wing to BAT probe (see text)	Constant	3.380	m
X	Estimator upwash model	Eq. (15)	Varies	°
$dC_L/d\theta_c$	Regression slope	Fig. 2a	$7.93 \times 10^{-2*}$	(°) ⁻¹
α_0	Attack angle at zero lift ($C_L = 0$)	Fig. 2a	-5.47*	°
α_f	Free-stream attack angle	θ_c	Varies	°
α_i	Upwash-affected attack angle	Eq. (2)	Varies	°
ε_θ	Pitch offset INS-GPS and probe	Eq. (15)	0.12*	°
θ_c	θ_i corrected	Eq. (12)	Varies	°
θ_i	Pitch angle by INS-GPS	Observed	Varies	°
χ	Vertical separation angle probe to wing (see text)	Constant	-17.1	°

* Calibration result.

$$U_a = M \sqrt{\gamma R_m T}. \quad (8)$$

A wind sensor must report undistorted flow, but an airplane must distort the flow to generate lift and thrust. These conflicting requirements are best treated by centering the probe on the fuselage as far forward of the wings and engine as possible. There, the distortion is well approximated as a vortex about the axis of the wings, the strength of which is found from the lift required to maintain flight (Crawford et al. 1996). Kalogiros and Wang (2002) found also a secondary contribution from the (tractor) engines of their Twin Otter. Assuming such influence to be negligible with a small pusher airplane like the Sky Arrow, the wing-bound vortex dominates. This vortex's tangential velocity experienced at the probe has speed w_u , called lift-induced "upwash," given by Crawford et al. (1996):

$$w_u = C_u U_a (\alpha - \alpha_0), \quad (9)$$

where C_u is the dimensionless aircraft-specific sensitivity coefficient, α_0 is the angle of attack at zero lift, and U_a is the true airspeed, the magnitude of $\hat{\mathbf{U}}_a$. The direction is perpendicular to the line of length x_{wp} connecting the wing's center of pressure (COP) to the probe's tip. This line makes an angle χ with the roll axis perpendicular to

the wingspan and passing just above the COP. For a wing of chord (i.e., width) c , the COP is distance $c/4$ aft of the leading edge. The C_u and α_0 need to be determined by a dedicated flight maneuver, but C_u can be estimated in advance by determining the wing's aspect ratio (A) and the nondimensional distance (n) from the COP of the wing to the tip of the probe (Crawford et al. 1996) as follows:

$$C_u \approx \frac{2A}{\pi n(2 + A)}. \quad (10)$$

The aspect ratio is determined from the airplane's geometry: $A = b^2/S$, where b is the wingspan and S is the wing area. The nondimensional distance n is x_{wp}/c . The parameters' values for the Sky Arrow are given in Table 1. Together with Eq. (10), this results in a theoretical value of 0.206 for C_u .

The wind components with respect to the aircraft ($\hat{u}_a, \hat{v}_a, \hat{w}_a$), in the coordinates of the airplane, are calculated as follows:

$$\begin{bmatrix} \hat{u}_a \\ \hat{v}_a \\ \hat{w}_a \end{bmatrix} = \frac{U_a}{D} \begin{bmatrix} -1 \\ \tan\beta \\ \tan\alpha_i \end{bmatrix} + w_u \begin{bmatrix} \sin\chi \\ 0 \\ -\cos\chi \end{bmatrix}, \quad (11)$$

$$D = (1 + \tan^2\alpha_i + \tan^2\beta)^{1/2}.$$

TABLE 2. Calibration results (calib set) with the initial set provided by theory except for α_0 (“init set”), including references to calibration maneuvers and equations in main text.

Parameter	Description	Equation	Maneuver (section 2b)	Init set	Calib set	Unit
C_u	Sensitivity coefficient upwash	(9)	1	0.206*	0.191	—
C_α	Sensitivity coefficient attack angle	(3)	1	0.250**	0.253	—
C_β	Sensitivity coefficient sideslip angle	(3)	2	0.250**	0.237	—
α_0	Attack angle at zero lift	(9)	1	N/A	−5.47	°
ε_q	Dynamic pressure correction	(4)	3	1	1.154	—
ε_r	Temp recovery factor	(7)	N/A	1	0.82	—
ε_θ	Pitch offset between INS-GPS and BAT	N/A	3	0	0.12	°
ε_ϕ	Roll offset between INS-GPS and BAT	N/A	N/A	0	0	°
ε_ψ	Heading offset between INS-GPS and BAT	N/A	3	0	−1.51	°

* Aircraft specific.

** Specific to sensor position and design.

For the Sky Arrow, χ is -17.1° , determined from its airframe dimensions (see before). Here, α is replaced by α_i , because the first right-hand-side term gives the upwash-contaminated relative velocity. The upwash is then removed by the second term. Note that, when $\chi \neq 0^\circ$, the upwash also includes a forewash component. Table 2 summarizes all calibration parameters presented here and their theoretical values.

Because the magnitude of each vector (i.e., $\hat{\mathbf{U}}_a$ and $\mathbf{U}_p$) is on the order of 10 times the magnitude of the turbulent wind, small errors in these vectors' magnitude or direction will corrupt the computed turbulent wind. At the Sky Arrow's nominal airspeed of 36 m s^{-1} , a 1° error in either vector's direction will result in an error in turbulent winds of $\frac{5}{8} \text{ m s}^{-1}$. These errors can be found and corrected, but the smaller the error, the cleaner the correction. Hence, precision and attention to detail are important in all aspects of the operation: instrument mounting, testing, calibration flight, and operational flight.

b. In-flight calibration

To determine the values of the calibration constants, four dedicated in-flight calibration maneuvers have been used here, which are mainly modified versions of the ones described by Bögel and Baumann (1991) and Leise and Masters (1993). A maneuver to determine the offset in the roll angle ε_ϕ , is not included here. Instead, ε_ϕ is set to 0° , similar to Hall et al. (2006). The following sections focus on the basic implementation of each maneuver and their relationship to the different calibration parameters. For a detailed description, see supplementary material Part A.

1) ACCELERATION–DECELERATION MANEUVER

The “acceleration–deceleration maneuver” determines the sensitivity coefficient (C_α) for the angle of attack, the angle of zero lift (α_0), and the sensitivity coefficient (C_u) of lift-induced upwash. During the maneuver, the aircraft is

kept straight and level at constant pressure altitude, while the airspeed is varied gradually down and then up to create a continuous change in the pitch angle.

This maneuver can also determine the temperature recovery factor ε_r (Lenschow 1986) as an alternative to the “racetrack maneuver” (Leise and Masters 1993). Here, as it is of minor importance at the typical low Mach numbers of the Sky Arrow, ε_r has simply been set to 0.82 according to a similar temperature probe (Gioli et al. 2006).

2) YAW MANEUVER

The “yaw maneuver” determines the sensitivity coefficient (C_β) for the angle of sideslip. It is executed by varying the yaw angle to the left and to the right while maintaining constant altitude and airspeed with zero roll. These uncoordinated turns generate sideslip in a snake-like movement of the aircraft's flight path. It is additionally used to determine the quality of the calibration with regard to the horizontal wind components (Lenschow 1986).

3) BOX MANEUVER

The “box maneuver” replaces the commonly used “reverse-heading maneuver” (Lenschow 1986) for calibrating the correction factor (ε_q) for dynamic pressure and the offset (ε_ψ) in the heading signal due to the misalignment between the BAT probe's relative wind sensors and the INS–GPS instrument. The flight path is a box of which the four straight legs are flown in cardinal headings for about 2 min at constant pressure altitude. It assumes a uniform horizontal wind field, which we considered justifiable for the relatively small region covered by a SERA flying at 36 m s^{-1} . The reverse-heading maneuver attempts to follow the same air on its return path in an effort to minimize the effect of heterogeneity in the wind field. It was not tested for this calibration set, but a trial is planned. The box maneuver is also used here

to give insight into the quality of the calibration (see later).

4) PITCH MANEUVER

The “pitch maneuver” is used to determine the quality of the calibration with regard to the vertical wind velocity according to Lenschow (1986). The aircraft is flown straight at constant mean pressure altitude, while the pitch angle is varied to create sinusoidal movements of the nose tip at different frequencies.

3. Implementation

The pusher propeller of the Sky Arrow allows the main scientific instruments to be mounted at the nose tip in front of the aircraft, that is, in minimally disturbed airflow. This makes the aircraft especially suited for flux measurements. Typical operating airspeed for surface flux measurements is 36 m s^{-1} at constant altitude mainly chosen between 20 and 80 m AGL (e.g., Kirby et al. 2008; Vellinga et al. 2010), but lower altitude has also been reported (6 m AGL, Zulueta et al. 2011). For vertical profiles, the aircraft has a nominal climb rate of 2.5 m s^{-1} and a ceiling of 4115 m.

a. Mobile flux platform

The MFP system in this study is one of the latest generations (section 1). The BAT probe is identical to Hacker and Crawford (1999), except the central pressure port has the same diameter as the rest of the holes. In previous versions, it was larger to host a microbead thermistor inside. Because of attenuation issues, the temperature sensor was moved outside the probe to its current position on top of the hemisphere (Fig. 1). The probe along with accelerometers at CG measures accelerations, wind angles via Honeywell absolute pressure (XCX15ANQ for $\bar{p}_r$) and differential pressure transducers (DCXL20DN for p_0 , and DCXL05DN for δp_y and δp_z), and temperature at high frequency ($>20 \text{ Hz}$). The accelerations and velocity data are used to upsample through frequency-blending GPS and attitude data, both obtained at 10 Hz by a GPS (NovAtel RT-20 ProPak) and a self-aligning INS-GPS instrument (Systron Donner C-Migits III; discontinued). For concentrations of CO_2 and H_2O , a LI-COR 7500 infrared gas analyzer (IRGA) is installed in the supporting part of the BAT probe.

In addition, height above ground level and earth surface temperature are measured by a Riegl LD90-3300 and a Heitronics KT15.85 IIP, respectively, from the bottom of the fuselage. Finally, net radiation (Kipp & Zonen NR Lite), and incoming and outgoing photosynthetically active radiation (LI-COR LI-190 quantum sensor) are installed at one end of the horizontal stabilizer. Other Sky

Arrow 650 ERA may be slightly different in one or more of these instruments.

The original temperature probe on our BAT probe was replaced by a slightly different one, which is smaller than the one mainly found on other Sky Arrow aircraft (see French et al. 2001; Garman et al. 2006). It basically consists of a thermocouple (Omega T-type COCO-002; $\varnothing$ 0.050 mm) with an instrumentation amplifier and thermocouple cold junction compensator on a monolithic chip (Analog Devices AD595). This allows a direct temperature readout without the need to convert output voltages to degrees Celsius during the postprocessing stage in contrast to the original sensor. Spectral analyses show that its frequency response follows the power law ($-5/3$ line) till the Nyquist frequency (25 Hz) is reached. We assume that due to its reduced dimensions, it affects the airflow less than the original.

b. Calibration flight

The calibration flight analyzed here took place on a cloudless day above land to the east of Teuge Airport ($52^\circ 14' \text{N}$, $06^\circ 03' \text{E}$, code EHTE) in the Netherlands on 21 April 2009. The area in the direct vicinity of the flight path has virtually no orography and is dominated mainly by agricultural land, grasslands, scattered forests and trees lines, and scattered houses to a few small cities.

The flight has been executed as outlined in supplementary material Part A at an average altitude of 1040 m ($\sigma = 7 \text{ m}$) above mean sea level (MSL). At the start of the flight at 0708 UTC (=LT - 2 h), a sharp temperature inversion was measured by the aircraft at about 160 m MSL, jumping from about 7.1° to about 10.5°C . A layer of temperature lapse was found between about 440 and 1005 m MSL, which is the residual boundary layer from the day before. Above 1005 m MSL, the temperature decreased again from about 11.2° to about 7.3°C . Hence, the calibration flight has been performed far above the local atmospheric boundary layer and just above the residual boundary layer. From the 0000 and 1200 UTC radiosoundings at the Royal Netherlands Meteorological Institute (KNMI, De Bilt, the Netherlands), which is about 70 km to the southwest, wind speeds of 4 m s^{-1} and wind direction of 72° at the average altitude and time of flight were determined by interpolation. Neither aircraft profile nor radiosoundings show signs of significant wind shear or wind veering at average flight altitude. So, we assume (and felt) no disturbance was present during the calibration flight.

c. Operational flight

To investigate the sensitivity of the fluxes to the uncertainty in calibration parameters (section 6), we used data from one of our weekly flights in the Netherlands in

2008 that were part of a major measurement campaign lasting one full seasonal cycle. All flights of that campaign have been performed in the same manner, following some basic rules with regard to aircraft handling, meaning that the bank angle did not exceed 20° , the turn rate did not exceed 3° s^{-1} , and flight movements were done as smoothly as possible.

The selected flight for the sensitivity tests took place on 25 July 2008, of which the part between 1130 and 1230 UTC has been used for analysis, because fluxes are the most stable around this part of the day. The flight track was located over the central part of the Netherlands (roughly between $52^\circ 17' \text{N}$, $05^\circ 19' \text{E}$ and $52^\circ 50' \text{N}$, $06^\circ 11' \text{E}$) with mainly agricultural areas and grasslands, and some farmhouses spread in between. The average wind speed and direction at 10 m AGL of three nearby stations were 5 m s^{-1} and 90° , respectively. Further weather conditions during flight were fairly good with increasing cloudiness (maximum 4/8 cloud cover), temperatures of about 25°C , and relative humidity of about 55%.

d. Processing software

The data collected with the onboard logging software (see Hall et al. 2006) is saved in Network Common Data Form (netCDF) format. We have developed the Wageningen University and Research Centre's mobile flux platform toolbox for MATLAB (WURMFP toolbox) to do all the essential postprocessing tasks, including analysis of calibration parameters, standard postprocessing of raw data into fluxes, and further spatial analysis in relation to any physiographical map [e.g., land use, satellite fraction of absorbed photosynthetically active radiation (fAPAR)].

The toolbox is an integrated set of tools partially based on already existing code, but with several extra features and enhancements. For retrieving the wind components from the pressure hemisphere, the toolbox contains a migrated code of the UNIX program MakePOD (Eckman et al. 1999) developed by ARL of NOAA. The code for calculating airborne fluxes in the toolbox has been based on code written by B. Gioli of Istituto di Biometeorologia del Consiglio Nazionale delle Ricerche (IBIMET-CNR, Firenze, Italy). Furthermore, the toolbox contains several additional subsets of tools, each focusing on a particular area of data analysis and processing, like flight path analysis, utility scripts, tools for plotting and exporting, and system checks. It has been built to work generically with most old and new MFP systems installed on current Sky Arrow aircraft that are operational today.

The toolbox has basic features to determine all calibration parameters. For more information, see supplementary material Part B.

4. Calibration

The goal here is to find a set of values for all calibration constants (section 2) that minimize the variance of wind speed (σ_U) and wind direction ($\sigma_{U_{\text{dir}}}$), and get the mean vertical wind component ($\bar{w}$) close to zero. In theory, if atmospheric conditions are ideal, then actual wind direction and speed are about constant and their variances are minimal, while mean vertical velocity is about zero ($\bar{w} \simeq 0 \text{ m s}^{-1}$). If the pilot handling is smooth and stable, and all calibration parameters are determined correctly, then computed winds would also match such characteristics. In practice, however, this is hard to achieve. As this is to fine-tune the parameters in a set of equations of our potential flow model, hereafter we call it "model calibration." With regard to the WURMFP toolbox, the script <mfppwind> is used (section 3d).

The basic principle is to go through the calibration constants in a particular order while varying one or more constant(s) at a time within a small range around the theoretical value. This is repeated in iterative sessions (in practice four to six cycles) until the values of all calibration constants are stationary. Ideally, the optimized constants will be close to their initial, theoretical values as given in Table 2. We have investigated a few different approaches to find out which one would be most efficient to reach an optimal set of parameters. Here, the resulting order from those tests is presented. In addition, we have tested to see whether different initial sets of calibration constants would result in different optimal sets to exclude the possibility that local minima are reached instead of a universal minimum. In all of those tests (not presented here) a universal minimum was reached. With regard to the toolbox, the optimal set of calibration parameters can be determined with script <mfppwind>, after having set model parameters with <setmfppwind> and <setmfppwind>.

First, the angle α_0 and the sensitivity coefficients C_u and C_w , all defined in section 2, have been determined simultaneously from the acceleration-deceleration maneuver following Crawford et al. (1996). For this, the coefficient of lift $C_L = L/(q S)$ is calculated, assuming that the wing's total lift (L) equals the aircraft's weight, including actual payload (i.e., $L = m g_0$). This is plotted on θ_c , which is the indicated pitch angle by the INS-GPS instrument (θ_i) corrected by the vertical motion of the probe in earth coordinates (w_p) and the offset in the pitch angle (ε_θ), as represented by the formula

$$\alpha_f \equiv \theta_c = \theta_i - \varepsilon_\theta - \left(\frac{w_p}{U_a} \frac{180}{\pi} \right). \quad (12)$$

During the initial iteration step, ε_θ is unknown and is, therefore, set to its theoretical value of 0° .

FIG. 2. Result of determining (a) C_u and α_0 by setting C_α (here: 0.253) to fit the regression line (b) to be of unit slope, according to our version of the upwash model of Crawford et al. (1996) with data from the acceleration–deceleration maneuver.

Using θ_c as a proxy for the free-stream attack angle (α_f) instead of θ_i results in a better linear relationship with C_L and, hence, better calibration method. This is a fundamental difference compared to the model of Crawford et al. (1996), who have used θ_i throughout the model. Crawford et al. (1996) found the relationship between C_L and θ_i to be linear up to $\theta_c = 5^\circ$. We assume the same for θ_c here (Fig. 2a), where its slope, determined by simple linear regression, has the empirical relationship

$$C_u \equiv \frac{dC_L}{d\theta_c} (n\pi^2)^{-1}. \quad (13)$$

As the BAT probe is (only) vertically displaced with regard to the roll axis of the Sky Arrow, a more generic form of C_u is used here to account for the resulting forewash from this displacement as follows:

$$C_{u\chi} = C_u [\cos\chi + (\theta_i - \varepsilon_\theta) \sin\chi], \quad (14)$$

where χ is the angle of vertical displacement of -17.1° (section 2a).

Knowing $C_{u\chi}$ allows for computation from known quantities of an estimator X for the upwash-contaminated α_i of Eq. (2). Using the vertical component of Eq. (1) with

$\hat{r} = 0$, and incorporating Eqs. (2)–(11), the following relation can be derived, similar to Crawford et al. (1996):

$$\alpha_i = X + B, \quad X = \frac{\theta_c}{1 - C_{u\chi}}, \quad B = \frac{-C_{u\chi}\alpha_0}{1 - C_{u\chi}}. \quad (15)$$

Coefficient C_α in Eq. (2) is adjusted until the relation between α_0 and X yields unit slope (Fig. 2b), which has been determined to be very close to its theoretical value of 0.250, namely, 0.253.

After this, final values for α_0 and C_u can be obtained. By extrapolation of the regression line in Fig. 2a, the pitch angle (θ_c) at zero lift ($C_L = 0$) is found to be -5.47° , which is equal to the attack angle at zero lift, α_0 . The resulting C_u of 0.191 from Eq. (13) during the final iterative step quite closely matches the theoretical value (0.206) given by the relatively simple upwash model [Eq. (10)], reflecting the simplicity of the flow distortion at the location of the probe on this airplane. Table 1 summarizes all the direct variables and constants necessary for use of the model with their dependencies to other variables, equations, and figures.

Next, the offset ε_θ has been determined with data of one or more straight sections of the calibration flight, for which the airplane was flying level at constant pressure altitude. The purpose here is to find the value of ε_θ , for

FIG. 3. Result of determining the sensitivity coefficient for sideslip (C_β) by finding the minimum of the sum of variances of the horizontal wind components ($\sigma_u + \sigma_v$) with data from the yaw maneuver, showing (top) the time series of wind speed (U), (middle) the time series of wind direction (U_{dir}), and (bottom) the sum of variances. Black lines in top two graphs show U and U_{dir} at optimal value $C_\beta = 0.237$, dark gray lines are below optimum, and light gray lines above optimum.

which the mean vertical wind component ($\bar{w}$) is zero by varying the offset within the typical range of $\pm 3^\circ$. The individual straight sections of the box maneuver are ideal for this. From the results (not shown), the average offset is calculated, which serves as input in the upwash model during the next iterative step. The final iterative step resulted in an offset of 0.12° .

The sensitivity coefficient C_β has been determined from data of the yaw maneuver by finding the minimum of the sum of the variances of the two wind components ($\sigma_u + \sigma_v$). Figure 3 shows the result of the final iterative session. The top two graphs that depict the wind speed (U) and the wind direction (U_{dir}) show the influence of this sideslip parameter very clearly by the dark gray lines (below optimal) and light gray lines (above optimum). The four different stages of sideslip are clearly shown for the yaw turn to the left (sample range 1–5000) by the staircase pattern, while during the yaw turn to the right, these stages are less distinct and show more disturbance in the signal. The black lines in both graphs correspond with the optimal value of C_β found in the bottom graph at 0.237. This value is acceptable, but the possible change of wind direction and/or speed during the implementation of the maneuver may have changed it from the theoretical value of 0.250. In addition, the pilot

handling may further have degraded the result as the heading signal shows (not presented here) a clear turn from north-northeast to west-northwest during the left yaw turn and a weak turn from west-northwest to north-northwest during the right yaw turn. It should be noted that a right yaw maneuver is more difficult to perform with the Sky Arrow because of the interaction of the propeller-generated vortex with the tail of the aircraft.

The correction factor ε_q and the offset ε_ψ have been both separately determined from the straight sections of the box maneuver by finding the minimum variances for horizontal wind direction ($\sigma_{U_{\text{dir}}}$) and wind speed (σ_U). For this, ε_q was varied around its theoretical value of 1 within a range of ± 0.2 , and ε_ψ was varied around its theoretical value of 0° with a margin of $\pm 3^\circ$. As example, Fig. 4 shows results for ε_ψ from the final iterative session of the calibration. Figure 4a clearly shows the influence of ε_ψ on both the magnitude of the wind direction and the magnitude of the wind speed. Dark gray lines correspond with values of ε_ψ smaller than its optimal value, and light gray lines above the optimum. The straight legs of the box are clearly visible by the four discrete steps in the data, which virtually disappear as ε_ψ reaches the minimum values of $\sigma_{U_{\text{dir}}}$ and σ_U (black lines). Note that U_{dir} seemed to have changed slightly with a few degrees

FIG. 4. As in Fig. 3, but (b) for determining the offset in heading (ϵ_ψ) at (top) minimum σ_U and (bottom) minimum $\sigma_{U_{dir}}$, using data from box maneuver. (a) Depiction of its effect on (top) wind speed and (bottom) direction.

during the implementation of the maneuver. This suggests nonhomogeneous (or nonsteady) conditions within the scale of the box track. Figure 4b shows $\sigma_{U_{dir}}$ and σ_U with minima of ϵ_ψ both at -1.51° . The offset may be related to a misalignment between the airplane's longitudinal axis and the INS–GPS unit, since the latter is automatically aligned on the ground at standstill before every flight by pointing the airplane's nose in the cardinal northern direction. The same method is used for determining ϵ_q , and ϵ_q shows a similar behavior (not presented here). Here, the minima did not coincide, and the average (i.e., $\bar{\epsilon}_q = 1.154$) of each individual corresponding ϵ_q has been selected.

Finally, the model calibration has been verified here in two ways. First, by analysis of the pitch maneuver and the yaw maneuver according to Lenschow (1986). Figure 5a shows results with time series of the probe's vertical velocity (w_p) and the vertical wind component (w) for the pitch maneuver. According to Lenschow (1986), less than 10% of w_p is allowed to be found in w . The graph clearly shows four different stages of pitching frequencies during the maneuver. For the two lower frequencies, our model performs very well. There is virtually no signature in w of the aircraft's pitch motion. The two higher frequencies, however, clearly violate

Lenschow's criterion, probably because the current upwash model gives no account to vertical acceleration. Treatments of vertical acceleration have been proposed by Kalogiros and Wang (2002) and Garman et al. (2008). We intend to incorporate these into our future model (and toolbox).

Figure 5b shows time series of the horizontal wind components (u , v) and the lateral wind component in aircraft coordinates ($\hat{v}_a$) for the yaw maneuver. The stepwise changes in sideslip, clearly visible in $\hat{v}_a$, should be absent from the horizontal wind components. From $\hat{v}_a$, the aircraft can be seen to have made a yaw turn to the left (without coordinated roll) in four steps, followed by a similar turn to the right. After about sample 5000, the aircraft encountered some disturbance. Though the disturbance is clearly visible in all traces, the pilot's intentional shifts in sideslip are not reflected in the horizontal wind. Note also the lack of ripple in u and v as $\hat{v}_a$ goes through some oscillations in lining up on the chosen sideslip angle during the left yaw turn. Except for the section with the strongest sideslip and through the recovery back to coordinated flight (samples 1600–2300), u and v are not affected to such an extent that it violates the 10% criterion.

Another way that has been used here to verify the quality of the calibration is by plotting wind vectors and

FIG. 5. Quality check of calibration according to Lenschow (1986), for which vertical wind component (w) should be $<10\%$ of vertical aircraft velocity (w_p) (a) for pitch maneuver and (b) for yaw maneuver for which horizontal wind components (u , v) should each be $<10\%$ of lateral wind component (v_a). High-frequency pitching violates this criterion because of shortcomings in the current upwash model.

calculating basic statistics for the whole box maneuver, including turns. Figures 6a and 6b show such a plot, when using the initial and calibrated set of parameters (Table 2), respectively. As there is no theoretical value available for α_0 , the value of -3.5° from the Sky Arrow of Hall et al. (2006) was taken for the initial set. The standard deviation for wind direction, $\sigma_{U_{\text{dir}}}$, is 1.8° for the calibrated set compared to 24° for the initial set, and the standard deviation of wind speed, σ_U , is 0.19 m s^{-1} for the calibrated set compared to 1.5 m s^{-1} for the initial set. The average vertical wind speed is much closer to zero ($\bar{w} = -0.01 \text{ m s}^{-1}$) for the calibrated set than for the default set ($\bar{w} = 0.37 \text{ m s}^{-1}$). Note that the sharp turns virtually do not appear in the calibrated wind box, while they are more visible in the noncalibrated one. The calibrated wind box shows a large ripple in the wind direction at one end of the east leg. This short change of direction can also be found in Fig. 4. Basic trend analysis by linear regression shows that during the course of the maneuver, the overall wind direction seemed to have changed from 55° at the beginning to 57° at the end of the maneuver with a slight increase in wind speed ($\Delta U = 0.15 \text{ m s}^{-1}$). These nonstationary conditions can affect the quality of the calibration and may have occurred during the whole calibration flight.

The tests depicted in Figs. 5 and 6 demonstrate a successful calibration, except for showing some limitations

in the current upwash model as noted earlier. Notably from Fig. 6, the horizontal wind is little affected by the sharp turns at the corners of the box, though there seems to be signs of nonstationary conditions during the calibration flight.

5. Data processing and quality control

a. Data processing

Using calibrated parameters, our standard processing includes three basic processing stages in order to calculate flux data from raw data.

In the first stage, the raw data need to be checked for outliers (spikes). We use two filters to detect outliers in each signal: a physical range filter and a margin around a zero-phase moving average. The margin is defined based on a percentage of the physical range. The physical range is usually set for all signals to the widest range to cover most atmospheric conditions. If a spike is found, then it is manually checked for physical plausibility. Spikes found implausible are removed and their gaps filled simply by linear interpolation. Gap filling is necessary because during the next processing stage, some signals are blended in frequency domain, which does not allow data gaps. In practice, spikes are rare and generally consist of a single sample. Should a continuous block of two or more outliers be found, the block is flagged

FIG. 6. Wind vectors and some basic statistics for box maneuver of calibration flight 21 Apr 2009 with (top) initial set and (bottom) calibrated set of parameters (Table 2).

as a bad data block (section 5b). In WURMFP toolbox, script `<mfpraw>` facilitates this processing step (section 3d).

In the second stage, wind components, true airspeed, and (dry) air density are calculated. First, the residual offset errors between the BAT probe accelerometer data and the CG accelerometer data are removed. Depending on the temperature sensor used (section 3a), offset errors between the slow and fast temperature sensor need to be corrected. Then the attitude, velocity, and position signals are upsampled from 10 to 50 Hz by frequency blending each of them with the 50-Hz accelerometer signals separately. We do not correct for time lag due to sensor separation between pressure hemisphere, temperature probe, and open-path gas analyzer because it cannot be resolved: at 50 Hz and nominal airspeeds of 36 m s^{-1} , sensor separation less than 72 cm cannot be accounted for. However, the LI-7500 gas analyzer has a relatively long internal electronic time lag that needs correction to synchronize the CO_2 and H_2O signal with the other signals. Then the wind component, true airspeed, and air density are calculated using the previously determined parameters. Finally, the same filters from the previous stage are used here. They act as an alert system for spikes slipped through earlier filtering due to human

error (i.e., misconfiguration of filters or misinterpretation of outliers). For instance, outliers in attitude data can have an adverse effect on the calculated wind components. If outliers are found, then outliers are flagged. All can be realized with script `<mfwind>` of the toolbox.

In the final stage, spatially averaged fluxes of momentum, energy, and carbon dioxide are calculated for transect flights. For profile flights, temperature, humidity, wind vectors, and carbon dioxide concentrations are calculated as time averages nominally over 1 s. First, flights are analyzed to filter out events that may adversely affect the results, such as sharp turns, turns with large banks angles (greater than 20°), passage of smoke plumes, rain, etc. Next, for fluxes, covariances are computed on data blocks (windows) representing flight sections of typically 2–8 km, and several standard corrections are applied. The wind components (u , v , w) are rotated as is common practice at flux towers (Aubinet et al. 2000), so that the average value of the lateral and the vertical wind components equals zero ($\bar{v} = 0$, $\bar{w} = 0$). Here, vertical rotation is minimal since after model calibration, data practically exhibit zero vertical wind. Then, wind components are corrected based on ground speed variations as reported by Crawford et al. (1993) to compensate for undersampling during updrafts (=lower airspeeds) and oversampling during downdrafts (=higher airspeeds). Next, fluxes are calculated and corrected for density effects due to heat and water vapor transfer [Webb–Pearman–Leuning (WPL) correction; Webb et al. 1980]. A frequency response correction, generally used to correct fluxes obtained by tower-based eddy covariance systems, is not applied (see later). The net radiation and the incoming and outgoing photosynthetically active radiation are corrected for systematic misalignment of sensor to nominal pitch angle during flight. At this stage, several data quality filters have additionally been implemented, the use of which will be presented in the next section. All computational steps here, including filters and corrections, are incorporated into toolbox script `<mfflux>`.

b. Quality control

During all three processing stages, several quality checks and standard data corrections are used. During those stages, each 50-Hz sample is marked with a quality class (0 = unchecked, 1 = highest quality to 8 = lowest quality, and 9 = rejected). By default, they are marked with quality class 0. Data that pass all quality filters are marked with quality class 1. If, however, two or more successive samples are replaced by interpolated values, then quality class 8 is assigned to these samples. A single interpolated sample is regarded as a good approximation of the true value and is marked with quality class 2.

Some instruments (e.g., INS–GPS) provide additional information about signal strength or reception and standard deviations about the data, which are translated to the quality flagging system (0–9) used here.

Data with quality class 8 or 9 are discarded before the computation of covariances and other statistics in the final postprocessing stage. If more than 25% of the flagged data (with class 8 or 9) within an averaging window is filtered out, then no fluxes are calculated for that specific window. No advanced algorithm has been implemented to analyze the pattern of removed data (i.e., large sample blocks vs scattered single samples). The stationarity test (see next) acts here as a basic alternative but only for averaging windows that will be calculated. In practice, this threshold of 25% is rarely reached, as typically less than 1% of a window is flagged with class 8 or 9.

Three filters are used to classify the retrieved flux data for each window. First, a stationarity test is performed on uncorrected fluxes (no WPL correction), where the average flux of a window is compared to the average flux of the same window but subdivided into a number of segments. The window is divided into three segments, instead of the usual 5 or 6 segments for 30-min windows at flux tower sites (Aubinet et al. 2000; Foken et al. 2004). Tests (not presented) showed that more than three segments would often filter out >50% of the flux data because of the short window length (4000 m ~114 s) we used. [Column A in Table 4 shows the classification of data quality for the stationarity test according to Foken et al. (2004).] For instance, if the average flux of all three segments deviates 22% from the flux of the whole window, then the flux of the whole window falls in quality class 2.

Second, an integral turbulence characteristics (ITC or I) test, coming from the flux tower community (Thomas and Foken 2002; Foken et al. 2004; Mauder and Foken 2004), is performed for each averaging window. The test follows the idea that the ratio of the standard deviation of a turbulent parameter to its turbulent flux is nearly constant or a function of atmospheric stability. The observed integral turbulence characteristics for any wind component or scalar (e.g., temperature) can be determined by the ratio between the standard deviation of a wind component or scalar and the normalizing factor derived from its turbulent flux. In the case of wind component, w , the equation will be

$$I_{\text{obs},w} = \frac{\sigma_w}{u_*}, \quad (16)$$

where u_* is the friction velocity. For the sensible heat flux in the surface layer, the equation is

TABLE 3. Coefficients for modeling integral turbulence characteristics with Eq. (18), similar to Foken et al. (2004) and Mauder and Foken (2004), where $\zeta \equiv (z - d)/L$.

Parameter	Stability range	c_1	c_2
σ_w/u_*	$0.4 < \zeta$	2.0	1/8
	$-0.2 < \zeta < 0.4$	see Eq. (19)	
	$\zeta < -0.2$	2.0	1/8
σ_u/u_*	$0.4 < \zeta$	4.15	1/8
	$-0.2 < \zeta < 0.4$	see Eq. (20)	
	$\zeta < -0.2$	4.15	1/8
σ_T/T_*	$1 < \zeta$	1.0	-1/3
	$0.02 < \zeta < 1$	1.4	-1/4
	$-0.062 < \zeta < 0.02$	0.5	-1/2
	$-1 < \zeta < -0.062$	1.0	-1/4
	$\zeta < -1$	1.0	-1/3

$$I_{\text{obs},T} = \frac{\sigma_T}{T_*}, \quad (17)$$

where $T_* = -(\overline{w'T'_{\text{pot},v}})_s/u_*$ (Stull 1988), for which we use the airborne potential temperature T_{pot} instead of the virtual potential temperature $T_{\text{pot},v}$. The ITC can be modeled according to

$$I_{\text{model},x} = c_1 \left(\left| \frac{z-d}{L} \right| \right)^{c_2}, \quad (18)$$

where L is now the Obukhov length and $(z - d)$ is the aerodynamic height, for which we choose the measurement height above ground level. Constants c_1 and c_2 are listed in Table 3 and depend on the atmospheric stability, $\zeta = (z - d)/L$. For neutral conditions, the following two equations with regard to wind components w and u are used instead (Johansson et al. 2001; Thomas and Foken 2002):

$$I_{\text{model},w} = 0.21 \times \ln \left(\frac{f z_+}{u_*} \right) + 3.1, \quad (19)$$

$$I_{\text{model},u} = 0.44 \times \ln \left(\frac{f z_+}{u_*} \right) + 6.3, \quad (20)$$

where $z_+ = 1$ m, which is introduced to get the resulting expression dimensionless; f is the Coriolis parameter; and u_* is the friction velocity.

The ITC test parameter, defined as $I = |(I_{\text{model},x} - I_{\text{obs},x})/I_{\text{model},x}|$, determines the second quality flag assigned to the flux of particular window. For u_* , I for vertical wind (I_w) and longitudinal wind (I_u) are both separately calculated, of which the larger determines the quality flag. The same is done for H but now I_w and I_T are used. For $f\text{CO}_2$ and λE , only I_w is calculated. As stated by Foken et al. (2004) and Mauder and Foken

TABLE 4. Classification of data quality for airborne flux data by stationarity test (A) and ITC test (B) and by physical range test (C) similar to Foken et al. (2004).

Class	A	B	C
0	N/A	Test not applicable*	N/A
1	0.00–0.15	0.00–0.15	Within range*
2	0.16–0.30	0.16–0.30	N/A
3	0.31–0.50	0.31–0.50	N/A
4	0.51–0.75	0.51–0.75	N/A
5	0.76–1.00	0.76–1.00	N/A
6	1.01–2.50	1.01–2.50	N/A
7	2.51–5.00	2.51–5.00	N/A
8	5.01–10.0	5.01–10.0	N/A
9	>10.0	>10.0	Out of range*

* Refers to the text.

(2004), the ITC test cannot be resolved for $H < 10 \text{ W m}^{-2}$ at neutral conditions (here: $-0.062 < z/L < 0.02$) because the model is not well defined for this range of H . In that case, the window is flagged with class 0 (=unchecked) and the test is ignored when combining the results into an overall quality flag (see later). Quality classification following the ITC test can be found in column B of Table 4.

Finally, a physical range filter is used to assign a third quality flag to the flux data. If a flux is not within a realistic physical range, then it is flagged with quality class 9. Otherwise, it is put in quality class 1. For $f\text{CO}_2$, the selected range is between -50 and $50 \mu\text{mol m}^{-2} \text{ s}^{-1}$. By default, for H and λE , the accepted range is 0 – 800 W m^{-2} because the PH-WUR is normally flown during daytime at nominal flight levels of 80 m , when only positive energy fluxes are assumed to occur. Depending on one's (scientific) interest and the type of experiment (e.g., flying stacked transects to study flux divergence), these ranges may need to be adjusted.

The results from the three test filters are combined into an overall quality flag for airborne flux data according to Table 5. This follows the approach of Foken et al. (2004), except that the third test about the horizontal orientation of the sonic anemometer is here replaced by the physical range test (column C in Tables 4 and 5) and that we deem quality classes 1–6 to be acceptable for scientific use, addressing surface flux quantification (see also section 7).

All these quality checks are included in the three main scripts of the WURMFP toolbox: <mfpraw>, <mfprwind> and <mfprflux> (section 3d).

6. Sensitivity analysis

Here, we analyze the relevance of the model calibration for the derived fluxes. The first sensitivity test consisted of calculating contiguous 4-km-averaged

TABLE 5. Overall classification of data quality for airborne flux data combining individual tests from Table 4; similar to Foken et al. (2004).

Class	Label	A	B*	C
1	High (“scientific use”)	1	1–2	1
2		2	1–2	1
3		1–2	3–4	1
4	Medium (“general use”)	3–4	1–2	1
5		1–4	3–5	1
6		5	≤ 5	1
7	Low (“indicative”)	≤ 6	≤ 6	1
8		≤ 8	≤ 8	1
9	Rejected	**	**	**

* Test B is ignored, if data value is flagged 0.

** If a test flags a data value with class 9, then it is rejected.

windows of fluxes (with quality class ≤ 6) for a part of a normal flight on 25 July 2008 (section 3) by adding an error of $\pm 10\%$ to the calibrated value of each calibration parameter alternately (“calib set” in Table 2); except for ε_ϕ , for which 10% of its typical range ($\pm 0.3^\circ$) was taken as its lower and upper bounds. From these, the root-mean-square error (RMSE) has been determined. Table 6 shows that RMSEs are, overall, between 0.0% and 8.5% . The uncertainty of the temperature recovery factor (ε_r) does not significantly contribute to an error in any of the fluxes with a maximum of 0.8% for H . It affects the least λE and u_* , while the RMSE for $f\text{CO}_2$ and H is still below 1% . With regard to the latter two fluxes, ε_ψ has minimal effect, and it affects λE somewhat in the order of ε_r . It supports our decision not to calibrate our temperature probe (section 7). Note that u_* is mainly affected by the sign of the added error. Its RMSE decreases slightly with a positive added error, except for C_u and ε_q . Most other fluxes are only affected when errors are added to C_α , ε_q , and ε_r , which decrease with a positive added error. The sign has no significant effect on the three offset parameters (ε_ϕ , ε_θ , ε_ψ), whose influence stay well below 1% . However, results for ε_ϕ may be troubled by the fact that our current model does not account for vertical sensor separation (section 7). Calibration parameters that have the largest effect on the airborne fluxes are the ones directly related to correcting horizontal wind speed (ε_q) and the vertical wind component (C_u and C_α). The sensitivity coefficient (C_α) for the angle of attack has the largest effect on $f\text{CO}_2$ ($\sim 7.3\%$), H ($\sim 8.2\%$), and λE ($\sim 8.2\%$), whereas ε_q has the largest effect on u_* ($\sim 8.0\%$).

During the test, we noticed that fluxes for some of the 4-km windows are less responsive to variations in any of

TABLE 6. RMSE for fluxes of contiguous 4-km-averaged data windows of flight 25 Jul 2008 (with quality class ≤ 6) by adding an error of $\pm 10\%$ to each calibrated model parameter separately (Table 2).

Parameter	Error (%)	RMSE (%)			
		$f\text{CO}_2$	H	λE	u_*
C_u	-10	3.2	3.2	3.5	2.8
	10	3.2	3.2	3.5	3.0
C_α	-10	7.3	8.3	8.3	6.1
	10	7.2	8.0	8.1	5.6
C_β	-10	0.6	0.7	0.5	1.8
	10	0.6	0.7	0.5	1.7
α_0	-10	1.7	1.3	1.7	1.6
	10	1.7	1.3	1.7	1.4
ε_q	-10	5.8	6.9	6.6	7.4
	10	4.3	7.0	5.2	8.5
ε_r	-10	0.1	0.8	0.01	0.3
	10	0.2	0.8	0.02	0.0
ε_ϕ^*	± 10	0.3	0.3	0.2	0.5
ε_θ	± 10	0.2	0.1	0.2	0.5
ε_ψ	± 10	0.04	0.03	0.03	0.5

* Since its optimum is 0, ε_ϕ was varied over $\pm 0.3^\circ$, which is 10% of its typical range.

the calibration parameters. Figure 7a shows an example for $f\text{CO}_2$ and its corresponding absolute errors in Fig. 7b along the flight track, where C_α is varied $\pm 10\%$ about its calibrated value (0.2535). The windows (1, 8, 16, 19, 28, and 29) that show small errors are less sensitive to changes

in C_α and, consequently, the calculated wind field. Similar behavior was found for other model parameters.

To test the overall interactions between the parameters, a second sensitivity test compares fluxes from the same dataset as in the previous test that are calculated with the calib set (Table 2) against those calculated with the same set but with an added error of $\pm 10\%$ to all members of the set. Table 7 sums up the results of this second test. Note that also here the sign of the error determines the extent of change in the fluxes, where an error of -10% results in larger fluxes, except for u_* . However, an error of $+10\%$ shows somewhat smaller variances, except for H . When comparing the RMSEs against Table 6, the overall interaction between the parameters becomes clear. The previous analysis showed that C_u , C_α and ε_q have the largest effect on the fluxes. Note, for instance, that in the first test, when only an error of $\pm 10\%$ was added to C_α , this resulted in an RMSE for H of about 8.2%. Here, H got a significant smaller RMSE of 3.3% or 4.9%, depending on the sign of the error, although C_α remains an error of $\pm 10\%$. This shows that, next to C_α , other parameters additionally affect the outcome. In case of ε_q affecting u_* , the interaction with other calibration parameters seems less evident, because the RMSE of both tests differs less.

Although these sensitivity tests are based only on one case (i.e., one airborne dataset), they show indications that an overall error of $\pm 10\%$ in all members of a calibrated

FIG. 7. Effect of the sensitivity coefficient for angle of attack (C_α) on contiguous 4-km-averaged windows of carbon dioxide fluxes (a) with $\pm 10\%$ error as in Table 6 for flight 25 Jul 2008 with (b) relative error for the same windows. Negative flux in (a) is carbon uptake at surface, and dashed line shows results using a calibrated C_α .

TABLE 7. Basic statistics for fluxes of the same dataset as Table 6, but now calculated with the calib set (Table 2) and with two contaminated sets of the calib set that contain an added error ($\pm 10\%$) to all members simultaneously. Negative carbon dioxide flux means carbon uptake at surface.

Parameter set		$f\text{CO}_2$ ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	H (W m^{-2})	λE (W m^{-2})	u_* (m s^{-1})
Calib set	mean	−14.3	61	282	0.54
	σ	5.9	20	108	0.13
Calib set with error of −10%	mean	−14.6	62	288	0.54
	σ	5.9	20	109	0.14
	rmse	0.5	2	10	0.03
	RMSE %*	3.2	3.3	3.5	6.1
Calib set with error of +10%	mean	−14.1	59	277	0.53
	σ	5.9	22	107	0.13
	rmse	0.4	3	9	0.04
	RMSE %*	3.0	4.9	3.2	7.4

* Relative to the calib set.

set of model parameters may result in an overall RMSE of about 4% in the flux magnitudes, but it may be larger (up to about 8%) when only C_α or ε_q has an error to the same degree. A negative error to the whole parameter set results in larger fluxes.

7. Discussion

Our results show a successful model calibration of an airborne MFP system, as it improves the observed wind field and slightly reduces flux magnitudes. Systematic quality control may further help to select flux data for subsequent analysis of land–atmosphere exchange processes. However, a few aspects need further discussion.

Flow effects due to thrust of the Sky Arrow's pusher propeller has not been investigated here. A propeller creates a region of low pressure ahead of it, which extends forward for some distance. Its influence is greatest along the thrust line, which for the Sky Arrow goes directly above the probe. We expect the effect to be minimal because of the reduced power of a small airplane and because the pusher configuration moves the propeller farther from the probe than a tractor configuration would. Moreover, the power setting is kept as nearly constant as possible during standard operational flights. To support our view, it would be worthwhile to investigate it in a similar fashion as Kalogiros and Wang (2002).

The flow model used here does not account for the presence of a temperature sensor mounted on the pressure hemisphere. From wind tunnel tests, Garman et al. (2006) found that their temperature sensor distorted the flow around the hemisphere, giving the probe an asymmetrical response to changes in the angle of attack. We assume that our temperature sensor has no significant effect on the flow distortion because of its smaller size.

Moreover, our sensor is located on top of and behind the hemisphere, and our aircraft has a positive angle of attack in straight-and-level flight at nominal fuel load. Thus, the sensor interferes with downdrafts rather than with updrafts. The latter are generally stronger and account for a larger portion of the flux. More detailed characterization of the BAT probe with its temperature sensor would require tests in a wind tunnel following Garman et al. (2006).

The calibration results show that our model is capable of removing virtually all aircraft motion from the data, especially for our standard operational flights, which are flown as level as possible at constant airspeed and avoid sharp turns. Only in extreme cases of sideslip and/or angle of attack are its limitations clearly visible. With regard to w , treatments of vertical accelerations need to be incorporated into the model following proposed methods by Kalogiros and Wang (2002) and Garman et al. (2008). In addition, extreme cases of sideslip cannot be filtered out at this stage. The method used here for determining C_β might be the reason for that, instead of the model itself. We are currently developing a more robust method similar to Crawford et al. (1996) that, we hope, determines C_β more accurately.

Our calibration does not yet include a method to determine the offset ε_ϕ in the roll angle because our flow model does not currently account for the approximately 1-m vertical separation between the BAT probe on the nose and the roll axis centered between the wings (Fig. 1). During the calibration flight, circles at the standard turn rate (120 s) were flown clockwise and counterclockwise, maintaining about 15° of roll in each. The circles were flown with sufficient skill that the relative velocity ($\hat{u}_a$, $\hat{v}_a$, $\hat{w}_a$) was constant around each. The circle flown clockwise was well coordinated (nearly constant $\hat{v}_a \simeq 0 \text{ m s}^{-1}$), while the circle flown counterclockwise uncoordinated

though steady (nearly constant $\hat{v}_a \simeq -2 \text{ m s}^{-1}$). The clockwise (coordinated) circle fits Lenschow's 10% criteria, while the other produces wind fluctuations (u, v) that are fully 35% of the associated variations in the probe's velocity ($\hat{u}_p, \hat{v}_p$). Apparently, our model can handle large roll angles in coordinated flight but not yet in uncoordinated flight. The mean roll angle therefore is as close as possible to zero during all operational flights. Data taken when the roll is larger than 20° are filtered out from the flux calculations. Our sensitivity tests showed no large effect of ε_ϕ on the airborne fluxes under this protocol.

The temperature recovery factor (ε_r) can be obtained by the racetrack maneuver (Leise and Masters 1993) or, alternatively, by the acceleration–deceleration maneuver (Lenschow 1986). Our operational flights, however, maintain a Mach number of about 0.1, at which the effect of compression on the temperature is minimal. Hence, ε_r has a marginal role [Eq. (7)], as our sensitivity analysis confirms (section 6), and we simply set a value of 0.82 (B. Gioli 2006, personal communication) from another Sky Arrow (Gioli et al. 2006) with a temperature probe design similar to the one studied by Garman et al. (2006).

Although atmospheric conditions seemed ideal during the calibration flight (section 3b), results from the box maneuver suggest a degree of nonstationarity or heterogeneity in the wind field, which may have had a slight effect on the determination of ε_ψ and ε_q . In a future calibration, we plan to test the reverse-heading maneuver as described by Lenschow (1986) to see if it better handles wind field heterogeneity.

The model calibration has been done manually, optimizing parameters in a particular order one parameter at a time and repeating this process until the whole set converged on a final result. Most calibration parameters (except α_0) have a known theoretical value from which to start, thus speeding convergence. In principle, the calibration procedure could be automated and (arguably) made more objective through a more sophisticated algorithm, such as a downhill simplex routine. However, implementing this in a robust way may be complicated, so we have avoided it for now. It may find the optimum for a whole set more rapidly but hardly more accurately.

Corrections for sensor separation and for flux loss at high frequencies have not been implemented in our postprocessing procedure. Our temperature sensor's power spectrum shows the theoretical $-5/3$ slope up to the Nyquist frequency (25 Hz, 1.5 m). Moreover, our operational flight altitudes average about 80 m AGL, considerably above the usual flux towers. Cospectra show that eddies of at least 4-m scale (10 Hz) contribute most to the flux at these altitudes, whereas the contribution of small eddies is negligible. Corrections for sensor separation are negligible because all sensors—wind, temperature,

moisture, and CO_2 —are within 0.3 m, about 10% of a 10-Hz eddy's scale.

Corrections could also be applied to account for low-frequency flux loss, but it would need to be a function of the window size used in the flux computations. Earlier tests with the Sky Arrow by IBIMET-CNR following the approach of Vickers and Mahrt (2006) on multi-resolution decomposition based on wavelet theory did not give any expected results (not presented here). Such a correction was not included in our postprocessing procedure, and future work will determine the averaging length by ogive analysis.

The stationarity test and the ITC test used in this study are usually part of a QC–QA strategy for tower-based flux data (Aubinet et al. 2000) but not for airborne MFP systems. Classification of these tests are based on knowledge gained on filtering time blocks of 30 min of (mainly) 10-Hz ground flux data, whereas our contiguous 4-km windows contain on average only about 114 s of 50-Hz data. Selecting an appropriate window length is a compromise between capturing the largest relevant eddies requiring a large window and resolving the scale of the heterogeneous surface by using a window as short as possible. This could be overcome by using new approaches recently published by Kirby et al. (2008) and Hutjes et al. (2010), which will be subject for study in future work. In the stationarity test, we have reduced the number of subwindow segments from the usual five or six segments to only three segments in order to avoid too short segments that may be subject to variations associated with the largest eddies, rather than to trends in the mean turbulent field. If and how this number of subwindow segments should vary (increase) with the total window length is not clear. Similar to the issue of optimal overall window size, ogive analyses could be helpful here.

With regard to flux data at towers, Foken et al. (2004) consider data within quality classes 1–3 as “suitable for scientific use.” Systematic research into the utility of such quality control systems for airborne flux data has to our knowledge not been published. For the present, we assume quality classes 1–6 to be acceptable for scientific use, depending of the particular objective of the experiment.

Our sensitivity analysis assesses the perturbation of the flux under variation of each calibrated parameter about its optimum value as well as under simultaneous variation of all calibrated parameters. The analysis applied to carbon dioxide flux showed greater sensitivity to parameter variation in some 4-km windows than in others (Fig. 7). This phenomenon may further indicate the level of data quality in individual windows. To help make our quality tests more robust, we plan to apply turbulence quadrant analysis and its related statistics for the relative contribution of updrafts and downdrafts and

countergradient movements. We speculate that this might be a better or at least an additional indicator of quality of fluxes from short windows. Our sensitivity analysis has been a basic analysis of the interactions between the parameters and therefore we plan to investigate higher sophisticated schemes.

The relevance of appropriate quality assurance procedures stems from two facts: 1) given the relatively high costs of airborne measurements, one is tempted to retain as much of the gathered data as possible and to relax any criteria for discarding expensive data; and 2) airborne flux data exhibit more patterns, which are physically hard to interpret, than do tower-based flux datasets. We regard our assessment of the quality of airborne flux data as an attempt like that of Vickers and Mahrt (1997) to create a systematic method that links the airborne and tower flux community by using similar methods of QC–QA. However, more study is needed of the behavior and utility of these QC–QA tests for airborne flux data.

8. Conclusions

The first objective of the present study has been to present a procedure of model calibration of a MFP system installed on the Sky Arrow 650 ERA with a nose-mounted BAT probe. In this respect, adapted versions of the potential flow model and the upwash model have been presented, including a description of the performed in-flight calibration maneuvers. Based on data from a recent calibration flight, an objective procedure for determining the set of model calibration parameters has been presented. Analyses show that the model preforms well with the parameter values determined by our procedure. Although the difference in the average wind field between the two sets is not large ($\Delta \bar{U}_{\text{dir}} = 1^\circ$; $\Delta \bar{U} = 0.4 \text{ m s}^{-1}$), the calibrated set of parameters produces a more stable, measured wind field for the box maneuver ($\bar{w} = -0.01 \text{ m s}^{-1}$; $\sigma_U = 0.19 \text{ m s}^{-1}$; $\sigma_{U_{\text{dir}}} = 1.8^\circ$) than the initial theoretical set of parameters ($\bar{w} = 0.37 \text{ m s}^{-1}$; $\sigma_U = 1.5 \text{ m s}^{-1}$; $\sigma_{U_{\text{dir}}} = 24^\circ$).

Second, the necessary elements for calculating airborne fluxes have been presented, with a major focus on our implemented methods of QC–QA that, to the authors' knowledge, have not been used with or published about airborne flux data before. These methods use assessments of stationarity and integral turbulence characteristics adopted from the tower flux community to provide objective criteria for selection or filtering large datasets, which we propose here for use on airborne fluxes. However, more research is needed to prove the validity of these methods to be used on airborne flux data, especially in relation to the length of the averaging window used to compute fluxes.

The final objective has been to assess the relevance of model calibration to the computation of fluxes of carbon dioxide, energy, and momentum. A sensitivity analysis was performed assessing the effect on the resulting fluxes of varying individual calibration parameters by $\pm 10\%$, using a section of a standard operational flux flight. The fluxes were found insensitive to variations in the temperature recovery factor (RMSE $< 0.8\%$), as expected because of the low Mach number. Calibration parameters that have the strongest effect are those that directly affect the vertical wind component: the sensitivity coefficients for the angle of attack (RMSE: 5.6–8.3%) and for lift-induced upwash (RMSE: 2.8–3.5%), and the correction factor for dynamic pressure (RMSE: 4.3–8.5%). A second test varying all model parameters simultaneously by $\pm 10\%$ produced an overall RMSE of 3.0–7.4% in airborne fluxes, where reductions of the parameter magnitude lead to increased fluxes and vice versa.

In summary, we have given here a complete overview of our procedures of calibration for a typical Sky Arrow with a nose-mounted BAT probe, and the postprocessing and QC–QA of the data coming from its MFP system to retrieve airborne fluxes. Our MATLAB toolbox, called WURMFP toolbox, facilitates this in one consistent software package. As these procedures have a generic character, we expect them to benefit at least those researchers using a SERA with a nose-mounted BAT probe in studies of air–surface interaction.

Acknowledgments. This study was jointly supported by the Dutch national research program “Climate Changes Spatial Planning” (www.climatechangesspatialplanning.nl) and by the Strategic Knowledge Development Program of Wageningen UR on Climate Change (Kennisbasis thema Klimaatverandering), funded by the Dutch ministry of Economic Affairs, Agriculture and Innovation. We thank Martin Blokland of the Aero Company Vliegsschool Teuge BV (ACVT, the Netherlands) for carefully performing the flight maneuvers. Many thanks go to Franco Miglietta at IBIMET-CNR for providing a training opportunity at his institute, which was funded by the Netherlands Organisation of Scientific Research (NWO). Last but not least, many thanks go to Bert Holtslag and Pavel Kabat at the Wageningen University and the reviewers for helping to improve the manuscript.

REFERENCES

- Aubinet, M., and Coauthors, 2000: Estimates of the annual net carbon and water exchange of forests: The EUROFLUX methodology. *Advances in Ecological Research*, A. H. Fitter and D. G. Raffaelli, Eds., Vol. 30, Academic Press, 113–175.
- Bögel, W., and R. Baumann, 1991: Test and calibration of the DLR Falcon wind measuring system by maneuvers. *J. Atmos. Oceanic Technol.*, **8**, 5–18.

- Brown, E. N., C. A. Friehe, and D. H. Lenschow, 1983: The use of pressure fluctuations on the nose of an aircraft for measuring air motion. *J. Climate Appl. Meteor.*, **22**, 171–180.
- Ciais, P., P. Rayner, F. Chevallier, P. Bousquet, M. Logan, P. Peylin, and M. Ramonet, 2010: Atmospheric inversions for estimating CO₂ fluxes: Methods and perspectives. *Climatic Change*, **103**, 69–92.
- Crawford, T. L., and R. J. Dobosy, 1992: A sensitive fast-response probe to measure turbulence and heat flux from any airplane. *Bound.-Layer Meteor.*, **59**, 257–278.
- , R. T. McMillen, R. J. Dobosy, and I. MacPherson, 1993: Correcting airborne flux measurements for aircraft speed variation. *Bound.-Layer Meteor.*, **66**, 237–245.
- , R. J. Dobosy, and E. J. Dumas, 1996: Aircraft wind measurement considering lift-induced upwash. *Bound.-Layer Meteor.*, **80**, 79–94.
- , G. H. Crescenti, and J. M. Hacker, 2001: Small environmental research aircraft: The future of airborne geoscience. Preprints, *11th Symp. on Meteorological Observations and Instrumentation*, Albuquerque, NM, Amer. Meteor. Soc., 5.6. [Available online at https://ams.confex.com/ams/annual2001/techprogram/paper_17684.htm.]
- Dobosy, R. J., and Coauthors, 2013: Calibration and quality assurance of an airborne turbulence probe in an aeronautical wind tunnel. *J. Atmos. Oceanic Technol.*, **30**, 182–196.
- Dolman, A. J., A. D. Culf, and P. Bessemoulin, 1997: Observations of boundary layer development during the HAPEX-Sahel intensive observation period. *J. Hydrol.*, **189**, 998–1016.
- , C. Gerbig, J. Noilhan, C. Sarrat, and F. Miglietta, 2009: Detecting regional variability in sources and sinks of carbon dioxide: A synthesis. *Biogeosciences*, **6**, 1015–1026.
- Dumas, E., S. B. Brooks, and J. Verfaillie, 2001: Development and testing of a Sky Arrow 650 ERA for atmospheric research. Preprints, *11th Symp. on Meteorological Observations and Instrumentation*, Albuquerque, NM, Amer. Meteor. Soc., 5.8. [Available online at https://ams.confex.com/ams/annual2001/techprogram/paper_18395.htm.]
- Eckman, R., 1999: Computation of flow angles and dynamic pressure on BAT probe. NOAA/ARL Field Research Division Tech. Note, 4 pp.
- , T. Crawford, E. Dumas, and K. Birdwell, 1999: Airborne meteorological measurements collected during the model validation program (MVP) field experiments at Cape Canaveral, Florida. NOAA Tech. Memo. ARL/ATDD-233, 54 pp.
- Foken, T., M. Goeckede, M. Mauder, L. Mahrt, B. Amiro, and W. Munger, 2004: Post-field data quality control. *Handbook of Micrometeorology: A Guide for Surface Flux Measurement and Analysis*, X. Lee, W. Massman, and B. Law, Eds., Atmospheric and Oceanographic Sciences Library, Vol. 29, Kluwer Academic Publishers, 181–208.
- French, J. R., T. L. Crawford, R. C. Johnson, and O. R. Cote, 2001: A high-resolution temperature probe for airborne measurements. Preprints, *11th Symp. on Meteorological Observations and Instrumentation*, Albuquerque, NM, Amer. Meteor. Soc., 5.10. [Available online at <https://ams.confex.com/ams/annual2001/webprogram/Paper17705.html>.]
- Garman, K. E., and Coauthors, 2006: An airborne and wind tunnel evaluation of a wind turbulence measurement system for aircraft-based flux measurements. *J. Atmos. Oceanic Technol.*, **23**, 1696–1708.
- , P. Wyss, M. Carlsen, J. R. Zimmerman, B. H. Stirm, T. O. Carney, R. Santini, and P. B. Shepson, 2008: The contribution of variability of lift-induced upwash to the uncertainty in vertical winds determined from an aircraft platform. *Bound.-Layer Meteor.*, **126**, 461–476.
- Gioli, B., and Coauthors, 2004: Comparison between tower and aircraft-based eddy covariance fluxes in five European regions. *Agric. For. Meteorol.*, **127**, 1–16.
- , F. Miglietta, F. P. Vaccari, A. Zaldei, and B. De Martino, 2006: The Sky Arrow ERA, an innovative airborne platform to monitor mass, momentum and energy exchange of ecosystems. *Ann. Geophys.*, **49**, 109–116.
- Hacker, J. M., and T. L. Crawford, 1999: The BAT-probe: The ultimate tool to measure turbulence from any kind of aircraft (or sailplane). *Tech. Soaring*, **23**, 42–45.
- Hall, P. G., E. J. Dumas, and D. L. Senn, 2006: NOAA ARL mobile flux platform instrumentation integration on University of Alabama Sky Arrow environmental aircraft. NOAA Tech. Memo. ARL-257, 41 pp.
- Hutjes, R., O. Vellinga, B. Gioli, and F. Miglietta, 2010: Disaggregation of airborne flux measurements using footprint analysis. *Agric. For. Meteorol.*, **150**, 966–983.
- Isaac, P. R., 2004: Estimating surface-atmosphere exchange at regional scales. Ph.D. thesis, Flinders University of South Australia, 310 pp.
- Johansson, C., A.-S. Smedman, U. Högström, J. G. Brasseur, and S. Khanna, 2001: Critical test of the validity of Monin–Obukhov similarity during convective conditions. *J. Atmos. Sci.*, **58**, 1549–1566.
- John, J. E. A., 1984: *Gas Dynamics*. 2nd ed. Allyn and Bacon, Inc., 426 pp.
- Kalogiros, J. A., and Q. Wang, 2002: Aerodynamic effects on wind turbulence measurements with research aircraft. *J. Atmos. Oceanic Technol.*, **19**, 1567–1576.
- Kirby, S., R. Dobosy, D. Williamson, and E. Dumas, 2008: An aircraft-based data analysis method for discerning individual fluxes in a heterogeneous agricultural landscape. *Agric. For. Meteorol.*, **148**, 481–489.
- Leise, J. A., and J. M. Masters, 1993: Wind measurements from aircraft. NOAA/Aircraft Operations Center Rep., 166 pp.
- Lenschow, D. H., 1986: Aircraft measurements in the boundary layer. *Probing the Atmospheric Boundary Layer*, D. H. Lenschow, Ed., Amer. Meteor. Soc., 39–55.
- Mahrt, L., 1998: Flux sampling errors for aircraft and towers. *J. Atmos. Oceanic Technol.*, **15**, 416–429.
- Mann, J., and D. H. Lenschow, 1994: Errors in airborne flux measurements. *J. Geophys. Res.*, **99** (D7), 14 519–14 526.
- Mauder, M., and T. Foken, 2004: Documentation and instruction manual of the eddy covariance software package TK2. Universität Bayreuth Work Results 26, 45 pp.
- Sarrat, C., and Coauthors, 2009: Mesoscale modelling of the CO₂ interactions between the surface and the atmosphere applied to the April 2007 CERES field experiment. *Biogeosciences*, **6**, 633–646.
- Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*. Atmospheric Oceanographic Sciences Library, Vol. 13, Kluwer Academic Publishers, 670 pp.
- Thomas, C., and T. Foken, 2002: Re-evaluation of integral turbulence characteristics and their parameterisations. [Available online at https://ams.confex.com/ams/BLT/techprogram/paper_43325.htm.]
- Traub, L., and O. Rediniotis, 2003: Analytic prediction of surface pressures over a hemisphere-cylinder at incidence. *J. Aircr.*, **40**, 645–652.
- van Heerwaarden, C. C., J. Vilà-Guerau de Arellano, A. F. Moene, and A. A. M. Holtslag, 2009: Interactions between dry-air entrainment,

- surface evaporation and convective boundary-layer development. *Quart. J. Roy. Meteor. Soc.*, **135**, 1277–1291.
- Vellinga, O. S., B. Gioli, J. A. Elbers, A. A. M. Holtslag, P. Kabat, and R. W. A. Hutjes, 2010: Regional carbon dioxide and energy fluxes from airborne observations using flight-path segmentation based on landscape characteristics. *Biogeosciences*, **7**, 1307–1321.
- Vickers, D., and L. Mahrt, 1997: Quality control and flux sampling problems for tower and aircraft data. *J. Atmos. Oceanic Technol.*, **14**, 512–526.
- , and —, 2006: A solution for flux contamination by meso-scale motions with very weak turbulence. *Bound.-Layer Meteor.*, **118**, 431–447.
- Webb, E., G. Pearman, and R. Leuning, 1980: Correction of flux measurements for density effects due to heat and water vapor transfer. *Quart. J. Roy. Meteor. Soc.*, **106**, 85–100.
- Williams, A., and D. Marcotte, 2000: Wind measurements on a maneuvering twin-engine turboprop aircraft accounting for flow distortion. *J. Atmos. Oceanic Technol.*, **17**, 795–810.
- Zulueta, R. C., W. C. Oechel, H. W. Loescher, W. T. Lawrence, and K. T. Paw U, 2011: Aircraft-derived regional scale CO₂ fluxes from vegetated drained thaw-lake basins and interstitial tundra on the Arctic coastal plain of Alaska. *Global Change Biol.*, **17**, 2781–2802.