Aircraft-based observations of air–sea turbulent fluxes around the British Isles

RMetS
Royal Meteorological Society

Aircraft-based observations of air–sea turbulent fluxes around the British Isles†

Peter A. Cook* and Ian A. Renfrew

Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences, University of East Anglia, Norwich, UK

*Correspondence to: Peter A. Cook, Department of Meteorology, University of Reading, Earley Gate, PO Box 243, Reading, RG6 6BB, UK. E-mail: p.a.cook@reading.ac.uk

Observations of turbulent fluxes of momentum, heat and moisture from low-level aircraft data are presented. Fluxes are calculated using the eddy covariance technique from flight legs typically ∼40 m above the sea surface. Over 400 runs of 2 min (∼12 km) from 26 flights are evaluated. Flight legs are mainly from around the British Isles although a small number are from around Iceland and Norway. Sea-surface temperature (SST) observations from two on-board sensors (the ARIES interferometer and a Heimann radiometer) and a satellite-based analysis (OSTIA) are used to determine an improved SST estimate. Most of the observations are from moderate to strong wind speed conditions, the latter being a regime short of validation data for the bulk flux algorithms that are necessary for numerical weather prediction and climate models. Observations from both statically stable and unstable atmospheric boundary-layer conditions are presented. There is a particular focus on several flights made as part of the DIAMET (Diabatic influence on mesoscale structures in extratropical storms) project.

Observed neutral exchange coefficients are in the same range as previous studies, although higher for the momentum coefficient, and are broadly consistent with the COARE 3.0 bulk flux algorithm, as well as the surface exchange schemes used in the ECMWF and Met Office models. Examining the results as a function of aircraft heading shows higher fluxes and exchange coefficients in the across-wind direction, compared to along-wind (although this comparison is limited by the relatively small number of along-wind legs). A multi-resolution spectral decomposition technique demonstrates a lengthening of spatial scales in along-wind variances in along-wind legs, implying the boundary-layer eddies are elongated in the along-wind direction. The along-wind runs may not be able to adequately capture the full range of turbulent exchange that is occurring because elongation places the largest eddies outside of the run length.

Key Words: surface turbulent fluxes; exchange coefficients; aircraft observations; boundary-layer turbulence; DIAMET

Received 30 August 2013; Revised 27 January 2014; Accepted 5 February 2014; Published online in Wiley Online Library

1. Introduction

The turbulent exchange of momentum, heat and moisture across the air–sea interface is an important contributor to the development of weather systems and a key component of the climate system. Thus it must be accounted for in numerical weather prediction and climate prediction models. This exchange is primarily subgrid-scale and so is parametrized via surface exchange (bulk flux) parametrization schemes. These schemes are semi-empirical and require observations of turbulent exchange and bulk meteorological properties to allow a tuning of the algorithm, typically by estimating exchange coefficients (e.g. Fairall *et al.*, 1996, 2003; Andreas *et al.*, 2012). Due to the turbulent nature of the observations a large amount of random scatter is inherent in any such observational dataset, meaning that a relatively large amount of data is required in order to obtain mean (or median) values for this tuning. Over the last few decades a large amount of data has been assembled (e.g. Fairall *et al.*, 2003; Andreas *et al.*, 2012; Vickers *et al.*, 2013), such that for low to moderate wind speeds there is now reasonable consensus amongst the majority of bulk flux algorithms commonly used. However as wind speeds increase this consensus breaks down – the bulk algorithms diverge – and the physical situation is complicated by the effects of sea-surface waves, swell, wave breaking, white caps,

†The copyright line in this article was amended on 1 May 2014 after original online publication.

Table 1. The 32 flights with significant low-level legs over the sea used in this study.

Campaign	Flight	Date	No of runs	Altitude (m)	Missing data	U_{10N} (m s^{-1})	Stability	Usable runs UW	WT	WQ
GFDEx *Irminger and Iceland Seas*	B268*	21/02/07	11	43	Heim. ARIES Q	22–25	U	10	8	0
	B271*	25/02/07	22	40	Heim. ARIES	5–16	U	19	22	22
	B276*	5/03/07	43	35	ARIES	16–20	U	43	43	43
	B277*	6/03/07	6	40	ARIES	18–22	U	6	6	6
	B278*	9/03/07	30	35	ARIES	15–20	U	29	29	29
VOCALS *Southeast Pacific*	B411	30/10/08	20	30	ARIES	4–8	U	8	14	17
	B417	9/11/08	10	40	ARIES	5–10	U	4	7	10
	B419	12/11/08	15	40	ARIES	4–6	U	15	9	15
T-NAWDEX	B488*	24/11/09	7	35	ARIES	18–19	S	7	7	2
CONSTRAIN	B499*	20/01/10	7	40	ARIES	9–11	U	4	7	7
RONOCO	B567*	18/01/11	4	38	Heim. ARIES Q	10–10	U	4	4	0
	B568	19/01/11	4	34	ARIES Q	5–6	U	2	4	0
COALESC	B574+	24/02/11	15	33	Q	6–8	S	7	3	0
	B578*+	3/03/11	7	34	Q	6–8	U	4	7	0
	B579*	5/03/11	9	34	Heim. Q	7–8	U	9	8	0
VACAR / CAVIAR	B586*+	18/03/11	9	34		6–9	U	7	4	9
DIAMET	B650*+ **IOP3**	23/09/11	11	35		8–12	S	10	11	11
	B652*+ **IOP4**	26/11/11	49	50		13–18	S	49	47	24
EXMIX	B653*+	27/11/11	27	40	ARIES Q	16–18	U	27	27	0
			21	80	ARIES			21	18	16
DIAMET	B656*+ **IOP6**	1/12/11	2	50		20–25	U	2	2	2
			8	75				8	8	7
EXMIX	B659+	10/12/11	7	30		4–6	U	1	7	7
PIK & MIX	B665*+	19/01/12	10	37		15–23	U	10	9	9
	B667*	22/01/12	9	80		11–13	U	6	5	8
	B668*	23/01/12	4	80	Heimann	11–13	U	2	2	1
MEVALI *Norwegian Sea*	B680*+	12/03/12	6	40		16–17	U	4	6	6
	B681*+	13/03/12	8	35		12–15	U	4	8	8
	B684*+	16/03/12	10	40		9–14	U	8	7	5
DIAMET	B695*+ **IOP11b**	10/05/12	25	40		11–13	U	25	25	22
			8	80				8	7	8
MAMM	B721*	23/07/12	15	40		8–10	S	15	13	11
ClearFlo	B724*	30/07/12	13	30		7–10	U	10	8	12
NCAS-Elgin	B727*	15/08/12	1	50	ARIES	8–13	S	1	1	1
			20	80				20	13	8
ACCACIA *Norwegian Sea*	B763*	26/03/13	8	35		12–18	U	8	8	8

The columns tabulate: the campaign (note the location is around the British Isles unless stated otherwise), flight number, date, number of 2 min runs, mean altitude, missing data (both the Heimann and the ARIES system measure SST – see section 3.2; Q indicates missing Lyman-alpha measurements of specific humidity), the range of U_{10N} (10 m neutral reference height wind speed), the stability of the boundary layer (S = Stable or U = Unstable), and the number of usable runs due to the quality of the covariances ($u'w'$, $w't'$ and $w'q'$). Only 26 of the flights are used for the turbulence flux estimations, marked with a *; the other 6 flights have generally low wind speeds and so their data is only included in Figure 11(b). The flights marked with a + are used for the SST comparison.

sea spray and interactions between low-level winds and these surface effects (e.g. Yelland and Taylor, 1996; Yelland *et al.*, 1998; Banner *et al.*, 1999; Andreas and DeCosmo, 2002; Drennan *et al.*, 2003; Fairall *et al.*, 2003; Perrie *et al.*, 2005; Persson *et al.*, 2005; Andreas, 2011). These conditions are precisely those encountered during extratropical cyclones and hurricanes.

In recent years several studies have focused on these high wind speed conditions; for example, Persson *et al.* (2005) present ship-based eddy covariance and inertial dissipation fluxes from the central North Atlantic; French *et al.* (2007), Drennan *et al.* (2007) and Zhang *et al.* (2008) present aircraft-based covariance flux observations from hurricanes; Powell *et al.* (2003) infer exchange coefficients from dropsonde observations of hurricanes; Donelan *et al.* (2004) estimate exchange coefficients from high wind speed wave-tank experiments; Petersen and Renfrew (2009) present aircraft-based covariance fluxes during barrier winds and tip jets; and Raga and Abarca (2007) present aircraft-based covariance fluxes during gap winds. These studies, along with recent reviews (Andreas *et al.*, 2012; Vickers *et al.*, 2013) have started to conclude that there is a 'roll off' in the momentum exchange coefficient for increasing ten-metre wind speed, perhaps as surface waves are flattened off, although there is still some uncertainty in the details of this effect, and indeed in the role of waves at moderate to strong winds too (e.g. Fairall *et al.*, 2003). Recently Andreas *et al.* (2012) presented a new algorithm which has this roll-off feature, is well-validated by a very large dataset and also seems to correspond with wind-wave theory (Moon *et al.*, 2007; Mueller and Veron, 2009). However, further examination of this new algorithm is required. Current understanding of the exchange coefficients for heat and moisture at higher wind speeds is also uncertain and hampered by a lack of observations (e.g. Fairall *et al.*, 2003), although there is increasing evidence that sea spray becomes critical to the transfer of both heat and salt (e.g. Andreas, 2011).

Under high wind speed conditions the air–sea turbulent fluxes of momentum are large, as are the fluxes of heat and moisture if the sea–air temperature difference is significant. Consequently these fluxes are a significant sink of momentum and a significant source or sink of energy and moisture to developing storms such as hurricanes and extratropical cyclones. The DIAMET project (Diabatic influence on mesoscale structures in extratropical storms) is a major UK consortium that aims to improve the understanding and prediction of extratropical storms through a complex programme of observation, parametrization development, data assimilation and numerical weather prediction studies. The overarching theme is on the role that diabatic processes play in developing the meso-to-convective scale structures within storms. Particular foci for the programme have been the role of latent heating through condensation and evaporation, and the role of air–sea fluxes in providing sources/sinks of heating. Recently a number of idealized modelling studies have indicated that the atmospheric boundary layer (ABL) plays a key role in dictating mesoscale structure within developing extratropical cyclones

Table 2. The flights with ABL legs used to estimate turbulence profiles. All flights, bar B568, also have surface-layer legs (see Table 1).

Campaign	Flight	Date	No of runs	Altitude (m)	Stability
VOCALS	B417	9/11/08	5	150	U
	B419	12/11/08	7	160	U
			4	315	
T-NAWDEX	B488	24/11/09	29	300	S
RONOCO	B567	18/01/11	29	160	U
			5	230	
	B568	19/01/11	8	150	U
COALESC	B574	24/02/11	11	160	S
	B578	3/03/11	18	150	U
			6	240	
	B579	5/03/11	4	160	U
			5	300	
EXMIX	B653	27/11/11	4	150	U
			5	300	
EXMIX	B659	10/12/11	16	150	U
			17	300	
MEVALI	B684	16/03/12	13	150	U
NCAS-Elgin	B727	15/08/12	2	150	S
			2	310	

(e.g. Adamson *et al.*, 2006; Beare, 2007; Boutle *et al.*, 2007, 2010; Plant and Belcher, 2007). Testing this role for the ABL in real case-studies is more challenging, but new potential vorticity (PV) budget techniques have been developed for use in both idealised and real numerical modelling studies (e.g. Gray, 2006; Chagnon and Gray, 2009; Chagnon *et al.*, 2013) which allow sources and sinks of PV to be quantified, thus allowing the role of, for example, ABL processes on storm development to be clearly ascertained. A major objective for the DIAMET project is to examine mesoscale structure development in real storms from both an observational and a modelling perspective, using the PV-budget approach to examine the reasons for the development. A necessary step then is the calculation of air–sea turbulent and ABL fluxes, as these are required for validating this PV-budget approach.

In this study we present a compilation of 410 runs from 26 flights over several years. A further 71 runs from 6 flights in low wind speed conditions are also processed, but are not examined in detail. There is a particular focus on flights made during the DIAMET campaigns of 2011–2012. The analysis methodology closely follows that of Petersen and Renfrew (2009, hereafter PR2009). Estimates of turbulent fluxes and exchange coefficients are presented and compared to a number of bulk flux algorithms. In addition an examination of along- versus across-wind differences and the variation of fluxes with height is included.

2. Theory

The eddy covariance method uses high-frequency measurements of the wind velocity components, temperature and humidity to estimate the fluxes of momentum (τ), sensible heat (SH) and latent heat (LH) for a particular time interval or run:

$$\tau = \overline{\rho}\sqrt{\overline{u'w'}^2 + \overline{v'w'}^2}, \quad (1)$$

$$SH = \overline{\rho} c_p \overline{w'\theta'}, \quad (2)$$

$$LH = \overline{\rho} L_v \overline{w'q'}. \quad (3)$$

Here u', v', w', θ' and q' are perturbations of the wind components, potential temperature and humidity from the run average, where $\theta = T + \gamma z$ is a function of the air temperature T and the altitude z. ($\overline{\rho}$) is the run average air density, $c_p = 1004\ \mathrm{J\,kg^{-1}\,K^{-1}}$ is the specific heat capacity for dry air, $L_v = 2.5 \times 10^6\ \mathrm{J\,kg^{-1}\,K^{-1}}$ is the latent heat of vaporization, and $\gamma = 0.00975\ \mathrm{K\,m^{-1}}$ is the adiabatic lapse rate. Following Donelan (1990), a small correction to the momentum flux is made to account for the fact that the surface layer is not an exact constant flux layer. This correction is typically a few per cent and is generally, but not always, applied in such studies. Examination of a few cases, where we have profiles of fluxes in similar conditions, suggests this correction is worth using (not shown).

Numerical weather and climate prediction models parametrize these fluxes as a function of 'bulk' meteorological parameters usually using exchange coefficients to relate the bulk values to the fluxes (e.g. Fairall *et al.*, 1996, 2003). Typically these exchange coefficients are determined for neutral stability and then stability corrections are applied using standard stability correction functions. Our treatment here follows that detailed in PR2009. In brief, the bulk flux equations are:

$$\tau = \overline{\rho} C_{DN} (U_{10N} - U_S)^2, \quad (4)$$

$$SH = \overline{\rho} c_p C_{HN} (U_{10N} - U_S)(\theta_S - \theta_{10N}), \quad (5)$$

$$LH = \overline{\rho} L_v C_{EN} (U_{10N} - U_S)(q_S - q_{10N}), \quad (6)$$

where C_{DN}, C_{HN} and C_{EN} are the neutral exchange coefficients for drag (momentum), heat and evaporation (moisture) respectively, U_{10N}, θ_{10N} and q_{10N} are values at the neutral ten-metre reference level and the subscript S denotes surface values. Note U_S, the surface velocity, is assumed zero. In regions of strong currents this assumption may incur errors of order 10% – see Zhai *et al.* (2012) for a discussion of the implications here – however around the British Isles this error should be smaller. Calculating the exchange coefficients and neutral 10 m variables are achieved through the evaluation of surface roughness lengths and scaling parameters using standard equations and stability correction factors in the usual manner, as detailed in PR2009.

The vast majority of turbulence runs are at ∼40 m altitude. However there are 70 runs at ∼80 m altitude (Table 1). An examination of flights with both ∼40 and ∼80 m runs showed that when fluxes, friction velocities (u_*), roughness lengths (z_0), and 10 m neutral wind speeds (U_{10N}) were calculated, the U_{10N} values from the ∼80 m runs systematically underestimated U_{10N} (by 10% on average) when compared to the U_{10N} values from the ∼40 m runs (for the same meteorological conditions). This is a consequence of the actual boundary-layer profiles having little change in the wind speed between ∼40 and ∼80 m. Hence for all the runs at ∼80 m the U_{10N} values were increased by 10% while the calculated fluxes were not altered. The greater U_{10N}values then reduced the calculated bulk flux coefficients at ∼80 m (C_{DN}, C_{HN} and C_{EN}).

3. Methodology

3.1. The research aircraft and instrumentation

The data analysed here were all obtained from the UK's Facility for Airborne Atmospheric Measurements' (FAAM's) BAe-146, jointly operated by the Met Office and the Natural Environment Research Council (NERC). This versatile four-engine jet aircraft is able to fly at a minimum safe altitude down to 100 feet (∼35 m) above the sea surface for straight and level runs at its standard science speed of 200 knots (∼100 m s^{-1}). The FAAM has been in operation since 2004, although many of the instruments and expertise were transferred from its predecessor in the UK, the Met Office's Hercules C130 (e.g. Nicholls, 1978). Further details on the aircraft and its capability are described in e.g. Renfrew *et al.* (2008). The key instruments for this study include the five-port pressure measurement system on the nose of the aircraft which, along with static pressure ports and the inertial navigation unit (INU) system, provides wind velocity components at 32 Hz. The turbulence probe requires frequent calibration and checks, for example, carried out by specific calibration flight manoeuvres (P. R. A. Brown, personal communication). Some discussion of this can be found in PR2009, who state that overall the uncertainty in horizontal wind measurements is estimated

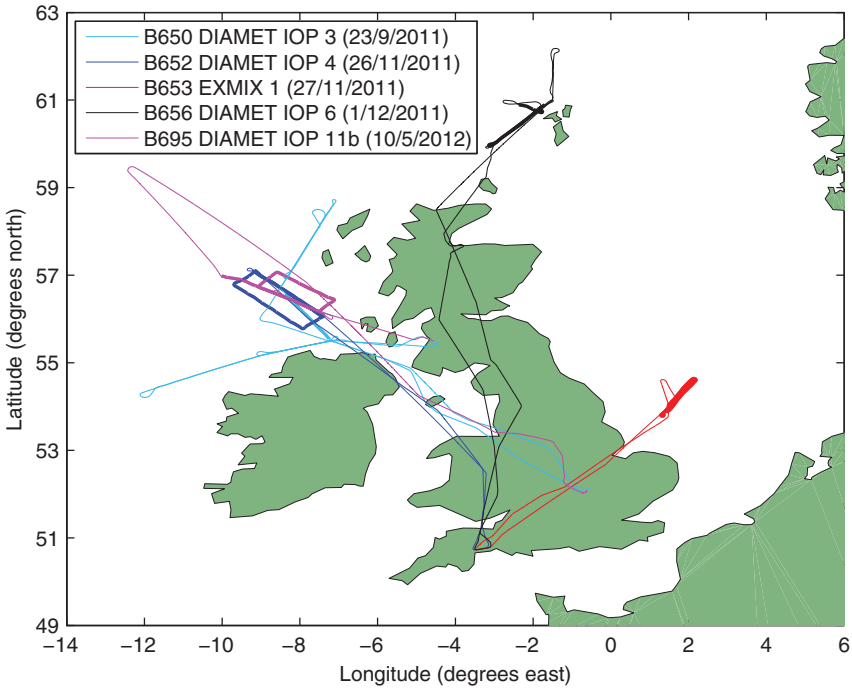

Figure 1. Flight tracks (thin lines) of selected missions, with the low-level components overplotted (thick lines). All the DIAMET project flights with low-level legs are included.

to be $< \pm 0.5$ m s^{-1} (and <0.27 m s^{-1} in the calibration flight highlighted). A Rosemount 102BL provides temperature at 32 Hz, but due to the Rosemount housing the instrument response rate is slower than quoted. To partly alleviate this problem a filtering algorithm, following Inverarity (2000) and MacCarthy (1973), is applied which improves the response to around 7 Hz. The temperature measurement uncertainty is ± 0.3 °C (at 95% confidence) for a typical clear-air measurement, with relative errors <0.01 °C. A Lyman-alpha hygrometer provides specific humidity with an uncertainty of ± 0.15 g kg^{-1}. The aircraft's altitude during low-level runs is from a radar altimeter which records at 2 Hz and has an uncertainty of $\pm 2\%$ below 760 m (2500 ft), so that at 40 m the uncertainty is $<\pm 1$ m. Further details on instrument accuracy and basic quality control can be found in Renfrew *et al.* (2008) and specifically for turbulence measurements in PR2009.

3.2. *The dataset*

The data analysed in this study are from 32 flights of the BAe-146 between 2007 and 2013 (Table 1). It is a compilation of all appropriate flights from the first decade of BAe-146 use. The data come from a relatively small number of flights when significant periods of straight low-level legs were part of the mission, i.e. when air–sea flux or below cloud-base legs were a key objective of the field campaign. In total, 481 runs have been analysed over a wind speed range of U_{10N} from 5 to 24 m s^{-1} and with most data in the moderate to strong wind speed range of 8–20 m s^{-1}. A subset of six flights which have relatively low wind speeds and fluxes are not included in the main analysis – the data quality is poorer and this study focuses on moderate to high wind speeds – although they are noted in Table 1 and used for the SST comparison and later in Figure 11(b). Typically the legs were flown at 35–40 m above the sea surface, with a minority flown at 80 m. None of these turbulence data has been analysed before with the exception of the first five flights in Table 1 which are from the Greenland Flow Distortion Experiment (GFDEx) – see PR2009. Some of the flights also have higher ABL legs (Table 2) providing a further 190 runs where mean covariances and hence ABL fluxes can be calculated. However, surface fluxes and near-surface variables (e.g. U_{10N}) cannot be reliably evaluated from these legs.

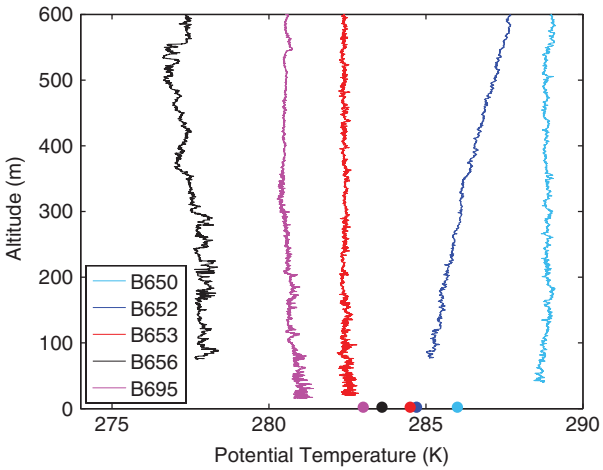

Figure 2. Profiles of potential temperature from the closest descent or ascent to the low-level runs on five DIAMET flights. Mean sea-surface temperatures are also shown as filled circles.

A number of flights from 2011 to 2012 are from the DIAMET field campaigns and Figure 1 shows several flight tracks from around the British Isles, illustrating typical flight patterns. The DIAMET flights are a particular focus in some parts of this study, and furthermore the fluxes from these flights are being used in case-study investigations of the associated storms.

Figure 2 shows potential temperature profiles from these five flights along with mean SST values. Flights B650 and B652 took place in stable boundary-layer conditions, while flights B653, B656 and B695 were in unstable conditions.

3.3. *Sea-surface temperature measurements*

According to PR2009 the largest source of measurement uncertainty in their analysis was in estimating the sea-surface temperature (SST). They used the FAAM's downward pointing Heinmann radiometer, which measures upwelling infrared radiation in the range 8–15 μm at 4 Hz, to obtain an SST. Here a surface emissivity must be specified (set at 0.987) and the reflected downwelling radiation must be accounted for, in their case neglected, leading to an accuracy of ± 0.7 K in SST. This is assuming the Heimann is calibrated in flight – a procedure that

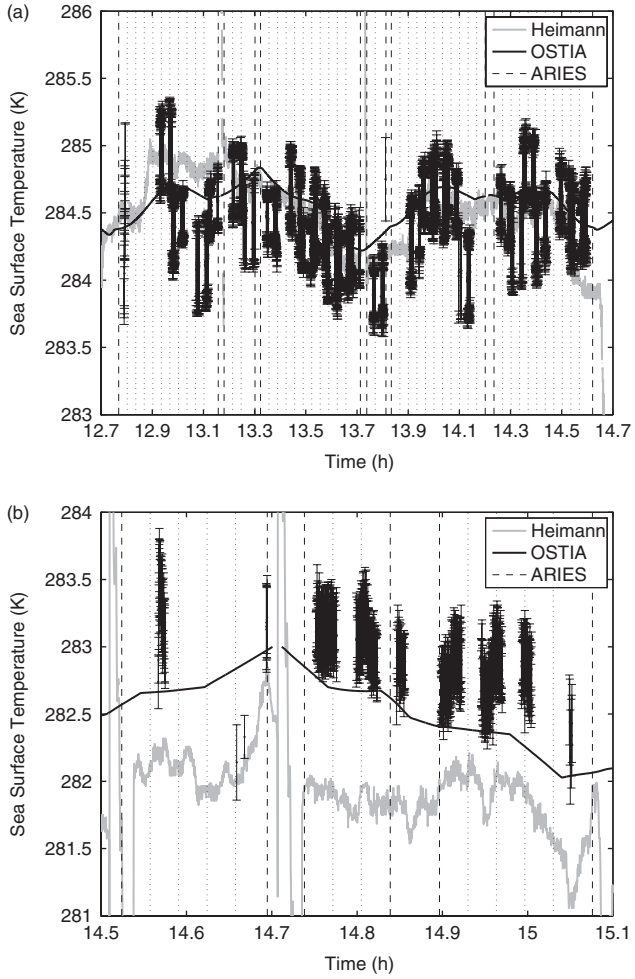

Figure 3. Time series of sea-surface temperature (SST) with measurements from the Heimann infrared thermometer (grey line), the ARIES interferometer (error bars) and the OSTIA satellite analyses (black line). Vertical dashed lines indicate the start and end of low-level legs, while vertical dotted lines are plotted between each run (every 2 min). Panel (a) is from B652 (DIAMET IOP4) on 26 November 2011 and panel (b) is from B665 on 19 January 2012.

does not always occur. In Renfrew *et al.* (2009a) the Heimann SST from GFDEx was compared against the OSTIA SST (the Operational Sea-surface Temperature and sea Ice Analysis) which is the current global analysis product from the Met Office, available at 1/20° resolution once a day. The comparison was reasonable – the correlation coefficient was 0.9, the linear regression slope was 0.78, there was a bias of 1.7 K and the root-mean-square (r.m.s.) error was 1.6 K – although worse than the aimed-for accuracy of an r.m.s. of 0.8 K (Stark *et al.*, 2007). The high-latitude location of the GFDEx data may have contributed to the greater r.m.s. error.

Here we have made use of SST from the Heimann and OSTIA again, as well as the ARIES interferometer where available (see Table 1). The Airborne Research Interferometer Evaluation System (ARIES) instrument measures in the range 3–18 μm and can be rotated from downward to upward pointing during a flight, thus allowing for an excellent estimate of downwelling radiation and hence a more accurate retrieval of SST (Wilson *et al.*, 1999). The retrieval of SSTs follows Newman *et al.* (2005) and incorporates both down and upwelling measurements, so should be accurate to ±0.3 K (S. Newman, personal communication).

Figure 3 illustrates co-located SST measurements from the Heimann and ARIES instruments and the OSTIA analysis for two flights. In both examples there is a good correspondence between the three sets of measurements – the SST gradients are captured in all the data, although with higher resolution in the Heimann (∼25 m) compared to the OSTIA (∼6000 m). The intermittent pattern of the ARIES measurements is due to the sampling procedure noted above, i.e. the gaps are when the instrument is pointing up. These gaps mean that ARIES measurements alone cannot provide a continuous time series of SST measurements for surface flux calculations. Instead the (most accurate) ARIES measurements are used here to provide validation for the Heimann and OSTIA SSTs. In Figure 3(a) all three measurements of SST agree to within error estimates; in Figure 3(b) the measurements are offset from one another, with ARIES ∼0.5 K higher than the OSTIA, which is ∼0.5 K higher than the Heimann. Other SST time series comparisons (not shown) present similar reasonably good spatial correspondences, but show a variety of small offsets between the three estimates.

Figure 4 shows scatter plots of SST from all available data (12 flights, see Table 1). Figure 4(a) shows that the ARIES and OSTIA measurements compare well overall – the r.m.s. error is 0.43 K, the linear regression slope is 0.79 and the bias (0.12 K) is small (see Table 3). Making the assumption that the ARIES instrument is accurate to within its stated ±0.3 K, this suggests that the OSTIA analysis is also reasonably accurate and reliable. There are a few outlier points, from flight B574, although most data from this flight correspond well. Figure 4(b) shows that the ARIES and Heimann measurements can compare very well

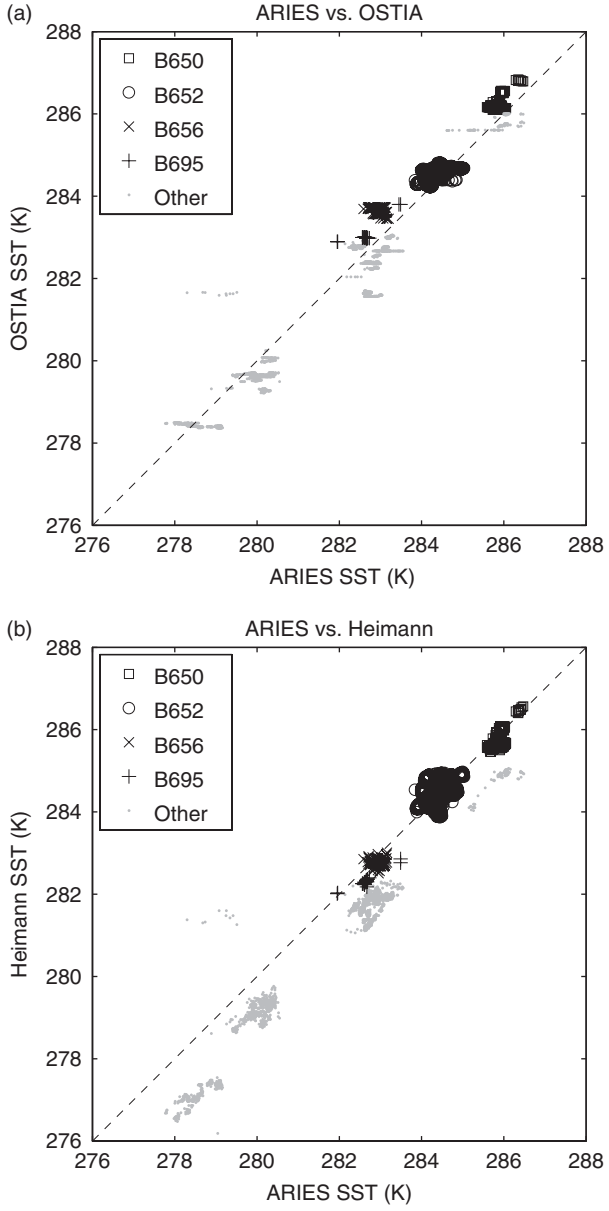

Figure 4. Comparisons of co-located SST: (a) ARIES versus OSTIA measurements and (b) ARIES versus Heimann measurements. In total over 4000 measurements (at one-second intervals) from 12 flights are shown, with four DIAMET flights highlighted via different symbols. The dashed line marks the 1:1 correspondence.

Table 3. Comparison statistics for co-located measurements of SST from the Heimann infrared thermometer (radiometer), an ARIES spectral derivation (interferometer) and the OSTIA satellite-based analysis; the columns show: mean correlation coefficient, mean slope of a linear regression, mean bias and the root-mean-square error.

Comparison	Correlation coefficient	Slope	Bias (K)	RMS error (K)
Heimann vs. OSTIA	0.58	0.90	−0.75	0.79
ARIES vs. Heimann	0.59	0.93	0.81	0.87
ARIES vs. OSTIA	0.45	0.79	0.12	0.43

(e.g. flights B650, B652 and B656), but often have a systematic offset of around −1 K. It seems likely that in some flights the Heimann did not undergo a calibration, leading to a significant bias. A linear regression slope of 0.93 (Table 3) and the high degree of clustering seen in Figure 4(b) implies this bias can be corrected for on a flight-by-flight basis by a simple offset. On flights without co-located ARIES measurements such a bias may still exist and still require a correction. Hence our approach has been to use the reasonably accurate and reliable OSTIA analysis to estimate a constant offset for each flight, which is then applied to the higher-resolution Heimann measurements. The corrected Heimann SSTs are then used in our surface flux calculations. Note in four flights the Heimann measurements are not available (Table 1), so in these OSTIA SSTs are used instead.

3.4. *Flux calculation procedure*

The turbulent fluxes are calculated from runs with minimal changes in heading or altitude. Each run should sample approximately homogeneous conditions and also be long enough to include several wavelengths of all turbulent eddies, so the run length is a compromise. Here we follow PR2009 in choosing runs of 2 min (∼12 km) – the last run on any leg includes the remaining time so these may be up to 4 min. Runs of ∼12 km appear reasonable in the context of previous studies (see Mahrt (1998) for a discussion), longer than the 4 km runs used by Vickers *et al.* (2013), but shorter than some other aircraft studies which use full leg lengths – for example, of 14–54 km in French *et al.* (2007). It is usually assumed that the maximum scaling of the turbulent eddies is approximately the depth of the ABL, i.e. up to 2–3 km at most, so 12 km should be sufficiently long to capture most turbulent transfer (PR2009), although we will come back to this assumption later. The turbulence calculations are carried out at 32 Hz resolution (∼3 m) with all data either resampled (q) or interpolated (SST, altitude) to this resolution. All turbulent variables are linearly detrended for each run before the fluxes are calculated using Eqs (1)– (3) this mostly removes the mesoscale structures but the majority of the turbulence is on much smaller spatial scales than the run length and so is unaffected.

3.5. *Quality control*

A careful quality-control procedure is followed for each flux run, e.g. French *et al.* (2007) and PR2009. In brief this involves checking that power spectra of all turbulent variables (u, w, θ, q, where u is the along-wind velocity component) have a well-defined decay slope (close to $k^{-5/3}$ for wave number k). Then, checking that the cumulative summation of the covariances of

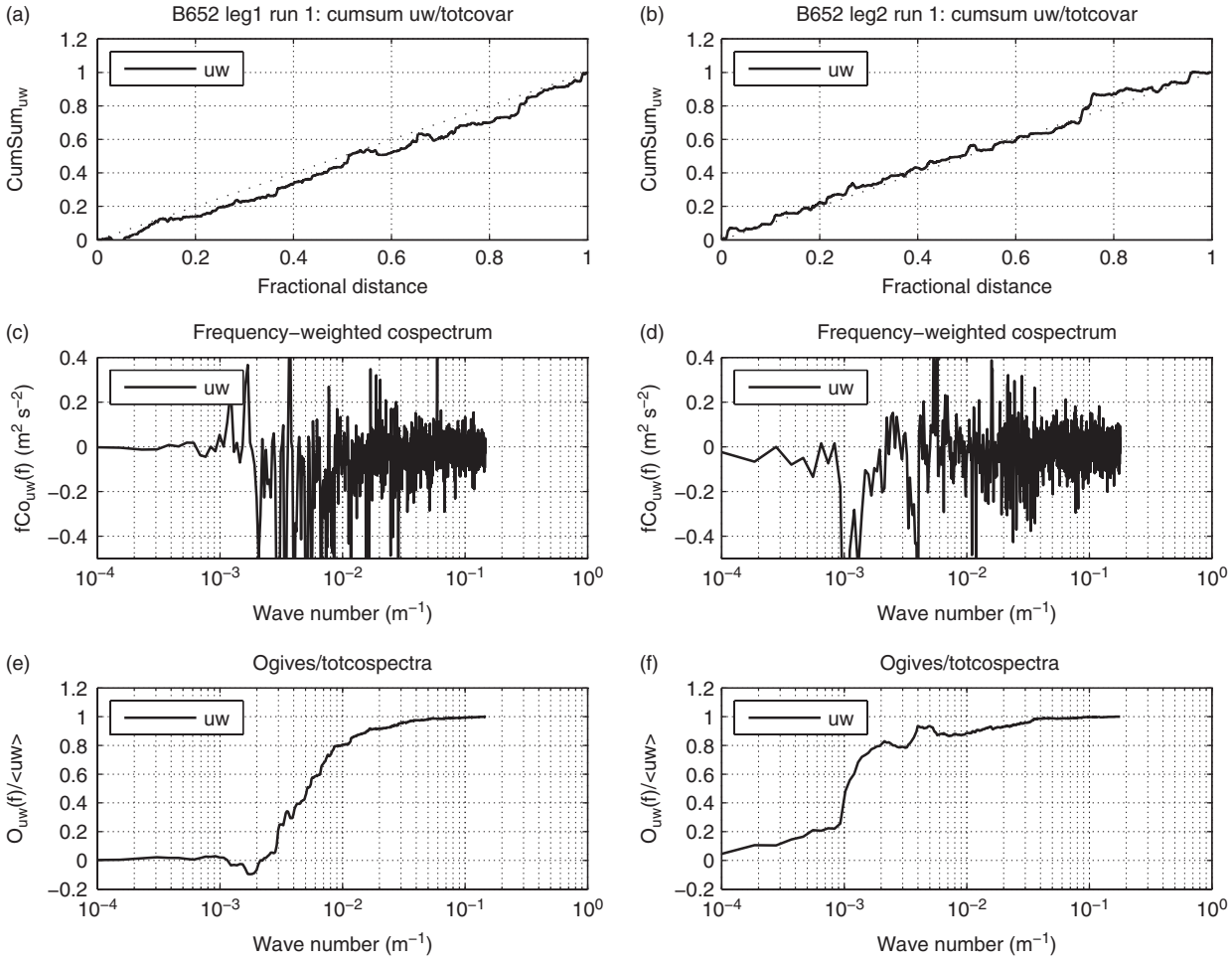

Figure 5. Quality control of the along-wind covariances $u'w'$ for two runs from flight B652. The left column shows an across-wind run, the right column shows an along-wind run. The panels show (a), (b) the cumulative summation as a function of fractional distance; (c), (d) the frequency-weighted co-spectra as a function of wave number; and (e), (f) the ogives (integrated co-spectra) as a function of wave number. The cumulative summation is normalised by the total covariance and the ogives by the total co-spectra.

w and u, θ or q are close to a near-constant slope; checking the co-spectra of the covariances have little power at wave numbers smaller than about 10^{-4} m^{-1}; and checking the cumulative summation of the co-spectra are shaped as ogives (S-shaped) with flat ends. Figure 5 provides an example of along-wind and across-wind components from 'good' momentum flux runs: this illustrates that most of the covariance is between ∼100 and 1000 m, justifying a run choice of 12 km and the resolution of the measurements. Note that across-wind covariances are 1–3 orders of magnitude smaller than the along-wind covariances so these are not used to discard runs. The vast majority of flux runs pass these quality-control checks, with the numbers of usable runs noted in Table 1.

4. Turbulent fluxes in the surface layer

The high-frequency wind, temperature and humidity data from the accepted low-level flux runs are used to calculate surface turbulent fluxes via Eqs (1)– (3), as described in section 2. Figure 6 shows the sensible heat (SH) flux as a function of $U_{10N}(\theta_s - \theta_a)$, where subscripts s and a denote surface and air, and the latent heat (LH) flux as a function of $U_{10N}(q_s - q_a)$. Presented in this manner the data should be linearly proportional if the bulk flux algorithms Eqs (5)–(6) hold, and if the coefficients are constant. Figure 6 shows that this is generally the case, although there is considerable scatter in the fluxes, especially at higher wind speeds. Figure 6(a) shows a clear linear correspondence; the air–sea temperature differences are, at most, around −10 K so the SH flux is limited to ∼300 W m^{-2}. Most of the SH flux is positive – a flux of heat from the ocean into the atmosphere –but there are some runs of negative heat flux such as in B650 (DIAMET IOP3). A linear regression line is fitted to the data giving SH = 5 W m^{-2} when this passes through $U_{10N}(\theta_s - \theta_a) = 0$, so very close to the zero value the bulk algorithm predicts. This implies our measurements are of high quality and, in particular, the SST corrections used are appropriate.

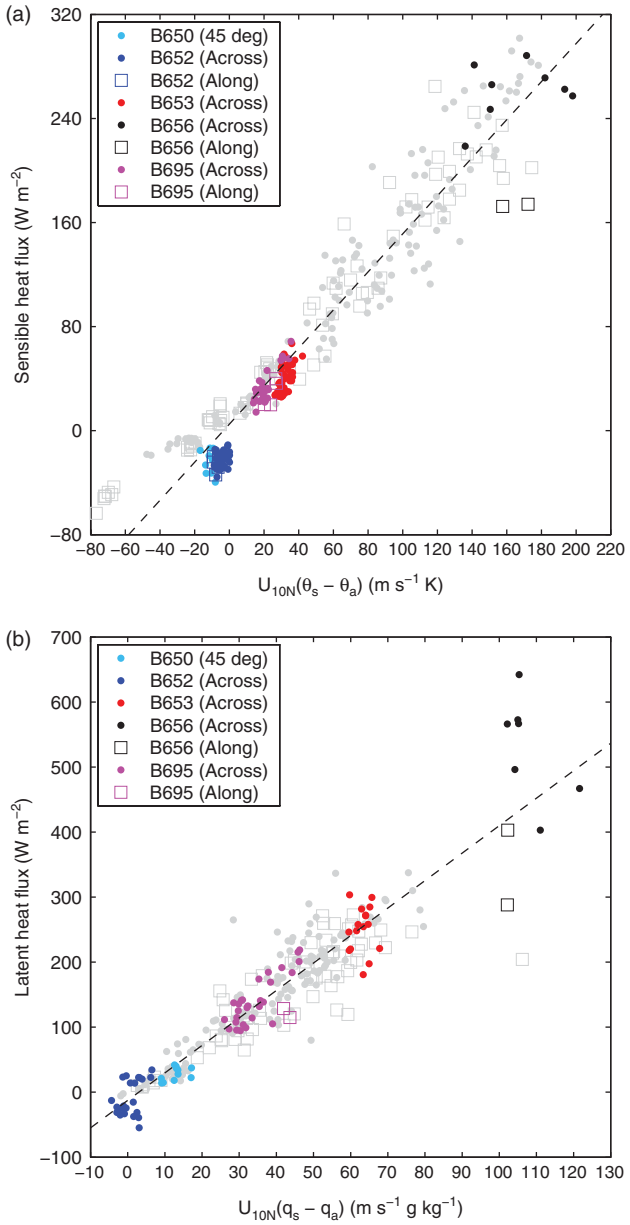

Figure 6. Turbulent flux observations from all low-level runs. Panels show (a) SH flux versus $U_{10N}(\theta_s - \theta_a)$ and (b) LH flux versus $U_{10N}(q_s - q_a)$. A linear regression line is shown for each panel. Across-wind runs (dots) are distinguished from along-wind runs (squares). Several DIAMET flights are highlighted.

Figure 6(b) shows LH flux is also mainly positive – a flux of heat out of the ocean associated with evaporation – although some runs have a negative flux. Again there is a reasonable linear relationship and the regression line has LH = −13 W m^{-2} at $U_{10N}(q_s - q_a) = 0$, further confirmation of the quality of our measurements. There is considerable scatter in flight B656 (DIAMET IOP6) in the unstable very high wind speed conditions, and here the LH fluxes are amongst the highest ever directly observed – similar values are shown by Grossman and Betts (1990) and Raga and Abarca (2007). A scatter plot of U_{10N}^2 versus wind stress (not shown) illustrates that most of these data fall in the range $U_{10N} \approx 8$–20 m s^{-1} with associated stress of up to ∼1.5 N m^{-2}. The highest stresses (up to 3 N m^{-2}) are associated with two flights B268 and B656 where flight-level winds were extraordinarily high, the ABL was very turbulent and the scatter in stress is very large – as would be expected as the sampling error scales with the flux (Donelan, 1990). Flight B268 samples an easterly tip jet off Greenland (Renfrew *et al.*, 2009b).

5. Bulk flux algorithms

Neutral exchange coefficients have been derived and are presented as a function of U_{10N} in Figure 7. There is considerable scatter in all of the exchange coefficients consistent with the random sampling error inherent in such observations and unaccounted-for physical effects such as surface wave interactions and sea spray. The ranges of exchange coefficients seen here are similar to previous studies (e.g. see Fig. 8 in PR2009). Some outlier values of C_{HN} in Figure 7(b) are for runs where the air–sea temperature difference is very small and the calculation is not well-posed. The observed increase in C_{DN} with U_{10N} is similar to many previous studies; for example Vickers *et al.* (2013) categorise their results as C_{DN} increasing linearly for the range U_{10N} from 10 to 20 m s^{-1}. Figure 7(d) plots u^* against U_{10N}, as advocated by Andreas *et al.* (2012). The data cluster reasonably well for much of the wind speed range, becoming more scattered for the highest winds.

The inherent scatter in turbulent flux estimates means the data must be placed into bins for a comparison to bulk flux algorithms. We have used bins every 2 m s^{-1} between 6 and 24 m s^{-1}, with bin means ±1 standard deviation. Table 4 details these quantities and notes the number of data points in each bin. The bulk flux algorithms are those of (i) Smith (1988); (ii) the Coupled Ocean–Atmosphere Response Experiment (COARE) 3.0 algorithm (Fairall *et al.*, 2003); (iii) the European Centre for Medium-range Weather Forecasts (ECMWF) algorithm (see http://www.ecmwf.int/research/ifsdocs/); (iv) the Met Office algorithm (Edwards, 2007); and (v) the Andreas *et al.* (2012) algorithm for momentum only. Note the ECMWF algorithm lies underneath the Met Office algorithm for C_{DN}. The ECMWF algorithm uses a Charnock constant of 0.018 in its uncoupled models, with this value provided by the wave model in its coupled models.

In general, the bulk flux algorithms correspond well with the observations. They lie within the error bars, i.e. within ±1

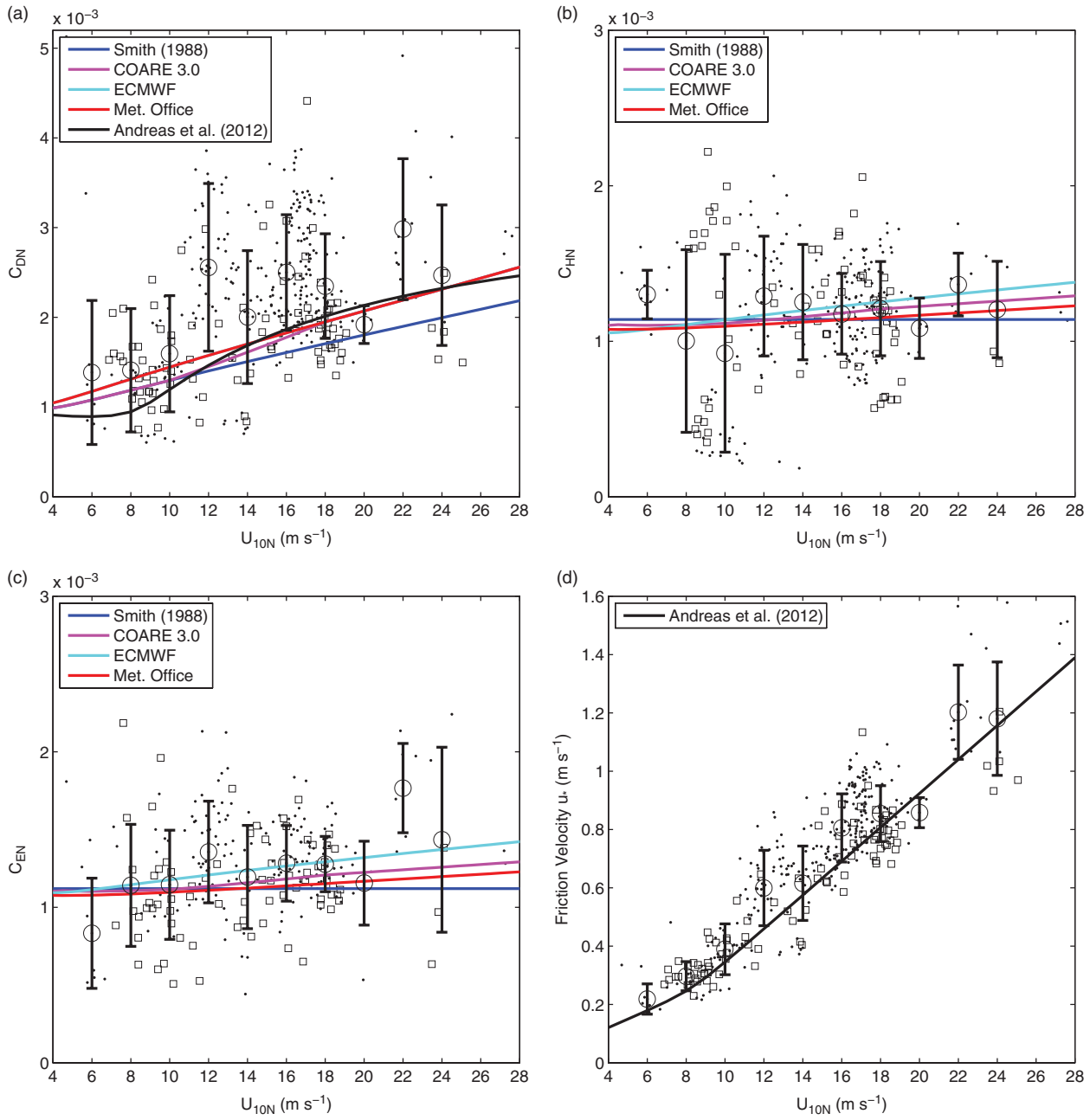

Figure 7. The 10 m neutral exchange coefficients for (a) momentum flux (C_{DN}); (b) sensible heat flux (C_{HN}); and (c) latent heat flux (C_{EN}), as a function of the 10 m neutral wind speed (U_{10N}). Panel (d) shows friction velocity (u_*) versus U_{10N}. Across-wind runs (dots) are distinguished from along-wind runs (squares). Error bars showing the mean and standard deviation for 2 m s^{-1} bins are plotted (see Table 4 for details). Several bulk flux algorithm relationships are overlaid.

standard deviation of the bin mean, with one or two exceptions. However, for momentum all of the curves are below the bin-means. Comparing to previous studies, the observed mean C_{DN}'s are higher than some studies (e.g. Fairall *et al.*, 2003; Persson *et al.*, 2005), but not that different from others (e.g. Vickers *et al.*, 2013). The range of observations is generally greater than the spread of the flux algorithms, making it difficult to draw conclusions about their performance, although it does appear that the Smith (1988) C_{DN} relationship may be a worse fit than the others for high wind speeds. The Smith (1988) algorithm uses a modified Charnock relation: $z_0 = \alpha_c \frac{u*^2}{g} + b\frac{\upsilon}{u*}$ with a Charnock constant $\alpha_c = 0.011$, g the gravitation constant, b the 'smooth flow' constant (often $b = 0.11$) and υ the dynamic viscosity. In contrast the Met Office and ECMWF algorithms set $\alpha_c = 0.018$ and the COARE 3.0 algorithm linearly increases α_c from 0.011 to 0.018 as U_{10N} increases from 10 to 18 m s^{-1}. Our results suggest the latter approaches are more appropriate. One controversial feature in such algorithms has been a flattening off and decrease in C_{DN} for very high wind speeds (>20 m s^{-1}) as implied by the Andreas *et al.* (2012) curve for example. Our results indicate a downturn between the 22 and 24 m s^{-1} bins, although there are relatively few data points in these bins, so this feature is not that well defined.

Figure 7(b) shows C_{HN} and suggests a very good correspondence between the algorithms and the bin-mean observations. Again there is considerable spread in the observations, but all of the algorithms are close to the bin-means over the entire ranges of wind speeds. For higher U_{10N} there is some support for a slightly elevated C_{HN} consistent with the COARE 3.0, ECMWF and Met Office algorithms. The conclusions are rather similar for C_{EN}, with all the algorithms broadly consistent with the bin-mean observations and some support for an elevated C_{EN} for the highest wind speeds (>20 m s^{-1}). However it should be noted there are relatively few data points in these higher U_{10N} bins and our results are not inconsistent with a constant C_{HN} or C_{EN}.

Figure 7(d) shows the friction velocity (u^*) against U_{10N} and compared to the Andreas *et al.* (2012) algorithm. Here the correspondence is very good. It is noticeable that the spread appears reduced when plotting the momentum observations in this way.

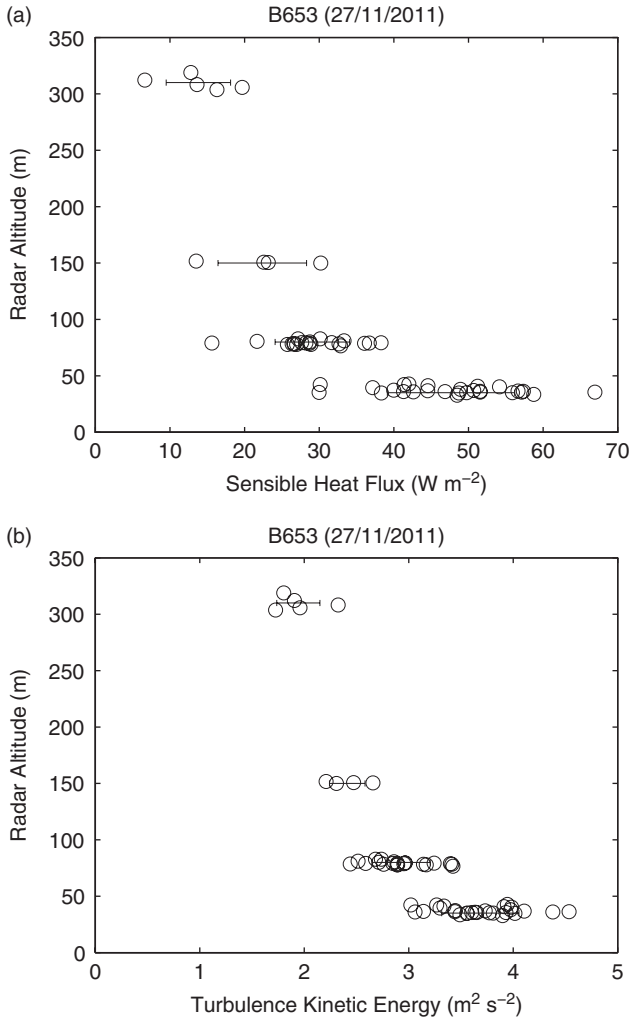

Figure 8. Profiles of (a) sensible heat flux and (b) turbulent kinetic energy for flight B653 (27 November 2011). Circles show run averages, horizontal lines show the mean and standard deviation for each altitude range.

Measured fluxes and turbulent kinetic energy (TKE) generally decrease with altitude in the ABL, with the greatest rate of decline at lower altitudes. Figure 8 shows the sensible heat flux and TKE values from flight B653 which had runs at four altitude ranges (although recall that only the runs below 100 m are used in the main study). Note that in some studies interpolation from measurements at relatively high altitudes have been used to estimate surface fluxes. It is clear from the profiles shown here that this may not be justified. Profiles of fluxes through the ABL will be examined in future work on particular case-studies, but are not examined in general here.

6. Across- and along-wind variability

In Figures 6 and 7, runs that are across the wind (the majority) are distinguished from those that are along the wind. Careful examination reveals that generally the along-wind runs have lower fluxes and lower exchange coefficients than the across-wind runs for the same U_{10N}. Such a difference may indicate an instrumental problem, or a difference due to aircraft sampling that is dependent on the meteorological conditions. Comparisons of mean wind and variance components in the calibration flights do not show such differences, so suggest that this feature is not an instrument problem. Rather as it occurs in only some flights, we suggest it is related to the meteorological conditions, as has been found in a few previous studies (e.g. Nicholls, 1978; Nicholls and Readings, 1981; Chou and Yeh, 1987; Kalogiros and Wang, 2011).

To investigate this further, power spectra of across- and along-wind legs have been examined, for example legs 1 and 2 from flight B652 (Figure 9). Each plotted spectra is the mean of the spectra from the individual runs on that leg. This flight was in stable conditions with U_{10N} = 13–18 m s^{-1}, although similar plots in unstable conditions with similar U_{10N} were also carefully examined (e.g. flight B695, not shown) and illustrated similar features. The left-hand panel (a) shows along-wind velocity spectra, the right-hand panel (b) shows across-wind velocity spectra. All the velocity spectra show a well-defined decay in the inertial subrange ($k \sim 5 \times 10^{-3}$ to 10^{-1} m^{-1}) that closely follows a $k^{-5/3}$ power law. However at smaller wave numbers (k between 1 and 5×10^{-3} m^{-1}, i.e. scales of 200–1000 m) there is some divergence of the curves, most obviously in the along-wind velocity spectra, with significantly more power in the across-wind leg than the along-wind leg. At the very smallest wave numbers ($k < 10^{-3}$ m^{-1}, i.e. scales >1000 m) the curves cross and there is more power in the along-wind legs. The same pattern is seen in other legs in these flights and in other flights (e.g. B695 and B656, not shown). The co-spectra and ogives for these along-wind runs also show some power at the longest wavelengths, although it is arguable whether they should fail the quality-control check.

In short, the across-wind legs contain more velocity variance at scales of 200–1000 m, whereas the along-wind legs contain less velocity variance at these scales and imply some power is shifted to longer scales (greater than 1 km) which may not be fully captured in our 12 km runs.

The differences in fluxes and spectra in along-wind and across-wind legs appear to be due to the sampling pattern of the aircraft under certain meteorological conditions. To characterise the observed differences we employ a multi-resolution decomposition technique that is able to attribute variance to particular scales (Howell and Mahrt, 1997; Vickers and Mahrt, 2003). Essentially an aircraft run is progressively divided into smaller and smaller sub-runs – analogous to a sequence of high-pass filters being applied – with estimates then made of the velocity variance at the scales associated with these sub-runs. Normally the runs are divided into 2, 4, 8, 16, etc. sub-runs, but in this study finer divisions are used: 2, 3, 4, 5, 6, 8, 10, 12, 15, etc. to accurately determine the spatial scales, even though this means that the points on the spectra are not independent. The result is an estimate of velocity variance at each particular time- or length-scale. Here multi-resolution decomposition is applied to the across- and along-wind legs from numerous flights. Figure 10 illustrates the results for B652 (stable conditions), the same flights and legs as Figure 9, and B695 (unstable conditions). Generally the along-wind velocity variances Figure 10(a), (c) are greater than the across-wind variances (b), (d). Focusing on the left panels (a), (c), the across-wind legs have more power at shorter time- and length-scales – the peak power is between 100 and 1000 m and

Table 4. Average neutral exchange coefficients (multipied by 10^3) in 2 m s^{-1} bins, as well as standard deviations (stdev) for each bin, and the number of data points in each in bin (see also Figure 7).

U_{10N}	C_{DN}			C_{HN}			C_{EN}		
bins	Mean	stdev	Number	Mean	stdev	Number	Mean	stdev	Number
4	–	–	1	–	–	1	–	–	1
6	1.39	0.80	8	1.30	0.16	7	0.83	0.35	7
8	1.41	0.69	28	1.00	0.59	16	1.14	0.39	14
10	1.59	0.65	40	0.92	0.64	31	1.14	0.35	26
12	2.56	0.93	52	1.29	0.39	43	1.36	0.33	51
14	2.00	0.74	42	1.25	0.37	22	1.19	0.33	27
16	2.50	0.64	87	1.18	0.26	63	1.28	0.24	50
18	2.35	0.58	90	1.21	0.30	82	1.28	0.18	60
20	1.92	0.21	8	1.08	0.19	8	1.16	0.27	7
22	2.98	0.79	11	1.36	0.20	6	1.77	0.29	6
24	2.47	0.78	9	1.20	0.31	4	1.43	0.59	5
26	–	–	1	–	–	0	–	–	0
28	–	–	3	–	–	3	–	–	2

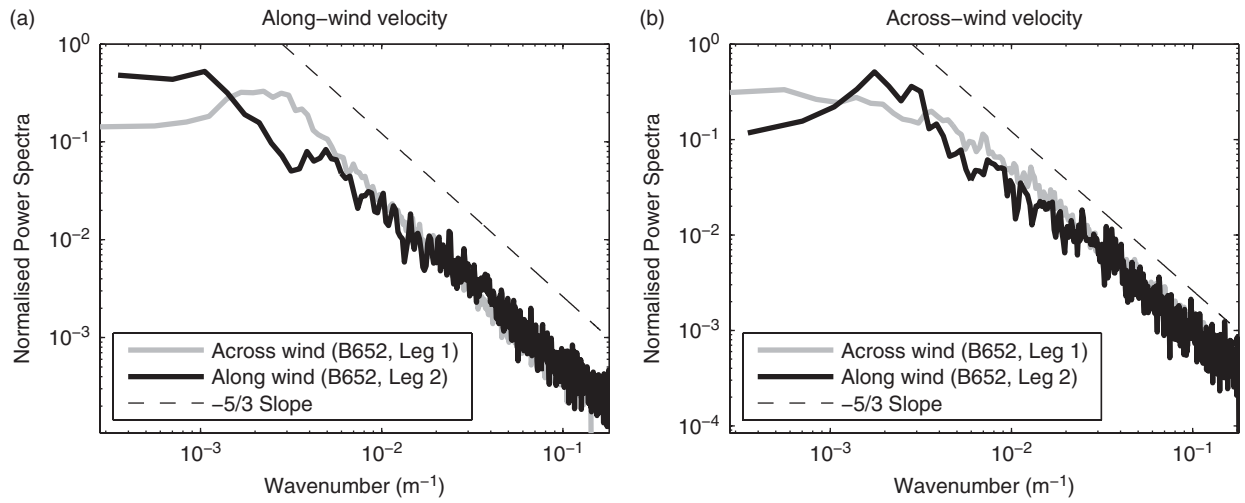

Figure 9. Normalised power spectra of horizontal wind components for flight B652. The left-hand panel (a) shows along-wind velocities, the right-hand panel (b) shows across-wind velocities. Each panel shows an across-wind leg (grey line) and an adjacent along-wind leg (black line). The spectra are averages over several runs for the legs indicated, i.e. B652 leg 1 has 11 12 km runs and leg 2 has 3 runs. A dashed line marks the $k^{-5/3}$ slope that is expected in the inertial subrange.

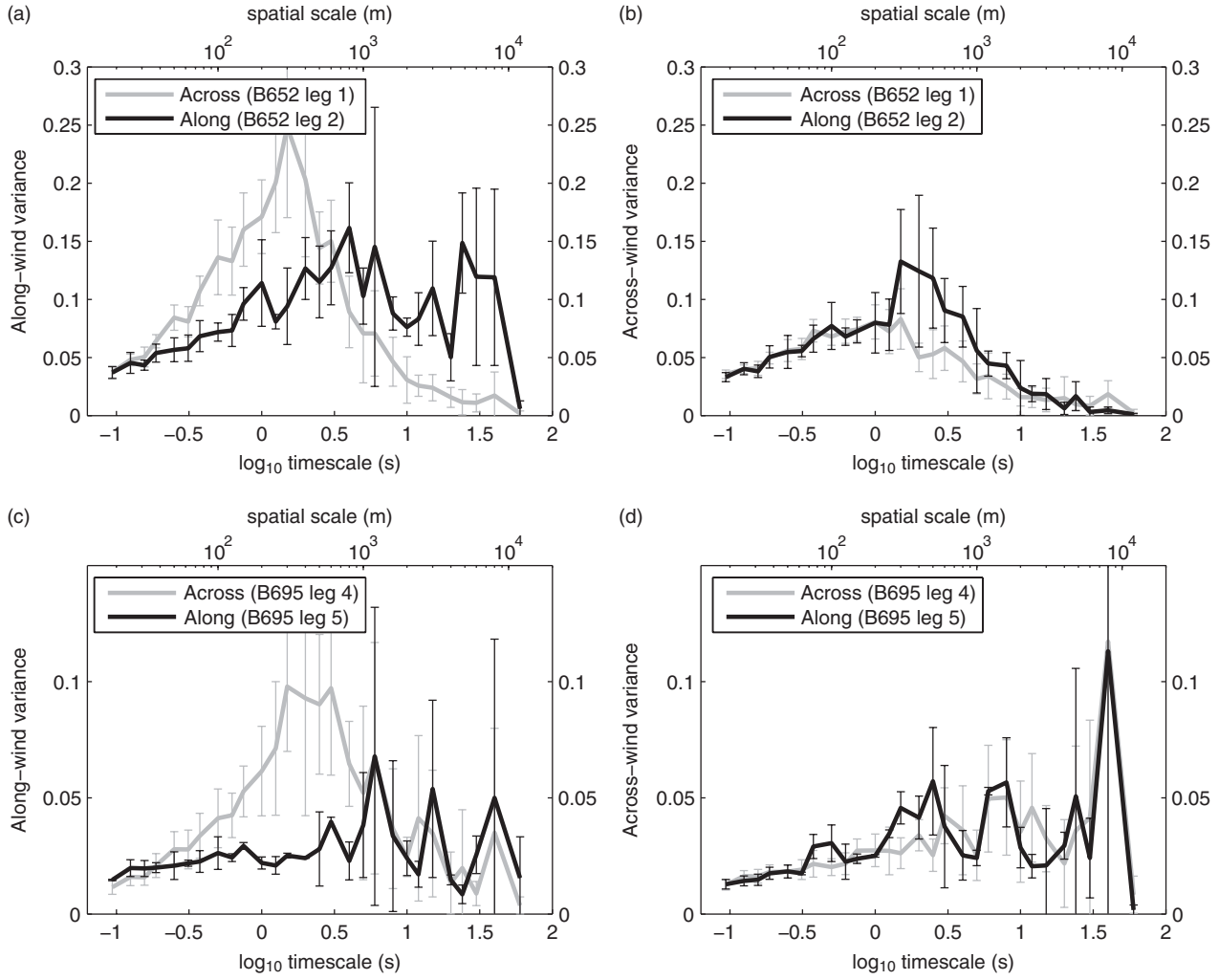

Figure 10. Variance of horizontal wind components as a function of scale for flights (a), (b) B652 and (c), (d) B695. The variance scales are determined via a multi-resolution decomposition technique (see text for details) and are shown on both a $\log_{10}$ time-scale (bottom axis) and on a spatial scale (top axis). The left-hand panels (a), (c) show along-wind velocity variances, the right-hand panels (b), (d) show across-wind velocity variances. Each panel shows mean variances (for several runs) for an across-wind leg (thin line) and for an adjacent along-wind leg (thick line), with the error bars indicating the standard deviation of variances at this scale. Note the legs illustrated here are the same as those in Figure 9. Recall B652 was flown in stable conditions, while B695 was in unstable conditions.

drops off rapidly for scales longer than 1000 m in both flights. In contrast, in the along-wind legs the peak power is between 1000 and 10 000 m, suggesting organisation of the ABL turbulence in the along-wind direction at these scales. In both cases the power does drop off for the longest time-scale bin. To test this further, multi-resolution variances have been calculated for 4 min runs in the along-wind direction (not shown). In both these cases the longer scales (up to 24 km) both show a further decrease in variance which suggests ever-decreasing amounts of turbulent flux would be added for longer runs. However, the limited length of the legs means this cannot be known with confidence, and furthermore, very long runs may break the requirement for

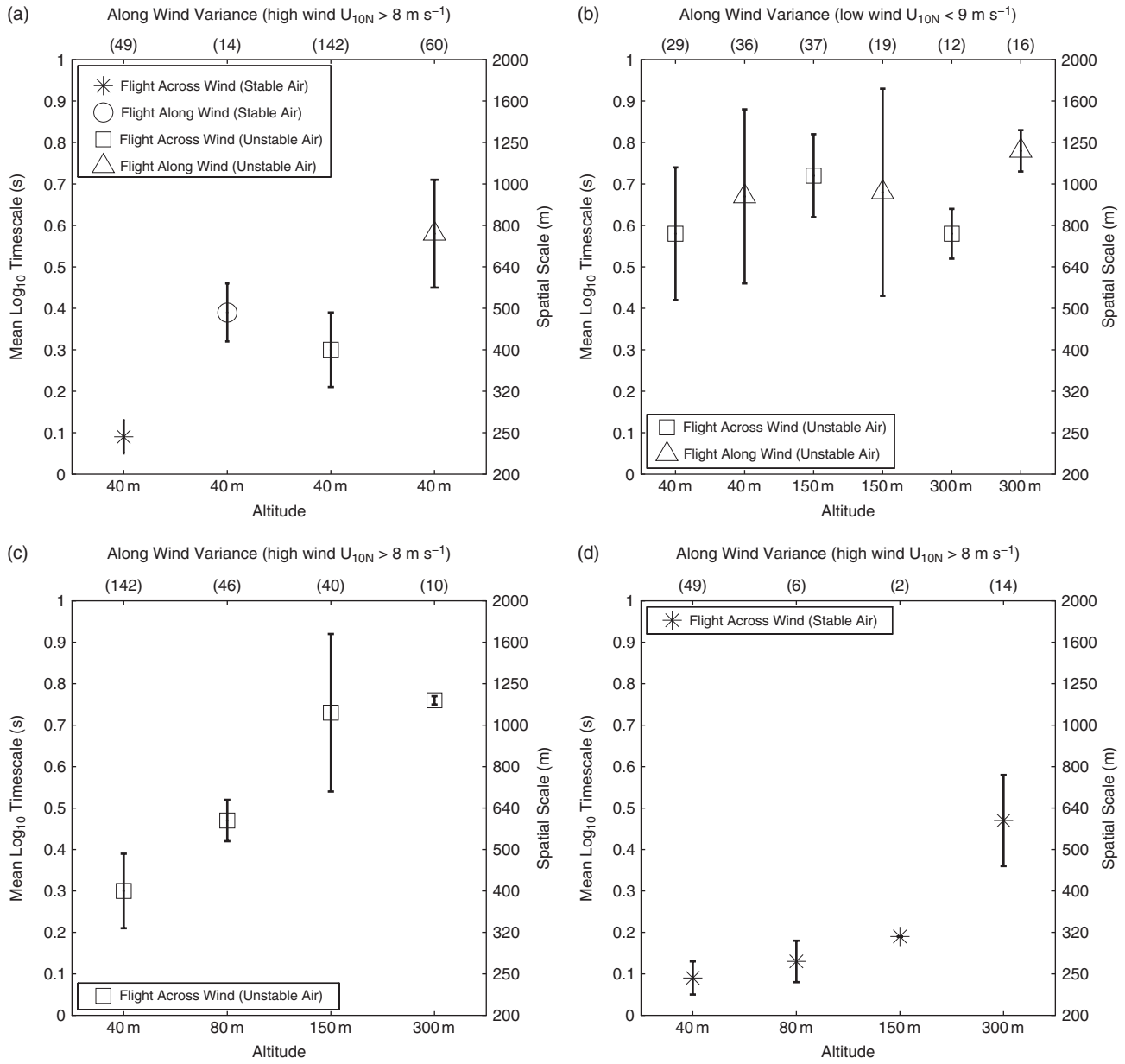

Figure 11. Summary of along-wind variance scales for selected meteorological conditions. The mean and standard deviation are shown. Panel (a) is for moderate to strong winds ($U_{10N} > 8$ m s^{-1}) and shows across- and along-wind runs in stable conditions, plus across- and along-wind runs in unstable conditions, with all runs at 40 m. Panel (b) is for weak winds ($U_{10N} < 9$ m s^{-1}) and shows across- and along-wind runs in unstable conditions versus height. Panel (c) is for $U_{10N} > 8$ m s^{-1} and across-wind runs in unstable conditions, while panel (d) is for $U_{10N} > 8$ m s^{-1} and across-wind runs in stable conditions, both showing eddy scale vs. height. The number of data points in each category is noted at the top of the plot.

homogenous ABL conditions. In B695 there is a notable peak in across-wind variance at around the 8000 m scale.

Mahrt (1998) and Mann and Lenschow (1994) discuss the scales of flux transportation in some detail, stating that a 'significant fraction of the flux' may be transported on surprisingly long scales by relatively weak mesoscale motions that are strongly correlated with the variable being transported. They highlight the need for long flux-sampling runs (10's to 100's of km) but go on to point out this is (in practice) impossible because of inhomogeneities in the ABL and aircraft limitations. In our dataset the majority of our runs are across-wind precisely because of these sampling strategies. However, as we have along-wind runs it is worthwhile trying to make use of them both for flux estimates and also in characterising the ABL features being encountered, while accepting that they provide only a partial picture of turbulence transport.

Differences in the across- and along-wind runs are summarised in Figure 11. This shows the mean time- and length-scales of along-wind velocity variances, calculated using the multi-resolution decomposition technique, for several meteorological categories. Here velocity variances are used as a metric for the scale of flux-carrying turbulent eddies (the co-spectra $u'w'$ and $v'w'$ were also examined and these showed broadly the same patterns). Figure 11(a) shows all 40 m runs for unstable and stable conditions with moderate to strong wind speeds. For both stability conditions the variances are significantly longer in the along-wind runs than they are in the across-wind runs, reinforcing the fact that larger scales are being measured in the along-wind runs. The aspect ratio is approximately 2:1, i.e. eddies are on average twice the size when flying along-wind. So at 40 m height, eddies are (on average) 250 by 500 m for stable conditions and 400 by 800 m for unstable conditions. Clearly the eddy scales are larger for unstable conditions.

Figure 11(b) shows along-wind velocity variances for low wind speed and unstable conditions; there are not enough data from low wind speed and stable conditions to merit examination. (Note Figure 11(b) shows data from flights without a '*' in Table 1, along with some higher altitude legs noted in Table 2). Here the variances in the along-wind runs are similar to those in the across-wind runs and furthermore there is no significant

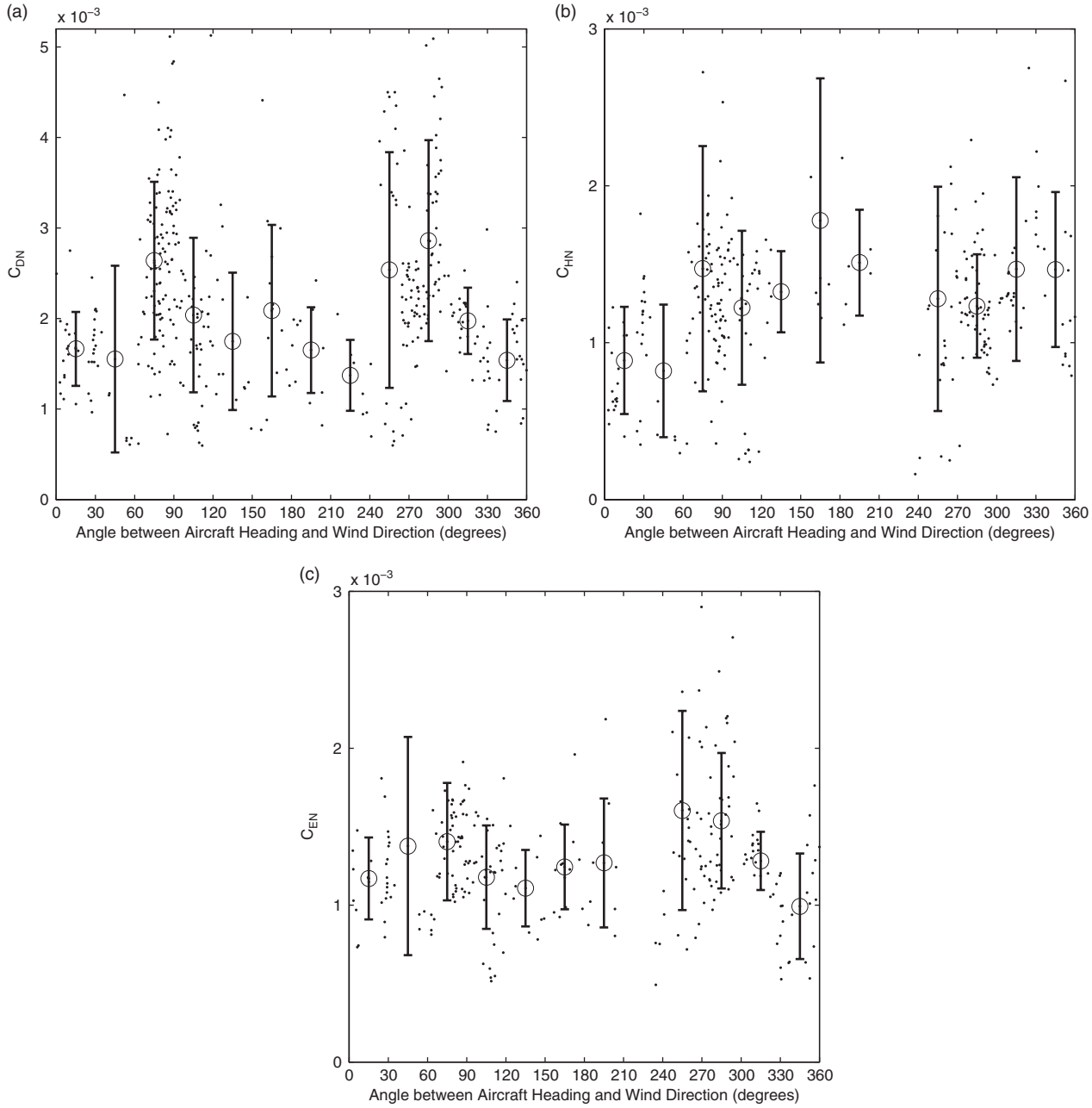

Figure 12. The 10 m neutral exchange coefficients for (a) momentum flux (C_{DN}); (b) sensible heat flux (C_{HN}); and (c) latent heat flux (C_{EN}) as a function of the difference angle between the aircraft heading and the mean wind direction. All observations are plotted as dots. Error bars showing the mean and standard deviation in 30° bins are overlaid.

change in scale with height. In other words, for low wind speeds the turbulent eddies are approximately isotropic and so there is no difference in the velocity variance due to sampling direction.

Figure 11(c) and (d) show along-wind velocity variances as a function of height for unstable and stable conditions (respectively) and moderate-to-strong winds. The flights used for the ABL data are indicated in Table 2; note these data are not used to estimate turbulent fluxes at the *surface*, as the sampling error associated with turbulent flux estimations scales with $z^{1/2}$ (e.g. Drennan *et al.*, 2007). However, profiles of turbulent quantities are of interest in their own right, for example, (i) to confirm assumptions used to extrapolate fluxes down to the surface; (ii) to examine the morphology of ABL eddies; (iii) to quantify ABL sources or sinks of heating; and (iv) to validate models and the PV-budget approach being used in DIAMET case-studies. In regards to (i) we have examined profiles of stress with and without the Donelan (1990) correction and find that in most cases with suitable observations the correction makes the stress more constant with height (as intended). Points (iii) and (iv) will be expanded upon in subsequent papers.

In unstable conditions there is a clear increase in mean eddy scale (with height) from around 400 to 1100 m on average. The increase is well-defined and monotonic between heights of 40 and 150 m, while the increase from 150 to 300 m is small and makes use of fewer data points. In stable conditions there is also a clear increase in mean eddy scale (with height) from around 250 to 600 m, although this result is based on relatively few data points, especially at heights of 80 and 150 m. Comparing the two cases, mean eddy sizes for statically unstable conditions are approximately double those of statically stable conditions at all heights.

The spectral and velocity variance scale analysis suggests an elongation of turbulent eddies in the along-wind direction when wind speeds are moderate to strong. This may be, for example, in the form of ABL roll vortices leading to cloud streets (e.g. Chou and Yeh, 1987; Renfrew and Moore, 1999). A check of relevant satellite imagery confirms this explanation in some cases (e.g. B268, Renfrew *et al.*, 2009b), but in other cases mid-level cloud shields the ABL and it is not possible to see cloud streets. So while the turbulence remains homogeneous, our results suggest it is not exactly isotropic, but is directionally organised by the wind.

In the previous section it was established that the morphology of ABL eddies (i.e. elongation in the along-wind direction) was affecting the turbulent flux estimates; that for along-wind runs, the fluxes were underestimated due to the shift of the spectra to longer length-scales. Figure 7 shows the observations as across- or along-wind and it appears there is a difference in exchange coefficient too, although it is not that clear due to the inherent scatter in the data. To examine this further, Figure 12 shows the exchange coefficients as a function of the difference angle between the aircraft and the mean wind for each run. Also shown are bin averages (and standard deviations) in 30° bins. Broadly speaking all three of the exchange coefficients are higher when the difference angle is nearer 90° and 270° (across the wind). This is clearest for C_{DN} and C_{EN} and is less apparent for C_{HN}. However, this result is influenced by the distribution of the difference angle – most of the runs are across-wind and there are some bins with relatively few data in (e.g. near 180°). In short, this corroborates our earlier discussion that the across-wind runs are capturing the full turbulent fluxes, while the along-wind runs can underestimate this turbulent exchange.

7. Conclusions

Eddy covariance observations of turbulent air–sea fluxes from low-level aircraft legs have been presented and analysed. A comparison of a number of bulk flux algorithms demonstrates these are generally consistent with the observations. It is not possible to distinguish which of the COARE 3.0 algorithm or those used by the ECMWF or Met Office numerical weather and climate predictions models are a better fit to the observations. However, there is some evidence that the algorithm of Smith (1988), which uses a lower Charnock constant, corresponds less well to the observations.

The new algorithm of Andreas *et al.* (2012) for momentum corresponds well to the observations and, when plotted as u^* versus U_{10N} the scatter in the observations appears reduced.

There is a difference in velocity variances and turbulent fluxes between legs flown across-wind and those flown along-wind. The along-wind legs do *not* capture all of the variance or flux; the ends of the spectra are shifted to surprisingly long scales (>12 km). Even lengthening the runs to 4 min (24 km) still does not capture all of the variance. A multi-resolution spectral technique is used to show that the turbulent eddies tend to be elongated in the downwind direction – with an aspect ratio of approximately 2:1 – for both unstable and stable conditions and moderate to strong wind speeds. Mean eddies are typically twice the size for unstable conditions, compared to stable conditions, and increase in scale with height more rapidly too.

One consequence of this heading-dependent result should be a re-evaluation of turbulent flux observations, especially for moderate-to-strong winds. Using aircraft to make observations, the run lengths are limited by logistics and the requirement of homogeneous conditions. The majority of our runs are across-wind, as has been common practice. However, some along-wind runs are inevitable if other objectives are part of the research flight, and care is needed in the interpretation of these runs.

The purpose of calibrating bulk flux algorithms has been for their use in numerical weather and climate prediction models where the grid boxes are assumed to be larger than the turbulence-carrying scales. In short-range operational forecasting, this assumption is now starting to break down. The Met Office operational forecast over the British Isles has a grid resolution of 1.5 km at present and so is able to resolve features down to several km in scale. Consequently these models may be resolving that part of the air–sea flux carried by motion on these scales. Here we have shown that in the along-wind direction there is a flux on scales of 5 km plus, in other words, the traditional mesoscale gap (between around 1 and 10 km) is not always there. This also raises the possibility of incorrectly augmenting turbulence flux transfer in models, as these are accounted for by the subgrid-scale parametrization and then enhanced by any resolved flux transfer. The possibility of such 'double counting' needs further investigation.

Acknowledgements

The authors would like to acknowledge all at the Facility for Airborne Atmospheric Measurements, DirectFlight and Avalon for their critical role in obtaining these observations. We would also like to thank the following for their helpful comments and mission scientist contributions to this work: Phil Brown, Stuart Newman and John Edwards at the Met Office, Axel Wellpott and Alan Woolley at FAAM, Doug Parker, John Methven, Geraint Vaughan, Sue Gray, Keith Bower and Phil Rosenberg from the DIAMET project. Further thanks for the comments from two reviewers which have improved this manuscript. This study has been funded by the DIAMET project (NE/I005293/1), part of NERC's Storm Risk Mitigation programme.

References

Adamson DS, Belcher SE, Hoskins BJ, Plant RS. 2006. Boundary-layer friction in midlatitude cyclones. *Q. J. R. Meteorol. Soc.* **132**: 101–124.

Andreas EL. 2011. Fallacies of the enthalpy transfer coefficient over the ocean in high winds. *J. Atmos. Sci.* **68**: 1435–1445.

Andreas EL, DeCosmo J. 2002. The signature of sea spray in the HEXOS turbulent heat flux data. *Boundary-Layer Meteorol.* **103**: 303–333.

Andreas EL, Mahrt L, Vickers D. 2012. A new drag relation for aerodynamically rough flow over the ocean. *J. Atmos. Sci.* **69**: 2520–2537.

Banner ML, Chen W, Walsh EJ, Jensen JB, Lee S, Fandry C. 1999. The Southern Ocean waves experiment. Part I: Overview and mean results. *J. Phys. Oceanogr.* **29**: 2130–2145.

Beare RJ. 2007. Boundary layer mechanisms in extratropical cyclones. *Q. J. R. Meteorol. Soc.* **133**: 503–515.

Boutle IA, Beare RJ, Belcher SE, Plant RS. 2007. A note on boundary-layer friction in baroclinic cyclones. *Q. J. R. Meteorol. Soc.* **133**: 2137–2141.

Boutle IA, Beare RJ, Belcher SE, Brown AR, Plant RS. 2010. The moist boundary layer under a mid-latitude weather system. *Boundary-Layer Meteorol.* **134**: 367–386.

Chagnon JM, Gray SL. 2009. Horizontal potential vorticity dipoles on the convective storm scale. *Q. J. R. Meteorol. Soc.* **135**: 1392–1408.

Chagnon JM, Gray SL, Methven J. 2013. Diabatic processes modifying potential vorticity in a North Atlantic cyclone. *Q. J. R. Meteorol. Soc.* **139**: 1270–1282.

Chou SH, Yeh EN. 1987. Airborne measurements of surface layer turbulence over the ocean during cold air outbreaks. *J. Atmos. Sci.* **44**: 3721–3733.

Donelan MA. 1990. Air–sea interaction. In *The Sea*, LeMehaute B, Hanes DM. (eds.) *Ocean Engineering Science*. 9: 239–292. Wiley-Interscience: New York, NY.

Donelan MA, Haus BK, Reul N, Plant WJ, Stiassne M, Graber HC, Brown OB, Saltzman ES. 2004. On the limiting aerodynamic roughness of the ocean in very strong winds. *Geophys. Res. Lett.* **31**: L18306, doi: 10.1029/2004GL019460.

Drennan WM, Graber HC, Hauser D, Quentin C. 2003. On the wave age dependence of wind stress over pure wind seas. *J. Geophys. Res.* **108**: 8062, doi: 10.1029/2000JC000715.

Drennan WM, Zhang JA, French JR, McCormick C, Black PG. 2007. Turbulent fluxes in the hurricane boundary layer. Part II: Latent heat flux. *J. Atmos. Sci.* **64**: 1103–1115, doi: 10.1175/JAS3889.1.

Edwards JM. 2007. Oceanic latent heat fluxes: Consistency with the atmospheric hydrological and energy cycles and general circulation modeling. *J. Geophys. Res.* **112**: D06115, doi: 10.1029/2006JD007324.

Fairall CW, Bradley EF, Rogers DP, Edson JB, Young GS. 1996. Bulk parameterization of air–sea fluxes for tropical ocean-global atmosphere coupled-ocean atmosphere response experiment. *J. Geophys. Res.* **101**: 3747–3764.

Fairall CW, Bradley EF, Hare JE, Grachev AA, Edson JB. 2003. Bulk parameterization of air–sea fluxes: Updates and verification for the COARE algorithm. *J. Clim.* **16**: 571–591.

French JR, Drennan WM, Zhang JA, Black PG. 2007. Turbulent fluxes in the hurricane boundary layer. Part I: Momentum flux. *J. Atmos. Sci.* **64**: 1089–1102.

Gray SL. 2006. Mechanisms of midlatitude cross-tropopause transport using a potential vorticity budget approach. *J. Geophys. Res.* **111**: D17113, doi: 10.1029/2005JD006259.

Grossman RL, Betts AK. 1990. Air–sea interaction during an extreme cold air outbreak from the eastern coast of the United States. *Mon. Weather Rev.* **118**: 324–342.

Howell JF, Mahrt L. 1997. Multiresolution flux decomposition. *Boundary-Layer Meteorol.* **83**: 117–137.

Inverarity GW. 2000. Correcting airborne temperature data for lags introduced by instruments with two-time-constant responses. *J. Atmos. Oceanic Technol.* **17**: 176–184.

Kalogiros J, Wang Q. 2011. Aircraft observations of sea-surface turbulent fluxes near the California coast. *Boundary-Layer Meteorol.* **139**: 283–306, doi: 10.1007/s10546-010-9585-x.

McCarthy J. 1973. A method for correcting airborne temperature data for sensor response time. *J. Appl. Meteorol.* **12**: 211–214.

Mahrt L. 1998. Flux sampling strategy for aircraft and tower observations. *J. Atmos. Oceanic Technol.* **15**: 416–429.

Mann J, Lenschow DH. 1994. Errors in airborne flux measurements. *J. Geophys. Res.* **99**: 14519–14526, doi: 10.1029/94JD00737.

Moon IJ, Ginis I, Hara T, Thomas B. 2007. A physics-based parameterization of air–sea momentum flux at high wind speeds and its impact on hurricane intensity predictions. *Mon. Weather Rev.* **135**: 2869–2879, doi: 10.1175/MWR3432.1.

Mueller J, Veron F. 2009. Nonlinear formulation of the bulk surface stress over breaking waves: Feedback mechanisms from air-flow separation. *Boundary-Layer Meteorol.* **130**: 117–134.

Newman SM, Smith JA, Glew MD, Rogers SM, Taylor JP. 2005. Temperature and salinity dependence of sea surface emissivity in the thermal infrared. *Q. J. R. Meteorol. Soc.* **131**: 2539–2557.

Nicholls S. 1978. Measurements of turbulence by an instrumented aircraft in a convective atmospheric boundary layer over the sea. *Q. J. R. Meteorol. Soc.* **104**: 653–676.

Nicholls S, Readings CJ. 1981. Spectral characteristics of surface layer turbulence over the sea. *Q. J. R. Meteorol. Soc.* **107**: 591–614.

Perrie W, Andreas EL, Zhang W, Li W, Gyakum J, McTaggart-Cowan R. 2005. Sea spray impacts on intensifying midlatitude cyclones. *J. Atmos. Sci.* **62**: 1867–1883.

Persson POG, Hare JE, Fairall CW, Otto WD. 2005. Air–sea interaction processes in warm and cold sectors of extratropical cyclonic storms observed during FASTEX. *Q. J. R. Meteorol. Soc.* **131**: 877–912.

Petersen GN, Renfrew IA. 2009. Aircraft-based observations of air–sea fluxes over Denmark Strait and the Irminger Sea during high wind speed conditions. *Q. J. R. Meteorol. Soc.* **135**: 2030–2045.

Plant RS, Belcher SE. 2007. Numerical simulation of baroclinic waves with a parameterized boundary layer. *J. Atmos. Sci.* **64**: 4383–4399, doi: 10.1175/2007JAS2269.1.

Powell MD, Vickery PJ, Reinhold TA. 2003. Reduced drag coefficient for high wind speeds in tropical cyclones. *Nature* **422**: 279–283.

Raga GB, Abarca S. 2007. On the parameterization of turbulent fluxes over the tropical eastern Pacific. *Atmos. Chem. Phys.* **7**: 635–643.

Renfrew IA, Moore GWK. 1999. An extreme cold air outbreak over the Labrador Sea: Roll vortices and air–sea interaction. *Mon. Weather Rev.* **127**: 2379–2394.

Renfrew IA, Moore GWK, Kristjánsson JE, Ólafsson H, Gray SL, Petersen GN, Bovis K, Brown PRA, Føre I, Haine T, Hay C, Irvine EA, Lawrence A, Ohigashi T, Outten S, Pickart RS, Shapiro M, Sproson DAJ, Swinbank R, Woolley A, Zhang S. 2008. The Greenland flow distortion experiment. *Bull. Am. Meteorol. Soc.* **89**: 1307–1324.

Renfrew IA, Petersen GN, Sproson D, Moore GWK, Adiwidjaja H, Zhang S, North R. 2009a. A comparison of aircraft-based surface-layer observations during high wind speed conditions over Denmark Strait and the Irminger Sea with meteorological analyses and QuikSCAT winds. *Q. J. R. Meteorol. Soc.* **135**: 2046–2066.

Renfrew IA, Outten SD, Moore GWK. 2009b. An easterly tip jet off Cape Farewell, Greenland. Part I: Aircraft observations. *Q. J. R. Meteorol. Soc.* **135**: 1919–1933.

Smith SD. 1988. Coefficients for sea surface wind stress, heat flux and wind profiles as a function of wind speed and temperature. *J. Geophys. Res.* **93**: 15467–15472.

Stark JD, Donlon CJ, Martin MJ, McCulloch ME. 2007. OSTIA: An operational, high resolution, real time, global sea surface temperature analysis system. *OCEANS 2007 – Europe*: 331–334. Institute of Electrical and Electronics Engineers: Piscataway, NJ, doi: 10.1109/OCEANSE.2007.4302251.

Vickers D, Mahrt L. 2003. The cospectral gap and turbulent flux calculations. *J. Atmos. Oceanic Technol.* **20**: 660–672.

Vickers D, Mahrt L, Andreas EL. 2013. Estimates of the 10-m neutral sea surface drag coefficient from aircraft eddy-covariance measurements. *J. Phys. Oceanogr.* **43**: 301–310, doi: 10.1175/JPO-D-12-0101.1.

Wilson SHS, Atkinson NC, Smith JA. 1999. The development of an airborne infrared interferometer for meteorological sounding studies. *J. Atmos. Oceanic Technol.* **16**: 1912–1927.

Yelland M, Taylor PK. 1996. Wind stress measurements from the open ocean. *J. Phys. Oceanogr.* **26**: 541–558.

Yelland MJ, Moat BI, Taylor PK, Pascal RW, Hutchings J, Cornell VC. 1998. Wind stress measurements from the open ocean corrected for airflow distortion by the ship. *J. Phys. Oceanogr.* **28**: 1511–1526.

Zhai X, Johnson HL, Marshall DP, Wunsch C. 2012. On the wind power input to the ocean general circulation. *J. Phys. Oceanogr.* **42**: 1357–1365, doi: 10.1175/JPO-D-12-09.1.

Zhang JA, Black PG, French JR, Drennan WM. 2008. First direct measurements of enthalpy flux in the hurricane boundary layer: The CBLAST results. *Geophys. Res. Lett.* **35**: L14813, doi: 10.1029/2008GL034374.