Proofreading

The most efficient (and highly encouraged) way to correct the page proofs and return them to AMS is to electronically annotate the PDF file containing your proof **using the annotation tools** described below in Adobe Acrobat and **send it via e-mail to pdfproofs@ametsoc.org**. Instructions for the PDF annotation process appear on the following pages. If necessary, you may instead print out your page proofs, make your corrections on the pages, and fax or express mail them to AMS. The fax number and delivery address are located at the end of this document.

Proofread the proofs carefully, as this will be your only chance to see your article before publication. **Pay special attention to color figures (if any), and characters such as Greek letters and mathematical symbols**. Make sure that everything is typeset the way that you want it as the corrected proofs will not be returned to you.

Return of your e-proof to AMS signifies that you have thoroughly read and approve of the proof, and thereby authorize publication based on the proof (except where you have identified errors or indicated desired changes within the proof file, or clearly itemized these changes in the email to pdfproofs@ametsoc.org). Adjustments for AMS style may be made prior to publication.

Alterations

Answer author queries at the appropriate place in the text and correct any errors you find. **Authors are allowed up to 10 new textual edits at proof stage at no additional charge** (copy and technical editing errors are not charged to authors at any time)**. Additional edits will be charged at the rate of $4.50 per edit, and changes to figures are more expensive (~$20 per black and white figure and ~$30-$50 per color figure).**

The layout in the page proof is considered final unless changes are essential. Every effort should be made to balance text and figure alterations so that deletions are compensated by nearby additions, and figure substitutions should be of the same size and shape, if possible. Please also note that material changes to the *science* in the manuscript at this stage will require further peer review, substantially increasing the time to publication. **The editors reserve the right to accept or reject proposed alterations.**

Figures

The figures that appear in your page proofs are lower resolution than the final printed article. The primary AMS technical editor for your journal will receive photoprints and the original figures for the purposes of checking the figure quality. If you have specific concerns about any of your figures, especially if any color figures do not appear in color, please describe them in your return. Please verify that each figure properly corresponds with the figure captions and citations in the text. If there are corrections that need to be made, send a new electronic version in eps or tiff format of the corrected figure with your corrected proofs. Include your name, the figure number, and the manuscript number.

Reprints

For manuscripts accepted starting in January 2009, a PDF file of the final published paper will be sent to the corresponding author at no extra charge in lieu of hard-copy reprints. If you would still like to order hard-copy reprints, please use the reprint order form (available at http://www.ametsoc.org/PUBS/journals/reprintorder.pdf).

Page and Color Charges

Please note that when your paper was accepted you were sent a link to the Journals Page and Color Charge form (http://www.ametsoc.org/PUBS/pagecolorcharge.pdf). This form should have been returned to AMS at that time. If you have not completed and returned the form, please contact Christine Keane at AMS Headquarters *immediately* (617-227-2426, ext. 254 or ckeane@ametsoc.org) to ensure publication of your paper.

Thank you very much for your cooperation, and thank you for publishing with the AMS.

Author instructions.doc

Annotating PDFs

ams_annotating_pdfs1.1; April 09, 2009; full release

1. Introduction

eProof files are self-contained PDF documents for viewing on-screen and for printing. They contain all appropriate formatting and fonts to ensure correct rendering on-screen and when printing hardcopy. DJS sends eProofs that can be viewed, annotated, and printed using the free version of Acrobat Reader 7 (or greater). These eProofs are "enabled" with commenting rights, therefore they can be modified by using special markup tools in Acrobat Reader that are not normally available unless using the Standard or Professional version.

The screen images in this document were captured on a PC running Adobe Acrobat Reader version 8.1.0. Though some of the images may differ in appearance from your platform/version, basic functionality remains similar. At the time of this writing, Acrobat Reader v9.1 is freely available and can be downloaded from: http://get.adobe.com/reader/

2. Comment & Markup toolbar functionality
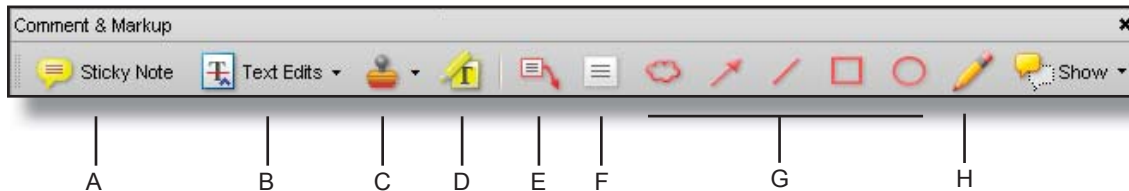
A. Sticky Note tool; B. Text Edits tool; C. Stamp tool; D. Highlight Text tool; E. Callout tool; F. Text Box tool; G. Various Object tools; H. Pencil tool

A. Show the Comment & Markup toolbar

The Comment & Markup toolbar doesn't appear by default. Do one of the following:

- Select View > Toolbars > Comment & Markup.
- Select Tools > Comment & Markup > Show Comment & Markup Toolbar.
- Click the Review & Comment button in the Task toolbar, and choose Show Comment & Markup Toolbar.

To add or remove tools for this toolbar, right-click the toolbar and select the tool. Or, select Tools > Customize Toolbars.

B. Select a commenting or markup tool

Do one of the following:

- Select a tool from the Comment & Markup toolbar.
- Select Tools > Comment & Markup > [tool].

Note: After an initial comment is made, the tool changes back to the Select tool so that the comment can be moved, resized, or edited. (The Pencil, Highlight Text, and Line tools stay selected.)

C. Keep a commenting tool selected

Multiple comments can be added without reselecting the tool.
Select the tool to use (but don't use it yet).

- Select View > Toolbars > Properties Bar.
- Select Keep Tool Selected.
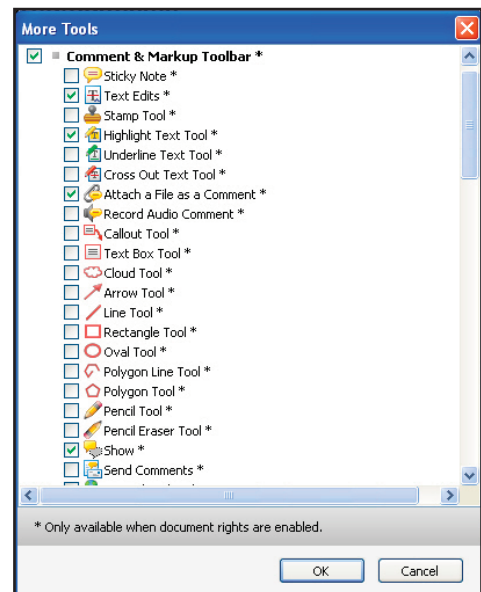
*Choose Tools > Customize Toolbars to remove unnecessary items from the toolbar
(see Section 7 for suggested toolbar layout)*

3. The Properties bar

The Properties bar can be used to format text and select options for individual tools.

To view the Properties bar, do one of the following:

- Choose View > Toolbars > Properties Bar.
- Right-click the toolbar area; choose Properties Bar.
- Select [Ctrl-E]

4. Using the comment and markup tools

To *insert*, *delete*, or *replace* text, use the **Text Edits** tool. Select the Text Edits tool, then select the text with the cursor (or simply position it) and begin typing. A pop-up note will appear based upon the modification (e.g., inserted text, replacement text, etc.). Use the Properties bar to format text in pop-up notes. A pop-up note can be minimized by selecting the ⊟ button inside it.
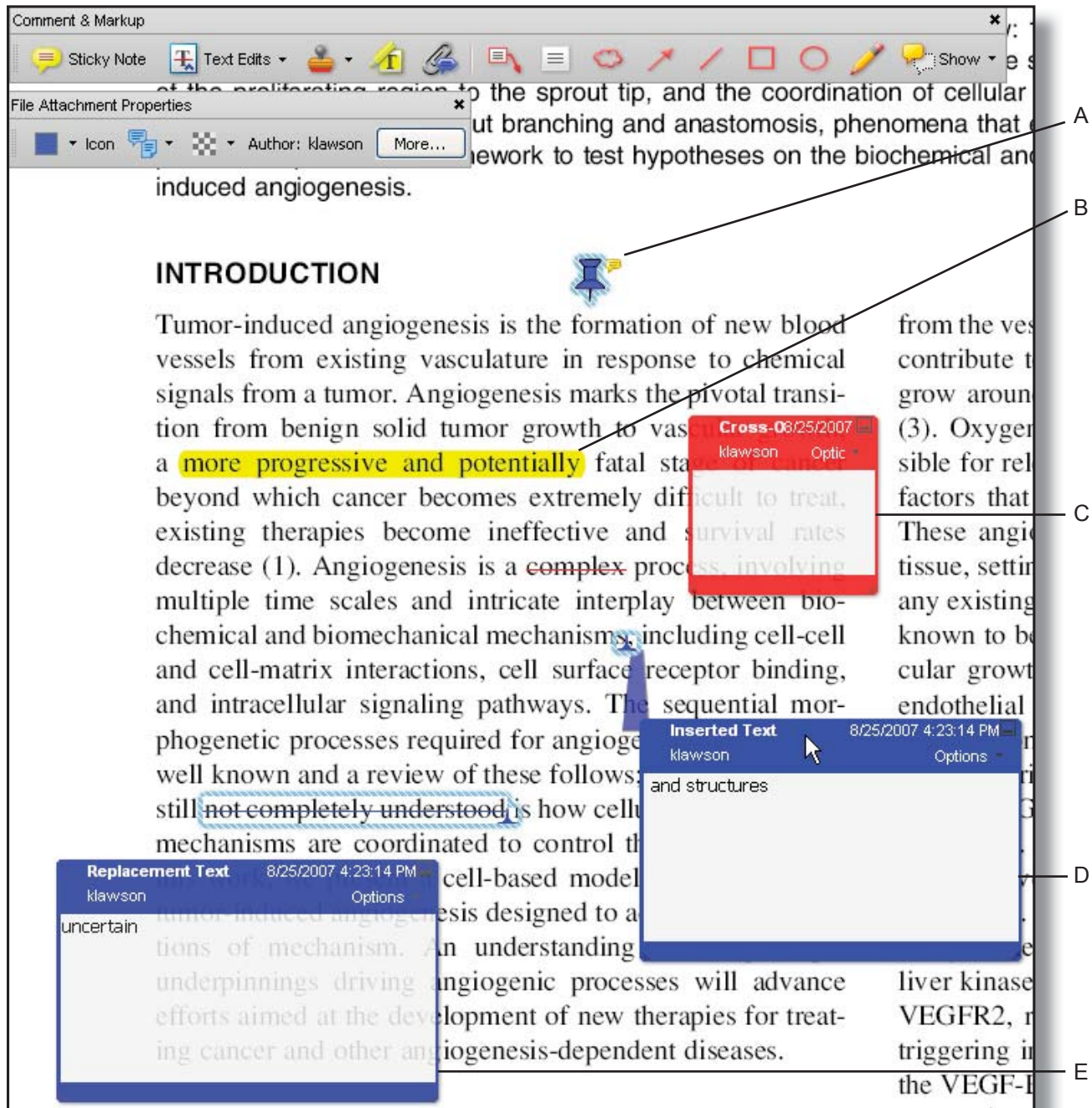
A. Attached file; B. Highlighted text; C. Crossed-out (strike-through) text; D. Inserted text; E. Replaced text

5. Inserting symbols or special characters

An 'insert symbol' feature is not available for annotations, and copying/pasting symbols or non-keyboard characters from Microsoft Word does not always work. Use angle brackets < > to indicate these special characters (e.g., <alpha>, <beta>).

6. Editing near watermarks and hyperlinked text

eProof documents often contain watermarks and/or hyperlinked text. Selecting characters near these items can be difficult using the mouse alone. To edit an eProof which contains text in these areas, do the following:

- Without selecting the watermark or hyperlink, place the cursor near the area for editing.
- Use the arrow keys to move the cursor beside the text to be edited.
- Hold down the shift key while simultaneously using arrow keys to select the block of text, if necessary.
- Insert, replace, or delete text, as needed.

7. Summary of main functions

Insert text - Use Text Edits tool (position cursor and begin typing)
Replace text - Use Text Edits tool (select text and begin typing)
Delete text - Use Text Edits tool (select text and press delete key)
Highlight text - Use Highlight Text tool (select text)
Attach a file - Use the Attach a File with Comment tool (select tool, position cursor and click mouse, select file)
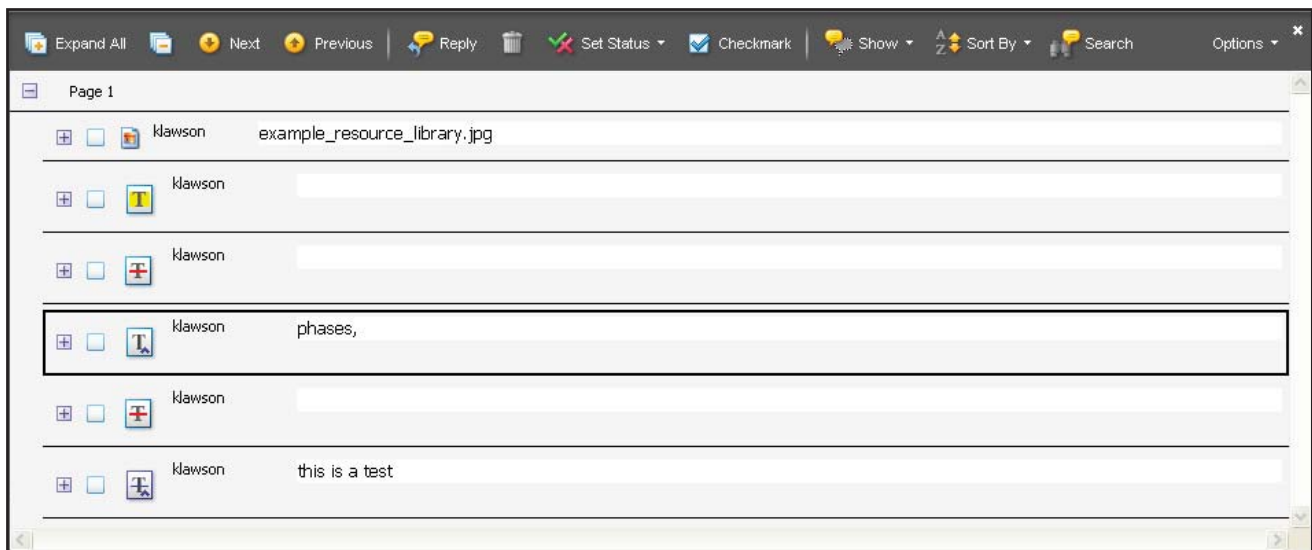

Suggested toolbar layout

8. Reviewing changes

To review all changes, do the following:

- Select the Show button on the Comment & Markup toolbar.
- Select Show Comments List.

    Note: *Selecting a correction in the list will highlight the corresponding item in the document, and vice versa.*

Use the Comments list to review all changes

AUTHOR QUERIES

DATE __6/12/2009__
JOB NAME __JTECH__
JOB NUMBER __0__

ARTICLE __jtechO667__
QUERIES FOR AUTHORS __COLBO AND WELLER__

PLEASE ANSWER THE AUTHOR QUERIES WHERE THEY APPEAR IN THE TEXT.

AU1: Please verify corresponding author address.

AU2: Please verify addition of sentence in order to cite Tables 2–9; otherwise, Tables 4–9 are not cited in the text and appropriate sequential citations of these tables should be added to the text.

AU3: Please provide first initials and year of personal communication for Payne.

AU4: Please provide expansion for DXD, if available.

AU5: Please verify edit of citation Folland (1998) to Folland (1988) to match supplied reference; otherwise, corresponding reference is not cited in the text and should be deleted.

AU6: Please provide first initials and year of personal communication for Yuter.

AU7: Please provide year of personal communication for C. Fairall.

AU8: Please verify edit of citation Jin et al. (2000) to Jin et al. (2002) to match supplied reference; otherwise, corresponding reference is not cited in the text and should be deleted.

AU9: Change ok?

AU10: Change correct?

AU11: What is T here?
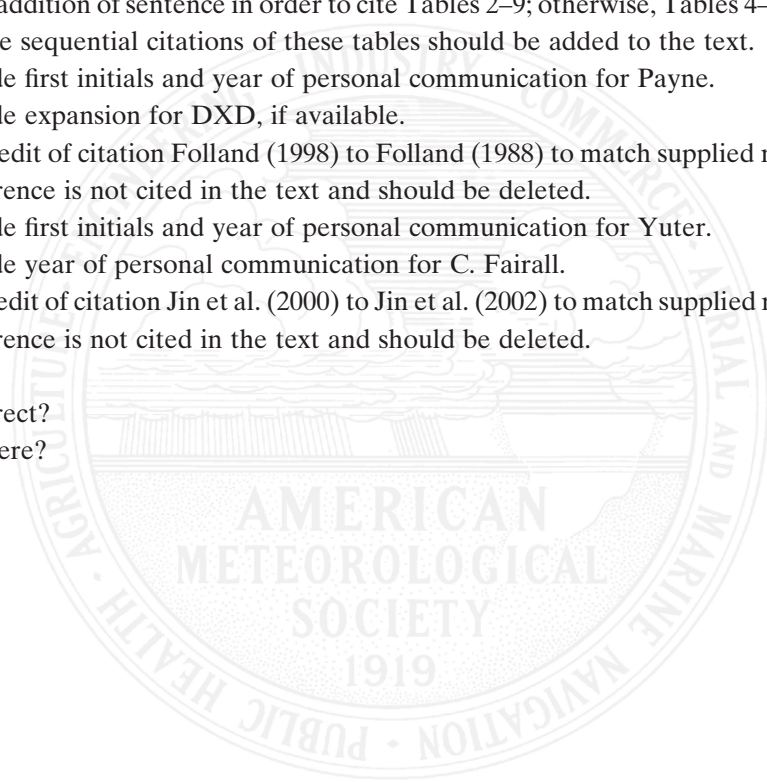
Accuracy of the IMET Sensor Package in the Subtropics

Keir Colbo

Department of Oceanography, Dalhousie University, Halifax, Nova Scotia, Canada

Robert A. Weller

Physical Oceanography Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

(Manuscript received 18 August 2008, in final form 1 February 2009)

ABSTRACT

The accuracies of the meteorological sensors (air temperature, relative humidity, barometric pressure, near-surface temperature, longwave and shortwave radiation, and wind speed and direction) that compose the Improved Meteorological (IMET) system used on buoys at long-term ocean time series sites known as ocean reference stations (ORS) are analyzed to determine their absolute error characteristics. The predicted errors are compared to in situ measurement discrepancies and other observations (direct flux shipboard sensors) to confirm the predictions. The meteorological errors are then propagated through bulk flux formulas and the Coupled Ocean–Atmosphere Response Experiment (COARE) algorithm to give predicted errors for the heat flux components, the freshwater flux, and the momentum flux. Absolute errors are presented for three frequency bands [instantaneous (1-min sampling), diurnal, and annual]. The absolute uncertainty in the annually averaged net heat flux is found to be 8 W m^{-2} for conditions similar to the current ORS deployments in the subtropics.

1. Introduction

The Improved Meteorological (IMET) sensor suite is a package for measuring surface meteorological variables at sea (Hosom et al. 1995) and observing the variables necessary to compute, from bulk formulas, the surface fluxes of heat, freshwater, and momentum. A standard deployment consists of two independent IMET packages, each with the following sensors: air temperature, sea surface temperature (SST), barometric pressure, relative humidity (RH), wind speed and direction, precipitation, and incoming shortwave and longwave radiation. The sensors of an IMET package are each placed as close as possible to their specific signal conditioning and analog-to-digital conversion circuitry, and the combinations of sensors and their respective signal conditioning electronics are packaged as discrete modules. These modules accept power and provide digital outputs (RS-485 or RS-232); they are connected to a common logger and satellite link. The signal conditioning electronics sample the sensors and compute 1-min averages. One module, for air temperature and humidity, has two sensors collocated in one module. All other modules have one sensor. The most common platform for deployment is a surface buoy, but IMET packages have also been placed on research and voluntary observing ships (VOS). On some deployments, additional modules not wired to a datalogger and internally powered are added for redundancy; these are referred to as stand-alone modules.

The IMET sensor suite is being deployed on buoys that are ocean reference stations (ORS) to collect data to be used to look at climate variability and to verify weather and climate models. Thus, it is important to understand the characteristics of the sensor errors and the accuracies of the observations. For a detailed introduction to issues associated with such validation and intercomparison, we refer the reader to the final report of the World Climate Research Programme/Scientific Committee for Oceanographic Research (WCRP/SCOR) Working Group on Air–Sea Fluxes (Taylor 2000). Here, we focus on the performance of the IMET sensor suite and the errors in its measurements and computed fluxes.

Corresponding author address: Dr. Robert A. Weller, Woods Hole Oceanographic Institution, Clark 204A MS 29, Woods Hole, MA 02540.
E-mail: rweller@whoi.edu

The error will be defined as the component of the total measured error that is not correctable after recovery. Postrecovery corrections can be, for example, based on additional calibrations. Our consideration of error and accuracy will be broken down into a sequence that progresses from the laboratory to the in situ measurements and also separates the uncertainty into an absolute and a variable component, where the variable component appears in the point-to-point measurements but cancels (or partially cancels) in daily or longer averages. The accuracies of the basic observed quantities (e.g., thermopile voltage and case and dome temperature in a longwave radiation sensor) and the uncertainties associated with deriving the desired measurement (e.g., incoming longwave radiation) by using equations and empirical calibration constants are discussed. Data from a series of yearlong deployments with pre- and post-deployment calibrations allow us to quantify drift, and data from comparisons with shipboard sensors during several days at the beginning and end of the deployments allow us to examine performance in the field. The focus is on absolute error.

The paper has the following structure: section 2 looks at all the sensors individually and creates a table that summarizes the error characteristics of each sensor. Section 3 examines sensors at sea on the same platform, as well as different platforms, and compares the observed errors to the predictions. Section 4 considers how these errors propagate through the heat flux calculations. Section 5 provides a discussion, including thoughts on future improvements, and summarizes the important conclusions.

2. Specific sensors

We concentrate here on the two most recent IMET buoy deployments. The first is the Stratus deployment, in which a 3-m discus buoy has been deployed at 20°S, 85°W since October 2000 (with annual recoveries of the old and deployments of the new buoy and instrumentation). The second is the Northwest Tropical Atlantic Station (NTAS), which has maintained a similar buoy at 15°N, 51°W since March 2001. The buoy superstructure supports an open platform for mounting sensors above the main deck (Fig. 1), with a vane attached to one of the uprights. This vane orients the buoy into the wind in winds greater than 2 m s^{-1}. The wind and the relative humidity–air temperature modules are mounted in the front of the sensor platform. In the central portion of the platform are the barometric pressure sensors and the siphon rain gauges. Aft, above the vane, are the radiation sensors on a platform raised so that the radiometers are not shaded by other instruments, except at very low sun angles when they themselves may create

shading. Typical instrument heights above the waterline are given in Table 1. Listings of the various sources of error for each sensor, along with estimates of the instantaneous, daily, and annual absolute errors, are shown in Tables 2–9.

a. Longwave radiation (Eppley PIR)

Longwave radiation is a challenging measurement; there have been intercomparison studies of different instruments (Philipona et al. 2001; Barton et al. 2004). Our module uses an Eppley precision infrared radiometer (PIR) with a modified aluminum case and shield adapted for the marine environment (Fairall et al. 1998). Incoming longwave radiation is computed from the measured thermopile voltage and the dome and case temperatures; Fairall et al. (1998) state that the accuracies of these measurements are 10 μV, 0.22 K, and 0.1 K, respectively. The dome temperature has a higher uncertainty than the case for two reasons: it experiences larger gradients (Philipona et al. 1995) and the thermal contact between the silicon dome and the thermistor is hindered by the epoxy. Translating these uncertainties through the equation for incoming longwave radiation leads to uncertainties in the incoming longwave radiation of ±2.7, ±2.6, and ±2.3 W m^{-2} (see Payne and Anderson 1999). The laboratory calibration at Woods Hole Oceanographic Institution (WHOI; Payne and Anderson 1999) relates the signal derived from the thermopile voltage and the thermistors to the radiation from a blackbody and involves presoaking the instrument over a cold (hot) bath and then transferring to a blackbody suspended in a hot (cold) temperature controlled tank. This procedure is repeated at a series of blackbody temperatures (0.1°, 5°, 10°, 30°, 40°, and 50°C) and two constants are determined for a linear fit.

The results of a series of extended calibrations illustrate the calibration errors (Fig. 2). The main panel and bottom inset show the difference between the blackbody longwave radiation and the longwave radiation determined from the ensemble of calibrations for each ensemble member and for three different instruments. The top inset shows the mean discrepancy (essentially the time mean of the bottom inset) versus the blackbody temperature of that calibration. The initial discrepancy over the first 2 min indicates periods when the sensor is not yet seated in the blackbody. A conservative estimate of the error stemming from the calibration coefficients would be 1.5 W m^{-2}. This is reduced from the value of Payne and Anderson (1999) because we find the coefficients covary ($r = 0.5$) at greater than 99% confidence. This correlation accounts for 30% of the variance of the calibration coefficients in time for a given instrument.
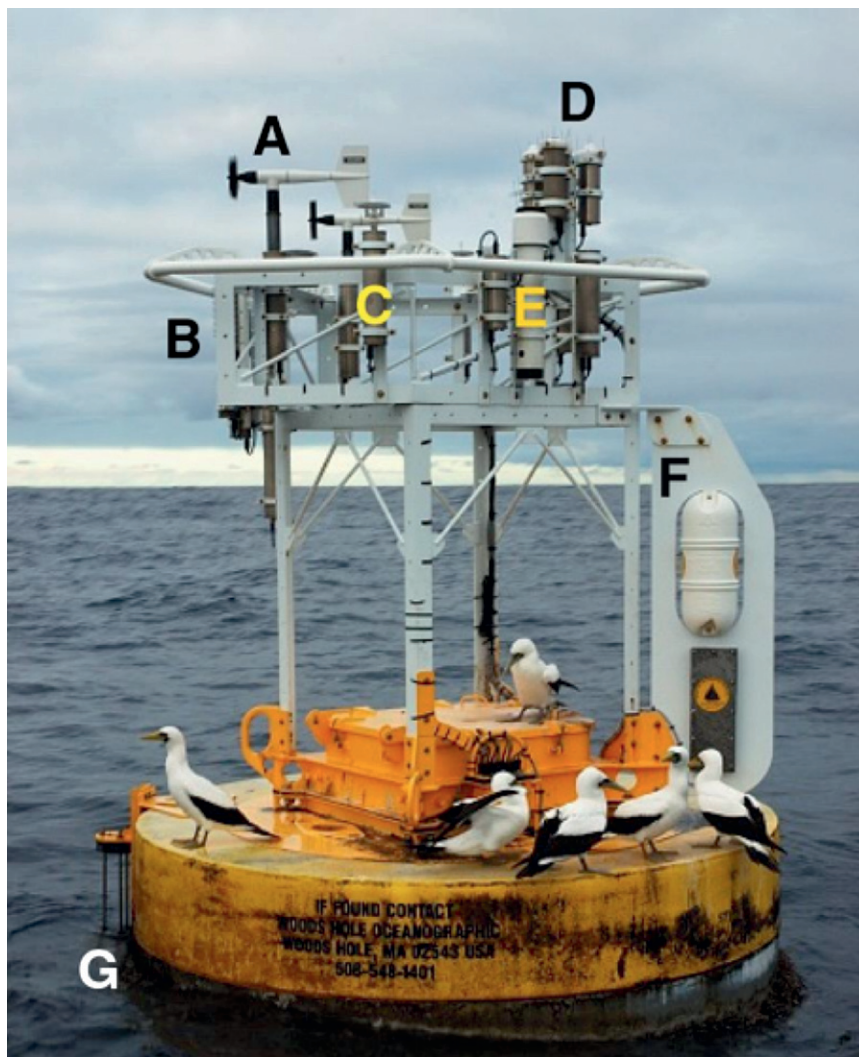
FIG. 1. The current version of the WHOI IMET buoy showing the location of the (a) wind sensors, (b) air temperature–humidity (mainly hidden), (c) barometric pressure, (d) radiation sensors, (e) rain gauge, (f) orienting vane, (g) and floating SST. The two primary near-surface temperature sensors are at 1-m depth and mounted on the buoy understructure.

Another source of error can be the stability of the module's electronics, including temperature dependence. Conveniently for the current applications, the ambient temperature during the deployment has been typically within a few degrees of the ambient room temperature during calibration. Thus, we expect temperature dependence to introduce little error. We have also examined the stability of the amplifier that boosts the output of the thermopile. To address this, we have deployed additional stand-alone longwave modules and compared them against shipboard longwave sensors. These tests pointed to an amplifier in early modules whose offset changed when power was applied; a new, stable amplifier has been introduced. As another check,

we look for periods during the burn-in (our name for the outdoor testing phase that is after laboratory calibration yet before deployment, a period of several weeks), when the sensors have sat in fog. This provides

TABLE 1. Instrument heights above the waterline for the Stratus 4 deployment.

Sensor	Height (m)
Wind speed and direction	2.97
Air temperature–RH	2.57
Barometric pressure	2.37
Precipitation	2.73
Solar radiation	3.16
Longwave radiation	3.16

Fig. 1 live 4/C

TABLE 2. The estimated components of the instantaneous error for longwave sensors, and the expected total error for the 1-min measurements, daily averages, and annual averages.

Precision	Lab calibration	Drift	Field errors	Tot
Dome temperature: 0.1°C Case temperature: 0.1°C Thermopile: 10 μV	Coef: 1.5 W m^{-2} Noise: 0.5 W m^{-2}	2 W m^{-2}	Tilt: <2 W m^{-2} T Gradients: 4 W m^{-2} Salt spray: <1 W m^{-2} Solar: <1% SW↓	Instant: 7.5 W m^{-2} Daily: 4 W m^{-2} Annual: 4 W m^{-2}

[AU11]

an additional test for the absolute accuracy of the longwave sensor, because it should be equal to the value calculated from the surrounding air temperature. The above calculation typically agrees within 1 W m^{-2}, which is compatible with an air temperature error of ±0.2 K.

There are several important sources of error that occur when radiometers are deployed on a buoy. The four most important ones are due to solar contamination, tilt effects (both mean and time varying), thermal gradients in the dome and case temperatures, and dome contamination at sea (e.g., salt spray crystallization, bird guano, etc.). Solar leakage in the longwave sensor is often a very large, noticeable error and is either picked up in the predeployment phase or else would render the data highly suspect (see, e.g., Payne and Anderson 1999, Fig. 10). Other authors (e.g., Pascal and Josey 2000) have worried about lower levels of incident solar radiation passing through the longwave dome. Pascal and Josey (2000) found that for four different Eppley PIRs the domes passed 0.7%, 1.1%, 1.1%, and 2.4% of the incident shortwave. During our predeployment burn-in procedure, the instruments are mounted outside on the roof for several weeks and their data are examined for evidence of shortwave leakage. The same check is done with the field data. Only one of the eight Stratus and four NTAS longwave modules showed signs of shortwave leakage, and the postcorrection determined that this module was passing about 0.8% of incoming shortwave radiation.

Tilt affects the radiometers by rotating their field of view away from the vertical. This effect is not as important for longwave radiation as it is for shortwave radiation because the source is typically more diffuse and the contrast between sea surface temperature and cloud temperature is often not dramatic. Two types of tilt are sources of error: one is a mean tilt and the other

is the rocking of the buoy by the surface gravity wave field. Assuming the worst case, with clear skies for maximum air–sea contrast (400 W m^{-2} outgoing versus 320 W m^{-2} incoming) and a mean tilt of 2° gives a +1.75 W m^{-2} error in the measured incoming longwave radiation. In a region like the Stratus mooring with persistent cloud cover, this error is <0.5 W m^{-2}. The swell-induced tilting is a smaller error than that due to the mean tilt, particularly for longwave radiation, where the signal is generally diffuse. MacWhorter and Weller (1991) studied the effect of mean tilt and rocking on shortwave radiometers. Assuming that shortwave and longwave thermopiles have the same time response, we used their results to estimate a tilting error of 0.75 W m^{-2} for longwave radiation. Because of the diffuse nature of the incoming longwave radiation, it seems that this is likely an overestimate. We suggest that the overall error from tilting is <2 W m^{-2} for Stratus and NTAS. It should be noted that considerable effort is made to level the radiometers on the buoys with respect to the anticipated waterline and that the mooring line underneath the buoys is under high enough tension so that the buoys tend to slide up and down waves rather than rock back and forth.

The existence of thermal gradients within the case and the dome because of differential heating has engendered significant research (Philipona et al. 1995). Although these will effect the instantaneous measurements, they should have no effect on the long-term averages, unless the gradients have a preferred orientation with respect to the thermistor (i.e., the thermistor is always on the shaded side of the dome). A buoy is not fixed; because of the inclusion of an orientating vane on a buoy, we might expect in steady winds to find that there is a preferred exposure of the dome thermistors to the sun. Using previously published measurements of

TABLE 3. As in Table 2, but for shortwave sensors.

Precision	Lab calibration	Drift	Field errors	Tot
0.1 W m^{-2}	2 W m^{-2}	<2 W m^{-2}	Tilt: <2% Temperature gradients: 1–2 W m^{-2} Salt spray: <1 W m^{-2}	Instant: 20 W m^{-2} (more in broken cloud) Daily: 6 W m^{-2} Annual: 5 W m^{-2}

TABLE 4. As in Table 2, but for humidity sensors.

Precision	Lab calibration	Drift	Field errors	Tot
0.01% RH	Linear: 0.16% RH Cubic: 0.1% RH	Linear: 0.9% RH Cubic: 0.9% RH	<95% RH: ±1% RH Heating: 3% RH (in low winds)	Instant: 1% RH (3% RH in low wind) Daily: 1% RH (3% RH in low wind) Annual: 1% RH

the thermal gradients in the dome (Philipona et al. 1995), we calculate that the maximum instantaneous error is 4 W m^{-2}. Finally, the accumulation of salt spray or other opaque materials on the dome of the radiometer could conceivably scatter radiation and affect the measurement. Our radiometers are recalibrated after deployment without first cleaning the domes. The postcalibrations of radiometers with uncleaned domes versus those with cleaned domes show no noticeable impact of dome exposure. At the same time, the comparison of pre- and postcalibrations points to an estimate of drift at 2 W m^{-2}.

The various sources of error are listed in Table 2, along with estimates of the instantaneous, daily, and annual absolute errors. Here, it is assumed that the individual errors are uncorrelated and that some of them cancel or partially cancel in the longer averages. In particular, it is assumed that the thermal gradients, although contributing 4 W m^{-2} to the point-to-point error, only contribute 2 and 1 W m^{-2} for daily and annual averages, respectively; it is also assumed that shortwave leakage contributes $+10$, $+2$, and $+2$ W m^{-2} to the instantaneous, daily, and annual absolute errors, respectively, of which 70% is postcorrectable.

b. Shortwave radiation (Eppley PSP)

The Eppley precision spectral pyranometer (PSP) is superficially similar to the longwave radiometer, except that it lacks dome and case thermistors and has a double glass dome. The single output is the thermopile voltage whose accuracy is equivalent to 0.1 W m^{-2}. The gain on the shortwave amplifier is two orders of magnitude smaller than on the longwave amplifier. We rely on the manufacturer's calibration with one minor adjustment. All the instruments are compared with traceable standard instruments on the roof at WHOI. After the burn-in, the data from the instruments are compared and a

simple linear correction is applied to the test instrument. This correction is typically small, about 2–3 W m^{-2}. After a one-year deployment, most of the shortwave radiometers that are postcalibrated are found to differ from the rooftop standards by about 2–3 W m^{-2}. It is difficult to determine if this 2–3 W m^{-2} error is due to slow instrument degradation, actual contamination of the dome at sea, or to uncertainty in the previous calibration. Therefore, 2 W m^{-2} has been assigned to both the calibration uncertainty and the annual drift (or deployment contamination).

The primary field errors for the shortwave radiometers are due to tilt effects and to thermal gradients within the dome. Thermal convection is supposed to be reduced by the double dome construction. However, land-based measurements, where a radiometer is shaded and then exposed to the sun, have shown that there are still residual effects. Bush et al. (2000) have shown that these errors are small for our application of the radiometer ($\pm$1–2 W m^{-2}) and are also likely to cancel in the average. Tilt is potentially the most serious source of error. The time-varying effect resulting from waves is less important than mean tilt. This is because the buoys have small pitch and roll magnitudes. Also, the sun is near zenith at noon for our deployment locations. From MacWhorter and Weller (1991), for estimated underestimation because of rocking of $\pm$10°, the percent error will be -0.5% in the daily average. Of greater concern is a small tilt of the buoy. Extrapolating from MacWhorter and Weller (1991), the error for a 2° tilt is estimated to be 2% in the incoming solar radiation. However, if the orientation of the mean tilt varies in time with respect to the zenith angle of the sun, this error will sometimes increase and sometimes decrease the measured solar radiation. In the long time mean, it contributes a source of error similar to that from wave motions [$O(0.5\%)$]. In calculating the daily and annual

TABLE 5. As in Table 2, but for air temperature sensors.

Precision	Lab calibration	Drift	Field errors	Tot
0.02 K	<0.03 K	0.05 K	0.4 K (wind = 3 m s^{-1}) 0.7 K (wind = 2 m s^{-1}) >1 K (wind < 1 m s^{-1}) Radiation: 0.2 K	Instant: 0.2 K (more in low wind) Daily: 0.1 K (more in low wind) Annual: 0.1 K

TABLE 6. As in Table 2, but for barometric pressure sensors.

Precision	Lab calibration	Drift	Field errors	Tot
0.01 mb	0.06 mb	1.5 mb (max) 0.2 mb (after postcalibration)	Temperature: 0.1 mb Wind: <0.1 mb (for wind < 10 m s^{-1})	Instant: 0.3 mb Daily: 0.2 mb Annual: 0.2 mb

average, we make use of the fact that the average solar radiation value is much less [$O(200$ W m^{-2})] than the peak values. The annual value is also reduced over the daily value because we assume that the tilt error is actually better than 2% for much of the time when the seas are fairly calm.

c. Air temperature (Rotronic MP-100F)

A platinum resistance thermometer adjacent to the humidity sensor measures air temperature. The manufacturer states an accuracy of 0.2°C with a repeatability of 0.1°C. All air temperature sensors are routinely calibrated at WHOI before and after deployment. The calibration fit is accurate to <0.03 K and the observed annual drift has never exceeded 0.05 K. Field error can stem from inadequate ventilation. In low winds, the sensor cavity forms its own microclimate where convective and radiative effects can become important. To maximize ventilation, the air temperature–relative humidity modules are placed on the windward face of the buoy. In low winds, there is little natural ventilation and the air temperature sensor can read anomalously high (Anderson and Baumgartner 1998). The R. M. Young shields used on the module are specified to yield rms air temperature errors under solar radiation of 1080 W m^{-2} of 0.4°, 0.7°, and 1.5°C at wind speeds of 3, 2, and 1 m s^{-1}, respectively. The NTAS and Stratus deployments rarely experienced very low wind speeds (speed <2 m s^{-1} only 2.5% of the time). No correction was made to air temperatures; the Anderson and Baumgartner (1998) empirical correction suggests that this leads to a +0.03 K bias in the annual mean air temperature (for typical stratus conditions).

A more problematic source of error is the radiative forcing, either because of a small fraction of incoming solar radiation reaching the sensor or a temperature difference between the sensor and the multiplate radiation shield. Hubbard et al. (2001) found that, although the Gill shield does allow about 8% of the incoming

solar radiation into the sensor cavity, when placed above a grass surface, this slight positive shortwave forcing was partially offset by a negative longwave forcing. This was due to an average temperature difference between the inner surface of the shield and the air temperature sensor of about −0.5°C during the day (which is contrary to many other early papers that have assumed, although not measured, a positive longwave forcing). Lin et al. (2001) compare the normal operating temperature inside several radiation shields with that found from an energy balance thermocouple (EBTC). They found that, when radiative and convective effects were accounted for in the EBTCs, the air temperatures in different shields were comparable (±0.34°C daily average). Although not significantly different from 0 at 95% confidence intervals, Lin et al. (2001) showed data that were consistent with the listed manufacturer's error. Averaging over intervals with solar radiation >800 W m^{-2}, they found errors of 0.75°, 0.55°, and 0.3°C at 1.5, 2.5, and 3.5 m s^{-1}, respectively. These studies tend to show that the radiative effects on long time scales may average to a small value (<0.1 K) but the instantaneous measurements and the diurnal cycle may have more serious errors.

d. Humidity (Rotronic MP-100F)

The IMET system uses a relative humidity sensor in which the capacitance of a dielectric material varies as it adsorbs and desorbs water molecules. Early versions of such sensors were fragile and exhibited calibration drift. The newer sensors, although still delicate, are more stable over an annual deployment. The instrument resolution is 0.01% in relative humidity. Rotronic states that the sensors are accurate to ±1% RH, repeatable to 0.3% RH, and have a calibration stability of better than 1% RH per year.

All the IMET instruments are calibrated in a Thunder Scientific 2500 humidity chamber. In the past, calibration of sensors over salt solutions led to corrosion of sensor leads and premature sensor failures; that practice

TABLE 7. As in Table 2, but for SST sensors.

Precision	Lab calibration	Drift	Field errors	Tot
0.001 K	0.001 K	0.05 K 0.03 K (after correction)	Low wind: 0.1 K Cool skin: <0.02 K	Instant: 0.1 K Daily: 0.1 K Annual: 0.04 K

TABLE 8. As in Table 2, but for wind speed sensors.

Precision	Lab calibration	Drift	Field errors	Tot
			Tilt: <0.3%	Instant: max(1.5%, 0.1 m s^{-1}) more in low wind
0.002 m s^{-1} (i.e., 1 pulse min^{-1})	1%	+0.1 m s^{-1}	Sea state: uncertain Very low wind: ±1 m s^{-1}	Daily: max(1%, 0.1 m s^{-1}) Annual: max(1%, 0.1 m s^{-1})

was abandoned in favor of the humidity chamber. The instruments are subjected to relative humidities from 20% to 95% in 5%-RH increments. The new calibration is then determined from either a linear or cubic fit to the data. The calibration fit has a residual error of about 0.1% RH. Our calibration facility does not reliably perform above 95% relative humidity. As a test of the linearity near saturation, a calibration in this range was performed at Thunder Scientific. The test showed that the nonlinearity is not large, with maximum deviations of about 1% at 100% relative humidity (Payne XXXX, personal communication). However, these conditions are relatively rare in the data examined here, and the latent heat flux at these humidities is small, so errors in the heat and salt fluxes are negligible.

Calibration drifts in our instruments over the year are small (although potentially more than the manufacturers value of 1% RH). Figure 3 plots the difference between the humidity calculated from the old and new calibration coefficients at the time of recalibration (thin curves) for many different humidity sensors. The mean of the calibration changes (thick solid; one standard deviation is thick dashed) is indistinguishable from zero at 95% confidence, which indicates that sensor change is not systematic in time. The recalibrations are anywhere from several months to two years apart, but the magnitude of the calibration change does not appear to be strongly correlated with time.

It is hypothesized that the relative humidity sensors suffer from two forms of calibration change. One change is a gradual linear drift, presumably because of slow changes in the dielectric and the electronics. There are many deployments where the in situ comparisons of the buoy at sea show discrepancies that are consistent with a perfect precalibration linearly degrading toward the value at postcalibration. This is encouraging because it implies that a simple linear postcorrection would improve the data. On a small subset of the deployments

(about 20%), the humidity sensors show a second behavior, demonstrating an episodic change in calibration during the shipping process and perhaps pointing to some continuing sensitivity of the sensors to shock, vibration, or other conditions encountered in shipping. As a consequence, the initial in situ comparison of the buoy might show a large [$O(2\%$ RH)] shift in humidity in comparison to the shipboard sensors. Then, from this point onward, the sensor error evolves linearly (i.e., the deployment in situ comparison, recovery in situ comparison, and postcalibration error values are linear).

We do protect the sensors with a porous Teflon sleeve. This lowers the response time but has stopped sensor degradation resulting from exposure to marine air. Liquid does not penetrate the Teflon sleeve. The Teflon sleeve, rather than the sensor itself, largely governs sensor response times. The manufacturer states that there is a 12–15-s response time for the sensor. The sleeve slows the response time to approximately 1 min, but it is essential for excluding saltwater from the sensor and shedding salt crystals left by evaporation. We do observe differences between the ability of individual sensors to recover from very moist conditions (>95% RH); we hypothesize that the rate at which the water molecules leave the dielectric sensor varies from sensor to sensor. However, these performance differences are restricted to very moist conditions, and exposure to such conditions has not been observed to not impact the sensor performance at lower humidities, where pairs of sensors together track changes in humidity.

In Fig. 3, the circles represent the average difference between the calculated calibration curve (cubic on the left and linear on the right) and the data used to calculate the calibration. It is thus the systematic error of our calibration approach. The misfit explains most of the mean variance (i.e., thick solid line and circles are similar), including the odd spike at 35% RH. The calibration misfit is thus applied as a correction to the data

TABLE 9. As in Table 2, but for wind direction sensors.

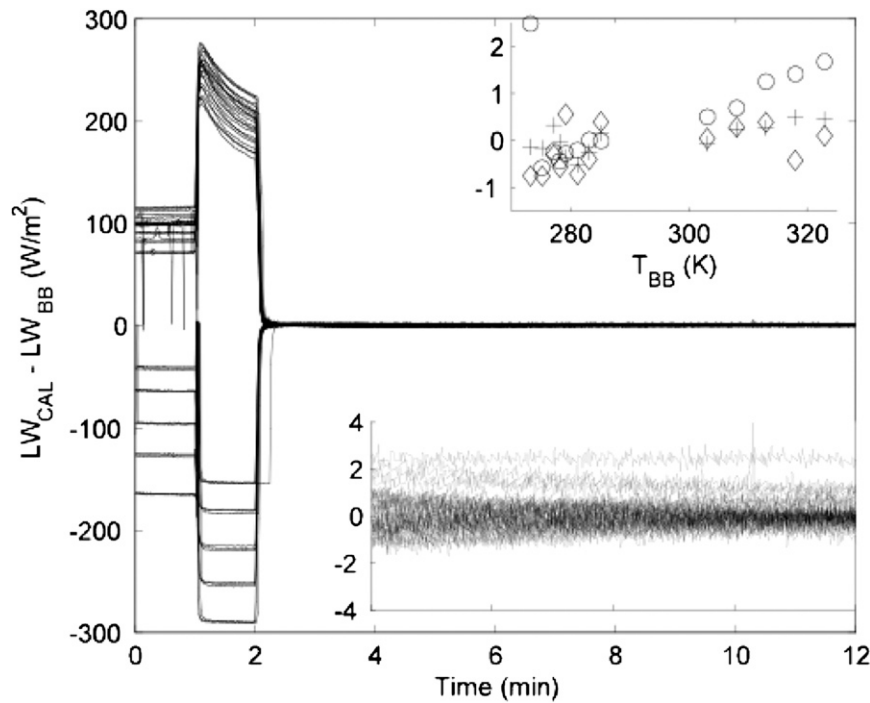| Precision	Lab calibration	Drift	Field errors	Tot
0.1°	1° (buoy spin 4°)	2°	Low wind: $O(1°)$ Flow distortion: <5°	Instant: 6° (more in low wind) Daily: 5° Annual: 5°

FIG. 2. (main) Discrepancy between the blackbody and calculated longwave radiations from a series of calibrations for three different sensors with (bottom right) expanded axes. (top right) Mean discrepancy as a function of blackbody temperature different symbols for the three different sensors.

after recovery. Also, there is little difference between the stability of the cubic and linear calibration curves. Both suffer from equal calibration drifts, which support the idea that the drift is due to a physical change of the sensor.

There are two possible field errors that can influence the humidity sensor: one is due to contamination of the dielectric sensor or Teflon shield and the other is due to self-heating effects in low wind. Postdeployment calibrations of the humidity sensors both with and without the Teflon sleeve show that it does not affect the calibration coefficients. We cannot distinguish sensor contamination effects, when the Teflon shield is used, from the calibration drift. Radiative heating in low winds is of concern for temperature measurements (Anderson and Baumgartner 1998); however, given an estimate of that temperature error, the observed relative humidity can be used to provide specific humidity and then estimate a corrected relative humidity.

e. Barometric pressure (AIR DB-1A, AIR DB-2A, and Heise DXD)

AU4   Newer IMET systems have switched to a Heise DXD from the Atmospheric Instrumentation Research (AIR) DB-1A and DB-2A as the barometric pressure module sensor. Initial indications are that the instruments have similar error characteristics. The resolution of the barometric pressure sensor is 0.01 mb. The sensors are calibrated in a DHI PPC2+ pressure generator by taking five readings at pressures between 980 and 1040 mb (in 10-mb steps), cycling first from low to high and then back to low pressure. Some hysteresis is noted. There is no mean bias between the sensor pressure and the reference pressure, and the standard deviation is 0.035 mb. The 90% confidence interval on the calibration will thus be 0.06 mb.

Comparing sensors on the same deployment shows that relative drift is fairly linear, indicating that the absolute drifts are probably linear and could be corrected by the postdeployment calibration. The average absolute drift is 0.58 mb with maximum and minimum drifts of 0.92 and -1.45 mb, respectively. Initial attempts at postcorrecting for drift have not been entirely successful. Although the relative instrument drift during the year is linear, correcting for a linear drift based on the postcalibration does not always improve the instrument agreement. This suggests that there may be some additional change in the calibration of individual sensors. However, the postcorrected pressures agree better than the uncorrected pressure for most deployments. The other major effect is due to wind, which is mitigated by the use of a Gill pressure port. Gill (1976)
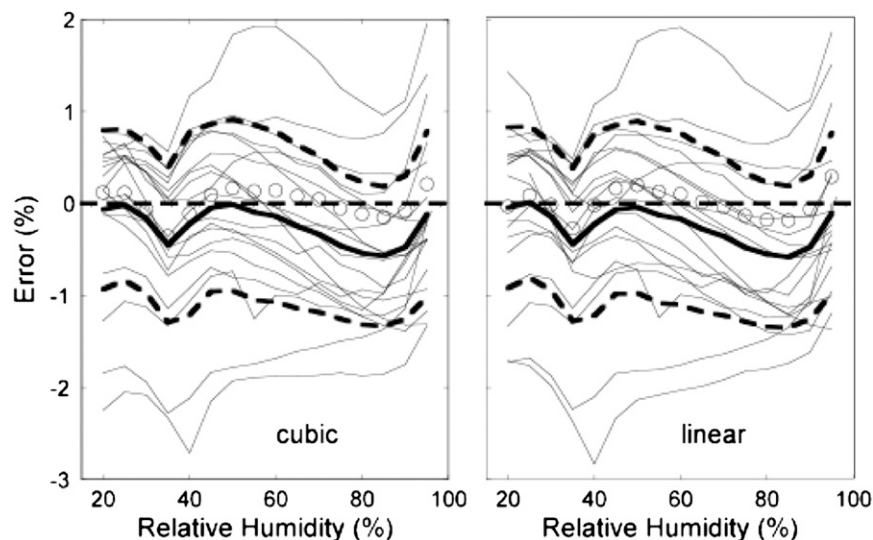
FIG. 3. The change in calibration in percent RH of various RH sensors (thin solid line), the mean calibration change (thick solid line), and one standard deviation (thick dashed line). Also shown is the initial misfit of the calibration curve to the calibration data (circles).

listed the error as 0.4 mb at 20 m s^{-1}, with a quadratic dependence on wind speed. For a typical Stratus wind speed of 7 m s^{-1}, the error would be 0.05 mb.

f. Sea surface temperature (SBE-39)

Present IMET systems use SeaBird Electronics (SBE) model 39 sensors for near-surface sea temperature. The sensor is very reliable; the main issue arises in extrapolating from the measurement depth to the sea surface. Our moorings measure near-surface temperature at a depth of 1 m. Almost always, this is within the turbulent mixed layer, and the surface extrapolation addresses the presence of the cool skin and the possibility of a thin warm layer in low winds (Fairall et al. 1996a). The Tropical Ocean and Global Atmosphere (TOGA) Coupled Ocean–Atmosphere Response Experiment (COARE) 2.6b flux routines that we use to calculate our fluxes and the skin temperature attempt to account for both of these processes. It is difficult to determine error bounds for the COARE algorithms. Because our near-surface measurement of temperature is close to the surface and we have high sampling rates to resolve the temporal evolution, we shall assume that they are reasonably accurate. In calculating the daily and annual averages, we have assumed that low wind contributes errors of 0.1 and 0.01 K, respectively.

g. Wind speed and direction (R. M. Young 5103)

A propeller–vane system from R. M. Young is used for the wind speed and direction. The propeller system is durable and does not suffer from the overspinning effect found in cup anemometers. The R. M. Young sensor has a signal of about 1 Hz per 0.1 m s^{-1} of wind speed. At low wind speeds, this is not realized because bearing drag renders the propeller reading unreliable below 1 m s^{-1}. At higher wind speeds, this translates into a resolution of 0.002 m s^{-1} over a 1-min average. The deviation of the speed calibration between similarly constructed instruments has been found to be indistinguishable from zero. Thus, the wind sensor speed is not routinely pre- and postcalibrated. However, there is a measurable dependence on the type of bearings (up to a 33% reduction in frequency at 1 m s^{-1}). Wind tunnel calibrations have been used to develop an empirical correction to the initial manufacturer's calibration nominal wind speed $(WS_{nom}) = 0.00 + 0.1021F$, where F is the frequency). This correction is applied equally across all sensors using the same type of bearings. The IMET propeller sensors use a nonstandard set of bearings with balls and races manufactured from the same grade of stainless steel to limit corrosion. These bearings are replaced after each deployment. Without routine calibration, it is difficult to determine the annual drift. Testing of a deployed anemometer after recovery showed some initial stiffness that quickly disappeared. This was attributed to corrosion of the bearings during the return shipment, when the propeller is fixed in position. After this thin corrosion was worn off in the first few runs, the measured response was within 0.1 m s^{-1} of the expected response, but it was generally higher for winds between 1 and 10 m s^{-1}. We infer that bearing

friction probably decreases during most of a deployment, leading to a $+0.1$ m s^{-1} shift over the deployment. This gradual increase in measured wind speed appears to be repeatable. The comparison with a shipboard anemometer (corrected for flow distortion and height) shows that where the discrepancy is indistinguishable from zero at a 95% confidence interval upon deployment, the comparison at the time of recovery has the buoy winds higher by 0.15 ± 0.05 m s^{-1}. Unfortunately, the timing of this bearing wear cannot be determined; it is assumed that the wind biases high by a percentage that increases linearly over the year. The wind direction is derived from orientation of the vane relative to the buoy and the absolute orientation of the buoy. Before deployment, the whole buoy is spun and orientation referenced to a surveyor's compass to calibrate the IMET compass. Based on these tests, we find that the wind direction is not accurate to more than 4°.

The main sources of field error are flow distortion and lack of response in very low wind conditions. We attribute the finding that the difference in flow direction between the two sensors tends to have a nonzero mean to flow distortion. It is common to observe steady differences of 5°–10°. Speeds are comparable between the two instruments (the difference in the annual mean is typically 2–3 cm s^{-1}), but this does not give us any indication of how flow distortion might influence the speed. Buoy motion can also affect wind observations by altering the angle of attack of the propeller or by superimposing an oscillatory platform motion. These contributions to error are small, because the mooring line tensions on the buoy bridle are large and the buoy pitch and roll is small. However, we expect that these errors will be small except for with high wind speeds and sea states (Zeng and Brown 1998), which are rare at the current mooring locations. Finally, accurate measurements in periods of very low wind are difficult because of both the response of the propeller system and the inability of the buoy to orient correctly.

h. Precipitation (R. M. Young 50201)

Precipitation is the hardest measurement to make because of the strongly intermittent spatial and temporal nature of rain. Thus, even a perfect point measurement may be unrepresentative of the surroundings. We do not address sampling errors, only the measurement errors. The gauge has a resolution of 0.1 mm and the manufacturer states an accuracy of 2% up to 25 mm h^{-1} and 3% up to 50 mm h^{-1}. Precipitation gauges, such as the ones we use, have been found to be biased because of flow distortion around the sensor. Raindrops tend to be accelerated over the opening of the gauge leading to systematically low measurements.

Koschemeider (1934) proposed one of the first empirical wind speed corrections for this effect. More recently, numerical simulations have been performed (Folland 1988; Nešpor and Sevruk 1999). Nešpor and Sevruk used a computational fluid dynamics model to examine the sensitivity of three different gauges to rain rate, wind speed, and droplet size distribution. Although their results are not easily summarized, it is clear that for small rain rates (<1 mm h^{-1}) all the gauges undersample by at least 10%. A recent intercomparison of several rain gauges on Kwajalein Island (KWAJEX) found that siphon gauges tend to undermeasure when compared to a disdrometer, particularly for small droplets (Yuter XXXX, personal communication). Siphon gauges can also have errors during periods of intense rain, because the sensor is inaccurate when emptying. Because the need to empty the gauge often corresponds to periods of heavy rain activity, some of this rainfall is not counted. This is not an issue with our datasets and can usually be corrected after deployment.

Assigning any error bars to IMET rain measurements is difficult (particularly in the stratus region, where the annual budget might be based of short showers plus intermittent drizzle). One positive note is that during the years when an acoustic rain gauge was deployed on the mooring line, there was a one-to-one relation between an observable signal in the siphon gauge and in the acoustic gauge. We conclude that if the local rainfall is dominated by periods of fairly steady rain (>3 mm h^{-1}) and if the wind speed is not consistently high (<15 m s^{-1}), then the wind-induced error in the gauge is less than 10% (Nešpor and Sevruk 1999).

3. Data comparisons

Sensor comparison is a crucial step in the verification of sensor accuracy. Past experiments have shown that instrument intercomparison can lead to improvements in instrument accuracy (Weller et al. 2004; Burns et al. 1999, 2000). The Stratus and NTAS deployments have generated more than seven years of side-by-side instrument comparisons. The deployment and recovery cruises have also yielded short time windows when up to six instruments are available for comparison (two IMETs on old buoy, two IMETs on new buoy, shipboard meteorological sensors, and shipboard direct-covariance measurements). The discrepancies between these different measurements will be compared to see if they are consistent with the error characteristics postulated in section 2.

a. Module comparisons

The long time series of collocated sensors enable us to examine the degradation of accuracy with time.

However, it is important to realize that the discrepancy between the 1-min sampling on the buoy should not necessarily agree with the listed instantaneous accuracies in the preceding tables. Because the modules are on the same platform, they will both experience some of the same errors (e.g., radiometer errors resulting from buoy motion). After the moorings are recovered, the data are downloaded and evaluated. Initial processing removes spikes and other clearly unphysical data points, although these are rare, and it checks for clock drifts. Although corrections are made for clock drift, these can only be done to the nearest minute. Thus, two modules could differ in their time window by up to 30 s. In a highly variable environment (e.g., quickly moving broken cloud), this could still lead to large sensor discrepancies.

1) LONGWAVE DATA

Three days of longwave (and shortwave) radiation from the first Stratus deployment are used to illustrate the typical agreement between two modules (Fig. 4). The top panels show an average day with some intermittent breaks in the cloud cover. The middle panels show a day with some broken cloud in the morning that becomes clear from approximately noon onward. The bottom panels show a day with dense, unbroken stratus. The leftmost plots show the incoming longwave signal from the two sensors (module 1 is black and module 2 is gray). The middle-left plots show a log plot of the absolute value of the longwave sensor difference for 1-min samples (gray dots) and 15-min averages (black circles). The right-hand panels are similar representations of incoming shortwave radiation (discussed later). The horizontal black line in each difference plot represents the previously determined uncertainty in the absolute value of the radiation. As mentioned before, the sensors should agree to better than the value in the previous section because they experience some of the error sources equally. In the case of longwave radiation, it is expected that the time-varying tilt error is the same between the two modules. The remaining sources of error are independent for the two modules. Thus, the expected discrepancy between the longwave sensors is 6 W m^{-2}.

The observed sensor agreement is generally consistent with the uncertainty determined in the previous paragraph, especially with some limited averaging to counter for variations in either the clocks or the microclimate between the two sensors. The data disagree the most during periods of strong insolation (i.e., when the longwave drops below about 360 W m^{-2}). A possible explanation is that the two domes have different shortwave pass characteristics. However, an examina-

tion of the longwave discrepancy between the two domes as a function of incoming shortwave radiation, for clear-sky conditions, shows that the difference in pass characteristics is indistinguishable from zero at 99% confidence. This could also imply that one of the instruments is experiencing stronger unresolved thermal gradients. This could be due to the orientation of the buoy, leaving the body of one instrument shaded while the other is exposed to direct sun. This is anticipated only during low sun angles. Some of the noticeable downward spikes in the longwave values are a known problem whereby the Argos satellite data transmitter interferes with the electronics. This is particularly clear in the bottom panel, where the hourly signal of the satellite transmission is evident. These points would be removed in the data processing, but we have chosen to present the raw signal, only adjusting for 1–2-min relative clock drifts.

2) SHORTWAVE DATA

Figure 4 also shows three days of shortwave measurements from the first Stratus deployment. The predicted instantaneous absolute error (±20 W m^{-2}) is indicated by the black line in the difference plots. A value of 20 W m^{-2} (about 2% of the maximum) is appropriate for the sensor-to-sensor comparison, even though the two sensors feel the same buoy motion. This is because the tilt error is probably dominated by small mean tilts [$O(1°)$], which could be different for the two shortwave sensors, even though they are hard mounted to the same platform. The time periods chosen are the same as for the longwave measurements earlier, except that the shortwave signal is only plotted during daylight hours. The three days in Fig. 4 represent a range of different scenarios from broken cloud (top), to mostly clear (middle, from noon onward), to dense unbroken cloud (bottom). During the instances of broken cloud, we see that the two modules can have large discrepancies (>100 W m^{-2}) owing to the nonsynchronicity of the 1-min samples. This indicates the importance of accounting for clock drift before examining sensor discrepancies. The 15-min averages (red dots) show much better agreement, which is generally within the expected uncertainty for point measurements of ±20 W m^{-2}.

The fact that the error is dominated by the tilt term is apparent in the bottom of Fig. 4. The maximum discrepancy between the sensors does not occur during the maximum in incoming solar radiation, at about 1400 LT. Rather, it occurs in the brief period near sunset (1700–1800 LT) when the sun breaks through the clouds (see the associated incoming longwave). This is because most of the day the dense clouds generate a diffuse source of incoming solar radiation from the whole sky
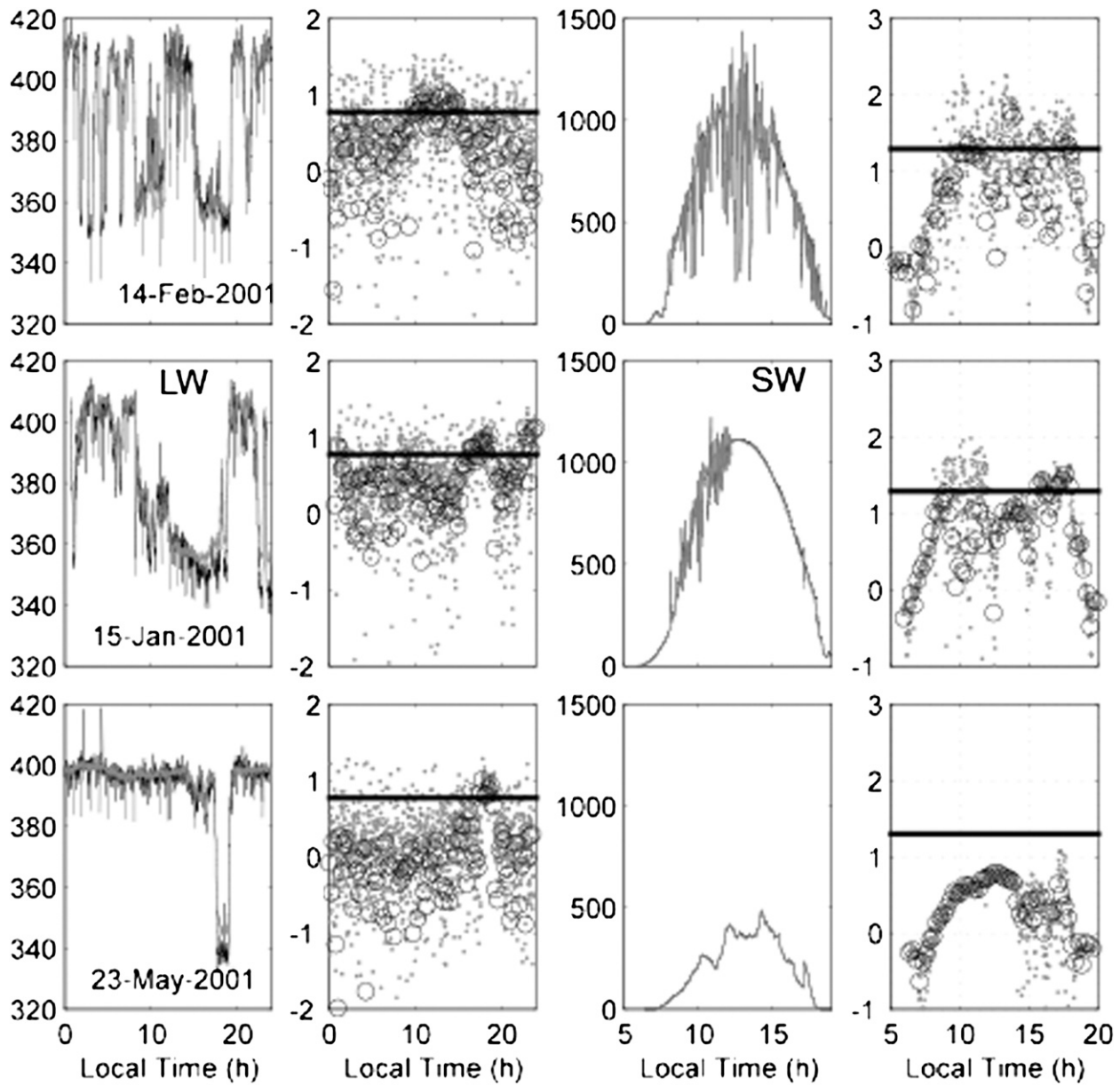
FIG. 4. Observed (left) longwave (W m^{-2}) and (left middle) shortwave (W m^{-2}) radiation for (top)–(bottom) three days during the Stratus 1 deployment. The raw observations show the 1-min data from the two modules [logger 1 (black) and logger 2 (gray)]. (right middle), (right) The difference between the two sensors with a vertical logarithmic scale [1-min data (gray dots) and 15-min averages (black circles)]. The horizontal black line in the difference plots is the expected error (mentioned in the text).

that is insensitive to tilt errors. By contrast, the brief period near sunset has a relatively weak but localized solar radiation source, for which tilt errors would be important.

3) HUMIDITY DATA

Three distinct segments of the humidity record are
F5   shown in Fig. 5 (note that these are distinct from the time periods used in the radiation plots but are the same

as for air temperature). The top panel shows a typical 24-h period from the data. The middle panel shows a 48-h period where the humidity gradually increases from fairly dry to more normal conditions. The bottom panel shows a period with rapid changes and a sustained period of high humidity. The relative humidity mismatch between sensors will be dominated by the calibration drift and thus should be 1% RH with little time variance on daily scales. In low wind, the modules should respond
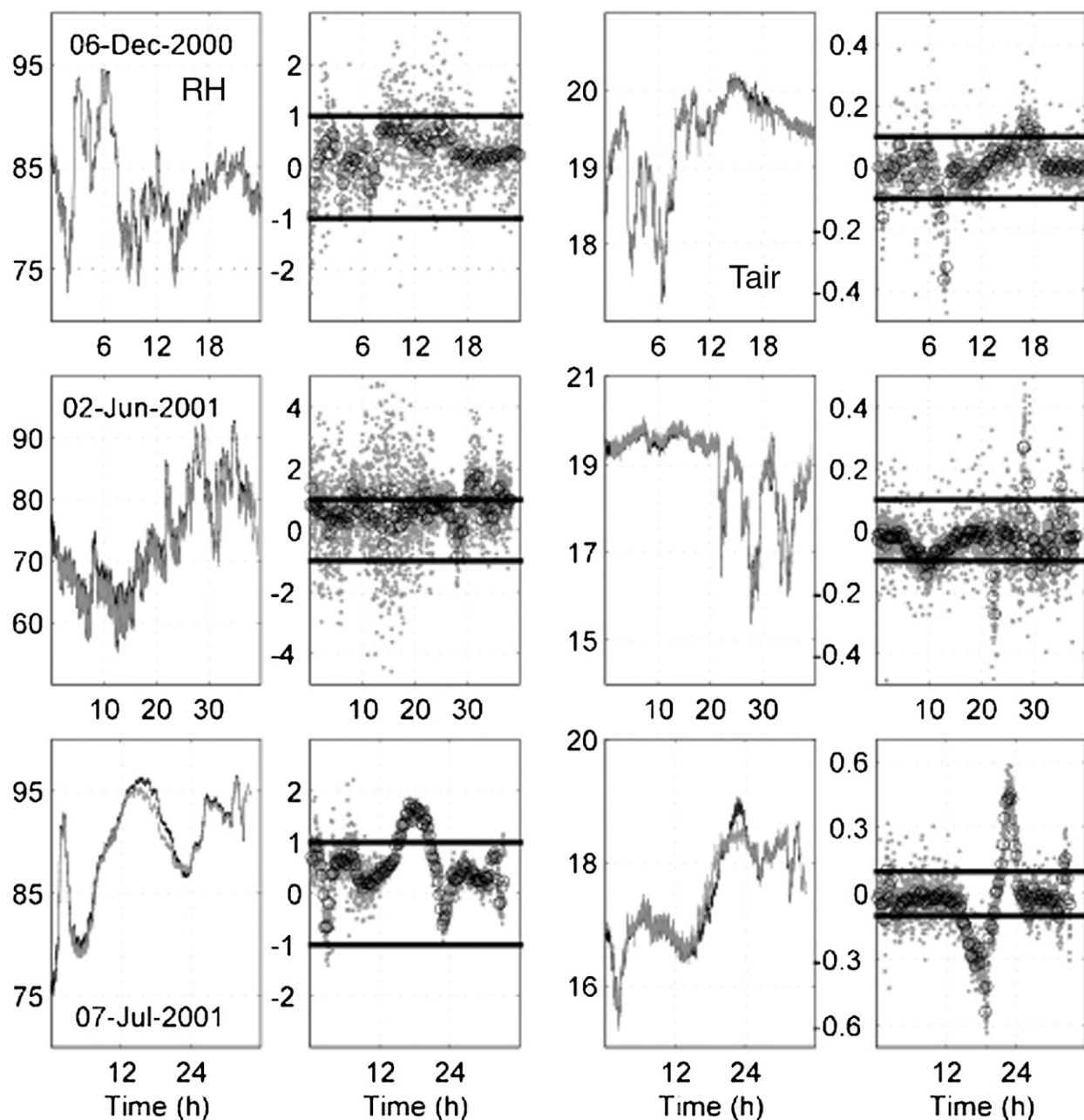
FIG. 5. As in Fig. 4, but for (left) RH (% RH) and (right) air temperature (°C). Also, the vertical axes in the difference plots are on a linear, not logarithmic, scale. The three days chosen are not the same as in Fig. 4.

similarly, which implies that the sensor mismatch should be largely invariant to wind speed.

The top panels in Fig. 5 show a normal summer day with higher humidities in the morning followed by slightly decreased values later in the day as the cloud cover burns off. The 1-min readings are generally bounded within the ±1% RH expected error because of calibration drift. Limited averaging improves the agreement by removing the additional variance result-

ing from slight relative drift in the individual module clocks. The middle panels represent a return from dry conditions to a more normal humidity. Again, the humidity values are bounded within the expected errors. The mismatch is worse than the top panel, possibly because of the later date of these data, indicating that the module calibrations are diverging in time. The bottom panel indicates a particularly difficult period in the data, with humidity near saturation and very low
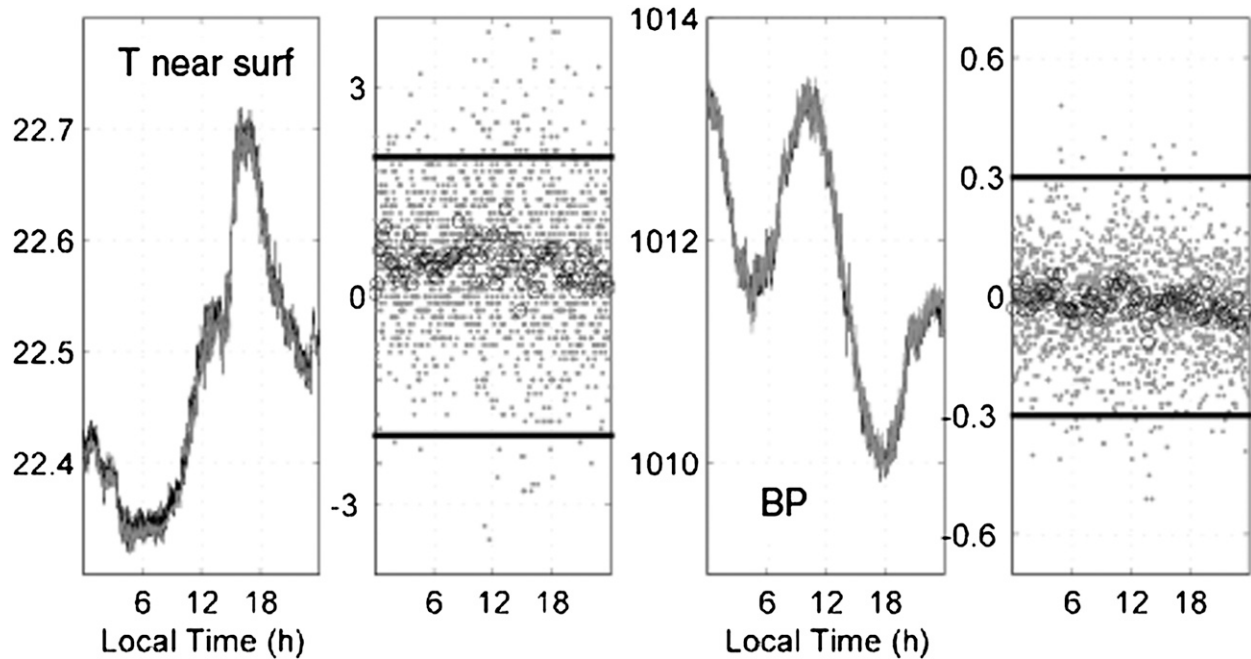
FIG. 6. As in Fig. 4, but for (left) near-surface temperature (°C) and (right) barometric pressure (mb). Also, the vertical axes in the difference panels use a linear, not logarithmic, scale. Only one day (14 Feb 2001) is displayed. Note that (left) is in °C ×100.

AU10

winds (<3 m s^{-1} from hours 10 to 30). The 1-min variance is greatly reduced from the above plots. Because the humidity signal itself has less high-frequency variance (at least at our resolvable frequencies), this would tend to confirm that much of the scatter in the 1-min sensor differences is aliasing of the high-frequency signal by clock drift. The main feature of Fig. 5 is the strong divergence of the two humidity values after a period of extended high humidity. This known problem was mentioned earlier. Both sensors could be equally affected by the low wind conditions of this period and so low wind-induced error cannot be assessed.

4) AIR TEMPERATURE DATA

Air temperature difference plots are also shown in Fig. 5. The difference between modules should largely be due to drift because the two modules share similar microclimates, except in low winds, where the buoy does not orient into the wind. Also, radiative forcing should be similar. A reasonable estimate of the sensor difference would be 0.1 K (during moderate winds), which is indicated by the black line in the right panels.

The sensors typically agree within this uncertainty at all times, at least in the top and middle panels. The periods during which the sensor mismatch exceeds expectations occur during strong temporal gradients in air temperature. These fluctuations are probably associated with stratus drizzle formation. The air temperature mismatch in the bottom panel shows a more serious and systematic error with variations of ±0.5 K. The period of these dramatic discrepancies occurs during very low winds (<2 m s^{-1}). It is thus possible that the observed difference between the two sensors is real. In this situation, the convective and radiative effects and flow distortion around the buoy become important because the vane is no longer capable of orienting the buoy effectively. The fact that the air temperature discrepancy returns to within the expected bounds when the wind speed is >3 m s^{-1} (between hours 12 and 30) adds weight to this interpretation.

5) BAROMETRIC PRESSURE DATA

The expected errors in pressure are dominated by the relative drifts in the calibrations and should be bounded by the 1-min error of 0.3 mb. The pressure error is remarkably monotonous, probably because the signal is dominated by the low frequencies. Barometric pressure and the mismatch between the two sensors (Fig. 6) show that the error is dominated by a simple bias in one or both of the sensors that is not adequately captured by the simple linear correction applied after postcalibration. Still, this bias is well within the predicted bounds for the 1-min error (0.3 mb). The 15-min averages are also all within the predicted daily and annual error of 0.2 mb.

F6

6) SEA SURFACE TEMPERATURE DATA

The near-surface sea temperature measurement itself is very precise, but the extrapolation to skin

temperature introduces error. However, that extrapolation to the skin temperature will not show up in differences between the measured in situ temperatures. This error should only be due to relative instrument drift plus some variability from small-scale spatial structure that will average to zero. The differences between the near-surface temperatures are shown for one day (Fig. 6). The variance between the sensors is small, with relative biases in the 15-min averages of <0.01 K. Even in periods where the rate of temperature change is quite fast (0.2 K in 15 min around 1500 LT), the relative error is still small.

7) WIND SPEED AND DIRECTION DATA

The wind speeds, directions, and respective differences are plotted for three time periods from the second year of the Stratus deployment that represent normal conditions (Fig. 7, top), low winds (Fig. 7, middle), and high winds (Fig. 7, bottom). We expect many errors in the field to be equal for the two sensors (at least in moderate winds where the buoy can effectively orient into the wind) and that the relative error should capture the minor calibration drifts between the two sensors, presumably because of differential bearing wear. Thus, sensor mismatch should be 0.1 m s^{-1}.

The sensor mismatch during the normal conditions (Fig. 7, top panel) shows that most of the 1-min signals and the 15-min averages lie within the ± 0.10 m s^{-1} expectation. The error is evenly distributed about zero (at least approximately), indicting that the mean bias is even smaller (≈ 0.02 m s^{-1}). The mismatch is considerably worse during low winds ($\sigma = 0.26$ m s^{-1} versus 0.12 m s^{-1} in normal conditions for the 1-min samples). This is interpreted as stemming from flow blockage blocking effect, because the buoy can no longer orient into the wind effectively. Even the 15-min averages are strongly dissimilar with differences of 25% of the mean. The high wind conditions (bottom panel) show a picture similar to the normal conditions, with slightly increased variance (although as a smaller percentage of the mean). Again, the mismatch shows little absolute bias (<0.04 m s^{-1}).

Wind directions and their differences are also shown in Fig. 7. The top and bottom panels show very similar behavior, with a significant absolute bias ($\approx 7°$) between the sensors but little variance ($\approx 3°$). The small degree of variance is due to the accuracy of the relative vane direction measurement as well as some real small-scale variability. The absolute error is due to two sources: uncertainty in the compass calibration and flow distortion around the buoy. The low wind conditions (middle panel) are particularly bad, which might be expected when the two sensors experience different amounts

of flow distortion and different degrees of blocking by the buoy.

8) SUMMARY

The standard deviation of the 1-mine difference time series, the standard deviation of a 15-min averaged difference time series, and the mean biases are shown in Table 10 for each sensor for five yearlong deployments. Precipitation is too uncertain to list in the table. We note that some of the mean differences in Table 10 (e.g., NTAS 1 humidity) do not represent the absolute uncertainty of the annual mean because we know which of the two sensors is wrong. In the case of NTAS 1 relative humidity, it was noted from the shipboard comparison after deployment that the second relative humidity sensor was anomalously low by about 3%.

If we assume that the annual linear bias between the instruments is correctable after recovery, then we can form probability distribution functions (pdfs) of the sensor mismatch, which have zero mean. Assuming that these must be symmetric, because sensor 1 and sensor 2 are arbitrary, we can form pdfs for each year as well as a multideployment average (Fig. 8). In each case in Fig. 8, the gray curve is the multideployment average pdf and the two numbers represent the 95% (left) and 50% (right) confidence intervals for the average. Thus, 95% of the wind speed records agree within 0.52 m s^{-1}, and 50% of the longwave sensors agree within 1.6 W m^{-2}. The assumption that we can remove a linear-in-time offset between the sensors is not always correct. These pdfs help characterize the time-varying errors but do not impact the error in the annual mean net heat flux (unless the errors are significantly correlated). Instead, the annual net heat flux is sensitive to the absolute biases in Table 10, particularly if these biases are uncorrectable (i.e., we do not know which instrument is correct). To better understand the sensor differences, we can examine their spectra (Fig. 9).

Several features in the plots stand out. The longwave radiation and wind direction errors are the worst (as a fraction of the total signal). Pressure, two years of SST, relative humidity, and wind direction all show evidence of bit noise swamping the difference signal at high frequencies. This is not entirely surprising because the modules are designed for bulk meteorological measurements and do not have the resolution necessary to measure the high-frequency fluctuations. Relative humidity shows diurnal variability in the sensor difference signal but not in the signal itself. This is a case of air temperature errors propagating into relative humidity errors. In cases where the error curves have a minimum near 1 cycle per hour (cph; solar and wind speed), the spectral increase toward higher frequencies is probably
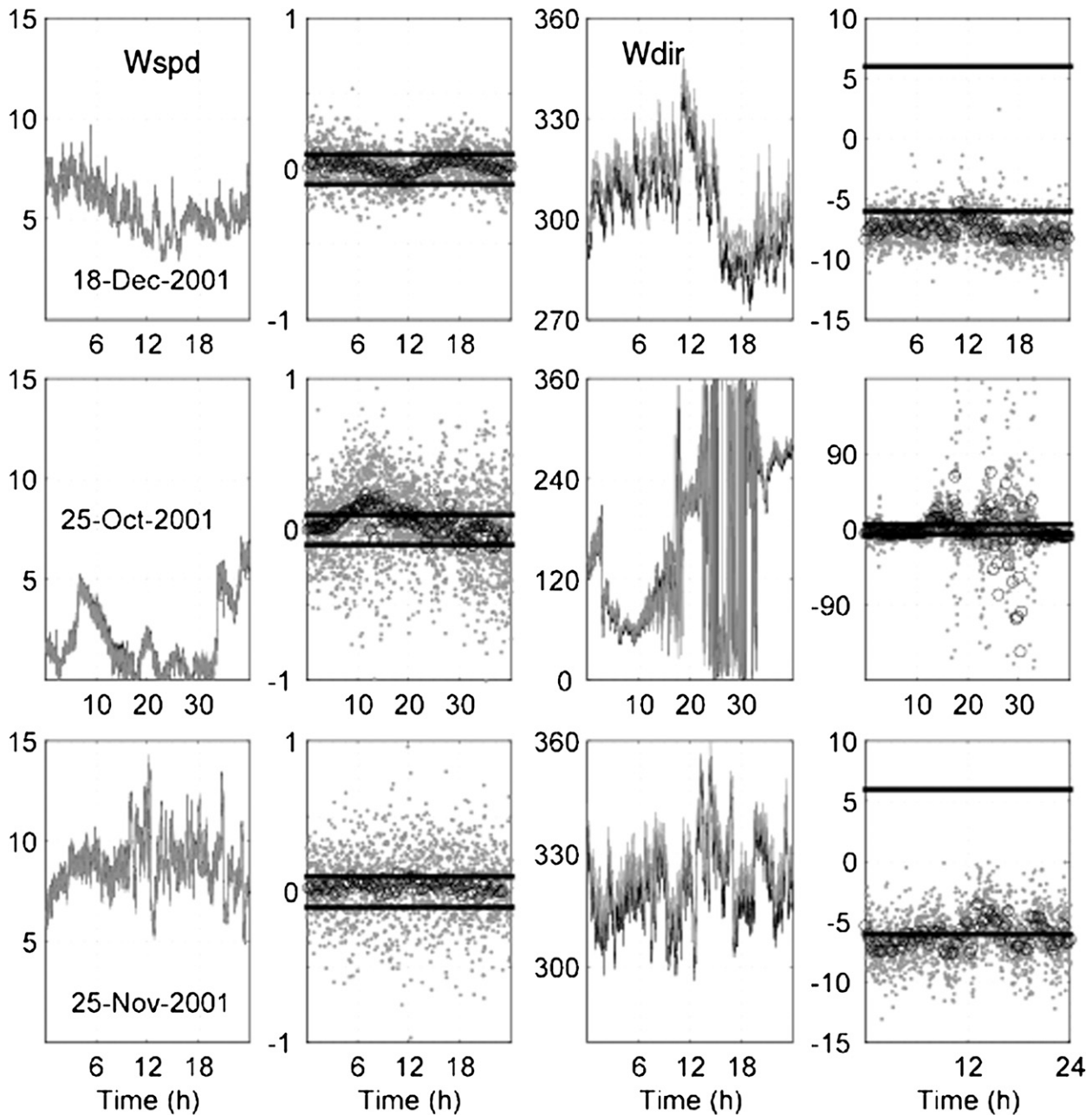
FIG. 7. As in Fig. 4, but for (left) wind speed (m s^{-1}) and (right) wind direction (°). Also, the difference plots are on a linear, not logarithmic, scale. The three days chosen are not the same as in Fig. 4.

the result of relative clock drift aliasing the highest-frequency signal to lower frequencies.

b. Different platform comparisons

Comparing the buoy measurements with those on other platforms provides an independent check of the buoy accuracy. However, one must be careful to account for discrepancies that are a result of errors on the new

platform. An example would be the known problem of flow distortion around ships resulting in errors in the ship-measured wind speed (Yelland et al. 2002). Additionally, the sensors used in the comparison must be as reliable as or better than those on the buoy.

On many of the Stratus deployment and recovery cruises (2001, 2003, and 2004), we have benefited from the presence of personnel from the National Oceanic and Atmospheric Administration (NOAA) Earth System

TABLE 10. The 1-min standard deviation, 15-min standard deviation, and mean of the module difference for each sensor type and for 5 annual deployments.

	Wind direction (°)	Wind speed (m s^{-1})	SST (K)	Pressure (mb)	Air temperature (K)	RH (%)	Shortwave radiation (W m^{-2})	Longwave radiation (W m^{-2})
Stratus 1	—	—	0.049	0.18	0.101	0.90	30.7	5.3
	—	—	0.005	0.08	0.025	0.28	2.6	2.3
	—	—	0.005	−0.05	−0.17	0.76	−2.4	0.0
Stratus 2	5.9	0.28	0.030	0.16	0.089	0.91	25.5	3.2
	1.4	0.03	0.003	0.03	0.010	0.11	4.3	0.3
	−6.0	0.01	0.003	0.05	0.155	0.45	−0.2	0.8
Stratus 3	4.6	0.28	0.014*	0.15	0.109	0.91	28.9	3.9
	1.4	0.05	0.004*	0.02	0.032	0.22	1.1	2.4
	3.2	0.02	−0.001*	−0.14	−0.036	−1.32	−2.1	6.7
NTAS 1	—	—	0.003*	0.16	0.042*	0.58*	14.2	3.6
	—	—	0.000*	0.12	0.021*	0.16*	1.3	1.7
	—	—	0.001*	0.51	−0.062*	3.32*	3.1	0.8
NTAS 2	2.2	0.22	0.005	0.14	0.062	0.70	33.9	3.1
	0.8	0.01	0.000	0.10	0.033	0.38	2.2	1.3
	−2.4	0.00	0.000	−0.21	0.055	0.57	−6.4	−1.5

* Record is less than one full year.

Research Laboratory Physical Sciences Division (PSD; Chris Fairall, Jeff Hare) who instrumented the ship with their own bulk and direct flux measurement systems. We also perform comparisons of shipboard and buoy measurements, at least 24 h in length, on each cruise: one with the new buoy after deployment and one with the year-old buoy before recovery. During these comparisons, the ship holds station a few hundred meters downwind of the buoy with its bow into the wind. Sample comparisons between the PSD measurements on the NOAA ship *Ronald H. Brown* and the Stratus 4 buoy are shown in Fig. 10. Note that some corrections have to be made before comparison. Wind speed has been adjusted for flow blocking and uplift around the ship. This is done using an empirically derived set of corrections obtained during the Joint Air–Sea Monsoon Interaction Experiment (JASMINE); C. Fairall (XXXX, personal communication) states that these corrections are similar to the computational fluid dynamics corrections for the NOAA ship *Ronald H. Brown* determined by Yelland et al. (2002). Air temperature, specific humidity and wind speed have also been height adjusted based on the stability-dependent Monin–Obukhov length, as in the COARE algorithm. Pressure is adjusted for height differences between ship and buoy. Longwave and shortwave radiation are not adjusted.

The buoy–ship disagreement is generally within the expected errors. Some variability within the first 2–3 h is the result of minor ship maneuvers. The air temperature (and relative humidity) on the ship is measured with an aspirated Vaisala temperature–relative humidity sensor, which was checked four times a day with a hand-held Assman psychrometer. The shipboard PSD sensor, standard ship IMET package, and handheld Assman psychrometer were found to agree within 0.09 K and 0.07 g kg^{-1}. The ensemble mean of the buoy air temperatures and specific humidities are also within this error, although the individual modules have a wider spread. One feature noted is the large biases seen in the early morning in the three air temperature sensors; this has been seen in multiple buoy–ship comparisons. The biases disappear when the cloud cover diminishes and is a radiative effect on the shipboard sensors. We postulate that this stems from shortwave forcing with the ship sensor, which uses a different form of radiation shielding, and the probability of stronger reflections from the ship's surface. Hubbard et al. (2001) found that the typical Gill radiation shields had twice the solar radiative forcing when placed over a white surface as opposed to either grass or a black surface.

The ensemble-average air temperature (three buoys plus height-corrected ship) for the period unaffected by the anomalous radiation (0300–1200 LT in Fig. 10) was taken as a best guess for the "true" value. The buoy air temperature observations were modified for a constant offset based on this true value. The magnitudes of the biases were thus −0.15, −0.03, 0.05, and 0.14 K. This form of ensemble averaging of observations has been shown to improve accuracy, primarily by accounting for minor calibration offsets (Weller et al. 2004). Observations on the year-end cruise yielded biases that were the same within 0.05 K.

Relative humidity shows moderate differences between the buoy and ship measurements (average differences of
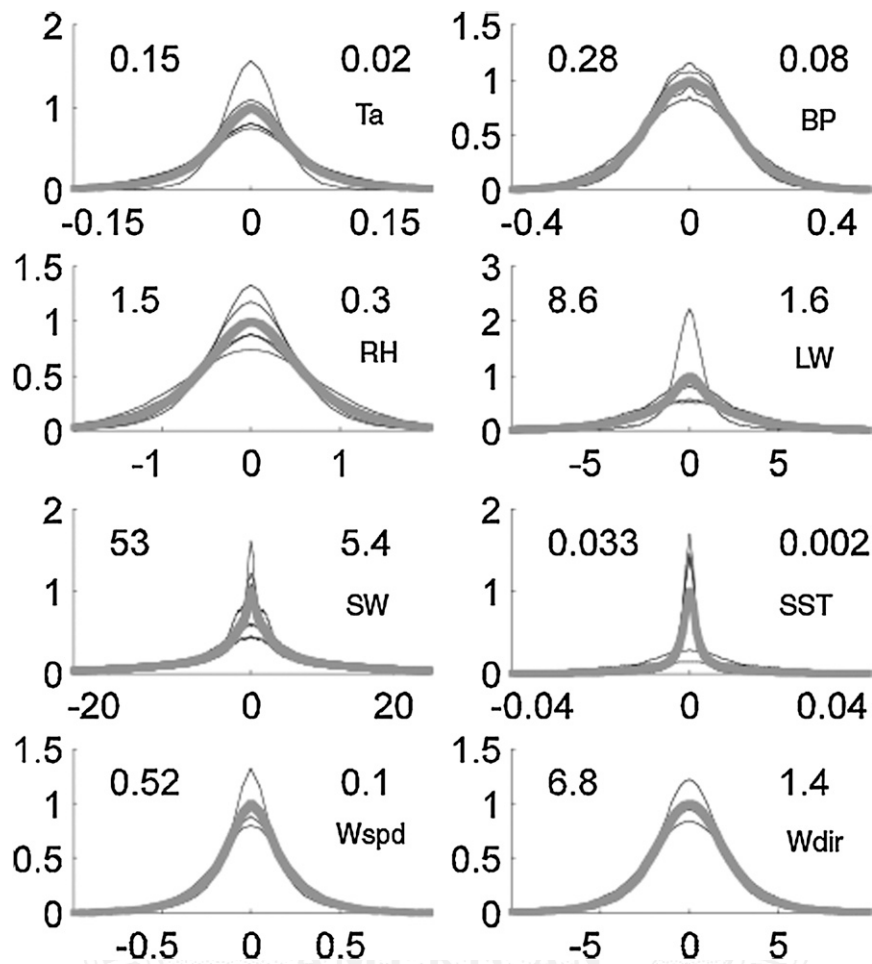
FIG. 8. Probability distribution functions for the zero-mean 1-min module differences for (left)–(right) and (top)–(bottom), respectively: air temperature (K) and barometric pressure (mb); RH (%) and longwave radiation (W m^{-2}); shortwave radiation (W m^{-2}) and SST (K); and wind speed (m s^{-1}) and wind direction (°). The multideployment mean (thick gray) is the average of all the available individual one-year deployments (thin black). The numbers in each panel represent the (left) 95% and (right) 50% confidence limits for each sensor.

0.38%, −0.57%, and 1.7% RH). The buoy–ship comparisons are valuable in that they allow us to flag particular sensors on the buoy (in this case the 1.7% RH mean discrepancy) for special consideration while validating the other two sensors. Wind speed is always above 4 m s^{-1}, which implies that the discrepancies are not related to ventilation. Specific humidity, ignoring the period when we feel the ship air temperature sensor is biased, is in good agreement with the ship values (0.08, −0.17, and −0.25 g kg^{-1}). Figure 11 shows the buoy–ship humidity difference for both the initial comparison and the final ship–buoy comparison for the three different buoy sensors. Also plotted is the difference between the old sensor calibration and the new calibration as determined at the time of postcalibration. Note that one sensor suffered damage on the return shipment and was not post-

calibrated. The stand-alone sensor (circles) had the best performance. Its behavior is consistent with a good precalibration that linearly drifted toward the postcalibration. This drift is also small compared to the other two sensors. Logger 1 (squares) shows little drift from the start of deployment, through the end of deployment, and up to postcalibration. However, it is quite far from its initial calibration (indicating episodic change on shipment to Chile) and has a calibration shift that is strongly humidity dependent. Logger 2 (diamond) has a large drift and lacks a postcalibration; consequently, the chosen relative humidity for this deployment is the stand-alone sensor with a linear temporal correction that shifts between the initial calibration and the postcalibration.

The shortwave radiation sensors are in close agreement (<3 W m^{-2} difference in the average). The main
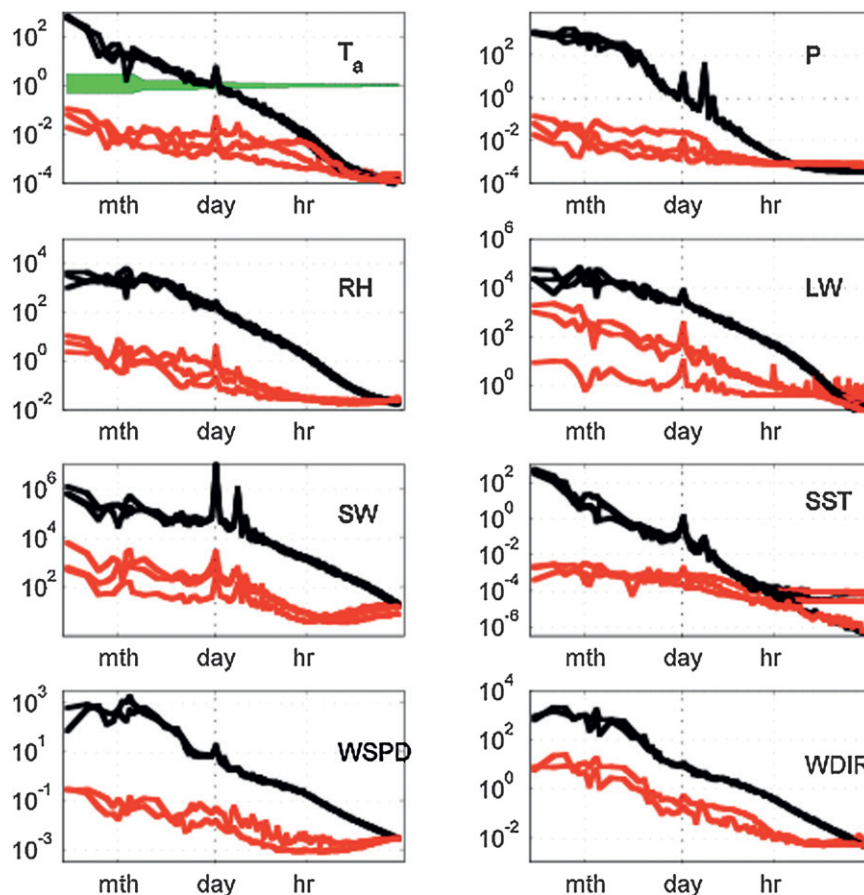
FIG. 9. The frequency spectra of the sensor difference (red) and the signal itself (black) for the first three years of stratus data. The green curve in (a) is the 95% confidence interval, which is the same for all subplots. The ordinate has units of x^2 cph, where x represents the units as listed in Fig. 8.

points of disagreement are during broken cloud, when spatial variability is aliased into the time record. Long-wave radiation sensors showed good agreement. Both the ship and the buoy use Eppley pyranometers and pyrgeometers, so some errors could be reproduced in both systems. Shipboard wind is measured with an IN USA sonic anemometer. Wind speed difference is variable but the daily averages of the buoy sensors are within 0.02 m s^{-1} of the ship wind daily average. Near-surface sea temperature (not shown) has a small mean bias (0.02 K) that is consistent with a weak near-surface temperature gradient (buoy measurement at 1 m versus ship measurement at 0.05 m), but it could also be due to existing spatial variability.

4. Flux errors

The errors calculated for the individual meteorological variables will combine to generate errors in each of the heat flux components and the net flux. For each component, we present the expected error by using a simple bulk formula. The errors in different components need not be uncorrelated. Examples include the time-varying tilt in the incoming solar and infrared radiation and the previously mentioned compensating errors in air temperature and relative humidity. A number of simulations using the more complex COARE algorithm determined that the air temperature–relative humidity correlation (e.g., Anderson and Baumgartner 1998) is the only one with a marked effect on the fluxes. Uncorrelated errors will also produce a slight bias, because the flux formulas are nonlinear. However, a simple experiment with the COARE algorithm shows that our instantaneous meteorological errors, if uncorrelated, would produce a mean error of <0.1 W m^{-2} in the annual average.

a. Net longwave heat flux

Net longwave errors are contributed by errors in measured incoming longwave (LW↓) and errors in sea
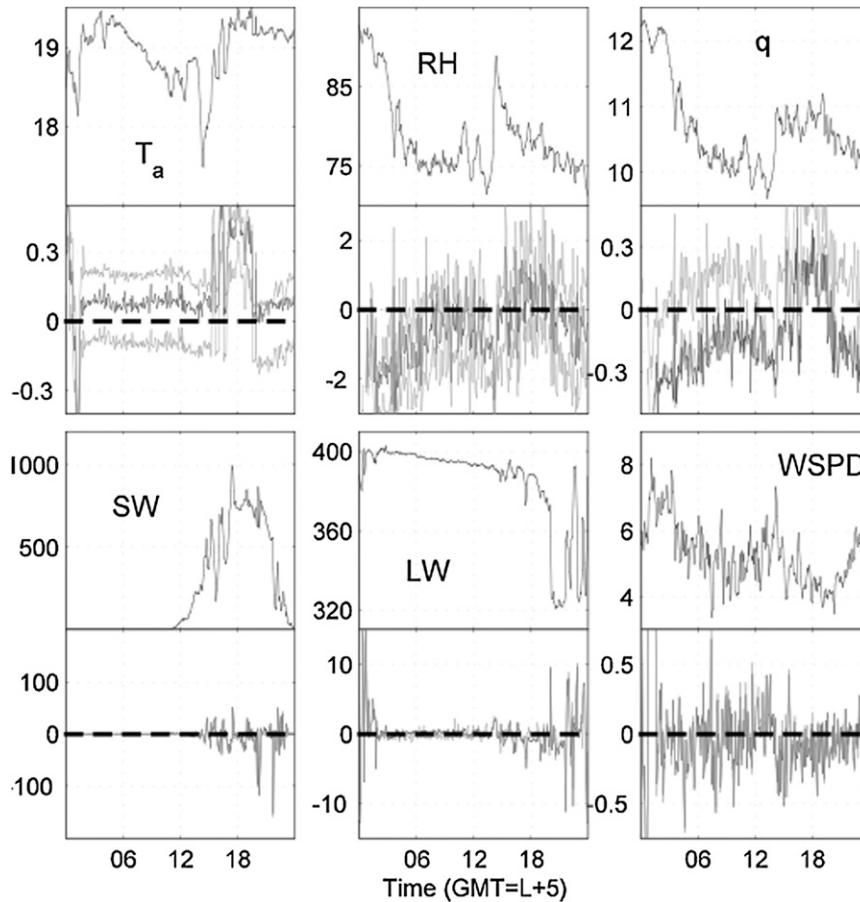
Fig. 9 live 4/C

FIG. 10. (top) Shipboard observations and (bottom) ship minus buoy differences during the 24-h comparison at the beginning of the Stratus 4 deployment. Sensors are (left)–(right) and (top)–(bottom), respectively: air temperature (°C), RH (%), and specific humidity (g kg^{-1}); and incoming shortwave radiation (W m^{-2}), incoming longwave radiation (W m^{-2}), and wind speed (m s^{-1}). On this deployment, there was an additional (i.e., third) RH–air temperature sensor on the buoy. The ship undertook some maneuvering during the first two hours, exaggerating the difference in some sensors.

surface skin temperature (T_{SSST}). The incoming longwave radiation had an rms bias of 1.4 W m^{-2} between the instruments, although most of this came from one deployment (Stratus 3). This mean bias is attributable to differences in the mean tilts, calibrations, and calibration drifts. This observable bias between sensors is less than we had previously surmised in section 2. We shall, therefore, take the more pessimistic approach that these five deployments are anomalous and that the real bias in the annual mean is 4 W m^{-2}. The high-frequency error was described in the pdf in Fig. 9, which had a 95% confidence interval of 8.6 W m^{-2}. This time-varying portion is due to unresolved temperature gradients, thermopile errors, and—potentially—dome contamination (although we have not found evidence of this latter effect). An rms near-surface temperature bias is 0.0015 K. However, this is easily overwhelmed by a bias in the

extrapolation to a skin temperature, and we assume a mean bias of 0.04 K. Shortwave leakage is an open question. We have shown that any leakage we experience is similar in the two sensors (within 0.2% of SW↓) on all the deployments. This would tend to support the idea of a universal constant for all the domes. Consequently, we can always postcorrect for this later. If the leakage is occurring at about 1% of SW↓, then this would add 2 W m^{-2} to the annual mean net longwave radiation, with an uncertainty of only 0.4 W m^{-2}. The error in net longwave (LW) is

$$\partial LW^2 = \varepsilon^2 \partial LW_{\downarrow}^2 + 16\varepsilon^2 \sigma^2 T_{SSST}^6 \partial T_{SSST}^2.$$

Assuming annual mean biases of 4 W m^{-2} for incoming longwave and 0.04 K for skin temperature, with T_{SSST} = 293 K, gives an annual mean bias for longwave of
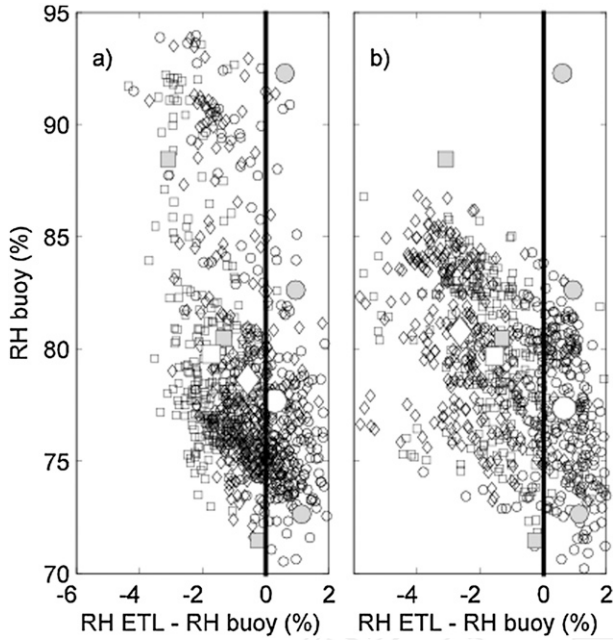
FIG. 11. Ship RH minus buoy humidity [logger 1 (square), logger 2 (diamond), and stand alone (circle)] at the (a) start and (b) end of Stratus 4. Large white symbols are the ensemble average over the ship–buoy comparison. Gray shaded symbols are the change in calibration as determined by the postcalibration.

$\partial LW = 3.9$ W m^{-2}. The mean error is dominated by the uncertainty in incoming longwave with skin temperature more than an order of magnitude smaller.

b. Net shortwave heat flux

The shortwave error depends on error in incoming shortwave and error in the albedo. We take the annual mean bias in incoming shortwave to be 5 W m^{-2} (a conservative value compared to the 1.6 W m^{-2} from Table 3), an albedo error to be 0.01 (see, e.g., Jin et al. 2002), our incoming shortwave to have an annual average value of 200 W m^{-2}, and the average albedo to be 0.058 (because most of the energy is input when the zenith angle is small). Because the error in the shortwave is simply

$$\partial SW^2 = \partial \alpha^2 SW_\downarrow^2 + (1 - \alpha)^2 \partial SW_\downarrow^2.$$

the annual mean bias in the flux is 5.1 W m^{-2} and it is dominated by the uncertainty in the incoming shortwave value.

c. Sensible heat flux

Sensible heat flux (H) errors are due to errors in the wind speed (U_{10}), ocean surface velocity (U_0), air (T_{10}) and sea surface (T_0) temperatures, and uncertainty in the Stanton number (S). For the subtropical sites under

discussion, surface ocean velocities are typically small, with means less than 0.05 m s^{-1}. In these trade wind conditions, U_{10} is two orders of magnitude larger and U_0 has been set to 0. This would not be true in strong boundary currents at the equator or in other situations where stronger surface currents and weaker winds would require including surface currents and their uncertainties in these error estimates. Assuming a simple bulk formula for sensible heat flux and assuming that wind speed and temperature errors are uncorrelated lead to an error estimate of

$$\left(\frac{\partial H}{H}\right)^2 = \frac{\partial S^2}{S^2} + \frac{\partial U_{10}^2 + \partial U_0^2}{(U_{10} - U_0)^2} + \frac{\partial T_{10}^2 + \partial T_0^2}{(T_{10} - T_0)^2}.$$

We will assume the following values for the biases in the annual mean: $\partial S/S = 0.1$, $\partial U_{10}/U_{10} = \partial U_3/U_3 = 0.01$, $\partial U_0 = 0.01$ m s^{-1}, $\partial T_{10} = 0.1$ K, and $\partial T_0 = 0.04$ K, using the uncertainty in the wind observed at 3 m as an estimate for that at 10 m. For mean meteorological conditions at the Stratus buoy ($S = 0.82 \times 10^{-3}$, $U_{10} = 6$ m s^{-1}, and $T_{10} - T_0 = 1$ K), this gives an annual mean bias in the sensible heat flux of $\partial H = 0.15H$. However, sensible heat flux at the buoy is always less than 10 W m^{-2}, so the error is less than 1.5 W m^{-2}. The error is dominated by the uncertainties in air temperature and the Stanton number, which are an order of magnitude larger than the SST or wind error.

d. Latent heat flux

Latent heat flux (L) is dependent on errors in wind speed (U_{10}), ocean surface velocity (U_0), specific humidity at saturation [$q_{10} = q(T_{10})$], RH, and uncertainty in the Dalton number (D). We shall ignore the very weak pressure dependence; a 50-mb change is approximately equivalent to a 0.5-K change in air temperature. The latent heat bulk formula error is similar to that for sensible heat flux with the form

$$\left(\frac{\partial L}{L}\right)^2 = \frac{\partial D^2}{D^2} + \frac{\partial U_{10}^2 + \partial U_0^2}{(U_{10} - U_0)^2}$$
$$+ \frac{RH^2 \partial q_{10}^2 + 0.96 \partial q_0^2 + q_{10}^2 \partial RH^2}{(RH q_{10} - 0.98 q_0)^2}.$$

Our assumptions for the mean biases over an annual cycle are $\partial D/D = 0.04$, $\partial U_{10}/U_{10} = 0.01$, $\partial RH = 0.015$, $\partial q_{10} = 0.10$ g kg^{-1} (assuming $\partial T_{10} = 0.10$ K), $\partial q_0 = 0.03$ g kg^{-1} (assuming $\partial T_0 = 0.04$ K), $\partial U_0 = 0.01$ m s^{-1}, RH = 0.75, $q_{10} = 13.6$ g kg^{-1}, $q_0 = 14.5$ g kg^{-1}, and $D = 1.3 \times 10^{-3}$. Together, these yield an annual mean bias in latent heat flux of $\partial L = 0.05L$. For typical conditions at the Stratus mooring, this is equal to 5 W m^{-2}. We note

TABLE 11. A summary of the accuracy of the annual average fluxes determined by the IMET buoy for the Stratus deployment.

	Net longwave	Net shortwave	Sensible heat	Latent heat	Momentum	Tot net
Percent error	10	2.5	15	5	10	20
Typical error	3.9 W m^{-2}	5 W m^{-2}	1.5 W m^{-2}	5 W m^{-2}	0.007 N m^{-2}	8 W m^{-2}

that the error is dominated by the uncertainty in the Dalton number, which is 5 times larger than the error resulting from biases in the air temperature and 3 times larger than uncertainty in the relative humidity. We have derived our assumed error for the Dalton number based on the TOGA COARE measurements (e.g., Fairall et al. 1996b, Fig. 2), where the overall discrepancy between the direct covariance and the COARE algorithm could be up to 0.07 but was less than 0.04 when averaged over typically observed wind speeds. One positive factor is that the Dalton number error is a postcorrectable bias in that the fluxes can be recalculated in the future by using the current best guess of the Dalton number. For instance, if the error was only $\partial D/D = 0.02$ in the future, then the fluxes would have an error of $\partial L = 0.04L$.

We do need to consider possible covariance between the relative humidity error (∂RH) and the air temperature error (∂T_{10}). Anderson and Baumgartner (1998, Fig. 7) note that the heating errors in collocated air temperature and relative humidity largely cancel in specific humidity. If we make the extreme assumption that air temperature errors are perfectly anticorrelated with relative humidity errors, then the bias in the annual mean latent heat is reduced to $\partial L = 0.04L$ (namely a reduction of 1 W m^{-2}). Assuming a more reasonable correlation of -0.7, gives a reduction of 0.5 W m^{-2} in the annual bias.

e. Net heat flux

The annual mean net heat flux bias can be determined from combining the individual terms assuming no covariance. With sensible, latent, longwave and shortwave biases given by 1.5, 4.5, 3.9, and 5.1 W m^{-2}, respectively, the net heat flux bias is 8.0 W m^{-2}. Although the details of the individual calculations are debatable, it seems unlikely that the annual mean bias exceeds 10 W m^{-2}. We have neglected the heat input resulting from rain, but this is not large at these sites.

f. Freshwater flux

The freshwater flux has two components: evaporation and precipitation. We do not have the data to support extensive discussion of precipitation error. Instead, we propose a rough error estimate of min(10%,10 cm) for the annual mean precipitation. Using the latent heat calculation, we can estimate that the evaporative freshwater flux is accurate to $\pm5\%$. At the Stratus site, this is equivalent to 6–7 cm year^{-1} of evaporation.

g. Momentum flux

Momentum flux has two components: magnitude and direction. The direction errors will be nearly identical to those for the wind speed direction, with minor uncertainty because of the underlying surface current direction. The magnitude depends on the relative wind speed and uncertainties in the drag coefficient. In principle, the drag coefficient is dependent on the atmospheric stability, so errors in air temperature and relative humidity can lead to errors in the drag coefficient. However, previously determined errors for air temperature and relative humidity (±0.1 K and 1%, respectively) induce an error of $<0.1\%$ in the wind stress. This is much less than the uncertainty in the drag coefficient, so we ignore these second-order errors:

$$\left(\frac{\partial \tau}{\tau}\right)^2 = \frac{\partial C_D^2}{C_D^2} + 4\frac{\partial U_{10}^2 + \partial U_0^2}{(U_{10} - U_0)^2}$$

We will assume the following values for the biases in the annual mean: $\partial C_D/C_D = 0.1$, $\partial U_{10}/U_{10} = \partial U_3/U_3 = 0.01$, $\partial U_0 = 0.01$ m s^{-1}, and $U_{10} - U_0 = U_{10} = 6$ m s^{-1}. This leads to $\partial \tau = 0.1\tau$, where the uncertainty arises almost entirely from the drag coefficient. In the Stratus dataset, this leads to a typical error in the wind stress of ±0.007 N m^{-2}. We have assumed a conservative accuracy on the drag coefficient ($\partial C_D = 0.1C_D$). A more optimistic assumption would lead to an almost linear improvement in the wind stress error.

5. Conclusions

The performance of the basic meteorological sensors used in the IMET system has been examined and their absolute accuracies have been determined though a combination of previous results, laboratory calibrations, and in situ instrument comparisons. The overall results are encouraging for use of the IMET system for climate research in the subtropics (see summary in Table 11). Knowledge of the error characteristics is essential for the proper analyses of the data, and publication of these results is seen as an essential accompaniment to release of the data. Furthermore, these results serve as a benchmark for future refinements of measurement capability.

It must be restated that the error characteristics summarized here are not necessarily applicable beyond the subtropics. There are a number of environmental conditions that would increase measurement errors and lead to the previously stated error of 8 W m^{-2} in the annual net heat flux being exceeded. Very high wind speeds and the related uncertainty in bulk formula could lead to large errors in latent heat flux. Additionally, high wind speeds produce salt spray and steep waves that degrade radiometer performance by increasing the buoy (or ship) motion. Conversely, very low winds, if persistent, would increase some errors. Extremes of temperature could also influence the measurements by causing currently unresolved and unnoticed temperature dependencies to become important. Furthermore, instrument performance in freezing and colder conditions is unknown. Finally, the solar zenith angle at higher latitudes can be large, even near noon. This increases the shortwave percentage error caused by radiometer movement and tilt. An additional module to observe and record orientation and movement would allow one to partially postcorrect for this movement.

The current work indicates several points where improvement could be made. As noted above, a dependable solid-state motion package should be installed on the buoys to monitor the movement. The package should distinguish wave accelerations from tilts, have an accuracy approaching 0.1°, and be capable of unattended operation for a year. An accurate knowledge of the buoy motion would enable better estimation of the motion-induced errors in shortwave (and to a lesser degree longwave) radiation. It might also allow a postcorrection for shortwave radiation following the results of MacWhorter and Weller (1991) and potentially reduce the conservative value of the uncertainty of annual mean incoming shortwave developed above by 1–2 W m^{-2}. In other locations where the sea state is often large, it may allow for postcorrection to within our given limits. The in situ–corrected longwave sensor is performing well, but continued attention to sensor electronics, sensor performance, and inspection of domes for leakage is warranted. The passage of shortwave radiation by the longwave domes needs to be clarified. Assuming that a longwave dome in good condition passes 1% of incoming shortwave radiation leads to an overestimation of longwave radiation by 2 W m^{-2} in the annual average. However, the instantaneous error could be up to 10 W m^{-2}, which is larger than desirable. A suite of calibrations with a large number of domes could help to better quantify the average transmittance of the longwave dome.

Air temperature errors are important, largely because of the small difference between sea surface and air temperatures. Consequently the air temperature error is more important in regions where the sensible heat flux is small. In regions of large flux, such as near the continental margins, the percentage error will be much smaller, although the magnitude of the error is probably similar. Low wind effects on the air temperature are well known and well researched. If the deployment is in a region where this is a likely problem, then the air temperature will need to be ventilated. Relative humidity errors have been surprisingly small (only contributing a 2% error to latent heat flux). Assuming a correlation with air temperature errors when calculating specific humidity leads to an even smaller error.

Some of the largest errors were due to uncertainties in the coefficients of the bulk formulas (Stanton and Dalton numbers). Ongoing work with combined turbulent and bulk flux measurements supported by wave measurements remains a need, especially in the low and high wind regimes that are not yet well sampled. An ongoing commitment is needed to field intercomparisons between different sensors on the same buoy and between the buoy measurements and those from ships spending time very near the buoy using time dedicated to the intercomparison task. The work done here depends on the existence of reliable in situ observations for comparison and sensor validation. The attended bulk and direct flux measurements made onboard the ship during the ship–buoy comparison period are crucial to the understanding of sensor performance and the nature of sensor degradation. This is particularly true for delicate sensors, such as relative humidity, which are sometimes damaged on recovery or during shipping and hence not available for postcalibration. Of course, care and attention must be devoted to the shipboard measurements; the reader is referred to the recent handbook on shipboard meteorological and flux measurements (Bradley and Fairall 2006).

The current version of the IMET sensor suite does a remarkable job of measuring the basic surface meteorology over the ocean from an unattended platform for periods of a year in the moderate conditions of the subtropics. Our estimate of the uncertainty in the annual average heat flux is ±8 W m^{-2}. It is also true that the error present in the diurnal average is not much greater than this annual average (except in cases of no wind and high insolation). One issue that we have not addressed as thoroughly is the relative error of the instruments. Although Fig. 9 presents some spectra, these are between two sensors on the same platform and thus do not address all the error sources. This would be a useful avenue of future exploration because it would allow for the placement of error bars not just on long-term means but also on features such as the diurnal cycle. Some of the work has been started in this paper, but a full examination of this would require a buoy

equipped with both an IMET sensor system and a set of direct-covariance measurements.

REFERENCES

Anderson, S. P., and M. F. Baumgartner, 1998: Radiative heating errors in naturally ventilated air temperature measurements made from buoys. *J. Atmos. Oceanic Technol.,* **15,** 157–173.

Barton, I. J., P. J. Minnett, K. A. Maillet, C. J. Donlon, S. J. Hook, A. T. Jessup, and T. J. Nightingale, 2004: The Miami 2001 infrared radiometer calibration and intercomparison: Part II: Shipboard results. *J. Atmos. Oceanic Technol.,* **21,** 268–283.

Bradley, F., and C. Fairall, 2006: A guide to making climate quality meteorological and flux measurements at sea. NOAA Tech. Memo. OAR PSD-311, 81 pp.

Burns, S. P., and Coauthors, 1999: Comparisons of aircraft, ship, and buoy meteorological measurements from TOGA COARE. *J. Geophys. Res.,* **104,** 30 853–30 883.

——, and Coauthors, 2000: Comparisons of aircraft, ship, and buoy radiation and SST measurements from TOGA COARE. *J. Geophys. Res.,* **105,** 15 627–15 652.

Bush, C. B., F. P. J. Valero, A. S. Simpson, and L. Bignone, 2000: Characterization of thermal effects in pyranometers: A data correction algorithm for improved measurement of surface insolation. *J. Atmos. Oceanic Technol.,* **17,** 165–175.

Fairall, C. W., E. F. Bradley, J. S. Godfrey, G. A. Wick, J. B. Edson, and G. S. Young, 1996a: Cool-skin and warm-layer effects on sea surface temperature. *J. Geophys. Res.,* **101,** 1295–1308.

——, ——, D. P. Rogers, J. B. Edson, and G. S. Young, 1996b: Bulk parameterization of air-sea fluxes for Tropical Ocean-Global Atmosphere Coupled-Ocean Atmosphere Response Experiment. *J. Geophys. Res.,* **101,** 3747–3764.

——, P. O. G. Persson, E. F. Bradley, R. E. Payne, and S. P. Anderson, 1998: A new look at calibration and use of Eppley precision infrared radiometers. Part I: Theory and application. *J. Atmos. Oceanic Technol.,* **15,** 1229–1242.

Folland, C. K., 1988: Numerical models of the raingauge exposure problem, field experiments and an improved collector design. *Quart. J. Roy. Meteor. Soc.,* **114,** 1485–1516.

Gill, G. C., 1976: Development and testing of a no-moving parts static pressure inlet for use on ocean buoys. NOAA Data Buoy Office Progress Rep. 01-6-038-115, 43 pp.

Hosom, D. S., R. A. Weller, R. E. Payne, and K. E. Prada, 1995: The IMET (Improved Meteorology) ship and buoy systems. *J. Atmos. Oceanic Technol.,* **12,** 527–540.

Hubbard, K. G., X. Lin, and E. A. Walter-Shea, 2001: The effectiveness of the ASOS, MMTS, Gill, and CRS air temperature radiation shields. *J. Atmos. Oceanic Technol.,* **18,** 851–864.

Jin, Z., T. P. Charlock, and K. Rutledge, 2002: Analysis of broadband solar radiation and albedo over the ocean surface at COVE. *J. Atmos. Oceanic Technol.,* **19,** 1585–1601.

Koschemeider, H., 1934: Methods and results of definite rain measurements. *Mon. Wea. Rev.,* **62,** 5–7.

Lin, X., K. G. Hubbard, E. A. Walter-Shea, J. R. Brandle, and G. E. Meyer, 2001: Some perspectives on recent in situ air temperature observations: Modeling the microclimate inside the radiation shields. *J. Atmos. Oceanic Technol.,* **18,** 1470–1484.

MacWhorter, M. A., and R. A. Weller, 1991: Error in measurements of incoming shortwave radiation made from ships and buoys. *J. Atmos. Oceanic Technol.,* **8,** 108–117.

Nešpor, V., and B. Sevruk, 1999: Estimation of wind-induced error of rainfall gauge measurements using a numerical simulation. *J. Atmos. Oceanic Technol.,* **16,** 450–464.

Pascal, R. W., and S. A. Josey, 2000: Accurate radiometric measurements of the atmospheric longwave flux at the sea surface. *J. Atmos. Oceanic Technol.,* **17,** 1271–1282.

Payne, R. E., and S. P. Anderson, 1999: A new look at calibration and use of Eppley precision infrared radiometers. Part II: Calibration and use of the Woods Hole Oceanographic Institution improved meteorology precision infrared radiometer. *J. Atmos. Oceanic Technol.,* **16,** 739–751.

Philipona, R., C. Frohlich, and C. Betz, 1995: Characterization of pyrgeometers and the accuracy of atmospheric long-wave radiation measurements. *Appl. Opt.,* **34,** 1598–1605.

——, and Coauthors, 2001: Atmospheric longwave irradiance uncertainty: Pyrgeometers compared to an absolute sky-scanning radiometer, atmospheric emitted radiance interferometer, and radiative transfer model calculations. *J. Geophys. Res.,* **106,** 28 129–28 141.

Taylor, P. K., Ed., 2000: Intercomparison and validation of ocean-atmosphere energy flux fields—Final report of the Joint WCRP/SCOR Working Group on Air-Sea Fluxes. WCRP-112, WMO/TD-No. 1036, 306 pp.

Weller, R. A., F. Bradley, and R. Lukas, 2004: The interface or air–sea component of the TOGA Coupled Ocean-Atmosphere Response Experiment and its impact on subsequent air–sea interaction studies. *J. Atmos. Oceanic Technol.,* **21,** 223–257.

Yelland, M. J., B. I. Moat, R. W. Pascal, and D. I. Berry, 2002: CFD model estimates of the airflow distortion over research ships and the impact on momentum flux measurements. *J. Atmos. Oceanic Technol.,* **19,** 1477–1499.

Zeng, L., and R. A. Brown, 1998: Scatterometer observations at high wind speeds. *J. Appl. Meteor.,* **37,** 1412–1420.