A Comparison of Surface Layer and Surface Turbulent Flux Observations over the Labrador Sea with ECMWF Analyses and NCEP Reanalyses

IAN A. RENFREW

British Antarctic Survey, Cambridge, United Kingdom

G. W. K. MOORE

Department of Physics, University of Toronto, Toronto, Ontario, Canada

PETER S. GUEST

Naval Postgraduate School, Monterey, California

KARL BUMKE

Institut für Meereskunde, Universität Kiel, Kiel, Germany

ABSTRACT

Comparisons are made between a time series of meteorological surface layer observational data taken on board the R/V *Knorr,* and model analysis data from the European Centre for Medium-Range Weather Forecasting (ECMWF) and the National Centers for Environmental Prediction (NCEP). The observational data were gathered during a winter cruise of the R/V *Knorr,* from 6 February to 13 March 1997, as part of the Labrador Sea Deep Convection Experiment. The surface layer observations generally compare well with both model representations of the wintertime atmosphere. The biases that exist are mainly related to discrepancies in the sea surface temperature or the relative humidity of the analyses.

The surface layer observations are used to generate bulk estimates of the surface momentum flux, and the surface sensible and latent heat fluxes. These are then compared with the model-generated turbulent surface fluxes. The ECMWF surface sensible and latent heat flux time series compare reasonably well, with overestimates of only 13% and 10%, respectively. In contrast, the NCEP model overestimates the bulk fluxes by 51% and 27%, respectively. The differences between the bulk estimates and those of the two models are due to different surface heat flux algorithms. It is shown that the roughness length formula used in the NCEP reanalysis project is inappropriate for moderate to high wind speeds. Its failings are acute for situations of large air–sea temperature difference and high wind speed, that is, for areas of high sensible heat fluxes such as the Labrador Sea, the Norwegian Sea, the Gulf Stream, and the Kuroshio. The new operational NCEP bulk algorithm is found to be more appropriate for such areas.

It is concluded that surface turbulent flux fields from the ECMWF are within the bounds of observational uncertainty and therefore suitable for driving ocean models. This is in contrast to the surface flux fields from the NCEP reanalysis project, where the application of a more suitable algorithm to the model surface-layer meteorological data is recommended.

1. Introduction

Deep convection in the ocean is a key component of the thermohaline circulation, governing both the location and timescale of vertical mixing processes at high latitudes. The occurrence of deep convection is a multistage process, requiring as a first step that the ocean be preconditioned into a state of neutral stratification.

For the open ocean, this occurs through the tilting of isopycnals to maintain thermal wind balance in the presence of a boundary current or through topographic forcing, either of which must be combined with cooling and evaporation into the lower atmosphere. In the second step, further losses of heat and moisture by the surface waters lead to a destabilization of the water column and the triggering of convective plumes within the preconditioned region. The third step is the sinking and spreading of the modified water mass and an eventual shutdown of convection (e.g., Killworth 1983; Marshall and Schott 1999). Deep convection in the open ocean is a

Corresponding author address: Ian Renfrew, British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, United Kingdom.
E-mail: i.renfrew@bas.ac.uk

process that is difficult to observe, due to its localized occurrence and rapid timescale, yet plays a crucial role in the thermohaline circulation and hence the climate system. To redress this situation a major international project has been initiated: The Labrador Sea Deep Convection Experiment (Lab Sea Group 1998). Its aim is "improving our understanding of the convective process in the ocean and its representation in models," through oceanographic and meteorological observations, theory, and modeling. The Labrador Sea was chosen due to its relative proximity to North America, compared to the other major convection sites of the Greenland–Iceland–Norwegian seas and the Weddell Sea. The choice was made additionally interesting by the somewhat ephemeral nature of Labrador Sea convection; there is evidence for significant decadal variability in the amount of convection that takes place (Dickson et al. 1996). Indeed it is the one region where convection shut down in a "realistic" coupled atmosphere–ocean climate modeling study carried out recently (Wood et al. 1999).

Modeling the ocean circulation in a realistic way relies on the faithful representation of atmosphere–ocean interactions. To be precise, the fluxes of momentum, heat, and moisture between the atmosphere and ocean must be accurate. Nowhere is this more critical than in convectively active regions where, as discussed above, the atmosphere directly forces the ocean dynamics. To generate accurate atmosphere–ocean fluxes requires first that the atmospheric boundary layer is realistic and second that the derivation of surface fluxes is carried out using a suitable model formulation. In this study we address these two requirements in turn. First, a time series of observations from a cruise of the R/V *Knorr* in the Labrador Sea during February and March 1997 is compared to model analyses data. The analyses are the operational analyses from the European Centre for Medium-Range Weather Forecasts (ECMWF), and the reanalyses from the National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) project. These are two of the leading numerical weather prediction centers in the world and two of the most common choices for providing surface flux fields to drive ocean models. In section 2 the observational and model data are described in detail. Section 3 details a surface layer meteorological data comparison, and following this, section 4 details a surface turbulent flux comparison. The differences in the surface heat flux fields warrant an explanation, and this leads to an investigation of surface flux algorithms in section 5. Conclusions are drawn in section 6.

2. Datasets

a. Observational data

A 40-day cruise in the Labrador Sea, from 6 February to 13 March 1997, was undertaken by the R/V *Knorr* as part of the field component of the Labrador Sea Ex-
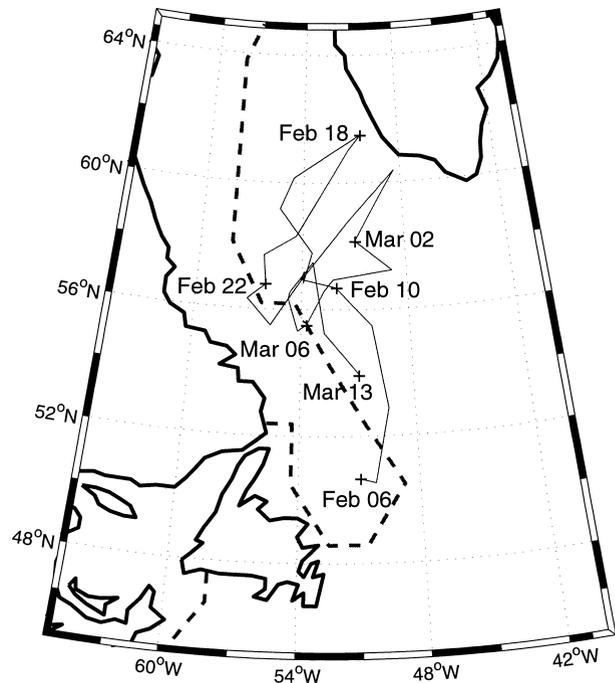

FIG. 1. The track of the R/V *Knorr* during the Feb and Mar 1997 Labrador Sea Experiment cruise. Overlain as a dashed line is the ice cover used by the NCEP model from 18 Feb 1997.

periment (Fig. 1). In companion to the comprehensive oceanographic field work that took place (e.g., Lab Sea Group 1998), a number of meteorological experiments were on board the *Knorr*. For example, a group from the Bedford Institute of Oceanography carried out turbulent heat and moisture flux measurements, and a group from the University of Kiel carried out turbulent heat flux and precipitation measurements (Bumke et al. 2002; 2001, manuscript submitted to *J. Phys. Oceanogr.*). In addition, a team of scientists released rawinsondes, typically every 6 hours and occasionally every 3 hours, and these data were transmitted via the Global Telecommunications System (GTS) to national forecast centers.

Throughout the cruise standard meteorological variables were automatically logged by the ship's Improved Meteorological (IMET) measurement system. These form the main dataset used in this study. Temperature, humidity, pressure, and wind velocity were measured using a thermistor, hygrometer, barometer, and propeller anemometer located on the foremast approximately 19.5 m above the sea surface. The ship's motion, as determined by a GPS navigation system, was subtracted from the anemometer-measured wind vector to obtain a true wind vector. Usually the ship was steaming between stations or performing oceanographic CTD measurements. In either case, the wind was usually from the bow and the ship's velocity vector was constant. Based on sensor comparisons and a numerical wind tunnel test (Moat and Yelland 1998) we estimate that when the

TABLE 1. Summary of surface layer meteorological data from the R/V *Knorr,* from the ECMWF, and from the NCEP as indicated.

	Slp (mb)	T_a (°C)	SST (°C)	U_{10} (m s^{-1})	q_a (g kg^{-1})	RH (%)	
Mean	996.4	−6.23	2.83	11.5	1.92	76.5	*Knorr*
	998.0	−7.54	1.15	11.6	1.63	71.5	ECMWF
	997.0	−6.19	1.49	11.9	2.30	93.4	NCEP
Standard deviation	10.6	3.65	1.06	3.6	0.77	8.2	*Knorr*
	10.3	3.90	1.29	3.8	0.71	8.5	ECMWF
	10.3	3.42	1.35	4.1	0.67	7.9	NCEP
Maximum	1019.0	3.70	4.56	20.1	4.85	96.9	*Knorr*
	1018.5	3.10	3.57	20.9	4.41	94.5	ECMWF
	1017.6	2.78	4.04	19.4	4.55	107.6	NCEP
Minimum	975.5	−13.39	−1.67	1.7	1.03	56.7	*Knorr*
	976.8	−16.14	−1.33	2.3	0.81	47.6	ECMWF
	975.9	−14.05	−1.71	0.4	1.28	63.9	NCEP

relative wind direction was within 20 degrees of the bow, the flow distortion effects were less than a few percent. There were times the wind was from the stern, and flow distortion effects may have been as large as 20%. The individual GPS measurements were despiked and averaged over 5 minutes. The resulting data were then manually edited to remove any remaining suspect data. The temperature, relative humidity, and wind vector sensors were checked every day and cleaned of snow and ice when required. Although sea-spray icing at lower levels was severe, it was not a problem for the meteorological sensors. Due to the generally high winds, ventilation of the temperature sensors was excellent. Heated air from the ship engine exhaust was almost always blown well clear of the thermometers, and during the few occasions that stack gas was coming in the direction of sensors, most of the heated air passed below the thermometers.

In this study, flow distortion around the *Knorr* was taken into account by a reduction in the wind measurement height by 1.3 m, following the report of Moat and Yelland (1998). The *Knorr* data have been interpolated down from the instrument heights (approximately 19.5

m) to the standard meteorological heights of 2 m and 10 m, using well-established surface layer similarity theory (i.e., Dyer 1974 or Smith 1988). Sea surface temperature (SST) was measured from an intake system 2 m below the surface. Due to the generally high winds and relatively small solar radiation, the intake sea temperature was representative of the surface skin temperature to within 0.5°C. Longwave and shortwave radiative fluxes were also measured.

The IMET data were sampled every second, with 15-s averages stored. For analysis purposes, records of 5-min averages were generated. For comparison with the model analyses, 6-hourly "instant" and "average" time series have been created. The 5-min data are used to create the instant time series. The same data are averaged into 6-h periods, starting at the reference time, to create the average time series. These data are used to compare with the surface flux fields from the ECMWF and NCEP, which are also 6-h averages accumulated over a model forecast. Note that this averaging may entail an aliasing in space if the R/V *Knorr* moved during this period. However, this does not greatly affect

TABLE 2. Comparison of surface layer meteorological data, comparing between the R/V *Knorr* and the ECMWF datasets, and the R/V *Knorr* and the NCEP datasets as indicated.

	Slp (mb)	T_a (°C)	SST (°C)	U_{10} (m s^{-1})	q_a (g kg^{-1})	RH (%)	*Knorr* vs
Correlation coefficient	0.99	0.95	0.19	0.79	0.96	0.64	ECMWF
	0.99	0.87	0.30	0.75	0.89	0.42	NCEP
Slope	0.97	1.02	0.23	0.82	0.89	0.66	ECMWF
	0.97	0.85	0.38	0.84	0.78	0.41	NCEP
Bias error	1.6	−1.31	−1.68	0.1	−0.29	−5.0	ECMWF
	0.6	0.04	−1.34	0.4	0.38	16.8	NCEP
Slope error	10.2	3.70	0.24	2.9	0.68	5.4	ECMWF
	10.2	2.96	0.40	3.0	0.59	3.3	NCEP
Random error	10.6	3.82	1.64	4.3	0.79	10.4	ECMWF
	10.6	4.01	1.66	4.5	0.82	10.8	NCEP
Total rms error	14.8	5.48	2.37	5.2	1.08	12.7	ECMWF
	14.7	4.98	2.17	5.4	1.09	20.3	NCEP

TABLE 3. Summary of surface turbulent flux data from the R/V *Knorr,* from the ECMWF, and from the NCEP as indicated. The R/V *Knorr* "observed" data are calculated from bulk algorithms, following Smith (1988) and DeCosmo et al. (1996), as described in the text.

	Sensible heat flux (W m^{-2})	Latent heat flux (W m^{-2})	Momentum flux (N m^{-2})	
Mean	183.8	136.2	0.28	*Knorr*
	208.0	150.4	0.37	ECMWF
	277.9	172.4	0.38	NCEP
Standard deviation	100.0	56.5	0.16	*Knorr*
	129.8	74.3	0.24	ECMWF
	177.2	89.4	0.25	NCEP
Maximum	406.5	243.5	0.72	*Knorr*
	479.5	287.3	1.11	ECMWF
	772.6	359.2	1.02	NCEP
Minimum	−43.7	−8.9	0.01	*Knorr*
	−56.1	−22.2	0.01	ECMWF
	−54.7	−29.3	0.01	NCEP

our results, as the ship moved a small distance (over any 6-h period) relative to the size of the model grid.

b. Model data

Two sets of model data are to be compared with the R/V *Knorr* data: the ECMWF operational analyses and the NCEP–NCAR reanalyses. These are both global analyses from state of the art atmospheric modeling and data assimilation systems, and so probably represent the global atmosphere as best we know it. A distinction between the datasets is that the ECMWF analyses are their operational products, whereas the NCEP products are from their reanalyses project (Kalnay et al. 1996). The idea behind the reanalyses projects was to generate a model-consistent dataset that would be ideal for short-term climatological studies. To this end, the operational atmospheric model and analysis schemes were "frozen" and a comprehensive reanalysis of all available data was carried out. In the NCEP–NCAR project the period of reanalysis was 1957 to the present, although clearly the

analysis quality will not be as good in the presatellite period (e.g., Kalnay et al. 1996). In contrast, the first ECMWF ReAnalysis (ERA-15) project was limited to a more recent period, 1979–93, so it stops short of the Labrador Sea Experimental period. The ECMWF model and analysis scheme for this period (Feb–Mar 1997) are very similar to that used during the ERA-15 project; indeed, to the best of our knowledge the boundary layer schemes are identical. The ECMWF operational analyses have the advantage of a resolution of T213 [about 100-km horizontal resolution; see ECMWF (1995) for more details]. The model used in the NCEP reanalysis was identical to the NCEP operational model that was active at the start of the reanalysis (11 Jan 1995), except at a lower resolution of T62 [about 210-km horizontal resolution; see Kalnay et al. (1996) for more details]. The model data have been extracted from the global analyses every 6 h, from 0600 UTC 6 February 1997 to 1800 UTC 13 March 1997, and so consist of time series of 143 data points. For each analysis time, the model data are extracted through a bilinear interpolation

TABLE 4. Comparison of surface turbulent flux data, comparing between the R/V *Knorr* and the ECMWF datasets, and the R/V *Knorr* and the NCEP datasets as indicated. The R/V *Knorr* "observed" data are calculated from bulk algorithms, following Smith (1988) and DeCosmo et al. (1996), as described in the text. The bulk flux comparisons in columns 4 and 5 employ the same algorithms to both the *Knorr* and model surface layer meteorological data.

	Sensible heat flux (W m^{-2})	Latent heat flux (W m^{-2})	Momentum flux (N m^{-2})	Bulk sensible flux (W m^{-2})	Bulk latent flux (W m^{-2})	*Knorr* vs
Correlation coefficient	0.93	0.90	0.82	0.91	0.87	ECMWF
	0.91	0.91	0.81	0.87	0.86	NCEP
Slope	1.21	1.18	1.27	0.99	0.90	ECMWF
	1.61	1.42	1.27	0.83	0.77	NCEP
Bias error	24.2	14.2	0.10	0.8	−7.5	ECMWF
	94.1	36.2	0.10	−18.2	−34.0	NCEP
Slope error	120.1	66.5	0.20	98.5	50.5	ECMWF
	160.0	80.6	0.20	82.1	43.3	NCEP
Random error	110.4	64.9	0.21	109.2	62.9	ECMWF
	124.4	67.8	0.21	110.1	61.9	NCEP
Total rms error	164.9	94.0	0.31	147.0	81.0	ECMWF
	223.4	111.4	0.31	138.6	82.9	NCEP

to the exact position of the *Knorr* using the surrounding four grid points. Thus, we have a time series extracted from the model analyses following the track of the *Knorr*.

It is perhaps pertinent at this point to note what observational data are assimilated by the two models. Modern numerical weather prediction systems typically carry out three-dimensional variational analyses of mass and wind fields, with these fields "balanced" to reduce spurious gravity waves. Separate analyses of temperature and humidity are also carried out. The greatest weights in the analysis procedure are given to upper air observations. In general, upper air observations of geopotential, wind, temperature, and humidity are used in the analyses. However, the ECMWF does not use temperature directly; instead the hydrostatic equation is solved, and the temperature data are used as a check. Surface data are also used in the mass and wind analysis, although it is generally only surface pressure data that are used. In addition, at the ECMWF surface winds over the ocean are generally assimilated, and in 1997 surface humidity observations were used in the humidity analysis (ECMWF 1995; P. Viterbo 2000, personal communication).

Upper air data from the rawinsondes released from the R/V *Knorr* were entered onto the GTS twice daily at 0130 and 1330 UTC, under call sign KCEJ. Typically four sondes per day were released. At the ECMWF 115 geopotential soundings and 112 wind soundings made it into the data assimilation system (F. Lalaurette 1999, personal communication), which is close to the expected four times daily for around 40 days. At the NCEP, "most" of the soundings made it into the data assimilation system (R. Kistler 1999, personal communication). Surface data from the *Knorr* were not entered onto the GTS, so they were not available for the model analyses.

It should be noted that the model surface layer variables used here are calculated by an interpolation between the lowest model level (around 30 m for the ECMWF model and 50 m for the NCEP model) and the surface, using a stability-dependent surface layer scheme of the same ilk as that used for the ship data. In this case, the ECMWF surface layer data are calculated from the analysis fields, whereas the NCEP surface layer data are calculated from 6-h forecasts (ECMWF 1995; Kalnay et al. 1996).

c. Sea surface temperatures and sea ice

A prescribed SST field is used as the lower boundary condition in the models and so SST strongly influences the 2-m temperatures, which are interpolated between the lowest model level and the surface. The SST field is determined through a mix of in situ measurements and infrared satellite data (e.g., Reynolds and Smith 1994). Unfortunately, remote areas like the Labrador Sea suffer through both a lack of ship data and the

tendency for prolonged cloud cover to reduce the amount of satellite data available. Where data are sparse the SST field is determined by a regression from the sea ice edge to the nearest available observation.

The model's sea ice masks are determined from passive microwave satellite data available routinely from the Special Sensor Microwave/Imager (SSM/I) instrument. The data are available as a sea ice concentration at a resolution of 25 km. However, at present both forecast centers only use a 0% or 100% flag for the ice cover and it has to be downgraded to the resolution of the model grid (e.g., ECMWF 1995; Kalnay et al. 1996). Figure 1 shows the track of the R/V *Knorr* in the Labrador Sea with the sea ice cover from the NCEP reanalysis on 18 February 1997 overlaid. A serious problem is clear as the *Knorr* is over model sea ice on several occasions during the cruise (the ECMWF model had the same problem). In reality, of course, the ship was in the marginal ice zone (MIZ), close to what one could define as an "ice edge." The mismatch in real and model sea ice cover meant that the model surface temperatures were too cold at these locations and hence the 2-m temperatures were also too cold. For the data comparison described in the following sections (i.e., the scatterplots and Tables 1–4), it was decided to neglect these data and compare only open ocean data. To this end, we ran a simple quality control check so that only model data where both models had SST values greater than $-1.8°C$ were used in the comparison. Below this temperature, the models deem the gridpoint ice covered. Due to our interpolation between grid points this effectively created an extremely simple MIZ. This reduces the datasets from 143 to 120 data points.

3. Surface layer data comparison

Figure 2 shows time series of R/V *Knorr*, ECMWF, and NCEP surface layer data every 6 hours, from 0600 UTC 6 February 1997 to 1800 UTC 13 March 1997. The panels show sea level pressure (slp), air temperature at 2 m (T_a), sea surface temperature (SST), wind speed at 10 m (U_{10}), specific humidity at 2 m (q_a), and relative humidity (RH) at 2 m. It is clear that the model analyses generally capture the magnitude and variation of the observational data. In particular, they accurately reproduce the broad-scale high and low slp readings that are associated with the passage of synoptic-scale weather systems. Hence, the models are generally accurate in reproducing the concomitant synoptic-scale variability in temperature and wind speed. There is a systematic difference in the sea surface temperatures, with several large temperature differences, where the ship is "over" model sea ice, as discussed above. Note that there are coincident large differences in T_a. In general, the model SSTs are too cold, and we suggest that this is due to the interpolation between the sea ice edge and nearest available observations. When the nearest observations are distant from the sea ice edge, the interpolation will
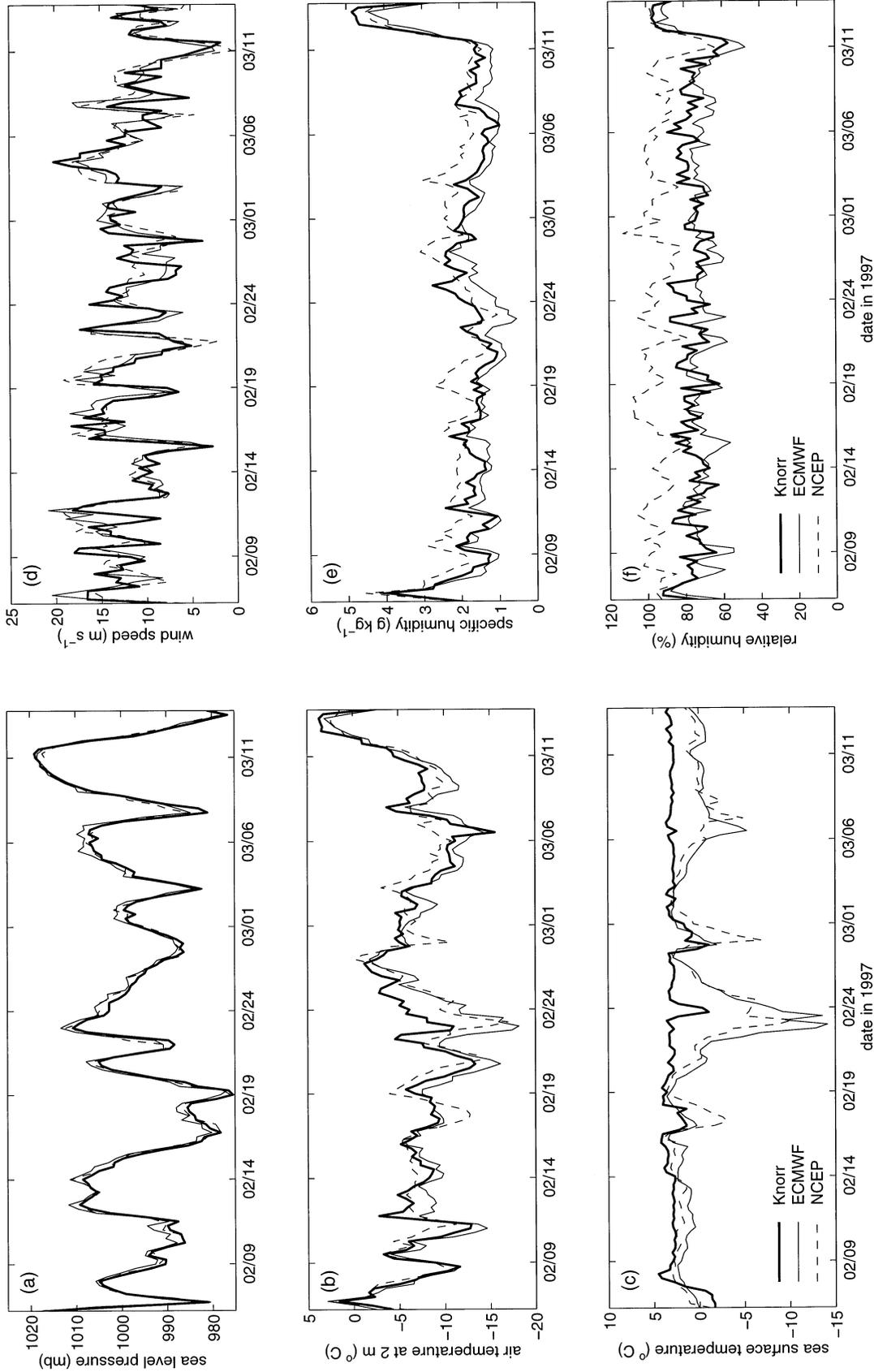
FIG. 2. Time series of *Knorr* observed data, ECMWF analysis data, and NCEP reanalysis data, from 6 Feb to 13 Mar 1997, with the data plotted as thick, thin, and dashed lines, respectively. (a) Sea level pressure (slp); (b) air temperature at 2 m (T_a); (c) sea surface temperature (SST); (d) wind speed at 10 m (U_{10}); (e) specific humidity at 2 m (q_a); and (f) relative humidity at 2 m (RH).

blur the SST gradient over that distance when, in reality, the gradient is strong in the immediate vicinity of the sea ice edge, as is evident from Fig. 2c. There is a large difference between the observed and NCEP model relative humidity, with the model often showing supersaturation (Fig. 2f).

A comprehensive comparison of the observed and modeled surface layer data is summarized in Tables 1 and 2, and in scatterplots of T_a, q_a, and U_{10} in Fig. 3. Note that this comparison uses the reduced dataset of 120 points as described in section 2c. The scatterplots show *Knorr* versus ECMWF or *Knorr* versus NCEP data as indicated. A linear regression line is shown, where the *Knorr* observations have been treated as the independent variable and the model analyses as the dependent variable. Table 2 summarizes some comparison statistics. The correlation coefficient (r) and the slope of the linear regression line indicate how well the data pairs match in a linear sense. The bias error quantifies any systematic model error. The slope error quantifies the departure from a linear relationship. The random error quantifies the random scatter in the comparison. The total error is the square root of the sum of squares of the component errors, which is also equal to the root-mean-square (rms) error.

Examining Tables 1 and 2 along with Figs. 2 and 3, it is clear that the slp is well modeled: $r = 0.99$ and the slope is 0.97 for both models. The model-analyzed slp both have small positive biases of 1.6 and 0.6 mb for ECMWF and NCEP, respectively. The accuracy of the analyzed sea level pressure is due to the model assimilation of the *Knorr* upper air rawinsonde data, as well as the inherent predictability of the pressure field; that is, there is less mesoscale and microscale variability in the pressure field.

Turning to the 2-m air temperature, the mean ECMWF temperature has a cold bias of $-1.31°C$, but a high correlation coefficient of 0.95 and a regression slope of 1.02. The mean NCEP temperature compares well, as the bias is only $0.04°C$, but the correlation coefficient is lower at 0.87, and the regression slope is 0.85. A possible explanation for the ECMWF's cold bias is the consistently cold sea surface temperatures (Fig. 2, Tables 1 and 2). Recall that T_a is interpolated between the lowest model level and the model surface. Comparing the sea surface temperatures, both the models are consistently too cold, with biases of $-1.68°$ and $-1.34°C$. We believe that this problem stems from the interpolation between satellite-determined SSTs, of which there are precious few in the Labrador Sea, to the sea ice edge, as discussed above. The mean SSTs are $2.83°$, $1.15°$, and $1.49°C$ for the *Knorr,* ECMWF, and NCEP data, respectively (Table 1). A second problem with the model's lower boundary is the limit of the 0% or 100% ice cover. This creates an unphysical situation, as boundary layer air in the model will cross a discontinuity in surface temperature, whereas in reality MIZs will be 10–100 km in width and heterogeneous,

with ice cover ranging between 1/10 and 10/10 on the scale of kilometers. The inclusion of an MIZ in a regional-scale model study of a cold-air outbreak produced a more realistic downstream boundary layer structure and more accurate surface heat fluxes when compared to observations (Pagowksi and Moore 2001; Renfrew and Moore 1999). The differences in the modeled boundary layer temperatures (with and without an MIZ) were 2–3 K, both over the MIZ and downstream for over 300 km. The differences in sensible and latent heat fluxes were 50–100 W m^{-2} in the MIZ and up to 50 W m^{-2} downstream.

Perhaps not surprisingly the wind speed plots show considerably more scatter than the other variables, although the mean and standard deviations compare well (Fig. 3, Tables 1 and 2). The ECMWF correlation and slope are respectively 0.79 and 0.82, and the NCEP correlation and slope are 0.75 and 0.84. The bias errors are both under 0.5 m s^{-1}. A similar low value of regression slope was also found in a comparison of ECMWF winds with buoy and ship data in Weller et al. (1998) and Roemmich et al. (2001). There seems to be increasing evidence of a systematic model underestimation of high wind speeds over the ocean. The relatively large scatter in the wind speed comparisons is due to the temporal and spatial variability of the wind field. This is illustrated by Renfrew et al. (1999), who show a plot of 5-min wind speed averages from the stationary R/V *Knorr,* along with regional model forecasts for the same period. The variations in wind speed are typically 1–3 m s^{-1} from one observation to the next, while the model wind speeds vary smoothly from one hour to the next. Due to the relative slowness of the R/V *Knorr* compared to meteorological timescales, it is not possible to evaluate the spatial variability using the ship data. Instead we refer to the aircraft observations of Renfrew and Moore (1999), which show considerable mesoscale variability, over a 380-km dropwindsonde cross section, and typically 1–4 m s^{-1} variations over hundreds of meters in their flight-level wind observations.

Focusing on the water vapor content and comparing the specific humidities, the ECMWF correlation and slope are 0.96 and 0.89, while for NCEP they are 0.89 and 0.78. There is a dry bias in the ECMWF data of -0.29 g kg^{-1} and a moist bias in the NCEP data of 0.38 g kg^{-1}. The ECMWF dry bias is comparable to the -0.2 g kg^{-1} that one would expect from the cold bias in T_a. The NCEP moist bias is due to the model's excessive relative humidity, rather than any T_a bias; indeed, the NCEP model is supersaturated on several occasions (Fig. 2f). In contrast, the ECMWF model performs well in terms of relative humidity.

The overestimate of relative humidity by the NCEP model is a cause for concern. There are several possible reasons: for example, too much water vapor transport into the region, too great a moisture flux out of the sea, or too little condensation and precipitation of water va-
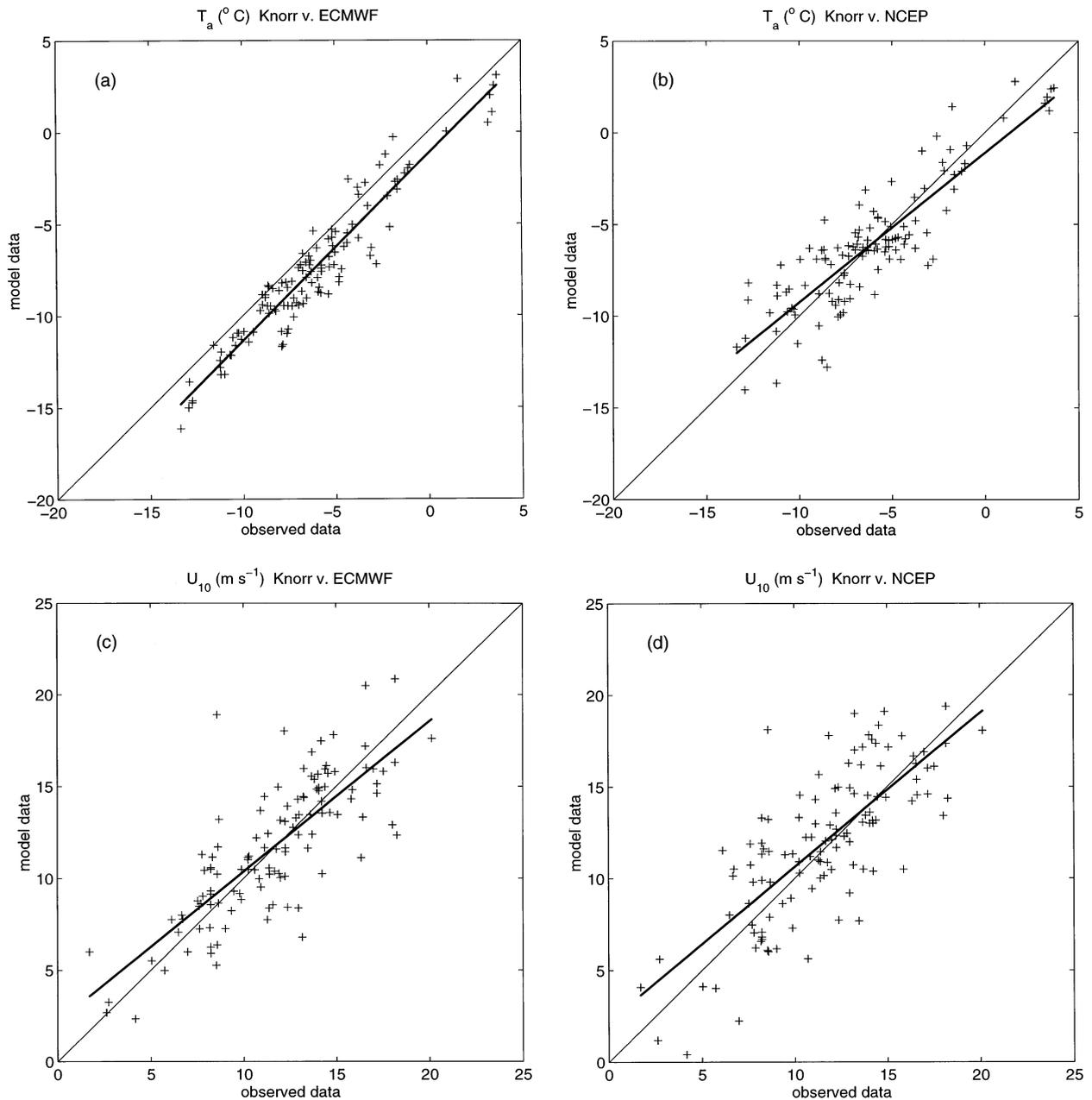
FIG. 3. Scatterplots of *Knorr* vs ECMWF and *Knorr* vs NCEP data as indicated: (a) and (b) compare T_a, while (c) and (d) compare U_{10}, and (e) and (f) compare q_a. A linear regression line is shown, where the observed data are assumed to be independent and the model data dependent.

por in the boundary layer. A thorough investigation the model's hydrological cycle is beyond the bounds of this study; however, we will briefly discuss some ideas. The region was well bounded by upper air stations, with an enhanced radiosonde release schedule due to the Fronts and Atlantic Storms Experiment (FASTEX) experiment. Therefore one would expect that the regional water vapor budget would be reasonably well constrained. Direct measurements of the moisture flux are not yet available from the *Knorr*. Comparing bulk latent heat flux esti-

mates and model latent heat fluxes (section 4) suggests that both models may be overestimating the flux of moisture out of the ocean; however a higher relative humidity will also act to reduce moisture fluxes, so the cumulative result is not clear. The cruise was dominated by cold-air outbreaks sweeping cold air across the relatively warm Labrador Sea and leading to shallow mesoscale convection in the form of cloud streets and mesoscale convective cells (Lab Sea Group 1998; Renfrew and Moore 1999). The shallow convection lead to snow
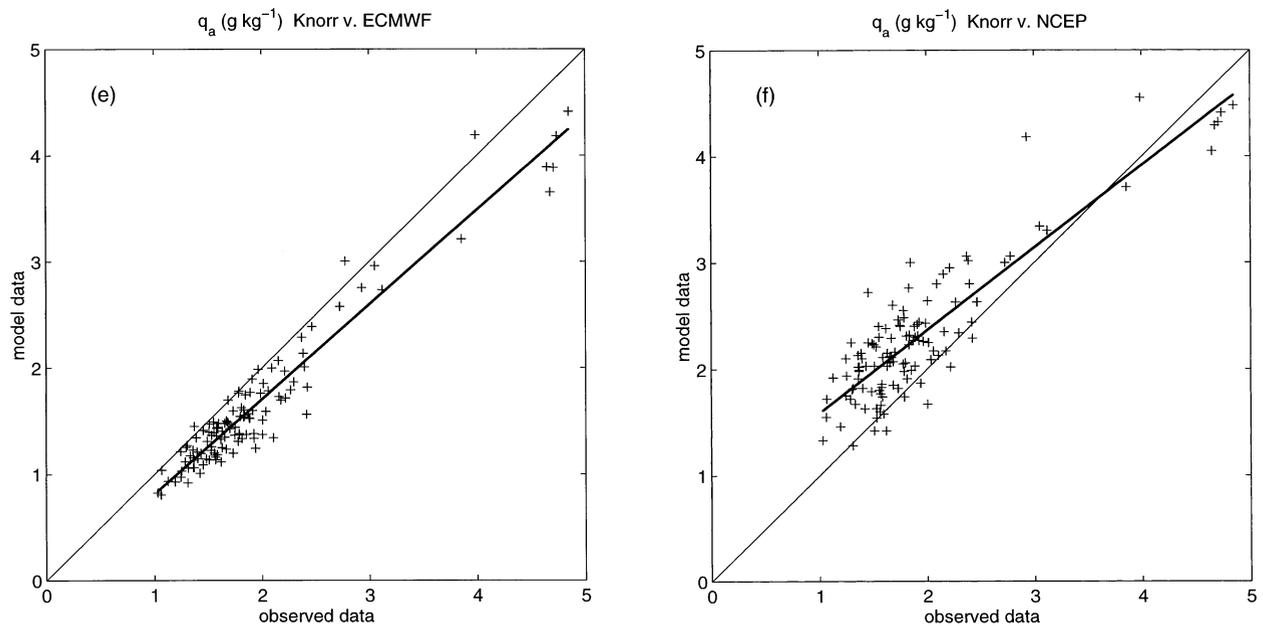
FIG. 3. (*Continued*)

virtually every day of the cruise, with an observed daily average of 3.2 mm and a total accumulation of 115 mm (Bumke et al. 2001, manuscript submitted to *J. Phys. Oceanogr.*). The NCEP model accumulation was 100 mm, an underestimate by around 15%, indicating that perhaps the models shallow-convection parameterization was underactive. The tops of the mesoscale (mostly liquid) clouds were typically −20 to −30°C, a temperature generally associated with the transition of supercooled water droplets to ice crystals. The ice crystals, once formed, will ''steal'' water vapor as they fall through the cloud, as the saturation vapor pressure with respect to ice is lower than with respect to liquid water. If this Bergeron–Findiesen process is not parameterized correctly, or does not occur frequently enough in the model, then it could explain any model overestimations in RH. In the cycle 48 version of the ECMWF model (active from September 1993) changes were made to decrease the moisture exchange coefficients to more closely fit experimental data. At the same time, it was realized that the model troposphere was too moist over the North Atlantic, and it was deduced that the model convection scheme was not active enough (e.g., Beljaars 1994). To remedy this, the closure scheme for shallow convection was modified to include a contribution from the surface sensible heat flux. This largely had the desired effect of drying out the boundary layer in shallow convective cold-air outbreak conditions (Beljaars 1994). This modification probably explains the ECMWF's generally good representation of the observed relative humidity time series. We suggest that the high relative humidity in the NCEP data may be caused by a similar problem and warrants further investigation.

In summary, the atmospheric surface layer in both models corresponds reasonably well with that observed from the R/V *Knorr*. The largest discrepancies were for sea surface temperatures (which influenced the 2-m air temperatures to some extent) and were due to the sparsity of satellite-determined SSTs and the crude sea ice mask employed in both models. The relative humidity in the NCEP model was consistently too high by around 15%–20%, although due to the cold temperatures the differences in RH do not equate to large differences in specific humidity. There was large scatter and a low regression slope in the 10-m wind speeds but no significant biases. In general, the ECMWF model compared better with observations than the NCEP model, perhaps partly due to its higher horizontal resolution, as well as the recently modified shallow convection scheme.

4. Surface turbulent fluxes

In terms of atmospheric forcing of the ocean, one is primarily interested in the surface fluxes of heat, moisture, and momentum at the atmosphere–ocean boundary, although it should be noted that radiative fluxes are also important in the surface energy budget. Turbulent surface heat and momentum fluxes were measured on board the ship by individuals from the Bedford Institute of Oceanography and the Institut für Meereskunde Kiel. However, these turbulent flux measurements are not available for the full length of the cruise and are also subject to strict quality control procedures, which means that a direct comparison between the turbulent flux data and model output is not straightforward. Instead, we have used the turbulent flux measurements to validate a well-established bulk flux algorithm, which we then use on the surface layer IMET data to generate a time
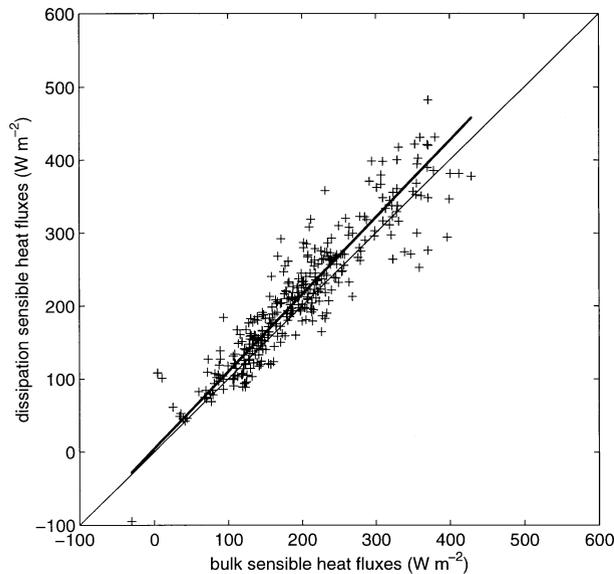
FIG. 4. A scatterplot comparing bulk sensible heat fluxes with turbulent sensible heat fluxes measured on board the R/V *Knorr* using the dissipation method. The bulk fluxes are calculated following the algorithm of Smith (1988) with the neutral exchange coefficients of DeCosmo et al. (1996). A linear regression line is overlain.

series of "observed" surface fluxes for comparison with the model fluxes. The bulk flux algorithm used is essentially that of Smith (1988), with the neutral exchange coefficients updated to those resulting from the Humidity Exchange over the Sea (HEXOS) experiments in the North Sea (DeCosmo et al. 1996). The formulation uses standard Businger–Dyer relations to correct for stability. For more details see section 5 and the appendix.

Figure 4 shows a validation of the bulk sensible heat fluxes against turbulent sensible heat fluxes calculated using the dissipation method. Details of the dissipation measurements and a comparison between these and eddy-correlation measurements can be found in Bumke et al. (2002). There is generally a good agreement between the data: the correlation coefficient is 0.92 and the linear regression slope is 1.06. There is a bias of -15 W m^{-2}, that is, the dissipation fluxes are on average a little higher than the bulk fluxes. This validation is about as good as one can expect, given the inherent inadequacies in the bulk flux method, and it justifies our choice of bulk algorithm. Note that Bumke et al. (2002) show that the algorithms of Anderson and Smith (1981) and Isemer and Hasse (1989) are also suitable surface heat flux parameterizations. The momentum results of Bumke et al. (2002) broadly concur with those of Smith (1980), the results of which are approximated by the momentum algorithm described in Smith (1988).

A comparison of surface flux time series is shown in Fig. 5, with parts (a) sensible heat flux (shfx), (b) latent heat flux (lhfx), and (c) surface momentum flux or stress (τ). The correspondence between the three time series is very good, as one would expect from the results of

the previous section (Fig. 2). However, there are large differences in magnitude between the sensible and latent flux estimates, especially during high heat flux events. The differences between the observed sensible and latent fluxes and the ECMWF fluxes are up to 30% during high heat flux events. The differences between the observed sensible and latent fluxes and the NCEP fluxes are even greater, up to 100% and 50%, respectively, during high heat flux events. The surface stresses generally correspond well, with any large differences between the model and observations due to differences in the 10-m wind speeds at that time (Fig. 2d).

Figure 6 shows scatterplots of observed versus model sensible heat flux and latent heat flux from the reduced dataset (120 points) described earlier. Table 3 summarizes the surface flux data, and Table 4 compares the observed data and model output. The correlation coefficients for the observed versus model comparisons are over 0.9 for the heat fluxes. This is a reflection of the reasonable level of accuracy in the surface layer variables, and an indication that the physics of the heat exchange is being modeled in a coherent way. However, the linear regression slopes in all the comparisons are more than 1, and in the case of the NCEP model the slopes are surprisingly large (1.61 and 1.42 for the shfx and lhfx, respectively). The positive slopes lead to bias errors of 24 and 14 W m^{-2} for the ECMWF fluxes and of 94 and 36 W m^{-2} for the NCEP sensible and latent heat fluxes. These are large biases given the observed dataset means of 184 and 136 W m^{-2}. The total rms differences for the sensible heat flux are 165 W m^{-2} for the ECMWF and 223 W m^{-2} for the NCEP, larger than the observed mean.

The correlation coefficients for the momentum flux comparisons are both around 0.8, the lower correlation reflecting the squared dependence on U_{10} (recall that U_{10} has a lower correlation than T_a and q_a). Both comparisons have slopes of 1.27 and biases of 0.1 N m^{-2}, 30% of the observed mean. Over the ocean, our bulk flux algorithm and both model algorithms rely on a modified Charnock relation to calculate the momentum roughness length (and hence the neutral drag coefficient) from the friction velocity and thus relate wind speed, taking into consideration the surface layer stability, to the surface stress. Therefore, the key parameter is the Charnock constant α_C. This is set as 0.011 in the Smith (1988) algorithm, as 0.014 in the NCEP algorithm, and as 0.018 in the ECMWF algorithm (for more details see the appendix). Lower α_C values appear to be more appropriate for low to moderate wind speeds or open ocean areas, with higher α_C values more appropriate for higher wind speeds or coastal areas. In reality the Charnock constant is not a constant, but is related to wave height and steepness (e.g., Komen et al. 1998; Taylor and Yelland 2001), so the choice of a single suitable constant α_C is somewhat arbitrary. The larger α_C in the two model formulations explains the higher mean surface stresses in the model data.
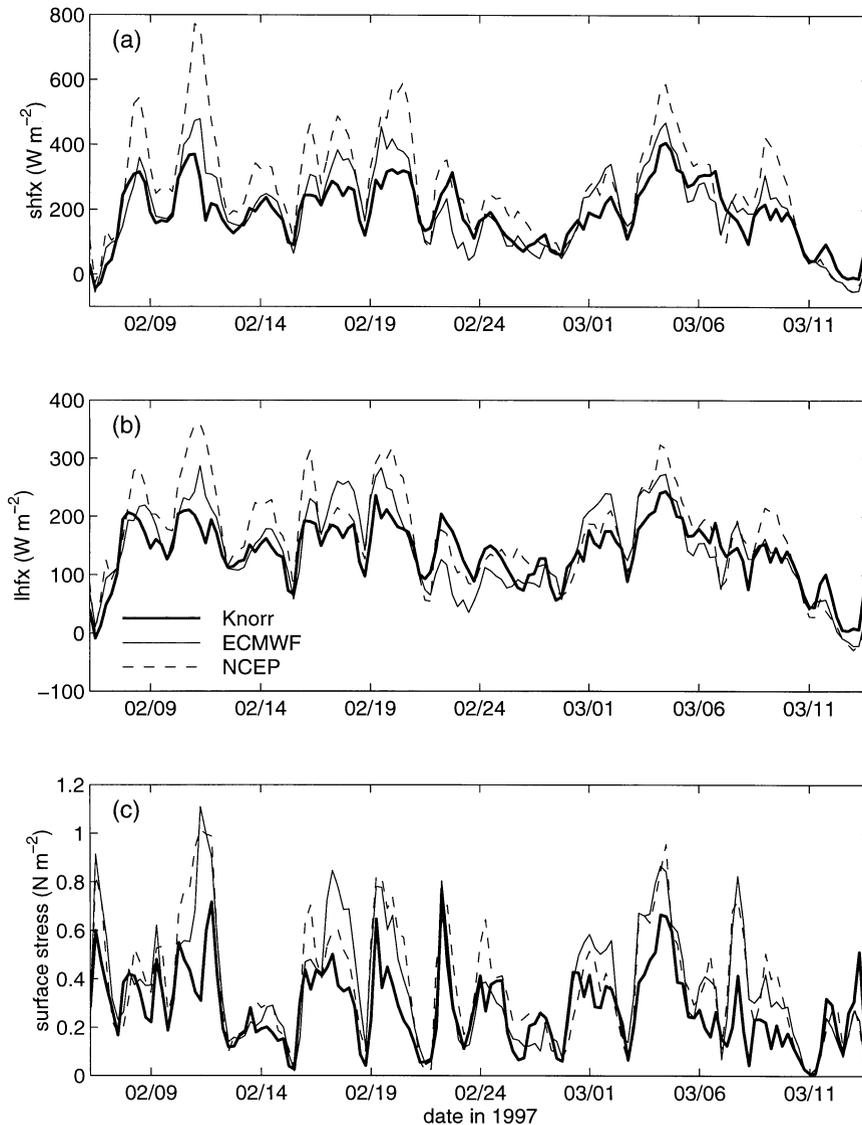
FIG. 5. Time series of *Knorr,* ECMWF, and NCEP data: (a) surface sensible heat flux (shfx); (b) surface latent heat flux (lhfx); and (c) surface wind stress (τ).

The large differences in the bulk observed heat fluxes and those from the ECMWF and NCEP models warrants further investigation. A first step was to recalculate surface heat fluxes using model surface-layer data and the bulk flux algorithm of Smith/DeCosmo et al. outlined above.[1] Comparing these fluxes would give an idea of the errors due to the differences in T_a, SST, q_a and U_{10} found in section 3. The results of this recalculation are shown in columns 4 and 5 of Table 4. In general, the correlation coefficients are around 0.9, the slopes are closer to 1, and the total rms errors are reduced. Recall that there were no significant biases in wind speed, hence the sensible heat flux comparisons are largely a reflection of the air–sea temperature difference,[2] with the regression slopes similar to those of T_a. The biases are only 1 and −18 W m^{-2} for ECMWF and NCEP respectively, significantly less than for the observed versus model comparison. The latent flux comparisons are largely a reflection of the difference between q_a and q_s

---

[1] It should be noted that these recalculated model fluxes would not be equal to model-output fluxes using this bulk algorithm within the model. The models do not calculate surface fluxes using, for example, 2-m temperature and 10-m wind; rather, the lowest model level and surface values are used to evaluate the surface fluxes at every time step. Hence, a different algorithm within the model would affect the surface layer evolution as well as the surface flux field.

[2] Note that the cold bias in the ECMWF data is not evident in the recalculated fluxes because the air–sea temperature differences are not biased.
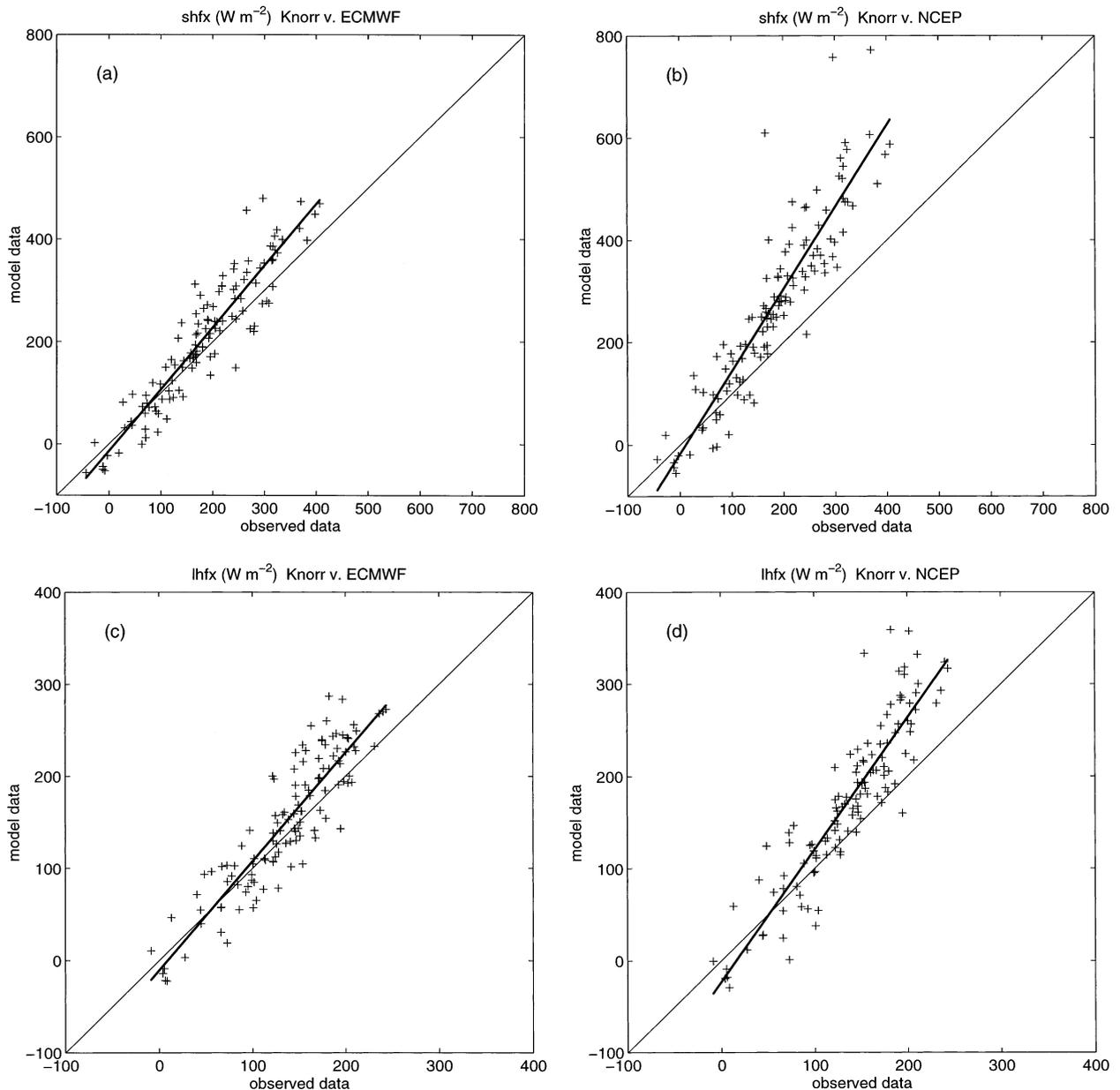
FIG. 6. Scatterplots of *Knorr* vs ECMWF and *Knorr* vs NCEP data: (a) and (b) compare shfx, while (c) and (d) compare lhfx. A linear regression line is shown, where the observed data are assumed to be independent and the model data dependent.

(SST), with regression slopes similar to those of q_a. The biases are -7 and -34 W m^{-2} for the ECMWF and NCEP data, respectively. The low latent heat flux using NCEP data is due to the high relative humidity.

In general the bias, slope, and rms errors for the re-calculated fluxes are smaller than those of the model output fluxes, and the linear regression fits have significantly improved. This therefore suggests that the largest source of error in the model heat fluxes is not from problems in the surface layer data, but rather in the calculation of the model surface fluxes. Several stud-

ies, recently published or in press, have carried out comparisons similar to the above. Bony et al. (1997) and Shinoda et al. (1999) compared NCEP reanalyses fluxes with satellite and in situ observations over the tropical ocean, and found small overestimates in the latent heat flux of order 10 W m^{-2}. Smith et al. (1999) compared NCEP fluxes with research vessel observations over several cruises in different parts of the World Ocean. Josey (2001) compared both ECMWF and NCEP data to buoy measurements in the northeast Atlantic and found 30–40 W m^{-2} total heat flux overestimates by
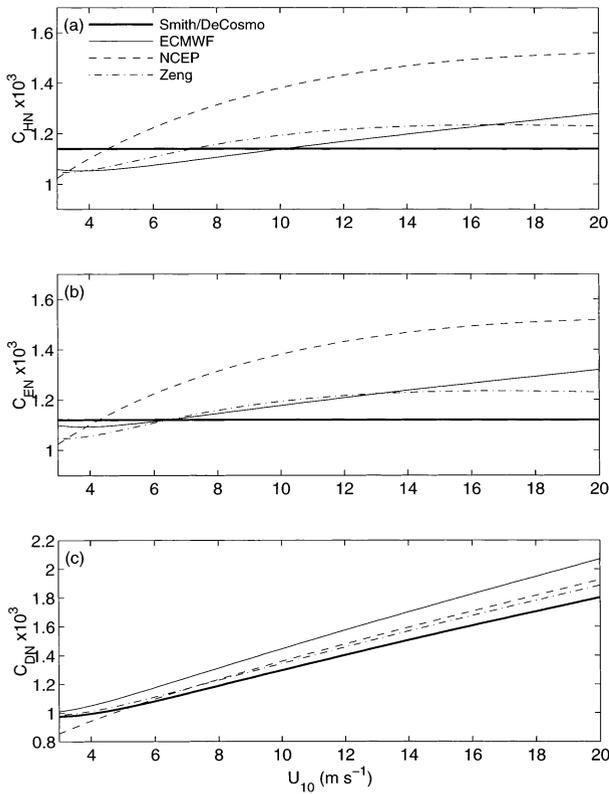
FIG. 7. Neutral transfer coefficients of heat, moisture, and momentum (C_{HN}, C_{EN}, and C_{DN}) as a function of 10-m wind speed. Results are from the bulk algorithms of Smith/DeCosmo (thick line), ECMWF (thin line), NCEP (dashed line), and Zeng et al. (dashed–dotted line).
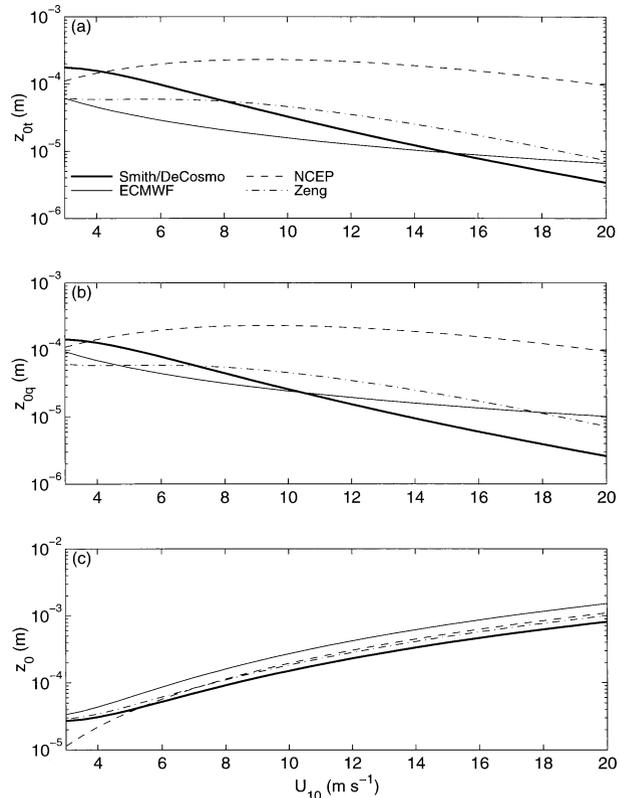

FIG. 8. Roughness lengths for heat, moisture, and momentum (z_{0b}, z_{0q}, and z_0) as a function of 10-m wind speed. Results are from the bulk algorithms of Smith/DeCosmo et al., ECMWF, NCEP, and Zeng et al., as indicated.

both models. In our comparison, the generally high wind speeds and unstable conditions highlight much larger differences in the surface heat fluxes, with model overestimates of up to 130 W m^{-2} in total heat flux. In the next section we show why this is the case through a detailed investigation of the model bulk algorithms.

5. Bulk algorithms for sea surface fluxes

The atmosphere–ocean fluxes of heat, moisture, and momentum are, by necessity, parameterized in numerical weather prediction and climate models. Most parameterizations rely on bulk algorithms that relate surface layer data to surface fluxes using formulas based on similarity theory and empirical relationships. The standard treatment involves the calculation of roughness lengths for heat, moisture, and momentum (z_{0t}, z_{0q}, z_0) from the observed wind speed, via an iteration routine involving a scaling temperature t_*, a scaling humidity q_*, and the friction velocity u_*. This allows calculation of the neutral transfer coefficients for heat, moisture, and momentum, that is, C_{HN}, C_{EN}, C_{DN}. To calculate the actual transfer coefficients (i.e., C_H, C_E, C_D), one must allow for the stability of the surface layer, and this is

done through an iteration routine involving stability correction factors or "ψ functions." The surface fluxes are then defined as

$$\tau = C_D \rho U_{10}^2 = \rho u_*^2 \qquad (1)$$

$$\text{shfx} = C_H \rho c_p U_{10}(T_{SST} - T_a) = -\rho c_p u_* t_* \qquad (2)$$

$$\text{lhfx} = C_E \rho L U_{10}(q_s - q_a) = -\rho L u_* q_*, \qquad (3)$$

where ρ is the density, c_p the specific heat at constant pressure for air, T_{SST} is the sea surface temperature, L is the latent heat of vaporisation, and q_s is the saturated specific humidity at T_{SST}. We use the convention that a positive surface flux is from the ocean into the atmosphere.

In this section we investigate the bulk algorithms used for the "observed" surface fluxes (i.e., the Smith/DeCosmo algorithm), the ECMWF model fluxes, the NCEP model fluxes, and that recommended by Zeng et al. (1998). An intercomparison with the Zeng et al. (1998) algorithm is included, as this was the algorithm adopted by the NCEP for their operational model as of 15 June 1998 (H.-L. Pan 2000, personal communication). The details of the bulk algorithms may be found in the appendix.
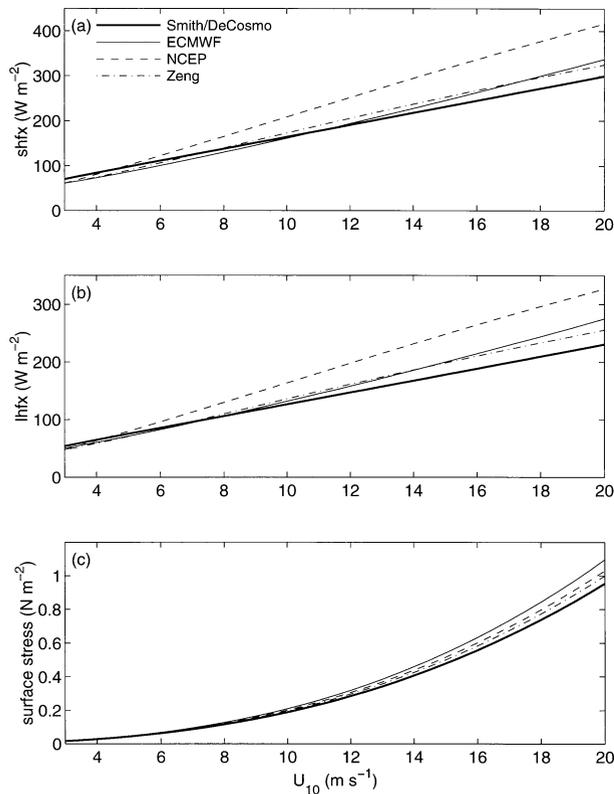
FIG. 9. Surface sensible heat flux, surface latent heat flux, and surface stress as a function of 10-m wind speed, for the cruise mean conditions (see text). Results are from the bulk algorithms of Smith/DeCosmo et al., ECMWF, NCEP, and Zeng et al., as indicated.
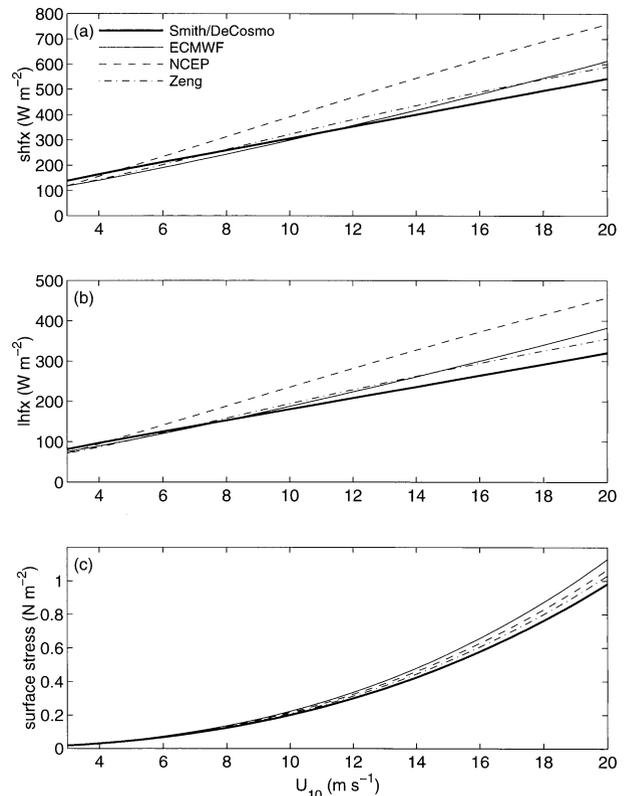
FIG. 10. Surface sensible heat flux, surface latent heat flux, and surface stress as a function of 10-m wind speed, for the cruise extreme conditions (see text). Results are from the bulk algorithms of Smith/DeCosmo et al., ECMWF, NCEP, and Zeng et al., as indicated.

Figure 7 shows the heat, moisture, and momentum exchange coefficients (C_{HN}, C_{EN}, C_{DN}) as functions of 10-m wind speed. The functional forms of the heat and moisture coefficients are similar for each algorithm, but quite different when comparing one algorithm to another. In particular, the NCEP algorithm is an obvious outlier to the others, giving much greater C_{HN} and C_{EN} values for moderate to strong wind speeds. The differences between the C_{DN} lines are due to the small differences in the Charnock coefficient, and at low wind speed the inclusion of the smooth flow regime (see appendix, Table A1).

A succession of observational programs have aimed to examine both the functional form of the bulk algorithms, and the coefficients that need to be prescribed. The consensus appears to be that the form of the Charnock formula–based drag coefficient is a valid first approximation for use in models. In other words, an increasing drag with wind speed, caused by generally steeper waves in high wind regimes, is appropriate (e.g., DeCosmo et al. 1996; Zeng et al. 1998). However, the same observational programs are essentially inconclusive and somewhat contradictory on the general forms of C_{HN} and C_{EN}. To quote DeCosmo et al., ''no significant variation with wind speed has been found for wind

speeds up to 18 m s^{-1} for C_{EN} and up to 23 m s^{-1} for C_{HN}.'' Hence the wide variety of parameterizations that are detailed here. Although they did not find a relationship between C_{EN} and U_{10}, DeCosmo et al. do find a statistically significant positive correlation between C_{EN} and $C_{DN}^{1/2}$. They explain this result as a correlation between the scatter in C_{DN} data and the variation in C_{EN} caused by physics that is not yet understood. The scatter in their results is large, however, and this result has so far been unsubstantiated.

Figure 8 shows the roughness lengths for heat, moisture and momentum plotted on a logarithmic scale, against wind speed and under neutral conditions. Note that our comments would be the same for unstable conditions. The z_{0t} and z_{0q} curves are all quite different: the majority decreasing with increasing U_{10}, to allow for the more rapidly increasing C_{DN} with U_{10} (Fig. 7), but with the NCEP curves almost constant. The DeCosmo et al. (1996) observations, and to a lesser extent those illustrated in Zeng et al. (1998), suggest that z_{0t} and z_{0q} should decrease with increasing wind speed, which implies that the NCEP algorithm is inappropriate. However, if there is a linear relationship between C_{EN} and $C_{DN}^{1/2}$, as hinted by the results of DeCosmo et al., this would imply a constant value of z_{0q} with wind speed

(if the intercept was zero). Given the contradictory observational evidence, it is difficult to establish which algorithms are correct from this figure. The z_0 curves are all very similar, with departures at low wind speed on the inclusion, or not, of the smooth flow regime. The observational data of DeCosmo et al. and Zeng et al. are compatible with the increasing z_0 with U_{10} curves.

How the various bulk algorithms compare in terms of surface fluxes is illustrated in Figs. 9 and 10. For this comparison the stability of the surface layer is taken into account through the inclusion of standard stability correction ψ functions (see appendix). Figure 9 shows the surface fluxes of sensible heat, latent heat, and momentum against U_{10} for the *mean* thermodynamic conditions observed on the R/V *Knorr* 1997 cruise (from Table 1). Figure 10 shows the same for the most *extreme* thermodynamic conditions experienced on the cruise (i.e., minimum T_a and minimum RH). Note that the SST is held fixed at the mean observed value for the cruise. For sensible heat fluxes in the mean conditions, the differences between the algorithms lead to differences of up to 40 W m^{-2} for moderate (10 m s^{-1}) wind speeds and up to 100 W m^{-2} for high (20 m s^{-1}) wind speeds. For the extreme conditions, there are differences of over 70 W m^{-2} at moderate wind speeds and up to 200 W m^{-2} at high wind speeds. For latent heat fluxes in the observed mean conditions, there are differences of up to 35 W m^{-2} for moderate wind speeds and 80 W m^{-2} for high wind speeds. For the extreme conditions, there are differences of up to 50 W m^{-2} at moderate wind speeds and 130 W m^{-2} at high wind speeds. For both mean and extreme conditions the differences in the surface stresses are more confined to the level of observational uncertainty.

The differences in the surface heat fluxes illustrated in Figs. 9 and 10 are clearly large enough to explain the differences seen in the comparison of the observed and model-extracted fluxes. The comparison of bulk algorithms for these wintertime cold-air outbreak conditions has highlighted problems with the NCEP algorithm used operationally till June 1998 and used in the reanalysis project, including the Reanalysis-2 partial rerun (Kanimitsu et al. 1999). The NCEP algorithm causes large overestimates of the surface heat fluxes and, from the evidence of Fig. 8, the functional form of the NCEP algorithm is questionable. In contrast the forms of the other three algorithms are broadly comparable and are consistent with the level of knowledge of the physics of air–sea heat and moisture exchange.

6. Conclusions

The surface heat, moisture and momentum fluxes between the atmosphere and ocean have been examined for the operational analyses from the ECMWF and the reanalyses from the NCEP. Our first step has been a comparison with surface layer meteorological observations from a winter cruise of the Labrador Sea by the

R/V *Knorr*. In general the models reproduce the observed surface layer well, with the higher-resolution ECMWF analyses performing better than the NCEP reanalyses. There are some significant errors, mainly related to the poor sea surface temperature field and the coarse treatment of sea ice, that is, the lower boundary conditions. For example, these lead to a $-1.31°$C cold bias in the ECMWF 2-m temperature. The NCEP model mean relative humidity was almost 20% higher than that observed, and both models had low regression slopes in 10-m wind speed.

It should be noted that the models ability to reproduce horizontal gradients in the meteorological fields has not been tested in this study. The slow speed of the ship relative to changes in the atmosphere precludes any such comparison. However, a comparison of aircraft observations (Renfrew and Moore 1999) and regional model data has shown the importance of the boundary layer parameterization schemes and the inclusion of a model marginal ice zone in reproducing the structure of the thermodynamic fields (Pagowski and Moore 2001). Neither the ECMWF nor the NCEP global models presently include a MIZ.

A time series of observed surface fluxes was calculated from the observed *Knorr* data, using a well-established bulk algorithm following Smith (1988) and DeCosmo et al. (1996). The bulk algorithm was independently validated with turbulent flux measurements from onboard measurements. Comparing this observed time series with the model surface flux fields has brought to light some systematic differences. There were only 0.1 N m^{-2} differences in the surface stress fields, which were primarily due to different prescriptions of the Charnock constant. However, there were large differences between the observed and modeled sensible and latent heat fluxes, especially during high heat flux events. Over the entire cruise (120 6-h averages) the mean sensible heat fluxes were 184, 208, and 278 W m^{-2} and the mean latent heat fluxes were 136, 150, and 172 W m^{-2}, for the R/V *Knorr,* ECMWF, and NCEP data, respectively. In other words, the ECMWF model overestimated the sensible and latent heat fluxes by 13% and 10%, while the NCEP model overestimated the heat fluxes by 51% and 27% respectively. We show that these differences are due to different heat and moisture roughness length formulations. In particular, we show that the NCEP reanalysis formulation is inappropriate for moderate to high wind speeds where the exchange coefficients are too large. This problem is most acute when air–sea temperature differences are large, that is, regions of high surface sensible heat flux: for example, areas prone to wintertime cold-air outbreaks such as the Labrador Sea, the Norwegian Sea, the Gulf Stream, and the Kuroshio (Josey et al. 1999). Note that the same roughness length formulation is used in the NCEP partial rerun Reanalysis-2.

There is an inherent uncertainty in relating surface layer fields to surface fluxes using bulk algorithms, due

to the approximation in the boundary layer physics that it entails. In a discussion of this, Garratt (1992) suggests an uncertainty in the neutral exchange coefficients of ±15%. With this in mind, we conclude that the surface flux fields from the ECMWF analyses are within the bounds of observational uncertainty and are therefore suitable as forcing fields for oceanic modeling. In contrast, we would conclude that the surface flux fields from the NCEP reanalyses are not suitable for driving ocean models, as the sensible and latent heat fluxes are likely to be systematically overestimated in moderate to high wind conditions. Instead we would recommend that the surface heat fluxes are recalculated using a more appropriate bulk algorithm.

APPENDIX

Transfer Coefficients and Bulk Flux Algorithms

The bulk algorithms compared in this study all make use of a modified Charnock formula (Charnock 1955) to relate the momentum roughness length (z_0) to the friction velocity (u_*) and hence wind speed over the ocean:

$$z_0 = \alpha_C \frac{u_*^2}{g} + b \frac{\nu}{u_*}, \tag{A1}$$

where α_C is the Charnock constant, b is a "smooth flow" constant, ν is the dynamic viscosity of air (a function of temperature, but usually set to a constant between 1.4 and 1.5×10^{-5} m s^{-1}), and g is the gravitation constant. The smooth flow limit is catered for by setting

TABLE A1. Coefficients used in the modified Charnock equation (A1) for the various bulk algorithms.

	Smith (1988)/ DeCosmo et al. (1996)	ECMWF	NCEP	Zeng et al. (1998)
α_C	0.011	0.018	0.014	0.013
b	0.11	0.11	0	0.11

b nonzero (e.g., Garratt 1992). Given U_{10}, Eq. (A1) is solved in an iteration loop with the logarithmic neutral wind profile:

$$U_{10} = \frac{u_*}{K} \ln\left(\frac{z}{z_0}\right), \tag{A2}$$

where $K = 0.4$ is the von Kármán constant and z is the measurement height (in this case 10 m). The neutral drag coefficient is then

$$C_{DN} = \frac{K^2}{\ln(z/z_0)^2}, \tag{A3}$$

following Smith (1988), for example. The above equations are common to all the bulk algorithms, with different values for the constants α_C and b (Table A1).

The neutral transfer coefficients for heat and moisture are calculated in a similar way:

$$C_{HN} = \frac{K^2}{\ln(z/z_0) \ln(z/z_{0t})} \tag{A4}$$

$$C_{EN} = \frac{K^2}{\ln(z/z_0) \ln(z/z_{0q})}, \tag{A5}$$

where z_{0t} and z_{0q} are the roughness lengths for temperature and humidity, respectively. The specification of these two scalar roughness lengths is the crucial difference between the bulk algorithms.

The bulk algorithm used on the *Knorr* observed data is based on that of Smith (1988), which fixes C_{HN} and C_{EN} as constant. The constants we have used are those suggested by the HEXOS results, as documented in the comprehensive study of DeCosmo et al. (1996):

$$C_{HN} = 1.14 \times 10^{-3} \tag{A6}$$

$$C_{EN} = 1.12 \times 10^{-3}. \tag{A7}$$

These were determined for largely moderate to high wind speeds and unstable conditions, so they are appropriate for our dataset. The coefficients are broadly consistent with the constants recommended in the Tropical Ocean Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE) study of Fairall et al. (1996), although they incorporate a "cool skin" effect.

The bulk algorithm employed by the operational model of the ECMWF (from 4 Aug 1993) and also used in their reanalysis project is detailed in Beljaars (1994) and Beljaars and Viterbo (1998). The roughness lengths for temperature and humidity are defined using equations analogous to those of a smooth flow:

$$z_{0t} = 0.62 \frac{\nu}{u_*} \tag{A8}$$

$$z_{0q} = 0.40 \frac{\nu}{u_*}, \tag{A9}$$

where ν is fixed at 1.5×10^{-5} m s^{-1}.

The bulk algorithm used in the NCEP suite of models and in the NCEP–NCAR reanalysis project is documented in Zeng et al. (1998). The roughness lengths for temperature and humidity are defined via

$$\ln\left(\frac{z_0}{z_{0t}}\right) = \ln\left(\frac{z_0}{z_{0q}}\right) = \frac{-1.076 + 0.7045 \ln(\text{Re}_*) - 0.058\,08[\ln(\text{Re}_*)]^2}{1 - 0.1954 \ln(\text{Re}_*) - 0.009\,999[\ln(\text{Re}_*)]^2},$$ (A10)

where the roughness Reynolds number is defined as $\text{Re}_* = u_* z_0 / \nu$.

The above NCEP algorithm was replaced on 15 June 1998 by that of Zeng et al. (1998) in the NCEP operational model (H.-L. Pan 2000, personal communication). The Zeng et al. bulk formula was developed from the TOGA COARE tropical ocean experiment, and is a simplification of the TOGA COARE algorithm described in Fairall et al. (1996). The roughness lengths for temperature and humidity are defined following the functional form of Brutsaert (1975, 1982), which is also recommended in Garratt (1992). It is based on scaling arguments and surface renewal theory (Liu et al. 1979). The functional form is then fit to observations such that

$$\ln\left(\frac{z_0}{z_{0t}}\right) = \ln\left(\frac{z_0}{z_{0q}}\right) = 2.67\text{Re}_*^{1/4} - 2.57,$$ (A11)

where Re_* is the roughness Reynolds number.

It has recently become commonplace (e.g., Fairall et al. 1996; Zeng et al. 1998) for the lower saturation vapour pressure of saline water to be taken into account when the surface latent heat fluxes are estimated, that is, by modifying q_s to $q_s = 0.98\ q_s(T_{\text{SST}})$ in Eq. (3), following, for example, Sverdrup et al. (1942). This modification is not used in either of the current model algorithms, nor was it incorporated by DeCosmo et al. (1996) when calculating their exchange coefficients. Hence, for the sake of a fair comparison we have not used it in calculating the *Knorr* bulk fluxes.

To calculate the surface fluxes of momentum, heat, and moisture the stability of the surface layer must be taken into account, and this is generally done through the inclusion of stability correction functions, or "ψ functions." The ψ functions are included in the iteration scheme used to calculate the scaling temperature t_*, the scaling humidity q_*, and the friction velocity u_*. For example, Eq. (A2) becomes

$$U_{10} = \frac{u_*}{K}\left(\ln\left(\frac{z}{z_0}\right) - \psi_m\right)$$ (A12)

with the addition of the ψ_m function for momentum. Similar equations apply for the calculation of t_* and q_*, based on Eqs. (2) and (3):

$$(T_{\text{SST}} - T_a) = \frac{t_*}{K}\left(\ln\left(\frac{z}{z_{0t}}\right) - \psi_t\right)$$ (A13)

$$(q_s - q_a) = \frac{q_*}{K}\left(\ln\left(\frac{z}{z_{0q}}\right) - \psi_t\right),$$ (A14)

where ψ_t is the stability correction for temperature and humidity. Equations (1)–(3) may then be used to calculate the surface fluxes and the transfer and drag coefficients. The stability correction functions for moderately stable and moderately unstable conditions are well established and used by all the bulk algorithms. They come from Monin–Obukhov similarity theory, and the standard formulas are the Businger–Dyer (or Kansas-type) relations (e.g., Paulson 1970; Businger et al. 1971; Dyer 1974). There are extensions to these formulas for highly stable conditions (e.g., Holtslag et al. 1990; Zeng et al. 1998) and highly unstable conditions (e.g., Kader and Yaglom 1990; Grachev et al. 1998; Zeng et al. 1998), and variations of these extensions are used in the model algorithms in operational mode. In the comparison discussed in section 5 only the standard Businger–Dyer functions are used, so that we are able to directly compare the effects of different scaling roughness formulations [i.e., (A6)–(A11)]. The effects of the highly stable and highly unstable correction factors are unimportant at the moderate to high wind speeds experienced over the Labrador Sea cruise.

REFERENCES

Anderson, R. J., and S. D. Smith, 1981: Evaporation coefficients for the sea surface from eddy flux measurements. *J. Geophys. Res.,* **86,** 449–456.

Beljaars, A. C. M., 1994: The impact of some aspects of the boundary layer scheme in the ECMWF model. *Proc. Seminar on Parameterization of Sub-grid Scale Processes,* Reading, United Kingdom, ECMWF, 125–161.

——, and P. Viterbo, 1998: The role of the boundary layer in a numerical weather prediction model. *Clear and Cloudy Boundary Layers,* A. A. M. Holtslag and P. G. Duynkerke, Eds., Royal Netherlands Academy of Arts and Sciences, 287–304.

Bony, S., Y. Sud, K. M. Lau, J. Susskind, and S. Saha, 1997: Comparison and satellite assessment of NASA/DAO and NCEP–NCAR reanalyses over tropical ocean: Atmospheric hydrology and radiation. *J. Climate,* **10,** 1441–1462.

Brutsaert, W. H., 1975: A theory for local evaporation (or heat transfer) from rough and smooth surfaces at ground level. *Water Resour. Res.,* **11,** 543–550.

——, 1982: *Evaporation into the Atmosphere—Theory, History and Applications.* Reidel, 299 pp.

Bumke, K., K. Uhlig, H. Berdt, and M. Clemens, 2001: Measurements of snow fall over the Labrador Sea. *J. Phys. Oceanogr.*, submitted.

——, U. Karger, and K. Uhlig, 2002: Measurements of turbulent fluxes of momentum and sensible heat over the Labrador Sea. *J. Phys. Oceanogr., 32,* 401–410.

Businger, J. A., J. C. Wyngaard, Y. Izumi, and E. F. Bradley, 1971: Flux-profile relationships in the atmospheric surface layer. *J. Atmos. Sci., 28,* 181–189.

Charnock, H., 1955: Wind stress on a water surface. *Quart. J. Roy. Meteor. Soc., 81,* 639–640.

DeCosmo, J., K. B. Katsaros, S. D. Smith, R. J. Anderson, W. A. Oost, K. Bumke, and H. Chadwick, 1996: Air–sea exchange of water vapor and sensible heat: The Humidity Exchange over the Sea (HEXOS) results. *J. Geophys. Res., 101,* 12 001–12 016.

Dickson, R., J. Lazier, J. Meinke, P. Rhines, and J. Swift, 1996: Long-term coordinated changes in convective activity of the North Atlantic. *Progress in Oceanography,* Vol. 38, Pergamon, 241–295.

Dyer, A. J., 1974: A review of flux-profile relationships. *Bound.-Layer Meteor., 7,* 363–372.

ECMWF, 1995: User guide to ECMWF products 2.1. Meteor. Bull. M3-2, 71 pp.

Fairall, C. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young, 1996: Bulk parameterization of air–sea fluxes for the Tropical Ocean–Global Atmosphere Coupled Ocean–Atmosphere Response Experiment. *J. Geophys. Res., 101,* 3747–3764.

Garratt, J. R., 1992: *The Atmospheric Boundary Layer.* Cambridge University Press, 316 pp.

Grachev, A. A., C. W. Fairall, and S. E. Larsen, 1998: On the determination of the neutral drag coefficient in the convective boundary layer. *Bound.-Layer Meteor., 86,* 257–278.

Holtslag, A. A. M., E. I. F. de Bruijn, and H.-L. Pan, 1990: A high resolution air mass transformation model for short-range weather forecasting. *Mon. Wea. Rev., 118,* 1561–1575.

Isemer, H. J., and L. Hasse, 1989: *The BUNKER Climate Atlas of the North Atlantic Ocean,* Vol. 2: *Air–Sea Interactions,* Springer-Verlag, 256 pp.

Josey, S. A., 2001: A comparison of ECMWF, NCEP–NCAR, and SOC surface heat fluxes with moored buoy measurements in the subduction region of the northeast Atlantic. *J. Climate, 14,* 1780–1789.

——, E. C. Kent, and P. K. Taylor, 1999: New insights into the ocean heat budget closure problem from analysis of the SOC air–sea flux climatology. *J. Climate, 12,* 2856–2880.

Kader, B. A., and A. M. Yaglom, 1990: Mean fields and fluctuation moments in unstably stratified turbulent boundary layers. *J. Fluid Mech., 212,* 637–662.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc., 77,* 437–471.

Kanimitsu, M., W. Ebisuzaki, J. Woolen, J. Potter, and M. Fiorino, 1999: An overview of NCEP/DOE reanalysis-2. *Proc. Second Int. Conf. on Reanalysis,* Reading, United Kingdom, ECMWF, 18.

Killworth, P. D., 1983: Deep convection in the world ocean. *Rev. Geophys. Space Phys., 21,* 1–26.

Komen, G., P. A. E. M. Janssen, V. Makin, and W. Oost, 1998: On the sea state dependence of the Charnock parameter. *Global Atmos. Ocean Syst., 5,* 367–388.

Lab Sea Group, 1998: The Labrador Sea Deep Convection Experiment. *Bull. Amer. Meteor. Soc., 79,* 2033–2058.

Liu, W. T., K. B. Katsaros, and J. A. Businger, 1979: Bulk parameterization of air–sea exchanges of heat and water vapor including the molecular constraints at the interface. *J. Atmos. Sci., 36,* 1722–1735.

Marshall, J., and F. Schott, 1999: Open-ocean convection: Observations, theory and models. *Rev. Geophys., 37,* 1–64.

Moat, B., and M. Yelland, 1998: Airflow distortion at instrument sites on the R/V *Knorr.* Southampton Oceanography Centre, 29 pp.

Pagowski, M., and G. W. K. Moore, 2001: A numerical study of an extreme cold-air outbreak over the Labrador Sea: Sea-ice air–sea interaction and the development of polar lows. *Mon. Wea. Rev., 129,* 47–72.

Paulson, C. A., 1970: The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *J. Appl. Meteor., 9,* 857–861.

Renfrew, I. A., and G. W. K. Moore, 1999: An extreme cold air outbreak over the Labrador Sea: Roll vortices and air–sea interaction. *Mon. Wea. Rev., 127,* 2379–2394.

——, ——, T. R. Holt, S. W. Chang, and P. Guest, 1999: Mesoscale forecasting during a field program: Meteorological support of the Labrador Sea Deep Convection Experiment. *Bull. Amer. Meteor. Soc., 80,* 605–620.

Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate, 7,* 929–948.

Roemmich, D., J. Gilson, B. Cornuelle, and R. Weller, 2001: Mean and time-varying meridional transport of heat at the tropical/subtropical boundary of the North Pacific Ocean. *J. Geophys. Res., 106,* 8957–8970.

Shinoda, T., H. Hendon, and J. Glick, 1999: Intraseasonal surface fluxes in the tropical western Pacific and Indian Oceans from NCEP reanalyses. *Mon. Wea. Rev., 127,* 678–693.

Smith, S. D., 1980: Wind stress and heat flux over the ocean in gale force winds. *J. Phys. Oceanogr., 10,* 709–726.

——, 1988: Coefficients for sea surface wind stress, heat flux, and wind profiles as a function of wind speed and temperature. *J. Geophys. Res., 93,* 15 467–15 472.

Smith, S. R., D. M. Legler, and K. V. Verzone, 1999: Quantifying uncertainties in NCEP reanalyses using high-quality research vessel observations. *Proc. Second Conf. on Reanalyses,* Reading, United Kingdom, ECMWF, 87.

Sverdrup, H. U., M. W. Johnson, and R. H. Fleming, 1942: *The Oceans: Their Physics, Chemistry and General Biology.* Prentice Hall, 1087 pp.

Taylor, P. K., and M. J. Yelland, 2001: The dependence of sea surface roughness on the height and steepness of the waves. *J. Phys. Oceanogr., 31,* 572–590.

Weller, R. A., M. F. Baumgartner, S. A. Josey, A. S. Fischer, and J. C. Kindle, 1998: Atmospheric forcing in the Arabian Sea during 1994–1995: Observations and comparisons with climatology and models. *Deep-Sea Res., 45,* 1961–1999.

Wood, R. A., A. B. Keen, J. F. B. Mitchell, and J. M. Gregory, 1999: Changing spatial structure of the thermohaline circulation in response to atmospheric CO_2 forcing in a climate model. *Nature, 399,* 572–575.

Zeng, X., M. Zhao, and R. E. Dickinson, 1998: Intercomparison of bulk aerodynamical algorithms for the computation of sea surface fluxes using TOGA COARE and TAO data. *J. Climate, 11,* 2628–2644.