Evaluation of Latent Heat Flux Fields from Satellites and Models during SEMAPHORE

DENIS BOURRAS* AND W. TIMOTHY LIU

NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

LAURENCE EYMARD

Centre d'Etude des Environnements Terrestre et Planétaires, Velizy-Villacoublay, France

WENQING TANG

NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California

(Manuscript received 5 October 2001, in final form 2 July 2002)

ABSTRACT

Latent heat fluxes were derived from satellite observations in the region of Structure des Echanges Mer–Atmosphère, Propriétés des Hétérogénéités Océaniques: Recherche Expérimentale (SEMAPHORE), which was conducted near the Azores islands in the North Atlantic Ocean in autumn of 1993. The satellite fluxes were compared with output fields of two atmospheric circulation models and in situ measurements. The rms error of the instantaneous satellite fluxes is between 35 and 40 W m^{-2} and the bias is 60–85 W m^{-2}. The large bias is mainly attributed to a bias in satellite-derived atmospheric humidity and is related to the particular shape of the vertical humidity profiles during SEMAPHORE. The bias in humidity implies that the range of estimated fluxes is smaller than the range of ship fluxes, by 34%–38%. The rms errors for fluxes from models are 30–35 W m^{-2}, and the biases are smaller than the biases in satellite fluxes (14–18 W m^{-2}). Two case studies suggest that the satellites detect horizontal gradients of wind speed and specific humidity if the magnitude of the gradients exceeds a detection threshold, which is 1.27 g kg^{-1} $(100\ km)^{-1}$ for specific humidity and between 0.35 and 0.82 m s^{-1} $(30\ km)^{-1}$ for wind speed. In contrast, the accuracy of the spatial gradients of bulk variables from models always varies as a function of the location and number of assimilated observations. A comparison between monthly fluxes from satellites and models reveals that satellite-derived flux anomaly fields are consistent with reanalyzed fields, whereas operational model products lack part of the mesoscale structures present in the satellite fields.

1. Introduction

Horizontal fields of latent heat flux (L_E) at the air–sea interface are used to carry out heat budgets of the upper ocean and to initialize ocean circulation models. The latent heat flux is often parameterized as a function of bulk variables, which are horizontal scalar wind (u_A), specific humidity of the air just above the ocean surface (q_A), sea surface temperature (SST), and Dalton number (C_E; Businger et al. 1971). At timescales ranging from instantaneous to 1 month, flux fields may be obtained from atmospheric circulation models (ACM) or satellite sensor data. ACMs produce fields that correspond to a physical and dynamical interpolation in space and time between observations. However, there are often large uncertainties in the analyses because assimilated data are sparse and the boundary layer models implemented in ACMs are inadequate.

Satellite observations may be used to derive flux fields with a resolution down to 0.3° × 0.3°. Moreover, it takes only 2–3 days to obtain global coverage. The disadvantage is that the accuracy of the satellite fluxes is not clearly known yet. Previous works indicated that the flux retrieval error ranged from 20 to 50 W m^{-2} at global scale, depending on the timescale considered (Esbensen et al. 1993; Chou et al. 1995, 1997; Schulz et al. 1997). Nevertheless, global-scale validation is difficult because of the insufficient number of reference observations, and because of their fluctuating quality (Blanc 1987; Weare 1989). In order to confirm the results found at global scale, it is therefore necessary to perform regional studies in which satellite fluxes are compared with accurate surface measurements. At re-

* Current affiliation: CETP–IPSL–CNRS, Velizy-Villacoublay, France.

Corresponding author address: Denis Bourras, CETP–IPSL–CNRS, 10–12 Ave. de l'Europe, 78140 Vélizy-Villacoubley, France. E-mail: denis.bourras@cetp.ipsl.fr

gional scale, the error of the instantaneous satellite fluxes is 30 W m^{-2} according to Schulz et al. (1997) and 50 W m^{-2} after Clayson and Curry (1996) and Curry et al. (1999).

Our understanding of the representativeness of the satellite fluxes is still inadequate for two reasons. First, with the exception of Schulz et al. (1997), who compared 1-yr time series of satellite and Ocean Weather Ship M fluxes (located at 66°N, 2°E), all previous studies at regional scale concern the Tropical Ocean and Global Atmosphere Coupled Ocean–Atmosphere Response Experiment (TOGA COARE; Webster and Lukas 1992). Accurate flux estimates are difficult to obtain over the TOGA COARE region because of the presence of large humidity and liquid water contents in the atmosphere, and because of convection. Next, only Curry et al. (1999) started to focus on the mesoscale variability of the satellite fluxes. These authors compared an instantaneous satellite flux field with interpolated aircraft measurements. They concluded that there was a good agreement between satellite and aircraft fluxes.

The present study is an attempt to assess the accuracy of the satellite fluxes over a region different from TOGA COARE, to evaluate the representativeness of the spatial variations of the satellite fluxes, and to show how these fluxes compare with fluxes from ACMs. Fluxes and bulk variables are derived from satellite data at the instantaneous scale and at 1-month scale. They are compared with ship and aircraft measurements collected during the Structure des Echanges Mer–Atmosphère, Propriétés des Hétérogénéités Océaniques: Recherche Expérimentale (SEMAPHORE) campaign (Eymard et al. 1996), which was conducted in the North Atlantic in 1993. The performances of two flux retrieval methods are analyzed, the bulk method pioneered by Liu and Niiler (1984, hereinafter LN) and a more recent method based on a neural network (Bourras et al. 2002). The satellite-derived flux fields are compared with operational analyses of the European Centre for Medium-Range Weather Forecasts (ECMWF) model and with reanalyses of the Météo France (French meteorological office) Action de Recherche Petite Echelle Grande Echelle model (ARPEGE; Giordani et al. 1998).

The datasets are presented in section 2, and the flux retrieval methods are described in section 3. Individual fluxes from satellites and models are compared with in situ measurements in section 4. Next, the spatial variability of the fluxes is analyzed (section 5), and conclusions follow.

2. Data and model fields

This section describes the satellite observations, in situ measurements, and model output fields used in this study. The data were collected during the intensive observation period of SEMAPHORE, from 15 October to 15 November 1993 and from 30° to 38°N and 20° to 28°W. The meteorological conditions were mainly anticyclonic during the experiment, and the mean sea level pressure was 1020 hPa.

a. Satellites

Two kinds of space-based observations were used, brightness temperatures (T_B) measured by the Defense Meteorological Satellite Program Special Sensor Microwave Imagers (SSM/I) *F10* and *F11*, and satellite-derived SSTs. The SSM/I provided a coverage of the SEMAPHORE region twice a day near 0600 and 1800 UTC. The T_B were measured in two polarizations and five frequencies ranging from 19.35 to 85.5 GHz. The measurements were interpolated on 0.3° × 0.3° grids, the resolution of which is halfway between the footprints of the low- and the high-frequency channels of the SSM/I.

The Jet Propulsion Laboratory (JPL) Physical Oceanography Distributed Active Archive Center (PO.DAAC) produces and distributes SST fields with a resolution of 0.09° × 0.09°. These fields are derived from observations of the National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR). The accuracy of the SST retrieval algorithm used is 0.02 ± 0.5°C (Kilpatrick et al. 2001). Two different PO.DAAC products were used, the daily and the weekly SSTs (Armstrong and Vazquez-Cuervo 2001). Although these SSTs are not instantaneous observations, they were used in combination with the SSM/I T_B to derive instantaneous fluxes. The reason is that the SST varies weakly over short timescales, except in case of strong meteorological events. It is, of course, more accurate to use daily SSTs instead of weekly SSTs to derive instantaneous fluxes. However, the daily SST fields often contain gaps because the AVHRR does not detect the SST through clouds. Because the SEMAPHORE region was often cloudy during the experiment, weekly SSTs were helpful to obtain a good coverage of the region. Figure 1, in which daily and weekly AVHRR SSTs are compared, shows the good agreement between the two kinds of data. In particular, the bias between daily and weekly SSTs is negligible (0.06°C). The slope of the linear fit between daily and weekly SSTs is 0.89, however, which implies that the range of the weekly SSTs is slightly smaller than the range of the daily SSTs. For this reason, daily SSTs are used by default in the following, except in sections 4a and 5a.

b. In situ data and model fields

The in situ data used are measurements performed onboard *Le Suroît*, a ship from Institut Francais de Recherche pour l'Exploitation de la Mer (IFREMER). For studying the spatial variability of the fluxes, airborne observations of a Focker 27 aircraft were also used. This aircraft, called *Avion de Recherche Atmosphérique et de Télédétection* (*ARAT*) belongs to the Centre National de la Recherche Scientifique (CNRS). Earlier compar-

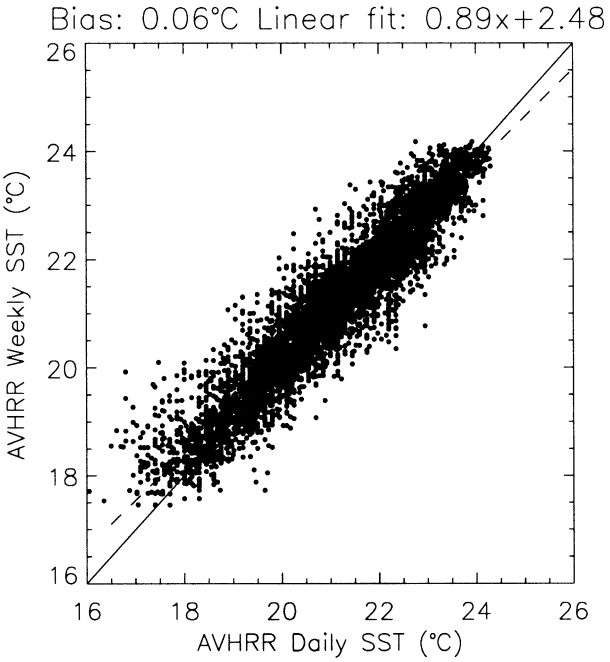

FIG. 1. Comparison between daily and weekly AVHRR-derived SSTs.

isons between observations from *Le Suroît*, two aircraft, radiosondes, and buoys have already shown good agreement in terms of bulk variables and fluxes (Eymard et al. 1996). The ship measurements used are the bulk variables u_A, q_A, and SST, plus air temperature just above the surface (θ_A) and sea level atmospheric pressure (p_S) that are also used to calculate the in situ bulk fluxes. The aircraft observations used are u_A and q_A only.

The ACM fields are ECMWF operational analyses and ARPEGE output fields (reanalyses, 3-h forecasts, and 6-h forecasts). The ECMWF analyses have a 1.125° × 1.125° resolution while the resolution of the ARPEGE fields is 0.16° × 0.16°. The analyses, reanalyses, and 6-h forecasts are available every 6 h from 0000 UTC, and the 3-h forecasts are available every 6 h from 0300 UTC. Giordani et al. (1998) used an optimal interpolation method to produce the reanalyses. They assimilated data from the Global Telecommunications System (GTS) as well as most of the SEMAPHORE buoy, radiosonde, aircraft, and ship data with a more important weight. This was intended to produce reanalyzed fields as close as possible to the data. The disadvantage is that the reanalyses often contain spatial artifacts because of the recent assimilation of data. For this reason, 3-h and 6-h ARPEGE forecasts (after reanalysis) are also considered for analyzing the spatial variability of the fluxes.

Although bulk fluxes were already available from ship data and from the two models, new fluxes were computed. The reason was to avoid possible differences in the flux calculation methods. Besides, the fluxes that are directly available from ACMs are not instantaneous fluxes, but fluxes integrated over 6-h intervals. The algorithm used to compute the new fluxes is described in Bourras (2000). It is based on the Dupuis et al. (1997) parameterization of the bulk coefficients and uses u_A, q_A, SST, p_S, and θ_A as inputs.

3. Satellite algorithms

Four algorithms were used to derive u_A, q_A, and L_E from the satellite data described in section 2a. The Goodberlet et al. (1990) and Schluessel et al. (1995, hereinafter S95) algorithms were used to obtain u_A and q_A, respectively. The latent heat flux retrieval methods used are the LN and the Bourras et al. (2002) methods.

The LN method consists in applying a bulk parameterization of the flux (1) to satellite estimates of bulk variables derived independently from each other:

$$L_E = \rho L_V C_E u_A (q_S - q_A). \quad (1)$$

In Eq. (1), ρ is the density of air and L_V is the vaporization latent heat constant (2.45×10^6 J kg^{-1}). Here q_S is the specific humidity against the sea surface. One assumes that q_S is 0.98 times the humidity at saturation, which is a known function of the SST. The SSTs used are the AVHRR products described in section 2a. Here C_E is 1.2×10^{-3}, for simplicity (Large and Pond 1982).

The Bourras et al. (2002) method is an artificial neural network (NN) that uses SSTs and physically based combinations of SSM/I T_B as inputs. The method was designed to derive instantaneous fluxes at 0.3° × 0.3° resolution. The NN was trained using a hybrid dataset that gathered SSM/I T_B, ECMWF-derived fluxes, and observations from the Tropical Atmosphere–Ocean (TAO) array of buoys. This dataset was based on 12 global fields randomly selected from 1997–98. Preliminary validation of the NN fluxes was carried out using data of three experiments, including SEMAPHORE. It was shown that the NN method was more accurate than the LN method, by 3–20 W m^{-2} in rms, depending on the dataset. Application of the NN method in the present study differs in several respects. For example, AVHRR SST data rather than in situ SST data are used. Also, the retrieval algorithm for q_A and the data selection criteria are different.

For both LN and NN fluxes, an algorithm described in Gerard and Eymard (1998) was used to track and discard all the SSM/I pixels for which the liquid water content was larger than 0.12 kg m^{-2}, because the liquid water present in the atmosphere affects the accuracy of the u_A and L_E estimates.

4. Intercomparisons of individual fluxes

In this section, instantaneous bulk variables (u_A, q_S, and q_A) and fluxes from satellites and models are validated against ship measurements. Next, the satellite fluxes are compared with ARPEGE and ECMWF fluxes, in order to confirm the results of the validation on a large number of points of comparison.

TABLE 1. Comparison of bulk variables and fluxes from satellites and models with ship measurements. The NN fluxes are satellite fluxes obtained using a neural network (Bourras et al. 2002). The LN fluxes are satellite fluxes derived using the Liu and Niiler (1984) method. The LNB fluxes are LN fluxes for which the bias in satellite q_A was removed. The LNTB fluxes are LN fluxes that account for ship measurements of θ_A and for satellite μ_A, q_S, and q_A (the bias in q_A was also removed for these fluxes). These results are based on 23 points of comparison.

Variable	Unit	Source	Correlation	Rms	Bias	Linear fit
μ_A	m s^{-1}	SSM/I	0.95	0.98	−0.13	$1.03x - 0.35$
—	—	ARPEGE	0.76	1.79	−0.41	$0.61x + 2.11$
—	—	ECMWF	0.86	1.53	−0.21	$0.91x + 0.35$
q_S	g kg^{-1}	AVHRR	0.98	0.26	−0.16	$0.79x + 2.99$
—	—	ARPEGE	0.96	0.30	−0.20	$0.84x + 2.20$
—	—	ECMWF	0.80	0.61	−0.15	$0.58x + 6.27$
q_A	—	SSM/I	0.88	0.94	2.46	$0.92x + 3.20$
—	—	ARPEGE	0.89	0.91	−0.22	$0.66x + 2.96$
—	—	ECMWF	0.91	0.78	0.36	$0.78x + 2.43$
L_E (NN)	W m^{-2}	Satellite	0.84	39.6	−58.8	$0.62x + 0.4$
— (LN)	—	—	0.86	36.1	−83.5	$0.66x - 28.3$
— (LNB)	—	—	0.90	33.0	−27.7	$0.91x - 13.7$
— (LNTB)	—	Hybrid	0.89	33.2	−5.8	$0.86x + 15.7$
L_E	—	ARPEGE	0.91	31.2	−14.1	$0.71x + 30.7$
—	—	ECMWF	0.87	35.9	−18.0	$0.78x + 15.9$

a. Validation of satellite and model variables

The ship measurements were selected within a 1-h interval around the hours of the SSM/I orbits and were averaged. Next, the pixels from the model fields the closest in time and space to the ship data were selected, and the satellite data available within a ±0.3° radius around the mean location of the ship were gathered and averaged. Hereinafter, bulk variables from satellites and models are first compared with ship data. Next, the validation is carried out in terms of fluxes.

1) BULK VARIABLES

The rms deviation between satellite-derived u_A and ship measurements is 0.98 m s^{-1} (Table 1), which is one-half the nominal error of the Goodberlet et al. (1990) algorithm. The u_A estimates are also accurate in terms of bias, the value of which is only −0.13 m s^{-1}. In contrast, ARPEGE and the ECMWF model have larger rms errors (1.79 and 1.53 m s^{-1}) and biases (−0.41 and −0.21 m s^{-1}) than the satellite in terms of u_A. Note the slope of the linear fit between ARPEGE u_A and ship data, which is only 0.61 (Fig. 2a, Table 1 line 2). Figure 2f shows that the satellite estimates of q_S compare well to the ship measurements in rms (0.26 g kg^{-1}) and bias (−0.16 g kg^{-1}). However, the satellite q_S are slightly underestimated at large q_S because of the use of weekly data, as mentioned in section 2a (the results of the comparison between daily and weekly SSTs presented in section 2a also apply in terms of q_S). The rms deviation between ECMWF and ship q_S is 0.61 g kg^{-1}, which is large and is 2 times the deviation between ARPEGE and ship q_S. This does not necessarily mean that ECMWF q_S are not accurate. Indeed, the ECMWF fields lack part of the variability of the ship measurements because of their low resolution. This is confirmed by the low slope of the linear fit between ECMWF and ship q_S (0.58).

The correlation coefficient and rms deviation between SSM/I and ship q_A are 0.88 and 0.94 g kg^{-1}, respectively, which is good (Fig. 2i). This success of the S95 algorithm in terms of rms is counterbalanced by a large bias (2.46 g kg^{-1}). In order to explain this bias, one has to recall that the SSM/I T_B are not sensitive to q_A but to the integrated water vapor content in the atmosphere W. Therefore, the S95 algorithm is in fact a statistical relationship between q_A and W, which means that the algorithm assumes the existence of a global mean vertical humidity profile. During SEMAPHORE, the humidity profiles had a singular shape because they were strong inversion profiles (Eymard et al. 1996). Thus, these profiles were different from a global mean profile, which explains the bias in Fig. 2i. Besides, strong inversions imply that there is almost no humidity above the top of the boundary layer, which potentially explains why the satellite q_A are overestimated. The rms deviation between ARPEGE (ECMWF) q_A and ship q_A is 0.91 g kg^{-1} (0.78 g kg^{-1}), which is comparable to the deviation between satellite and ship q_A. On the other hand, the biases in ARPEGE and ECMWF q_A are small, ~10% only of the bias in satellite q_A (Table 1). Note, however, that the slope of linear fit between satellite and ship q_A is surprisingly larger than the slopes between q_A from models and ship measurements, by 18%–40%.

2) FLUXES

The comparisons between fluxes from satellites, models, and the ship are presented in Fig. 3 and Table 1. The correlation between satellite and ship fluxes is ~0.85, which is comparable to the correlation between model and ship fluxes (~0.89). The rms errors of the two kinds of satellite fluxes are close, 36.1 W m^{-2} for

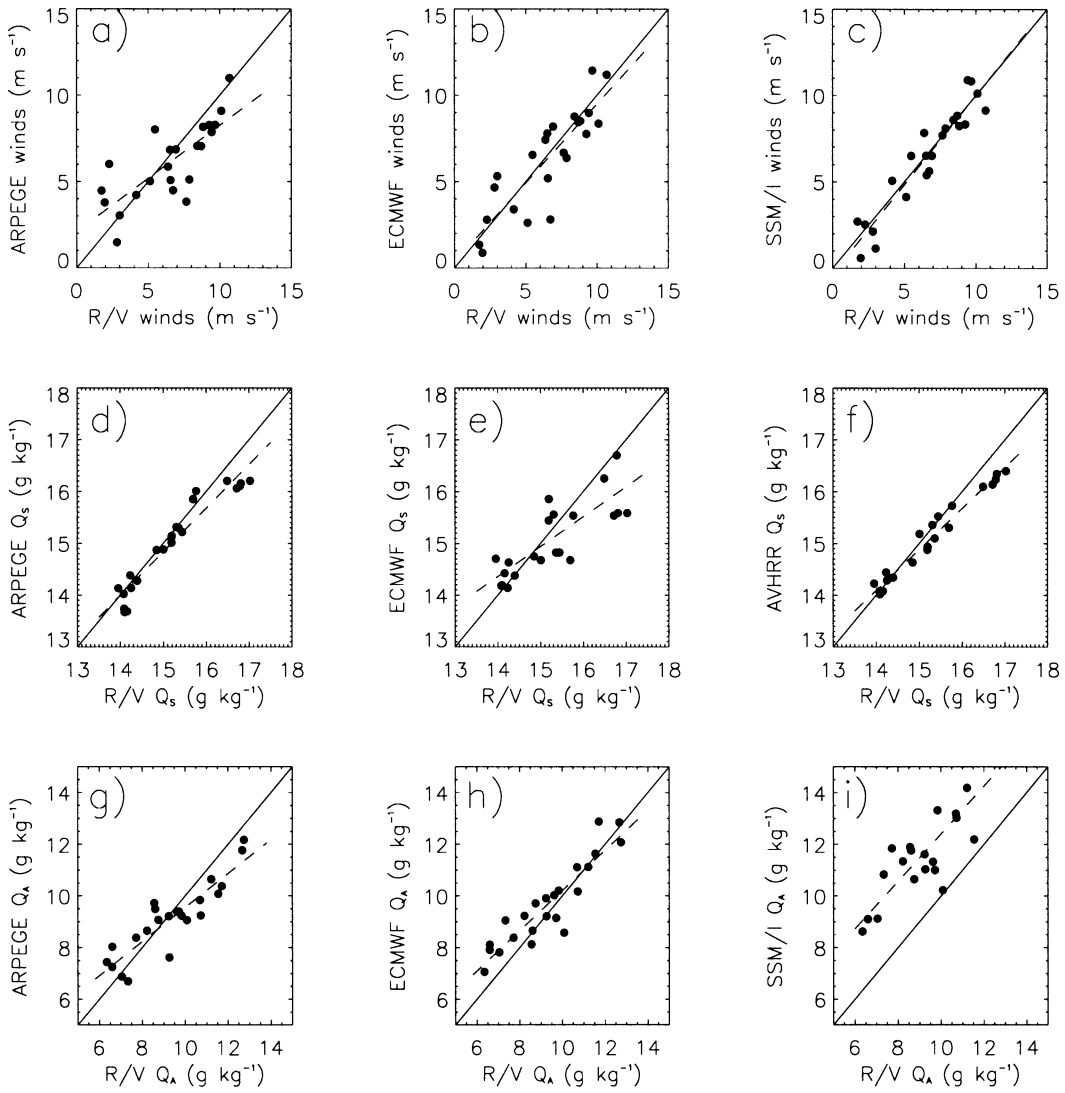

FIG. 2. Comparison of instantaneous SSM/I, ARPEGE, and ECMWF variables with ship measurements. The variables compared are (a)–(c) wind, (d)–(f) specific humidity q_S, and (g–i) q_A.

the LN fluxes and 39.6 W m^{-2} for the NN fluxes. The error in model fluxes is 31.2–35.9 W m^{-2}, which is only 0.2–8.4 W m^{-2} smaller than the error in satellite fluxes. The biases in model fluxes are between 14 and 18 W m^{-2} (or 5.6%–6.7%), which is small. In contrast, the satellite fluxes have large biases, which are -58.8 W m^{-2} for the NN fluxes and -83.5 W m^{-2} for the LN fluxes. These biases are mainly related to the bias in q_A (previous section), but also to the fact that satellite fluxes do not account for a variable C_E. This can be quantified in two steps. First, new LN fluxes that use nonbiased satellite q_A are computed (the value of the bias in q_A with respect to ship data is subtracted from the q_A estimates). These fluxes (referred as LNB fluxes) have a bias of -27.7 W m^{-2}, which is 55.8 W m^{-2} less than the bias of the LN fluxes (Table 1 line 13). Second, another set of LN fluxes (called LNTB) are computed using the bulk algorithm described in section 2b. The input variables are the satellite-derived u_A, q_S, and q_A (the bias of which was corrected as described above), plus the ship measurements of θ_A. The fact that θ_A was used means that a C_E that is a function of stability of the boundary layer is accounted for in the calculation of the fluxes. The bias of the LNTB fluxes is 5.8 W m^{-2}, as reported in Table 1. The biases of the LNB and LNTB fluxes eventually explain that ~55 W m^{-2} of the bias in LN fluxes are related to the bias in q_A, while ~20 W m^{-2} are related to C_E that is not derived from satellite data.

Figure 3 and Table 1 reveal that the slope of the linear fit between satellite and ship fluxes is only ~0.64. This implies that the range of satellite fluxes is smaller than the range of ship fluxes. The main reason is the bias in q_A. Indeed, the slope of the linear fit between LNB and

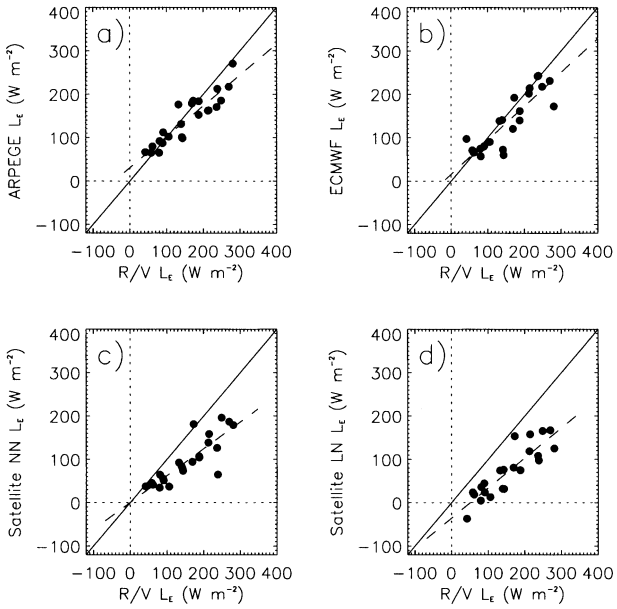

FIG. 3. Comparison of instantaneous SSM/I, ARPEGE, and ECMWF fluxes with ship fluxes. Satellite fluxes are estimated by means of (c) the Liu and Niiler (1984) method and (d) a neural network.

TABLE 2. Comparison of instantaneous fluxes from satellites and models. The results are based on 776 (9000) points for the comparisons between satellite and ECMWF (ARPEGE) products.

Sources compared	Correlation	Rms (W m^{-2})	Bias (W m^{-2})	Linear fit
NN–ARPEGE	0.76	50.2	−58.7	$0.74x + 12.9$
LN–ARPEGE	0.71	51.6	−88.1	$0.57x + 13.5$
NN–ECMWF	0.85	51.6	−71.1	$0.62x + 1.8$
LN–ECMWF	0.81	58.8	−100.9	$0.52x - 7.7$

ship fluxes is 0.91, which is even larger than the slope between model and ship fluxes (0.71–0.78). This could also be deduced from Eq. (1), which shows that a bias in q_A leads to a bias in L_E that consists of a constant plus a part that is proportional to the flux value. Therefore, a bias in q_A does not only imply a bias in L_E, but also reduces (or increases, depending on the sign of the bias) the slope of linear fit between satellite and ship fluxes.

b. Further comparisons between satellites and models

The satellite fluxes are compared with the ARPEGE and ECMWF fluxes in Fig. 4 and Table 2. The satellite flux fields were linearly interpolated at the resolution of the ECMWF fields for the comparison with the ECMWF

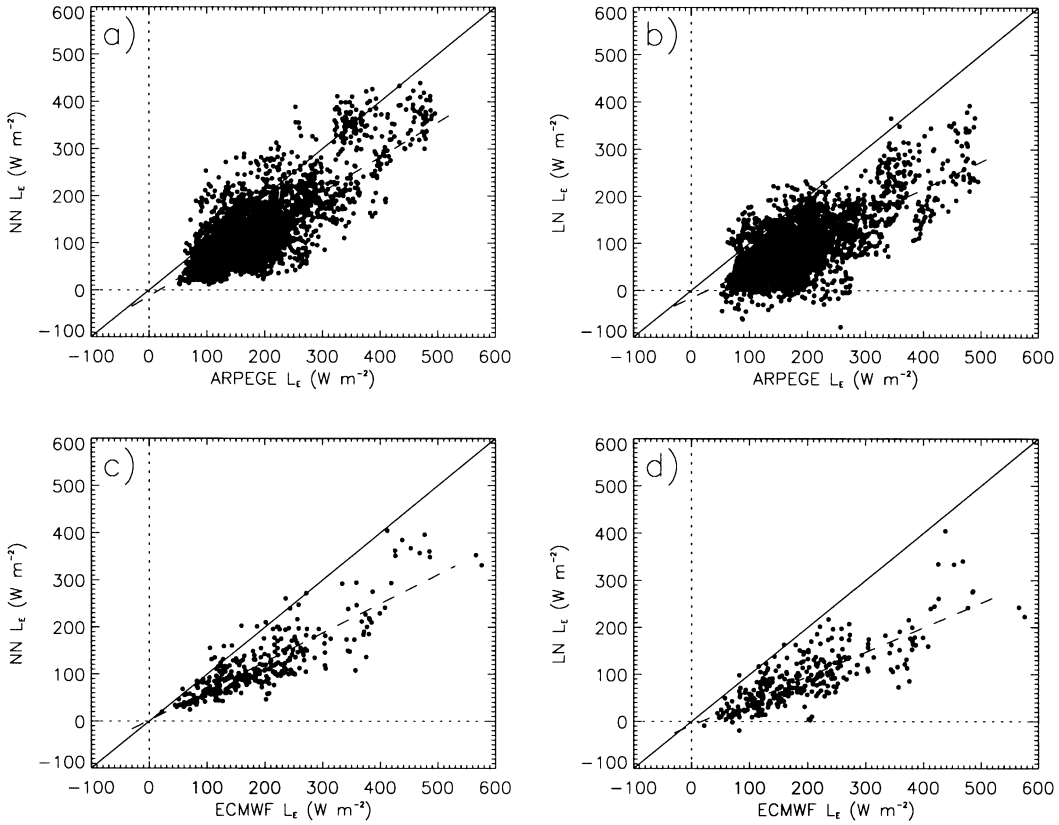

FIG. 4. Comparison between instantaneous fluxes from satellites and models. Satellite fluxes are (b),(d) LN and (a),(c) NN fluxes. The comparison is shown with respect to (a),(b) ARPEGE and (c),(d) ECMWF fields.

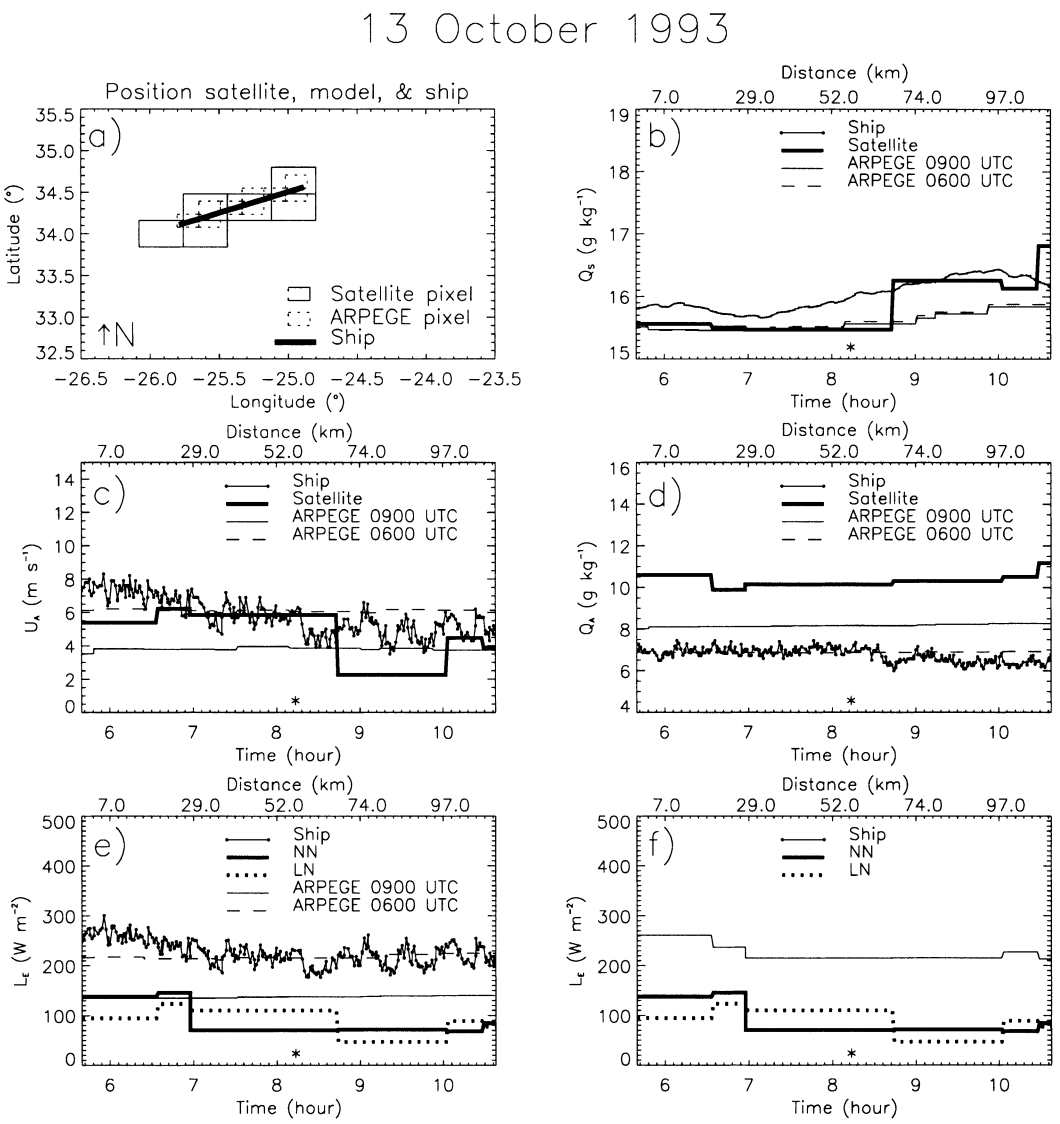

FIG. 5. Horizontal profiles of bulk variables and fluxes from satellites, ARPEGE, and the ship on 13 Oct 1993. (a) The path of the ship (thick line) and the corresponding selected satellite and model pixels (empty boxes). (f) Ship data were averaged over time intervals that match the resolution of the satellite fluxes. (b)–(f) Stars represent the hour of the satellite orbit.

fluxes. On the other hand, the ARPEGE fields were interpolated at the resolution of the SSM/I for the comparison with ARPEGE. For both datasets, the flux range is ~550 W m^{-2}, which is ~300 W m^{-2} more than in the comparison of section 4a. The correlation between satellite and model fluxes is 0.71–0.85, as reported in Table 2. This confirms that fluxes from models and satellites do not diverge dramatically from each other. The rms deviation between satellite and model fields, however, is 50–60 W m^{-2}, which is large. In order to analyze this rms deviation and deduce an estimate of the rms errors in satellite and model fluxes, we performed Monte Carlo simulations.

Based on the assumptions that the errors in satellite and model fluxes are random and follow normal distributions of same rms, one simulated the rms deviations between various artificial flux distributions. The process consists of the following steps. One builds two linear distributions of fluxes of 9000 elements ranging from 0 to 550 W m^{-2} each, to which separate random noises of same rms (R) are added. Note that the number of elements and flux range chosen here correspond to the features of the comparison dataset between satellite and ARPEGE fluxes. Next, one computes the rms deviation (R_0) between the two distributions. The experiment is repeated 3000 times, R being constant. Last, one computes the mean value (M_1) and rms deviation (R_1) from the series of R_0 obtained. The process described above is repeated for various R. Our simulations indicate that the average rms deviation M_1 between two linear dis-

TABLE 3. Comparison of spatial gradients of bulk variables and fluxes from SSM/I, ARPEGE, and the ECMWF model to ship measurements, on 13 Oct 1993. Here $\partial/\partial x$ is the spatial gradient expressed in units of considered variable per 100 km. The flux gradients are the slopes of the first-degree polynomial fit of either time series of ship data or spatial variations of model or satellite variables along the path of the ship.

Variable	Source	$\partial/\partial x$
μ_A (m s^{-1})	Ship	−3.03
—	ARPEGE 0600 UTC	−0.03
—	ARPEGE 0900 UTC	−0.02
—	SSM/I	−3.57
q_S (g kg^{-1})	Ship	0.75
—	ARPEGE 0600 UTC	0.43
—	ARPEGE 0900 UTC	0.48
—	AVHRR	1.08
q_A (g kg^{-1})	Ship	−0.75
—	ARPEGE 0600 UTC	0.05
—	ARPEGE 0900 UTC	0.21
—	SSM/I	0.16
L_E (W m^{-2})	Ship	−47
—	ARPEGE 0600 UTC	11
—	ARPEGE 0900 UTC	6
—	LN	−56
—	NN	−77

tributions that have random noises R of 35 W m^{-2} is 50 W m^{-2}. Because R_1 is negligible—that is, smaller than 1 W m^{-2}—50 W m^{-2} is an accurate estimate of the deviation. Because 50 W m^{-2} corresponds to the deviation observed between satellite and model fluxes, the simulations performed suggest that the rms errors in satellite and model fluxes found in section 4a were representative; that is, the rms errors in satellites and models fluxes are between 30 and 40 W m^{-2}.

Table 2 shows that the NN fluxes have a smaller rms deviation with respect to fluxes from models than the LN fluxes, by 1.4–7.2 W m^{-2}. Besides, the LN method produces a number of unrealistic negative fluxes because of the errors in satellite q_A and q_S (Fig. 4b). This happens when q_A is close to q_S, in which case any overestimation of q_A or underestimation of q_S may translate into negative fluxes, according to Eq. (1). This cannot happen with the NN fluxes, the training of which was done considering positive fluxes only. Over regions of large negative fluxes, this would obviously be a drawback of the NN method. The slope of the linear fit between the LN (NN) fluxes and the ARPEGE fluxes is 0.57 (0.74). As a result, the range of estimated fluxes is smaller than the range of model fluxes, like in section 4a. Note that the ARPEGE fluxes have a better fit to the satellite fluxes than the ECMWF fluxes in terms of slope of linear fit (Table 2), because the bias between satellite and ECMWF q_A (2.4 g kg^{-1}) is larger than the bias between satellite and ARPEGE q_A (2.2 g kg^{-1}).

5. Spatial variability of fluxes

In this section, one attempts to validate the spatial variations of satellite and model fluxes at the instantaneous scale. Next, as only two cases are available for validation, the spatial variability of the fluxes is further investigated at 1-month scale.

a. Instantaneous fluxes

Carefully selected ship and aircraft data were used for validation. For ship data it was necessary to assume Taylor's hypothesis, that is, the spatially averaged atmospheric variables from satellites or models correspond to time-averaged ship measurements. For this reason, all ship data available within ±150 min around the hours of the available satellite orbits were gathered. Next, the pixels of the satellite and ARPEGE fields the closest (<50 km) to the path of the ship were selected. Because the ship is a slow-moving platform, a number of cases were eliminated because they did not correspond to an extent in space of more than one satellite pixel. Cases for which the spatial gradients of fluxes were not significant (i.e., smaller than the rms error found in section 4) were also avoided. This procedure led us to select 13 October 1993 for validation against ship data, since the measured flux gradient was ~50 W m^{-2} (100 km)$^{-1}$ and the distance covered by the ship was ~100 km.

There was systematically a shift in time of more than 4 h between SSM/I orbits and *ARAT* cross sections. Therefore, data from these two platforms could not be compared in any cases but one, 16 November. This day, the environmental conditions were almost steady for more than 12 h because of the presence of a strong anticyclone (1025 hPa) SE of the region (35°N, 21°W). Subsequently, the spatial gradients of u_A and q_A along the aircraft cross section were positive and consistent from 0600 to 1800 UTC, according to ARPEGE (Fig. 6 and Table 4). The ECMWF analyses from 0600 to 1800 UTC confirm that at a larger scale the gradients of u_A and q_A were also all positive. Therefore, satellite data at 0729 UTC and 1831 UTC can be compared with *ARAT* measurements at 1400 UTC on 16 November.

1) CASE STUDY: 13 OCTOBER 1993

The magnitude of the flux gradient measured by the ship is −47 W m^{-2} (100 km)$^{-1}$ (Figs. 5e,f and Table 3). As 47 W m^{-2} is larger than the rms error in satellite and model fluxes found in sections 3–4 (30–40 W m^{-2}), the measured flux gradient should be detected by satellites and rendered by ARPEGE. Figures 5e,f reveal that the satellites indeed detect a decrease of the flux along the path of the ship. The estimated flux variation is −56 W m^{-2} (100 km)$^{-1}$ for the LN fluxes and −77 W m^{-2} (100 km)$^{-1}$ for NN fluxes, which is overestimated by 9–30 W m^{-2} with respect to ship measurements. On the other hand, ARPEGE sees a slightly positive flux gradient of 6 W m^{-2} (100 km)$^{-1}$ at 0900 UTC (3-h forecast), instead of a negative gradient.

Figures 5b–d show the horizontal gradients of the different bulk variables, which are of interest to understand the errors in satellite and model flux gradients. The gradients of q_S and u_A obtained from satellite are

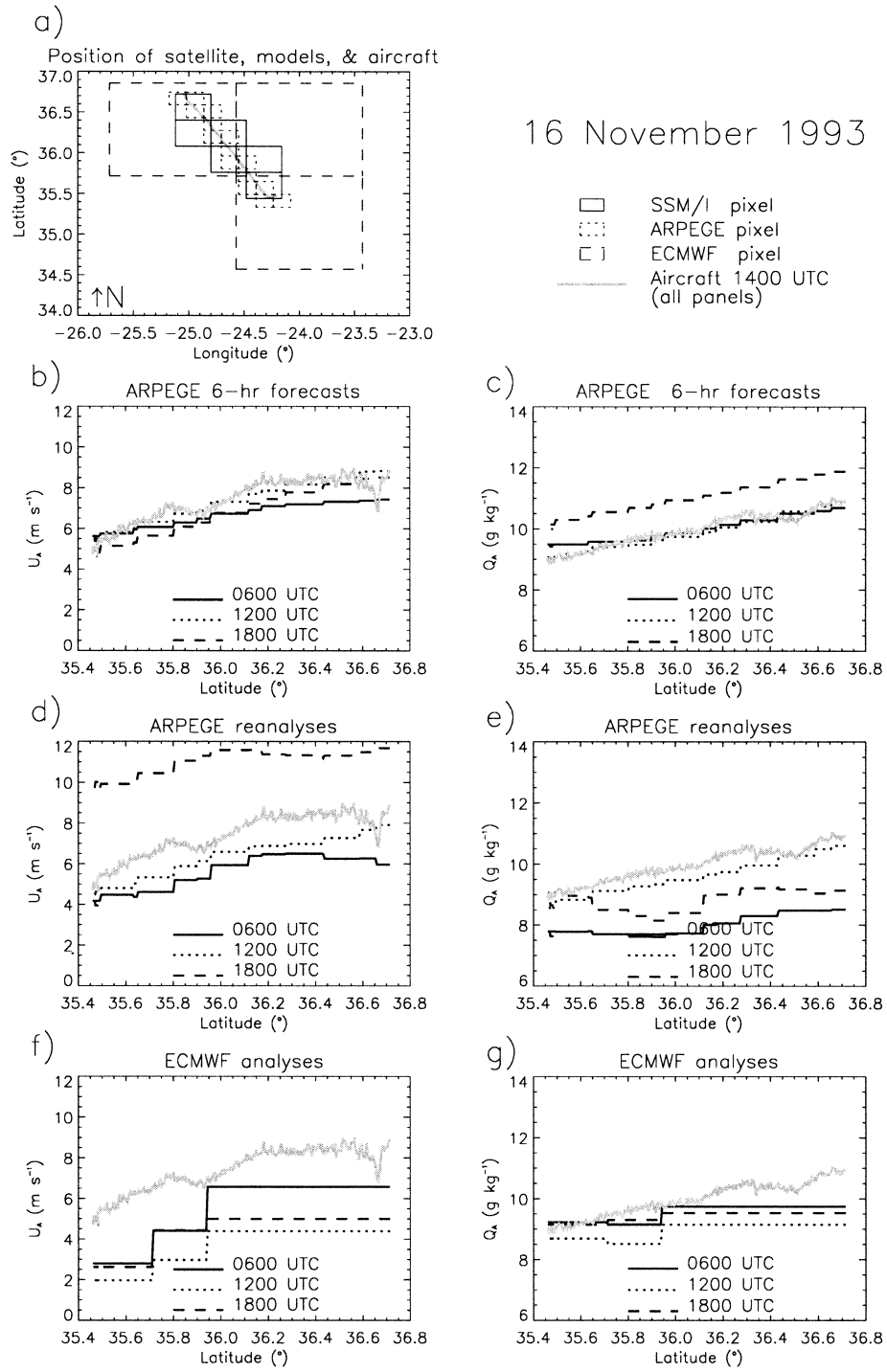

FIG. 6. Horizontal profiles of u_A and q_A from ARPEGE and the ECMWF model along the *ARAT* cross section of 16 Nov.

consistent with the in situ observations, as reported in Table 3. On the other hand, the SSM/I does not see any q_A gradient, although the measured q_A gradient is only 0.19 g kg^{-1} (or 20%) smaller than the rms error in satellite q_A. This has no consequence on the estimation of the flux gradient by the satellite, because the wind gradient is strong enough to induce an overall decrease of the flux along the path of the ship.

The ARPEGE q_S gradient at 0900 UTC [0.48 g kg^{-1} (100 km)$^{-1}$] has a good fit to the ship measurements [0.75 g kg^{-1} (100 km)$^{-1}$]. However, the gradients of ARPEGE u_A and q_A at 0900 UTC are negligible, which

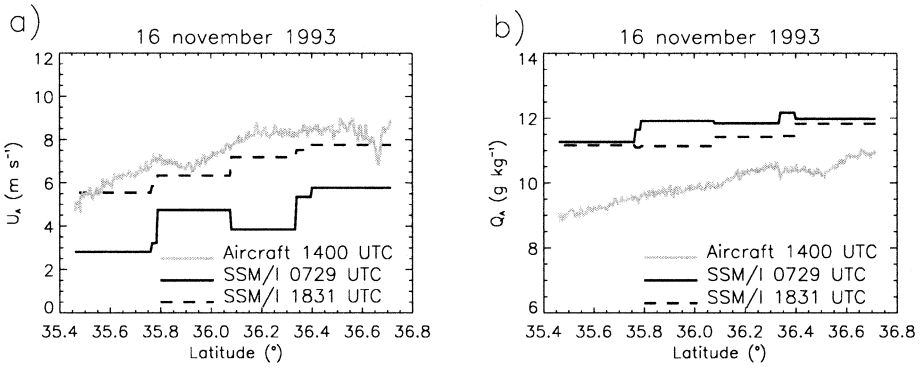

FIG. 7. Horizontal profiles of SSM/I-derived u_A and q_A along the *ARAT* cross section of 16 Nov.

explains why there is no negative gradient of ARPEGE flux at 0900 UTC. Because most of the SEMAPHORE data were assimilated into ARPEGE, one may have expected ARPEGE to render the measured gradients of u_A and q_A. The reasons are the following. First, ship data were assimilated at 0600 UTC only, which means that part of the ship observations (after 0600 UTC) were not/less accounted for in ARPEGE. Next, there were no other data assimilated in the close neighborhood of the ship at 0600 UTC. Last, the ship was close to the center of an anticyclone (1021 hPa at 37°N, 27°W), which possibly explains that humidity and wind gradients were consistent from 0600 to 0900 UTC.

2) CASE STUDY: 16 NOVEMBER 1993

According to Figs. 6 and 7 and Table 4, the SSM/I detects a positive gradient of u_A [1.91–2.21 m s^{-1} (100 km)$^{-1}$]. The gradients of ARPEGE u_A and aircraft u_A are 1.18–2.73 m s^{-1} and 2.32 m s^{-1} (100 km)$^{-1}$, respectively, which means that the satellite detects a wind gradient of 1.18–2.73 m s^{-1} (100 km)$^{-1}$. It is reasonable to give this result in terms of satellite pixels, that is, 0.35–0.82 m s^{-1} (30 km)$^{-1}$, since the gradient of satellite u_A is gradual along the aircraft cross section. This result means that the wind gradient is detected by the SSM/I, although it is smaller than the rms error in satellite u_A (0.98 m s^{-1}, section 4a).

The q_A gradients from ARPEGE 6-h forecasts and aircraft data are 0.9–1.20 and 1.27 g kg^{-1} (100 km)$^{-1}$, respectively, which is consistent (Fig. 6c). The SSM/I detects the increase in q_A but underestimates it by 0.7 g kg^{-1} (100 km)$^{-1}$ with respect to aircraft data (Fig. 7b). The ARPEGE reanalyses of q_A at 0600 and 1800 UTC were not considered in this comparison because their study revealed the presence of strong artifacts due to the assimilation of data (not shown). In Fig. 6d and Table 4, this translates into very small spatial gradients of reanalyzed q_A at 0600 and 1800 UTC [0.38–0.66 g kg^{-1} (100 km)$^{-1}$]. Note also that the analysis was not carried out in terms of fluxes for this case because it is questionable whether the bulk algorithm can be applied to measurements at 70 m, which was the altitude of the aircraft on 16 November.

TABLE 4. Spatial gradients of bulk variables from SSM/I, ARPEGE, the ECMWF model, and aircraft measurements on 16 Nov 1993. Here $\partial/\partial x$ is the spatial gradient expressed in units of considered variable per 100 km. The gradients correspond to the slope of linear fit to the data. The ARPEGE variables were interpolated at the resolution of the SSM/I fields. The resolution of the ECMWF products is 1.125° × 1.125°. The exact hours of the SSM/I observations are 0729 and 1831 UTC.

Variable	Source	$\partial/\partial x$	$\partial/\partial x$	$\partial/\partial x$	$\partial/\partial x$
Time (UTC)		0600	1200	1400	1800
μ_A (m s^{-1})	Aircraft			2.32	
—	ARPEGE reanalysis	1.71	2.22		1.18
—	ARPEGE 6-h	1.38	2.43		2.73
—	SSM/I	2.21			1.91
—	ECMWF	3.18	2.05		1.8
q_A (g kg^{-1})	Aircraft			1.27	
—	ARPEGE reanalysis	0.66	1.32		0.38
—	ARPEGE 6-h	0.9	1.24		1.20
—	SSM/I	0.57			0.56
—	ECMWF	0.51	0.48		0.32

b. Monthly fluxes

Thirty ARPEGE, ECMWF, and satellite fields of bulk variables and fluxes were averaged from 15 October 1993 to 15 November 1993. The comparison between these different fields is presented in Figs. 8 and 9 and Table 5. In Fig. 8, flux anomaly fields are presented instead of flux fields in order to highlight the spatial variations of the different fluxes. The rms deviations between satellite and model fluxes are 13–23 W m^{-2}. Despite these rms deviations being small, the horizontal structures of fluxes from models and satellites are not all identical (Fig. 8). In particular, satellite and ARPEGE flux fields reveal the presence of a couple of mesoscale circular structures corresponding to strong (34°N, 28°W) and weak fluxes (36°N, 24°W). These structures, referred to as A–B structure in the following, are not present in the ECMWF field, although they cover a 2° × 4° area. The A–B structure is related to the particular mesoscale configuration of q_S and u_A, which is shown

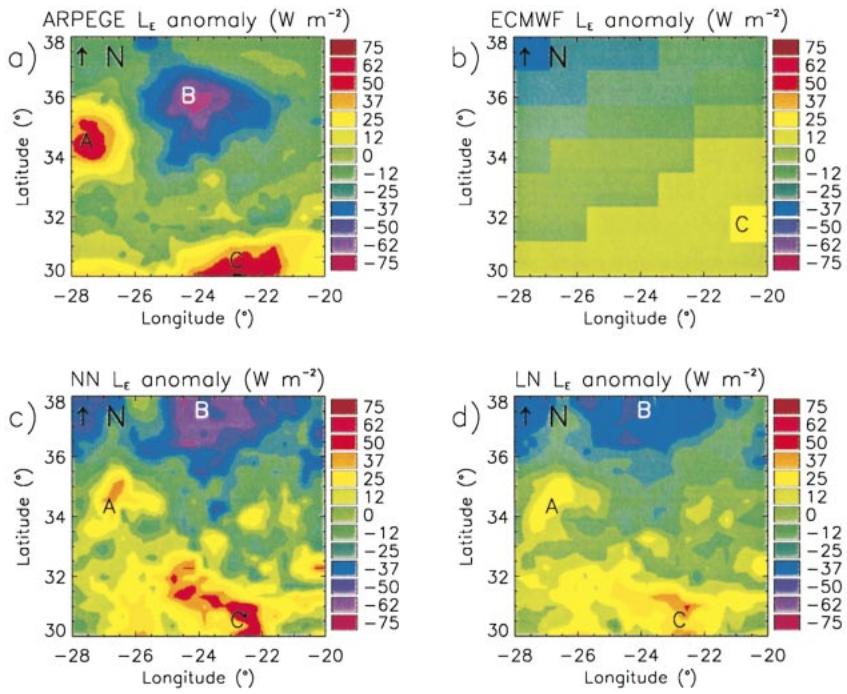

FIG. 8. Monthly flux anomaly fields from (a) ARPEGE, (b) the ECMWF model, and (c) and (d) satellites. The satellite fields were derived from a combination of AVHRR and SSM/I observations using the (d) Liu and Niiler (1984) method and (c) a neural network. Marks A, B, and C locate features of interest.

in Figs. 9a–f. Figures 9b,e reveal that the ECMWF q_S field almost resolves the A–B structure, while the ECMWF model does not render the associated wind structure. The reason is either a lack of assimilated data at this location or the use of too-smooth structure functions during the assimilation process (which set the actual resolution of the model). One common feature in the ARPEGE, ECMWF, and satellite fields is a structure of large fluxes at 31°N, 23°W (referred to as C in Figs. 8 and 9). This structure is related to large q_S values and winds stronger than the mean wind (Figs. 9a–f). Figures 8a,c reveal that the structure C present in the ARPEGE fields is shifted northward and deformed with respect to the satellite observations. On the other hand, structure C is reduced to one pixel in the ECMWF flux anomaly field, because the corresponding wind anomaly is weak (Figs. 8b and 9b). The LN and NN fluxes have comparable performances, as shown in Fig. 8c,d. However, the maximum LN flux anomaly (~90 W m^{-2}) is underestimated with respect to the maximum flux anomaly of the ARPEGE or NN fields (~140 W m^{-2}), because the bias in LN fluxes with respect to model fluxes is larger than the bias between NN fluxes and model fluxes (Table 5), like in section 4b.

6. Conclusions

This paper presented an attempt to validate latent heat flux fields from satellites and models in the context of SEMAPHORE. Two retrieval methods were used to derive flux fields from SSM/I and AVHRR observations. The satellite fluxes were compared with fluxes of the ECMWF operational model and the ARPEGE model, in which observations of the GTS as well as most of the SEMAPHORE data were assimilated.

Comparisons with individual ship measurements indicate that the rms error in satellite-derived fluxes is 35–40 W m^{-2}, which is close to the rms error in model fluxes (30–35 W m^{-2}). On the other hand, the bias in satellite fluxes is large (60–85 W m^{-2}) and implies that the range of estimated fluxes is smaller than the range of ship fluxes, by 34%–38%. Approximately 20 W m^{-2} of the bias in the satellite fluxes is caused by an inadequate description of the effects of variations in atmospheric stability on C_E, a situation that had not been fully elucidated by Bourras et al. (2002). However, most of the bias corresponds to a large bias in satellite q_A, which is related to the departure of the humidity profiles from those used in developing the retrieval methods. The biases in model fluxes are small (14–18 W m^{-2}) because models account for C_E and have small biases in q_A, u_A, and q_S.

The rms deviation between satellite and model fluxes is 50–60 W m^{-2}, which is large. However, this deviation corresponds to random errors in satellite and model fluxes of 35 W m^{-2} in rms, according to simple Monte Carlo simulations. This result suggests that the rms errors in satellite and model fluxes found with respect to ship

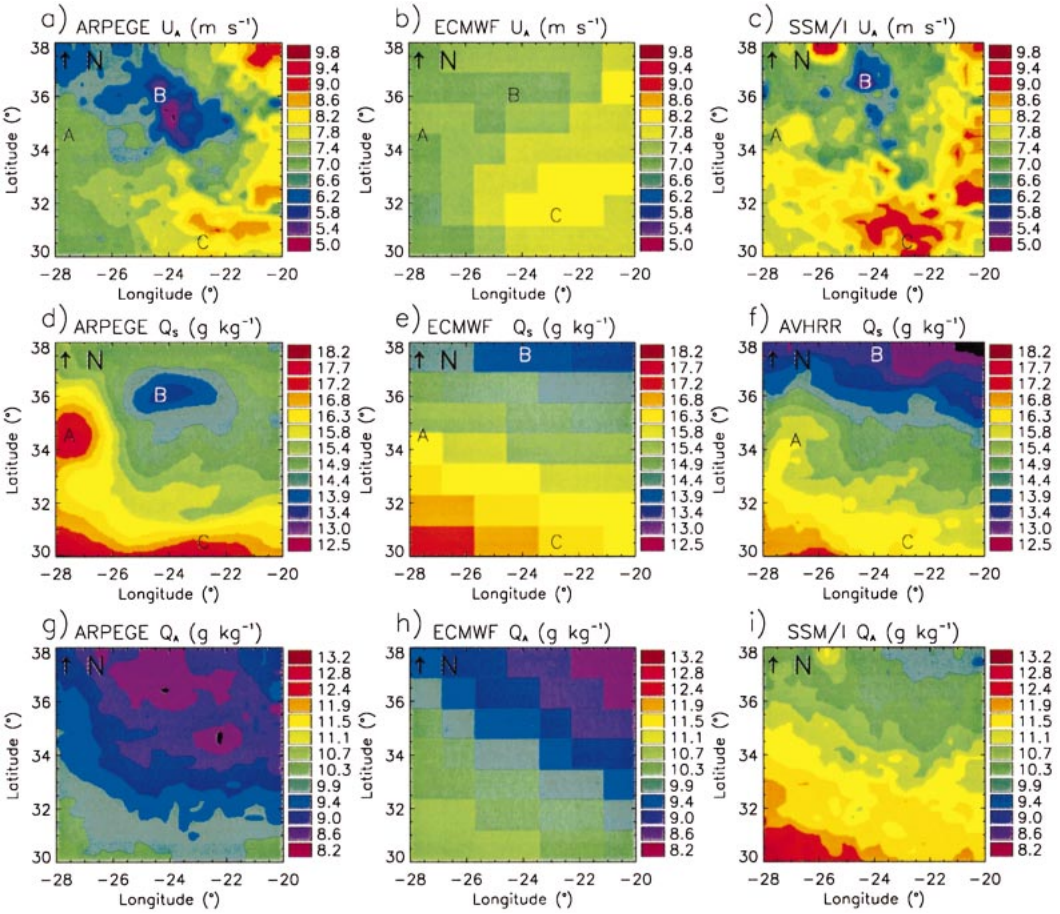

FIG. 9. Monthly fields of bulk variables from (a) ARPEGE, (b) the ECMWF model, and (c) and (d) satellites. The bulk variables are (a)–(c) u_A, (d)–(f) q_S, and (g)–(i) q_A. Marks A, B, and C locate features of interest.

data were representative. We therefore conclude that the rms accuracy of fluxes from satellite and models is 30–40 W m^{-2} or 20%–30% for SEMAPHORE, which does not include the error in reference fluxes. The performances of the two flux retrieval methods were similar in terms of rms and correlation. However, the bias in NN fluxes was smaller by 25–34 W m^{-2} than the bias in LN fluxes.

The spatial variations of fluxes and bulk variables from satellites and models were compared with ship and aircraft measurements on 13 October and 16 November. It was shown that the SSM/I detected—but underestimated—the spatial gradients of humidity if the magnitude of the gradients was larger than the rms error in q_A. The SSM/I detected gradients of u_A as small as 0.35–0.82 m s^{-1} (30 km)$^{-1}$, which is smaller than the rms error in u_A (0.98 m s^{-1}). In terms of fluxes, the satellite detected a gradient of -47 W m^{-2} (100 km)$^{-1}$ on 13 October, but overestimated it by 9–30 W m^{-2} (100 km)$^{-1}$. On the other hand, the accuracy of the gradients of fluxes and bulk variables from ARPEGE depended on the case under study. On 13 October, the gradients of u_A and q_A measured by the ship were not present in the ARPEGE fields, while ARPEGE had a good fit to

TABLE 5. Comparison between monthly SSM/I, ARPEGE, and ECMWF bulk variables and fluxes at 1.125° × 1.125° resolution. S-ECM (S-ARP) corresponds to the comparison between satellite and ECMWF (ARPEGE) fields.

	Correlation	Rms	Bias	Correlation	Rms	Bias
Fields compared	S-ECM	S-ECM	S-ECM	S-ARP	S-ARP	S-ARP
μ_A	0.62	0.56	0.16	0.77	0.50	0.58
q_S	0.94	0.34	−0.11	0.77	0.69	−0.36
q_A	0.93	0.27	1.50	0.89	0.35	1.81
L_E (LN)	0.74	13.0	−56.3	0.59	21.0	−58.8
L_E (NN)	0.66	19.7	−22.9	0.58	23.4	−25.2

aircraft measurements on 16 November. These results suggest that satellites detect flux gradients if the magnitude of the gradients exceeds a detection threshold, the accurate assessment of which will require further studies. In contrast, the accuracy of model fields always depends on the number and location of assimilated observations, as well as on the assimilation scheme used.

Monthly fluxes from models and satellites have a good fit in terms of rms deviation (13–23 W m^{-2}). Flux anomaly fields from ARPEGE and the satellites exhibit similar mesoscale structures, although the structures present in the ARPEGE fields are deformed and shifted with respect to the satellite observations. On the other hand, the ECMWF model does not resolve part of the mesoscale flux structures present in both ARPEGE and satellite fields. The reasons are the coarse resolution of the ECMWF model and possibly a lack of assimilated data. The fact that monthly anomaly fields of ARPEGE fluxes compare well to satellite observations is encouraging for further use of satellite fluxes, because reanalyses are generally not available to flux users.

Overall, our results show that satellites give useful results in terms of spatial variations of bulk variables and fluxes, but they have large biases during SEMAPHORE. On the other hand, fluxes from models have small biases but their accuracy in terms of spatial variations is fluctuating. Based on these conclusions, we would give the following directions to increase the accuracy of latent heat flux fields. In the long term, assimilation of satellite observations in ACMs, which is an ongoing effort at centers such as ECMWF, is a promising approach. For more immediate results and for regions where the conclusions of this paper apply, we would suggest to correct the bias in satellite fluxes with regional averages of q_A and C_E from in situ data, or model fields, for instance.

Acknowledgments. This study was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA). It was partly supported by the Physical Oceanography Program of NASA. The authors acknowledge the Satellite Active Archive at the National Oceanic and Atmospheric Administration, the JPL PO-DAAC, the ECMWF, IFREMER, P. Durand, and H. Giordani. The authors are grateful to P. Woiceshyn, D. Menemenlis, E. Armstrong, A. Bingham, and the *Journal of Applied Meteorology* reviewers for their helpful comments about this study.

REFERENCES

Armstrong, E. M., and J. Vazquez-Cuervo, 2001: A new global satellite-based sea surface temperature climatology. *Geophys. Res. Lett.,* **28,** 4199–4202.

Blanc, T. V., 1987: Accuracy of bulk-method-determined flux, stability and sea surface roughness. *J. Geophys. Res.,* **92,** 3867–3876.

Bourras, D., 2000: Calcul des flux turbulents au dessus des océans par la méthode bulk. Tech. Note RI-CETP/2/2000, Centre d'Etude des Environnements Terrestre et Planétaires, France, 54 pp. [Available from CETP, 10-12 av. de l'Europe, 78140 Vélizy, France.]

——, L. Eymard, and W. T. Liu, 2002: A neural network to estimate the latent heat flux over oceans from satellite observations. *Int. J. Remote Sens.,* **23,** 2405–2423.

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

Chou, S. H., R. M. Atlas, C. L. Shie, and J. Ardizzone, 1995: Estimates of surface humidity and latent heat fluxes over oceans from SSM/I data. *Mon. Wea. Rev.,* **123,** 2405–2425.

——, C. L. Shie, R. M. Atlas, and J. Ardizzone, 1997: Air–sea fluxes retrieved from Special Sensor Microwave Imager data. *J. Geophys. Res.,* **102,** 12 705–12 726.

Clayson, C. A., and J. A. Curry, 1996: Determination of surface turbulent fluxes for the Tropical Ocean–Global Atmosphere Coupled Ocean–Atmosphere Response Experiment: Comparison of satellite retrievals and in situ measurements. *J. Geophys. Res.,* **101,** 28 515–28 528.

Curry, J. A., C. A. Clayson, W. B. Rossow, R. Reeder, Y.-C. Zhang, P. J. Webster, G. Liu, and R.-S. Sheu, 1999: High-resolution satellite-derived dataset of the surface fluxes of heat, freshwater, and momentum for the TOGA COARE IOP. *Bull. Amer. Meteor. Soc.,* **80,** 2059–2080.

Dupuis, H., P. K. Taylor, A. Weill, and K. B. Katsaros, 1997: Inertial dissipation method applied to derive turbulent fluxes over the ocean during the Surface of the Ocean, Fluxes and Interactions with the Atmosphere/Atlantic Stratocumulus Transition Experiment (SOFIA/ASTEX) and Structure des Echanges Mer–Atmosphere, Properties des Heterogeneites Oceaniques: Recherche Experimentale (SEMAPHORE) experiments with low to moderate wind speed. *J. Geophys. Res.,* **102,** 21 115–21 129.

Esbensen, S. K., D. B. Chelton, D. Vickers, and J. Sun, 1993: An analysis of errors in Special Sensor Microwave Imager evaporation estimates over the global oceans. *J. Geophys. Res.,* **98,** 7081–7101.

Eymard, L., and Coauthors, 1996: Study of the air–sea interactions at the mesoscale: The SEMAPHORE experiment. *Ann. Geophys.,* **14,** 986–1015.

Gerard, E., and L. Eymard, 1998: Remote sensing of integrated cloud liquid water: Development of algorithms and quality control. *Radio Sci.,* **33,** 433–447.

Giordani, H., S. Planton, B. Bénech, and B. H. Kwon, 1998: Atmospheric boundary layer response to sea surface temperatures during the SEMAPHORE experiment. *J. Geophys. Res.,* **103,** 25 047–25 060.

Goodberlet, M. A., C. T. Swift, and J. C. Wilkerson, 1990: Ocean surface wind speed from the Special Sensor Microwave/Imager (SSM/I). *IEEE Trans. Geosci. Remote Sens.,* **28,** 823–828.

Kilpatrick, A., G. P. Podesta, and R. Evans, 2001: Overview of the NOAA/NASA Advanced Very High Resolution Radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *J. Geophys. Res.,* **106,** 9179–9197.

Large, W. G., and S. Pond, 1982: Sensible and latent heat flux measurements over the ocean. *J. Phys. Oceanogr.,* **12,** 464–482.

Liu, W. T., and P. P. Niiler, 1984: Determination of monthly mean humidity in the atmospheric surface layer over oceans from satellite data. *J. Phys. Oceanogr.,* **14,** 1451–1457.

Schluessel, P., L. Schanz, and G. Englisch, 1995: Retrieval of latent heat flux and long wave irradiance at the sea surface from SSM/I and AVHRR measurements. *Adv. Space Res.,* **16,** 107–116.

Schulz, J., J. Meywerk, S. Edwald, and P. Schlussel, 1997: Evaluation of satellite derived latent heat fluxes. *J. Climate,* **10,** 2782–2795.

Weare, B. C., 1989: Uncertainties in estimates of surface heat fluxes derived from marine reports over the tropical and subtropical oceans. *Tellus,* **41,** 1377–1416.

Webster, P. J., and R. Lukas, 1992: TOGA COARE: The Coupled Ocean–Atmosphere Response Experiment. *Bull. Amer. Meteor. Soc.,* **73,** 1377–1416.