A New Technique for Estimation of Surface Latent Heat Fluxes Using Satellite-Based Observations

Randhir Singh, P. C. Joshi, and C. M. Kishawal

Atmospheric Sciences Division, Meteorology and Oceanography Group, Space Applications Centre (ISRO), Ahmedabad, India

(Manuscript received 27 September 2004, in final form 31 January 2005)

ABSTRACT

Monthly mean surface latent heat fluxes (LHFs) over the global oceans are estimated using bulk formula. LHFs are computed using wind speed (\(U\)) from the Special Sensor Microwave Imager (SSM/I), sea surface temperature (SST) from the Advanced Very High Resolution Radiometer (AVHRR), and near-surface specific humidity. Near-surface specific humidity (\(Q_a\)) is estimated from SSM/I-observed precipitable water (\(W\)) and AVHRR-observed SST using a genetic algorithm (GA) approach. The GA-retrieved monthly mean \(Q_a\) has an accuracy of 0.80 ± 0.32 g kg\(^{-1}\) as compared with surface marine observations based on the Comprehensive Ocean–Atmosphere Data Set (COADS). The GA approach improves upon the surface specific humidity retrieval based on regression, the EOF approach, and is comparable to the artificial neural network technique.

The satellite-derived LHFs are compared with globally distributed surface marine observations to monthly averages of 1° × 1° latitude–longitude bins, during 1988–93. When GA-retrieved \(Q_a\) is used in the computation of satellite-derived latent heat fluxes (LHF\(_{GA}\)) the global mean rmse, bias, and correlation are 22 ± 8 W m\(^{-2}\), 5 W m\(^{-2}\), and 0.85, respectively, for monthly mean latent heat fluxes. The rmse in LHF are larger when \(Q_a\) is retrieved using regression and EOF approaches.

1. Introduction

Surface latent heat fluxes (LHFs) over the oceans represent an important aspect of the atmosphere–ocean interaction. LHFs have been a topic of active research in the past decades because they are required to study the heat budget of the upper ocean or to initialize ocean general circulation models. Also fluxes of energy, momentum, and moisture across the air–sea interface play an integral role in the earth’s climate. For example, the mean upper-ocean circulation is primarily driven by the winds, while evaporation from the sea surface and resulting release of latent heat drives much of the atmospheric circulation.

Conventionally, the LHF has been estimated from bulk formula that employs measurements of several kinds of physical variables. These physical variables are wind speed (\(U\)), sea surface temperature (SST), air temperature (\(T_a\)), saturation specific humidity (\(Q_s\)), and near-surface specific humidity (\(Q_a\)). Most of the studies carried out to estimate the evaporation fields on a global scale have used only historical ship reports (Bunker 1976; Hastenrath and Lamb 1977, 1979; Hsiung 1986). Since the temporal and spatial resolution of ship measurements is limited, their use in multiyear studies is doubtful (Simonot and Gautier 1989).

Alternative to using ship measurements is the use of the satellite observations. With the advent of satellite technology there are unique and complementary ways to compute the LHF. Satellite observations have high spatial and temporal resolution and hence are best suited for studying the large-scale phenomena. The satellite observations have now become more and more realistic owing to several improvements in sensors, their calibrations, and validations. Due to these advances in satellite observations, the combined use of various kinds of geophysical parameters observed by satellite is also possible. But in this case the near-surface specific humidity (\(Q_a\)), and air temperature (\(T_a\)) needed to compute LHF are not directly accessible by satellite measurements. Recently several global surface heat flux datasets have been derived from satellite data. These include the Japanese Ocean Flux datasets with...
Use of Remote sensing Observation data (J-OFURO; Kubota et al., 2002), the Hamburg Ocean–Atmosphere Parameters and fluxes from satellite data (HOAPS; Grabli et al. 2000), and version 2 of the Goddard Satellite-Based Surface Turbulent Fluxes (GSSTF2; Chou et al. 2003). Numerous investigators (Liu 1988; Eymard et al. 1989; Simonot and Gautier 1989; Esbensen et al. 1993; Jourdan and Gautier 1995, hereafter JG95; Chou et al. 1995; Schulz et al. 1997; Chou et al. 2003; Singh 2004) have looked into the problem of remote sensing of LHF. Most of these studies have demonstrated that application of Liu’s (1986, hereafter L86) empirical relationship between monthly averaged columnar water vapor (W) and near surface specific humidity (Qs) to SSM/I data gives very large spatial biases of LHF estimates.

To improve satellite estimation of LHF we have developed a method based on the genetic algorithm (GA) to retrieve Qs. The main objective of this paper is to describe the GA and to assess the impact of Qs derived from GA on global LHF estimations. The paper is organized as follows. In section 2, we present the datasets used. Section 3 introduces the method used to compute the LHF. Section 4 compares the accuracy of Qs retrievals. This is followed by intercomparison of satellite and in situ monthly mean LHF fields (section 6) and a discussion on the spatial and temporal variability of the LHF. Finally, conclusions are given in section 7.

2. Datasets

a. Surface marine data

In the present study we have used near-surface specific humidity (Qs), wind speed (U), and SST observations from the surface marine data (SMD; da Silva et al. 1994) to develop and evaluate the new methodology. The SMD data available at 1° × 1° latitude–longitude resolution for the period 1988–93. The SMD is based on ship reports compiled in the Comprehensive Ocean–Atmosphere Data Set (COADS). The SMD however has stricter quality-control criteria than COADS and applies bias corrections for several surface observations. A large number of studies (Chou et al. 1997, hereafter CSAA; Gautier et al. 1998; Jones et al. 1999) have been carried out using these data.

b. Satellite data

The satellite data used in the present study includes observations from the National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR), and the Special Sensor Microwave Imager (SSMI/I) on board the Defense Meteorological Satellite Program (DMSP) satellite. AVHRR Pathfinder (Brown et al. 1993) data consists of daily fields of gridded SST with a spatial resolution of 54 km² with data gaps over cloudy regions. We computed monthly averages of SST fields at 1° × 1° latitude–longitude resolution for the period 1988–93, using the daily values from Pathfinder data. Further the vertically integrated water vapor (W) and wind speed (U) fields are part of the global SSM/I ocean products (Wentz 1997). Monthly values were obtained by averaging daily values at 1° × 1° latitude–longitude grids for the period 1988–93.

To compare GA-retrieved Qs, with other existing Qs products we have used the Qs products of CSAA and Schlusse et al. (1995, hereafter SSE). These two products are available through GSSTF2 and HOAPS.

3. Methodology

The most common way to compute the LHF at the ocean–atmosphere interface is to use the aerodynamic bulk formula:

\[ \text{LHF} = \rho L C_L U (Q_s - Q_a), \]  

where ρ is the air density, L is the latent heat of evaporation, \( C_L \) is bulk transfer coefficient for water vapor (also called the Dalton number), U is the wind speed at a height of typically 10 m, \( Q_s \) is the saturation specific humidity at the surface, and \( Q_a \) is the near-surface specific humidity at the atmospheric measurement level (~10 m). Thus, the calculation of LHF using Eq. (1) implies knowledge of three key variables:

1) Surface wind speed (U) at 10-m height: In the present study it is derived from SSM/I measurements as discussed in the previous section.

2) Saturation specific humidity (Qs): The saturation specific humidity, Qs, is calculated from SST assuming saturation at the surface. The LHF depends on the skin SST. However, the AVHRR SST is derived from empirical relation developed with regressing the AVHRR radiance with the bulk SST (Chou et al. 2003). To compensate for the cool skin effects on LHF, Qs is estimated using the approximated (Chou et al. 2003) formula as

\[ Q_s = (0.622 e_s, P^{-1}) \]  

Here \( e_s \) is the saturation vapor pressure for pure water at the bulk SST, and P is the sea level pressure. The error due to the uncertainty of sea level pressure is on the order of 1% and could be neglected (Schulz et al. 1997). Further, \( Q_s \) is reduced by 2% to account for salinity effects (Bentamy et al. 2003) and is derived from

\[ Q_s = 0.98(0.622e_s/P). \]  


3) Near-surface specific humidity ($Q_a$): Several authors have investigated the estimation of $Q_a$ from satellite measurements. Use of total precipitable water ($W$) as an independent parameter is a common way to determine the near-surface specific humidity from satellite measurements. The correlation between these quantities depends heavily on the time scale considered. The method of L86 determines the monthly marine surface layer humidity with a simple polynomial regression of $Q_a$ versus $W$. This simple formula can be used with any retrieval algorithm that determines $W$. Unfortunately, errors in $Q_a$ originating from the $Q_a$–$W$ relation can result in large errors in the LHF. This was shown by Esbensen et al. (1993), who compared 1 yr of LHF derived from satellite data using the $Q_a$–$W$ relation and in situ estimates from COADS observations. The above study showed systematic error of over 2 g kg$^{-1}$ in the satellite estimates (L86) of monthly averages $Q_a$.

Schulz et al. (1993) developed a new method that first derives the integrated water vapor of the atmospheric boundary layer $W_b$ (using some artificial height of 500 m) and then deduces $Q_a$ with a simple linear regression from $W_b$. They showed that $W_b$ can be retrieved independently of $W$ and the correlation between $Q_a$ and $W_b$ is much higher than with $W$. The accuracy achieved for this retrieval was 1.7 g kg$^{-1}$ on the instantaneous time scale. SSE improved this technique slightly by obtaining $Q_a$ directly from brightness temperature, thus avoiding the error propagation that occurs in two-step methods. The standard error for this globally valid retrieval was 1.6 g kg$^{-1}$ for an instantaneous measurement. An alternative approach has been reported by CSAA. CSAA estimated $Q_a$ from the total precipitable water $W$ and boundary layer water vapor content $W_b$. Here $W$ was retrieved with the algorithms of Wentz (1989) and $W_b$ was retrieved with the algorithm of Schulz et al. (1993). The accuracy attained for $Q_a$ was not much different from Schulz et al. (1993). A recent paper by Jones et al. (1999) tried to retrieve monthly mean $Q_a$ from SSM/I measurements of total precipitable water, $W$, and SST analysis from the National Centers for Environmental Prediction (NCEP) using Artificial Neural Network (ANN) techniques. The global mean root-mean-square error (rmse) by the above technique was stated to be 0.77 ± 0.39 g kg$^{-1}$ for monthly mean $Q_a$.

To improve the satellite estimates of LHF, we have used GA to retrieve $Q_a$, SST from AVHRR and $W$ from the SSM/I sensor are used as input to the GA. The detail of the methodology is described in section 4.

Another problem with the bulk formulas is the choice of values for the coefficient $C_E$, which vary with the wind speed and stability of the atmosphere (Large and Pond 1982; Smith 1988). Based on the behavior of $C_E$ as a function of wind speed and air–sea temperature difference and assuming slightly unstable stratification ($T_a = SST - 1.25$ K) over the global oceans, along with the Hasse and Smith (1997) results, the $C_E$ is computed using the following fit (Bentamy et al. 2003):

$$10^3 C_E = a \exp[b(U + c)] + d/U + 1,$$

where $a = -0.146785$, $b = -0.292400$, $c = -2.206648$, $d = 1.611229$, and $U$ is the wind speed (m s$^{-1}$). The possible error due to the assumption of unstable stratification is generally less than 2.5% (Bentamy et al. 2003).

4. Algorithm development for $Q_a$ retrieval

a. Genetic algorithm

Modeling natural phenomena has been a standard practice in atmospheric sciences. Traditionally, modeling a physical system requires one to derive the relevant equations from first principles taking into account the physical laws that determine the system under consideration. Alternatively, when such an approach is not feasible due to some reasons, for example, when the perfect physics of the system is not well understood, or the required computing resources are not available, empirical laws governing the physical processes can be obtained by model-fitting approaches based on the observed variability of the system and its relationship with other parameters. Let us assume that there exists a smooth mapping function $P(.)$ that explains the relationship between a predicted variable $x$ and a set of dependent variables $[a, b, c, d, e, \ldots]$, so that

$$x = P[a, b, c, d, e, \ldots].$$

In various areas of geophysics, it is a standard practice to employ linear or nonlinear regression techniques in order to obtain the form of $P(.)$. However, the choice of a regression model is quite subjective, and it is difficult to ensure that a particular regression model provides the best possible explanation to the variance in a dataset. This has inspired the researchers to look for more objective data-fitting approaches like the GA. A GA is programmed to approximate the equation, in symbolic form, that best describes the relationship between independent and dependent parameters. The genetic algorithm considers an initial population of potential solutions, that is subjected to an evolutionary process, by selecting those equations (individuals) that
best fit the data. The strongest strings (made up from a combination of variables, real numbers, and arithmetic operators) choose a mate for reproduction whereas the weaker strings become extinct. The newly generated population is subjected to mutations that change fractions of information. The evolutionary steps are repeated with the new generation. The process ends after a number of generations, determined a priori by the user. The procedural details of genetic algorithm can be found in Szpiro (1997) and Alvarez et al. (2000). A brief description of genetic algorithm is as follows.

First, for a desired function \( f(x) \) [e.g., in Eq. (4)], a set of candidate equations for \( P(L) \) are randomly generated. An equation is stored in the computer as a set of characters that define the independent variables, \( a, b, c, d, e \ldots \) in Eq. (4), and four elementary arithmetic operators (+, −, ×, and ÷). A criterion that measures how well the equation strings perform on a training set of the data is its fitness to the data, defined as sum of the squared differences between data and independent parameter derived from the equation string. The strongest individuals (equations with best fits) are then selected to exchange parts of the character strings between them (reproduction and crossover) while individuals less fitted to the data are discarded. Finally, a small percentage of the equation strings’ most basic elements, single operators and variables, are mutated at random. The process is repeated a large number of times to improve the fitness of the evolving population of equations. The fitness strength of the best scoring equation is defined as

\[
R^2 = 1 - \frac{\Delta^2}{\sum (x_0 - \langle x_0 \rangle)^2},
\]

where \( \Delta^2 = \sum (x - x_0)^2 \), \( x \) is the parameter value estimated by the best scoring equation, \( x_0 \) is the corresponding “true” value, \( \langle x_0 \rangle \) is the mean of the true values of \( x \). Szpiro (1997) has shown the robustness of the GA to forecast the behavior of the one-dimensional chaotic dynamical system. Later, Alvarez et al. (2000) applied the GA to real physical systems and used this algorithm for the prediction of space–time variability of the SST in the Alboran Sea (the western Mediterranean Sea). Kishtawal et al. (2003) used GA for the prediction of seasonal monsoon rainfall over the Indian region.

b. Training

The development of the retrieval algorithm involves a number of steps. In the first step, all the available data (1988–93) of monthly mean \( W \) from SSM/I, and SST from AVHRR, and \( Q_a \) from SMD were divided into two subsamples. The partitioning of the data into training and validation was done in the same manner as Jones et al. (1999) who chose 2 yr belonging to contrasting climatic regimes (El Niño and La Niña) for validation of their algorithm. The 47-month period including January–June 1988, July 1989–June 1992, and July–November 1993 was designated as Sample I and is used to train the GA. The 24-month period of July 1988–June 1989 and July 1992–June 1993 was designated as Sample II and is used for evaluation of the retrieval algorithm.

Because the number of surface marine observations varies significantly in space and time, it is necessary to ensure that high-quality observations of \( Q_a \) are used to develop and evaluate the empirical relationship. Robust and statistically meaningful estimation of monthly mean \( Q_a \) requires that more than 20 observations (\( N \)) per month be used for averaging (e.g., Luther and Harrison 1984). We considered only those bins in the analysis for which at least 20 samples were used to produce monthly averages \( Q_a \) in SMD. As expected, such bins are colocated with major shipping routes. Furthermore, for training we randomly selected 20 000 points out of \( \approx 300 \, 000 \) points from Sample I data. The GA has the capability to learn the process from relatively smaller datasets compared to other optimization techniques (Szpiro 1997).

5. Assessment of retrieval accuracy of \( Q_a \)

a. Statistical comparison

The GA produces the following empirical equation for the computation of \( Q_a \):

\[
Q_a = \frac{a}{W + b} \left[ cWST + dW + \frac{eSST}{fW} - 7.55 \right] - 9.74 + 8.94,
\]

where SST and \( W \) represent the SST (°C) and vertically integrated water vapor (gm cm\(^{-2}\)), respectively. The values of constants (\( a, b, c, d, e, f \)) are given in Table 1. Figure 1a shows the population of observation points for different combinations of SST and \( W \), in a combined dataset (Samples I and II). Figure 1a shows that \( W \) ranges from 0 to 6 gm cm\(^{-2}\), while SST ranges from 0°

<table>
<thead>
<tr>
<th>Symbol</th>
<th>( a )</th>
<th>( b )</th>
<th>( c )</th>
<th>( d )</th>
<th>( e )</th>
<th>( f )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>1.0 g(^2) cm(^{-2}) kg(^{-1})</td>
<td>2.4 g cm(^{-2})</td>
<td>1.0 cm(^2) g(^{-1}) 0°C(^{-1})</td>
<td>−8.5 cm(^2) g(^{-1})</td>
<td>1.0 °C(^{-1})</td>
<td>1.0 cm(^2) g(^{-1})</td>
</tr>
</tbody>
</table>

TABLE 1. Value of constants.
The empirical Eq. (6) can be expected to be valid in this range of the SST and \( W \). Figure 1b shows the variation of \( Q_a \) as a function of SST (°C) and \( W \) (gm cm\(^{-2}\)) based on the empirical equation (6).

![Figure 1](image-url)  

**Fig. 1.** (a) The population of observation points for different combinations of SST (°C) and \( W \) (gm cm\(^{-2}\)) in the combined dataset (Samples I and II). (b) The variation of \( Q_a \) (g kg\(^{-1}\)) as a function of SST (°C) and \( W \) (gm cm\(^{-2}\)) based on the empirical equation (6).
mates of $Q_a$. The rmss between SMD and SSE are within 0.6 and 1.2 g kg$^{-1}$, with a large value (more than 1.5 g kg$^{-1}$) over the equatorial region, the Arabian Sea, the Bay of Bengal, the western Pacific Ocean, and over the Southern Hemisphere south of 30°S. The global mean rmss between SMD and SSE for monthly mean $Q_a$ is 1.2 ± 0.50 g kg$^{-1}$. The rmss between SMD and CSAA are between 0.9 and 1.5 g kg$^{-1}$ with a higher error (more than 1.5 g kg$^{-1}$) over the Southern Hemisphere south of 30°S, the Arabian Sea, the Bay of Bengal, and over the eastern equatorial Pacific Ocean. The global mean rmse between SMD and CSAA for monthly mean $Q_a$ is 1.4 ± 0.64 g kg$^{-1}$. Over a large part of the global oceans the rmse between SMD and L86 is between 0.9 and 1.5 g kg$^{-1}$. The large rmss (more than 1.5 g kg$^{-1}$) are found to the east of South America, South Africa, the Arabian Sea, the Bay of Bengal, the South Indian Ocean, and over the northwestern Atlan-

**Fig. 2.** Differences between observed and retrieved $Q_a$ (g kg$^{-1}$) for four methods: (a) GA, (b) SSE, (c) CSAA, and (d) L86 during the Sample II data.
tic and Pacific Oceans. The global mean rmse between SMD and L86 for monthly mean $Q_a$ is 1.7 ± 0.44 g kg$^{-1}$. The accuracy of retrieved $Q_a$ from the present algorithm is comparable to that obtained by the ANN techniques (Jones et al. 1999). Jones et al. (1999) took into account the mean seasonal bias along with $W$ and SST to retrieve $Q_a$. So, his method requires an extra parameter, which is not available directly from satellite observations. The present method uses only satellite-observed parameters, making it more convenient to use.

Further to visualize the advantage of the proposed algorithm (GA) over other algorithms, we have computed improvement parameter “η” which is defined as
\( \eta = \text{abs (other algorithm error)} - \text{abs (GA error)} \) for every grid point. The distribution of \( \eta \) is given in Fig. 4. The positive values of \( \eta \) indicate real improvement. With respect to SSE (Fig. 4a), GA shows large improvement over the Southern Hemisphere, the equatorial region of the western Pacific Ocean, the Arabian Sea, and over the Indian Ocean. The GA shows improvement almost over all the oceans when compared to CSAA (Fig. 4b) and L86 (Fig. 4c). The improvements are quite large particularly over the eastern boundaries of major continents. The percentage of grid points where \( \eta \) is positive are 74%, 94%, and 72% for SSE, CSAA, and L86, respectively.

b. Spatial and temporal variability in GA-retrieved \( Q_a \)

The previous section provides statistical results that showed the potential of the GA technique. In this sec-
tion, the large-scale features and space–time variability of GA-retrieved $Q_a$ are compared with surface marine observations. Equation (6) was applied to the entire record (Samples I and II) of $W$, the SST dataset (1988–93) to estimate $Q_a$.

1) Spatial Variability

The maps of mean $Q_a$ for January and July are presented in Figs. 5 and 6. Figs. 5a and 6a are obtained using satellite data (the GA algorithm), Figs. 5b and 6b are obtained with SMD data, and Figs. 5c and 6c are obtained by SMD – GA. These figures show the monthly maps for January and July averaged over the period 1988–93. The qualitative agreement between the maps is excellent and the location of characteristics features is well produced by GA in both the months. These maps show high concentration of humidity in the intertropical convergence zones and in the South Pacific convergence zone. High humidity is also observed over

![Fig. 5. Monthly averaged $Q_a$ (g kg$^{-1}$): (a) GA, (b) SMD observed, and (c) SMD – GA retrieved, during Jan 1988–93.](image-url)
the Maritime Continent in the Indonesian area and the subtropical regions of strong poleward moisture advection off the east coasts of the continents, especially noticeable in the summer hemisphere. Broad dry zones stretch from the eastern part of the subtropical high pressure zones toward the equatorial zones. Over most of the oceans the differences (Figs. 5c and 6c) between SMD-observed and GA-retrieved $Q_a$ are within $\pm 1$ g kg$^{-1}$. The differences are slightly higher over the northwestern Pacific and Atlantic Oceans during wintertime. The reasons for these large differences over these regions were discussed earlier.

2) TEMPORAL VARIABILITY

A good way to check the performance of the GA retrieval is a comparison to in situ time series in different geographical locations. Figure 7 shows the annual variation of $Q_a$ averaged over the Arabian Sea (10°–20°N, 60°–70°E), the northwestern Pacific (35°–45°N,
160°–170°E), and the northwestern Atlantic Ocean (35°–45°N, 25°–35°W). Annual variation of $Q_a$ derived by other methods is also presented in this figure for comparison. A summary of these comparisons is given in Table 2. These regions are chosen because in situ sampling is very high (see JG95 for more details) over these regions.

(i) Arabian Sea

The satellite-retrieved $Q_a$ by CSAA, SSE, and L86 have large positive bias in comparison to the SMD observations, particularly during wintertime (as shown in Fig. 7a). Maximum (3–4 g kg$^{-1}$) bias is observed in L86 retrieved $Q_a$. Compared to SMD observation, GA also underestimates $Q_a$ through out the year by 0.5–0.8 g kg$^{-1}$. Table 2 shows the statistical results of these algorithms.

(ii) Northwesern Pacific Ocean

Compared to SMD during wintertime (Fig. 7b), the GA- and CSAA-retrieved $Q_a$ have negative (0.5–1 g kg$^{-1}$) and positive (1–2 g kg$^{-1}$) biases, respectively. Large differences (3–4 g kg$^{-1}$) between $Q_a$ of L86 and SMD have been observed during the summer season. Over the northwestern Pacific Ocean (Table 2) SSE shows better results compared to other algorithms.

(iii) Northwesern Atlantic Ocean

Over the North Atlantic Ocean (Fig. 7c) good agreement between GA-retrieved $Q_a$ and SMD observation over the entire year is observed. CSAA and L86 also agree well with SMD-observed $Q_a$, but not better than GA. SSE and GA provide comparable estimates.

In summary GA provides superior accuracy of $Q_a$ retrieval over the Arabian Sea, while over other regions selected for comparison, the performance of GA is better than CSAA and L86 while it is comparable to SSE.

6. Satellite-derived LHF

In this section, we present and compare the satellite-derived monthly mean LHF with observed fluxes, computed using SMD data, through bulk formula [Eq. (1)]. There are two ways to compute monthly averages of LHF. One is by applying daily averages of basic parameters to bulk formula and then get monthly mean LHF by averaging daily values. Another is by applying monthly averaged basic parameters to the bulk formula. Some studies (Chou et al. 1997; Liu 1988; Bates 1991) indicate that differences are negligible in monthly fluxes computed by the above two methods. However Mahrt (1987) suggested some caution about the use of spatiotemporal averages of basic parameters for flux computation using the bulk formula because of the nonlinearity of the bulk coefficient. Four (GA, SSE, CSAA, and L86) estimates of $Q_a$ are used to obtain the LHF from satellite measurements. The LHF obtained using GA-retrieved $Q_a$ is termed as LHF$_{GA}$. To assess the improvement resulting from GA-retrieved $Q_a$ on satellite-derived LHFs, we calculated the LHFs from satellite observations again by replacing GA-retrieved $Q_a$ with other (SSE, CSAA, and L86) satellite estimates of $Q_a$ in Eq. (1). The rest of the parameters in Eq. (1) are same as they are in case of LHF$_{GA}$ computation. We used $Q_a$ of L86 in Eq. (1) to compute the LHF. This

### Table 2. Statistics of retrieved $Q_a$.

<table>
<thead>
<tr>
<th>Regions</th>
<th>Source</th>
<th>Correlation</th>
<th>Rmse</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabian Sea</td>
<td>GA</td>
<td>0.94</td>
<td>0.55</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>SSE</td>
<td>0.91</td>
<td>0.77</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>CSAA</td>
<td>0.91</td>
<td>1.36</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>L86</td>
<td>0.83</td>
<td>1.70</td>
<td>1.20</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>0.99</td>
<td>0.65</td>
<td>-0.60</td>
</tr>
<tr>
<td>Northwestern Ocean</td>
<td>SSE</td>
<td>0.99</td>
<td>0.59</td>
<td>0.24</td>
</tr>
<tr>
<td></td>
<td>CSAA</td>
<td>0.97</td>
<td>0.97</td>
<td>0.80</td>
</tr>
<tr>
<td></td>
<td>L86</td>
<td>0.97</td>
<td>2.20</td>
<td>-1.30</td>
</tr>
<tr>
<td></td>
<td>GA</td>
<td>0.99</td>
<td>0.52</td>
<td>-0.50</td>
</tr>
<tr>
<td>Atlantic Ocean</td>
<td>SSE</td>
<td>0.99</td>
<td>0.51</td>
<td>-0.34</td>
</tr>
<tr>
<td></td>
<td>CSAA</td>
<td>0.99</td>
<td>0.73</td>
<td>-0.40</td>
</tr>
<tr>
<td></td>
<td>L86</td>
<td>0.99</td>
<td>0.76</td>
<td>-0.44</td>
</tr>
</tbody>
</table>
is termed as \( \text{LHF}_{\text{L86}} \). The LHF computed using SSE and CSAA \( Q_a \) are termed as \( \text{LHF}_{\text{SSE}} \) and \( \text{LHF}_{\text{CSAA}} \), respectively. The SMD reported \( U, \text{SST}, \) and \( Q_a \), have been used to derive SMD latent heat flux (\( \text{LHF}_{\text{SMD}} \)). The exchange coefficients \( C_E \) computed using Eq. (3) is shown in Fig. 8. The \( C_E \) values range (Fig. 8) between 0.00152 and 0.00105 for wind speeds between 2 and 19 m s\(^{-1}\).

a. Statistical comparison

Figure 9 shows the differences between SMD and satellite-derived LHF for four methods: (Fig. 9a) \( \text{LHF}_{\text{GA}} \), (Fig. 9b) \( \text{LHF}_{\text{SSE}} \), (Fig. 9c) \( \text{LHF}_{\text{CSAA}} \), and (Fig. 9d) \( \text{LHF}_{\text{L86}} \), during the Sample II data period. Over most of the places the differences between \( \text{LHF}_{\text{SMD}} \) and \( \text{LHF}_{\text{GA}} \) are within \( \pm 20 \) W m\(^{-2}\). The higher differences (more than \( 20 \) W m\(^{-2}\)) occur over the western boundary currents, like the Gulf Stream and the Kuroshio Current. As we have seen earlier (Fig. 2a), over these regions the GA-retrieved \( Q_a \) is higher (1–2 g kg\(^{-1}\)) than SMD resulting in underestimation of \( \text{LHF}_{\text{GA}} \). Generally the differences between SMD and SSE are also within \( \pm 20 \) W m\(^{-2}\). Higher differences (more than \( 20 \) W m\(^{-2}\)) are found over the equatorial region, the central Pacific Ocean, and over the northwestern Atlantic Ocean. Over the equatorial region the SSE overestimates the \( Q_a \) by 1–2 g kg\(^{-1}\) as compared to SMD (Fig. 2b). This overestimation of \( Q_a \) by SSE resulted in lower values of \( \text{LHF}_{\text{SSE}} \). The differences between \( \text{LHF}_{\text{SMD}} \) and \( \text{LHF}_{\text{CSAA}} \) are much higher (20–40 W m\(^{-2}\)), particularly over the southern Indian Ocean, the South Pacific Ocean, the northwestern Atlantic Ocean, and the equatorial Pacific region. The geographical distribution of \( \text{LHF}_{\text{SMD}} - \text{LHF}_{\text{L86}} \) is almost similar to that of \( Q_a \) differences (Fig. 2d) maps. Large negative differences (more than \( -40 \) W m\(^{-2}\)) particularly to the east of South America, South Africa, and over the Arabian Sea are observed. Over these regions compared to SMD the L86 values are lower (Fig. 2d) by 3–4 g kg\(^{-1}\). Over the equatorial region and the northwestern Atlantic and Pacific Oceans the differences of more than \( 40 \) W m\(^{-2}\) are observed. Compared to SMD the L86 values are higher by 2–3 g kg\(^{-1}\) (Fig. 2d) over the northwestern Atlantic and Pacific Oceans and over the equatorial region. Figure 10 shows the rmsses between SMD and satellite-derived LHF for four methods: (Fig. 10a) \( \text{LHF}_{\text{GA}} \), (Fig. 10b) \( \text{LHF}_{\text{SSE}} \), (Fig. 10c) \( \text{LHF}_{\text{CSAA}} \), and (Fig. 10d) \( \text{LHF}_{\text{L86}} \), during the Sample II data period. Over a large part of the global ocean the rmse between \( \text{LHF}_{\text{SMD}} \) and \( \text{LHF}_{\text{GA}} \) is less than \( 30 \) W m\(^{-2}\). Slightly higher (30–45 W m\(^{-2}\)) rmse is found over the South Indian Ocean, the central Pacific Ocean, and over the northwestern Pacific and Atlantic Oceans. The global mean rmse for satellite-derived \( \text{LHF}_{\text{GA}} \) is \( 22 \pm 8 \) W m\(^{-2}\). The geographical distribution of the rmse of \( \text{LHF}_{\text{SSE}} \) is similar to that of \( \text{LHF}_{\text{GA}} \); however, in the case of \( \text{LHF}_{\text{SSE}} \) the areas covered by values in the range of 30–45 W m\(^{-2}\) is slightly larger compared to \( \text{LHF}_{\text{GA}} \). The global mean rmse for \( \text{LHF}_{\text{SMD}} \) is \( 29 \pm 10 \) W m\(^{-2}\). Over most of the oceanic regions the rmse are quite higher (30–45 W m\(^{-2}\)) for \( \text{LHF}_{\text{CSAA}} \) and \( \text{LHF}_{\text{L86}} \). In case of \( \text{LHF}_{\text{CSAA}} \) the rmsses are more than \( 45 \) W m\(^{-2}\) over the southern Indian Ocean, the South Pacific Ocean, the western Pacific Ocean, the northwestern Atlantic Ocean, and the equatorial Pacific regions. In the case of \( \text{LHF}_{\text{L86}} \) the rmsses are very large (more than \( 60 \) W m\(^{-2}\)) particularly over the southern Indian Ocean, the eastern equatorial Pacific Ocean, the eastern Atlantic Ocean, and over the Arabian Sea. The global mean rmsses for \( \text{LHF}_{\text{CSAA}} \) and \( \text{LHF}_{\text{L86}} \) are \( 37 \pm 13 \) and \( 41 \pm 13 \) W m\(^{-2}\), respectively.

b. Spatial and temporal variability in the satellite-derived LHF

1) Spatial variability

The monthly mean LHF estimates for January and July over the period 1988–93 are shown in Figs. 11 and 12. Figures 11a and 12a are obtained from satellite data (\( \text{LHF}_{\text{GA}} \)). Figs. 11b and 12b are obtained with SMD data (\( \text{LHF}_{\text{SMD}} \)). Figures 11c and 12c are obtained by \( \text{LHF}_{\text{SMD}} - \text{LHF}_{\text{GA}} \). The qualitative agreement between the two estimates is good in the sense that similar structures are found at the same location and at the same time during the year. The maximum LHFs (\( \sim 180 \) W m\(^{-2}\)) are generally found in the trade zones of both hemispheres (with larger fluxes in the winter), due to a larger \( Q_s - Q_a \) coupled with stronger wind. The higher...
Fluxes (~140 W m⁻²) are also found over the northwestern Pacific and Atlantic Oceans during January, where strong offshore winds carry cold, dry continental air over the warm Kuroshio Current and Gulf Stream. The higher fluxes (about 140–160 W m⁻²) are found over the Arabian Sea and the Bay of Bengal during July, due to higher wind speed. The equatorial region is associated with a small LHF (~50–75 W m⁻²), which may be attributed to small $Q_s - Q_a$ and low wind speed.

The flux decreases poleward because of the decrease in the air–sea humidity difference. We notice from Fig. 11c that the systematic positive LHF differences are mainly concentrated in high evaporation regions during January. In January (Fig. 11c), we observe two regions of the Northern Hemisphere where large differences are found. These regions are the northwestern Pacific and northwestern Atlantic Oceans. These large differences between LHF_{SMD} and LHF_{GA} may be due to the...
differences (Fig. 5c) between the SMD and GA retrieval. Over these regions GA-retrieved \( Q_a \) shows large (1–2 g kg\(^{-1}\)) negative bias in the retrieved \( Q_a \) resulting in the underestimation of LHF\(_{GA}\).

2) TEMPORAL VARIABILITY

A good way to check the performance of the single parameter retrieval and the combined method is validation with in situ time series in different geographical regions. In the following the results are analyzed for three regions (defined earlier) during 1988–93. A summary of these comparisons is given in Table 3.

(i) Arabian Sea

Figures 13 shows the annual variations of wind speed (\( U \)), saturation specific humidity (\( Q_s \)), and LHF. The
time series of near-surface specific humidity ($Q_a$) has already been discussed. The satellite-estimated wind is lower than SMD-observed wind speed throughout the year with pronounced differences ($\sim 2$ m s$^{-1}$) during the summer season. Similar results were obtained by JG95 for wind comparisons over the Arabian Sea. Satellite-derived $Q_a$ compares well with in situ estimates. The overall agreement between LHF$_{SMD}$ and LHF$_{GA}$ is good except during the winter and summer seasons when satellite-derived LHF$_{GA}$ is lower ($\sim 15$–$20$ W m$^{-2}$) than LHF$_{SMD}$. This underestimation of LHF$_{GA}$ is due to the underestimation of $U$ by SSM/I (Fig. 13a). This underestimation of the wind speed by SSM/I is partially offset by the underestimation (Fig. 7a) of $Q_a$ by GA; otherwise, an error of $2$ m s$^{-1}$ in $U$ would have produced large differences in LHF$_{GA}$. The differences between LHF$_{SMD}$ and other satellite-derived LHFs (LHF$_{CSAA}$, LHF$_{SSE}$, and LHF$_{L86}$) are large ($\sim 30$–$40$ W m$^{-2}$) in the winter and summer seasons. These large negative biases in the LHF$_{CSAA}$, LHF$_{SSE}$, and LHF$_{L86}$
are due to large positive bias in the satellite-derived (SSE, CSAA, and L86) $Q_a$ values (Fig. 7a).

(ii) Northwestern Pacific Ocean

Figure 14 is the same as Fig. 13, but over the northwestern Pacific Ocean. The satellite-derived winds are systematically lower than SMD. The differences are more pronounced during the summer season (also seen by JG95). Satellite derived $Q_a$ is always higher (~0.5 g kg$^{-1}$) compared to SMD-observed $Q_a$. During summer the LHF$_{GA}$ is close to LHF$_{SMD}$ even though the satellite-derived wind speed is less than the SMD-observed wind speed. The impact of the positive biases in the satellite-derived wind speed is not seen in LHF$_{GA}$ because of the negative bias in the satellite-derived $Q_a$ and a negligible bias in GA-retrieved $Q_a$. During the winter season the differences between LHF$_{GA}$ and LHF$_{SMD}$ are found on the order of 15–20 W m$^{-2}$. This is due to the negative bias on the order of 1 g kg$^{-1}$ in the GA-retrieved $Q_a$ and a negative bias in the satellite-derived $Q_a$. The satellite-derived LHF (LHF$_{SSE}$, LHF$_{CSAA}$, and LHF$_{L86}$) shows higher (30–40 W m$^{-2}$) fluxes compared
to SMD during the winter months. This is due to the $Q_a$ differences between SMD and these three satellite-based methods. The negative bias in satellite derived $Q_s$ is also contributing to these differences. Large differences between $LHF_{LS86}$ and $LHF_{SMD}$ have been observed, particularly in the summer season. During the summer, $LHF_{LS86}$ shows strong negative values. These negative values of $LHF_{LS86}$ are due to an overestimation of $Q_a$ by the LS86 method.

(iii) Northwestern Atlantic Ocean

Figure 15 is the same as Fig. 13, but over the northwestern Atlantic Ocean. The satellite-derived wind speed is always lower than the SMD observation, particularly during the summer season. The satellite-derived $Q_s$ is always higher than the SMD observations. The satellite-derived LHF$s$, except $LHF_{CSAA}$ and $LHF_{LS86}$ also compared well with $LHF_{SMD}$. Compared to $LHF_{SMD}$ during the winter season $LHF_{CSAA}$ shows

![Figure 13](http://journals.ametsoc.org/mwr/article-pdf/133/9/2692/4218633/mwr2993_1.pdf)

**Figure 13.** Time evolution of monthly mean (a) $U$ (m s$^{-1}$), (b) $Q_s$ (g kg$^{-1}$), and (c) LHF (W m$^{-2}$), over the Arabian Sea average during 1988–93.

![Figure 15](http://journals.ametsoc.org/mwr/article-pdf/133/9/2692/4218633/mwr2993_1.pdf)

**Figure 15.** Same as Fig. 13, but over the northwestern Atlantic Ocean average during 1988–93.

<table>
<thead>
<tr>
<th>Regions</th>
<th>Source</th>
<th>Correlation</th>
<th>Rmse</th>
<th>Bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabian Sea</td>
<td>LHF$_{GA}$</td>
<td>0.95</td>
<td>10</td>
<td>6.0</td>
</tr>
<tr>
<td>Pacific Ocean</td>
<td>LHF$_{SSE}$</td>
<td>0.81</td>
<td>20</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>LHF$_{CSAA}$</td>
<td>0.95</td>
<td>20</td>
<td>−12</td>
</tr>
<tr>
<td></td>
<td>LHF$_{LS86}$</td>
<td>0.70</td>
<td>26</td>
<td>−10</td>
</tr>
<tr>
<td></td>
<td>LHF$_{GA}$</td>
<td>0.99</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>Northwestern</td>
<td>LHF$_{SSE}$</td>
<td>0.97</td>
<td>20</td>
<td>−14</td>
</tr>
<tr>
<td></td>
<td>LHF$_{CSAA}$</td>
<td>0.97</td>
<td>43</td>
<td>−34</td>
</tr>
<tr>
<td></td>
<td>LHF$_{LS86}$</td>
<td>0.96</td>
<td>47</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>LHF$_{GA}$</td>
<td>0.97</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>Atlantic Ocean</td>
<td>LHF$_{CSAA}$</td>
<td>0.92</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>LHF$_{LS86}$</td>
<td>0.88</td>
<td>22</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>LHF$_{GA}$</td>
<td>0.89</td>
<td>18</td>
<td>10</td>
</tr>
</tbody>
</table>

**TABLE 3. Statistics of satellite-derived LHF.**

![Figure 14](http://journals.ametsoc.org/mwr/article-pdf/133/9/2692/4218633/mwr2993_1.pdf)

**Figure 14.** Same as Fig. 13, but over the northwestern Pacific Ocean average during 1988–93.
higher values while it shows lower values during the summer season. During the winter, differences as large as 40 W m\(^{-2}\) are observed between LHF\(_{SMD}\) and LHF\(_{CSAA}\). This is due to the combined effects of the negative bias in the satellite-derived \(Q_a\) and the positive bias in CSAA-derived \(Q_a\). Here LHF\(_{L86}\) shows higher positive bias during the summer season, which is due to the underestimation of the wind by SSM/I.

7. Conclusions

A new technique has been developed to estimate the global monthly mean LHF exclusively from satellite observations. LHF computation from satellite observations requires combined information from different spectral channels. The LHFs are derived from the surface winds from SSM/I (Wentz 1997), SST from AVHRR, and near-surface specific humidity retrieved using a GA that takes SSM/I-derived precipitable water vapor content and SST from AVHRR as input.

It is clearly shown that the use of GA leads to considerable improvement in the retrieval of near-surface specific humidity compared to L86, SSE, and CSAA methods. The global rmse in monthly mean \(Q_a\) is 0.80 ± 0.32 g kg\(^{-1}\). The differences between GA-retrieved \(Q_a\) and SMD-observed \(Q_a\) are slightly higher over the northwestern Pacific and Atlantic Oceans, particularly during wintertime. These regions are characterized by intense synoptic variability when the cold and dry continental air mass flows over warm ocean currents. The accuracy of the present method is comparable to best current method (the ANN approach) of \(Q_a\) retrieval, and this method has the advantage that it can be used exclusively with satellite observations using a simple empirical function.

Comparing the retrieved \(Q_a\) from other methods (SSE, CSAA, and L86) with SMD observations, it is found that the positive and negative biases of these methods are larger than that of GA. The SSE method shows strong (2–3 g kg\(^{-1}\)) negative and positive biases over the equatorial region and Arabian Sea, respectively. Over a large part of the global oceans the CSAA-derived \(Q_a\) differs by more than ±1 g kg\(^{-1}\), with higher differences over the Arabian Sea, the Bay of Bengal, and over the eastern equatorial Pacific Ocean. The L86 method shows strong positive biases (3–4 g kg\(^{-1}\)), particularly to the east of South America, South Africa, the Arabian Sea, the Bay of Bengal, and the South Indian Ocean. It shows negative biases (2–3 g kg\(^{-1}\)) over the northwestern Atlantic and Pacific Oceans and over the equatorial region.

The retrieved \(Q_a\) fields were used for the estimates of LHF using the bulk formula. The monthly averages of LHF's are compared with the SMD observation. The satellite-derived wind speed has a positive bias over all the oceans. The biases are more during summer season. The LHF determined from satellite data is temporally and spatially coherent. The rms difference between satellite- and SMD-derived LHFs shows that the two computations differ by about 22 W m\(^{-2}\) when GA-retrieved \(Q_a\) is used in LHF computations. Compared to LHF\(_{SMD}\), the LHF\(_{GA}\) has a positive bias of about 40–60 W m\(^{-2}\) over the western boundary currents like the Gulf Stream and the Kuroshio Current. This large positive bias in the LHF\(_{GA}\) is due to the negative bias in the GA-retrieved \(Q_a\) over these regions. Except for these two regions, the comparison between LHF\(_{GA}\) and LHF\(_{SMD}\) is excellent. Estimated LHF using \(Q_a\) derived from other methods (LHF\(_{SSE}\), LHF\(_{CSAA}\), and LHF\(_{L86}\)) shows larger biases and rmses than LHF\(_{GA}\). The comparisons of time series in the Arabian Sea and the northwestern Pacific and Atlantic Oceans have shown that temporal variability is very well represented by GA-based LHF. Overall the results show that the GA method is a promising technique for retrieving monthly mean \(Q_a\) and LHF.

In the future we plan to develop a modified algorithm that takes care of systematic regional biases present along the western continental boundaries.

Acknowledgments. The ISRO-Geosphere Biosphere Programme supported this study. The authors are especially grateful to Dr. M. S. Narayanan, the Group Director of the Meteorology and Oceanography Group and Dr. B. Simon for their support and encouragement. The authors would like to acknowledge the Physical Oceanography Data Archival Centre (PO.DAAC) at the Jet Propulsion Laboratory (JPL), the SMD active archive online at http://ingrid.ideo.columbia.edu, and the SSM/I ocean products available online at ftp.ssmi .com. The HOAPS data was obtained online at http:// www.mpimet.mpg.de/Depts/Physik/HOAPS/ and the GSSTF2 data was obtained online at http://daac.gsfc .nasa.gov/CAMPAIGN-DOCS/hydrology.

REFERENCES


