Assimilation of SSM/I and GOES Humidity Retrievals with a One-Dimensional Variational Analysis Scheme

GODELIEVE DEBLONDE, LOUIS GARAND, PIERRE GAUTHIER, AND CHRISTOPHER GRASSOTTI*

Aerospace Meteorology Division, Atmospheric Environment Service, Dorval, Quebec, Canada

(Manuscript received 14 March 1994, in final form 14 December 1994)

ABSTRACT

Total precipitable water (TPW) retrieved from Special Sensor Microwave/Imager (SSM/I) brightness temperatures and specific humidity retrieved from Geostationary Operational Environmental Satellite (GOES) radiances are assimilated using a one-dimensional (1D) variational analysis technique. The study is divided into two parts. First, collocations with radiosondes are performed to assess the quality of the satellite water vapor retrievals. Collocations are also performed with 6-h forecast fields. Second, SSM/I TPW and GOES specific humidity are assimilated using a 1D variational analysis technique that minimizes the error variance of the analyzed field.

A global collocation study over the oceans for SSM/I TPW retrievals and 6-h forecasts of TPW shows that the rmse (with respect to radiosondes) are, respectively, 4.7 and 5.0 kg m^{-2}. A separate collocation study over both the oceans and land for GOES retrieved TPW and 6-h forecasts of TPW yields rmse of 4.6 and 4.4 kg m^{-2}, respectively, in the midlatitudes and 6.8 and 5.9 kg m^{-2} in the Tropics.

The reduction of the 6-h forecast rmse when assimilating SSM/I TPW is 1 kg m^{-2}, which is a reduction of 20% in the rmse. When GOES retrievals of specific humidity are assimilated, the effective reduction is 0.6 kg m^{-2}. It is shown that in the upper levels of the troposphere (above 600 mb), the error reduction of specific humidity is largely due to the GOES retrievals, whereas in the lower troposphere (850 and 700 mb), the reduction is mostly due to the SSM/I TPW. This emphasizes the complementarity of the information contained at different wavelengths and the advantage of using multisensor retrievals in data analysis.

1. Introduction

Earlier studies that considered the importance of specifying the initial moisture field (Mintz 1964; Smagorinsky et al. 1970) with a large-scale GCM (acronyms listed in appendix A) showed that water vapor behaved like a passive tracer. However, recent studies have since shown that an accurate measure of the humidity field is necessary for proper initialization of high-resolution short- to medium-range NWP models, especially in the Tropics where the "spinup" time of the moisture field is of the order of 3 days rather than 12–15 h as in the midlatitudes. Medium-range forecasting in the midlatitudes is influenced by humidity generated in the Tropics. Also, precipitation forecasts are sensitive to initial moisture conditions (e.g., Illari 1989; Kuo et al. 1993).

Humidity observations are available through conventional observations (e.g., surface and upper-air observations or radiosondes, ship measurements, and aircraft reports). Because of the high spatial and temporal variability of the humidity field, point observations, if too sparse, will have a high error of representativeness. This is certainly the case in the Southern Hemisphere where the number of radiosonde launch sites is very small and its distribution uneven. Satellite retrievals of humidity are potentially advantageous when compared to radiosonde data because of their high temporal and spatial resolution (sometimes limited to oceanic areas only). The verification of satellite data retrieval algorithms is difficult to accomplish because they are generally compared with conventional data that are very sparse over oceans. Often, only the Northern Hemisphere radiosonde network over continents is adequate.

Satellite retrievals of moisture fields have been obtained from both microwave and infrared observations. Microwave observations at longer wavelengths are less affected by clouds and thus are a better tool to retrieve water vapor in cloudy atmospheres. The SSM/I, which is onboard of the polar-orbiting DMSP satellite, sees the same location on the globe about twice a day. About 60% of the globe is covered by one satellite in one day. There are gaps in the equatorial regions and the number of overpasses at the same location increases as one moves toward northern latitudes. Over the ocean sur-

* Current affiliation: Atmospheric and Environmental Research Inc., Cambridge, Massachusetts.

Corresponding author address: Dr. Godelieve Deblonde, Aerospace Meteorology Division, Atmospheric Environment Service, 2121 Trans-Canada Highway, 5th Floor, Dorval, PQ, H9P 1J3 Canada.

face, due to the low surface emissivity values, microwave brightness temperatures at the top of the atmosphere appear warm with respect to the ocean surface. Over land, the surface emissivity is higher and more variable and thus the retrieval of atmospheric water vapor is much more difficult. Retrievals of water vapor obtained from infrared sensors on the geostationary satellites are usually limited to cloud-free areas. Techniques that obtain humidity profiles based on cloud classification schemes are an exception to this. Infrared water vapor retrievals are also available over land and have a higher sampling frequency (30 min). Thus, a combination of both retrieval techniques (IR and microwave) would give access to the largest observation dataset. Both sensor types are used in this study to retrieve the moisture field.

Retrievals of atmospheric hydrological variables over the oceans (such as total precipitable water, which is equal to the column-integrated atmospheric water vapor, cloud liquid water, and precipitation rate) from DMSP SSM/I brightness temperatures have been used in the context of mesoscale analysis (e.g., Chang et al. 1993; Alliss et al. 1993), NWP [e.g., Isaacs et al. 1986 (review paper); Illari 1989; Nehrkorn et al. 1993], and in the production of global climatologies (e.g., Tjemkes and Stephens 1990; Liu et al. 1992; Bauer and Schluessel 1993). In this paper, the focus is on the retrievals of TPW and specific humidity q from DMSP and GOES satellite data.

TPW retrievals are obtained from SSM/I brightness temperatures using a statistical regression approach. The SSM/I has four channels (19.3, 22.2, 37.0, and 85.5 GHz), three of which have dual polarization. The 22.2-GHz channel has only a vertical polarization and samples the water vapor absorption line. The retrieval of TPW has the highest correlation with that channel (Schluessel and Emery 1990). The other channels usually reduce the rmse by compensating for varying surface wind speeds (which modify the surface emission) and/or liquid water content. A technical description and evaluation of the SSM/I for the F08 satellite may be found in Hollinger et al. (1990). Retrieval accuracies obtained from radiosonde collocations for TPW are usually less than 5.0 kg m^{-2} (Alishouse et al. 1990; Schmetz and van de Berg 1991, their Table 1; Colton 1993).

Retrieval techniques based on IR usually have similar or higher rmse for TPW than the microwave techniques (Schmetz and van de Berg 1991). Comparisons between IR and microwave retrievals have been performed previously by Tjemkes and Stephens (1990) and Liu et al. (1992). Both studies compared water vapor retrievals obtained from TOVS HIRS-2 and the SSM/I over oceans and showed the superiority of SSM/I retrievals. However, the technique to retrieve precipitable water from HIRS-2 used as a first guess a combination of eigenvectors obtained using radiosonde data (Smith and Woolf 1976). This may lead to fields that are too smooth especially in data sparse areas. The problem with cloud contamination also remains a difficult problem for most IR-based methods.

In this paper, retrievals of q are obtained using the so-called "Humsat" technique (humidity from satellites), which is based on a cloud-classification scheme developed by Garand (1993). Three channels of full disk *GOES-7* images are employed—the infrared channel (10.5–12 μm), the visible channel (0.55–0.75 μm) when available, and the water vapor channel (6.7 μm) also when available. The weighting function of the water vapor channel peaks near 400 mb. An advantage of such a method is that retrievals are available in cloudy atmospheres.

Assimilation of humidity fields to provide good analyses for NWP models has been explored before. Haydu and Krishnamurti (1981) developed an analysis scheme to obtain 3D fields of q (from 1000 to 500 mb). The scheme used microwave retrievals of TPW, radiosonde, and conventional surface observations of specific humidity. First, they calibrated the microwave TPW to that derived from the radiosondes and at the same time, using a relaxation technique, preserved the satellite-derived Laplacian. Thus, satellite moisture gradients were preserved. Second, assuming an exponential decay of q with height, the surface specific humidity was extrapolated upward under the constraint that its vertical integral be equal to the calibrated TPW. The major problem with this technique is the sparsity of radiosondes and conventional surface measurements of humidity, particularly in the Tropics.

Kuo et al. (1993) described a method where a linear correction is performed on the vertical profile of the trial field precipitable water to constrain TPW of the trial fields to equal that obtained from analyzed fields. Iterations were performed to ensure that supersaturation does not occur. They hoped that eventually the analyzed fields could be replaced by SSM/I retrievals of TPW to correct the vertical profile of specific humidity of the trial field. However, this approach does not weigh in an optimal manner the information contained in the retrieval and the trial field.

Illari (1989) compared the quality of precipitable water obtained from 1) 6-h forecasts with the ECMWF NWP model (T63 spectral model); 2) radiosondes; and 3) satellite retrievals. The satellite retrievals of precipitable water were obtained with an algorithm that uses TOVS channels. She found that all three fields have comparable accuracy and from a data assimilation experiment concluded that the inclusion of satellite data in the humidity analysis had a positive impact on the forecast and that the largest impact was for the tropical precipitation forecast.

In this study, satellite retrieved SSM/I TPW and specific humidity retrieved from GOES imagery are assimilated using a 1D variational analysis technique (Lorenc 1986, 1988; Daley 1991). This is a preliminary study for the 3D variational analysis method currently

being developed at CMC. In the variational approach, the analyzed field is obtained by combining a 6-h forecast with the retrieved satellite data (both with known error covariances) by minimizing the analysis error variance.

The study is divided into two parts: first, in section 2, an evaluation of the accuracy of the retrievals of total precipitable water and specific humidity profiles is done by comparing them with collocated upper-air radiosondes. The retrievals of TPW and q are also compared with 6-h forecasts or trial fields obtained with the CMC operational NWP model [global spectral model, T79 (see Ritchie 1991), which corresponds to a horizontal resolution of 250 km at the equator]. Second, in section 3, a 1D variational analysis system is used to obtain the best estimate for specific humidity profiles using all three types of data. Results of the collocation study and assimilation are presented. An example of joint assimilation of SSM/I and GOES humidity retrievals is also given. Finally, in section 4, conclusions are presented.

2. Methodology

a. *Retrieval of total precipitable water from DMSP SSM/I brightness temperatures*

Total precipitable water was computed from DMSP SSM/I brightness temperatures for two periods of 12 days each. The first dataset extended from 1 June to 12 June 1991 and was available for the two satellites F08 and F10. The second dataset extended from 1 March to 12 March 1992 and was available for the F10 and F11 satellites. The procedure described in Wentz (1991) was used to transform the antenna temperatures to brightness temperatures. Remapping was performed on a $0.3° \times 0.3°$ latitude–longitude grid by nearest neighbor resampling. Numerous statistical and physical-statistical algorithms are available to retrieve TPW. For simplicity, an empirically based algorithm was chosen. The most recent intercomparison between such algorithms may be found in Colton (1993). The rmse (with respect to radiosondes) varies between 2.6 and 6.0 kg m^{-2} according to the intercomparison performed by Petty and listed in Colton (1993). In our study, a nonlinear regression equation (Alishouse et al. 1990) was used to retrieve TPW from SSM/I brightness temperatures over the ocean. This algorithm was used because it has low rmse values and it is widely used in the literature (the algorithm was chosen as the calibration–validation algorithm to retrieve TPW, see Hollinger 1991). The regression equation was developed for a global dataset obtained from the F08 satellite and for a dataset that spanned almost a year (June 1987 to April 1988). Alishouse et al. (1990) obtained a rmse of 3.0 kg m^{-2} with 0.0 bias for the dependent dataset (dataset with which the empirical relationship was developed) and a rmse of 4.2 kg m^{-2} with a bias of -0.7 kg m^{-2} for an independent dataset (Hollinger 1991). The nonlinearity in the regression arises from a quadratic term for the 22.2-GHz channel. The equation is as follows:

$$\text{TPW (kg m}^{-2}) = c_0 + c_1 T_b^{19v} + c_2 T_b^{22v} + c_3 T_b^{37v} + c_4 (T_b^{22v})^2, \quad (1)$$

where T_b^f are brightness temperatures (K) for the frequencies f(GHz) with vertical polarization. The coefficients c_0–c_4 are, respectively, 232.89, -0.148596, -1.829125, -0.36954, and 0.006193. Regression equation (1) was developed in nonprecipitating areas. Therefore, a precipitation screen (Hollinger 1991) has to be applied prior to computing TPW. The precipitation screen (based on the 37.0-GHz vertical and horizontal polarization channels) is defined as follows. If

$$-11.7939 - 0.02727 T_b^{37v} + 0.09929 T_b^{37h} \leq 0 \text{ K}, \quad (2)$$

then TPW is computed, otherwise, it is not.

Collocations between SSM/I TPW pixels and radiosondes were performed to estimate the accuracy of the retrievals from (1). Collocations were global but restricted to be located between 60°N and 60°S. An SSM/I TPW pixel was chosen as a collocation pixel provided that both the distance to the radiosonde (<300 km) and the difference in observation time (<3 h) were minimized. Collocations were performed only for radiosondes that were launched from small islands to minimize land contamination. As in Alishouse et al. (1990), for an island to qualify as small, its area has to be less than 20% of the incident field of view of the 19-GHz channel (which has the lowest resolution).

Vertical profiles of q for 6-h trial fields obtained with the spectral global NWP model (Ritchie 1991) were also available at the location of the radiosondes. The data that is assimilated for initializing the forecast falls into three categories: upper-air data, satellite data, and surface observations. The upper-air data consists of radiosonde observations (temperature, dewpoint temperature, pressure, and geopotential height) and aircraft data (temperature, wind speed, and direction). The satellite data consists of satob observations (geostationary satellite observations of wind speed and direction at various levels) and satem observations (NOAA series satellite observations of temperature at various levels). Synoptic surface meteorological observations and fixed and drifting buoy observations form the surface observation dataset. Thus, over the oceans the only moisture data that are assimilated come from radiosonde observations and surface observations. For the trial fields and the radiosondes, TPW was computed from vertical profiles of q as follows:

$$\text{TPW (kg m}^{-2}) = \frac{1}{g} \int_{1000\text{ mb}}^{300\text{ mb}} q\,dp, \quad (3a)$$

where g is the gravitational constant and p is pressure. The integral was discretized using the trapezoidal rule and leads to the following equation:

$$\mathrm{TPW} = \frac{1}{2g} \sum_{j=1}^{6} (q_j + q_{j+1})(p_j - p_{j+1}), \quad (3b)$$

where j is the level index. The levels used in this study are 1000, 850, 700, 500, 400, and 300 mb. These levels coincide with the mandatory levels for radiosonde observations. There are two reasons why the upper limit of the integral is chosen to be 300 mb. First, the radiosonde reports rarely include humidity observations above that level, and second, the largest integrated water vapor amounts above 300 mb are less than 1% of the value measured below 300 mb.

b. *Retrieval of specific humidity from GOES imagery*

Humsat retrievals (Garand 1993) of DPD profiles from *GOES-7* were performed in the −60° and 60° latitude and 190° to 320° longitude window and on a 1° × 1° grid. Retrievals were obtained for 29 May 1991 to 12 June 1991 and 26 February 1992 to 12 March 1992 and overlap with the SSM/I datasets. DPD was converted to specific humidity q by using temperature profiles provided by the 6-h forecast. A brief description of the Humsat technique follows.

Three features are extracted when GOES radiances are available for both the VIS and IR channels. These are cloud-top pressure (which is determined by using the 6-h temperature trial field and the IR channel), cloud fraction (based on the albedo and temperature thresholds), and mean cloud albedo. In the case where only the IR channel is available, the mean cloud albedo is not extracted and cloud fraction is obtained by using an IR threshold based on the 6-h temperature trial fields at 850 and 1000 mb and also SSTs obtained from an operational analysis. The features are then identified with a cloud class. Each cloud class has an assigned profile of DPD (obtained from radiosonde observations), which is used as a first guess for the retrieval profile. TBWV and 6-h trial field surface temperatures are used to improve on the first guess. A description of how the cloud classes were obtained and how the first-guess profile was improved on follows.

Cloud classes were determined with a clustering algorithm based on the analysis of the features (Duda and Hart 1973, section 6.9). Nine classes were obtained for the VIS–IR scheme and seven classes for the IR-only scheme. Medians of DPD obtained from collocated radiosondes were assigned to each class. For the midlatitudes (35°–60°N), Meteosat and GOES DPD medians were averaged. For the Tropics, medians were obtained from the GOES retrievals only.

The first guess (based on class medians) is improved on by applying empirical, class-dependent, relationships that significantly reduce the rmse. First, the TBWV when available is used to estimate the humidity above 600 mb. Linear relationships between DPD and TBWV were determined at 500, 400, and 300 mb. These relationships are used at all latitudes. Second, an empirical relation between DPD and the 6-h temperature trial field at 1000 mb was developed and used to replace retrieved DPD at 1000 mb. The dynamic range of such a relation is too low in the Tropics; hence, the relation is only used in the midlatitudes or over desert areas. A "supermoist" class (defined by DPD values in the range of 1–3 K depending on the level) was added in the case where cloud-top pressure exceeds 150 mb. This class occurs mostly in the Tropics and is associated with convective precipitation.

A collocation study with radiosondes over Europe (for an independent dataset) showed that a bias corrected rmse of 4.5 kg m^{-2} for TPW was obtained. The dataset had a mean TPW of 22.9 kg m^{-2} and a standard deviation of 11.4 kg m^{-2}. A few improvements mainly concerned with data quality checking were recently added to the algorithm and are summarized in appendix B. At the CMC, vertical profiles of DPD are retrieved from GOES radiances operationally four times daily and were introduced into the optimal interpolation cycle in July 1993.

c. *Data assimilation formalism*

Specific humidity q was chosen to be the working variable of the 1D variational analysis of humidity and is linearly related to TPW via (3). We evaluate q at the following six pressure levels: p_k (k = 1, . . . , 6) = 1000, 850, 700, 500, 400, and 300 mb. The model state was restricted to where Humsat data were available. The aim of the 1D variational analysis is to provide the best estimate for the vertical profile of specific humidity by combining different sources of observations and a background field chosen to be the 6-h trial field obtained from a forecast model.

If one assumes that observations and background error have Gaussian error distributions, then minimizing the cost function $J(q)$ gives an estimate that minimizes the error variance of the analyzed field (Lorenc 1986, 1988). The cost function $J(q)$ for the 1D univariate variational analysis (q is the only variable that is assimilated) is given by

$$J(q) = \tfrac{1}{2}(\mathbf{q} - \mathbf{q}_b)^{\mathrm{T}}\mathbf{P}^{-1}(\mathbf{q} - \mathbf{q}_b) + \tfrac{1}{2}[\mathbf{H}(q) - \mathbf{z}]^{\mathrm{T}}\mathbf{O}^{-1}[\mathbf{H}(q) - \mathbf{z}]. \quad (4)$$

The first term on the rhs of (4) is the contribution to the cost function from the background field and the second term comes from the observations; $\mathbf{z}$ are the observations to be assimilated; and $\mathbf{H}(q)$ is the so-called forward model for the variable $\mathbf{z}$. A forward model produces the model equivalent of observations. Here $\mathbf{q}_b$ is the background field of specific humidity, and $\mathbf{P}$ and $\mathbf{O}$ are the error autocovariance matrices of the background field and observations, respectively. The computations involved in obtaining these matrices are

described in section 2d. The minimum of $J(q)$ is found by solving for the zero of the first derivative of $J(q)$, which is given by

$$J'(q) = \mathbf{P}^{-1}(\mathbf{q} - \mathbf{q}_b) + \mathbf{H}'^{\mathrm{T}}\mathbf{O}^{-1}[\mathbf{H}(q) - \mathbf{z}], \quad (5)$$

where $\mathbf{H}'$ is the tangent linear operator to $\mathbf{H}$ in the vicinity of $\mathbf{q}$.

If $\mathbf{H}(\mathbf{q})$ is nonlinear, then an iterative procedure would probably be required to find the zero of (5) (Gilbert and LeMaréchal 1989). In the case where $\mathbf{z}$ is equal to the SSM/I TPW, the forward model simply consists of the integral operator (3a), and thus $\mathbf{H}(q)$ is linear. In the case where $\mathbf{z}$ represents the Humsat $\mathbf{q}$ observations, then the forward model is the identity matrix. When $\mathbf{H}(q)$ is linear as is the case in our study and we define $\mathbf{H}(q)$ such that $\mathbf{H}(q) = \mathbf{H}\mathbf{q}$ [hence $\mathbf{H}'(q) = \mathbf{H}$], then an exact solution for the zero of (5) may be found (Lorenc 1986) and is given by

$$\mathbf{P}^{-1}(\mathbf{q}_a - \mathbf{q}_b) + \mathbf{H}^{\mathrm{T}}\mathbf{O}^{-1}[\mathbf{H}(\mathbf{q}_a - \mathbf{q}_b) + \mathbf{H}(\mathbf{q}_b) - \mathbf{z}] = 0, \quad (6a)$$

where $\mathbf{q}_a$ represents the field that minimizes the error variance of the analysis. The equation may also be rewritten as

$$\mathbf{q}_a - \mathbf{q}_b = (\mathbf{P}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{O}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{O}^{-1}[\mathbf{z} - \mathbf{H}(\mathbf{q}_b)] \quad (6b)$$

or

$$\mathbf{q}_a = \mathbf{q}_b + \mathbf{K}[\mathbf{z} - \mathbf{H}(\mathbf{q}_b)], \quad (7)$$

where

$$\mathbf{K} = (\mathbf{P}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{O}^{-1}\mathbf{H})^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{O}^{-1}. \quad (8)$$

After some algebraic manipulations, $\mathbf{K}$ may also be shown to be given by

$$\mathbf{K} = \mathbf{P}\mathbf{H}^{\mathrm{T}}(\mathbf{O} + \mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}})^{-1}, \quad (9)$$

where $\mathbf{K}$ is referred to as the gain matrix and $\mathbf{z} - \mathbf{H}(q_b)$ the innovation vector. The error autocovariance of the analyzed field $\mathbf{q}_a$ may also be computed and is shown below to be given by

$$\mathbf{P}_a = [J''(\mathbf{q})]^{-1}. \quad (10)$$

From (4), one has that

$$J''(\mathbf{q}) = \mathbf{P}^{-1} + \mathbf{H}^{\mathrm{T}}\mathbf{O}^{-1}\mathbf{H} \quad (11)$$

and (8) may be rewritten as

$$\mathbf{K} = J''^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{O}^{-1}. \quad (12)$$

On the other hand, one has (Ghil 1989):

$$\mathbf{P}_a = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{P} = (\mathbf{I} - \mathbf{J}''^{-1}\mathbf{H}^{\mathrm{T}}\mathbf{O}^{-1}\mathbf{H})\mathbf{P}. \quad (13)$$

Finally, using (11) in the above equation completes the derivation of (10).

In the assimilation performed in this study, the two sources of observations (for which the errors are not correlated), thus are SSM/I TPW and Humsat q, and it follows that $\mathbf{z}$, $\mathbf{O}$, and $\mathbf{H}(\mathbf{q})$ may be written as

$$\mathbf{z} = \begin{bmatrix} z_m \\ q_h \end{bmatrix}; \quad \mathbf{O} = \begin{bmatrix} \mathbf{Q}_m & 0 \\ 0 & \mathbf{O}_h \end{bmatrix}; \quad \mathbf{H}(q) = \begin{bmatrix} \mathbf{L}_m q \\ q \end{bmatrix}, \quad (14)$$

where z_m = SSM/I TPW, q_h is the Humsat retrieved specific humidity, $\mathbf{O}_m$ is the error variance for the SSM/I TPW, and $\mathbf{O}_h$ is the matrix of the error autocovariance for Humsat retrievals of q. Here $\mathbf{L}_m q$ is the vertical integral of q, and $\mathbf{L}_m$ is given by [see (3b)]

$$\mathbf{L}_m = \frac{1}{2g}(p_1 - p_2, p_1 - p_3, p_2 - p_4, p_3 - p_5, p_4 - p_6, p_5 - p_6). \quad (15)$$

d. Autocovariance error matrices

Because a 1D univariate analysis is performed, in the case of q, only vertical correlations of errors need to be computed. Since the errors vary with season and latitude, they were computed by latitude band (midlatitudes: 30°–60° and Tropics: 0.0°–30°) and for each dataset separately. The computation of error autocovariances by latitude band and season is a standard procedure that is performed in the optimal interpolation analysis at CMC (Mitchell et al. 1990, 1993).

By definition, errors have to be computed with respect to so-called "true" values. Frequently, however, the errors reported are with respect to radiosonde values. An example as to how these errors are related for TPW follows. Let us define e^2 as the (rmse)^2 between SSM/I TPW (TPW_m) and TPW obtained from radiosondes (TPW_r). Then e^2 may be written as follows

$$e^2 = \langle(\mathrm{TPW}_m - \mathrm{TPW}_r)^2\rangle, \quad (16)$$

where the brackets denote ensemble averaging. Substituting TPW_m in (16) by ($\mathrm{TPW}_m - \mathrm{TPW}_t$) and TPW_r by ($\mathrm{TPW}_r - \mathrm{TPW}_t$) where TPW_t is the true value of TPW, one obtains

$$e^2 = \langle(\mathrm{TPW}_m - \mathrm{TPW}_t)^2\rangle + \langle(\mathrm{TPW}_r - \mathrm{TPW}_t)^2\rangle - 2\langle(\mathrm{TPW}_m - \mathrm{TPW}_t)(\mathrm{TPW}_r - \mathrm{TPW}_t)\rangle. \quad (17)$$

The last term in the above equation vanishes because radiosonde errors are uncorrelated with TPW_m. Both e^2 and the second term on the rhs of the above equation may be computed, and thus the error of the microwave retrievals with respect to the true values could be obtained as a residual of the above equation. The rmse error of radiosondes was estimated to be 1 kg m^{-2} for the United States network (Alishouse et al. 1990). Radiosonde errors have been estimated in units of relative humidity by Illari (1989). In her study, radiosonde errors were estimated to be similar to those of the IR satellite retrievals obtained with TOVS.

In our study, we compare errors between three datasets: SSM/I TPW, Humsat q, and 6-h trial fields. In

TABLE 1. SSM/I, trial field, and assimilated total precipitable water collocation statistics (kg m^{-2}).

Sample size/No. of stations	Radiosonde mean	Radiosonde SD	Technique	SD	rmse	Bias
		Dataset: 1–12 June 91—F08				
163/27	39.0	12.9	SSM/I	12.1	4.3	1.0
			Trial field	12.7	4.7	1.1
			Assimilate	12.2	3.7	1.1
		Dataset: 1–12 June 91—F10				
242/37	37.0	15.0	SSM/I	14.3	5.4	1.1
			Trial field	15.0	4.6	1.0
			Assimilate	14.4	4.4	1.1
		Dataset: 1–12 March 92—F10				
182/35	31.2	14.0	SSM/I	13.0	4.8	−0.63
			Trial field	13.8	5.7	1.7
			Assimilate	13.1	4.4	0.25
		Dataset: 1–12 March 92—F11				
95/27	36.9	12.3	SSM/I	12.7	4.4	−1.3
			Trial field	12.6	5.2	1.1
			Assimilate	12.3	3.7	−0.41

all cases, the errors are computed with respect to radiosondes and therefore are comparable. Thus, P, O_m, and O_h were determined by collocation with radiosonde observations without removing the radiosonde error.

For the SSM/I TPW, the error variance of TPW needs to be computed since there is only one value per horizontal coordinate. For Humsat q and the 6-h forecasts of q, error autocovariance matrices have to be computed. Error statistics were obtained by collocating model and satellite retrievals with radiosonde observations. The error autocovariance matrices are computed as follows. Let $q_{s,ik}$ and $q_{r,ik}$ be the satellite and radiosonde specific humidity at collocation point i and level k and also define the error to be given by E_{ik}

TABLE 2. Humsat, trial field, and assimilated total precipitable water collocation statistics (kg m^{-2}).

Sample size/No. of stations	Radiosonde mean	Radiosonde SD	Technique	SD	rmse	Bias
		Dataset: 29 May 1991 to 12 June 1991 Midlatitudes				
809/80	24.8	11.9	Humsat	10.6	5.1	−0.37
			Trial field	11.3	4.5	0.90
			Assimilate	10.8	4.1	0.33
		Tropics				
238/34	40.6	9.3	Humsat	7.6	6.7	0.49
			Trial field	8.1	6.4	−0.24
			Assimilate	7.3	5.5	−0.25
		Dataset: 26 February 1992 to 12 March 1992 Midlatitudes				
251/29	13.5	8.3	Humsat	9.2	4.0	0.56
			Trial field	9.1	4.2	2.3
			Assimilate	8.8	3.5	1.5
		Tropics				
231/38	39.8	11.3	Humsat	10.4	6.9	−0.89
			Trial field	11.3	5.4	−0.84
			Assimilate	10.5	5.1	−0.86

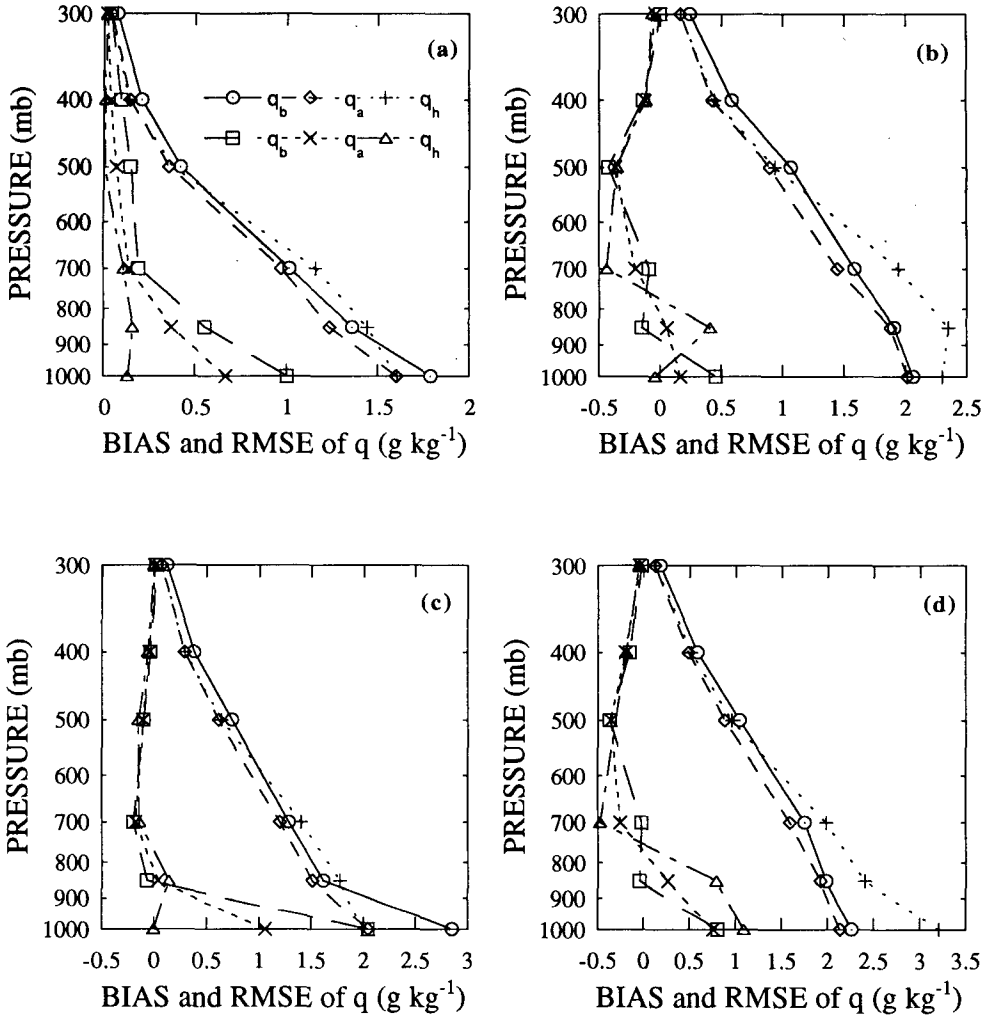


FIG. 1. Assimilation of Humsat specific humidity. Both the bias (the three left-most curves in each graph) and the rmse of the specific humidity are illustrated (the three right-most curves in each graph). Here q_b represents the background (6-h trial field) specific humidity, q_a represents the analyzed field, and q_h are the values obtained with the Humsat (or GOES) retrieval scheme: (a) midlatitudes, March 1992 dataset; (b) Tropics, March 1992 dataset; (c) midlatitudes, June 1991 dataset; and (d) Tropics, June 1991 dataset.

$= q_{s,ik} - q_{r,ik}$. Then the mean and autocovariance are given, respectively, by

$$\bar{E}_k = \frac{1}{N}\sum_{i=1}^{N} E_{ik} \tag{18a}$$

$$C_{km} = \frac{1}{N-1}\sum_{i=1}^{N} (E_{ik} - \overline{E_k})(E_{im} - \overline{E_m}). \tag{18b}$$

3. Results and discussion

a. *Retrieval of total precipitable water and specific humidity*

To evaluate the quality of moisture retrievals, three collocation studies with radiosondes were performed. The first study derived the rmse of SSM/I TPW values for collocated radiosondes, the second study derived the rmse for Humsat retrievals and radiosondes, and, finally, collocations were performed between SSM/I TPW, Humsat q, and radiosondes. In all three cases, statistics were also available for the 6-h trial fields interpolated from an NWP grid to the location of the radiosondes.

Table 1 summarizes the collocation analysis for the SSM/I TPW retrievals (over ocean only). In Table 1, the total number of collocations (sample size) and the number of different radiosonde launch sites are listed. The radiosonde mean and standard deviation are listed in the second and third column. The column entitled "technique" lists the statistics for the SSM/I TPW and the 6-h trial field. The mean of these fields may be obtained by adding the bias to the radiosonde mean.

TABLE 3a. Autocovariance error statistics for specific humidity ($g^2\ kg^{-2}$) (March 1992).

Pressure level	Mean error	Trial field mean error and error autocovariance matrix (**P**) Error autocovariance matrix 1000 mb	850 mb	700 mb	500 mb	400 mb	300 mb
1000 mb	0.64	0.85				(30°–60°N): N = 4246	
850 mb	0.13	0.23	0.53				
700 mb	0.036	0.049	0.098	0.27			
500 mb	0.047	0.019	0.023	0.029	0.058		
400 mb	0.038	0.0076	0.011	0.010	0.017	0.015	
300 mb	0.011	0.0033	0.0035	0.0018	0.0033	0.0034	0.0018
1000 mb	0.37	4.0				(0°–30°N): N = 1123	
850 mb	−0.27	0.88	3.2				
700 mb	−0.026	0.27	0.69	1.9			
500 mb	−0.10	0.089	0.24	0.31	0.62		
400 mb	0.0013	0.061	0.098	0.15	0.20	0.20	
300 mb	−0.0065	−0.0081	0.037	0.028	0.041	0.040	0.031
1000 mb	1.2	5.2				(0°–30°S): N = 375	
850 mb	0.91	1.2	4.3				
700 mb	0.41	0.12	1.2	3.4			
500 mb	−0.12	0.059	0.33	0.96	1.5		
400 mb	0.069	0.047	0.16	0.50	0.51	0.50	
300 mb	0.051	0.019	0.029	0.14	0.13	0.10	0.084
1000 mb	1.0	2.8				(30°–60°S): N = 290	
850 mb	0.60	0.84	3.1				
700 mb	0.45	0.071	0.55	2.0			
500 mb	0.27	0.12	0.29	0.18	0.55		
400 mb	0.16	0.044	0.040	0.048	0.15	0.13	
300 mb	0.079	0.0026	0.0070	0.012	0.024	0.023	0.015

The magnitude of the rmse (4.3–5.7 kg m^{-2}) with respect to radiosonde observations is similar to that obtained in other studies (see section 2a). The collocation time window and range chosen in our study was 3 h and 300 km, respectively. Decreasing the collocation time window and range to 2 h and 200 km (matches the one used in Alishouse et al. 1990), respectively, decreases the rmse by 0.5 kg m^{-2} for three of the datasets and 0.1 kg m^{-2} for the fourth (1–12 March 1992, F11 dataset). The sample sizes of the collocations are also reduced considerably (up to 60% depending on the dataset).

Data merging for different sensors were not performed because equations for the intercalibration between the different satellites were not available. However, differences between satellites were estimated by Hollinger in Isaacs et al. (1992) by comparing near coincident pixels in time for different satellites over selected regions. The differences in the mean that they obtained between F10 and F08 were less than 0.5 K for channels 19.3, 22.2, and 37 GHz. The differences between the F10 and F11 satellites were 1.6 K for the 22.2-GHz channel and less than 0.5 K for the 19.3- and 37.0-GHz channels.

The rmse for the SSM/I TPW are smaller than for the 6-h trial fields except for one of the datasets (Table 1). This shows a higher performance of the SSM/I. The trial field bias of TPW is positive for all datasets implying overestimated moisture. Actually, atmospheric water vapor (at least below 400 mb) is reasonably well predicted by numerical weather prediction models and in particular when the performance is compared with the forecast of cloud liquid water and surface rain rates.

The rmse of TPW values obtained with Humsat are listed in Table 2. Since more collocations were available for Humsat (both land and ocean cases were considered), the datasets were split by latitudinal bands. The rmse for Humsat TPW is lowest for the midlatitudes and is comparable to that of the trial fields (5.1 kg m^{-2} for the former compared with 4.5 kg m^{-2} for the latter and for the June dataset, these values are 4.0 kg m^{-2} compared with 4.2 kg m^{-2} for the March dataset). For the tropical datasets, the Humsat TPW rmse are 6.7 kg m^{-2} for the June dataset and 6.9 kg m^{-2} for the March dataset. The 6-h trial field performs slightly better than Humsat in the Tropics (Table 2). The bias of the trial field is positive in the midlatitudes and negative in the Tropics. The rmse values obtained for the midlatitudes are similar to the ones reported in Garand (1993) and summarized in section 2b. In the Tropics, the mean TPW is larger (Table 2) and thus a larger

TABLE 3b. Autocovariance error statistics for specific humidity ($g^2\ kg^{-2}$) (March 1992).

Pressure level	Mean error	Humsat mean error and error autocovariance matrix ($\mathbf{O}_h$) Error autocovariance matrix 1000 mb	850 mb	700 mb	500 mb	400 mb	300 mb
1000 mb	0.17	1.6				(30°–60°N): N = 223	
850 mb	0.065	0.26	1.56				
700 mb	0.012	0.079	0.41	1.13			
500 mb	−0.057	0.020	−0.015	0.070	0.089		
400 mb	−0.0053	−0.011	−0.015	0.0074	0.017	0.013	
300 mb	0.0081	0.0018	−0.0045	−0.0051	0.00005	0.00086	0.00054
1000 mb	0.22	5.3				(0°–30°N): N = 108	
850 mb	0.20	1.7	5.1				
700 mb	−0.011	0.58	0.79	2.5			
500 mb	−0.18	0.24	−0.21	−0.044	0.26		
400 mb	−0.23	−0.11	−0.15	0.040	0.052	0.076	
300 mb	−0.093	0.0038	−0.053	0.020	0.017	0.0089	0.021
1000 mb	−0.27	5.3				(0°–30°S): N = 123	
850 mb	0.59	1.7	5.6				
700 mb	−0.82	0.83	1.9	4.2			
500 mb	−0.51	0.099	0.24	1.0	1.1		
400 mb	−0.019	0.18	0.16	0.20	0.27	0.26	
300 mb	−0.038	0.0090	−0.00081	−0.0030	0.017	0.021	0.024
1000 mb	−0.21	10.4				(30°–60°S): N = 28	
850 mb	0.91	3.1	5.6				
700 mb	0.89	−0.23	0.17	2.37			
500 mb	0.48	−0.045	0.12	0.022	0.30		
400 mb	0.12	0.016	0.14	−0.16	0.041	0.058	
300 mb	0.042	0.0046	0.0045	0.028	−0.018	−0.0037	0.0064

rmse is to be expected. The explained variance (1 minus the ratio of the retrieval error variance to the observed variance) was 82% and 77% for the midlatitudes (June and March, respectively) and 47% and 63% for the Tropics (June and March, respectively). Thus, in the midlatitudes, the performance of Humsat is superior to that in the Tropics. This may be explained in part by the fact that the empirical equation that relates surface DPD to the 6-h temperature trial field is not used in the Tropics (section 2b).

Finally, collocation statistics of TPW for SSM/I, Humsat, and the 6-h trial field were also performed but are not shown because the number of samples is substantially reduced and thus quantitative conclusions cannot be made.

Even if Humsat TPW estimates are the least accurate, statistics of specific humidity by level (Fig. 1, variable q_h) show that Humsat retrievals of q have lower rmse than the 6-h trial fields (Fig. 1, variable q_b) for the 500-, 400-, and 300-mb levels. The reason that Humsat performs better there is mostly due to the WV channel whose weighting function peaks above 600 mb. For the midlatitude datasets, the retrieval of q at 1000 mb performs better than the 6-h trial field. This is not the case for the tropical datasets (Fig. 1b and 1d). As noted earlier, this is probably due to the fact that the empirical equation developed at the surface that relates DPD and the 6-h temperature trial field is not used in the Tropics (section 2b). Thus, the levels where Humsat performs worst are 700 and 850 mb.

A complementarity between SSM/I and Humsat humidity retrievals exists in the sense that Humsat retrievals of specific humidity are better than the trial fields at the higher levels, whereas for TPW, SSM/I retrievals are comparable (slightly better) to the trial fields and are superior in quality to Humsat. As most of the atmospheric humidity is concentrated in the lowest 2 km of the troposphere, the contribution of the SSM/I should be largest in the lowest levels.

b. Assimilation of TPW and specific humidity

Before one can perform the computation of q_a or the analyzed field in (7), one has to define the background error autocovariance matrix $\mathbf{P}$, and the analog for the satellite retrieval errors O_h and O_m. Error statistics were computed according to (18). Because there is a strong dependence of the error statistics with latitude and season, they were computed separately for each latitude band of 30° and for the June and March datasets.

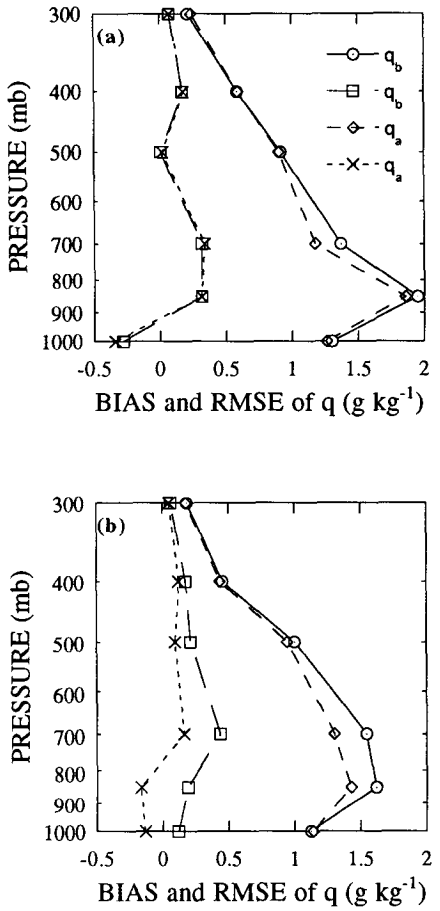


FIG. 2. Assimilation of SSM/I total precipitable water. Both the bias (the two left-most curves in each graph) and the rmse of the specific humidity are illustrated (the two right-most curves in each graph). Here q_b represents the background (6-h trial field) specific humidity and q_a represents the analyzed field. (a) June 1991 dataset obtained with the F08 satellite (results are similar for the F10 satellite), and (b) March 1992 dataset obtained with the F10 satellite (results are similar for the F11 satellite).

Background error statistics (difference between background field and radiosondes) were computed from a dataset that included both radiosondes and trial field values interpolated at the location of the radiosondes. The June dataset extended from 29 May 1991 to 12 June 1991 and the March dataset extended from 26 February 1992 to 12 March 1992. As expected, the magnitudes of the error autocovariances in the Tropics are similar for both the June dataset and March dataset. Table 3a illustrates the error statistics for March only. In the midlatitudes, the errors vary considerably with seasons. The smallest errors are obtained for the March northern midlatitude band and results from the fact that lower values of q occur during colder and hence drier periods (can contain less water vapor). The mean error at 1000 mb is positive for both the March (Table 3a) and June (not shown) datasets, which indicates that the surface 6-h trial field is too moist.

Humsat specific humidity error statistics for March are listed in Table 3b. The number of collocations is substantially smaller than for the background field, which in certain cases produces a very small sample size (e.g., 30°–60°S). This is because the background field errors are computed over the globe, whereas Humsat errors are computed over the GOES window. If one considers the diagonal terms only (variance), trial field errors usually are smaller than those for Humsat in the lower levels (1000, 850, and 700 mb) and are larger for the higher levels.

The error variance for the SSM/I retrievals (O_m) was computed for the same datasets as used in Table 1 by latitude bands of 30°. However, because of the small sample size, the number of cases for some midlatitude bands dropped to as low as 5. Therefore, an estimate for the error variance as a function of latitude and season had to be used. The estimate is based on a relationship for the relative uncertainty of TPW as a function of latitude and season (Li and Leighton 1993). In short, the magnitude of the error is related to the magnitude of the estimate. The details are given in appendix C.

Satellite retrievals of SSM/I TPW and Humsat q were assimilated by solving (7). Assimilation of SSM/I TPW was performed for the SSM/I datasets as given in Table 1. The assimilation reduced the TPW rmse with respect to that of the 6-h trial field (background field) by as little as 0.2 and as much as 1.5 kg m^{-2} as indicated in Table 1. Figure 2 illustrates the rmse reduction (as a function of pressure) between the analyzed field q_a and the background field q_b. The largest reductions occur at the 700- and 850-mb levels. There is almost no reduction in rmse at the surface (1000-mb level). The integral operator $\mathbf{L}_m$ based on the trapezoidal rule took on the following value: $\mathbf{L}_m = (1/2g)(150, 300, 350, 300, 200, 100)$ mb. Thus the weight for the surface level is one-half that of the next level up. This will limit corrections at the surface and implies that the vertical integral operator should be defined and tested carefully.

Assimilation of Humsat q was evaluated for the datasets presented in Table 2. The resulting TPW rmse was reduced by between 0.3 and 0.9 kg m^{-2}. This is about one-half the reduction by SSM/I TPW. However, for the assimilation of Humsat, both ocean and land surfaces are considered. Only a small percentage of the collocation sample is over ocean surfaces (14% of collocations for March and 26% for collocations for June). The q rmse profiles for the background field, the analyzed field and Humsat q are illustrated in Fig. 1. The rmse of the analyzed field are reduced at all levels but by only a small amount. The corrections that occur for the levels above 600 mb are mostly due to the contribution of the WV channel. The rmse of the background q at the surface is the largest as is typical over land surfaces. As is illustrated in Fig. 2, over ocean surfaces, the rmse of the 6-h trial field of q is smaller

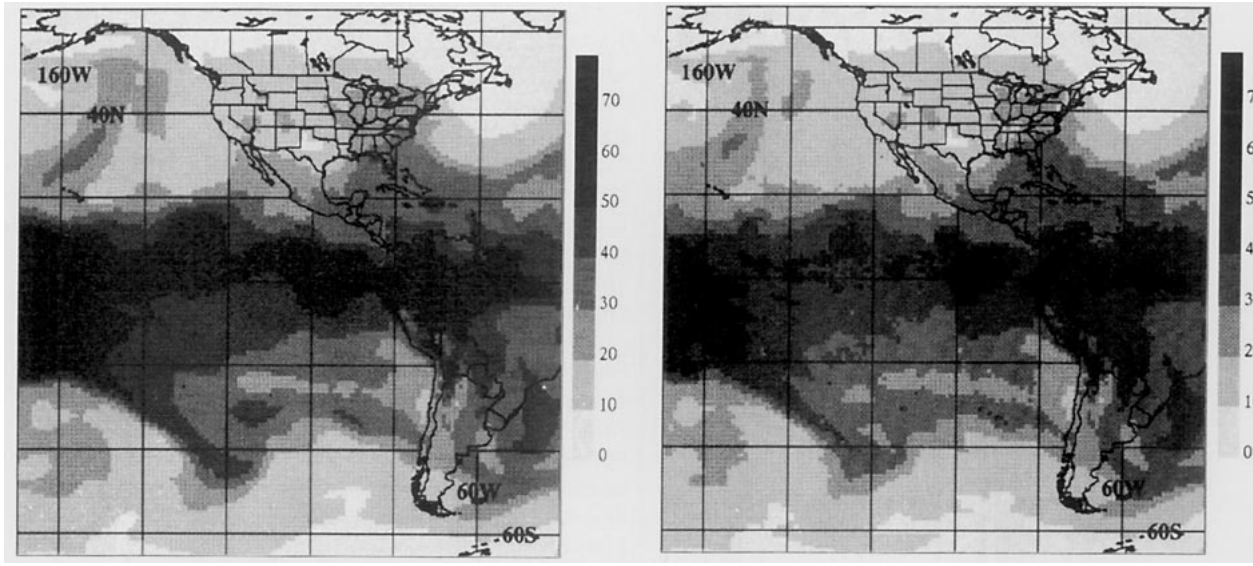


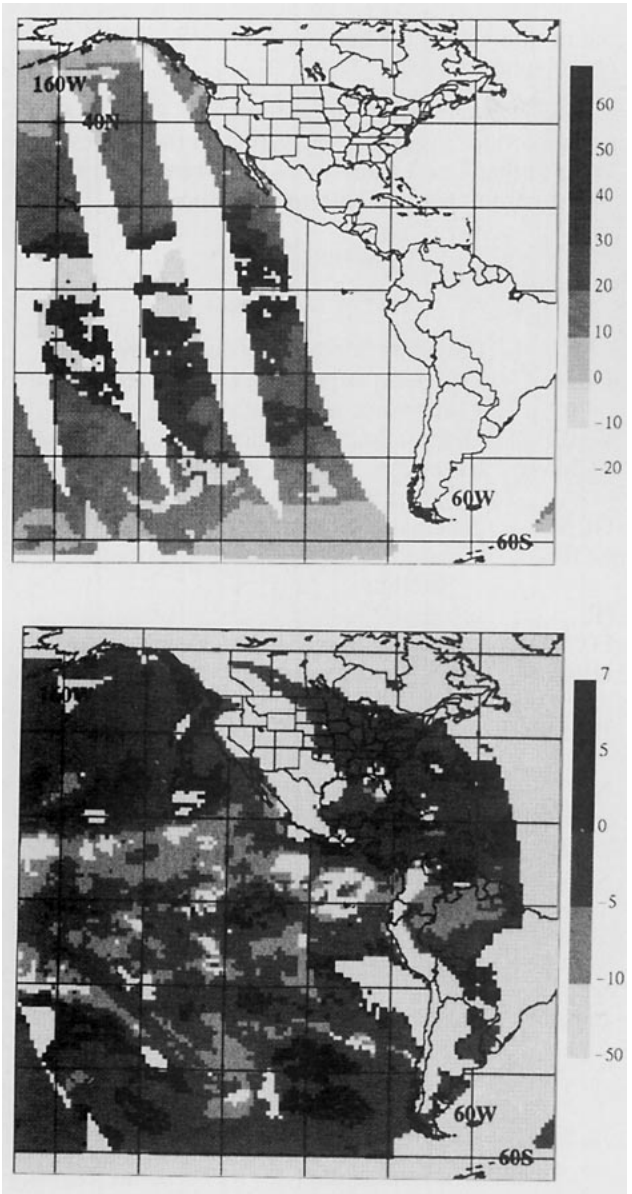
FIG. 3. Assimilation of SSM/I TPW and Humsat q for 7 March 1992 at 0600 UTC. The fields illustrated represent total precipitable water. Units are in kilograms per square meter, and the geographical grid spacing is 20°. (a) Background field (6-h trial field), and (b) analyzed field of TPW.

at the surface than at the 850- and 700-mb levels. This is typical over ocean surfaces and is due to the lower variability of q at the ocean surface.

An example of assimilation for 0600 UTC 7 March 1993 was also performed (Figs. 3 and 4) to illustrate the impact of assimilating (SSM/I TPW and Humsat q) on the resulting analyzed TPW field. The area analyzed extends between 60°S and 60°N latitude and 190° and 320° longitude (with a grid of 1.0°). The 6-h trial field of TPW is illustrated in Fig. 3a and is relatively smooth because of the fairly low resolution (T79) of the model. The ITCZ is noticeable with TPW values above 50 kg m^{-2}. Dry areas are found on the west coasts of South America and the United States. A large frontal system is located in the South Pacific Ocean (40°S, 120°W). Figure 4a illustrates the SSM/I TPW composite that was obtained by retaining pixels with times closest to 0600 UTC (±3 h) for both ascending and descending passes. As it turns out, this particular composite mostly contains F10 ascending orbits. The higher resolution (0.3°) SSM/I data were combined by averaging the 12 pixels nearest to each grid point of the 1° analysis grid. Values of −10.0 indicate missing values due to the application of the precipitation screen and are obviously not used for the analysis (section 2a). Values of −20 indicate missing data. The frontal system mentioned above is also present, but most of the data in the core is absent due to the precipitation screen. Humsat data illustrated in Fig. 4b was obtained within 30 min prior to 0600 UTC. The missing data in the Humsat image (indicated by −10 values) are due to the application of the rejection criteria described in appendix B. Essentially, data are missing over mountainous areas and where the differences between retrievals and the background field (or trial field) are very large. A broad region was rejected off the coast of Chile where Humsat detected low clouds but inferred too much moisture from the surface to 700 mb, an indication that these clouds are relatively thin.

The analyzed field of TPW (Fig. 3b) is a result of the assimilation of Humsat q and/or SSM/I TPW when available. If no satellite data are available, then the analyzed field is set equal to the background field. The difference between the analyzed field and the 6-h forecast is illustrated in Fig. 4c. Negative values of −50 indicate missing data. Most adjustments to the background field are bound between 0.0 and 10.0 kg m^{-2} (absolute values). Large areas of negative adjustments show that the background field of TPW is generally more humid. The frontal system in the South Pacific (mentioned above) is less humid for the analyzed field. Some of the drier zones in the background field are too dry (e.g., 40°S, 100°W). The intertropical convergence zone (ITCZ) in the model (Fig. 3a) appears to be rather uniformly moist as opposed to the satellite image (Fig. 4b), which indicates some clear sky breaks in the ITCZ. The analyzed field is not as smooth as the background field, which is desirable. There is variability at smaller horizontal scales in the satellite data. This effect may disappear somewhat under 3D variational analysis where horizontal error autocorrelations will also be considered.

Assessment of the errors using collocations cannot be performed for this case, because the number of collocations with radiosondes is too low. However, a



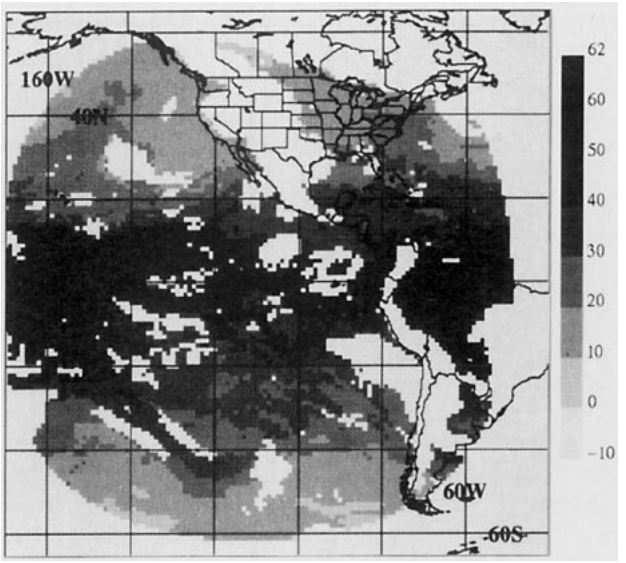


FIG. 4. Total precipitable water (units are in kilograms per square meter and the geographical grid spacing is 20°) for 0600 UTC 7 March 1992. (a) SSM/I TPW retrievals. (Missing values are represented by values of −20 and values of −10 illustrate where the precipitation filter was applied.) (b) Humsat q (integrated to give TPW) retrievals. (Missing values represented by values of −10.) (c) Analyzed field minus background field as illustrated in Figs. 3b and 3a, respectively. (Missing values represented by values of −50 kg m^{-2}.)

theoretical error estimate ($\mathbf{P}_a$) may be obtained via (10). Bias corrected rmse values for the analyzed field of TPW given by $\mathbf{L}_m\mathbf{P}_a\mathbf{L}_m^T$ [see (15)] are listed in Table 4. The theoretical bias corrected rmse for TPW nearly decreases by one-half with respect to the background field (Table 4, second column) when both satellite data sources are assimilated (fifth column). As expected, the assimilation of SSM/I TPW would lead to the largest decrease in rmse (third column) except for the Northern Hemisphere midlatitudes where it is comparable with the assimilation of Humsat q (fourth column). The errors for the 6-h trial field of q and Humsat q are not fully independent, because the 6-h temperature trial field was used to obtain the retrievals of Humsat q, and, therefore, (10) is not strictly applicable.

TABLE 4. Theoretical estimate of analyzed total precipitable water bias-corrected rmse (kg m^{-2}) (March 1992).

Latitude band	TPW_b	TPW_a (SSM/I)	TPW_a (Humsat)	TPW_a (SSM/I + Humsat)
30°–60°N	2.1	1.6	1.7	1.4
0°–30°N	5.4	3.3	3.7	2.7
0°–30°S	7.0	4.2	5.0	3.6
30°–60°S	5.0	3.0	3.6	2.6

4. Conclusions

A variational analysis study was performed by concentrating on a 1D variational analysis technique for specific humidity and total precipitable water. First, the accuracy of the satellite retrieved fields was esti-

mated by computing the rmse using collocated radiosondes. The retrieval of SSM/I TPW over the oceans had an rmse of 4.7 kg m^{-2}. This is somewhat smaller than the rmse error of the 6-h trial field (5.0 kg m^{-2}). The retrieval of DPD obtained from a combination of GOES radiances and ancillary data that was subsequently integrated to give TPW had an rmse of 4.6 kg m^{-2} (in the midlatitudes), while trial field TPW errors were 4.4 kg m^{-2}. In the Tropics the rmse were higher and, respectively, equal to 6.8 and 5.9 kg m^{-2}. These results are in agreement with prior findings. As expected, the rmse of q was smaller at the surface than at 700 and 850 mb over oceanic surfaces but not over land.

Assimilation of SSM/I TPW and Humsat q was first performed separately at the collocations with radiosondes. The results showed a reduction in the rmse (with respect to the 6-h trial field) for the total precipitable water of 1 kg m^{-2} when SSM/I TPW (limited to ocean surfaces) was assimilated and 0.6 kg m^{-2} when Humsat q was assimilated. The reduction in TPW rmse for the SSM/I occurred mainly at the levels of 700 and 850 mb. For the Humsat q assimilation, a small reduction in the rmse of q occurred everywhere. The corrections above 600 mb were mainly due to the effect of the WV channel in the retrievals. Thus, the SSM/I mainly determines humidity at the lower levels where most of the atmospheric water vapor is located, whereas Humsat provides important information at the higher levels and nonnegligible information at the surface in the midlatitudes and at other lower levels when SSM/I data are not available.

The multiple collocation study performed in our study (SSM/I, Humsat, and radiosondes) was not fully satisfactory because of the small sample size of radiosondes against which the data could be verified. An alternative to this problem would be to use larger datasets or to obtain vast amounts of ground data from intensive campaigns, for example.

In the future, we expect to implement changes in the Humsat algorithm that will yield improved specific humidity retrievals in the Tropics. Efforts are under way to develop a new retrieval algorithm that takes into account anticipated data from future GOES-Next satellites. We also hope to use sounding data from the DMSP SSM/T-2 (Nehrkorn et al. 1993), which will allow even better specification of tropospheric water vapor as well as TOVS humidity soundings.

This study was a necessary step toward a more realistic 3D variational analysis technique to assimilate both total precipitable water and vertical profiles of specific humidity. Such a system will produce more realistic global analyses of water vapor by explicitly accounting for the full 3D error covariances of the data and the background field. A 3D variational analysis system is currently being developed at the Atmospheric Environment Service of Canada for eventual implementation within the operational NWP analysis/forecast system.

Acknowledgments. The authors are grateful to C. Chouinard for insightful comments on the manuscript. We also thank two anonymous reviewers for their pertinent comments and helpful suggestions.

APPENDIX A

List of Acronyms

CMC	Canadian Meteorological Centre
DPD	Dewpoint depression (temperature minus dewpoint temperature)
DMSP	Defense Meteorological Satellite Program
ECMWF	European Centre for Medium-Range Weather Forecasts
GCM	general circulation model
GOES	Geostationary Operational Environmental Satellite
IR	infrared
ITCZ	intertropical convergence zone
HIRS	High-Resolution Infrared Radiometer Sounder (second instrument)
Humsat	specific humidity retrieval algorithm developed by Garand (1993)
NWP	numerical weather prediction
rmse	root-mean-square error
SSM/I	Special Sensor Microwave/Imager
SST	sea surface temperature
TOVS	*TIROS-N* Operational Vertical Sounder
TBWV	water vapor brightness temperature
TPW	total precipitable water
VIS	visible
WV	water vapor

APPENDIX B

Update of Humsat Retrieval Scheme

The main modifications to the original Humsat technique (Garand 1993) are the following.

1) An additional cloud class was added with mean cloud fraction of 85%, mean cloud-top pressure of 775 mb, and mean cloud albedo of 50%. The IR scheme now has 8 classes and the VIS–IR scheme has 10 classes.

2) Quality control was implemented. No profiles are sent to the analysis system in the following cases:

(a) The temperature profile as inferred from the 6-h trial field is nearly isothermal in the lower troposphere, making the determination of the cloud fraction difficult when visible data is not available. The rejection criterion is met when the 6-h temperature trial field at 1000 mb minus 10 K is colder than the 600-mb temperature. Arctic air masses are rejected by this criterion in winter, but virtually no rejections occur in the tropical band 30°S–30°N.

(b) The 6-h trial field (TPW_b) and Humsat estimates of TPW (TPW_h) differ by more than 2.1 times an estimate of the rmse of TPW_b. The method used to compute an estimate for the rmse of TPW_b is described in appendix C. Most rejections over oceans are due to that criterion (see Fig. 4b).

(c) Humsat cannot detect cases where the atmosphere is humid at high levels (300–500 mb), but very dry at lower levels (700 and 850 mb). When the 6-h trial field indicates such a structure and Humsat detects high-level humidity, the profiles are not used. This occurs in about 1% of the cases, essentially ahead of depressions where cirrus clouds prevail.

Humsat does make retrievals over mountains. However, in this study, we considered only regions where data were available at all the six lowest mandatory pressure levels.

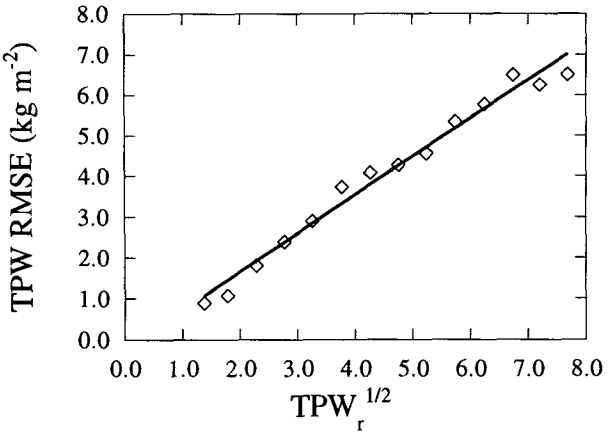


FIG. C1. Illustration of the validity of Eq. (C1) for the 6-h trial field of total precipitable water. The x axis represents binned values of $TPW_r^{1/2}$ [bin width of 0.5 (kg m^{-2})$^{1/2}$] for the radiosonde datasets used in this study (combined June 1991 and March 1992 data), and the y axis represents the TPW rmse of the 6-h trial field with respect to the radiosondes for each bin.

APPENDIX C

Estimate of TPW Error Variance

a. Determination of SSM/I TPW error variance

For a climatological dataset of total precipitable water, one may assume that the uncertainty or error on the estimate of daily values of TPW is given by the standard deviation (δ TPW) of daily TPW with respect to monthly mean TPW (m TPW). Based on the climatological data of Oort and Rasmusson (1971) and Oort (1983), Li and Leighton (1993) (their Fig. 20) have shown that

$$\delta\text{TPW} \approx a_0 + a_1(m\text{TPW})^{1/2}, \qquad \text{(C1)}$$

where the units of TPW are kilograms per square meter and a_0 and a_1 are constants equal to 0.0 and 2.2, respectively. Thus, $\delta\text{TPW}(\text{TPW})^{-1/2}$ does not depend on latitude or season provided a_0 is small.

In this study, we apply the same relationship to the rmse of the SSM/I retrievals. If we assume that the rmse in the Tropics is of the order of 5.0 kg m^{-2} and use the mean radiosonde TPW for March (43 kg m^{-2}) in (C1), then a_1 is estimated to be 0.75 (a_0 is assumed to be zero). Since the errors for the background field and Humsat retrievals were specified by latitude band and season, we do the same for the SSM/I TPW by applying (C1) with values for mTPW taken to be the latitude band means obtained from radiosonde observations. The resulting rmse estimates are listed in Table C1. The values obtained for the errors are comparable with those in Alishouse et al. (1990).

TABLE C1. Estimated values of rmse for SSM/I total precipitable water (kg m^{-2}).

	June 1991		March 1992	
Latitude band	Radiosonde mean	Estimated rmse (SSM/I)	Radiosonde mean	Estimated rmse (SSM/I)
30°–60°N	25.6	3.8	9.9	2.4
0°–30°N	43.0	5.0	30.3	4.1
0°–30°S	33.7	4.4	46.7	5.1
30°–60°S	17.8	3.2	26.6	3.8

b. Estimate of rmse of the 6-h trial field of total precipitable water

The same relationship as (C1) is shown to hold for the rmse of the 6-h trial field of precipitable water (Fig. C1) with $a_1 = 0.97$ and $a_0 = -0.24$. Note that the value of the intercept is small. The x axis in Fig. C1 represents binned values of $\text{TPW}^{1/2}$ [bin width of 0.5 (kg m^{-2})$^{1/2}$] for the radiosonde datasets used in this study (combined June 1991 and March 1992 data) and the y axis represents the TPW rmse of the 6-h trial field with respect to the radiosondes for each bin. Essentially then, TPW errors may be defined as a function of the square root of the absolute value of TPW, which avoids the need to specify the errors per latitude band and season. The absolute error decreases with latitude as expected, but the relative error increases.

REFERENCES

Alishouse, J. C., S. A. Snyder, J. Vongsathorn, and R. R. Ferraro, 1990: Determination of oceanic total precipitable water from the SSM/I. *IEEE Trans. Geosci. Remote Sens.*, **28,** 811–816.

Alliss, R. J., G. D. Sandlin, and S. W. Chang, 1993: Applications of SSM/I data in the analysis of Hurricane Florence (1988). *J. Appl. Meteor.*, **32,** 1581–1591.

Bauer, P., and P. Schluessel, 1993: Rainfall, total water, ice water, and water vapor over sea from polarized microwave simulations and Special Sensor Microwave/Imager data. *J. Geophys. Res.*, **98,** 20 737–20 759.

Chang, S. W., R. J. Alliss, S. Raman, and J.-J. Shi, 1993: SSM/I observations of ERICA IOP 4 marine cyclone: A comparison with in situ observations and model simulation. *Mon. Wea. Rev.*, **121,** 2452–2464.

Colton, M. C., 1993: *Proc. Shared Processing Network DMSP SSM/I Algorithm Symp.*, Monterey, CA, Fleet Numerical Oceanography Centre.

Daley, R., 1991: *Atmospheric Data Analysis.* Cambridge University Press, 457 pp.

Duda, R. O., and P. E. Hart, 1973: *Pattern Recognition and Scene Analysis.* John Wiley and Sons, 482 pp.

Garand, L., 1993: A pattern recognition technique for retrieving humidity profiles from Meteosat or GOES imagery. *J. Appl. Meteor.*, **32,** 1592–1607.

Gauthier, P., P. Courtier, and P. Moll, 1993: Assimilation of simulated wind lidar data with a Kalman filter. *Mon. Wea. Rev.*, **121,** 1803–1820.

Ghil, M., 1989: Meteorological data assimilation for oceanographers. Part I: Description and theoretical framework. *Dyn. Atmos. Oceans.*, **13,** 171–218.

Gilbert, J. C., and C. LeMaréchal, 1989: Some numerical experiments with variable-storage quasi-Newton algorithms. *Math. Program.*, **45,** 407–435.

Haydu, K. J., and T. N. Krishnamurti, 1981: Moisture analysis from radiosonde and microwave spectrometer data. *J. Appl. Meteor.*, **20,** 1177–1191.

Hollinger, J. P., 1991: DMSP Special Sensor Microwave/Imager calibration/validation. Final Report, Vol. 2.

——, J. L. Pierce, and G. A. Poe, 1990: SSM/I instrument evaluation. *IEEE Geosci. Remote Sens.*, **28,** 781–790.

Illari, L., 1989: The quality of satellite precipitable water content data and their impact on analyzed moisture fields. *Tellus.*, **41A,** 319–337.

Isaacs, R. G., R. N. Hoffman, and L. D. Kaplan, 1986: Satellite remote sensing of meteorological parameters for global numerical weather prediction. *Rev. Geophys.*, **24,** 701–743.

——, E. Kalnay, G. Ohring, and R. A. McClatchey, 1992: *Proc. First NMC/NESDIS/DOD Conf. on DMSP Retrieval Products,* NMC, 266 pp.

Kuo, Y.-H., Y.-R. Guo, and E. R. Westwater, 1993: Assimilation of precipitable water measurements into a mesoscale numerical model. *Mon. Wea. Rev.*, **121,** 1215–1238.

Li, Z., and H. G. Leighton, 1993: Estimation of SW flux absorbed at the surface from TOA reflected flux. *J. Climate.*, **6,** 317–330.

Liu, W. L., W. Tang, and F. J. Wentz, 1992: Precipitable water and surface humidity over global oceans from special sensor microwave images and European Centre for Medium-Range Weather Forecasts. *J. Geophys. Res.*, **97,** 2251–2264.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112,** 1177–1194.

——, 1988: Optimal nonlinear objective analysis. *Quart. J. Roy. Meteor. Soc.*, **114,** 205–240.

Mintz, Y., 1964: Very-long term global integration of the primitive equations of atmospheric motion. Tech. Note No. 66, World Meteorological Organization, 141–167.

Mitchell, H. L., C. Charette, C. Chouinard, and B. Brasnett, 1990: Revised interpolation statistics for the Canadian data assimilation procedure: Their derivation and application. *Mon. Wea. Rev.*, **118,** 1591–1614.

——, ——, S. J. Lambert, J. Halle, and C. Chouinard, 1993: The Canadian global data assimilation system: Description and evaluation. *Mon. Wea. Rev.*, **121,** 1467–1492.

Nehrkorn, T., R. N. Hoffman, J.-F. Louis, R. G. Isaacs, and J.-L. Moncet, 1993: Analysis and forecast improvements from simulated satellite water vapor profiles and rainfall using a global data assimilation system. *Mon. Wea. Rev.*, **121,** 2727–2739.

Oort, A. H., 1983: Global atmospheric circulation statistics. NOAA Professional Paper No. 14, U.S. Govt. Printing Office, Washington, DC, 180 pp.

——, and E. M. Rasmussen, 1971: Atmospheric Circulation Statistics. NOAA Professional Paper No. 5, U.S. Govt. Printing Office, Washington, DC, 232 pp.

Ritchie, H., 1991: Application of the semi-Lagrangian method to a multilevel spectral primitive-equations model. *Quart. J. Roy. Meteor. Soc.*, **117,** 91–106.

Schluessel, P., and W. J. Emery, 1990: Atmospheric water vapor over oceans from SSM/I measurements. *Int. J. Remote Sens.*, **11,** 753–766.

Schmetz, J., and L. van de Berg, 1991: Estimation of precipitable water from Meteosat infrared window radiances over sea. *Beitr. Phys. Atmos.*, **64,** 93–102.

Smagorinsky, J., K. Miyakoda, and R. Stricker, 1970: The relative importance of variables in initial conditions for dynamical weather prediction. *Tellus,* **22,** 141–157.

Smith, W. L., and H. M. Woolf, 1976: The use of eigenvectors of statistical covariance matrices for interpreting satellite sounding radiometer observations. *J. Atmos. Sci.*, **33,** 1127–1140.

Tjemkes, S. A., and G. L. Stephens, 1990: Intercomparison between microwave and infrared observations of precipitable water. *Proc. Fifth Satellite Meteorology and Oceanography, Conf.*, London, England, Amer. Meteor. Soc., 82–86.

Wentz, F. J., 1991: User's manual SSM/I antenna temperature tapes. Revision 1. RSS Tech. Rep. 120191, Remote Sensing Systems, Santa Rosa, CA.