

Estimating moisture transport over oceans using space-based observations

W. Timothy Liu and Wenqing Tang

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

Received 29 July 2004; revised 29 October 2004; accepted 21 January 2005; published 17 May 2005.

[1] The moisture transport integrated over the depth of the atmosphere (Θ) is estimated over oceans using satellite data. The transport is the product of the precipitable water and an equivalent velocity (\mathbf{u}_e), which, by definition, is the depth-averaged wind velocity weighted by humidity. An artificial neural network is employed to construct a relation between the surface wind velocity measured by the spaceborne scatterometer and coincident \mathbf{u}_e derived using humidity and wind profiles measured by rawinsondes and produced by reanalysis of operational numerical weather prediction (NWP). On the basis of this relation, Θ fields are produced over global tropical and subtropical oceans (40°N–40°S) at 0.25° latitude-longitude and twice daily resolutions from August 1999 to December 2003 using surface wind vector from QuikSCAT and precipitable water from the Tropical Rain Measuring Mission. The derived \mathbf{u}_e were found to capture the major temporal variability when compared with radiosonde measurements. The average error over global oceans, when compared with NWP data, was comparable with the instrument accuracy specification of space-based scatterometers. The global distribution exhibits the known characteristics of, and reveals more detailed variability than in, previous data.

Citation: Liu, W. T., and W. Tang (2005), Estimating moisture transport over oceans using space-based observations, *J. Geophys. Res.*, 110, D10101, doi:10.1029/2004JD005300.

1. Introduction

[2] The moisture transport integrated over the depth of the atmosphere (Θ) is an important branch of the global hydrological cycle. It redistributes moisture (and latent heat), with an overall effect of transforming the input from evaporation over oceans into the deposition as precipitation over land. Fluctuations in the Θ fields can significantly influence flood or drought, with important economic consequences. With appropriate spatial and temporal resolution, Θ fields can be used to estimate the regional differences between evaporation and precipitation and can help to identify major water sources or sinks. This is particularly important over the oceans, where measurements of surface freshwater flux are extremely sparse. Through the surface freshwater flux the atmosphere affects the circulation, salinity, and density stratification of the upper ocean [e.g., Lagerloef, 2002]. Estimating Θ fields over global oceans with sufficient resolutions, however, is a great challenge.

[3] Starr and White [1955] compiled zonally averaged Θ in the Northern Hemisphere using wind and humidity profiles in rawinsonde observations (RAOB) almost half a century ago. Since then, there have been many studies to improve the estimation of Θ using RAOB and to apply it in determining surface water flux over various regions and time periods [e.g., Starr and Peixoto, 1958; Rasmusson, 1967; Rosen et al., 1979; Peixoto and Oort, 1983; Bryan and Oort, 1984]. Another source of humidity and wind

profiles, from which Θ could be computed, is the analysis product of numerical weather prediction (NWP) models [e.g., Rasmusson and Mo, 1996; Trenberth and Guillemot, 1995]. The Θ computed from NWP products showed agreement with those from RAOB in the large-scale patterns, but significant uncertainties and differences remain.

[4] Satellite observations may provide better coverage than RAOB and are not affected by the deficiency in physics and parameterizations of numerical models. Wind profiles, however, are not measured from space at present, but the geostrophic wind shear (the vector difference between geostrophic winds at two levels) can be derived from the horizontal temperature gradients of the layers through the thermal wind equation. Slonaker and Van Woert [1999] computed the meridional moisture transport into Antarctica by first deriving the geostrophic winds at five levels relative to the surface wind measurements using the three-dimensional temperature fields from TIROS operational vertical sounder in the high southern latitudes. In this study, we describe a simple and practical method of estimating Θ from satellite observations, which avoids the geostrophic assumption and the coarse resolution of spaceborne atmospheric sounders.

2. Methodology

2.1. Principles

[5] Integration of the equation for conservation of atmospheric water vapor in the vertical gives

$$\frac{\partial W}{\partial t} + \nabla \cdot \Theta = E - P, \quad (1)$$

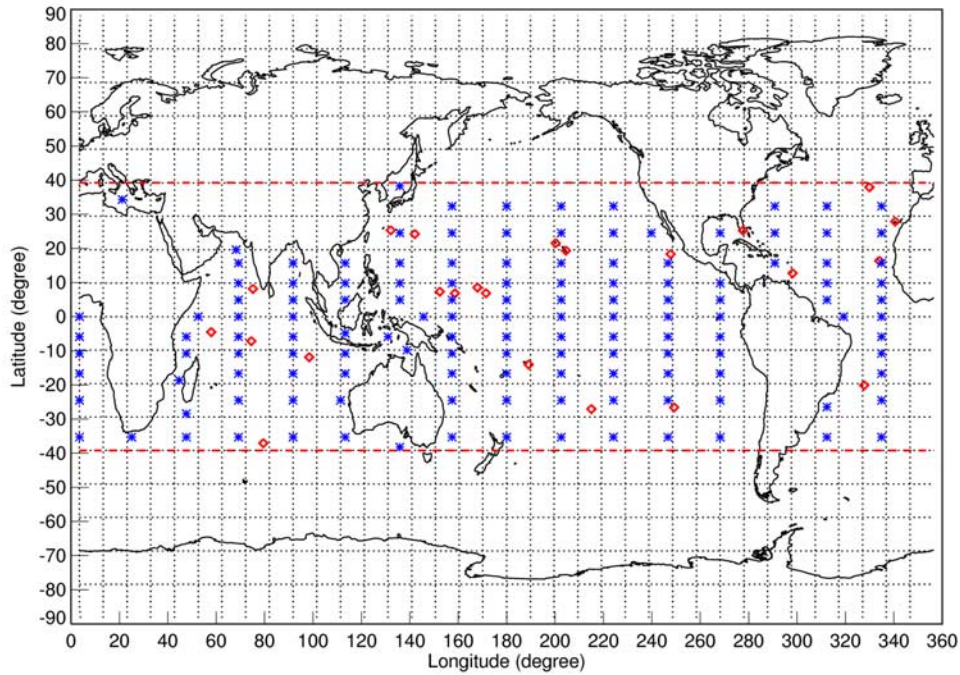


Figure 1. Locations of rawinsonde observations (RAOB) (red) and National Centers for Environmental Prediction (NCEP) (blue) data used in the artificial neural network training.

where

$$\Theta = \frac{1}{g} \int_0^{p_s} q \mathbf{u} dp \quad (2)$$

is the integrated moisture advection and

$$W = \frac{1}{g} \int_0^{p_s} q dp \quad (3)$$

is the precipitable water. In equations (2) and (3), g is the acceleration due to gravity, p is the pressure, p_s is the pressure at the surface, and q and \mathbf{u} are the specific humidity and wind vector at a certain level, respectively. E and P are the evaporation and precipitation at the surface, respectively, and t is time.

[6] The computation of Θ in equation (2) requires measurement of the vertical profiles of wind vector and humidity in the atmosphere, which traditionally come from RAOB. Over oceans, rawinsonde data are sparse. Liu [1993] proposed a method to estimate Θ using surface wind vector (\mathbf{u}_s) and W measured by space-based scatterometers and microwave radiometers.

[7] As suggested by Liu [1993], an equivalent velocity is defined as $\mathbf{u}_e = \Theta/W$, which is the depth-averaged velocity weighted by the humidity. Since W can be accurately measured by space-based microwave radiometers and relatively long records exist from the Special Sensor Microwave Imager, Tropical Rain Measuring Mission (TRMM) Microwave Imager (TMI) [e.g., Kummerow *et al.*, 2000], and others, the problem of estimating Θ is essentially the determination of \mathbf{u}_e . Our objective is to develop a quantitative relation between \mathbf{u}_e and \mathbf{u}_s , measured by the scatterometers [Liu, 2002]. The variability of vertical humidity distribution has been extensively studied [e.g., Liu *et al.*,

1991]. The dominant mode of variability was found to peak at the top of the boundary layer. Using W derived by Liu [1987] from the microwave radiometer on Nimbus 7 and using 850-mbar cloud drift winds as \mathbf{u}_e , Heta and Mitsuta [1993] have estimated $E - P$ in the tropical Pacific from the divergence of Θ . Their results were consistent with climatological features. The top of the boundary layer is approximately at 850 mbar, and wind at this level is closely related to surface level wind through similarity relations [e.g., Brown and Liu, 1982]. In the past, oceanographers used to get ocean surface wind stress by multiplying the geostrophic winds (derived from pressure gradient) by a constant factor and by turning them by a constant angle. Deriving a simple relation between \mathbf{u}_e and \mathbf{u}_s appears to be feasible.

2.2. Training Database

[8] A database that relates input \mathbf{u}_s and the output target \mathbf{u}_e was established first. The quality of the database plays a crucial role in the usefulness of the derived model function. The output target was a combination from two data resources: (1) RAOB, which was quality controlled to best represent in situ measurements, and (2) reanalysis of the National Centers for Environmental Prediction (NCEP), which was included to achieve a global representation. There were 23 RAOB stations available, scattered in the Pacific, Atlantic, and Indian oceans between 40°S and 40°N, and NCEP data at 135 grid points were selected (Figure 1). NCEP provides the most timely and openly available NWP reanalysis. One year (2000) of data from both resources was used, including all available RAOB measurements, and only two records every 5 days were used for NCEP, at 0000 and 1200 UT. For each record of the blended RAOB/NCEP target a spatial-temporal collocation procedure was used to extract corresponding ocean surface

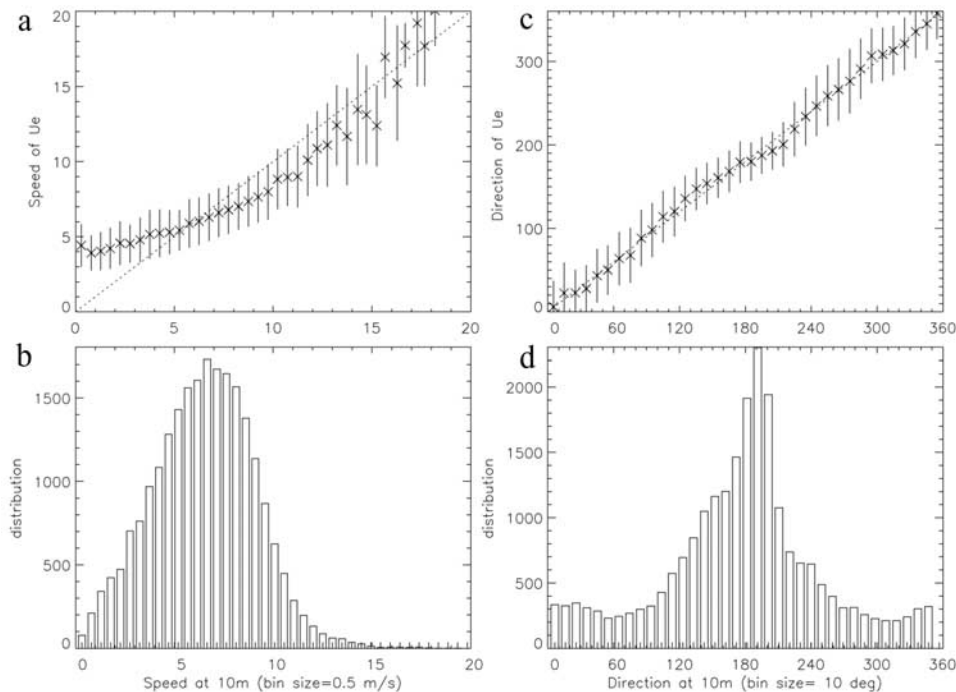


Figure 2. Equivalent velocity \mathbf{u}_e averaged in the corresponding bins of surface vector wind \mathbf{u}_s for (a) magnitude (bin size 0.5 m/s) and (b) direction (bin size 10°). Vertical bars indicate the standard deviation. (c and d) Distributions for magnitude and direction, respectively.

wind vectors \mathbf{u}_s from QuikSCAT. A 5-day moving average is applied to each parameter to filter out the high-frequency fluctuation. The established database contains a total of 22,925 records that are composed of $\sim 30\%$ from RAOB and 70% from NCEP. From the training data set plotted in Figure 2 the correlation between \mathbf{u}_s and \mathbf{u}_e in both magnitude and direction can be seen.

2.3. Artificial Neural Network

[9] The mapping from \mathbf{u}_s to \mathbf{u}_e embedded in the training database was simulated in an empirical model function through an artificial neural network (ANN) [Bishop, 1995]. The advantage of ANN over the multivariate regression method that was used by Liu [1993] is the ability to blend two sets of data (from NCEP and RAOB) together with various weights in addition to the ability to account for the nonlinearity between inputs and the outputs without requiring a functional form a priori. Figure 3 shows the structure of a neural network with input, output, and two hidden layers. Each layer contains a number of entities called neurons. Each neuron is basically a transfer function that produces an output signal based on weighted input signals plus a bias from the previous layer. The outputs from one layer become inputs to the next layer with another set of weights and biases. Neurons in the network are connected through these weights and biases. The so-called network training is actually a process that iteratively decides weights and biases to minimize the error between outputs produced by the network in response to the inputs and the corresponding training targets. Various network architecture, algorithms, and software tools exist for ANN. In this

study, the Neural Network Toolbox with MATLAB[®] [Demuth and Beale, 1998] was used to design, to develop, and to train the ANN.

[10] Through an iterative process of training, error assessment, modification, and training the final version of the network architecture used in this study is the multilayer feed forward network, with 7 neurons in the input layer and two hidden layers with 28 and 8 neurons each. The seven input variables are the (1) ocean surface wind speed, (2) direction, (3) latitude, and (4 and 5) sine and cosine of longitude and (6 and 7) sine and cosine of time. The values for time and longitude were represented in harmonic function to reflect the seasonal cycle and periodical condition. The output is the equivalent velocity in magnitude and direction. Finally, the matured model function was applied to \mathbf{u}_s from QuikSCAT [Liu, 2002] in conjunction with W measured by TMI [Wentz, 1997] to produce space-based estimation of Θ . The final product of Θ consists of twice daily maps in oceans from 40°S to 40°N at spatial resolution of 0.25° latitude by 0.25° longitude and covers the time period from August 1999 to August 2003.

3. Validation

[11] To assess the quality of the derived model function, a blended RAOB/NCEP data set from year 2001, independent from those used in model development, was assembled as the validation truth set $\{\mathbf{u}_e\}$. The data set to be validated, the derived equivalent velocity $\{\mathbf{u}'_e\}$ was generated by applying the model function to the ocean surface wind $\{\mathbf{u}_s\}$ extracted from QuikSCAT collocated in space and time with $\{\mathbf{u}_e\}$.

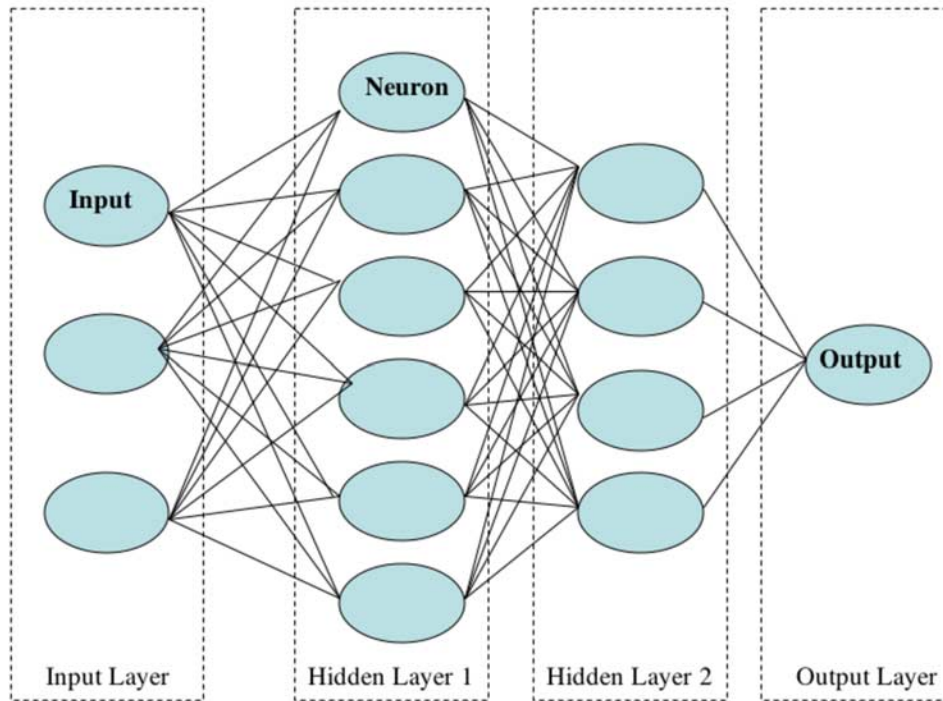


Figure 3. Diagram of a neural network with input, output, and two hidden layers. Each neuron represents a transfer function, and lines denote input and output signals between the corresponding neurons and weights and the associated biases.

The time series in Figure 4 shows that \mathbf{u}_e and \mathbf{u}'_e agree well in intraseasonal to seasonal variations. The station, being away from any landmass, is representative of oceanic conditions.

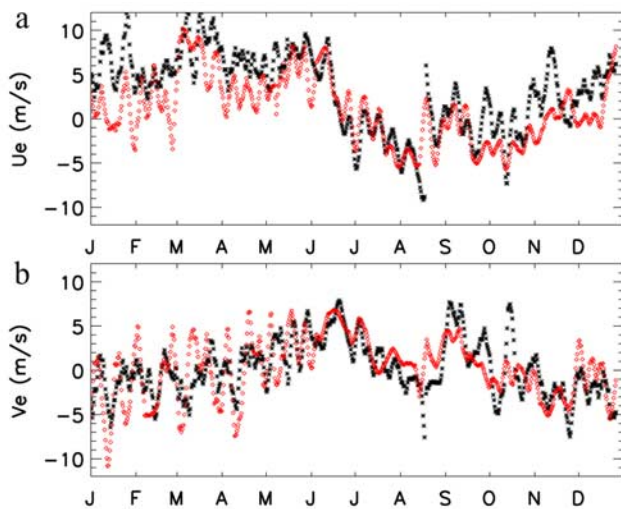


Figure 4. Time series of the (a) zonal (u_e) and (b) meridional (v_e) components of the equivalent velocity at Minamidaitojima (25.83°N, 131.23°E), computed from rawinsonde data (red) and derived from QuikSCAT winds (black) for the year 2001. A filter of 5-day moving average was applied to the twice daily data.

[12] There is a total of 82,218 records in the matched-up validation set $\{\mathbf{u}_e, \mathbf{u}'_e\}$, and the differences $\Delta = \mathbf{u}'_e - \mathbf{u}_e$, in magnitude and direction, were computed. Table 1 lists the validation results. The mean Δ are small, -0.45 m/s in magnitude and -2.14° in direction, and the standard deviation (SD) of the ensemble Δ for the direction appears to be large. The distributions of the mean and the SD of Δ , as a function of u_e , were examined. For speeds the mean Δ are positive and negative at weak and strong winds, respectively; the exclusion of weak winds causes the mean Δ to be more negative. Excluding weak winds decreases the SD of wind speed only slightly. The values of SD increase with wind speed, but the number of data decreases rapidly at strong winds. For wind direction the mean biases are small, almost negligible for $5 \text{ m/s} < u_e < 15 \text{ m/s}$. The values of SD are large at weak winds.

[13] To examine the large directional errors further, the direction of Δ was binned into two-dimensional space of the magnitude of \mathbf{u}_e and \mathbf{u}'_e and is presented in Figure 5. Contours represent the number of matchups in each two-dimensional bin. Those bins with number of matchups less than 0.05% of the total records were blacked out. The

Table 1. Statistical Results of $\mathbf{u}'_e - \mathbf{u}_e$

	Magnitude, m/s		Direction, deg		N
	Mean	SD	Mean	SD	
$u_e \geq 0$	-0.45	2.39	-2.14	52.37	79,976
$u_e \geq 3 \text{ m/s}$	-0.83	2.27	-1.73	45.50	68,705
$u_e \geq 5 \text{ m/s}$	-1.32	2.19	-1.00	39.10	52,051

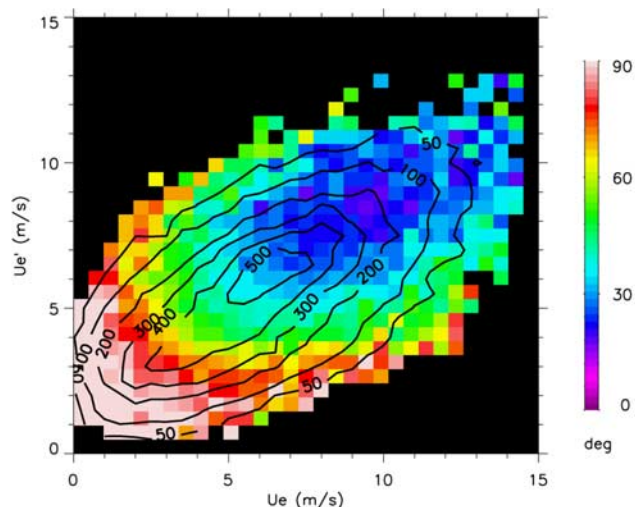


Figure 5. Standard deviation of directional differences between the equivalent velocities derived from QuikSCAT and collocated RAOB/NCEP data in the validation database. Contours represent the number of collocated pairs in bins of 0.5 m/s in the two-dimensional spaces. Bins with less than 0.05% of the total validation records are blacked out.

direction error clearly shows the model deficiency at low wind speed. In the range of medium to strong wind (≥ 5 m/s), directional SD is mostly less than 30° . It increases to 60° for u_e with magnitude between 3 and 5 m/s. The degradation of wind direction accuracy at weak winds for both space-based [e.g., Plant, 2000; Ebuchi et al., 2002] and in situ measurements is well known [e.g., Dobson et al., 1980].

4. Global Distribution

[14] The global pattern for satellite-derived Θ (Figure 6) is similar to NCEP results (not shown). The major feature is a year-round broad belt of easterly transport in the tropical ocean, interrupted by westerly transport in the Indian Ocean and vicinity. The locations and magnitudes of the maxima in the Pacific and Atlantic oceans are consistent with the results that Rosen et al. [1979] derived from 10-year rawinsonde data. The satellite estimation also clearly reveals the seasonal variation of monsoons in the Indian Ocean and the western Pacific Ocean.

[15] The winter and summer latitudinal variation of zonally averaged Θ shown in Figure 7 agrees in general with the climatology of Peixoto and Oort [1992], considering that the climatology was compiled over both land and

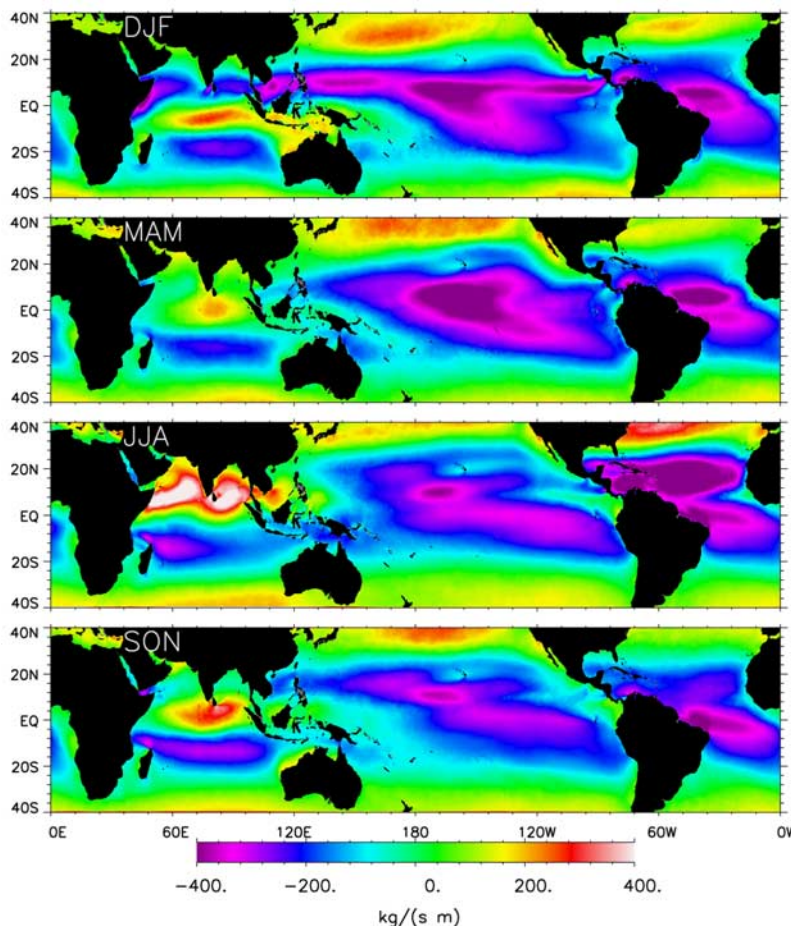


Figure 6a. Zonal component of the climatological seasonal means of Θ derived from QuikSCAT and Tropical Rain Measuring Mission (TRMM), averaged in the 3-months periods (as indicated) from September 1999 to August 2003.

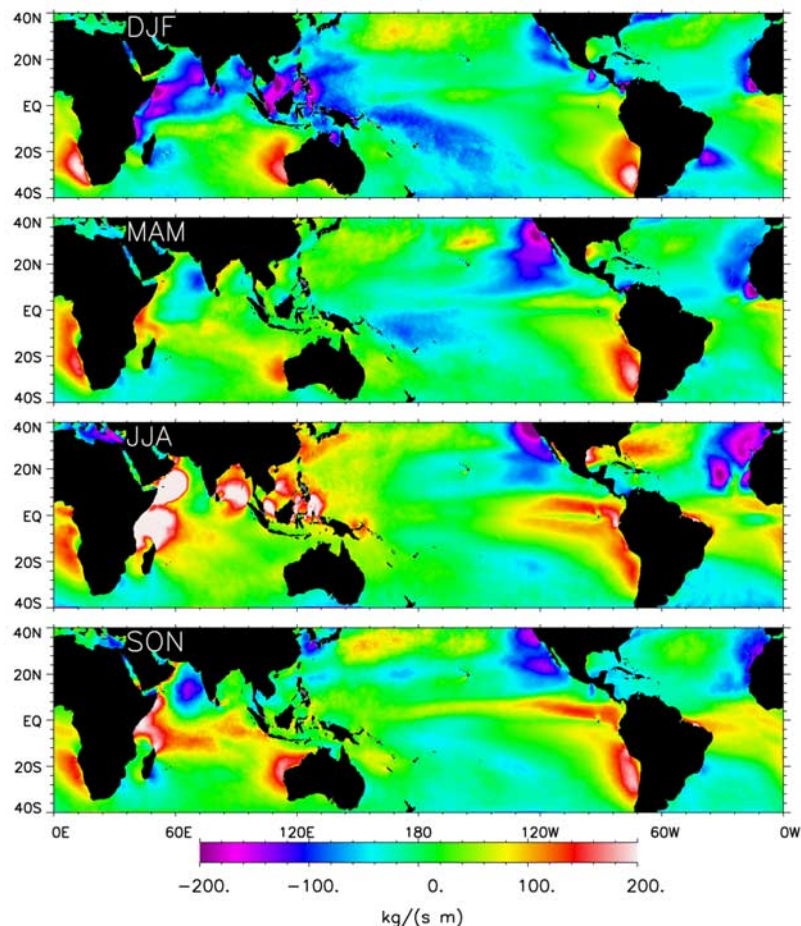


Figure 6b. Similar to Figure 6a but for meridional component.

ocean with most of the data from land. In boreal summer (June, July, and August), when the Indian and Asian monsoons are strong, rawinsonde data may underestimate the moisture transport in the tropical ocean. Satellite data add information on interannual variation and differences among ocean basins. There is little interannual change in the meridional moisture transport profiles except in the Pacific Ocean between the equator and 20°N. It shows that the southward moisture transport between 10°N and 20°N in the 2002–2003 winter is almost doubled in magnitude compared with 2 previous years. The anomalies may be related to the La Niña and El Niño episodes, which will be analyzed in future studies.

[16] The difference between Θ derived from satellite data and NCEP reanalysis can be detected in monitoring the oceanic influence of continental rainfall through monsoons. The computed Θ fields, as described at the end of section 2.3, were interpolated to some simplified coastal lines (straight-line approximation) of India. The components normal to these coastlines were computed (onshore advection is positive). A similar quantity was also computed from NCEP reanalysis data. Figure 8 shows that the annual cycles of Θ computed from satellite data are in phase with the precipitation integrated over the Indian subcontinent derived from TRMM. At the onset of the summer monsoon in May/June the magnitude of Θ directed into the continent is

higher than the total precipitation; the difference may indicate the transport of moisture overland to the north. The moisture influxes computed from NCEP reanalysis are much weaker during the summer monsoon, with the peaks occurring behind those of satellite-based Θ and precipitation.

5. Discussion

[17] A simple and innovative method of estimating Θ using space-based active and passive microwave sensors is presented. A fictitious \mathbf{u}_e (depth-averaged wind vector weighted by humidity) is introduced and related to surface wind vector. The relation is established by statistical tools, which compensate for the shortcoming in our knowledge of the physical processes. The relation may not work well when the surface winds are completely decoupled from the winds aloft. This practical methodology is developed with the awareness that simple methods, while useful, may not be sufficient for all purposes.

[18] In comparison with a validation data set composed of both radiosonde and NWP data the mean and standard deviation are found to be -0.8 and 2.3 m/s for magnitude and -1.7° and 45.5° for direction for \mathbf{u}_e larger than 3 m/s, as shown in Table 1. The instrument specifications of a series of space-based scatterometers are 2 m/s and 20° in

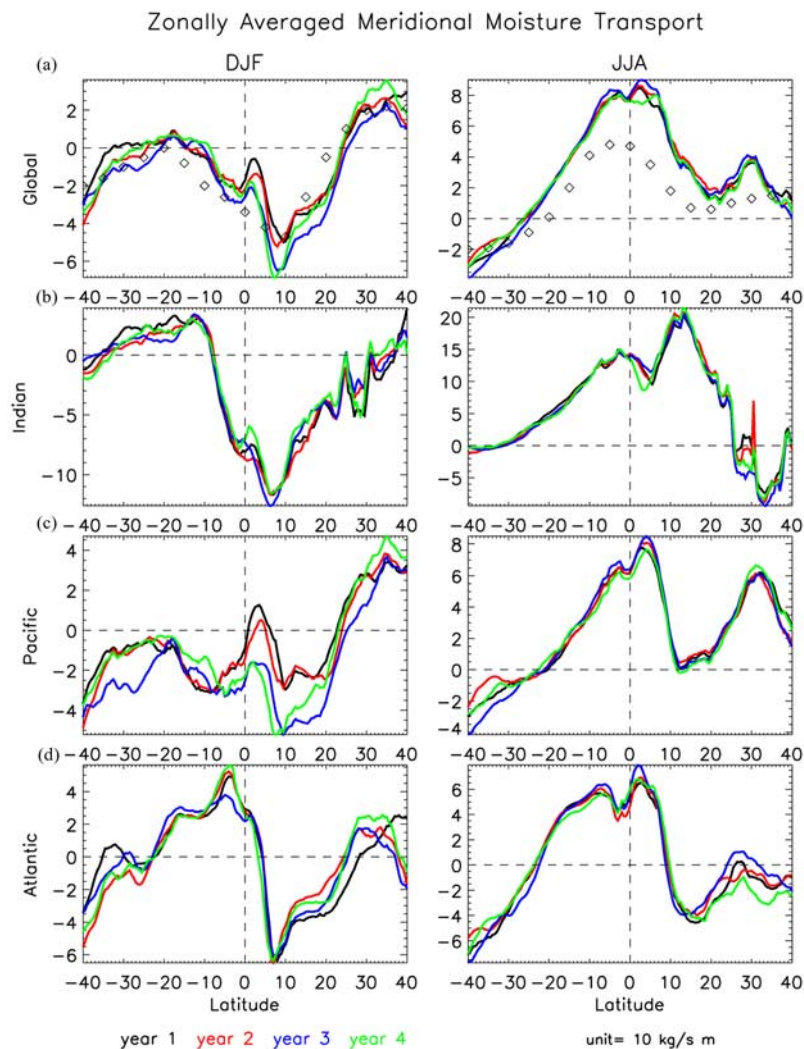


Figure 7. Meridional profiles for the zonally averaged meridional moisture transport over the (a) global, (b) Indian, (c) Pacific, and (d) Atlantic oceans, averaged for the months of (left) December, January, and February and (right) June, July, and August. Year 1 indicates the period from September 1999 to August 2000, year 2 indicates the period from September 2000 to August 2001, etc. Diamonds in Figure 7a represent 10-year climatological averages compiled by *Peixoto and Oort* [1992] using RAOB.

magnitude and direction for surface wind vector for the wind speed range of 3–20 m/s [*Naderi et al.*, 1991]; the standard deviations of the error of retrieved \mathbf{u}_e are only slightly higher. Comparisons with time series of \mathbf{u}_e at selected RAOB ocean stations show that the derived statistical relation reproduces major temporal variability. The limited validation effort in this study shows that, except for weak wind conditions, the derived transport field should be useful.

[19] At present, the model function is only developed for tropical and subtropical oceans within 40° of the equator. Effort is underway to extend it to higher latitudes. The computed Θ field captures the major large-scale features exhibited by the operational NWP products but with the advantage of improved resolution. An example of the large local difference between satellite- and NWP-derived Θ is also provided. Radiosonde data over oceans are too sparse to construct any meaningful large-scale spatial variability. The present evaluation is hampered by

lack of standards, and there will surely be continuous effort in improving and validating the model function. The best validation of space-based observations is, perhaps, through scientific applications, and the Θ fields are being applied to the studies of oceanic influence of continental hydrologic balance and to the studies of hydrologic balance of the upper ocean.

[20] Knowing the surface wind vector or the atmospheric circulation may not be sufficient to quantify Θ . The moisture transport is strongly influenced by but is not equivalent to atmospheric circulation. Because humidity is much higher at low altitude, Θ is strongly influenced by low-level winds. W generally decreases with increasing latitude, and Θ also emphasizes low-latitude flows. Water vapor is not a passive quantity advected by winds. Phase changes in atmospheric water vapor change the energy cycle and then the winds; Θ has clear feedback to atmospheric circulation. The method presented is a practical tool to estimate an important parameter at improved resolution and coverage.

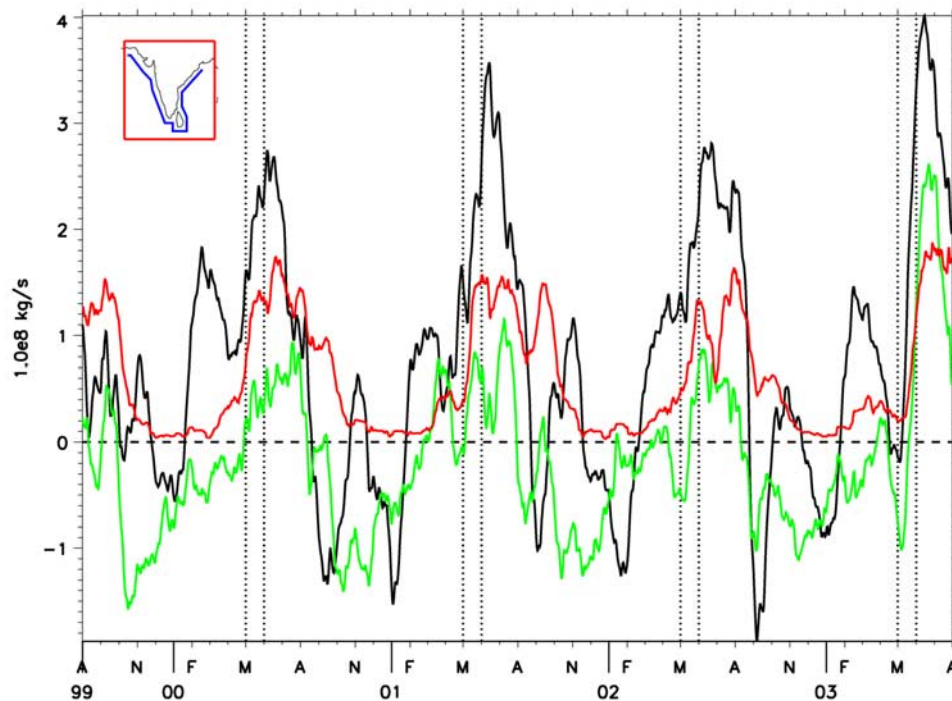


Figure 8. Time series of the moisture transport into the Indian subcontinent through simplified coastline, calculated from satellite data (black) and NCEP (green) and compared with total precipitation over land measured by TRMM (red).

[21] **Acknowledgments.** This study was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA. It was jointly supported by the Ocean Vector Wind, Precipitation Missions, and EOS/Aqua Programs of NASA. W. Tang was also supported by the Climate Data and Detection Program of NOAA.

References

- Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, 482 pp., Oxford Univ. Press, New York.
- Brown, R. A., and W. T. Liu (1982), An operational large-scale marine planetary boundary layer model, *J. Appl. Meteorol.*, *2*, 261–269.
- Bryan, F., and A. Oort (1984), Seasonal variation of the global water balance based on aerological data, *J. Geophys. Res.*, *89*, 11,717–11,730.
- Demuth, H., and M. Beale (1998), *Neural Network Toolbox, MATLAB user's guide, version 3*, MathWorks, Inc., Natick, Mass. (Available online at http://www-ccs.ucsd.edu/matlab/pdf_doc/nnet/nnet.pdf.)
- Dobson, F., L. Hasse, and R. Davis (Eds.) (1980), *Air-Sea Interaction: Instruments and Methods*, chap. 1–4, Springer, New York.
- Ebuchi, N. H., H. C. Graber, and M. J. Caruso (2002), Evaluation of wind vectors observed by QuikSCAT/SeaWinds using ocean buoy data, *J. Atmos. Oceanic Technol.*, *19*, 2049–2062.
- Heta, Y., and Y. Mitsuta (1993), An evaluation of evaporation over the tropical Pacific Ocean as observed from satellites, *J. Appl. Meteorol.*, *32*, 1242–1247.
- Kummerow, C. J., et al. (2000), The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit, *J. Appl. Meteorol.*, *39*, 1965–1982.
- Lagerloef, S. E. G. (2002), Introduction to the special section: The role of surface salinity on upper ocean dynamics, air-sea interaction and climate, *J. Geophys. Res.*, *107*(C12), 8000, doi:10.1029/2002JC001669.
- Liu, W. T. (1987), El Niño atlas, Nimbus-7 microwave radiometer data, *JPL Publ.*, 87-5, 68 pp.
- Liu, W. T. (1993), Water vapor and greenhouse warming, in *Remote Sensing of the Ocean*, edited by S. F. Jones, Y. Sugimori, and R. W. Stewart, pp. 214–219, Seibutsu Kenkyusha, Tokyo.
- Liu, W. T. (2002), Progress in scatterometer application, *J. Oceanogr.*, *58*, 121–136.
- Liu, W. T., W. Tang, and P. P. Niiler (1991), Humidity profiles over the ocean, *J. Clim.*, *4*, 1023–1034.
- Naderi, F. M., M. H. Freilich, and D. G. Long (1991), Spaceborne radar measurement of wind velocity over the ocean—An overview of the NSCAT scatterometer system, *Proc. IEEE*, *79*, 850–866.
- Peixoto, J. P., and A. H. Oort (1983), The atmospheric branch of the hydrological cycle and climate, in *Variations in the Global Water Budget*, edited by A. Street-Perrott, M. Beran, and R. Ratcliffe, pp. 5–65, Springer, New York.
- Peixoto, J. P., and A. H. Oort (1992), *Physics of Climate*, Am. Inst. of Phys., New York.
- Plant, W. J. (2000), Effects of wind variability on scatterometry at low wind speeds, *J. Geophys. Res.*, *105*, 16,899–16,910.
- Rasmusson, E. M. (1967), Atmospheric water vapor transport and the water balance of North America, *Mon. Weather Rev.*, *95*, 403–426.
- Rasmusson, E. M., and K. Mo (1996), Large-scale atmospheric moisture cycling as evaluated from NMC global analysis and forecast products, *J. Clim.*, *9*, 3276–3297.
- Rosen, R. D., D. A. Salstein, and J. P. Peixoto (1979), Variability in the annual fields of large-scale atmospheric water vapor transport, *Mon. Weather Rev.*, *107*, 26–37.
- Slonaker, R. L., and M. L. Van Woert (1999), Atmospheric moisture transport across the Southern Ocean via satellite observations, *J. Geophys. Res.*, *104*, 9229–9249.
- Starr, V. P., and J. P. Peixoto (1958), On the global balance of water vapor and the hydrology of deserts, *Tellus*, *10*, 188–194.
- Starr, V. P., and R. M. White (1955), Direct measurement of the hemispheric poleward flux of water vapor, *J. Mar. Res.*, *14*(3), 217–225.
- Trenberth, K. E., and C. J. Guillemot (1995), Evaluation of the global atmospheric moisture budget as seen from analyses, *J. Clim.*, *8*, 2255–2272.
- Wentz, F. J. (1997), A well-calibrated ocean algorithm for special sensor microwave/imager, *J. Geophys. Res.*, *102*, 8703–8718.

W. T. Liu and W. Tang, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA. (liu@pacific.jpl.nasa.gov)