

# Multimodel Analysis of Energy and Water Fluxes: Intercomparisons between Operational Analyses, a Land Surface Model, and Remote Sensing

RAGHUVVEER K. VINUKOLLU

*Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey, and Swiss Re, Armonk, New York*

JUSTIN SHEFFIELD AND ERIC F. WOOD

*Department of Civil and Environmental Engineering, Princeton University, Princeton, New Jersey*

MICHAEL G. BOSILOVICH

*NASA GSFC Global Modeling and Assimilation Office, Greenbelt, Maryland*

DAVID MOCKO

*NASA GSFC Global Modeling and Assimilation Office, Greenbelt, Maryland, and Science Applications International Corporation, McLean, Virginia*

(Manuscript received 13 September 2010, in final form 14 June 2011)

## ABSTRACT

Using data from seven global model operational analyses (OA), one land surface model, and various remote sensing retrievals, the energy and water fluxes over global land areas are intercompared for 2003/04. Remote sensing estimates of evapotranspiration (ET) are obtained from three process-based models that use input forcings from multisensor satellites. An ensemble mean (linear average) of the seven operational (mean-OA) models is used primarily to intercompare the fluxes with comparisons performed at both global and basin scales. At the global scale, it is found that all components of the energy budget represented by the ensemble mean of the OA models have a significant bias. Net radiation estimates had a positive bias (global mean) of  $234 \text{ MJ m}^{-2} \text{ yr}^{-1}$  ( $7.4 \text{ W m}^{-2}$ ) as compared to the remote sensing estimates, with the latent and sensible heat fluxes biased by  $470 \text{ MJ m}^{-2} \text{ yr}^{-1}$  ( $13.3 \text{ W m}^{-2}$ ) and  $-367 \text{ MJ m}^{-2} \text{ yr}^{-1}$  ( $11.7 \text{ W m}^{-2}$ ), respectively. The bias in the latent heat flux is affected by the bias in the net radiation, which is primarily due to the biases in the incoming shortwave and outgoing longwave radiation and to the nudging process of the operational models. The OA models also suffer from improper partitioning of the surface heat fluxes. Comparison of precipitation ( $P$ ) analyses from the various OA models, gauge analysis, and remote sensing retrievals showed better agreement than the energy fluxes. Basin-scale comparisons were consistent with the global-scale results, with the results for the Amazon in particular showing disparities between OA and remote sensing estimates of energy fluxes. The biases in the fluxes are attributable to a combination of errors in the forcing from the OA atmospheric models and the flux calculation methods in their land surface schemes. The atmospheric forcing errors are mainly attributable to high shortwave radiation likely due to the underestimation of clouds, but also precipitation errors, especially in water-limited regions.

## 1. Introduction

Fundamental to World Climate Research Programme's (WCRP) Global Energy and Water Cycle Experiment (GEWEX) is to use regional-to-continental datasets,

pertaining to water and energy cycles, to improve coupled model prediction through evaluation of models ranging from regional coupled models, numerical weather prediction (NWP) models, and climate models. Numerical weather prediction is based on initial atmospheric and land states from operational analyses (OA) in which forecast models assimilate large amounts of observational and satellite-derived information to provide the best possible initial condition for the forecast.

---

*Corresponding author address:* Eric Wood, CEE Department, EQUAD-Olden Street, Princeton University, Princeton, NJ 08544.  
E-mail: efwood@princeton.edu

Quantifying the errors in the energy and water budgets from OA is necessary for further improving the skill in predicting weather and climate from daily to seasonal time scales. In the past decade, the Coordinated Enhanced Observing Period (CEOP; Bosilovich and Lawford 2002; Koike 2004; Lawford et al. 2006), based on efforts initiated by the GEWEX Hydrometeorological Panel (GHP), has been seeking to archive data related to land–atmospheric processes from various data sources, including observations, remote sensing, and numerical model outputs, with the aim of understanding and predicting local- to continental-scale hydroclimates. Toward achieving this objective, CEOP has archived model output from eight numerical weather prediction and data assimilation centers (NWPCs) at the World Data Center for Climate (WDCC) managed by the Max Planck Institute for Meteorology (MPIM) and the German Climate Computing Center (DKRZ). These datasets were recently reprocessed to common data structures for 2003/04 [Multimodel Analysis for CEOP (MAC)] by Bosilovich et al. (2009). This provides a common framework that allows an evaluation of the performance of individual models with the expectation of model improvements based on our improved understanding of their physical and feedback processes.

Previously, many studies have focused on assessing the land and atmosphere water and energy budget components of OA and historical reanalyses at point-to-regional scales. Several studies have focused on evaluating model output location time series (MOLTS) against in situ data from flux towers and field experiments (Betts et al. 2006, 1998; Hirai et al. 2007; Rikus 2007; Yang et al. 2007). These studies focused on different surface energy and water budget variables but there was a common consensus that there are significant differences between the in situ data and the model output. Some of these differences include too-high values of nighttime surface latent and sensible heat flux (Betts et al. 2006); too-high evaporation estimates due to soil water nudging to remove low-level humidity errors (Betts et al. 1998); high biases in modeled annual precipitation values (Rikus 2007), which could further lead to excess evaporation; and evaporation exceeding precipitation, which could be associated with model spinup (Yang et al. 2007). Comparisons against in situ data can be problematic because small-scale spatial variability contributes to differences between patch-scale, in situ observations and coarse-resolution (tens to hundreds of kilometers) model output (e.g., Bosilovich and Lawford 2002). At regional scales, many studies have focused on assessing the water and energy budget components (Betts et al. 1999, 2003a,b, 2005, 2009; Fernandes et al.

2008; Karam and Bras 2008; Luo et al. 2007; Szeto 2007) mainly for large-scale river basins such as the Amazon, Mackenzie, and Mississippi. They reported that the evapotranspiration (ET) estimates from the analysis fields of the OA models were significantly higher than that from observation-forced land surface models and those inferred from observations of precipitation ( $P$ ) and streamflow ( $Q$ ). Evaluations of the water budgets of model reanalyses showed significant reduction in the bias in ET compared to OA models, but ET was sensitive to the errors in the precipitation, runoff, and/or soil moisture nudging terms (Fernandes et al. 2008; Su et al. 2006).

However, no studies exist that compare turbulent and/or evaporative fluxes from operational NWP models with observation-based datasets at continental-to-global scales. This can be partly associated with the lack of observational datasets at large scales, although this has been addressed to some extent in recent years by the increasing availability of satellite-remote-sensing-based datasets of ET and turbulent heat fluxes at global scales (e.g., Fisher et al. 2008; Mu et al. 2007; Vinukollu et al. 2011). The objective of the current study is to assess the accuracies of NWP analyses by intercomparing the various components of the energy and water budgets with the current best estimates that have been developed under the auspices of the National Aeronautics and Space Administration (NASA) Energy and Water Study (NEWS) and GEWEX. These estimates include remote-sensing-based data and output from observation-forced land surface modeling—the latter of which provides budget-constrained estimates. We evaluate seven operational models from forecast centers around the world over 2003/04 at basin-to-global scales. Section 2 provides detailed information regarding the OA models and the comparison datasets considered for this study. Some information on the data processing needed for comparisons is also presented in section 2. An extensive evaluation of the fluxes at basin and global scales is given in section 3. Finally, section 4 provides a summary of the current work and suggestions toward improving our representation and understanding of the energy and water cycles at large scales.

## 2. Data and methods

### *a. Operational analysis model output*

Data from eight international OA have been collected and reprocessed through the GEWEX CEOP Project for the period October 2002–December 2004 (Bosilovich et al. 2009). The complete list of contributing data centers and the corresponding references for

TABLE 1. The seven OA datasets, contributing data centers and the corresponding references used for the current study.

Dataset name	NWP center	Land scheme	Reference(s)
BMRC	Bureau of Meteorology Research Center, Australia	Bucket scheme, no vegetation types	Rikus (2007)
CPTEC	Centro de Previsão de Tempo e Estudos Climáticos (the Center for Weather Forecasts and Climate Studies), Brazil	Simplified Simple Biosphere (SSiB) scheme, 13 vegetation types	Chou et al. (2007)
ECPC-RII	Experimental Climate Prediction Center, United States	Oregon State University LSM scheme, 12 vegetation types (no monthly variation)	Ruane and Roads (2007)
ECPC-SFM	Experimental Climate Prediction Center, United States	Oregon State University LSM scheme, U.S. Geological Survey (USGS) monthly, 12 vegetation types	Ruane and Roads 2007
JMA	Japan Meteorological Agency	SSiB scheme, 13 vegetation types	Hirai et al. (2007)
MSC	Meteorological Services of Canada	Interface Soil–Biosphere–Atmosphere (ISBA) scheme, 22 vegetation types	Cote et al. (1998); Belair et al. (2005, 2008)
NCEP	National Centers for Environmental Prediction, United States	Noah scheme, 12 vegetation types	Ek et al. (2003)

the OA datasets are given in Table 1. Data for surface meteorology, radiation (energy), and water budget components have been combined and reprocessed at a 1.25° spatial resolution and at 6-hourly, daily, and monthly scales. For the current study, we use the daily averaged analyses from seven of the eight OA models. The Met Office (UKMO) model output has many missing days of data for the considered variables and thus is excluded from the current analyses. Complete details regarding the processing and the development of MAC can be found in Bosilovich et al. (2009).

### b. Remote sensing products

Remote sensing estimates of turbulent fluxes were derived by Vinukollu et al. (2011) using three process-based models: Surface Energy Balance System (SEBS; Su 2002), a Penman–Monteith formulation (PM; Monteith 1965) with revised stomatal resistance formulation (RSPM; Mu et al. 2007), and a Priestly–Taylor-based approach (PT; Fisher et al. 2008; Priestley and Taylor 1972). Radiation, meteorology, and vegetation inputs for the above process models are derived from data from radiometric sensors on board the NASA Earth Observing System (EOS) PM (*Aqua*), National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) polar orbiting satellites, and the Surface Radiation Budget, version 3.0 (SRB V3.0; Stackhouse et al. 2000) project. Vinukollu et al. (2011) compared the predicted surface sensible and latent heat fluxes to tower observations, which gave mean (temporal) correlations of 0.54 and 0.57 and a corresponding mean error (RMS difference) of 40 and 24 W m<sup>-2</sup>, respectively. The ET estimates were also evaluated at regional and continental scales against an inferred estimate from observed  $P$  and  $Q$ , with

correlations  $> 0.70$  across the models and error (RMS) in the range of 120–200 mm yr<sup>-1</sup>.

Precipitation estimates are obtained from four different sources: multisatellite-based precipitation estimates from the Global Precipitation Climatology Project (GPCP; Adler et al. 2003; Huffman et al. 2001), Tropical Rainfall Measuring Mission (TRMM) multisatellite precipitation product (TMPA; Huffman et al. 2007), and gauge-observation-based analyses from the University of East Anglia Climate Research Unit Time Series, version 3.0 (CRU TS3.0; Mitchell and Jones 2005) and Global Precipitation Climatology Center (GPCC; Rudolf and Schneider 2005). The GPCP precipitation product is based on existing multisatellite estimates of precipitation and bias-corrected rain gauge measurements provided by GPCC.

Currently, there exists no remotely sensed or observation-based gridded runoff product at continuous time scales over the land surface. To incorporate the runoff term for evaluation of the water budget components, we use large-basin streamflow observations and a climatological product, both available from the Global River Discharge Center (GRDC). The climatological product is derived by disaggregating streamflow observations to gridded runoff fields using the water balance model (WBM) of Fekete et al. (2000). Two separate climatological products were considered: one based on disaggregated in situ streamflow measurements over a set of selected basins and a composite global runoff field that combines the disaggregated streamflow observations with output from the WBM directly for regions without streamflow observations. The basin streamflow observations were only available for 2003/04 for six basins (out of nine analyzed) and so the climatological values were used otherwise.

### c. Land surface model (LSM) output

Land surface hydrological models are widely used to understand the land surface water and energy budgets at regional scales. Some of these models serve as the land scheme in general circulation models (GCMs). One such model is the Variable Infiltration Capacity (VIC) model (Cherkauer et al. 2003; Liang et al. 1996, 1994), which solves the water and energy balance at the land surface, thus acquiring closure for both. The model uses a variable infiltration capacity curve for each computational grid cell, which parameterizes the soil storage capacity as a probability distribution to which the partitioning of precipitation into infiltration and runoff is related.

For the current study, the VIC model estimates are taken from the updated global,  $1.0^\circ$  simulation of Sheffield and Wood (2007), which was forced by a hybrid observation–reanalysis meteorological dataset (Sheffield et al. 2006). The forcing dataset and the simulation have been extended to 1948–2006 using the latest versions of the CRU monthly temperature and precipitation datasets (TS3.0) and the SRB monthly downward short- and longwave radiation product (V3.0). The model parameters have also been updated via calibration based on spatially disaggregated streamflow observations (Sheffield et al. 2009a) using the sparse-grid calibration strategy of Troy et al. (2008). The VIC data are not computed over Antarctica.

We argue that a well-calibrated, observationally forced LSM will provide reasonable estimates of annual and seasonal water budgets and by inference, reasonable estimates of energy fluxes. Nevertheless, large errors do exist regionally because of poor observational data, insufficient calibration, and model structural error. On the other hand, the remote sensing retrievals can provide near-continuous spatial coverage and for regions where ground data are sparse, although the retrieval algorithms introduce errors in the same way as LSMs. Recent studies by Jiménez et al. (2011) and Mueller et al. (2011) showed large spread (up to  $25 \text{ W m}^{-2}$ ) in the annual latent heat flux estimates ( $LE$ ) between remote sensing retrievals, LSMs, and reanalyses. The remote sensing (RS) and calibrated LSM data therefore provide complementary information and help capture the uncertainties derived from the modeling approach and input data.

### d. Data processing

For the comparisons, all datasets were regridded (using a box-averaging method) to the resolution of the OA data ( $1.25^\circ$ ). The remote-sensing-based turbulent fluxes contain missing data owing to the presence of clouds that

obscure some of the retrievals. Comparisons involving the remote-sensing-based data must therefore consider the missing values. For this, we filter the OA-based fluxes using the available remote sensing fluxes at the daily scale. These daily values are further averaged and/or summed to monthly and annual scales for further comparisons.

## 3. Results

### a. Global comparisons

For the global comparisons, we start by evaluating the individual components of the energy balance, followed by comparisons of the individual components of the radiation budget at the land surface, meteorological variables, and precipitation. In all of the comparisons, the mean of the OA, referred to as mean-OA, and range of the OA models will be compared instead of the individual model data. However, any specific characteristic of a particular model will be discussed as needed. Considering the asymmetry of land and the strong seasonal differences between the Northern and Southern Hemispheres, we consider comparisons over latitudinal bands (averaged over longitude). Although the land-mass contributes to more than 40% in the Northern Hemisphere, only  $\sim 20\%$  of the Southern Hemisphere is covered by land. Furthermore, when comparing climatic variables, zonal mean values are of interest considering the relative uniformity in weather patterns in the east–west directions. While considering the energy components at the annual scale, we use cumulative flux units ( $\text{MJ m}^{-2} \text{ yr}^{-1}$ ) instead of the more usual convention of  $\text{W m}^{-2}$ . This is important for understanding the budget closures and the relative magnitudes of the individual terms (Fasullo and Trenberth 2008). For reference, a value of  $100 \text{ MJ m}^{-2} \text{ yr}^{-1}$  corresponds to approximately  $3.2 \text{ W m}^{-2}$  or  $40 \text{ mm yr}^{-1}$  in terms of water equivalent.

#### 1) NET RADIATION ( $R_{\text{net}}$ )

Figure 1a shows the seasonal and annual net radiation for the OA, RS, VIC model, and the SRB. The lines represent the meridional distribution of the mean zonal values of  $R_{\text{net}}$ , with the horizontal bars (in top-right corner plot) representing the fraction of land at each latitude. The gray shading shows the range of the OA models and their mean (mean-OA) is represented by the black line. For net radiation, the authors consider the SRB dataset (blue line) as the best observational estimate available based on Raschke et al. (2006). Note the distinct seasonal and annual cycle of  $R_{\text{net}}$  across the land surface. Although a general agreement among the datasets

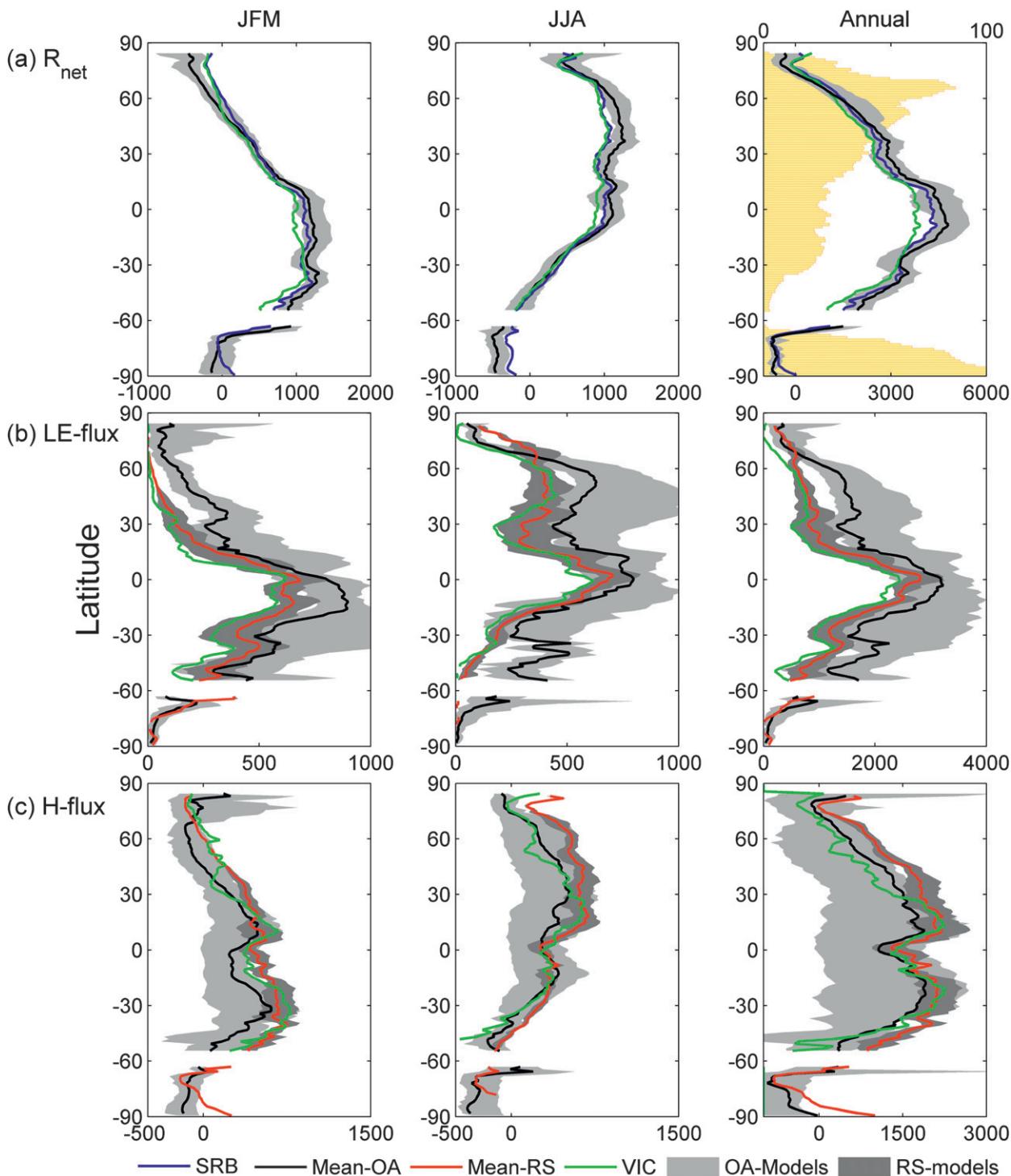


FIG. 1. (left) Winter, (middle) summer, and (right) annual (a) total net radiation, (b) latent heat flux, and (c) sensible heat flux as an average across the latitudinal bands and the years 2003 and 2004 ( $\text{MJ m}^{-2} \text{yr}^{-1}$ ). Mean-OA is the ensemble mean of the OA models. Also shown in right panel of (a) is the fraction of land per latitude band.

can be found in the seasonality for the two hemispheres, the mean-OA estimates tend to be higher than the SRB estimates on an annual basis. The mean-OA estimates have a global positive bias of  $\sim 234 \text{ MJ m}^{-2} \text{ yr}^{-1}$  ( $\sim 10\%$  of mean annual net radiation) with respect to SRB, with the Experimental Climate Prediction Center's seasonal forecast model (ECPC-SFM) OA model being an outlier. This result is consistent with the analysis by Ruane and Roads (2007), which showed that the diurnal net radiative fluxes from the ECPC-SFM model during July–September are significantly higher than observations over the continental United States (CONUS). The authors attributed the bias in the fluxes to problems in the cloud and albedo parameterizations of the model. The main outliers among the seven analyses and the regions of largest bias are the Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) (high latitudes;  $>50^\circ\text{N}$ ), ECPC Reanalysis II (ECPC-RII) (midlatitudes;  $25^\circ\text{--}35^\circ\text{N}$ ,  $30^\circ\text{--}50^\circ\text{S}$ ), and ECPC-SFM (tropics;  $15^\circ\text{S--}15^\circ\text{N}$ ). Details regarding the individual variables affecting the net radiation are discussed in section 3a(5). Within the spread of the OA models, it is observed that the Bureau of Meteorology Research Centre (BMRC) (in the tropics and extratropics) and the Meteorological Service of Canada (MSC) analyses closely reproduce the SRB estimates with an annual bias of 463 and  $-42 \text{ MJ m}^{-2} \text{ yr}^{-1}$ , respectively. The net radiation values reported are positive downward.

To understand the bias that exists in the net radiation, we looked at the latitudinal profiles (not shown) of the components of net radiation: downward shortwave radiation ( $\text{SW}_\downarrow$ ), upward shortwave radiation ( $\text{SW}_\uparrow$ ), downward longwave radiation ( $\text{LW}_\downarrow$ ), and upward longwave radiation ( $\text{LW}_\uparrow$ ). We found that the main source of bias in  $R_{\text{net}}$  is from the  $\text{SW}_\downarrow$ , which has a bias of  $740 \text{ MJ m}^{-2} \text{ yr}^{-1}$  as compared to the SRB estimates. These results are consistent with the findings by previous studies (Cess et al. 1995; Garratt and Prata 1996; Ramanathan et al. 1995; Wild and Roeckner 2006; Wild et al. 1995, 2001), where the authors indicate that GCMs (including those in assimilation mode; i.e., reanalysis models) overestimate  $\text{SW}_\downarrow$  by up to  $786 \text{ MJ m}^{-2} \text{ yr}^{-1}$  because of the underestimation of cloud shortwave absorption. Furthermore, the overestimation of net radiation is affected by the underestimation of  $\text{LW}_\uparrow$ , although this is cancelled out somewhat by the underestimation of the  $\text{LW}_\downarrow$ . The biases in the  $R_{\text{net}}$  components for the OA models compared to the RS (SRB) estimates are 740, 380,  $-387$ , and  $-127 \text{ MJ m}^{-2} \text{ yr}^{-1}$  for  $\text{SW}_\downarrow$ ,  $\text{SW}_\uparrow$ ,  $\text{LW}_\downarrow$ , and  $\text{LW}_\uparrow$ , respectively.

Estimates of  $R_{\text{net}}$  from the VIC model are also plotted in Fig. 1a, which is based on incoming short- and longwave radiation at the surface from the SRB dataset. As

expected, this closely follows the SRB data, except at high latitudes where it is biased slightly low, possibly because of differences in surface temperature or albedo estimates between VIC and SRB. It is also noted that the most differences in  $R_{\text{net}}$  radiation between the VIC and RS estimates are found in the tropics and are likely due to differences in surface albedo and emissivity. VIC uses surface emissivity of 1 for estimation of  $\text{LW}_\uparrow$ .

## 2) LATENT HEAT FLUX

Even though a general agreement in the seasonality for the Northern and Southern Hemispheres can be found for  $LE$  (Fig. 1b), a wide range of values (light gray shade) are seen across the various OA models. It is noted that the spread in  $LE$  is more than that in  $R_{\text{net}}$ . The three process-model-based estimates from remote sensing are plotted in darker gray shading, and their mean (mean-RS) is plotted in red.

Comparison of the mean annual  $LE$  estimates of mean-OA to those from RS and VIC shows high spatial correlation ( $\tau > 0.84$ , where  $\tau$  is Kendall's tau), but significant bias exists—a positive bias of 470 (543)  $\text{MJ m}^{-2} \text{ yr}^{-1}$  relative to the RS (VIC) estimates. This suggests a poor partitioning of the surface energy budget by most OA models. Visual inspection of the plots shows overestimation of  $LE$  by almost all of the analyses. However, the spread among the analyses is dominated at the higher end by the ECPC-RII analysis. Analysis by Ruane and Roads (2007), which compared the diurnal cycle of water and energy over CONUS using data from ECPC-RII, ECPC-SFM, and the North American Regional Reanalysis (NARR), showed that the  $LE$  values from the ECPC-RII analysis had a higher diurnal amplitude and mean than the observations over the southern Great Plains. They further concluded that the bias in the  $LE$  estimates from ECPC-RII were consistent across the CONUS. This outlier was also noted by Bosilovich et al. (2009). One factor affecting the ECPC-RII-based  $LE$  estimates could be their fixed vegetation cover (as compared to monthly varying vegetation), which further affects surface characteristics like surface albedo and roughness length. The BMRC analysis, which uses a fixed bucket-based interactive scheme to model soil moisture (Rikus 2007), matches very closely to the RS estimates in the tropics ( $30^\circ\text{S--}30^\circ\text{N}$ ). Analysis of the BMRC precipitation in the midlatitudes [section 3a(5)] shows that it is higher than the gauge-corrected, remote sensing estimates, which also may be the cause of the bias in the mid- and high latitudes. Global mean monthly  $LE$  time series (not shown) showed that all of the OA models overpredict  $LE$  globally.

### 3) SENSIBLE HEAT FLUX

Sensible heat flux ( $H$ ) is a major component of the energy cycle and is mainly dependent on the temperature gradient between the surface and the air above. It is one of the major contributors to the structure and diurnal cycle of the boundary layer. In OA models,  $H$  is coupled to the cloud-base height and the cooling processes (radiative and evaporation of rainfall) in the sub-cloud layer (Betts et al. 2005). Comparisons of the mean-OA net shortwave radiation and sensible heat flux (not shown) reveals a high positive correlation (global mean  $\tau = 0.85$ ) between the estimates. A similar result was found by Betts et al. (2005) using the 40-yr European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-40). This suggests that any bias in the net shortwave radiation will positively impact the  $H$  estimates.

Comparison of  $H$  estimates from the OA models, RS, and VIC is plotted in Fig. 1c. Although the OA models replicate the seasonal cycle, the seasonal peak in mean-OA has a time lag of about 1–2 months as compared to the RS estimates (not shown). The VIC  $H$  shows major differences (negative bias relative to the RS estimates) in the Northern Hemisphere summertime and a minimum in May (not shown), which can be attributed to the springtime snowmelt and precipitation processes combining to give higher  $LE$  (and thus lower  $H$ ) estimates. The global mean value of  $H$  from the OA models is  $880 \text{ MJ m}^{-2} \text{ yr}^{-1}$  (RS is  $1247 \text{ MJ m}^{-2} \text{ yr}^{-1}$  and VIC is  $720 \text{ MJ m}^{-2} \text{ yr}^{-1}$ ). As is consistent with the  $R_{\text{net}}$  and  $LE$  estimates, the  $H$  estimates from the ECPC-RII analysis form the significant outlier (bottom of the OA models spread) except at the high latitudes ( $>65^\circ\text{N}$ ) in the Northern Hemisphere. Although the positive bias in the  $LE$  estimates from the ECPC-RII cancel out most of the negative bias in the  $H$  estimates, the high latitudes suffer from a high energy balance closure problem [shown later in section 3a(4)] at the annual scale. Comparisons among the seven analyses show that the CPTEC analysis closely represents the RS estimates, while the MSC analyses are closer to the VIC estimates. Although the BMRC analysis falls within the spread of the RS for the  $R_{\text{net}}$  and  $LE$  estimates (in the tropics), it shows significant differences in  $H$  as compared to the RS and VIC estimates. This may be because the BMRC analysis does not consider any variation in land cover and that the surface roughness lengths over land are prescribed constant.

Global maps of the differences in the  $R_{\text{net}}$ ,  $LE$ , and  $H$  estimates between the OA models and remote sensing are shown in Figs. 2–4, respectively. The ECPC-RII analysis stands out in terms of overpredicting (underpredicting)

the latent (sensible) heat flux estimates. The largest differences in the surface fluxes are over central Africa (all models) and eastern side of North America (six of the seven OA models). Although the bias is reduced in most places with the multimodel mean, we find that the bias is reduced even more (not shown) by removing the ECPC-RII model.

### 4) SURFACE ENERGY RESIDUAL AND SOIL HEAT FLUX

Figure 5 shows the seasonal and annual energy residual over land (i.e., difference between the available net radiation and the surface latent and sensible heat fluxes). Although the residual should be equal to the soil heat flux ( $G$ ) (and snowpack-related heat transfer, which we assume can be neglected as it is less than 1% globally and 5% for high latitudes of mean annual net radiation, as estimated from the VIC data), it is important to note that most OA models do not close the energy budget (Yang et al. 2007) because of their assimilation schemes. The residual will be a combination of the soil heat flux and a nonclosure term. This nonclosure term is a result of the addition or subtraction of energy and water to the atmosphere during the assimilation of observations by the OA models and is related to the forecast error of the background model. Also, it is to be noted that only a few models estimate the soil heat flux explicitly and almost none report the nonclosure term. Within the framework of the current paper, we will report this term as an energy residual term unless otherwise noted.

Results show a distinct seasonal variation in the mean-OA estimates of the energy residual. At the annual scale, it is expected that the soil heat flux should be close to zero. It is found that the mean-OA estimates tend to be close to zero, suggesting that the models achieve energy balance closure reasonably well, except for the nonclosure term mentioned above. However, as pointed out earlier, the  $LE$  ( $H$ ) estimates from OA models are significantly higher (lower) than the remote sensing estimates, suggesting that the partitioning of  $R_{\text{net}}$  into surface fluxes in the OA models needs improvement. This will be discussed further in the basin-scale energy and water budget comparisons (section 3b). The VIC and remote sensing estimates used in this study have forced energy budget closure associated with the respective models. The mean global energy residual ( $R_{\text{net}} - LE - H$ ) or soil heat flux for the mean-OA, remote sensing, and VIC model are 87, 23, and  $307 \text{ MJ m}^{-2} \text{ yr}^{-1}$ , respectively. For VIC, the high annual soil heat flux values are likely related to the use of a fixed temperature bottom soil layer boundary condition over high latitudes. It is important to note that

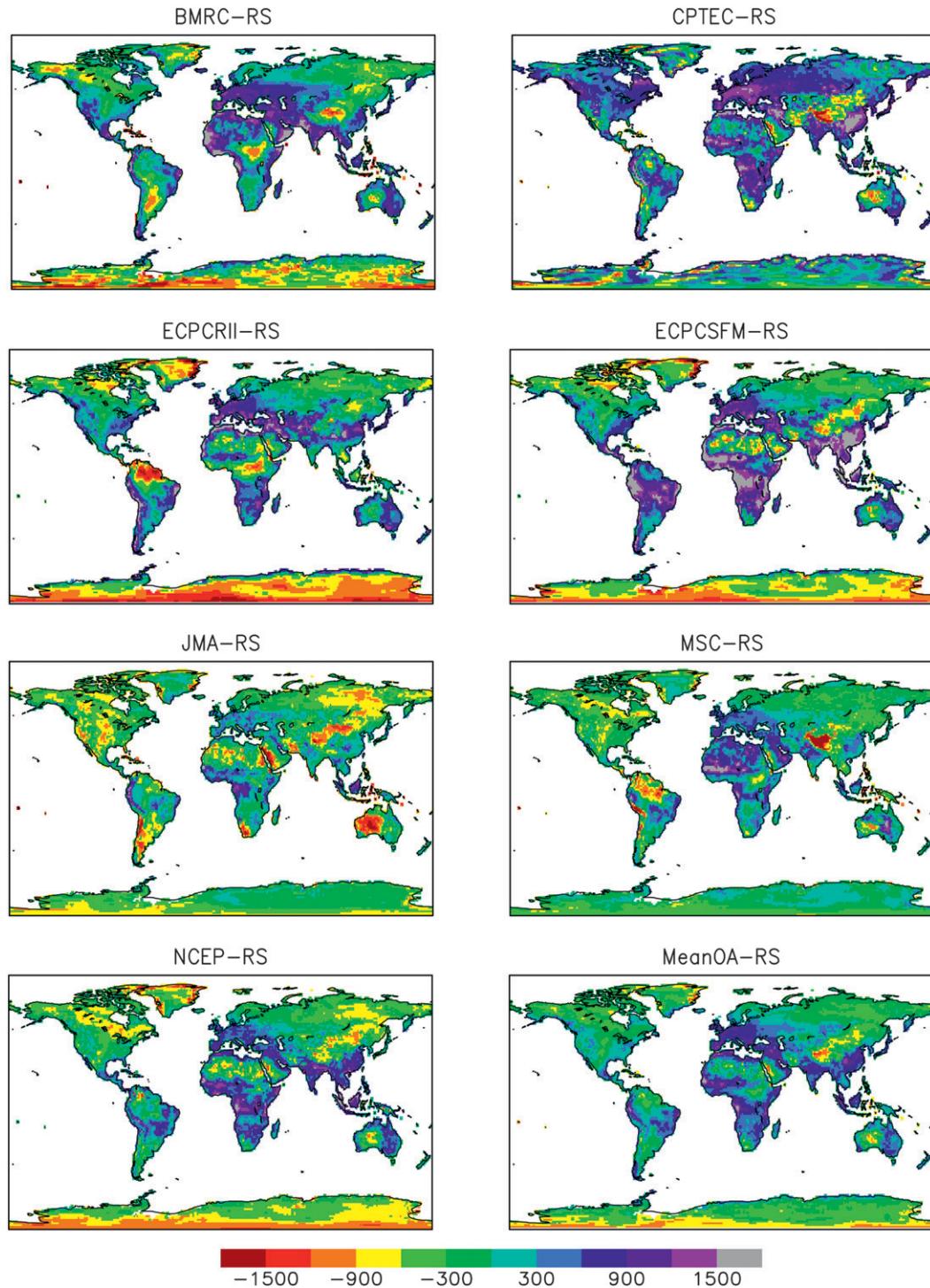


FIG. 2. Annual differences in net radiation ( $\text{MJ m}^{-2} \text{yr}^{-1}$ ) between the seven operational analysis models and the remote-sensing-based SRB estimates.

the remote sensing estimates of soil heat flux do not consider the effect of snowpack over the soil surface. The spikes in the soil heat flux over the high latitudes by the OA models can be associated with regions with

thin snowpack, where the OA models estimate unrealistic snow and upper-soil-layer temperatures, leading to the spikes in the modeled  $G$  values (Hinkelman et al. 1999). Note also the outlier of the VIC energy

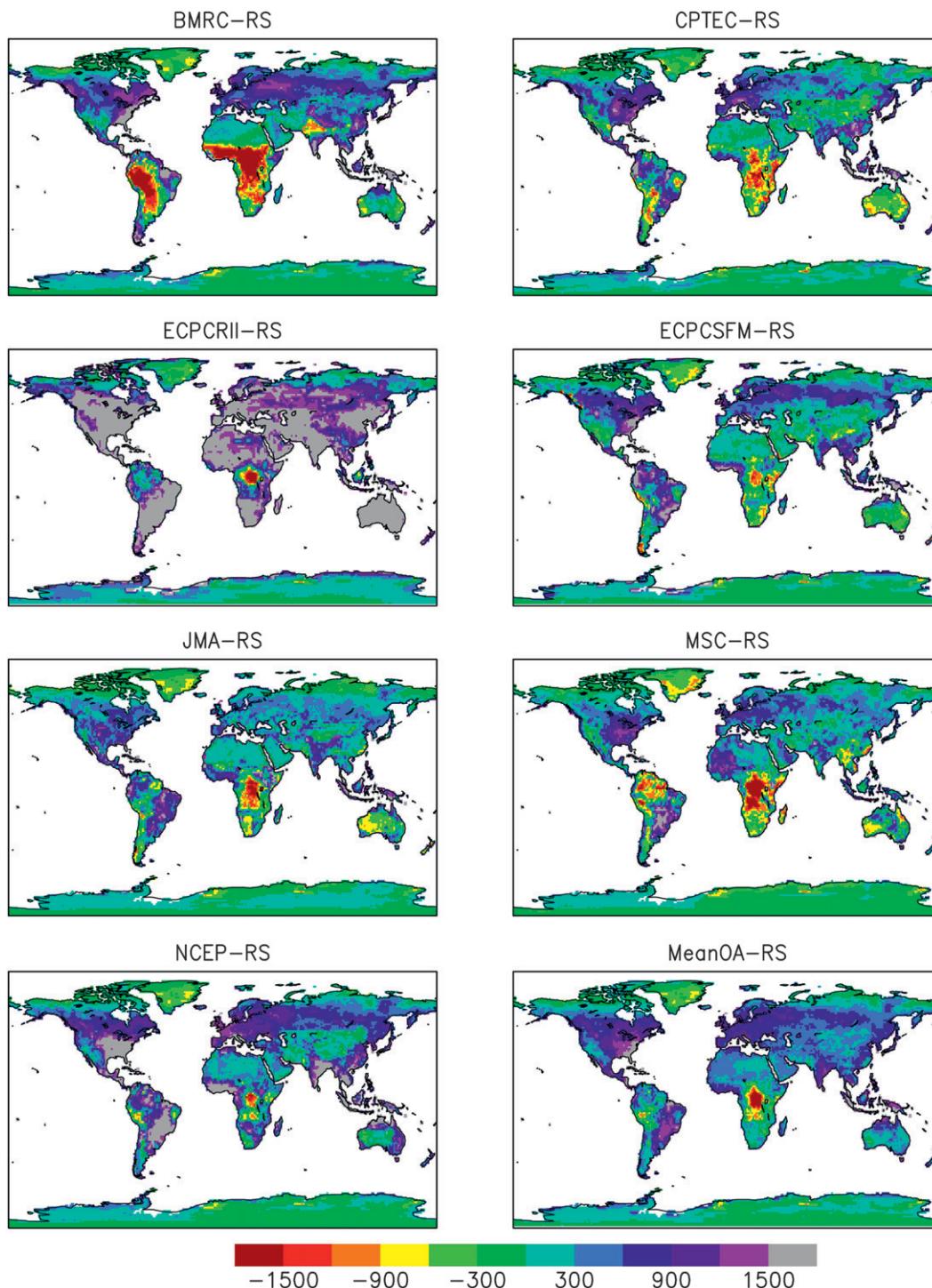


FIG. 3. Annual differences in the latent heat flux ( $\text{MJ m}^{-2} \text{yr}^{-1}$ ) between the various operational models and the mean of remote-sensing (SEBS/PM/PT)-based flux estimates.

residual, and thus the soil heat flux, around  $55^{\circ}\text{S}$  latitude. This value is associated with low sensible heat flux estimates over a few grid points (possibly over glaciers) in Chile.

#### 5) PRECIPITATION AND $P$ - $ET$ ANOMALIES

Precipitation estimates (Fig. 6a) from remote sensing (GPCP; gauge corrected to the GPCC) and two gauge

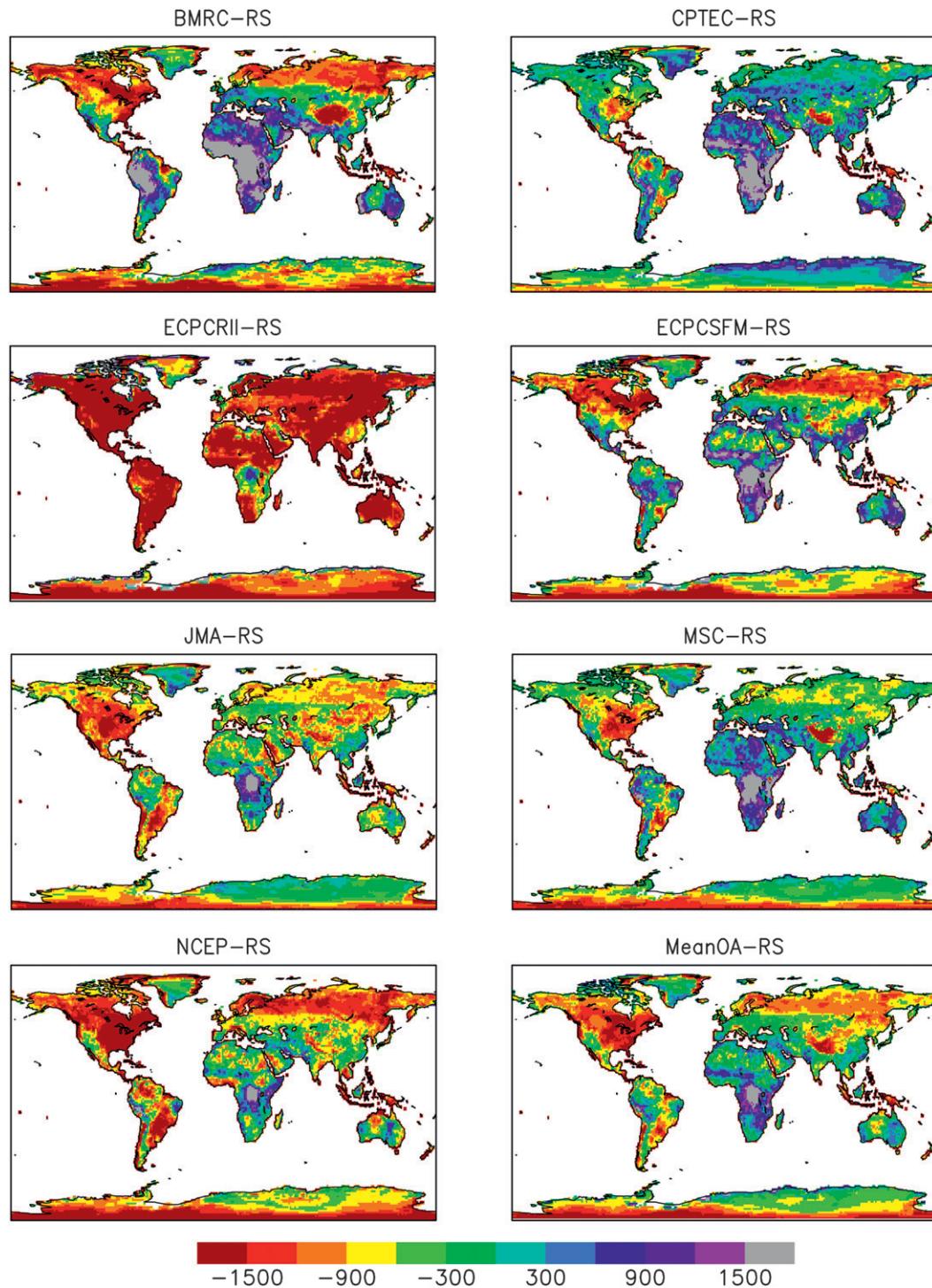


FIG. 4. Same as Fig. 3 but for sensible heat flux.

products (CRU and GPCC) were considered. Note that the VIC simulations use precipitation based on the CRU data product. The mean-OA estimates of precipitation agree well with the gauge-observed and remote sensing estimates except for over the tropics ( $20^{\circ}\text{S}$ – $20^{\circ}\text{N}$ ), where

the OA models are higher. However, it is difficult to assess the confidence level of observational datasets considering that uncertainties also exist in the gauge precipitation estimates. Rudolf and Schneider (2005) point out the two major sources of error in the GPCC

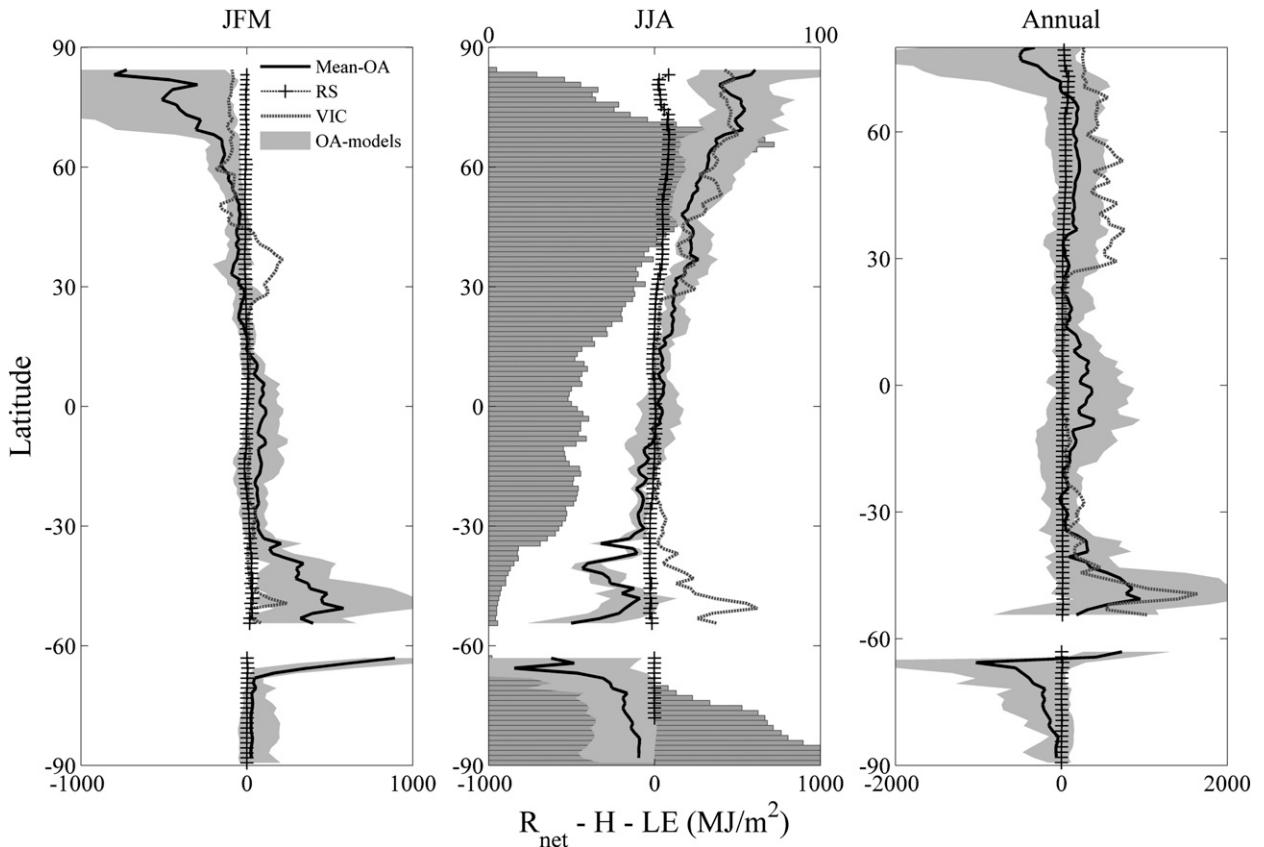


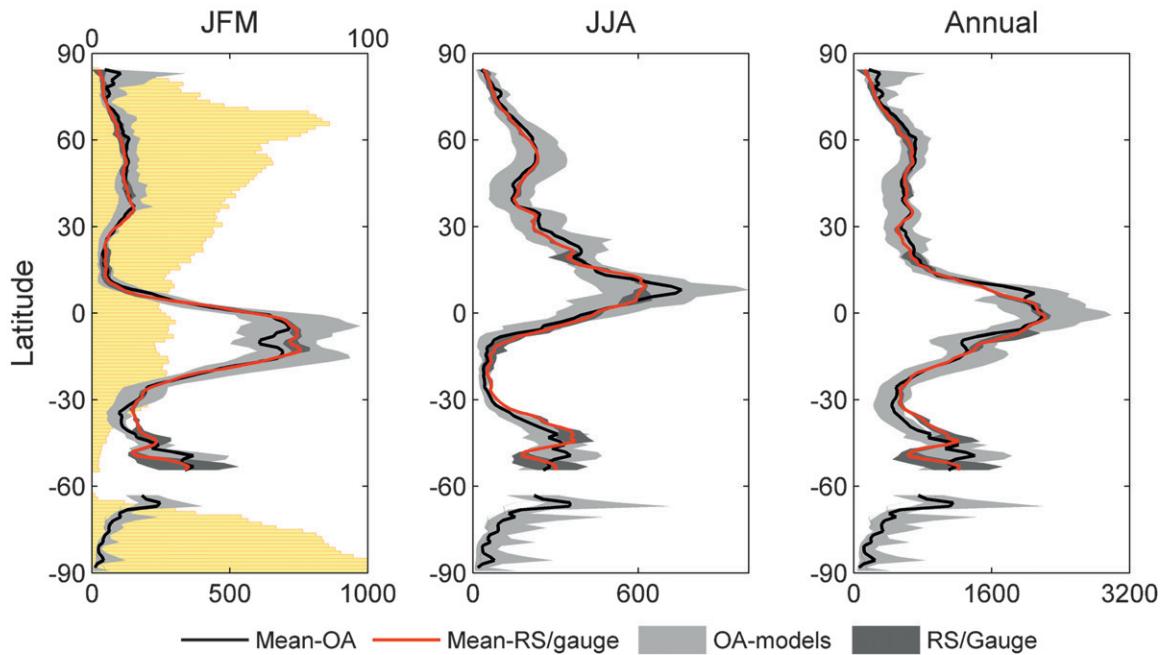
FIG. 5. (left),(middle) Seasonal and (right) annual total residual of the energy fluxes (i.e.,  $R_{\text{net}} - LE - H$ ) as obtained using operational analysis, land surface model (VIC), and remote sensing. Data plotted as in Fig. 1. The fraction of land per latitude band is also shown in the middle panel.

dataset, which also applies to the CRU data: 1) gauge undercatch error caused by evaporation out of the gauge and aerodynamic effects and 2) stochastic sampling error over regions with a sparse gauge network. They quote that the sampling errors, based on the number of stations within a grid cell, could range from  $\pm 7\%$  to as high as 40%. The GPCP dataset uses about four times as many gauges per unit area as the CRU dataset, although the CRU gauge densities are generally of the order of at least 25 gauges per  $10^6 \text{ km}^2$ , which is about the threshold at which spatial sampling errors become stable (Oki et al. 1999). The uncertainties at large scales are therefore small and this is borne out by the low spread in the estimates in Fig. 6. Given this, the mean-OA precipitation estimates in the tropics could be well within the range of errors of the observed precipitation estimates.

To evaluate the  $P$  and  $ET$  estimates from each of the individual OA models, we calculate the mean global precipitation, evapotranspiration, and their ratio ( $ET/P$ ) and compare them with the remote sensing estimates (Table 2) and the values provided by Trenberth et al. (2009). Note that the estimates from Trenberth et al.

(2009) are for 2000–04. The precipitation and especially evapotranspiration estimates from all the OA models are overpredicted with the mean bias of  $0.14 \text{ mm day}^{-1}$  for  $P$  and  $0.55 \text{ mm day}^{-1}$  for  $ET$  (based on the mean estimates of three remote sensing products). The significant bias in the evapotranspiration estimates from the OA models leads to high values of  $ET/P$ , which have a mean bias of 0.30 compared to the mean-RS (mean of CRU, GPCC, and GPCP) estimates. Note that two of the seven OA models (ECPC-RII and ECPC-SFM) estimate  $ET/P$  ratios of greater than and close to one over the entire land surface, which contradicts the conservation of mass that  $ET < P$  over land. Approximately 35% of the rainfall over land is attributable to marine evaporation driven by winds while the remaining 65% comes from evaporation over land (Chahine 1992). The problem of  $ET > P$  in the models could be related to model spinup such that excess moisture is available for evaporation above what would normally be available from precipitation (Yang et al. 2007). As was discussed in the previous sections, the ECPC-RII and the National Centers for Environmental Prediction

## (a) Precipitation [mm/year]



## (b) P-E [mm/year]

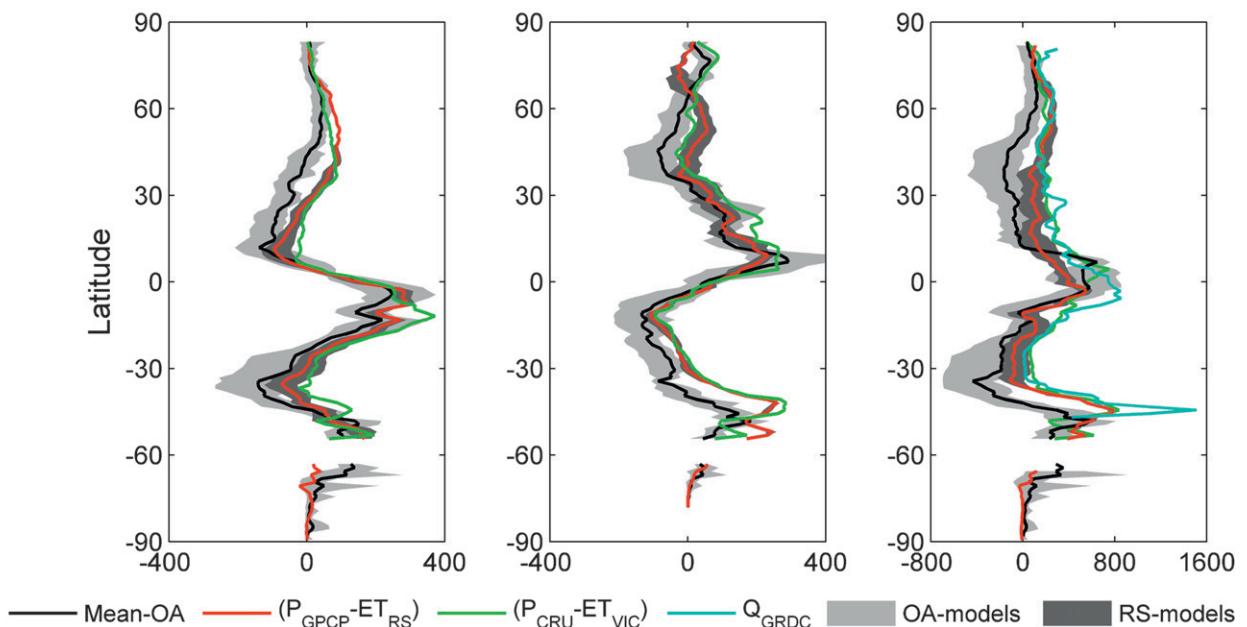


FIG. 6. (left),(middle) Seasonal and (right) annual (a) total  $P$  and (b)  $P - ET$  anomalies ( $\text{mm yr}^{-1}$ ) plotted as an average across the latitudinal bands. Precipitation data from OA models, remote sensing (GPCP/TRMM), and gauge observations (CRU/GPCC) are plotted. A climatological runoff ( $Q$ ) product from GRDC is also plotted to compare against the  $(P - ET)$  anomalies. The fraction of land per latitude band is shown in the left panel of (a).

TABLE 2. Global averages of annual mean ET and  $P$  ( $\text{mm day}^{-1}$ ), and the ratio ( $ET/P$ ) as obtained using the various datasets. Note that the VIC model uses the CRU precipitation dataset.

Dataset	$P$	ET	ET/ $P$
BMRC	2.26	1.60	0.71
CPTEC	2.28	1.65	0.72
ECPC-RII	2.60	2.81	1.08
ECPC-SFM	1.82	1.75	0.96
JMA	2.28	1.70	0.74
MSC	2.19	1.56	0.71
NCEP	2.61	2.08	0.80
Mean-OA	2.29	1.88	0.82
$P_{\text{GPCP}}/E_{\text{SEBS}}$	2.43	1.55	0.64
$P_{\text{GPCP}}/E_{\text{PM}}$		1.07	0.44
$P_{\text{GPCP}}/E_{\text{PT}}$		1.37	0.56
Mean RS	2.43	1.33	0.55
CRU/VIC	2.50	1.12	0.45
TFK2009	2.43	1.36	0.56

(NCEP) models clearly overestimate evaporation. Although the ECPC-RII model does not use a soil moisture nudging process, the models assimilate precipitation (pentad scale) observations to compute soil moisture. This assimilation process, as pointed out by Kanamitsu et al. (2002a,b) and Maurer et al. (2001), introduces a soil moisture correction (addition or subtraction at each time step), which is analogous to the nudging process. Note that the VIC simulation does not include Antarctica. Excluding Antarctica, global mean values of precipitation from CRU, GPCC, and GPCP are 911, 912, and 870  $\text{mm yr}^{-1}$ , respectively. Although CRU and GPCC (gauge based) estimates have errors associated with gauge undercatch (Rudolf and Schneider 2005), these products also show high uncertainty (positive bias) in regions with no gauges, more specifically in the tropics. This may be one reason for the differences between the remote sensing (GPCP) and the gauge-based (CRU and GPCC) products. Evaporation estimates from VIC are biased low in comparison to the mean of the remote sensing estimates; however, in many regions we found (not shown) that the PM-based remote sensing estimates were closer to VIC, which also uses Penman–Monteith as the basis for estimating evaporation.

Figure 6b shows the  $P - ET$  distribution, averaged across the longitudes, for January–March (JFM), June–August (JJA), and the annual cycle. Note the seasonal shift around the equator in the peak positive and negative values, which is related to the monsoons. On an annual basis, it is expected that the mean global value is greater than zero (i.e., runoff to the oceans). However, as noted in Table 2, two of the seven OA models show evaporation over land greater (or very close to) precipitation, suggesting an influence of model spinup or soil moisture nudging. Furthermore, when considering the

terrestrial water balance, annual  $P - ET$  should balance runoff over the long term (Oki et al. 1995; Seneviratne et al. 2004; Yeh et al. 1998). Thus, we also plot in Fig. 6b the long-term climatological runoff from the composite observation-model GRDC data (section 2b). Although the errors associated with the GRDC runoff dataset is unknown, it is likely that the errors are greatest in ungauged regions (e.g., central Africa) and for headwaters of large basins, where the influence of the observation data from the outlet gauging station is lowest and short-term errors are more influential. However, the meridional trend correlates ( $\tau = 0.52$ ) with the  $P - ET$  estimates from remote sensing. The VIC estimates are calibrated to GRDC runoff for a number of large basins globally and thus correlate well with the GRDC composite data ( $\tau = 0.72$ ). It is important to note that for a short time period (two years in this case: 2003/04) the assumption of  $P - ET = Q$  is not totally valid. Overall, the OA models reproduce the latitudinal profile of the estimates from RS, VIC, and GRDC, but there is tendency to underestimate  $P - ET$  in midlatitudes.

Figure 7 shows global maps of  $P - ET$  from the seven OA models, mean-OA, VIC, and mean-RS. The ECPC-RII model stands out with negative  $P - ET$  values for most of the midlatitudes. Although the mean-OA estimate captures the major climatic zones, it is biased low, especially in the Southern and Northern Hemisphere subtropics to midlatitude dry regions (see maps in Fig. 7), because of excessive evaporation relative to precipitation. One notable difference is the east–west gradient of the continental United States, which is only subtly represented by most of the OA models.

## 6) SUMMARY OF GLOBAL WATER AND ENERGY BUDGETS

To summarize the surface energy and water budget components over the land surface, we tabulate (Table 3) the global mean annual estimates of all the components from the various models. We also include the flux estimates of Trenberth et al. (2009), referred to hereafter as TFK2009, for the period March 2000–May 2004. The TFK2009 estimates are based on various satellite and reanalyses products and so are also prone to error, but provide another estimate to help quantify the uncertainties. Considering the energy balance, we find that TFK2009 estimates of net radiation are lower than the SRB estimates, however the difference ( $120 \text{ MJ m}^{-2} \text{ yr}^{-1}$ ;  $<1\%$  of the annual net radiation) is well within the range of errors ( $314\text{--}472 \text{ MJ m}^{-2} \text{ yr}^{-1}$ ; Zhang et al. 2007) expected in a net radiation dataset, which is linked to errors in input datasets like air temperature and specific humidity. With that being said, the energy residual ( $R_{\text{net}} - LE - H$ ; considering a negligible annual soil heat

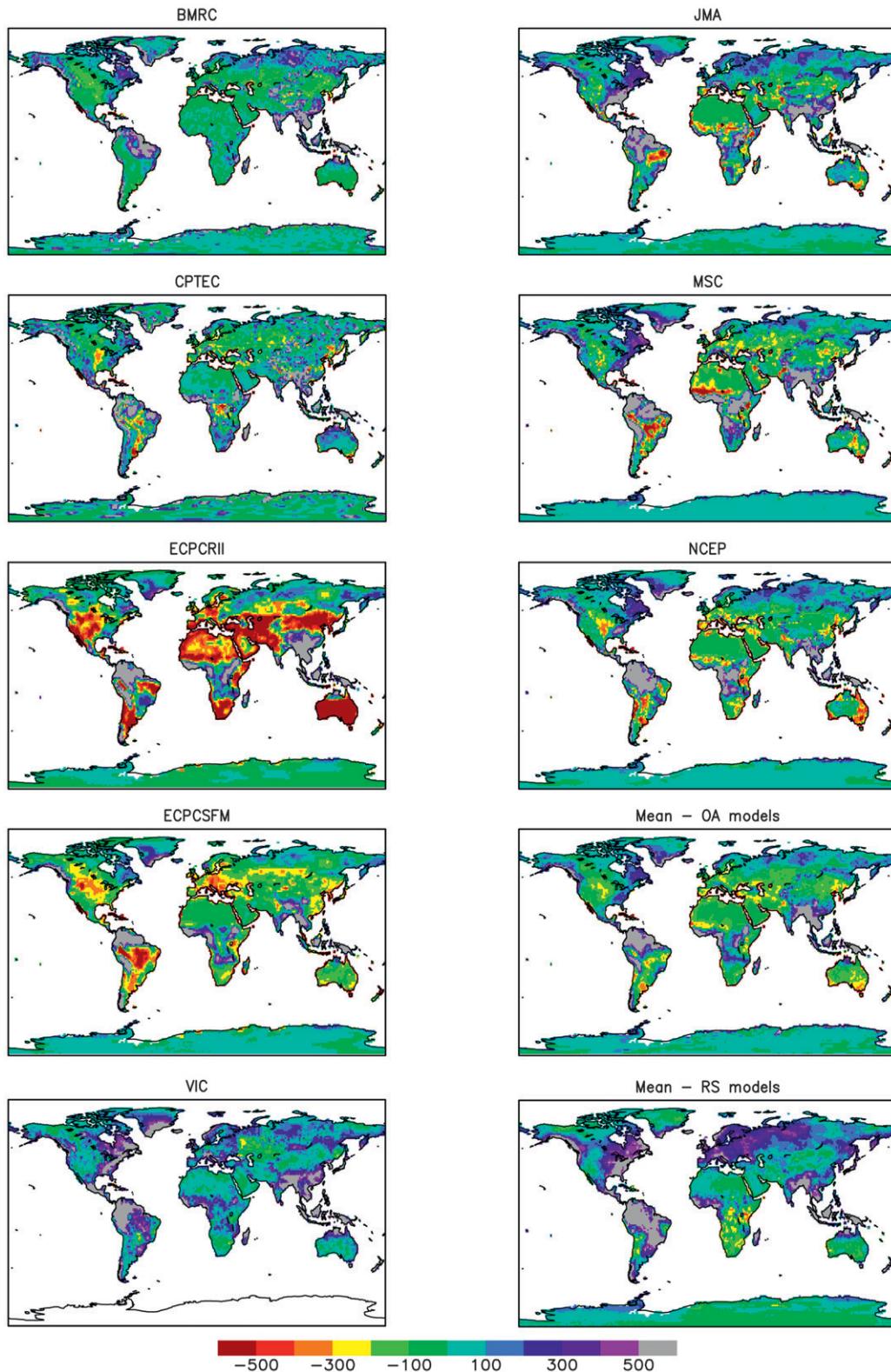


FIG. 7. Annual (mean over 2003/04) anomalies of precipitation and evaporation (i.e.,  $P - ET$ ) for each of the OA models, their mean, mean of remote sensing estimates, and the VIC land surface model output ( $\text{mm yr}^{-1}$ ).

TABLE 3. Comparison of annual estimates of energy and water fluxes over land. VIC estimates do not cover Antarctica.

Global land	SW <sub>↓</sub>	Albedo	SW <sub>↑</sub>	LW <sub>↓</sub>	LW	R <sub>net</sub>	LE flux	H flux	P	ET	Q
	MJ m <sup>-2</sup> yr <sup>-1</sup>		—	MJ m <sup>-2</sup> yr <sup>-1</sup>						mm yr <sup>-1</sup>	
SEBS RS							1387	792		566	
PM RS	5749	0.23	1339	9835	12 034	2211	954	1225	887	389	329
PT RS							1227	951		501	
Mean RS	5749	0.23	1339	9835	12 034	2211	1189	989	887	485	329
BMRC	6492	0.26	1700	9614	11 947	2459	1429	1228	824	583	—
CPTEC	6425	0.24	1559	9337	11 379	2824	1480	1413	833	604	281
ECPC-RII	6362	0.27	1712	9596	11 847	2398	2513	-182	948	1025	113
ECPC-SFM	7000	0.27	1890	9287	11 857	2542	1566	1116	665	639	22
JMA	6644	0.27	1796	9005	11 793	2060	1519	693	833	620	—
MSC	6288	0.24	1503	9501	11 743	2543	1397	1146	798	570	62
NCEP	6333	0.30	1874	9590	11 760	2288	1862	532	951	760	354
Mean-OA	6506	0.26	1719	9419	11 761	2445	1681	849	836	686	166
VIC	—	—	—	—	—	2527	1002	1232	911	409	405
TFK2009	5825	0.21	1223	9574	12 085	2091	1214	851	887	495	286

flux) as estimated by mean-OA, RS, and TFK2009 are -85, 33, and 26 MJ m<sup>-2</sup> yr<sup>-1</sup>. As mentioned above, the OA models overestimate  $R_{net}$  and  $LE$ , which tend to cancel each other and give the low estimate of the energy residual. It is found that the remote sensing estimates used in this study (Vinukollu et al. 2011) are close to the energy budget estimates of TFK2009.

For the water budget ( $P$ ,  $ET$ , and  $Q$ ), Table 3 indicates that the  $P$  and  $ET$  estimates from remote sensing match well with the TFK2009 estimates. As is expected, the  $Q$  estimates diverge, considering that the  $Q_{RS}$  is obtained from a climatological product that considers a different time period (1901–2000). Of the five OA models [excludes BMRC and Japan Meteorological Agency (JMA)] that provide runoff estimates, we find that CPTEC and best represent the observational estimates provided by TFK2009 based on discharge-to-ocean estimates.

### b. Basin comparisons

Nine major basins (see Table 4 and Fig. 8) are used to intercompare the datasets at regional scales. The basins were selected to cover a wide range of hydroclimates across the six inhabited continents. We distinguish the basins based on climatic zones and also classify them by

the primary limitation (energy/water or demand/supply; Table 4). The major focus of this section is to understand the partitioning of precipitation into the three water budget components: river discharge, evapotranspiration, and change in total water storage. Similarly, we also look at the partitioning of net radiation into the respective heat flux components (i.e., soil heat, sensible heat, and latent heat fluxes). We assume that total water storage changes and soil heat are near zero. In the current study, we use the runoff observations from GRDC where available. For basins where data is unavailable, the GRDC climatological data are used.

Figure 9 shows the partitioning of annual precipitation into  $ET$  and  $Q$  for the nine basins. Diagonal lines represent mean annual  $P$  (Fig. 9) and are shown for the mean-OA, CRU, and GPCP. If the point representing  $ET$  versus  $Q$  coincides with the  $P$  line, the assumption of no change in storage is valid, but only if the dataset closes the water budget (which is the case for the VIC model). Thus, offsets of  $ET$  and  $Q$  from the  $P$  line can be attributed to either nonclosure of the water balance, a change in storage, or both. Because of the assimilation of observations, budget closure in the OA models is not guaranteed. For budget analyses based on independent

TABLE 4. Nine global basins considered in this study, their gauge location, dominant climate, upstream area, and primary limitation on ET.

River basin	Gauge location	Climate	Upstream area (km <sup>2</sup> )	Primary limitation
Amazon	Obidos, Brazil	Tropical	4 618 746	Energy
Amur	Komsomolsk, Russia	Arctic	~1 730 000	Energy
Ganges	Bahadurabad, Bangladesh	Midlatitude rainy	~1 000 000	Energy
Mekong	Pakse, Laos	Tropical	~545 000	Energy
Mississippi	Vicksburg, United States	Midlatitude rainy	2 964 254	Energy
Murray–Darling	Australia	Semiarid	~1 000 000	Water
Niger	Lokoja, Niger	Semiarid (north) tropical savannah (south)	2 209 300	Water
Ob	Salekhard, Russia	Arctic	~2 430 000	Energy
Parana	Corrientes, Argentina	Midlatitude rainy	~2 300 000	Water

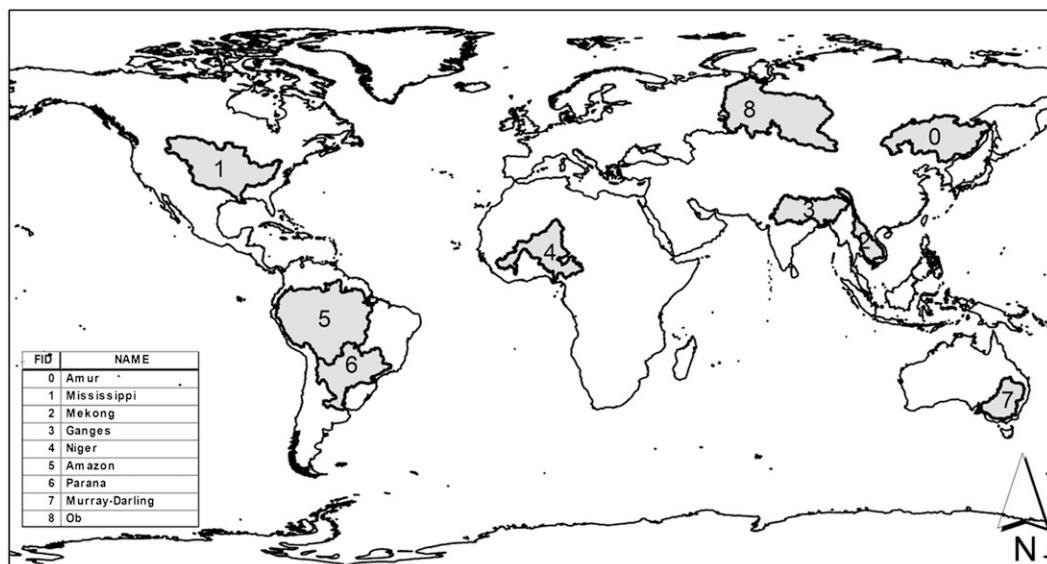


FIG. 8. Spatial location of the nine global basins considered for the current study. Basins were selected so as to represent most climatic zones across the earth's surface.

remote sensing products, uncertainty in the retrieved products for each budget component results in non-closure of the water budget (Gao et al. 2010; Sheffield et al. 2009b).

Similar to the water budget partitioning, we evaluate the energy balance over the basins using the net radiation and the surface turbulent heat fluxes. Figure 10 shows the partitioning of the net radiation into the sensible and latent heat fluxes. Although the partitioning should include the available energy ( $R_{\text{net}} - G$ ), we assume that the soil heat flux term is small compared to the other components of the energy budget and negligible at annual time scales. Also, since most OA models do not estimate  $G$ , we do not include the  $G$  term to be consistent. The diagonal lines represent net radiation and the individual symbols indicate the partitioning between sensible and latent heat fluxes. Symbols below the constant  $R_{\text{net}}$  line suggest positive soil heat storage (and/or an energy imbalance in the model/dataset).

The energy and water budgets are summarized for all basins in Figs. 9 and 10 and in Table 5, but we focus the discussion below on two basins, namely Amazon and Mississippi River basins. These two basins were selected because of availability of observational data, their large sizes, and differing climates. Note that the remote sensing estimates force closure of their energy balances, but not for the water budget.

### 1) AMAZON RIVER BASIN

Over the Amazon River basin (ARB), all three precipitation datasets are in close agreement for annual

precipitation. The mean daily rainfall over the Amazon from the three data sources is 4.91 (mean-OA), 5.60 (RS), and 5.65 (VIC/CRU)  $\text{mm day}^{-1}$ . Costa and Foley (1997) and Marengo (2005) reported an all-Amazonia rainfall estimate from rain gauge observations as approximately  $5.8 \text{ mm day}^{-1}$  for different periods between 1920 and 1992. They also pointed out that although the remote sensing gauge-corrected GPCP estimates agreed qualitatively, there were differences of up to  $0.6 \text{ mm day}^{-1}$  (mean annual precipitation) as compared to the rain-gauge-based observed products. Considering the interannual differences and that the remote sensing and gauge observations show minor differences, it can be concluded that the RS and VIC estimates are well within the error estimates of precipitation over the ARB, while the OA models clearly underestimate their precipitation analysis. Correlations (Kendall's tau; not shown) between the precipitation estimates were greater than 0.7; however, differences between the mean-OA and RS estimates were observed to be as high as  $2.13 \text{ mm day}^{-1}$  during the rainy season.

Net radiation estimates over the ARB show significant differences, which is consistent with the global estimates as reported above. Net radiation from the mean of the OA models is overestimated with amplitudes as high as  $136 \text{ MJ m}^{-2} \text{ month}^{-1}$ , which is a direct effect of the excess  $\text{SW}_{\downarrow}$  radiation estimated by the OA models. Furthermore, we find (not shown) that most of the bias is centered on July–April, which includes the period of highest cloud cover. As reported in previous studies (Cess et al. 1995; Garratt and Prata 1996; Ramanathan

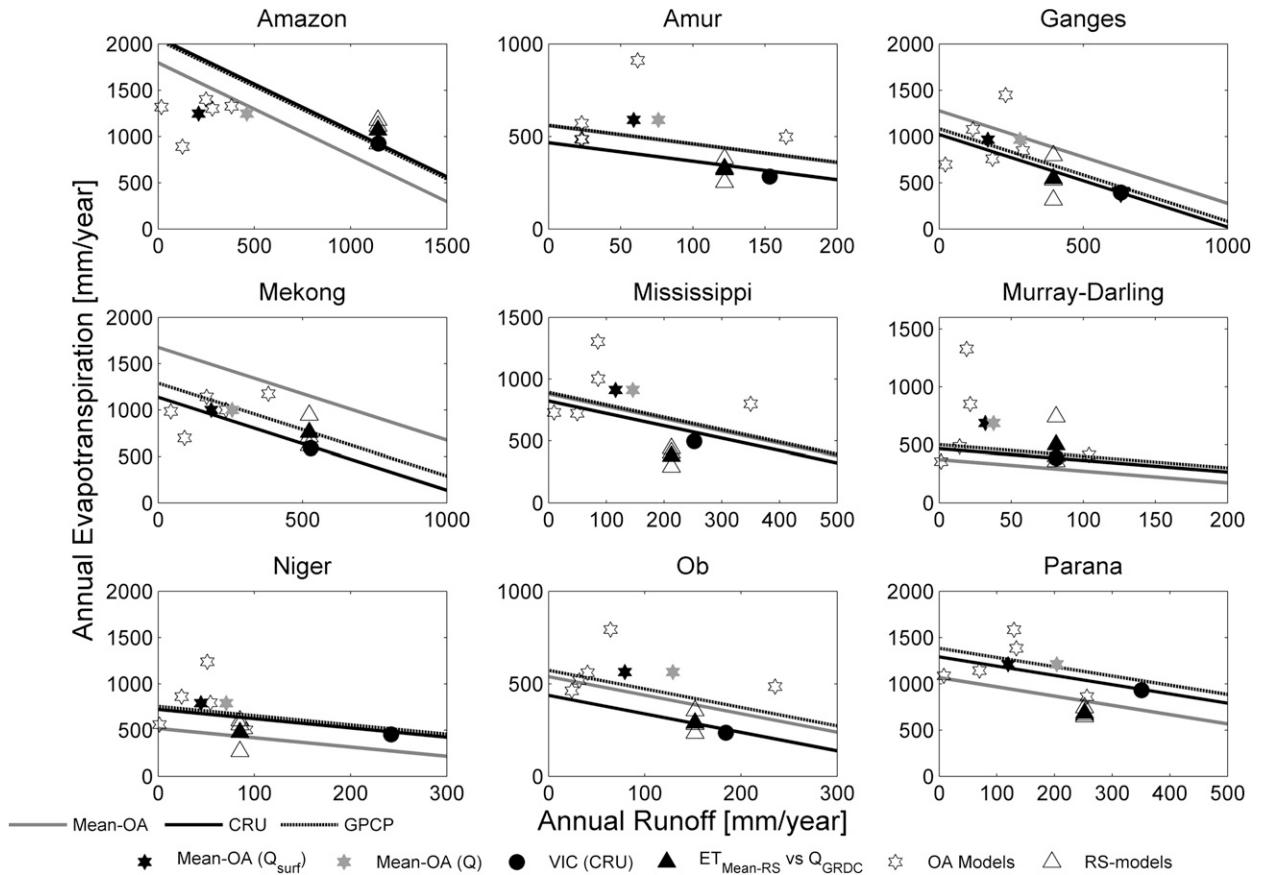


FIG. 9. Partitioning of  $P$  from OA models, VIC, remote sensing (GPCP), and gauge observations (CRU) in the nine basins considered for the current study. Diagonal lines represent constant precipitation. Note that if the model symbol lies above the precipitation line, then the model represents a negative storage change;  $Q_{GRDC}$  represents the GRDC streamflow observations for 2003/04 where available and an observed climatological runoff product when not. Gray shaded marker of the OA models represents the mean-OA estimate when the baseflow component is included.

et al. 1995; Wild and Roeckner 2006; Wild et al. 1995, 2001), most GCMs overestimate surface shortwave radiation because of the underestimation of cloud shortwave absorption. Although the RS and VIC estimates have similar amplitudes ( $\sim 86 \text{ MJ m}^{-2} \text{ month}^{-1}$ ), VIC estimates are consistently biased low, which is a result of the low surface temperature (not shown), and thus higher  $LW_{\uparrow}$  radiation.

Based on various observational-based studies (Callede et al. 2002; Costa and Foley 1997; Ramillien et al. 2006) over the Amazon, the mean annual basin-scale ET estimates were in the range of  $3.3\text{--}3.7 \text{ mm day}^{-1}$ , which corresponds to a mean of  $\sim 1250 \text{ mm yr}^{-1}$ . Although this value is derived from data for different time periods, it is considerably higher than the remote sensing (mean RS =  $1070 \text{ mm yr}^{-1}$ ) and VIC ( $924 \text{ mm yr}^{-1}$ ) estimates, but well represented by the mean of the OA models ( $1246 \text{ mm yr}^{-1}$ ). Considering a mean value of  $3.5 \text{ mm day}^{-1}$  as representative over the Amazon River

basin based on previous studies, it could be concluded that the OA models (with the exception of the MSC analysis:  $890 \text{ mm yr}^{-1}$ ) are in good agreement with the in situ observations. One possible explanation that can be associated with the underestimation of ET by RS is that the leaf area index (LAI) data that are used in the process models are saturated, thus reducing the canopy conductance values over forest land cover, and do not include the understory LAI. Furthermore, both VIC and RS have less radiation available for evapotranspiration than the OA estimates. Canopy interception is also a large component of total ET in the Amazon (Miralles et al. 2010), which may be underestimated by VIC and the RS data. Unfortunately, data for canopy interception was not available for the OA models.

The OA models only report a surface runoff component (not subsurface baseflow) and so a direct comparison with observed streamflow data is not possible. Therefore, we estimate the baseflow component for the

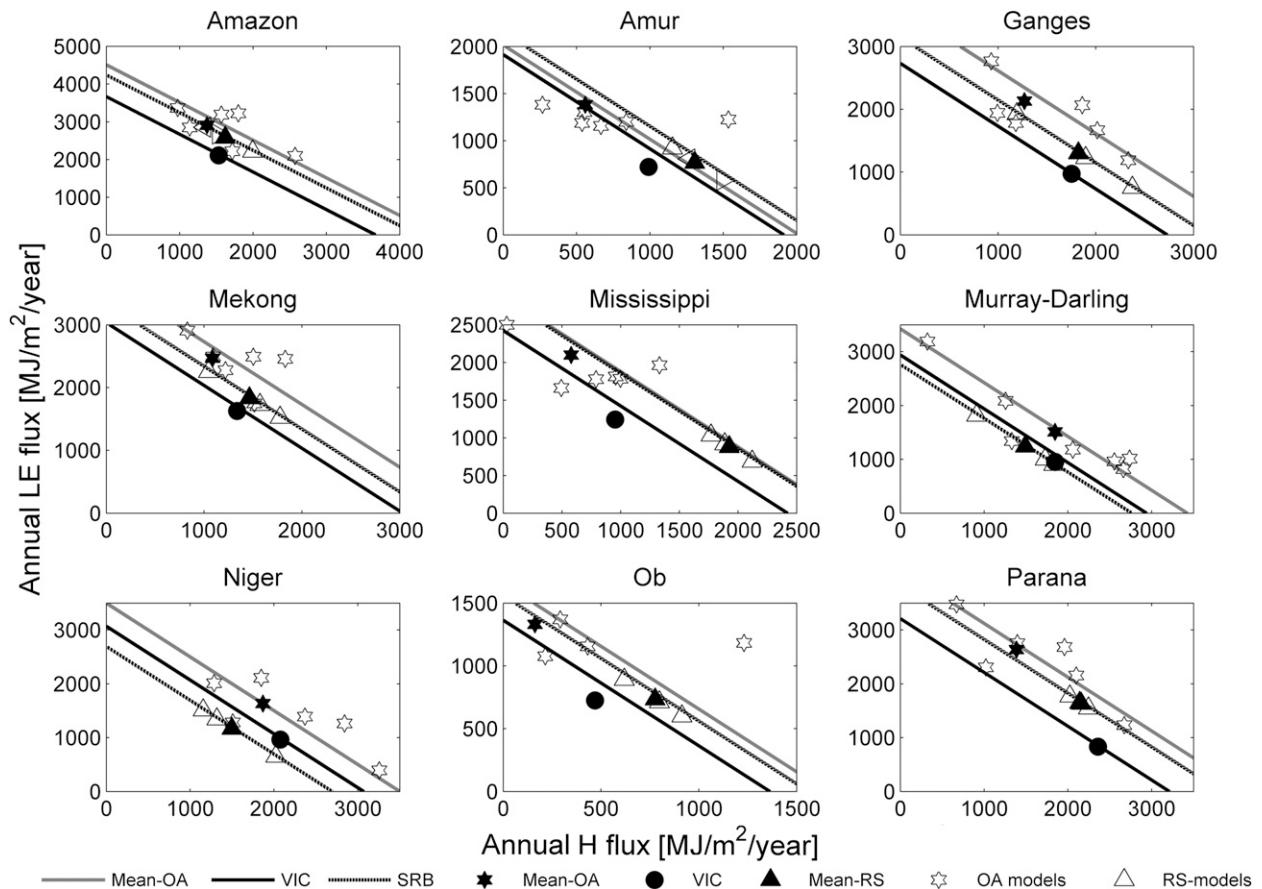


FIG. 10. Same as Fig. 9, but for  $R_{\text{net}}$ . Symbols that fall below the line show positive  $G$  values. Note that remote sensing and land surface model data show forced closure of energy balance.

OA models using the baseflow-to-surface-flow ratio from the VIC simulation. Runoff observations over the ARB, based on measured discharges for the Amazon, Xingu, and Tocantins Rivers, has been estimated as  $2.9 \text{ mm day}^{-1}$  (Marengo 2005). This value is close to the runoff estimates of  $3.13 \text{ mm day}^{-1}$  and  $2.7 \text{ mm day}^{-1}$  obtained from the GRDC runoff observations and the VIC model output, respectively. The surface runoff component of total runoff from the VIC model is  $1.43 \text{ mm day}^{-1}$ , indicating that the surface runoff from all the OA models (mean-OA =  $0.6 \text{ mm day}^{-1}$ ) is underestimated.

With the above results in mind, we further look at the energy balance partitioning. Figure 10 confirms the differences in the net radiation, with the mean-OA estimate higher than the remote sensing and VIC estimates. However, as pointed out in the global analysis, there seems to be a reasonable closure for the OA models. The mean-OA combination of sensible and latent heat flux lies very close to (below) the net radiation line, with the difference ( $231 \text{ MJ m}^{-2} \text{ yr}^{-1}$ ) associated

with the annual soil heat flux and energy balance error. This suggests that the energy budget from the OA models has 1) a bias in the net radiation flux and 2) improper partitioning of the surface fluxes. The improper partitioning was further confirmed by comparison of Bowen ratio values over the nine basins: the mean-OA is biased low by  $>0.5$  over seven of the nine basins, including the ARB. A summary of the statistics of the mean-OA against the other datasets is provided in Table 5. The correlations ( $\tau$ ) of the various components among the nine basins reveal that the Amazon River basin has the lowest values, except for the precipitation estimates. The above analysis suggests that there is large uncertainty in our understanding of the hydrology of the ARB.

## 2) MISSISSIPPI RIVER BASIN

The Mississippi River basin (MRB) is the region in which GEWEX began performing energy and water budget studies in the early 1990s. Considering the extensive gauge network over the basin, precipitation

TABLE 5. Statistics of the mean-OA  $P$ , ET,  $Q$ , and  $R_{\text{net}}$  compared to various observational datasets over the nine basins listed in Table 4. Considering that no gridded runoff product is available over the 2003/04 period, surface runoff is evaluated against a simulation of the VIC land surface model, which is calibrated to observed streamflow data. The third header row shows the comparison datasets for each variable. Kendall's tau correlations are calculated using monthly data.

Basin	RMSE								
	$P$ (mm day <sup>-1</sup> )	ET (mm day <sup>-1</sup> )			$Q$ (mm day <sup>-1</sup> )	$R_{\text{net}}$ (MJ m <sup>-2</sup> day <sup>-1</sup> )	$H$ (MJ m <sup>-2</sup> day <sup>-1</sup> )		
	vs GPCP	vs SEBS	vs PM	vs PT	vs VIC	vs SRB	vs SEBS	vs PM	vs PT
Amazon	0.43	0.29	0.30	0.25	0.17	0.47	0.70	1.02	1.03
Amur	0.22	0.27	0.47	0.34	0.12	0.90	0.86	0.88	1.39
Ganges	0.92	0.26	0.43	0.79	0.10	0.56	0.94	0.92	1.45
Mekong	0.52	0.23	0.34	0.76	0.14	0.42	0.62	0.81	1.10
Mississippi	0.21	0.43	0.48	0.39	0.11	0.72	0.72	0.70	0.85
Murray–Darling	0.19	0.59	0.56	0.37	0.04	0.99	1.34	0.96	2.24
Niger	0.21	0.19	0.21	0.29	0.02	0.41	0.79	0.62	0.82
Ob	0.20	0.35	0.43	0.32	0.19	0.68	0.95	1.19	1.65
Parana	0.40	0.38	0.40	0.59	0.03	0.66	0.86	0.90	1.07
Average	0.37	0.33	0.40	0.46	0.10	0.65	0.86	0.89	1.29
Basin	Bias								
	$P$ (mm day <sup>-1</sup> )	ET (mm day <sup>-1</sup> )			$Q$ (mm day <sup>-1</sup> )	$R_{\text{net}}$ (MJ m <sup>-2</sup> day <sup>-1</sup> )	$H$ (MJ m <sup>-2</sup> day <sup>-1</sup> )		
	vs GPCP	vs SEBS	vs PM	vs PT	vs VIC	vs SRB	vs SEBS	vs PM	vs PT
Amazon	-0.68	-0.10	0.08	0.60	-2.55	0.74	0.49	0.05	-1.23
Amur	-0.01	0.66	0.86	0.51	-0.26	-0.73	-2.33	-2.83	-1.97
Ganges	0.53	0.21	1.52	0.94	-1.26	1.41	0.43	-2.81	-1.36
Mekong	1.06	0.11	1.01	0.76	-0.94	0.95	0.33	-1.89	-1.26
Mississippi	-0.03	1.22	1.54	1.13	-0.37	0.12	-3.29	-4.06	-3.06
Murray–Darling	-0.35	0.73	0.60	-0.32	-0.13	1.48	-0.40	-0.10	2.19
Niger	-0.65	0.17	1.07	0.29	-0.54	2.31	1.98	-0.25	1.69
Ob	-0.09	0.73	0.86	0.53	-0.29	0.23	-1.85	-2.17	-1.35
Parana	-0.87	1.11	1.05	0.84	-0.63	0.96	-1.92	-1.79	-1.26
Average	-0.12	0.54	0.95	0.58	-0.77	0.83	-0.73	-1.76	-0.85
Basin	Kendall's $\tau$								
	$P$	ET			$Q$	$R_{\text{net}}$	$H$		
	vs GPCP	vs SEBS	vs PM	vs PT	vs VIC	vs SRB	vs SEBS	vs PM	vs PT
Amazon	0.85	-0.14	0.01	-0.34	0.28	0.77	0.57	-0.17	0.05
Amur	0.91	0.85	0.81	0.88	0.43	0.93	0.80	0.78	0.61
Ganges	0.88	0.86	0.64	0.07	0.62	0.87	0.66	0.73	0.45
Mekong	0.91	0.81	0.70	0.09	0.72	0.89	0.56	0.38	0.06
Mississippi	0.78	0.80	0.81	0.80	0.43	0.93	0.80	0.80	0.75
Murray–Darling	0.83	0.39	0.40	0.64	0.46	0.86	0.75	0.83	0.48
Niger	0.93	0.82	0.74	0.69	0.89	0.79	0.41	0.65	0.46
Ob	0.70	0.81	0.78	0.79	0.45	0.91	0.74	0.67	0.59
Parana	0.88	0.77	0.67	0.59	0.93	0.86	0.62	0.57	0.54
Average	0.85	0.66	0.62	0.47	0.58	0.87	0.66	0.58	0.44

estimates are less prone to errors than for the Amazon. Precipitation estimates from the three analyses (mean-OA, RS, and VIC–CRU) agree well over the MRB (Fig. 9) with values of 880, 892, and 822 mm yr<sup>-1</sup> from mean-OA, RS, and VIC/CRU, respectively. We consider that, over the MRB, the best estimates are provided by the North American Land Data Assimilation System (NLDAS; Mitchell et al. 2004). Over the MRB, the NLDAS reports values of approximately 814 mm yr<sup>-1</sup>, which is less than 0.20 mm day<sup>-1</sup> difference compared to the mean-OA

estimates. Errors in the GPCP estimates could be related to the gauge undercatch and orographic effects associated with the GPCC estimates (Adam and Lettenmaier 2003; Adam et al. 2006; Huffman et al. 2001; Rudolf and Schneider 2005). The mean-OA estimates of precipitation shows a difference of approximately +66 mm yr<sup>-1</sup> (0.18 mm day<sup>-1</sup>); the offset mainly due to the ECPC–R11 (948 mm yr<sup>-1</sup>) and NCEP (956 mm yr<sup>-1</sup>) analyses.

Discharge observations are available from the Vicksburg gauge managed by the U.S. Army Corps

of Engineers (USACE). Converting these discharge observations to runoff gives an annual estimate of  $213 \text{ mm yr}^{-1}$ . The Mississippi is highly managed, but out-of-basin extractions are rather small, with most dams in the upper reaches (primarily the Missouri River) for flood control and energy, while along the main stem the dams are mostly low head and related to navigation. Thus, dam operation and extractions have minimal impact on the annual observed flows. The GRDC observation-based climatological estimates of runoff ( $203 \text{ mm yr}^{-1}$ ) agrees well with the observations considering the interannual differences that could be expected over the different time periods of the datasets. Although the VIC-estimated runoff is based on a calibrated simulation, annual runoff ( $\sim 252 \text{ mm yr}^{-1}$ ) is higher than observed runoff. For the OA models, we apply the ratio of surface to total runoff as estimated from the VIC data to the mean-OA surface runoff estimate. The estimated total runoff for the ensemble mean (mean-OA) is  $146 \text{ mm yr}^{-1}$ , which is an underestimate as compared to the GRDC and VIC estimates.

One of the best available estimates of evapotranspiration over the MRB is available from the VIC NLDAS dataset of Troy et al. (2008), which is calibrated to streamflow observations from a large set of unmanaged basins across the CONUS. Based on their simulations, the annual ET over the MRB for the period 2003/04 is  $540 \text{ mm yr}^{-1}$ . This is higher than the VIC estimates considered in this study (VIC GLOBAL;  $497 \text{ mm yr}^{-1}$ ) by Sheffield and Wood (2007). However, as noted earlier, the precipitation and other input forcings for the two datasets differ (VIC GLOBAL uses CRU  $P$  and VIC NLDAS uses NDLAS  $P$ ), and also since the calibrated runoff from the VIC GLOBAL simulations is higher than the observed, the impact is seen as an underestimation of ET. It is also observed that all the remote sensing retrievals underestimate ET over the MRB. Two reasons can be associated with this underestimation: 1) saturated values of remote-sensing-based LAI leading to low conductance values and thus lower evaporation, and 2) low bias in canopy evaporation, which is a result of the low bias in the remote-sensing-based precipitation (and the parameterization of canopy evaporation).

ET estimates for the OA models show a wider spread over the MRB as compared to the ARB, with a range of  $721\text{--}1304 \text{ mm yr}^{-1}$  and a model mean of  $910 \text{ mm yr}^{-1}$ . The two outliers of the OA models that offset the mean are ECPC-RII ( $1304 \text{ mm yr}^{-1}$ ) and NCEP ( $1001 \text{ mm yr}^{-1}$ ). Removing the above two analyses reduces the mean (mean-OA) to approximately  $750 \text{ mm yr}^{-1}$ . This estimate is still higher than the VIC NLDAS estimates, and as pointed out before for the ARB, the reason for this overestimation is twofold: 1) the precipitation

estimates are higher than the NLDAS, which are considered as the best available estimates; and 2) the soil moisture nudging process used in the operational models nudges the models toward their climatology. The bias (compared to RS estimates) in the precipitation (radiation) estimates is approximately  $211 (20) \text{ mm yr}^{-1}$ , while the ET estimates are biased high by  $532 \text{ mm yr}^{-1}$ . Clearly the nudging process in the operational models has a huge impact on the water budget closure. However, the largest impacts are seen in the ECPC-RII and NCEP analyses.

Energy partitioning over the MRB shows a significant spread, with  $H$  estimates ranging from  $-553 \text{ MJ m}^{-2} \text{ yr}^{-1}$  (ECPC-RII) to  $1330 \text{ MJ m}^{-2} \text{ yr}^{-1}$  (CPTEC). Similar to the results from the ARB, the mean-OA estimate of soil heat flux ( $R_{\text{net}} - LE - H = 202 \text{ MJ m}^{-2} \text{ yr}^{-1}$ ) is high for an annual flux. Removing the ECPC-RII and NCEP analyses reduces the soil heat flux estimate to  $165 \text{ MJ m}^{-2} \text{ yr}^{-1}$ , which is less than  $50 \text{ MJ m}^{-2} \text{ yr}^{-1}$ . This confirms the observation by Bosilovich et al. (2009) that adding better (high correlation, lower error) members to an ensemble reduces the error of the ensemble, yet adding members with lower skill does not significantly degrade the ensemble while the better members are in place.

### 3) OTHER BASINS

The two basins discussed above are energy-limited basins. Although few studies have concentrated on the water-limited basins considered in this study, we find some distinct features that are comparable to the energy-limited basins. Firstly, precipitation estimates by the OA models were biased low (compared to the RS and VIC estimates), which is different from the high bias (except for over Amazon) seen in the energy-limited basins. In general, the energy-limited basins showed better agreement in the runoff estimates between the datasets. Table 5 shows the statistics of the various energy and water components as compared to the RS/observational estimates.

## 4. Summary and conclusions

One of the main objectives of the CEOP Water and Energy Simulation and Prediction (WESP) group is to address the question: what is our skill in predicting hydroclimatological water and energy budgets? Toward answering this question, this study serves as an evaluation of the predictive skill (for energy and water fluxes) of seven general circulation models (GCMs) run in operational assimilation mode against one LSM output and remote sensing retrievals at regional-to-global scales. The current study also serves as an extension to

the work by Yang et al. (2007), who compared MOLTS from five GCMs and three Global Land Data Assimilation System (GLDAS) LSMs to CEOP EOP-3 tower observations. The time period of the comparisons (2003/04) was dictated by the availability of the processed OA data (Bosilovich et al. 2009) and future work should extend this to a longer time period that better represents the mean climate and captures more variability and extreme events. Nevertheless, 2003/04 was not particularly outstanding globally in terms of ENSO activity, although events such as the European heat wave and above-average West African rainfall in 2003, and continuing drought in the U.S. west and eastern Australia through 2004, were important regionally (Levinson and Waple 2004; Levinson 2005).

At global scale, the mean-OA energy fluxes have significant biases relative to the RS and LSM data. Net radiation is biased high, because of overestimation of  $SW_{\downarrow}$  and underestimation of  $LW_{\uparrow}$ , mainly in the tropics. A positive bias ( $463 \text{ MJ m}^{-2} \text{ yr}^{-1}$ ) in latent heat flux, relative to the RS data, is balanced by a negative bias ( $-367 \text{ MJ m}^{-2} \text{ yr}^{-1}$ ) in the sensible heat flux. However, the biases in the ensemble mean and spread were mainly caused by two of the models (ECPC-RII and NCEP). All the OA models consistently overestimated (underestimated) latent (sensible) heat flux over the eastern part of the North American continent. There is an opposite bias over central Africa, although the uncertainties in the observational data and spread in the RS estimates are much higher here. Finally, the annual energy balance ( $R_{\text{net}} - LE - H$ ) was represented well by the OA models with a mean-OA residual of  $87 \text{ MJ m}^{-2} \text{ yr}^{-1}$  globally, which is likely due to a nonclosure term associated with model forecast error. This suggests that although the energy balance is reasonable in the OA models, the partitioning of the fluxes needs further improvement. This is also true for the RS and VIC models, which show quite different partitioning in some regions.

Comparison of the OA model precipitation with gauge-based data and remote sensing retrievals showed better agreement than the energy fluxes. The largest differences were in the deep tropics ( $15^{\circ}\text{S}$ – $15^{\circ}\text{N}$ ), although the uncertainties in the gauge-based datasets are higher because of interpolation of data from fewer stations. Comparison of  $ET/P$  and  $P - ET$  showed that the two analyses from the EPCP had significant bias, with  $ET/P$  ratios close to or greater than 1, which is physically unrealistic.

Over the Amazon River basin, OA model precipitation estimates are well correlated ( $\tau = 0.85$ ) with the gauge-based and remote sensing estimates, with a negative bias of  $0.70 \text{ mm day}^{-1}$ . Net radiation was also highly correlated but with a high positive bias ( $273 \text{ MJ m}^{-2} \text{ yr}^{-1}$ ),

due mainly to overestimation of  $SW_{\uparrow}$  radiation during July–April. OA model ET estimates were higher and out of phase with the remote sensing and VIC estimates, however other studies have indicated that the mean-OA (annual) estimates of ET are in good agreement with observations. The runoff estimates by most OA models are underestimated significantly ( $>700 \text{ mm yr}^{-1}$ ). Over the Mississippi, the three precipitation analyses (mean-OA, RS, and gauge observations) agreed well. Comparison with a fourth precipitation product (NLDAS, which is a high-resolution gauge-radar analysis with orographic adjustments for the continental United States; Cosgrove et al. 2003) showed that the CRU data were reasonable and the mean-OA estimates were within the errors associated (bias  $< 0.20 \text{ mm day}^{-1}$ ) with a precipitation product. The ECPC-RII and NCEP models were outliers over the MRB, mainly in their estimates of sensible heat flux. The bias in the OA ET is more than twice the bias in precipitation estimates, which may be attributed to the nudging process and the model spinup errors that have been a priori recognized as affecting the water budget closure (Maurer et al. 2001). Energy budget partitioning shows significant scatter among the surface fluxes, with a large error associated with the annual soil heat flux ( $>200 \text{ MJ m}^{-2} \text{ yr}^{-1}$ ). The ensemble mean is significantly improved by the addition of a skillful analysis model, but is not significantly degraded by the addition of lower skill models provided the better models were in place.

The errors in the OA-simulated surface fluxes are due to a combination of errors from their land surface scheme and the atmospheric model/assimilation system. Errors in the representations of clouds (e.g., for the ECPC-SFM) are a source of error in the radiation and may be linked to biases in the precipitation, although globally the models represent precipitation reasonably well. In water-limited regions, biases in the precipitation are more important and will induce biases in soil moisture and ET, but it is more likely that the land scheme induces biases in ET either directly or via its representation of the dynamics of soil moisture (e.g., Sheffield et al. 2012). The OA model land schemes are summarized in Table 1. Some OA models use simple bucket hydrology schemes with a single vegetation type, while the majority uses soil–vegetation–atmosphere transfer (SVAT) schemes of varying levels of complexity. Each scheme uses different vegetation and soil characteristics and spatial distributions (including specification of albedo and emissivity values), which will induce further differences in their estimates. Offline comparisons of land schemes have shown that no one model outperforms the others for all components of the water and energy budgets (e.g., Mitchell et al. 2004; Dirmeyer et al. 2006; Mueller et al. 2011), but SVATs tend to do better at

simulating energy fluxes and hydrological-orientated LSMs tend to do better at simulating runoff, soil moisture, and snow (Xia et al. 2011, manuscript submitted to *J. Geophys. Res.*). Model uncertainties (spread among models) have previously been found to be largest for soil moisture in comparison to variables more directly tied to the meteorological forcings such as surface temperature or net radiation, with evapotranspiration falling somewhere in between (Dirmeyer et al. 2006). It is therefore expected that the type and complexity of the land scheme play an important role in the level of errors in the surface fluxes, especially in moisture-limited regions where soil moisture is a dominant control. However, the model with the simplest scheme (BMRC) and the model with probably the most complex scheme (MSC) have the least error in net radiation. On the other hand, the BMRC does poorly at representing sensible heat flux and the MSC is an outlier in terms of ET over the Amazon. Furthermore, overestimation of ET by the ECPC-SFM model exemplifies that the role of the forcings can also be dominant. It therefore appears to be difficult to distinguish the role of the land scheme relative to the atmospheric/assimilation model without the use of offline experiments with the land schemes.

Betts (2004) addresses the importance of global models to understand the land–atmosphere climate system at local-to-global scales. He further points out the importance of the ET as “One way to encapsulate hydrometeorology is to ask what controls ET . . . .” Results from the current study support the importance of the ET process by analyzing the biases that exist in the OA estimates of latent heat flux, which lead to improper partitioning of the energy and water budgets. These biases can have a significant impact on the forecasting capability of models by enhancing the precipitation due to the overestimation of evapotranspiration. For the Amazon River basin, one important finding is the low correlation between the remote sensing and OA estimates (except for precipitation) of the various energy/water components. This questions our understanding of the energy and water cycles over this major basin. This work also demonstrates the challenge in evaluating predictions of the water and energy budgets at regional-to-global scales, whether they are from remote sensing observations and retrievals, in situ observations (given their sparse nature), LSMs, or OA models. This is an ongoing challenge to GEWEX since one of their scientific goals is to use remote sensing observations and retrievals to improve OA models.

*Acknowledgments.* This work was jointly supported by NASA Grants NNG04GQ32G “A Terrestrial Evaporation Data Product Using MODIS Data,” NNX08AN40A

“Developing Consistent Earth System Data Records for the Global Terrestrial Water Cycle,” and NNX09AK35G “Development and Diagnostic Analysis of a Multi-decadal Global Evaporation Product.” The operational model data for the current study were obtained from the authors of Bosilovich et al. (2009), which is part of the Coordinated Energy and Water Observation Project (CEOP, GEWEX), NASA Langley Research Center Atmospheric Science Data Center, Global Precipitation Climatology Center (GPCC), MEDIAS-France, and the Global River Discharge Center (GRDC).

#### REFERENCES

- Adam, J. C., and D. P. Lettenmaier, 2003: Adjustment of global gridded precipitation for systematic bias. *J. Geophys. Res.*, **108**, 4257, doi:10.1029/2002JD002499.
- , E. A. Clark, D. P. Lettenmaier, and E. F. Wood, 2006: Correction of global precipitation products for orographic effects. *J. Climate*, **19**, 15–38.
- Adler, R. F., and Coauthors, 2003: The Version-2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979–present). *J. Hydrometeorol.*, **4**, 1147–1167.
- Betts, A. K., 2004: Understanding hydrometeorology using global models. *Bull. Amer. Meteor. Soc.*, **85**, 1673.
- , P. Viterbo, and A. C. M. Beljaars, 1998: Comparison of the land-surface interaction in the ECMWF reanalysis model with the 1987 FIFE data. *Mon. Wea. Rev.*, **126**, 186–198.
- , J. H. Ball, and P. Viterbo, 1999: Basin-scale surface water and energy budgets for the Mississippi from the ECMWF reanalysis. *J. Geophys. Res.*, **104D**, 19 293–19 306.
- , —, M. Bosilovich, P. Viterbo, Y. Zhang, and W. B. Rossow, 2003a: Intercomparison of water and energy budgets for five Mississippi subbasins between ECMWF reanalysis (ERA-40) and NASA Data Assimilation Office fvGCM for 1990–1999. *J. Geophys. Res.*, **108**, 8618, doi:10.1029/2002JD003127.
- , —, and P. Viterbo, 2003b: Evaluation of the ERA-40 surface water budget and surface temperature for the Mackenzie River basin. *J. Hydrometeorol.*, **4**, 1194–1211.
- , —, —, A. G. Dai, and J. Marengo, 2005: Hydrometeorology of the Amazon in ERA-40. *J. Hydrometeorol.*, **6**, 764–774.
- , M. Zhao, P. A. Dirmeyer, and A. C. M. Beljaars, 2006: Comparison of ERA40 and NCEP/DOE near-surface data sets with other ISLSCP-II data sets. *J. Geophys. Res.*, **111**, D22S04, doi:10.1029/2006JD007174.
- , M. Köhler, and Y. C. Zhang, 2009: Comparison of river basin hydrometeorology in ERA-Interim and ERA-40 reanalyses with observations. *J. Geophys. Res.*, **114**, D02101, doi:10.1029/2008JD010761.
- Bosilovich, M. G., and R. Lawford, 2002: Coordinated Enhanced Observing Period (CEOP) International Workshop. *Bull. Amer. Meteor. Soc.*, **83**, 1495–1499.
- , D. Mocko, J. O. Roads, and A. Ruane, 2009: A multimodel analysis for the Coordinated Enhanced Observing Period (CEOP). *J. Hydrometeorol.*, **10**, 912–934.
- Callede, J., J. L. Guyot, J. Ronchail, M. Molinier, and E. De Oliveira, 2002: The River Amazon at Obidos (Brazil): Statistical studies of the discharges and water balance. *Hydrol. Sci. J.*, **47**, 321–333.

- Cess, R. D., and Coauthors, 1995: Absorption of solar radiation by clouds: Observations versus models. *Science*, **267**, 496–499.
- Chahine, M. T., 1992: The hydrological cycle and its influence on climate. *Nature*, **359**, 373–380.
- Cherkauer, K. A., L. C. Bowling, and D. P. Lettenmaier, 2003: Variable infiltration capacity cold land process model updates. *Global Planet. Change*, **38**, 151–159.
- Chou, S. C., J. A. Marengo, C. P. Derczynski, P. V. Waldheim, and A. O. Manzi, 2007: Comparison of CPTEC GCM and Eta model results with observational data from the Rondonia LBA reference site, Brazil. *J. Meteor. Soc. Japan*, **85**, 25–42.
- Cosgrove, B. A., and Coauthors, 2003: Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. *J. Geophys. Res.*, **108**, 8842, doi:10.1029/2002JD003118.
- Costa, M. H., and J. A. Foley, 1997: Water balance of the Amazon Basin: Dependence on vegetation cover and canopy conductance. *J. Geophys. Res.*, **102D**, 23 973–23 989.
- Cote, J., S. Gravel, A. Methot, A. Patoine, M. Roch, and A. Staniforth, 1998: The operational CMC-MRB Global Environmental Multiscale (GEM) model. Part I: Design considerations and formulation. *Mon. Wea. Rev.*, **126**, 1373–1395.
- Dirmeyer, P. A., X. Gao, M. Zhao, Z. Guo, T. Oki, and N. Hanasaki, 2006: GSWP-2: Multimodel analysis and implications for our perception of the land surface. *Bull. Amer. Meteor. Soc.*, **87**, 1381–1397.
- Ek, M. B., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*, **108**, 8851, doi:10.1029/2002JD003296.
- Fasullo, J. T., and K. E. Trenberth, 2008: The annual cycle of the energy budget. Part I: Global mean and land–ocean exchanges. *J. Climate*, **21**, 2297–2312.
- Fekete, B., C. Vörösmarty, and W. Grabs, 2000: Global, composite runoff fields based on observed river discharge and simulated water balances. Global Runoff Data Centre Rep., 120 pp. [Available online at <http://www.grdc.sr.unh.edu/html/paper/ReportUS.pdf>.]
- Fernandes, K., R. Fu, and A. K. Betts, 2008: How well does the ERA40 surface water budget compare to observations in the Amazon River basin? *J. Geophys. Res.*, **113**, D11117, doi:10.1029/2007JD009220.
- Fisher, J. B., K. P. Tu, and D. D. Baldocchi, 2008: Global estimates of the land–atmosphere water flux based on monthly AVHRR and ISLSCP-II data, validated at 16 FLUXNET sites. *Remote Sens. Environ.*, **112**, 901–919.
- Gao, H., Q. Tang, C. R. Ferguson, E. F. Wood, and D. Lettenmaier, 2010: Estimating the water budget of major U.S. river basins via remote sensing. *Int. J. Remote Sens.*, **31**, 3955–3978.
- Garratt, J. R., and A. J. Prata, 1996: Downwelling longwave fluxes at continental surfaces—A comparison of observations with GCM simulations and implications for the global land surface radiation budget. *J. Climate*, **9**, 646–655.
- Hinkelman, L. M., T. P. Ackerman, and R. T. Marchand, 1999: An evaluation of NCEP Eta model predictions of surface energy budget and cloud properties by comparison with measured ARM data. *J. Geophys. Res.*, **104D**, 19 535–19 549.
- Hirai, M., T. Sakashita, H. Kitagawa, T. Tsuyuki, M. Hosaka, and M. Oh'izumi, 2007: Development and validation of a new land surface model for JMA's operational global model using the CEOP observation dataset. *J. Meteor. Soc. Japan*, **85**, 1–24.
- Huffman, G. J., R. F. Adler, M. M. Morrissey, D. T. Bolvin, S. Curtis, R. Joyce, B. McGavock, and J. Susskind, 2001: Global precipitation at one-degree daily resolution from multisatellite observations. *J. Hydrometeorol.*, **2**, 36–50.
- , and Coauthors, 2007: The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55.
- Jiménez, C., and Coauthors, 2011: Global intercomparison of 12 land surface heat flux estimates. *J. Geophys. Res.*, **116**, D02102, doi:10.1029/2010JD014545.
- Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002a: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Amer. Meteor. Soc.*, **83**, 1631–1643.
- , and Coauthors, 2002b: NCEP dynamical seasonal forecast system 2000. *Bull. Amer. Meteor. Soc.*, **83**, 1019–1037.
- Karam, H. N., and R. L. Bras, 2008: Climatological basin-scale Amazonian evapotranspiration estimated through a water budget analysis. *J. Hydrometeorol.*, **9**, 1048–1060.
- Koike, T., 2004: The Coordinated Enhanced Observing Period—An initial step for integrated global water cycle observations. *WMO Bull.*, **53**, 115–121.
- Lawford, R., and Coauthors, 2006: U.S. contributions to the CEOP. *Bull. Amer. Meteor. Soc.*, **87**, 927–939.
- Levinson, D. H., Ed., 2005: State of the climate in 2004. *Bull. Amer. Meteor. Soc.*, **86**, S1–S86.
- , and A. M. Waple, Eds., 2004: State of the climate in 2003. *Bull. Amer. Meteor. Soc.*, **85**, S1–S72.
- Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.*, **99D**, 14 415–14 428.
- , —, and —, 1996: One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model. *J. Geophys. Res.*, **101D**, 21 403–21 422.
- Luo, Y., E. H. Berbery, K. E. Mitchell, and A. K. Betts, 2007: Relationships between land surface and near-surface atmospheric variables in the NCEP North American Regional Reanalysis. *J. Hydrometeorol.*, **8**, 1184–1203.
- Marengo, J. A., 2005: Characteristics and spatio-temporal variability of the Amazon River basin water budget. *Climate Dyn.*, **24**, 11–22.
- Maurer, E. P., G. M. O'Donnell, D. P. Lettenmaier, and J. O. Roads, 2001: Evaluation of the land surface water budget in NCEP/NCAR and NCEP/DOE reanalyses using an off-line hydrologic model. *J. Geophys. Res.*, **106D**, 17 841–17 862.
- Miralles, D. G., J. H. Gash, T. R. H. Holmes, R. A. M. de Jeu, and A. J. Dolman, 2010: Global canopy interception from satellite observations. *J. Geophys. Res.*, **115**, D16122, doi:10.1029/2009JD013530.
- Mitchell, K. E., and Coauthors, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.*, **109**, D07S90, doi:10.1029/2003JD003823.
- Mitchell, T. D., and P. D. Jones, 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *Int. J. Climatol.*, **25**, 693–712.
- Monteith, J. L., 1965: Evaporation and the environment. *Symp. Soc. Exp. Biol.*, **19**, 205–234.
- Mu, Q., F. A. Heinsch, M. Zhao, and S. W. Running, 2007: Development of a global evapotranspiration algorithm based on

- MODIS and global meteorology data. *Remote Sens. Environ.*, **111**, 519–536.
- Mueller, B., and Coauthors, 2011: Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations. *Geophys. Res. Lett.*, **38**, L06402, doi:10.1029/2010GL046230.
- Oki, T., K. Musiak, H. Matsuyama, and K. Masuda, 1995: Global atmospheric water-balance and runoff from large river basins. *Hydrol. Processes*, **9**, 655–678.
- , T. Nishimura, and P. Dirmeyer, 1999: Assessment of annual runoff from land surface models using Total Runoff Integrating Pathways (TRIP). *J. Meteor. Soc. Japan*, **77**, 235–255.
- Priestley, C. H. B., and R. J. Taylor, 1972: On the assessment of surface heat flux and evaporation using large-scale parameters. *Mon. Wea. Rev.*, **100**, 81–82.
- Ramanathan, V., B. Subasilar, G. J. Zhang, W. Conant, R. D. Cess, J. T. Kiehl, H. Grassl, and L. Shi, 1995: Warm pool heat budget and shortwave cloud forcing: A missing physics? *Science*, **267**, 499–503.
- Ramillien, G., F. Frappart, A. Güntner, T. Ngo-Duc, A. Cazenave, and K. Laval, 2006: Time variations of the regional evapotranspiration rate from Gravity Recovery and Climate Experiment (GRACE) satellite gravimetry. *Water Resour. Res.*, **42**, W10403, doi:10.1029/2005WR004331.
- Raschke, E., S. Bakan, and S. Kinne, 2006: An assessment of radiation budget data provided by the ISCCP and GEWEX-SRB. *Geophys. Res. Lett.*, **33**, L07812, doi:10.1029/2005GL025503.
- Rikus, L., 2007: Validating basic surface variables in the Australian Bureau of Meteorology model with CEOP EOP3 in-situ data. *J. Meteor. Soc. Japan*, **85**, 73–97.
- Ruane, A. C., and J. O. Roads, 2007: The diurnal cycle of water and energy over the continental United States from three reanalyses. *J. Meteor. Soc. Japan*, **85**, 117–143.
- Rudolf, B., and U. Schneider, 2005: Calculation of gridded precipitation. *Proc. 2nd Workshop of the Int. Precipitation Working Group (IPWG)*, Monterey, CA, EUMETSAT, 231–247.
- Seneviratne, S. I., P. Viterbo, D. Luthi, and C. Schar, 2004: Inferring changes in terrestrial water storage using ERA-40 reanalysis data: The Mississippi River basin. *J. Climate*, **17**, 2039–2057.
- Sheffield, J., and E. F. Wood, 2007: Characteristics of global and regional drought, 1950–2000: Analysis of soil moisture data from off-line simulation of the terrestrial hydrologic cycle. *J. Geophys. Res.*, **112D**, D17115, doi:10.1029/2006JD008288.
- , G. Goteti, and E. F. Wood, 2006: Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *J. Climate*, **19**, 3088–3111.
- , K. M. Andreadis, E. F. Wood, and D. P. Lettenmaier, 2009a: Global and continental drought in the second half of the twentieth century: Severity–area–duration analysis and temporal variability of large-scale events. *J. Climate*, **22**, 1962–1981.
- , C. R. Ferguson, T. J. Troy, E. F. Wood, and M. F. McCabe, 2009b: Closing the terrestrial water budget from satellite remote sensing. *Geophys. Res. Lett.*, **36**, L07403, doi:10.1029/2009GL037338.
- , B. Livneh, and E. F. Wood, 2012: Representation of terrestrial hydrology and large scale drought of the Continental United States from the North American Regional Reanalysis. *J. Hydrometeorol.*, in press.
- Stackhouse, P. W., S. K. Gupta, S. J. Cox, M. Chiacchio, and C. Mikovitz, 2000: The WCRP/GEWEX Surface Radiation Budget Project Release 2: An assessment of surface fluxes at 1 degree resolution. *IRS 2000: Current Problems in Atmospheric Radiation*, W. L. Smith and Y. M. Timofeyev, Eds., A. Deepak Publishers, 439–459.
- Su, F., J. C. Adam, K. E. Trenberth, and D. P. Lettenmaier, 2006: Evaluation of surface water fluxes of the pan-Arctic land region with a land surface model and ERA-40 reanalysis. *J. Geophys. Res.*, **111**, D05110, doi:10.1029/2005JD006387.
- Su, Z., 2002: The Surface Energy Balance System (SEBS) for estimation of turbulent heat fluxes. *Hydrol. Earth Syst. Sci.*, **6**, 85–99.
- Szeto, K. K., 2007: Assessing water and energy budgets for the Saskatchewan River Basin. *J. Meteor. Soc. Japan*, **85**, 167–186.
- Trenberth, K. E., J. T. Fasullo, and J. Kiehl, 2009: Earth's global energy budget. *Bull. Amer. Meteor. Soc.*, **90**, 311–323.
- Troy, T. J., E. F. Wood, and J. Sheffield, 2008: An efficient calibration method for continental-scale land surface modeling. *Water Resour. Res.*, **44**, W09411, doi:10.1029/2007WR006513.
- Vinukollu, R. K., E. F. Wood, C. R. Ferguson, and J. B. Fisher, 2011: Global estimates of evapotranspiration for climate studies using multi-sensor remote sensing data: Evaluation of three process-based approaches. *Remote Sens. Environ.*, **115**, 801–823.
- Wild, M., and E. Roeckner, 2006: Radiative fluxes in the ECHAM5 general circulation model. *J. Climate*, **19**, 3792–3809.
- , A. Ohmura, H. Gilgen, and E. Roeckner, 1995: Validation of general circulation model radiative fluxes using surface observations. *J. Climate*, **8**, 1309–1324.
- , —, —, J. J. Morcrette, and A. Slingo, 2001: Evaluation of downward longwave radiation in general circulation models. *J. Climate*, **14**, 3227–3239.
- Yang, K., and Coauthors, 2007: Initial CEOP-based review of the prediction skill of operational general circulation models and land surface models. *J. Meteor. Soc. Japan*, **85**, 99–116.
- Yeh, P. J. F., M. Irizarry, and E. A. B. Eltahir, 1998: Hydroclimatology of Illinois: A comparison of monthly evaporation estimates based on atmospheric water balance and soil water balance. *J. Geophys. Res.*, **103D**, 19 823–19 837.
- Zhang, Y., W. B. Rossow, and P. W. Stackhouse Jr., 2007: Comparison of different global information sources used in surface radiative flux calculation: Radiative properties of the near-surface atmosphere. *J. Geophys. Res.*, **112**, D01102, doi:10.1029/2005JD007008.