Reconciling Land–Ocean Moisture Transport Variability in Reanalyses with \( P – ET \) in Observationally Driven Land Surface Models

FRANKLIN R. ROBERTSON
Earth Science Office, NASA Marshall Space Flight Center, Huntsville, Alabama

MICHAEL G. BOSILOVICH
Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland

JASON B. ROBERTS
Earth Science Office, NASA Marshall Space Flight Center, Huntsville, Alabama

(Manuscript received 13 May 2016, in final form 30 August 2016)

ABSTRACT

Vertically integrated atmospheric moisture transport from ocean to land [vertically integrated atmospheric moisture flux convergence (VMFC)] is a dynamic component of the global climate system but remains problematic in atmospheric reanalyses, with current estimates having significant multidecadal global trends differing even in sign. Continual evolution of the global observing system, particularly stepwise improvements in satellite observations, has introduced discrete changes in the ability of data assimilation to correct systematic model biases, manifesting as nonphysical variability. Land surface models (LSMs) forced with observed precipitation \( P \) and near-surface meteorology and radiation provide estimates of evapotranspiration (ET). Since variability of atmospheric moisture storage is small on interannual and longer time scales, \( \text{VMFC} = P – ET \) is a good approximation and LSMS can provide an alternative estimate. However, heterogeneous density of rain gauge coverage, especially the sparse coverage over tropical continents, remains a serious concern.

Rotated principal component analysis (RPCA) with prefiltering of VMFC to isolate the artificial variability is used to investigate artifacts in five reanalysis systems. This procedure, although ad hoc, enables useful VMFC corrections over global land. The \( P – ET \) estimates from seven different LSMS are evaluated and subsequently used to confirm the efficacy of the RPCA-based adjustments. Global VMFC trends over the period 1979–2012 ranging from 0.07 to \(-0.03\) mm day\(^{-1}\) decade\(^{-1}\) are reduced by the adjustments to \(-0.016\) mm day\(^{-1}\) decade\(^{-1}\), much closer to the LSM \( P – ET \) estimate (0.007 mm day\(^{-1}\) decade\(^{-1}\)). Neither is significant at the 90\% level. ENSO-related modulation of VMFC and \( P – ET \) remains the largest global interannual signal, with mean LSM and adjusted reanalysis time series correlating at 0.86.

1. Introduction

Moisture transport to land from the global oceans is a crucial process linking the global water and energy cycles and is also at the heart of societal concerns regarding terrestrial water availability, food security, exposure to extreme weather events, and climate change. Recent best estimates of the net atmospheric transport of water to land (Rodell et al. 2015) put the climatological amount at \( 45.8 \pm 6.7 \times 10^3 \) km\(^3\) yr\(^{-1}\), or about 40\% of precipitation falling over land. The remainder of land precipitation arises from moisture recycling via evapotranspiration (ET; Eltahir and Bras 1994; Trenberth 1999; Bosilovich and Schubert 2002). The variability of this transport and its potential long-term trend at regional scales are emerging as a prime concern. Tropical circulations linked to sea surface temperature (SST) variability exert first-order controls on the delivery of water to land by virtue of El Niño–Southern Oscillation (ENSO) events (Ropelewski and Halpern 1987; Dai and Wigley 2000; Gu et al. 2007; Robertson et al. 2014). Midlatitude storm-track changes embodying teleconnections with tropical forcing also
have significant variations at higher latitudes. Over longer time scales Pacific decadal variability [PDV or Pacific decadal oscillation (PDO); e.g., Power et al. 1999; Dai 2013; Lyon et al. 2014] and other basin-scale phenomena [e.g., the Atlantic multidecadal oscillation (AMO); Enfield et al. 2001; Sutton and Hodson 2005; Ting et al. 2011] also modulate moisture transport. Anthropogenic radiative forcing changes and the consequent hydrologic cycle effects are expected to produce regional variations, encapsulated in the “wet get wetter and dry get drier” paradigm (Chou and Neelin 2004) wherein hydrologic extremes are expected to increase. As yet, evidence for this behavior in observational datasets is weak at best (Greve et al. 2014). There is also substantial uncertainty as to trends in soil moisture dryness depending on diagnostic approaches and choice of observed precipitation and surface meteorological forcing (Dai 2011; Sheffield et al. 2012; Trenberth et al. 2014). Untangling the role of these varied water and energy cycle mechanisms and their relationship to moisture transport continues to be a challenging task.

Moisture transport syntheses are routinely produced by reanalysis efforts (e.g., Kalnay et al. 1996; Onogi et al. 2007; Saha et al. 2010; Dee et al. 2011; Rienecker et al. 2011; Kobayashi et al. 2015) that blend diverse measurements of wind, moisture, and temperature as well as other observations with first-guess estimates from model short-term forecasts. While reanalyses effectively reconcile observations with physically based dynamical models, there are a number of practical problems that result in moisture transport fields typically having substantial systematic time-dependent biases (Trenberth et al. 2011; Robertson et al. 2011; Lorenz and Kunstmann 2012; Trenberth and Fasullo 2013; Robertson et al. 2014). The root of the difficulty lies in the fact that model physics (e.g., moist convective parameterizations, turbulence, and radiation) each have shortcomings so that assimilating models have biased climatologies. Once initialized, model forecasts (first guesses for analyses) “drift” toward a preferred state that differs from reality. But input data streams that correct this drift are nonstationary in the sense that observing system data densities, and satellite observations especially, have a time-dependent ability to correct the model first-guess fields. Therefore discrete biases develop in water and energy fluxes and transports.

For reanalyses the vertically integrated atmospheric moisture budget over land grid points is

\[
\frac{\partial W_v}{\partial t} = \text{VMFC} - P + \text{ET} + \text{ANA};
\]

that is, vapor plus condensate \(W_v\) increases as the result of vertically integrated atmospheric moisture flux convergence (VMFC) and ET and is depleted by precipitation \(P\). In reanalyses, the analysis increment (ANA) represents the departure of the forecast from the analysis divided by the temporal length of the corrector step—specifically in the case of MERRA and MERRA version 2 (MERRA-2), the forcing needed to drive the evolving reanalysis to the final analysis in a 6-h corrector step. Ideally this term should be randomly distributed about zero. Instead, time-dependent biases typically characterize ANA, reflecting bias contained in each of the physical terms in (1). Many previous studies (Trenberth and Guillemot 1998; Lorenz and Kunstmann 2012; Trenberth and Fasullo 2013) suggest that VMFC estimates have more consistency among reanalyses than precipitation minus evaporation \((P - E)\) derived from the model physics. But still, significant global land trends in VMFC were found by these studies.

These issues are seen in Fig. 1a, which shows reanalysis VMFC monthly anomalies around their respective annual mean. Trends over the period 1979–2012 range between \(-0.03\) and \(0.08\) mm day\(^{-1}\) decade\(^{-1}\) (Table 1) and represent roughly from \(-2.0\%\) to \(5.0\%\) of the climatological annual means. These trends are difficult to justify physically given recent estimates of \(P\) and ET changes (New et al. 2001; Jung et al. 2010) over recent decades. Discontinuities in the satellite record, in particular with the beginning of Special Sensor Microwave Imager (SSM/I) series in July 1987, the Advanced Microwave Sounding Unit A (AMSU-A) in late 1998, and Atmospheric Infrared Sounder (AIRS) in 2002 are known to link with abrupt changes in water and energy fluxes (Bosilovich et al. 2011; Trenberth et al. 2011; Robertson et al. 2011, 2014; Bosilovich et al. 2015).

LSMs and other related diagnostic models constrained by observations of precipitation, near-surface atmospheric variables, and radiation offer an independent estimate of terrestrial \(P - \text{ET}\). In these observationally constrained models, the water budget has the following form:

\[
\frac{\partial W_T}{\partial t} = P - \text{ET} - \text{RO},
\]

where column terrestrial water \(W_T\) (soil plus vegetation) is sustained by \(P\) but depleted by runoff (RO) and ET. Through efforts at the same institutions involved in global data assimilation (e.g., NASA, NCEP, and ECMWF) and various internationally coordinated programs [e.g., the Global Land Data Assimilation System (Rodell et al. 2004), the Global Soil Wetness Project (GSWP) and its successors (GSWP-2 and GSWP-3; Dirmeyer et al. 1999, 2006), Water and Global Change (WATCH; Harding et al. 2011), and Trends and Drivers of Regional Sources and Sinks of Carbon Dioxide...
reasonably mature diagnoses of \( P - ET \) are now available. These syntheses of land surface state and fluxes have facilitated the quantitative study of droughts and hydrologic variability— their scale, intensity, and some assessment of changes on a continental and global basis (e.g., Koster et al. 2010, 2011; Wisser et al. 2010; Haddeland et al. 2011; Sheffield et al. 2012; van Dijk et al. 2014). Figure 1b shows corresponding \( P - ET \) monthly anomalies from a number of these sources, along with their ensemble mean. To the extent that atmospheric moisture storage anomalies on monthly time scales are small, reanalysis VMFC and \( P - ET \) should be equivalent. At interannual-to-near-decadal time scales the agreement between these two quantities is reasonably good with systematic deficits (excess) moisture transport to land and smaller (larger) \( P - ET \) during El Niño (La Niña) events. El Niño events in 1982/83, 1986/87, 1991/92, and 1997/98 coincide with anomalously weak ocean–land moisture transport. After the turn of the century only the 2004/05 and 2009/10 events are prominent. At longer scales, however, the large trends in many of the reanalyses (0.08 and 0.07 mm day\(^{-1}\) decade\(^{-1}\) for MERRA and CFSR) are not shared by the LSMs, whose mean trend is 0.007 mm day\(^{-1}\) decade\(^{-1}\). Against the prominent interannual and longer excursions the mean LSM global land trend is not significant at the 0.90 level.

Further evidence for the likelihood of small trends in these budget components comes from the time series of annual global runoff shown in Fig. 2 (Dai et al. 2009; updated by Dai 2016). The values shown here are now in terms of millimeters per year since the RO values are aggregated over water years (October through September). The RO trend (0.26 mm yr\(^{-1}\)) is roughly

<table>
<thead>
<tr>
<th>LSM</th>
<th>( P - ET ) trend</th>
<th>( VMFC^* ) trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>MERRA-Land</td>
<td>0.023 (±0.019)</td>
<td>0.073 (±0.026)</td>
</tr>
<tr>
<td>MERRA-2</td>
<td>0.002 (±0.019)</td>
<td>0.003 (±0.026)</td>
</tr>
<tr>
<td>ERA-Interim/Land</td>
<td>0.001 (±0.017)</td>
<td>0.081 (±0.020)</td>
</tr>
<tr>
<td>GLDAS-2 Noah</td>
<td>0.022 (±0.012)</td>
<td>-0.030 (±0.018)</td>
</tr>
<tr>
<td>ORCHIDEE</td>
<td>0.009 (±0.017)</td>
<td>0.074 (±0.025)</td>
</tr>
<tr>
<td>CLM4C</td>
<td>0.023 (±0.015)</td>
<td></td>
</tr>
<tr>
<td>MPI-BGC</td>
<td>0.026 (±0.022)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reanalysis</th>
<th>( VMFC^* ) trend</th>
</tr>
</thead>
<tbody>
<tr>
<td>MERRA</td>
<td>0.073 (±0.026)</td>
</tr>
<tr>
<td>MERRA-2</td>
<td>0.003 (±0.026)</td>
</tr>
<tr>
<td>ERA-I</td>
<td>0.081 (±0.020)</td>
</tr>
<tr>
<td>JRA-55</td>
<td>-0.030 (±0.018)</td>
</tr>
<tr>
<td>CFSR</td>
<td>0.074 (±0.025)</td>
</tr>
</tbody>
</table>

TABLE 1. Trend statistics (mm day\(^{-1}\) decade\(^{-1}\)) for VMFC\(^*\) over land for various reanalyses and \( P - ET \) for LSM members over the period 1979–2012. In parentheses are errors calculated using lag-1 statistics to account for serial autocorrelation.
LSM monthly anomalies from Fig. 1 have been summed over water year intervals beginning in October 1979.

one-half that of $P - ET$, so that (2) implies a net continental water storage trend over the 30-plus-year period. This estimate is likely quite uncertain, although recent work by Reager et al. (2016) finds storage rates of 0.71 ($\pm 0.20$) mm yr$^{-1}$ over the period 2002–14 from GRACE measurements. The relevant point here is that the independent RO and $P - ET$ estimates both provide evidence that large multidecadal trends in reanalysis VMFC are exaggerated. Taken together with the known inconsistencies introduced by the changing observing system and other observational evidence against such large global trends, the VMFC decadal trends should be treated with considerable skepticism.

The objective of this paper is to explore reanalysis VMFC discrepancies with independent LSM-based estimates in more detail. Specifically, 1) we aim to characterize and quantify observing system influences that produce nonphysical trends in reanalysis VMFC trend over global land. We explore some of the regional patterns of variability, noting how sensitive global VMFC is to regional uncertainties. 2) In this process, we consider the $P - ET$ record of several observationally constrained LSMs as a surrogate for validation of VMFC. But uncertainties in forcing data as well as the model formulations are still important error sources (Jiménez et al. 2011; Mueller et al. 2013). Thus, we examine differences among the $P - ET$ estimates and evaluate their utility as a means of reanalysis validation. 3) We then show that using rotated empirical orthogonal function (REOF) analysis, along with some prefiltering, artificial steps, and trends induced by changing satellite data streams can be largely isolated and removed.

2. Data

Our investigation depends primarily on monthly mean data from global reanalyses and observationally constrained land surface models. Since available fields are at different native or archived grid resolutions, we interpolated all data to a 1.0° latitude–longitude resolution. Unless otherwise noted, all variability estimates are anomalies that were calculated by removing from the total fields a monthly resolved climatology for the respective datasets at each grid point over the period January 1979 through December 2010. Some minor departures from these dates are noted in the discussion below.

a. Reanalyses

VMFC is calculated from five state-of-the-art reanalysis projects—the NASA Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011) and an updated version, MERRA-2 (R. Gelaro et al. 2016, unpublished manuscript; Molod et al. 2015; Takacs et al. 2015); the European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim, hereinafter ERA-I; Dee et al. 2011); the new Japanese 55-year Reanalysis (JRA-55) produced by the Japanese Meteorological Agency (Kobayashi et al. 2015) that extends from 1958 to 2012; and the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR; Saha et al. 2010). For ERA-I and JRA-55 northward and eastward components of vertically integrated moisture transport were available and the horizontal flux divergences of these quantities were computed. For MERRA and MERRA-2 the divergence of the vertically integrated transport was archived as a standard product. CFSR VMFC has been derived by Trenberth et al. (2011) and was obtained directly from the National Center for Atmospheric Research (see Table 2).

In addition to the references noted here more indepth documentation of these reanalysis products, assimilating models, and data used can be found online (https://reanalyses.org/). Some of the more salient details of these reanalyses are provided in Table 2.

b. Land surface models

Products generated by LSMS rely on forcing data from reanalysis output as a first guess but, crucially, incorporate in situ observations, gauge precipitation, and some satellite data to relax substantially the biases of these initial estimates (e.g., Dirmeyer et al. 1999; Sheffield et al. 2006; Weedon et al. 2011). Still, there remain uncertainties whose origin and characteristics can be complex; thus, evaluation and validation of ET (and sensible heating and soil moisture) estimates is an ongoing process with assessments that have targeted model formulation and input forcing (Kato et al. 2007; Badgley et al. 2015) and used river discharge, GRACE, and field data (Zaitchik et al. 2010; Rodell et al. 2011) for validation. Intercomparison and validation efforts within larger collaborative efforts of LandFlux (Jiménez et al. 2011; Mueller et al. 2013) and the Water Model Intercomparison Project (WaterMIP; Haddeland et al. 2011; Harding et al. 2011) are generating
needed quantitative perspectives on flux and surface state uncertainties.

Seven different estimates of $P - ET$ are used in this study, each constrained by observational forcing: three are from land surface hydrology models—the Global Land Data Assimilation System [version 2 (GLDAS-2); Rodell 2004] initiative; MERRA-Land (Reichle et al. 2011), the MERRA-2 land component (Reichle and Liu 2014) that runs as part of the assimilation but uses an observation corrected precipitation analysis; and ERA-I global land surface dataset (ERA-Interim/Land; Balsamo et al. 2015). Two of the models—the Organizing Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE) model (Krinner et al. 2005) and the Common Land Model 4.0 for carbon (CLM4C; Oleson et al. 2010; Lawrence et al. 2011)—are dynamic global vegetation models used in the TRENDY initiative (Sitch et al. 2013), a contribution to the regional carbon cycle assessment and processes (RECCAP; Canadell et al. 2011). Finally, a diagnostic ET estimate from the Max Planck Institute for Biogeochemistry (MPI-BGC) flux dataset (Jung et al. 2009, 2010) uses a machine-learning methodology to scale up eddy covariance measurements from FLUXNET (Baldocchi et al. 2001). A surface energy balance constraint is combined with absorbed photosynthetically active radiation data derived from SeaWiFS (Gobron et al. 2007). For consistency with the derivation of this ET dataset we combine it with GPCC version 6 (v6) precipitation to construct $P - ET$ gridded values. Details and further references for these products are given in Table 3.

c. Other data

From the CMIP5 AMIP archive (http://cmip-pcmdi.llnl.gov/cmip5/data_portal.html) we selected five datasets for analysis that roughly sample the breadth of model diversity: the GFDL HiRAM-C180, GISS-E2-R, HadGEM2-A, MIROC5, and MRI-CGCM. Their monthly $P - ET$ fields are available for the period January 1979 through December 2008. Global runoff as presented by Dai et al. (2009) [updated by Dai (2016)] are used in Fig. 2. These data are used for comparison of time series behavior.

3. Isolating nonphysical changes in reanalysis VMFC

In the process of generating reanalyses, ANA in (1) yields information on the mismatch between the first guess (model forecast) and the observations. Schubert and Chang (1996) employed a least squares analysis of the projection of the physical terms in the moisture

---

Table 2. Summary characteristics of reanalysis datasets used in this study. (Acronym expansions are available online at http://www. ametsoc.org/PubsAcronymList.)

<table>
<thead>
<tr>
<th>Attributes</th>
<th>References</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERA-I IFS, cy31r2 (2006), 80 km (T255 spectral) grid with 60 vertical levels (Jan 1979–Dec 2012)</td>
<td>Dee et al. (2011); Dee and Uppala (2009); Simmons et al. (2010)</td>
<td>Four-dimensional variational data assimilation (4-DVar) system with adaptive estimation of satellite bias correction. RTTOV radiation operator, many revised analysis and physics improvements over ERA-40 (e.g., humidity and O₃) (<a href="http://apps.ecmwf.int/datasets/datasets/">http://apps.ecmwf.int/datasets/datasets/</a>).</td>
</tr>
<tr>
<td>MERRA GEOS-5.2.0 AGCM, 0.5° × 0.667° grid, 72 vertical levels (Jan 1979–Dec 2010)</td>
<td>Rienecker et al. (2011); Trenberth et al. (2011); Bosilovich et al. (2011); Robertson et al. (2011)</td>
<td>Three-dimensional variational data assimilation (3D-Var) Gridpoint Statistical Interpolation (GSI) analysis system with incremental analysis update (IAU) (<a href="http://disc.sci.gsfc.nasa.gov/uis/search/%22MERRA-2%22">http://disc.sci.gsfc.nasa.gov/uis/search/%22MERRA-2%22</a>).</td>
</tr>
<tr>
<td>MERRA-2 GEOS-5.12.4 AGCM, 0.5° × 0.667° grid, 72 vertical levels (Jan 1980–Dec 2012)</td>
<td>Molod et al. (2015); Takacs et al. (2015); Kleist et al. (2009)</td>
<td>GSI with IAU; significant additional new satellite data assimilated, now conserves dry air mass (<a href="http://disc.sci.gsfc.nasa.gov/daac-bin/FTPSubset.pl?LOOKUPID_List=MAIMCPASM">http://disc.sci.gsfc.nasa.gov/daac-bin/FTPSubset.pl?LOOKUPID_List=MAIMCPASM</a>).</td>
</tr>
<tr>
<td>CFSR GFS coupled atmosphere–ocean–land surface–sea ice system, T382L64 (2009) (Jan 1979–Dec 2009)</td>
<td>Saha et al. (2010); Wang et al. (2010); Trenberth and Fasullo (2013)</td>
<td>3D-Var GSI system, Atmospheric and Environmental Research (AER) radiation, Noah LSM with MOM; ocean is 0.25° at the equator, extending to a global 0.5° beyond the tropics, with 40 levels (<a href="http://www.cgd.ucar.edu/cas/catalog/reanalysis/index.html">http://www.cgd.ucar.edu/cas/catalog/reanalysis/index.html</a>).</td>
</tr>
<tr>
<td>Model</td>
<td>Attributes</td>
<td>Forcing</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td>CLM4C</td>
<td>Five primary subgrid land-cover types (glacier, lake, wetland, urban, and vegetated) in each grid cell. The vegetated portion of a grid cell is divided into patches of plant functional types with separate energy and water calculations.</td>
<td>Based on a merged product of Climatic Research Unit (CRU) observed monthly 0.5° analysis (version 3.0, 1901–2009; New et al. 2001) and the high temporal fidelity NCEP reanalysis forcing.</td>
</tr>
<tr>
<td>ERA-Interim/Land</td>
<td>Hydrology Tiled ECMWF Scheme of Surface Exchanges over Land (HTESSEL); 80-km resolution with 3-h integration steps.</td>
<td>ERA-I near-surface meteorology and radiation. ERA-I precipitation is rescaled using GPCP version 2.1.</td>
</tr>
<tr>
<td>GLDAS-2 Noah</td>
<td>one-dimensional column model, which can be executed in either coupled or uncoupled mode. Governing equations of the physical processes of the soil–vegetation–snowpack medium.</td>
<td>Updated Sheffield et al. (2006) forcing dataset based on the NCEP–NCAR reanalysis near-surface meteorological variables; GPCP, TRMM precipitation, and SRB radiation; CRU meteorological data are used to correct biases.</td>
</tr>
<tr>
<td>MERRA-2</td>
<td>Updated version of catchment model used in MERRA.</td>
<td>Precipitation constraints comprising anomalies from CMAP V0011 and RT pentad product plus GPCP version 2.1 climatology. Near-surface meteorology and radiation are from MERRA-2.</td>
</tr>
<tr>
<td>MPI-BGC</td>
<td>Machine-learning methodology, “model tree ensembles,” to upscale eddy covariance (EC) measurements from FLUXNET (Baldocchi et al. 2001) to a 0.5° monthly product.</td>
<td>AVHRR NDVI data SeaWiFS fPAR; CRU near-surface temperature.</td>
</tr>
</tbody>
</table>
budget onto ANA to infer errors in the GEOS reanalysis budget terms. Robertson et al. (2014) used the results of a principal component analysis (PCA) applied to MERRA ANA to regress out artifacts in the water and heat budget flux terms. That study showed that nonphysical modes of variability, due largely to increasing amounts of satellite data and their ability to counter model biases, can have prominent regional to global structure. Since ANA is not readily available for all of the reanalyses used in this study we use PCA in conjunction with prefiltering to identify these nonphysical components in the reanalysis VMFC.

The form of the continental moisture budget equation we analyze is

$$ VMFC = P - ET + Res, $$

where each term is a monthly gridpoint anomaly. Here we have subsumed all uncertainties regarding the moisture storage term and the remaining imbalance into a residual term (Res). This framework differs from (1) since we are taking $P - ET$ from the LSMs, a source independent of the reanalyses. How effectively can we then dissect reanalysis VMFC into its physical part and that due to observing system effects?

Simple PCA provides a compact treatment of variance contributions by mutually orthogonal modes in terms of the spatial coherence of the variability (EOFs) and associated temporal variability (PCs). Successive modes explain the maximum amount of remaining variance. However, individual PCA modes cannot generally be equated with specific sources of variability. Also, we cannot guarantee that physical signals and assimilation artifacts are collected into separate modes. After examining raw VMFC EOFs and PCs from a PCA decomposition for each reanalysis it was noted that while apparent nonphysical variability dominated the leading few modes, these were typically mixed with additional ENSO signals.

To identify the artifacts more clearly we first prefiltered the VMFC. The first eight PCs of an EOF analysis of Global Precipitation Climatology Centre (GPCC) precipitation were used to largely remove VMFC physical signals at each grid point via principal component regression:

$$ VMFC_{pf} = VMFC - \sum_{i=1}^{m} \text{cov}(VMFC, PC_i)PC_i, $$

where $VMFC_{pf}$ is the prefiltered VMFC and $PC_i$ is the $i$th PC of the GPCC precipitation. To the extent that ET covaries with $P$, we can think of this step as removing VMFC covarying with $P$ and ET. This approach has two attributes: First, it does not add any source of physical variability to VMFC since by construction it only removes VMFC signals that project onto $P$ variability. This would not be the case if we just subtracted the $P$ anomalies from those of VMFC. Second, since we are first removing much of the physical VMFC signal, this minimizes the likelihood that any subsequent analysis of $VMFC_{pf}$ will mistakenly identify physical low-frequency behavior or trends as being artifacts.

We then take a conservative approach of using just the leading few modes of a rotated PCA of $VMFC_{pf}$ as representing the bulk of the artifact signals. We chose to rotate the modes (i.e., make linear combinations of them) to collect regional variability into fewer leading modes. Using the varimax constraint (Richman 1986), we rotated the leading 10 modes. The raw PCs were scaled by $-1/(\text{eigenvalues})^{1/2}$ before input to the rotation matrix so as to preserve orthogonality among the rotated PCs yet relax that constraint for the rotated EOF patterns. The product of the rotated PC (RPC) time series and the REOFs recovers each mode’s contribution to VMFC variability. The resulting “artifact” modes can then be subtracted from the raw VMFC leaving $VMFC^*$, the estimated physical variability:

$$ VMFC^* = VMFC - \sum_{j=1}^{n} \text{REOF}_{jpf} \text{RPC}_{jpf}, $$

where $\text{REOF}_{jpf}$ and $\text{RPC}_{jpf}$ are the $j$th REOF and RPC of $VMFC_{pf}$. Inserting (5) into (3) we have

$$ VMFC = VMFC^* + \sum_{j=1}^{n} \text{REOF}_{jpf} \text{RPC}_{jpf} $$

$$ = P - ET + Res. $$

Since $VMFC_{pf}$ is determined by an ad hoc procedure that only minimizes the presence of true physical variability, we cannot regard this whole signal as being the artifact to be removed. Thus, a more conservative approach is to use only the leading modes that have some obvious relationship to changes in the assimilated data streams. To the extent that $VMFC^*$ and $P - ET$ agree, the Res term is explained by the sum of the “artifact” modes and other nonsystematic VMFC and $P - ET$ errors. Results of this analysis and the methods to determine the number of modes used along with a sensitivity analysis are presented in section 7.

4. Regional VMFC and $P - ET$ contributions to global land averages

Regional interannual signals and trends

To determine the extent to which variability shown in Figs. 1a,b is manifest regionally, we first examine maps of root-mean-square (RMS) monthly mean VMFC
anomalies (Figs. 3a–e) and trends (Figs. 3f–j) for each reanalysis over the period 1979–2012. Anomalies are departures from the respective monthly varying climatologies. For comparison, the RMS and trend of the ensemble LSM $P - ET$ and reanalysis VMFC are shown in Fig. 4. The anomalies were composited before the RMS and trend were calculated. In addition the MERRA and MERRA-2 values were averaged to form one sample so as maintain diversity of the reanalysis systems and LSMs considered.

In general terms, the RMS values of ERA-I, JRA-55, and MERRA (Fig. 3) agree reasonably with those of the LSM ensemble mean (Fig. 4). MERRA-2 and CFSR RMS values are a factor of 2 or more greater in many places over tropical continents. One striking feature is that except for MERRA all reanalyses have larger RMSs over west-central Africa compared to that of the LSM ensemble mean. LSM values are typically less than 1.0 mm day$^{-1}$ there, whereas reanalysis values exceed 1.5–2.0 mm day$^{-1}$ over broad areas and are frequently much larger. There are great observational challenges over tropical continents, especially Africa, not only for the radiosonde density but also for rain gauge and surface atmospheric measurements. Over much of South America, reanalysis VMFC variability agrees well with that of the LSMs, again with MERRA-2 and CFSR being much larger. Variability over the headwaters of the Amazon basin is slightly stronger in the LSM versus

**Fig. 3.** Statistics for monthly mean VMFC anomalies for various reanalyses over the period 1979–2012: (a)–(e) RMS of deviations (mm day$^{-1}$) and (f)–(j) trends (mm day$^{-1}$ decade$^{-1}$).
the reanalysis ensembles. A separate center of strong variability common to the reanalyses and LSMs (Figs. 4a,b) is present over the La Plata basin, a region of strong convective activity and storm-track origination, both modulated by ENSO.

Reanalysis VMFC trends show regional structure that does not average out in the ensemble mean and contrasts in many areas with ensemble LSM trends (Figs. 3a–e and Figs. 4c,d). Strong downward trends exceeding 1 mm day$^{-1}$ decade$^{-1}$ dominate central Africa in all reanalyses except for CFSR, which is strongly positive. These signals are much weaker in the LSMs. Upward trends exceeding 1 mm day$^{-1}$ decade$^{-1}$ in the reanalyses are seen in East Africa but are not found in the LSMs. The Maritime Continent and upper reaches of the Amazon basin trend upward in all reanalyses in agreement with the LSMs. Negative VMFC tendencies extend from southern Brazil through the La Plata basin but are much less organized than the negative $P - ET$ values from the LSMs. Somewhat surprising is the lack of agreement between the reanalyses in terms of trends over the United States. In this area of dense observational data, LSM trends are near zero but ERA-I and JRA-55 have strong downward trends. In another data-rich region over northern and eastern Europe, the upward VMFC trends extend across all reanalyses but are very weak in the LSMs.

One might wonder whether variability of moisture storage (1) might be large enough to explain some of the discrepancies between LSMs and the reanalyses. We calculated monthly mean $\partial W_a/\partial t$ from MERRA-2 using 1-h data since this intramonthly time resolution provides the largest-amplitude signal. Monthly mean, atmospheric storage anomalies averaged over global land (not shown) are an order of magnitude less than VMFC and $P - ET$ values shown in Figs. 1a,b. However, regional monthly mean RMS $W_a$ tendencies can reach near 0.50 mm day$^{-1}$ in subtropical regions and are not negligible for constructing moisture budgets over many areas. Nevertheless, these RMS tendencies are much smaller over tropical continents and cannot explain the reanalysis–LSM discrepancy over central Africa. Furthermore, regional trends in moisture storage are negligible and cannot explain the differing trends for VMFC compared to $P - ET$ in Figs. 3 and 4.

5. Analysis of regional VMFC errors

The results of section 4 reinforce our assertion that the differences between the LSM $P - ET$ and the reanalyses’ VMFC variability on longer than interannual time scale are attributable to systematic reanalysis errors that have largest expression over tropical regions. We now examine several of these specific regions where the LSM $P - ET$ and reanalysis VMFC differ so strongly.

a. Western equatorial Africa

Based on the trend differences in Figs. 3 and 4 we examine the region extending from 10°W to 20°E and from 5°S to 5°N. Time series of VMFC for the reanalyses and mean LSM $P - ET$ is shown in Fig. 5a. Each reanalysis shows a distinct change in behavior near the end of 1988. ERA-I drops sharply in a steplike manner as does JRA-55. MERRA amplitudes decrease by half but with far less evidence of a change in mean value. Attributing these changes to a specific data stream is difficult. SSM/I ingest began in late 1987 for the reanalyses,
but these effects have to be indirect since the radiances are not used over land. The transition between NOAA-9 and NOAA-11 MSU data also occurs in late 1988. Near the end of 1998 MERRA and JRA-55 show a return to a higher-amplitude seasonal oscillation with opposite polarity of the pre-1988 period. The MERRA VMFC anomalies also begin a drop over the next 5–7 years. CFSR shows a pronounced increase in 2002 consistent with the beginning of AIRS data. Characteristic of each dataset is a change in the annual cycle at the end of 1988 and again in 1998. The annual cycle phase shifts between these periods, and its amplitudes decrease in the 1988–98 period. Clearly, these signals are not physical.

Some insight into MERRA’s behavior can be gleaned from Fig. 5b. In addition to the MERRA VMFC anomalies here we also plot the first two PCs of the global, vertically integrated moisture ANA, which is the forcing needed to drive the forecast as close as possible to the analysis of observations. These PCs are analogous to those plotted in Fig. 7 of Robertson et al. (2011). PC1 minus PC2 also almost exactly recaptures the VMFC time series. The systematic, nonrandom behavior of the two PCs is evidence of systematic changes in the ability of the data to correct the model forecast first guess. PC1 shows a small but clearly visible drop in late 1987 coincident with SSIM/I availability. Further distinct drops in late 1998 and again in late 2000 correspond to NOAA-15 and NOAA-16 AMSU-A data onset. PC2 carries the main signal of the seasonal cycle change. The distinct reduction in seasonal cycle amplitude between 1992 and late 1998 corresponds to the tenure of NOAA-12. This PC2 behavior is common to all three reanalyses except CFSR, suggesting that some combination of satellite-induced changes in the moisture analysis are significant and that the temporal stability of the reanalysis moisture budgets in this region is unreliable.

b. Coastal Ecuador and Colombia

The coastal Ecuador and Colombia region we examined is land area encompassed within the boundaries of equator−10° N and 80°–70° W. The choice of this area is motivated by examination of the PCA results to be discussed in section 7 (not shown). Clearly ERA-I is the outlier here (Fig. 5c), particularly with respect to the jump in VMFC in the spring of 2004. This zero-order change appears associated with the assimilation of METAR surface pressure reports beginning at this time (Fig. 11 in Dee et al. 2011). It is unclear exactly why these data produce this effect, but assimilation of surface pressures that are in disagreement with the first-guess forecast could cause analyzed mass changes that subsequently affect the divergent wind, vertical motion, and, ultimately, moisture transport fields. These effects can then propagate some distance before being damped. Despite the small area of influence this large near-zero-order change in VMFC is the primary signal of ERA-I RPC2. In this region we also see some evidence of a jump in VMFC in the spring of 2004. This zero-order change appears associated with the assimilation of ATOVS and AIRS data. The exact attribution of the VMFC changes or JRA-55 is not yet clear. Except for this ERA-I problem after NH spring 2004, the agreement between the LSM mean and the reanalyses is quite good after about 1990.

c. Central United States

A large downward trend in ensemble mean VMFC over the central United States was noted in Figs. 3 and 4c, driven primarily by ERA-I and JRA-55. The time series in Fig. 5d shows a distinct downward transition for these two reanalyses of approximately 0.6 mm day−1 at the end of 1994. Since VMFC climatological values are of comparable size in this region (Fig. 6a) this represents a significant change in the moisture balance. Although there is also a decrease in P − ET from the LSMS before and after 1995, that change is much more gradual. A change in sources of conventional data from historical archives to the ECMWF operational feed beginning in 1995 (Uppala et al. 2005) may explain its distinct VMFC decrease in ERA-I. JRA-55 incorporated this same data as used by ECMWF (Kobayashi et al. 2015), which may explain the similar discontinuity in that system. Complicating this interpretation is the fact that 1995 also marks the transition from NOAA-I1 to NOAA-I4 sounder coverage. Because of the diurnal drift of the afternoon satellite equatorial crossing times there was approximately a 3-h diurnal cycle difference between these two sensors. Bosilovich et al. (2015) have analyzed VMFC behavior in MERRA but note that this shift occurs near 2000 and may be more related to the assimilation of ATOVS and AIRS data. The exact attribution of the VMFC changes to different data streams remains to be settled, yet it is clear that even over data-rich regions such as the continental United States significant VMFC artifacts exist.

6. LSM P − ET signals and uncertainties

Although observationally constrained LSMS offer a physically consistent estimate of terrestrial water balance, there exist uncertainties stemming not only from model physics formulation but also from the quality of the forcing data (Badgley et al. 2015). Precipitation datasets (e.g., GPCC, Global Precipitation Climatology
FIG. 5. Time series of reanalysis VMFC and ensemble LSM anomalies (mm day$^{-1}$) over (a) equatorial Africa, (c) coastal Colombia and Ecuador region, and (d) central United States. (b) MERRA VMFC anomaly (cyan) and first two global PCs of vertically integrated moisture increment over equatorial Africa.
Project (GPCP), and others] differ in sampling, gauge undercatch, and data quality. More problematic is near-surface meteorology and radiative forcing. These variables are taken from reanalyses also but are bias adjusted using surface observations and satellite radiative fluxes (e.g., Sheffield et al. 2006; Weedon et al. 2011). Given the contrasts noted between reanalysis VMFC and LSM \( P - ET \) it is necessary to assess uncertainties in the LSMs to further quantify their credibility vis-à-vis the reanalyses.

Recent work by the LandFlux evaluation community (LandFlux-EVAL) has highlighted uncertainties in LSMs, diagnostic retrievals, and reanalyses. In an initial assessment of flux estimates over the 1993–95 period, Jiménez et al. (2011) find ET uncertainties on the order of 0.50–0.70 mm day\(^{-1}\) relative to an annual mean of about 1.60 mm day\(^{-1}\). Mueller et al. (2013) extend this study in developing a baseline time series of flux estimates including interannual variability and trends. They attribute much of this uncertainty to differences in precipitation forcing used, the influence of water-limited ET regimes, and interception by vegetation. On the other hand, there also exists a fair degree of sensitivity to model formulation evident when LSMs are run using identical forcing data (Schlosser and Gao 2009). Lipton et al. (2015) point to ET differences between satellite driven diagnostic approaches and LSMs noting sensitivities to surface parameters and LSM forcing precipitation. Still, Mueller et al. (2013) find realistic interannual variations in ET from composites of these methods are present, including an upward trend between 1989 and 1997 followed by a downward ET trend during the 1998–2005 period. This behavior is most robust for the LSMs and echoes the earlier results of

---

**Fig. 6.** (a) Ensemble mean climatological \( P - ET \) (mm day\(^{-1}\)). (b) S/N ratio for LSMs (see text for details on calculation). (c) The \( P - ET \) ACC time series of each LSM with the ensemble mean.
Jung et al. (2010). One common finding from these and other studies is that no single model can be regarded as sufficient and that multiple models with alternative forcing offer the most reliable syntheses of fluxes.

LSM ensemble mean climatological mean $P - ET$ patterns (Fig. 6a) look very much like precipitation climatologies with large values over the Amazon basin, Maritime Continent, and Southeast Asia. Storm tracks impinging on the west coasts of North America and Chile are present. Still, quantitative differences exist among mean $P$ and ET climatological means (not shown). Despite the climatological uncertainties, the global anomalies in Fig. 1b show good coherence. To assess more deeply the character of the $P - ET$ anomalies comprising the ensemble estimate we examine two statistical metrics. The mean signal-to-noise ratio (S/N) is defined as $S/N = \frac{\sigma^2_{LSM}}{\sigma^2_{LSM}}$, where $\sigma^2_{LSM}$ is the square of the ensemble mean monthly $P - ET$ anomaly and $\sigma^2_{LSM}$ is the mean of the individual squared departures of the $P - ET$ anomalies from the ensemble mean monthly anomaly. The S/N diagnostic (Fig. 6b) is a local measure of uncertainty among the LSM members in defining $P - ET$ monthly anomalies. Densely populated and gauged areas of the eastern United States, Europe, and China have systematically high values ranging from 5 to 8. The periphery of Australia and South America also show values in the range from 3 to 5. Deserts (the Sahara and those in central Asia and the interior of Australia) have the lowest values owing to sporadic rain as well as a dearth of gauges. S/N values within key tropical precipitation regimes of Brazil, New Guinea, and central Africa are typically 3 or less and likely suffer most directly from insufficient gauge density. These values are for 1.0° resolution data, and it is important to keep in mind that spatial averaging to coarser resolution of several degrees enhances these numbers significantly.

For a more global skill metric we use the anomaly correlation coefficient (ACC) first introduced by Miyakoda et al. (1972):

$$ACC = \frac{\langle p_{i}m_{j}\rangle}{\left(\langle p_{i}^{2}\rangle\langle m_{j}^{2}\rangle\right)^{1/2}}.$$ (7)

Here, $p_{i}$ and $m_{j}$ are the $P - ET$ anomalies at grid point $(i, j)$ at any given time for any dataset and the ensemble mean anomaly is defined by $m$. The angle brackets denote area-weighted averaging over all grid points $(i, j)$ within the 60°N–60°S land domain. ACC values measure spatial pattern fidelity as a function of time. Because we lack independent $P - ET$ validation on these scales, these diagnostics are more a measure of $P - ET$ sensitivity to input data and model formulation than of accuracy. Results for the seven different LSM $P - ET$ datasets relative to the six-member ensemble mean are given in Fig. 6c. (As noted earlier we average the MERRA-Land and MERRA-2 $P - ET$ values as input to the LSM ensemble mean.) MPI-BGC and ORCHIDEE both use GPCC precipitation. ERA-I uses GPCP precipitation, which is strongly tied to GPCC gauge data but also differs because of an adjustment to deal with undercatch of gauges. Thus, these three precipitation forcing datasets dominate the six-member ensemble mean. Values are reasonably stable and generally lie in the range from 0.75 to 0.95. Experience has shown that values above 0.60 generally are indicative of agreement on the synoptic scale. Thus, from a global coherence perspective, the datasets are similar in their spatial patterns. Sensitivity to the precipitation forcing has a significant influence. Accordingly, MPI-BGC and ORCHIDEE ACC values are each strongly correlated with the ensemble mean; ERA-Interim/Land also shows high correlations. CLM4C ACC values tend to decline in time, especially after 2000. The much smaller number of precipitation gauges used in the CRU time series version 3.10 (TS3.10) product forcing for CLM4C has been shown to lead to a systematic overestimation of precipitation since the mid-1990s (Trenberth et al. 2014; Dai and Zhao 2016). This appears to influence the lowering CLM4C correlations with time. MERRA-Land and MERRA-2 $P - ET$ values are systematically low, although this happens in part because the Climate Prediction Center (CPC) unified precipitation (CPCU) and CPC Merged Analysis of Precipitation (CMAP) data represent one sample compared to effectively three GPCC forcing sets. There are notable departures though with MERRA-Land before 1982 and GLDAS-2 Noah LSM in 2006. These periods reflect outliers that originate in the precipitation forcing (CPCU and the amalgam of datasets that are used for Noah LSM).

ACC calculations performed separately for $P$ and ET (not shown) yielded uniformly high correlations for the former ($>0.80$) while those for ET averaged between 0.5 and 0.8. The lower ET correlations likely reflect the varied physical formulations among the models and the uncertainties in radiative forcing and near-surface moisture and temperature.

7. Isolating artifacts via REOF analysis (assessing space–time variability)

To the extent that we believe the mean LSM $P - ET$ trends (Fig. 4b) the differences with the VMFC trends in Figs. 3a–e and Fig. 4c indicate the regional trend errors or artifacts inherent in the raw reanalyses. In this section we now determine how effectively the rotated principal
component analysis (RPCA) methodology can be used to capture these effects in a few modes.

a. Adjustment effects on regional trends

In section 3 we outlined the methodology of applying RPCA to the quantity VMFC$^{pf}$ in order to identify the leading structures and temporal variability of artificial variability induced by changes in observing system input. For most of the reanalyses a single RPCA mode, $n = 1$, identified the globally averaged trends characterized as steplike transitions. However, it was found that typically three modes were needed to effectively capture regional trend artifacts. This determination was made by visually inspecting the RPCs and EOFs of each reanalysis VMFC$^{pf}$ and confirming that those modes contained RPC “discontinuities or steps” that coincided with satellite changes such as SSM/I, AMSU, or AIRS. We also confirmed that these modes made changes that reduced the regional discrepancy between the trend patterns of VMFC in Fig. 3 and the ensemble LSM $P − ET$ in Fig. 4d. Only the MERRA and MERRA-2 reanalyses showed that additional modes were significant in changing regional trends. Thus, we applied $n = 5$ for MERRA and MERRA-2 and $n = 3$ for the others as constituting the signal of changes induced by evolving assimilation data input. The sensitivity of trend patterns to inherent subjectivity of this selection process is discussed below.

This new VMFC estimate for each reanalysis, VMFC*, can then be compared to $P − ET$ of the ensemble mean LSMS. Although VMFC* is now not formally independent of $P − ET$, none of the $P$ variability has modified VMFC*. The effects of these potential adjustments are presented in Fig. 7. Figure 7, left, contains the area-averaged VMFC signal (black) and diagnosed area average of the artifact that must be removed (red line). Figure 7, right, contains the trends in VMFC*, the adjusted reanalyses after this artifact signal has been removed. Figure 7, bottom, is the ensemble mean trend of VMFC*. Note the geometric progression of the color scale. (Recall that we have averaged the MERRA and MERRA-2 data and considered it as one of four reanalysis systems.)

It is clear from Fig. 7 (left, red lines) that VMFC$^{pf}$ collects a large amount of trend or low-frequency variability. For example, the impact of AIRS after 2002 in elevating CFSR VMFC is quite apparent. SSM/I availability after 1987 changes ERA-I and JRA-55 noticeably. The large ERA-I VMFC jump in 2004 over coastal Ecuador and Colombia (Fig. 4c) has a significant global impact. The artifact time series for each reanalysis also has high-frequency signals since significant observing system changes such as SSM/I, AMSU, and AIRS also affect the VMFC annual cycle. After adjusting the reanalyses at each grid point the spatial trend patterns (Fig. 7, right) and the ensemble mean (Fig. 7, bottom) are much smaller compared to those in Fig. 3 and show significant changes in structure. Over central Africa the amplitudes of VMFC* and $P − ET$ decreases are much more consistent. Both ensemble VMFC* and LSM $P − ET$ (Fig. 4d) trends hint at a tendency for increases in moisture convergence to the south over Zambia and Angola and north over portions of the Sahel. This pattern suggests perhaps increasing annual latitudinal excursion of the ITCZ in these regions. MERRA and CFSR VMFC* no longer have huge upward trends over Australia, and the large VMFC increases over East Africa common to all reanalyses have been removed. The Amazon basin shows increased moisture convergence over time with associated reductions over southern Brazil. There remain differences though with the LSMS positioning the $P − ET$ increases over the headwaters region and the reanalyses having the upward VMFC* trends more toward the east. An interesting aspect of this analysis is the relatively small fraction of VMFC$^{pf}$ total variance needed to explain these regional trend artifacts: JRA-55 (18.09%), ERA-I (14.15%), MERRA (19.08%), MERRA-2 (15.29%), and CFSR (11.61%). This indicates that the bulk of VMFC$^{pf}$ variability does not project onto regional VMFC$^{pf}$ trends. The leading mode for each reanalysis largely explains the global average land trend and contributes typically about 6%. Since the prefiltering removes much of the physical signal we interpret this remaining VMFC$^{pf}$ as predominantly error in higher-frequency VMFC regional signals.

The global land area-averaged VMFC* trends (60°N−60°S) shown in Fig. 8 are now each much reduced. The ensemble mean value (Fig. 8b) is now 0.016 (±0.13) mm day$^{−1}$ decade$^{−1}$ over the period 1979−2012. This result is much closer to the mean trend of the LSMS [0.007 (±0.010) mm day$^{−1}$ decade$^{−1}$] and the AMIP models [0.012 (±0.016) mm day$^{−1}$ decade$^{−1}$].

Perhaps the most significant outcome of the VMFC adjustment is the improved agreement with LSM $P − ET$ in terms of regional trend patterns and their amplitudes. With these adjustments the pattern correlation of the trends increases from 0.41 to 0.55. Two applications of a nine-point spatial filter were used prior to determining the correlations that raised the correlations by about 0.09. To gauge how sensitive this result is to our selection of modes we repeated the analysis with one-, two- and three-mode-only corrections. The resulting pattern correlations for the ensemble VMFC* with the LSM mean were 0.048, 0.48, and 0.50, respectively. Higher modes ($n > 5$) were examined, but
their relationship to satellite changes was not clear; in keeping with a conservative approach to making corrections, these were not used. Individual pattern trend correlations were as follows: 0.48 for JRA-55, 0.61 for ERA-I, 0.43 for MERRA, 0.44 for MERRA-2, and 0.38 for CFSR. Although ERA-I exceeded the ensemble mean, the value for an ensemble with ERA-I removed is 0.52, thus supporting the value of an ensemble strategy as generally providing more skill than individual ensemble members.

Fig. 7. (left) Time series of globally averaged (60°N–60°S) reanalysis VMFC (black) and area-averaged corrections (red) (mm day$^{-1}$). Niño 3.4 SST multiplied by 0.1 is plotted as gray shading with inverted scale on right-hand side. (right) Trends (mm day$^{-1}$ decade$^{-1}$) in VMFC$^*$ over the period 1979–2012 for various reanalyses after corrections have been applied. Compare to Figs. 3a–e, which shows the uncorrected trends. (bottom) Ensemble mean corrected VMFC$^*$ trends.
The agreement in trend patterns and amplitudes between VMFC* and the LSMs (Fig. 7, bottom, and Fig. 4d) is therefore quite improved over the raw VMFC. Studies seeking to explain decadal changes and trend patterns like these have consistently pointed to SST variations as important controls on regional hydrologic anomalies even if details of patterns, seasonality, and intensity remain unresolved. The AMO has been found to influence rainfall over the Sahel (Folland et al. 1986; Giannini et al. 2003), northeastern Brazil (Hastenrath and Greischar 1993; Folland et al. 2001), and the United States (Enfield et al. 2001). Gloor et al. (2013) note the effects of Atlantic SST changes on the upward trend in wet season rainfall over Amazonia since 1980. Low-frequency ENSO-like behavior of Pacific SSTs (Power et al. 1999; Zhang et al. 1997) has been argued as forcing for global monsoon variations (Wang et al. 2012). Positive phases of the PDO, the North Pacific component of this SST variability, are associated with an increase in precipitation in the central and northern parts of the Amazon but decrease in the southern parts (Marengo 2004). These PDV teleconnections are global as evidenced by Lyon and DeWitt (2012), who have shown that recent spring declines in East African rainfall are tied to cold eastern tropical Pacific SSTs. Asefi-Najafabady and Saatchi (2013) have noted a continued downward trend in precipitation over central Africa by merging CRU and TRMM data, although Washington et al. (2013) strongly caution against reliance on any precipitation dataset in this part of Africa.

b. Adjustment effects on interannual variability

Time series of globally averaged VMFC for the individual corrected reanalyses are given in Fig. 8a, with Fig. 8b showing time series area averages of ensemble corrected reanalyses and the mean LSMs. Comparing Fig. 8a to Fig. 1a the reduced trends reveal more consistent interannual VMFC signals among the reanalyses, and the relationship between interannual VMFC* and Niño-3.4 SST anomalies is much clearer. Ensemble mean reanalyses and LSMs (Fig. 8b) correlate well (correlation is 0.86). Ensemble averaging over multiple AMIP experiments reduces internal atmospheric variations (i.e., “weather noise”) that cannot represent the correct deterministic signals that were observed (Bretherton and Battisti 2000). Thus, the remaining AMIP signal is only that component forced by SST. Differences in AMIP P − ET or VMFC anomaly response structure to SST anomalies are also present. These factors lower the AMIP correlation with the LSMs and VMFC* (correlation is 0.64 and 0.52, respectively). The agreement between the three datasets thus confirms the significant role that interannual SST variations play in land–ocean moisture exchange.

Fig. 8. (a) Time series of individual corrected reanalysis VMFC* global land area average (mm day$^{-1}$; 60°N–60°S). (b) The P − ET (mm day$^{-1}$) from individual LSMs (black) and mean VMFC* from corrected reanalyses (red) and AMIP models (cyan). A 3-month running smoothing is applied. Niño-3.4 SST anomalies are plotted in gray shading with inverted scale (°C) on the right y axis.
To explore the degree to which spatial VMFC patterns have been affected, we assess the changes in VMFC (equivalently, $P - ET$) patterns via ACC—this time between the individual corrected reanalyses and the ensemble LSM (Fig. 9). The ACC of the ensemble mean raw reanalyses is also plotted (black dotted line), which indicates that on average the adjustment process has slightly degraded VMFC agreement with $P - ET$ (less than 0.05 on average). Here we see that on an individual reanalysis basis, the ACC is typically only 0.35–0.60. Again, some of this limitation is local to regional details of the ensemble LSM $P - ET$ values. But the skill of the ensemble mean corrected reanalysis exceeds that of the individual members.

An indication of where VMFC and $P - ET$ agreement has changed can be gleaned from local correlations between their time series (Fig. 10). There is excellent agreement in locations of dense station sampling and significant rainfall but there is a strong resemblance between Fig. 10a and the S/N estimates of Fig. 6b. This indicates the significant limitations of the LSM signals likely produced by rain gauge sparse density. This lack of station coverage means that in some areas even where the corrected reanalyses are improved the LSMs have such poor ability to discern signals that they cannot confirm this. Changes in correlation with the LSM ensemble $P - ET$ compared to the raw reanalyses is shown in Fig. 10b. There are areas of improvement as well as reduced agreement. Many areas in the tropics are improved, but the sparsely gauged areas in Africa and the headwaters of the Amazon are not. In central Asia agreement with the LSMs shows both strong positive and negative changes in VMFC–LSM agreement.

The results of Figs. 9 and 10 might raise concerns about the effect the REOF adjustments have on interannual signals. To check this we composited VMFC and $P - ET$ anomalies based on warm Niño-3.4 SST anomaly maxima during boreal winters 1982/83, 1986/87, 1991/92, 1994/95, 1997/98, 2004/05 and 2009/10 (not shown). Pattern correlations of VMFC and VMFC* with $P - ET$ were 0.82 and 0.79, respectively. All three composites had regional anomaly patterns the canonical precipitation anomalies first isolated by Ropelewski and Halpert (1987) and more recently by Camberlin et al. (2001), Grimm (2003), Hendon (2003), and Malhi and Wright (2004). We conclude from these results that interannual variability related to SST forcing is not significantly altered by our RPCA adjustments.

8. Conclusions

In this study we have sought to characterize the uncertainties in estimating variations in moisture transport...
of moisture from ocean to land. Reanalyses and LSMs offer two nearly independent methodologies for estimating components of the atmospheric water budget. On seasonal and longer time scales moisture transport should be equivalent to net precipitation minus evaporation. Although reanalyses offer VMFC estimates determined from dynamical modeling constraints on observations, the episodic introduction of new data sources, particularly satellite data streams, has introduced serious time dependent biases. LSMs offer similar physically based constraints on precipitation and surface meteorological forcing in determining \( P \) – ET. Although these forcings also have their own uncertainties, our assessment shows that they offer a strong quantitative assessment of VMFC issues. We have also shown that RPCA diagnostics, although ad hoc, can be applied to adjust the raw VMFC estimates. A posteriori, these error reductions are justified by improved agreement of regional trends between the LSM \( P \) – ET and VMFC regional trends. Our findings can be summarized as follows:

(i) The large trends in near-global mean moisture convergence over land during the period 1979–present in reanalyses are predominantly an artifact due to changes in assimilated data streams and the ability of those data streams to correct model biases. These biases differ among reanalyses because of their differing physical parameterization formulations and aspects of the data assimilation methodology. Averaged over the ensemble LSMs a small net positive trend in \( P \) – ET (0.007 mm day\(^{-1}\) decade\(^{-1}\)) is found but is only significant at 90% confidence.

(ii) Corrections to VMFC \((P \text{–} ET)\) using RPCA with prefiltering to identify the nonphysical signal are effective in removing many of the problems and substantially enhance the agreement in regional \( P \) – ET trends during the 1979–2012 period. RPCA-based adjustments also result in an improvement in trend field correlation from 0.41 to 0.55. Simple PCA is likely to fail as signals of the artifacts and those of physical variability are mixed. The decision on how many modes are needed to represent artifact structure in any reanalysis is subjective and depends on cross-referencing RPCA results with assimilated observational data stream metadata.

(iii) Interannual ENSO-related variations and their decadal-scale modulation are highly consistent between the LSMs and adjusted VMFC time series (correlation is 0.86) and composite El Niño \( P \) – ET and VMFC patterns (correlation is 0.84). Although the adjustments are not needed to detect these interannual signals (Fig. 1a), the agreement between VMFC and \( P \) – ET interannual variability is more evident in time series plots (Fig. 8b).

(iv) Despite uncertainties inherent with observationally constrained LSMs, these syntheses can help identify and corroborate problems associated with reanalysis data changes. The sparseness and uneven sampling of precipitation gauging in remote areas (e.g., tropical continents, especially central Africa) are a significant uncertainty in estimating interannual variability. However, corrections to near-surface meteorology and radiative forcing are important (Ngo-Duc et al. 2005) and need additional scrutiny.

(v) CMIP5 AMIP experiments, despite having somewhat distorted VMFC patterns not directly studied here, also corroborate our estimated VMFC* corrections. Although not encumbered with the effects of changing observing systems, these experiments can only confirm the role of SST forcing since internal atmospheric variability is only at best stochastically consistent with the historical record.

How do we envision the utility of the present results? The broader problem of reconstructing water and energy budget variability, whether from reanalyses assimilating observations or from diagnostic methods (e.g., Pan et al. 2012; van Dijk et al. 2014; Rodell et al. 2015; L’Ecuyer et al. 2015), requires identifying and accounting for time dependent biases. The present
results are a step in that direction in that they would facilitate combining VMFC and \( P - ET \) estimates in a diagnostic approach. From a broader perspective additional opportunities are apparent for indirect checks and estimates of VMFC in (1) and (2). Satellite retrievals of \( W_T \), \( P \), and ET are one direct means to check VMFC. The accuracy of these retrievals varies according to space and time. Robertson et al. (2014) have recently shown that existing \( P \) and especially \( E \) estimates over global oceans have serious uncertainties for the purpose of climate variability studies. Although retrieval physics errors contribute to these problems, the intercalibration of multiple sensors and temporal changes in global sampling are also important issues. Improvements are actively being pursued. The lack of robust passive microwave remote sensing before late 1987 on which these retrievals are based is a limitation.

The purpose of climate variability studies. Although improvements are actively being pursued. The lack of robust passive microwave remote sensing before late 1987 on which these retrievals are based is a limitation. From the terrestrial side using (2) it is possible to determine changes in \( W_T \) directly via the Gravity Recovery and Climate Experiment (GRACE) satellite mission (Tapley et al. 2004) and to recover a \( P - ET \) estimate with RO measurements from river and streamflow gauges. Although GRACE measurements are a unique resource for enabling this approach, those data exist only since 2003.

Another source of information may come from reduced observation reanalyses that, in addition to SST, sea ice, and radiative constituents, also assimilate surface pressure observations (Compo et al. 2011) and marine wind speeds (Poli et al. 2015). While these less robust data constraints also perhaps minimally enforce actual synoptic weather realism, discontinuities in their multidecadal records appear to be less of a problem than those in the satellite record. These limitations are offset by the property that the SST (and surface pressure) forcing is largely free from the more discrete changes in atmospheric observing systems. Conceivably these integrations could also be run with observed land surface forcing as was applied in MERRA-2.

Ideally, improved model physical parameterizations and removal of data stream biases would mean that analysis increments or innovations would essentially be unbiased and normally distributed. However, model physics improvements (e.g., AGCM convective, turbulence, and cloud parameterizations and LSM soil, vegetation, and water routing formulations) are long-term development efforts, and our discontinuous data streams, particularly from satellites, will always present a time-varying capability to correct assimilating model errors in water and energy fluxes. Continued reevaluation of these modeling, retrieval, and in situ resources is necessary to narrow uncertainties in quantifying climate variability.

Acknowledgments. This work was sponsored in large part by the NASA Energy and Water Cycle Study (NEWS) program (Grant WBS281945)—Dr. Jared Entin, Program Manager. The authors wish to acknowledge the institutions and individuals involved in development and production of the many datasets listed here. (For references and data access, see Tables 1 and 2.) The authors also gratefully acknowledge the many individuals who discussed, enabled data access to, and contributed datasets used in this investigation: Drs. Steven Sitch and Sam Levis for access to the CLM4C data under the auspices of the TRENDY initiative and Drs. Graham Weedon and Ben Poulter for the ORCHIDEE integrations from the WFDEI component of the WATCH program. Dr. Markus Reichstein facilitated access to the MPI-BGC satellite-based ET estimates. MERRA and MERRA-2 are distributed by the Goddard Earth Sciences (GES) Data and Information Services Center (DISC). JRA-55 data were obtained from the project website (http://jra.kishou.go.jp/JRA-55/index_en.html). CFSR VMFC fields were obtained from the National Center for Atmospheric Research Climate Analysis Section data holdings (http://www.egd.ucar.edu/cas/catalog/newbudgets/index.html). We thank Dr. Aiguo Dai for providing the updated global river runoff time series.

REFERENCES


——, J. D. Chern, D. Mocko, F. R. Robertson, and A. M. da Silva, 2015: Evaluating observation influence on regional water


Sutton, R. T., and D. L. Hodson, 2005: Atlantic Ocean forcing of
Sitch, S., and Coauthors, 2013: Trends and drivers of the regional-scale