Reconciling land/ocean moisture transport variability in reanalyses with P-ET in observationally-driven land surface models

Franklin R. Robertson¹, Michael G. Bosilovich² and Jason B. Roberts¹

¹NASA / Marshall Space Flight Center, Earth Science Office
MSFC, AL 35812

²GSFC Global Modeling and Assimilation Office
Greenbelt, MD 20771

Corresponding author: Franklin R. Robertson
(pete.robertson@nasa.gov)
tel: (256) 961-7836

Submitted to J. Climate (MERRA-2 Special Collection)

Aug 22, 2016
Abstract

Vertically-integrated atmospheric moisture transport from ocean to land, VMFC, is a dynamic component of the global climate system but remains problematic in atmospheric reanalyses with current estimates having significant multi-decadal global trends differing even in sign. Regional VMFC trends over continents are especially uncertain. Continual evolution of the global observing system, particularly step-wise improvements in satellite observations, has introduced discrete changes in the ability of data assimilation to correct systematic model biases, manifesting as non-physical 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, 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 pre-filtering of VMFC to isolate the artificial variability is used to investigate artifacts in five reanalysis systems. This procedure, though ad hoc, enables useful VMFC corrections over global land. 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 d\(^{-1}\) decade\(^{-1}\) are reduced by the adjustments to 0.016 mm d\(^{-1}\) decade\(^{-1}\), much closer to the LSM P-ET estimate (0.007 mm d\(^{-1}\) decade\(^{-1}\)). Neither is significant at the 90 percent 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. 2014) put the climatological amount at $45.8 \pm 6.7 \times 10^3 \text{ km}^3\text{ yr}^{-1}$, 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 (Ropeleski and Halpert, 1987; Dai and Wigley, 2000; Gu et al, 2007; Robertson et al, 2014). Mid-latitude storm track changes embodying teleconnections with tropical forcing also have significant variations at higher latitudes. Over longer time scales Pacific Decadal Variability (PDO / PDV), (e.g. Power et al, 1999; Dai 2013; Lyon et al, 2013) and other basin scale phenomena (e.g. the Atlantic Multi-decadal Oscillation, AMO, Enfield et al, 2001; Sutton and Hodgson, 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 / dry get drier” paradigm (Chou and Neelin, 2004) wherein hydrologic extremes are expected to increase. As yet, evidence for this behavior in observational data sets 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 and 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 which 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 non-stationary 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,

$$\frac{\partial W_a}{\partial t} = \text{VMFC} - P + \text{ET} + \text{ANA},$$ (1)

i.e., $W_a$, vapor plus condensate increases as the result of vertically-integrated atmospheric moisture flux convergence (VMFC), evapotranspiration, 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-2, the forcing needed to drive the evolving reanalysis to the final analysis in a 6h 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 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 Figure 1a which shows reanalysis VMFC monthly anomalies around their respective annual mean. Trends over the period 1979 to 2012 range between -0.03 to 0.08 mm d^{-1} decade^{-1} (Table 1) and represent roughly -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 (SSMI) series in July 1987, the Advanced Microwave Sounding Unit-A (AMSU A) in late 1998, and Atmospheric Interferometer Sounder (AIRS) in 2002 are known to link with abrupt changes in water and energy fluxes (Bosilovich et al, 2011; Trenberth et al, 2011 and 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-ET. In these observationally-constrained models, the water budget has the form

\[
\frac{\partial w_T}{\partial t} = P - ET - RO
\]
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 3 [Dirmeyer et al, 1999, 2006]; Water and Global Change, WATCH, [Harding et al, 2011]; Trends and Drivers of Regional Sources and Sinks of Carbon Dioxide, TRENDY, [Sitch et al, 2013]), 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, 2009, 2011; Wisser et al, 2010; Haddeland et al, 2011; Sheffield et al, 2012; van Dijk et al, 2013). 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 of (excess) moisture transport to land and smaller (larger) P-ET during El Niño (La Niña) events. El Niño events in 1982/1983, 1986/1987, 1991/1992, 1997/1998 coincide with anomalously weak ocean to land moisture transport. After the turn of the century only the 2004/2005 and 2009/2010 events are prominent. At longer scales though the large trends in many of the reanlysises (0.08 and 0.07 mmd^-1 decade^-1 for MERRA and CFSR) are not shared by the LSMs whose mean trend is 0.007 mmd^-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 Figure 2, (Dai, 2009; updated by Dai, 2016). The
values shown here are now in terms of mmy$^{-1}$ since the RO values are aggregated over water years
(OCT through SEP). The RO trend (0.26 mmy$^{-1}$) is roughly 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 (+/- .20)
mmy$^{-1}$ over the period 2002-2014 from GRACE measurements. The relevant point here though is
that the independent RO and P-ET estimates both provide evidence that large multi-decadal 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 non-physical 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 (Jimenez et al, 2011
and 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 pre-filtering, 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 degree latitude by 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 data sets at each gridpoint over the period Jan 1979 through Dec 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 (Gelaro et al. 2016; Molod et al, 2015; Takacs et al, 2015); the European Centre for Medium Range Weather Forecasting (ECMWF) Interim Reanalysis, ERA-I, (Dee et al. 2011); the new 55-year reanalysis produced by the Japanese Meteorological Agency, JRA-55, (Kobayashi et al, 2015) that extends from 1958 to 2012; and the National Centers for Environmental Prediction Climate Forecast System Reanalysis, NCEP / CFSR (Saha et al. 2010).

For ERA-I and JRA55 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.
In addition to the references noted here more in-depth documentation of these reanalysis products, assimilating models and data used can be found at 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, 2006; 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 (Jimenez et al, 2011; Mueller et al, 2013) and WaterMIP (Haddeland, 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 initiative, GLDAS-2 (Rodell, 2004); MERRA-Land (Reichle et al, 2011; 2012); 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-Interim Land (Balsamo et al, 2015). Two of the models ORCHIDEE (Organising Carbon and Hydrology In Dynamic Ecosystems) model (Kriner et al, 2005) and the Common Land Model, CLM4C, (Oleson et al, 2010; Lawrence et al, 2011) are dynamic global vegetation models used in the Trends and Drivers of Regional Sources
and Sinks of Carbon Dioxide (TRENDY) initiative (Sitch et al., 2013), a contribution to the REgional Carbon Cycle Assessment and Processes (RECCAP; Canadell et al., 2013). Finally, a diagnostic ET estimate from the MPI-BGC flux data set (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. 2006). For consistency with the derivation of this ET data set we combine it with GPCC 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 CMIP-5 AMIP archive (http://cmip-pcmdi.llnl.gov/cmip5/data_portal.html) we selected five data sets 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, 2009 (updated by Dai, 2016) were used in Figure 2. These data are used for comparison of time series behavior.

3. Isolating non-physical changes in reanalysis VMFC

In the process of generating reanalyses, the analysis increment, 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 budget onto ANA to infer errors in the GEOS reanalysis budget terms. Robertson et al. (2014) used the results of a principal components analysis (PCA) applied to MERRA ANA to regress out artifacts in the water and heat budget flux terms. That study showed that non-physical
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 pre-filtering to identify these non-physical components in the reanalysis VMFC.

The form of the continental moisture budget equation we analyze is

\[ VMFC = P - ET + Res, \] (3)

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. 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. Nor can we guarantee physical signals and assimilation artifacts to be collected into separate modes. After examining raw VMFC EOFs and PCs from a PCA decomposition for each reanalysis it was noted that while apparent non-physical variability dominated the leading few modes, these were typically mixed with additional ENSO signals.
To identify the artifacts more clearly we first pre-filtered the VMFC. The first eight PCs of an EOF analysis of Global precipitation Climatology Center (GPCC) precipitation were used to largely remove VMFC physical signals at each grid point via principal component regression:

\[
VMFC_{pf} = VMFC - \sum_{i=1}^{m} \text{cov}(VMFC, PC_i) \cdot PC_i, \tag{4}
\]

where \(VMFC_{pf}\) is the pre-filtered VMFC and \(PC_i\) is the \(i\)th PC of the GPCC precipitation. To the extent that ET co-varies with \(P\), we can think of this step as removing VMFC co-varying 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 (Richmond, 1986), we rotated the leading 10 modes. The raw PCs were scaled by \(-1/\sqrt{\text{eigenvalues}}\) 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 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,
where $\text{REOF}_j^\text{pf}$ and $\text{RPC}_j^\text{pf}$ are the $j$th REOF and RPC of VMFC$^\text{pf}$. Inserting (5) into (3) we have

$$\text{VMFC} = \text{VMFC}^\ast + \sum_{j=1}^{n} \text{REOF}_j^\text{pf} \cdot \text{RPC}_j^\text{pf} = \text{P} - \text{ET} + \text{Res}. \quad (6)$$

Since VMFC$^\text{pf}$ is determined by an ad hoc procedure that only minimizes the presence of true physical variability we can’t 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$^\ast$ and P-ET agree, the Res term is explained by the sum of the “artifact” modes and other non-systematic 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

a. Regional Interannual Signals and Trends

To determine the extent to which variability shown in Figure 1a, b is manifest regionally, we first examine maps of root mean square (RMS) monthly mean VMFC anomalies (Figure 3, left panel) and trends (right panel) 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 Figure 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 (Figure 3) agree reasonably with those of the LSM ensemble mean (Figure 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 d⁻¹ there whereas reanalysis values exceed 1.5 to 2.0 mm d⁻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 the reanalysis ensembles. A separate center of strong variability common to the reanalyses and LSMS (Figure 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 (Figure 3 right, Figure 4c, d). Strong downward trends exceeding 1 mmd⁻¹decade⁻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 mmd⁻¹decade⁻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 U.S. 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 Eq. (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-hour data since this intra-monthly 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 Figure 1a, b. However, regional monthly mean RMS $W_a$ tendencies can reach near 0.50 mm/day 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 Figures 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 Figures 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 Figure 5a. Each reanalysis shows a distinct change in behavior near the end of 1988. ERA-I drops sharply in a step-like manner as does JRA55. 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. SSMI 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 NOAA9 and NOAA11 MSU data also occurs in late 1988. Near the end of 1998 MERRA and JRA55 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 data set 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 1988-1998 period. Clearly, these signals are not physical.

Some insight into MERRA’s behavior can be gleaned from Figure 5b. In addition to the MERRA VMFC anomalies here we also plot the first two PCs of the global, vertically-integrated moisture analysis increment, 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 Figure 7 of Robertson et al. (2011). PC1 minus PC2 also almost exactly recaptures the VMFC time series. The systematic, non-random 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 SSMI availability. Further distinct drops in late 1998 and again in late 2000 correspond to NOAA15 and NOAA16 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 NOAA12. 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 / Colombia

The coastal Ecuador / Colombia region we examined is land area encompassed within the
boundaries of Eq. to 10N and 80W to 70W. 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 (Figure
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
(Figure 11 in Dee et al, 2011). It is unclear as to 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 jump-like behavior for
ERA-I VMFC anomalies in late 1987 although the presence of what appear to be ENSO-related
signals in 1982/83 and 1986 complicate the actual magnitude of changes. Why this VMFC
increase is more pronounced for ERA-I than in MERRA or JRA55 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 U.S.

A large downward trend in ensemble mean VMFC over the central U.S. was noted in Figures 3 and 4c, driven primarily by ERA-I and JRA-55. The time series in Figure 5d shows a distinct downward transition for these two reanalyses of approximately 0.6 mmd⁻¹ at the end of 1994. Since VMFC climatological values are of comparable size in this region (Figure 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 11 to NOAA 14 sounder coverage. Because of the diurnal drift of the PM satellite equatorial crossing times there was approximately a 3h 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 U.S. 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 from the quality of the forcing data (Badgley et al, 2015). Precipitation data sets (e.g. GPCC,
Global Precipitation Climatology Project [GPCP] and others) differ in sampling, gauge under-
catch, and data quality. More problematic is near-surface meterorology and radiative forcing.
These variables are taken from reanalyses too 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-EVAL community has highlighted uncertainties in LSMs,
diagnostic retrievals and reanalyses. In an initial assessment of flux estimates over the 1993-1995
period, Jimenez et al. (2011) find ET uncertainties of order 0.50 to 0.70 mmd$^{-1}$ relative to an
annual mean of about 1.60 mmd$^{-1}$. 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
downward ET trend of during the 1998-2005 period. This behavior is most robust for the LSMs
and echoes the earlier results of 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 (Figure 6a) look very much like precipitation climatologies with large values over the Amazon Basin, Maritime Continent, and S.E. Asia. Storm tracks impinging on the west coasts of N. 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 Figure 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

\[ S/N = \frac{\sigma_{\text{LSM}}^2}{\sigma_{\text{LSM}}^2} \]

where \( \sigma_{\text{LSM}}^2 \) is the square of the ensemble mean monthly P-ET anomaly and \( \overline{\sigma_{\text{LSM}}^2} \) is the mean of the individual squared departures of the P-ET anomalies from the ensemble mean monthly anomaly. The S/N diagnostic (Figure 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 US, 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 of 3 to 5. Deserts (the Sahara, 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 degree 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):

\[ ACC \equiv \frac{\langle p_{\text{m}} m_{\text{u}} \rangle}{\left( \langle p_{\text{m}}^2 \rangle \langle m_{\text{u}}^2 \rangle \right)^{1/2}} \] (7)
Here, $p_{ij}$ are the P-ET anomalies at gridpoint $i, j$ at any given time for any data set and the ensemble mean anomaly is defined by $m$. The angle brackets denote area-weighted averaging over all gridpoints $i, j$ within the $60^\circ$ N/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 data sets relative to the six-member ensemble mean are given in Figure 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 usesGPCP precipitation which is strongly tied to GPCC gauge data but also differs because of an adjustment to deal with under-catch of gauges. Thus these three precipitation forcing data sets dominate the six member ensemble mean. Values are reasonably stable and generally lie in the range of 0.75 to .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 data sets 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-I 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 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 "Unified" (CPCU) and CPC Merged Analysis of Precipitation (CMAP) precipitation 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 in 2006. These periods reflect outliers that originate in the precipitation forcing (CPCU, and the amalgam of data sets that are used for Noah).

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 (Figure 4b) the differences with the VMFC trends in Figures 3 (right hand side) and 4c indicate the regional trend errors or artifacts inherent in the raw reanalyses. In this section we now determine how effectively the 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 step-like 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$^{hf}$ and confirming that those modes contained RPC “discontinuities or steps” that coincided with satellite changes such as SSMI, AMSU or AIRS. We also confirmed that these modes made changes that reduced the regional discrepancy between the trend patterns of VMFC in Figure 3 and the ensemble LSM P-ET in Figure 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. Though 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 Figure 7. The left panel contains the area-averaged VMFC signal (black) and diagnosed area-average of the artifact that must be removed (red line). The right panel contains the trends in VMFC*, the adjusted reanalyses after this artifact signal has been removed. At the bottom of Figure 7 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 Figure 7 (left panel, red lines) that VMFC* 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. SSMI availability after 1987 changes ERA-I and JRA-55 noticeably. The large ERA-I VMFC jump in 2004 over coastal Ecuador / Colombia (Figure 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 SSMI, AMSU and AIRS also affect the VMFC annual cycle. After adjusting the reanalyses at each gridpoint the spatial trend patterns (Figure 7, right panel) and the ensemble mean (bottom) are much smaller compared to those in Figure 3 and show significant changes in structure. Over central Africa the amplitudes of VMFC* and LSM P-ET decreases are much more consistent. Both ensemble VMFC* and LSM P-ET (Figure 4d)
trends hint at a tendency for increases in moisture convergence to the south over Zambia / 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\textsuperscript{pf} total variance needed to explained these regional trend artifacts: JRA55 (18.09%), ERA-I (14.15%), MERRA (19.08%), MERRA-2 (15.29%) and CFSR (11.61%). This indicates that the bulk of VMFC\textsuperscript{pf} variability does not project onto regional VMFC\textsuperscript{pf} trends. Each reanalysis’ leading mode largely explains the global average land trend and contributes typically about 6%. Since the pre-filtering removes much of the physical signal we interpret this remaining VMFC\textsuperscript{pf} as predominantly error in higher frequency VMFC regional signals.

The global land area-averaged VMFC* trends (60° N/S) shown in Figure 8 are now each much reduced. The ensemble mean value (Figure 8b) is now 0.016 (+/- 0.13) mmd\textsuperscript{-1}decade\textsuperscript{-1} over the period 1979 to 2012. This result is much closer to the mean trend of the LSMs (0.007 [+/ -0.010] mmd\textsuperscript{-1}decade\textsuperscript{-1}) and the AMIPs (0.012 [+/- 0.016] mmd\textsuperscript{-1}decade\textsuperscript{-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 9-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
1-, 2- and 3-mode only corrections. The resulting pattern correlations for the ensemble mean VMFC* with the LSM mean were .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: JRA-55 = 0.48, ERA-I = 0.61, MERRA = 0.43, MERRA-2 = 0.44 and CFSR = 0.38. Though 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.

The agreement in trend patterns and amplitudes between VMFC* and the LSMs (Figure 7, bottom and Figure 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), northeast Brazil (Hastenrath and Greisher, 1990; Folland et al, 2001) and the U.S. (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, 1999) 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 (2016) have noted a continued downward trend in
precipitation over central Africa by merging CRU and TRMM data, though Washington et al. (2013) strongly caution against reliance on any precipitation data set 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 Figure 8a with Figure 8b showing time series area-averages of ensemble corrected reanalyses and the mean LSms. Comparing Figure 8a to Figure 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 (Figure 8b) correlate well (cor = 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 (Battisti and Bretherton, 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* (cor = 0.64 and cor = 0.52, respectively). The agreement between the three data sets thus confirms the significant role that interannual SST variations play in land / ocean moisture exchange.

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 (Figure 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 to 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 (Figure 10). There is excellent agreement in locations of dense station sampling and significant rainfall but there is a strong resemblance between Figure 10a and the S/N estimates of Figure 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 Figure 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 Figures 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 Nino 3.4 SST anomaly maxima during boreal winters 1982/1983, 1986/1987, 1991/1992, 1994/1995, 1997/1998, 2004/2005 and 2009/2010 (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. Though 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. Though 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, though 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 due to 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 mmd⁻¹decade⁻¹) is found, but is only significant at 90% confidence.

ii. Corrections to VMFC (P-ET) using RPCA with pre-filtering to identify the non-physical 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 cor=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 (cor = 0.86) and composite El Niño P-ET and VMFC patterns (cor = 0.84). Though the adjustments are not needed to detect these interannual signals (Figure 1a), the agreement between VMFC* and P-ET interannual variability is more evident in time series plots (Figure 8b).

iv. Despite uncertainties inherent with observationally-constrained LSMs, these syntheses can help identify and corroborate more 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. Though 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 $W_a$, P and ET are one direct means. 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. Though retrieval physics errors
contribute to these problems the inter-calibration 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. From the terrestrial side, eq. (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 with RO measurements from river and streamflow gauges, recover a P-ET estimate. Though
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 multi-decadal
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; 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 re-evaluation 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 Studies (NEWS)
Program, Dr. Jared Entin, Program Manager. The authors wish to acknowledge the institutions
and individuals involved in development and production of the many data sets 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, and contributed data sets used in this investigation:
Drs. Steven Sitch and Sam Levis for access to the CLM4C data under the auspices of the
TRENDY initiative, 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.cgd.ucar.edu/cas/catalog/newbudgets/index.html). We thank Dr.
Aiguo Dai for providing the updated global river runoff time series.


______, 2013: The influence of the inter-decadal Pacific oscillation on US precipitation during...


Harding, R., M. Best, E. Blyth, S. Hagemann, P. Kabat, L. M. Tallaksen, T. Warnaars, D. Wiberg,
knowledge of the terrestrial global water cycle. *J. Hydrometeor.*, 12, 6, pp. 1149–1156.


Hendon, H.H., 2003. Indonesian rainfall variability: Impacts of ENSO and local air-
sea interaction. *J. Climate*, 16(11), 1775-1790.


Jung, M., M. Reichstein, and A. Bondeau, 2009: Towards global empirical upscaling of
FLUXNET eddy covariance observations: Validation of a model tree ensemble approach
using a biosphere model. *Biogeosci. Discuss.*, 6, 5271–5304, doi:10.5194/bg-6-5271-
2009.

Jung, M., and Coauthors, 2010: Recent decline in the global land evapotranspiration trend due to

White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo,

the GSI into the NCEP Global Data Assimilation System. *Wea. and Forecasting*, 24, 1691–


Taylor, K. E., R. J. Stouffer, G. A. Meehl, 2012: An Overview of CMIP5 and the experiment


Washington, R., R. James, H. Pearce, W. M. Pokam and W. Moufouma-Okia, 2013: Congo Basin rainfall climatology: can we believe the climate models? Phil. Trans. Roy. Soc., B, 368(1625 (20120296)): 1-7


Table 1. Trend statistics (mmd\(^{-1}\)decade\(^{-1}\)) for VMFC* over land for various reanalyses and P-ET for LSM members over the period 1979 to 2012. In parentheses are errors calculated using lag-one statistics to account for serial autocorrelation.

<table>
<thead>
<tr>
<th>Land Surface Model P-ET Trend</th>
<th>MERRA-Land</th>
<th>MERRA-2</th>
<th>ERA-I Land</th>
<th>GLDAS NOAH</th>
<th>ORCHIDEE</th>
<th>CLM4C</th>
<th>MPI-BGC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.023 (+/-0.019)</td>
<td>-0.002 (+/- 0.019)</td>
<td>-0.001 (+/- 0.017)</td>
<td>0.022 (+/-0.012)</td>
<td>0.009 (+/- 0.017)</td>
<td>0.023 (+/-0.015)</td>
<td>0.028 (+/-0.022)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reanalysis VMFC Trend</th>
<th>MERRA</th>
<th>MERRA-2</th>
<th>ERA-I</th>
<th>JRA-55</th>
<th>CFSR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.073 (+/- 0.026)</td>
<td>0.003 (+/- 0.026)</td>
<td>0.081 (+/- 0.020)</td>
<td>-0.030 (+/- 0.018)</td>
<td>0.074 (+/- 0.025)</td>
</tr>
</tbody>
</table>
Table 2. Summary characteristics of reanalysis data sets used in this study.

<table>
<thead>
<tr>
<th>Attributes</th>
<th>References</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERA-I</td>
<td>Cy31r2, 2006 IFS 80 km (T255 spectral) grid with 60 vertical levels. Jan1979-Dec2012.</td>
<td>Dee et al. 2011; Dee and Uppala 2008; 2009; Simmons et al, 2010 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, O3).</td>
</tr>
<tr>
<td>CFSR</td>
<td>GFS 2009 T382L64 coupled atmosphere-ocean-land surface- sea ice system, Jan1979-Dec2009.</td>
<td>Saha et al. 2010; Wang et al, 2010; Trenberth and Fasullo, 2013 3DVar GSI system, AER Radiation, Noah LSM with MOM ocean model Ocean is 0.25° at the equator, extending to a global 0.5° beyond the tropics, with 40 levels.</td>
</tr>
</tbody>
</table>
### Table 3. Observationally constrained Land Surface Models providing P and ET in this study.

<table>
<thead>
<tr>
<th>Model</th>
<th>Attributes</th>
<th>Forcing</th>
<th>Comments and references</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CLM4C</strong></td>
<td>Five primary sub-grid land cover types (glacier, lake, wetland, urban, 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 Climate Research Unit (CRU) observed monthly 0.5 analysis (v3.0, 1901–2009; New et al, 2000) and the high temporal fidelity NCEP reanalysis forcing.</td>
<td>Land model for the National Center for Atmospheric Research (NCAR) Community Earth System Model and the Community Atmosphere Model. (Oleson et al, 2010; Lawrence et al, 2011; Weedon et al, 2011). <a href="http://www.esm.ucar.edu/models/clm/">http://www.esm.ucar.edu/models/clm/</a></td>
</tr>
<tr>
<td><strong>ERA-I Land</strong></td>
<td>Hydrology-Tiled ECMWF Scheme for Surface Exchanges over Land (HTESTSEL). 80 km res with 3h integration steps.</td>
<td>ERA-Interim near surface meteorology and radiation. ERA-I precipitation is rescaled using GPCP v2.1.</td>
<td>Updated physics including soil hydrology, a new snow scheme, multi-year satellite based vegetation climatology. (Balsamo et al, 2012, 2015). <a href="http://www.ecmwf.int/en/research/climate-reanalysis/era-interim/land">http://www.ecmwf.int/en/research/climate-reanalysis/era-interim/land</a></td>
</tr>
<tr>
<td><strong>GLDAS-2 Noah</strong></td>
<td>1-D 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 precip, SRB Radiation. CRU meteorological data used to correct biases.</td>
<td>Development began 1993 through a collaboration of investigators from public and private institutions, spearheaded by the National Centers for Environmental Prediction. (Chen et al, 1996; Koren et al. 1999; Rodell, 2004). <a href="http://ldas.gsfc.nasa.gov/gldas">http://ldas.gsfc.nasa.gov/gldas</a></td>
</tr>
<tr>
<td><strong>MERRA-2</strong></td>
<td>Updated version of catchment model used in MERRA.</td>
<td>Precipitation constraints comprised of anomalies from CMAP V0011 and RT pentad product plus GPCP v2.1 climatology. Near surface meteorology and radiation from MERRA-2.</td>
<td>MERRA 0.5 deg, hourly time series of precipitation (background) are constrained to have the same daily totals as constraining CMAP/GPCP data. (Reichle and Liu, 2014). <a href="http://gmao.gsfc.nasa.gov/reanalysis/MERRA-2">http://gmao.gsfc.nasa.gov/reanalysis/MERRA-2</a></td>
</tr>
<tr>
<td><strong>MPI-BGC</strong></td>
<td>Machine-learning methodology. “model tree ensembles”, to upscale eddy covariance (EC) measurements from FLUXNET (Baldocchi et al. 2001) to a 0.5 degree monthly product</td>
<td>AVHRR NDVI data SeaWiFS fPAR; CRU near-sfc temperature</td>
<td>ET estimates use GPCCV6 precip in classification step (Jung et al., 2009, 2010). We thus use GPCCV6 precip to make consistent P-ET. <a href="https://www.bgc-jena.mpg.de/bgi/index.php">https://www.bgc-jena.mpg.de/bgi/index.php</a></td>
</tr>
<tr>
<td><strong>ORCHIDEE</strong></td>
<td>Solves water-energy-carbon budget. Represents ecosystem in terms of a range of Plant Functional Types using big leaf approach. Computes its own phenology.</td>
<td>ERA-I fields constrained with CRU 3.2 Sfc air temp and GPCCV6 monthly precip Forcing precipitation, air temperature, wind, solar radiation, humidity and atmospheric CO2</td>
<td>SECHIBA land-surface scheme, which is dedicated to the surface energy and water balances, and the carbon and vegetation model STOMATE. Kriner et al, 2005; Weedon et al. (2012); Sitch et al, 2013; Poulter (personal communication). <a href="http://forge.ipsl.jussieu.fr/orchidee/wiki">http://forge.ipsl.jussieu.fr/orchidee/wiki</a></td>
</tr>
</tbody>
</table>
Figure Captions

Figure 1. (a) Time series of global land area-average (60°N/S) vertically-integrated moisture convergence anomalies (VMFC) from various reanalyses. Units are mm d⁻¹ (left hand side scale). Niño 3.4 SST anomalies (°C) are shaded gray with inverted scale on the right hand side. (b) Same as above but for P-ET from individual LSMs. Ensemble mean is shown in black. Units are mm d⁻¹. For both time series a three month running smoother has been applied for display purposes.

Figure 2. Annual mean anomalies of LSM ensemble mean P-ET and globally-integrated streamflow from Dai (2016). Units are mm yr⁻¹. LSM monthly anomalies from Figure 1 have been summed over water year intervals beginning in October 1979.

Figure 3. Statistics for monthly mean VMFC anomalies for various reanalyses over the period 1979-2012. Left (a-e): RMS of deviations (mm d⁻¹). Right (f-j): Trends (mm d⁻¹ decade⁻¹).

Figure 4. Ensemble mean statistics for reanalysis and LSM monthly anomalies over the period 1979-2012. Left panel: RMS (mm d⁻¹). Right Panel: Trends (mm d⁻¹ decade⁻¹).

Figure 5. Time series of reanalyses VMFC and ensemble LSM anomalies (mm d⁻¹) over (a) Equatorial Africa (c) Coastal Colombia / Ecuador region, and (d) Central U.S. (c) Shows MERRA VMFC (red) and first two global PCs of vertically-integrated moisture increment.

Figure 6. (a) Ensemble mean climatological P-ET (mm d⁻¹). (b) S / N for LSMs (see text for details on calculation). (c) P-ET ACC time series of each LSM with the ensemble mean.
Figure 7. (Left) Time series of globally averaged, 60° N/S, reanalysis VMFC (black) and area-averaged corrections (red). Niño 3.4 SST x 0.1 is plotted as gray shading with inverted scale on right hand side. (Right) Trends (mmd⁻¹ decade⁻¹) in VMFC* over the period 1979-2012 for various reanalyses after corrections have been applied. Compare to Figure 3, right which are the uncorrected trends. (Bottom) Ensemble mean corrected VMFC* trends.

Figure 8. (a) Time series of individual corrected reanalysis VMFC* global land area-average (60°N/S). (b) P-ET from individual LSMs (black) and mean VMFC* from corrected reanalyses (red) and AMIPs (cyan). A three-month running smoothing is applied. Units are mmd⁻¹. Niño 3.4 SST anomalies are plotted in gray shading with inverted scale (deg C) on right.

Figure 9. Anomaly correlations, ACC, between individual corrected reanalysis VMFC* and ensemble mean LSM P-ET. ACC for the ensemble corrected and uncorrected reanalyses with the ensemble LSMs is shown by the solid (dotted) black lines, respectively. A running 3-month smoother has been applied to each time series for display.

Figure 10. Local correlations (1979-2012): (a) Ensemble mean adjusted reanalysis VMFC* and ensemble mean LSM P-ET, (b) Adjusted VMFC* correlation with LSMs minus raw reanalysis ensemble reanalyses VMFC correlation with ensemble LSM P-ET. Note different color scales.
Figure 1. (a) Time series of global land area-average (60°N/S) vertically-integrated moisture convergence anomalies (VMFC) from various reanalyses. Units are mmd$^{-1}$ (left hand side scale). Niño 3.4 SST anomalies (°C) are shaded gray with inverted scale on the right hand side. (b) Same as above but for P-ET from individual LSMs. Ensemble mean is shown in black. Units are mmd$^{-1}$. For both time series a three month running smoother has been applied for display purposes.
Figure 2. Annual mean anomalies of LSM ensemble mean P-ET and globally-integrated streamflow from Dai (2016). Units are mm yr$^{-1}$. LSM monthly anomalies from Figure 1 have been summed over water year intervals beginning in October 1979.
Figure 3. Statistics for monthly mean VMFC anomalies for various reanalyses over the period 1979-2012. Left (a-e): RMS of deviations (mmd⁻¹). Right (f-j): Trends (mmd⁻¹decade⁻¹).
Figure 4. Ensemble mean statistics for reanalysis and LSM monthly anomalies over the period 1979-2012. Left panel: RMS (mmd\(^{-1}\)). Right Panel: Trends (mmd\(^{-1}\)decade\(^{-1}\)).
Figure 5. Time series of reanalyses VMFC and ensemble LSM anomalies (mm d$^{-1}$) over (a) Equatorial Africa (c) Coastal Colombia / Ecuador region, and (d) Central U.S. (b) Shows MERRA VMFC (red) and first two global PCs of vertically-integrated moisture increment.
Figure 6. (a) Ensemble mean climatological P-ET (mmd⁻¹). (b) S/N for LSMs (see text for details on calculation). (c) P-ET ACC time series of each LSM with the ensemble mean.
Figure 7. (Left) Time series of globally averaged, 60° N/S, reanalysis VMFC (black) and area-averaged corrections (red). Niño 3.4 SST x 0.1 is plotted as gray shading with inverted scale on right hand side. (Right) Trends (mm d⁻¹ decade⁻¹) in VMFC* over the period 1979-2012 for various reanalyses after corrections have been applied. Compare to Figure 3, right which are the uncorrected trends. (Bottom) Ensemble mean corrected VMFC* trends.
Figure 8. (a) Time series of individual corrected reanalysis VMFC$^*$ global land area-average ($60^\circ$N/S). (b) P-ET from individual LSMs (black) and mean VMFC$^*$ from corrected reanalyses (red) and AMIPs (cyan). A three-month running smoothing is applied. Units are mmd$^{-1}$. Niño 3.4 SST anomalies are plotted in gray shading with inverted scale (deg C) on right.
Figure 9. Anomaly correlations, ACC, between individual corrected reanalysis VMFC* and ensemble mean LSM P-ET. ACC for the ensemble corrected and uncorrected reanalyses with the ensemble LSMs is shown by the solid (dotted) black lines, respectively. A running 3-month smoother has been applied to each time series for display.
Figure 10. Local correlations (1979-2012): (a) Ensemble mean adjusted reanalysis VMFC* and ensemble mean LSM P-ET, (b) Adjusted VMFC* correlation with LSMs minus raw reanalysis ensemble reanalyses VMFC correlation with ensemble LSM P-ET. Note different color scales.