On the Use of a Water Balance to Evaluate Interannual Terrestrial ET Variability

EUNJIN HAN

International Research Institute for Climate and Society, Earth Institute, Columbia University, Palisades, New York

WADE T. CROW

Hydrology and Remote Sensing Laboratory, Agricultural Research Service, USDA, Beltsville, Maryland

CHRISTOPHER R. HAIN

Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland

MARTHA C. ANDERSON

Hydrology and Remote Sensing Laboratory, Agricultural Research Service, USDA, Beltsville, Maryland

(Manuscript received 11 September 2014, in final form 30 January 2015)

ABSTRACT

Accurately measuring interannual variability in terrestrial evapotranspiration ET is a major challenge for efforts to detect trends in the terrestrial hydrologic cycle. Based on comparisons with annual values of terrestrial evapotranspiration $\overline{\text{ET}}$ derived from a terrestrial water balance analysis, past research has cast doubt on the ability of existing products to accurately capture $\overline{\text{ET}}$ variability. Using a variety of $\overline{\text{ET}}$ estimates, this analysis reexamines this conclusion and finds that estimates of $\overline{\text{ET}}$ variations obtained from a land surface model are more strongly correlated with $\overline{\text{ET}}$ independently acquired from thermal infrared remote sensing than $\overline{\text{ET}}$ derived from water balance considerations. This tendency is attributed to significant interannual variations in terrestrial water storage neglected by the water balance approach. Overall, results demonstrate the need to reassess perceptions concerning the skill of $\overline{\text{ET}}$ estimates derived from land surface models and show the value of accurate remotely sensed ET products for the validation of interannual ET.

1. Introduction

There has been a great deal of recent interest in the development of large-scale terrestrial evapotranspiration ET datasets for climate applications. These products can be derived via a range of remote sensing, modeling, and data assimilation approaches (Mueller et al. 2011). However, evaluating large-scale ET products at interannual time scales remains a major challenge (Zhang et al. 2012). The classical approach for verifying such products is comparison against a terrestrial water balance calculation. The instantaneous terrestrial water balance is typically based on equating changes in terrestrial water storage ΔTWS (mm) with the net sum of precipitation accumulation P, horizontal runoff flow Q, and evapotranspiration losses:

$$\Delta\text{TWS} = P - Q - \text{ET}. \tag{1}$$

This balance holds within any spatial control volume; however, it is commonly applied to discrete hydrologic units so that Q can be equated with observed streamflow at a basin outlet. Summing (1) over annual time periods—indicated using the overbar notation—leads to

$$\overline{\Delta\text{TWS}} = \overline{P} - \overline{Q} - \overline{\text{ET}}. \tag{2}$$

Therefore, based on (2), annual evapotranspiration $\overline{\text{ET}}$ can be estimated as

$$\overline{\text{ET}} = \overline{P} - \overline{Q} - \overline{\Delta\text{TWS}}. \tag{3}$$

At annual time scales and above, $\overline{\Delta\text{TWS}}$ is commonly assumed to be zero. Therefore, the classical water

Corresponding author address: Eunjin Han, International Research Institute for Climate and Society, Earth Institute, Columbia University, 61 Route 9W, Monell Building, Palisades, NY 10964-1000.
E-mail: eunjin@iri.columbia.edu

DOI: 10.1175/JHM-D-14-0175.1

TABLE 1. List of moderately sized, unregulated catchments used in the analysis.

Medium-scale basin No.	USGS station No.	USGS station name	Basin size (km^2)	Mean annual precipitation* ($mm\ yr^{-1}$)
1	07144780	North Fork Ninnescah River above Cheney Reservoir, KS	2049	704
2	07144200	Lower Arkansas River at Valley Center, KS	3402	777
3	07152000	Chikaskia River near Blackwell, OK	4891	825
4	07243500	Deep Fork near Beggs, OK	5210	891
5	07147800	Walnut River at Winfield, KS	4855	908
6	07177500	Bird Creek Near Sperry, OK	2360	954
7	06908000	Blackwater River at Blue Lick, MS	2924	1069
8	07196500	Illinois River near Tahlequah, OK	2492	1124
9	07019000	Meramec River near Eureka, MO	9766	1164
10	07052500	James River at Galena, MO	2568	1202
11	07186000	Spring River near Wace, MO	2980	1206
12	07056000	Buffalo River near St. Joe, AR	2148	1229
13	06933500	Gascondade River at Jerome, MO	7356	1256
14	07067000	Current River at Van Buren, MO	4351	1304
15	07068000	Current River at Doniphan, MO	5323	1309

* Note that mean annual precipitation is the temporal (2002–12) mean of spatially averaged annual NLDAS-2 precipitation.

balance WB approach for estimating $\overline{ET}$ is based on measuring $\overline{P}$ and $\overline{Q}$ and applying (3) under the assumption that $\overline{\Delta TWS} = 0$.

By comparing decadal trends in WB-based $\overline{ET}$ with trends derived from independent $\overline{ET}$ estimates derived from modeling and remote sensing, Zhang et al. (2012) emphasized the inability of many model- and remote sensing–based products to accurately capture interannual $\overline{ET}$ trends in humid climates. Likewise, Jung et al. (2010) validated $\overline{ET}$ derived via the spatial interpolation of ground flux tower observations using independent $\overline{ET}$ estimates derived from a catchment-scale water balance analysis. However, at least in very large hydrologic basins, recent studies on the Gravity Recovery and Climate Experiment (GRACE) have called into question the classical assertion that $\overline{\Delta TWS}$ can be safely neglected at annual time scales (Zeng et al. 2012, 2014).

Here, we seek to intercompare multiple large-scale $\overline{ET}$ products with the aim of developing an improved strategy for validating their interannual variability. Results are based on 1) $\overline{ET}$ from the Noah land surface model $\overline{ET}_{Noah}$, 2) $\overline{ET}$ from the remote sensing–based Atmosphere–Land Exchange Inverse (ALEXI) energy balance model $\overline{ET}_{ALEXI}$, and 3) $\overline{ET}$ from a WB approach that neglects annual changes in terrestrial water storage $\overline{ET}_{WB}$. For an additional analysis, $\overline{\Delta TWS}$ from GRACE data and $\overline{ET}$ estimates from the interpolation of ground-based flux towers using the model tree ensemble (MTE) algorithm of Jung et al. (2009) $\overline{ET}_{MTE}$ are also considered. Given that past work has already examined mutual biases in these products (Hain et al. 2015), our focus here is on the correlation (at zero lag) of interannual $\overline{ET}$ anomalies.

2. Methods

The analysis is divided into two scales. The first scale is defined by an east–west transect of 15 medium-sized ($\sim 40^2$–$100^2\ km^2$) unregulated basins within the U.S. southern Great Plains (SGP) region (Table 1). The second scale is consistent with five much larger ($\sim 500^2$–$1000^2\ km^2$) major basins within the Mississippi River system (Table 2). See Fig. 1 for a map of all basins. The 15 medium-sized basins described in Table 1 were selected

TABLE 2. List of larger river basins within the Mississippi River system used in the analysis.

Major Mississippi subbasin	USGS station No.	USGS station name	Basin size (km^2)	Mean annual precipitation* ($mm\ yr^{-1}$)
Missouri River	06934500	Missouri River at Hermann, MO	1 347 556	537
Arkansas River	07263450	Arkansas River below Little Rock, AR	409 201	699
Red River	07344370	Red River at Spring Bank, AR	153 906	809
Upper Mississippi River	07022000	Mississippi River at Thebes, IL	496 016	876
Ohio River	03611500	Ohio River at Metropolis, IL	527 557	1200

* Note that mean annual precipitation is the temporal (2002–12) mean of spatially averaged annual NLDAS-2 precipitation.

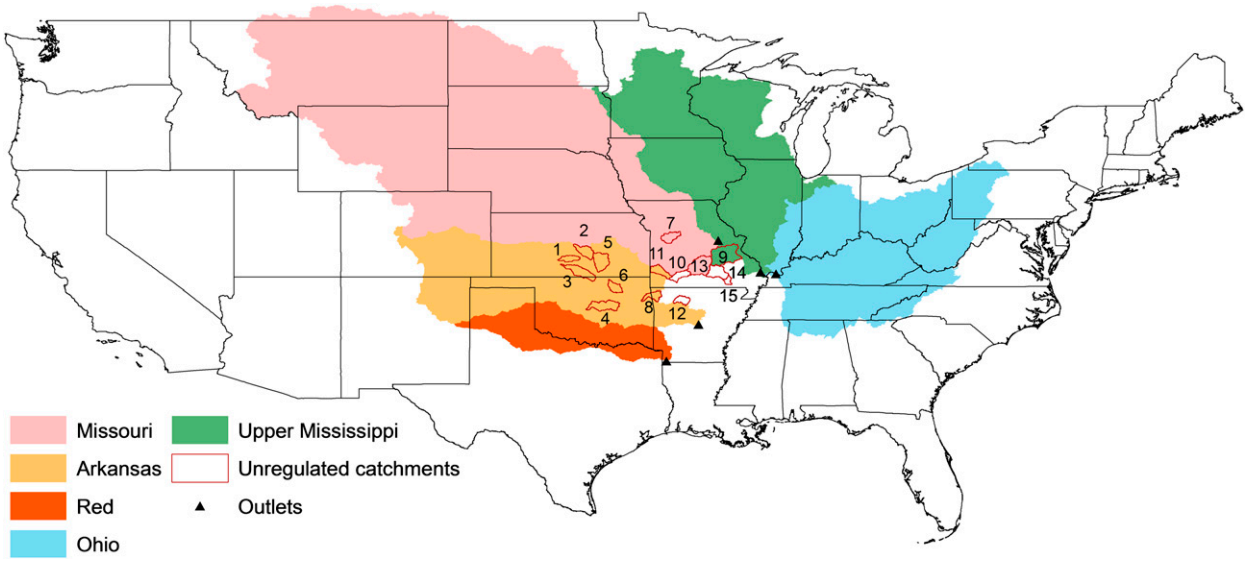

FIG. 1. Map of the 15 unregulated medium-scale basins (red outlines) and the five Mississippi River system major subbasins used in the analysis. See Tables 1 and 2 for basin details.

based on a screening analysis by the Model Parameter Estimation Experiment (MOPEX) to remove basins with poor rain gauge coverage and/or excessive human regulation/impoundment of streamflow. In addition, an attempt was made to select medium-scale basins that span the strong east–west precipitation gradient across the SGP and receive a relatively low fraction of their annual precipitation as snowfall. Naturally, the impact of human streamflow regulation cannot be neglected within the major basins examined.

3. Data

a. *Water balance*–ET

As described in (3), $\overline{\mathrm{ET}}_{\mathrm{WB}}$ estimates were derived from the difference between annual observed streamflow and precipitation where $\overline{\Delta \mathrm{TWS}}$ is assumed to be zero. In particular, daily streamflow volumes at individual basin outlets listed in Tables 1 and 2 (and mapped in Fig. 1) were obtained from the U.S. Geologic Survey (USGS), normalized by basin drainage areas, and aggregated to (calendar year) annual values. Annual precipitation was based on the temporal aggregation of terrain-corrected daily rain gauge observations collected from the National Centers for Environmental Prediction (NCEP) Climate Prediction Center (CPC) and processed onto a 0.125° grid as part of phase 2 of the North American Land Data Assimilation System (NLDAS-2). More details on the NLDAS-2 project and meteorological forcing datasets can be found in Mitchell et al. (2004) and Xia et al. (2012). Following (3), $\overline{\mathrm{ET}}_{\mathrm{WB}}$ was calculated for calendar years 2002–12.

b. *Noah*–ET

The $\overline{\mathrm{ET}}_{\mathrm{Noah}}$ product was based on the temporal aggregation of hourly ET predictions acquired from a 0.125°-resolution Noah land surface model simulation driven by NLDAS-2 meteorological forcing data. The NLDAS-2 hourly precipitation forcing dataset is based on the disaggregation of daily NCEP CPC data using available ground-based rain radar observations. The Noah model is a one-dimensional, physically based land surface model that calculates surface state and flux variables using prognostic energy and water balance equations. Total ET is calculated by summing up hourly Noah predictions of 1) direct evaporation from the surface soil, 2) direct evaporation of canopy-intercepted precipitation, 3) transpiration via plant root uptake of water, and 4) sublimation. Annual averages were then obtained by summing hourly ET within calendar years 2002–12 and spatially averaging these 0.125° summations over all basin domains indicated in Fig. 1. More information about the Noah model version implemented in NLDAS-2 (version 2.8) is given in Chen et al. (1996), Chen and Dudhia (2001), and Ek et al. (2003). Note that since $\overline{\mathrm{ET}}_{\mathrm{Noah}}$ and $\overline{\mathrm{ET}}_{\mathrm{WB}}$ are both derived (in part) from NLDAS-2 precipitation data, they cannot be considered wholly independent estimates.

c. *ALEXI*–ET

Unlike $\overline{\mathrm{ET}}_{\mathrm{Noah}}$ and $\overline{\mathrm{ET}}_{\mathrm{WB}}$, the ALEXI surface energy balance model produces ET using thermal infrared (TIR) remote sensing data without any precipitation input

(Anderson et al. 2011). ALEXI was processed at a spatial resolution of 10 km over the period of 2003–12, forced with meteorological inputs from the North American Regional Reanalysis (NARR; Mesinger et al. 2006), TIR land surface temperature from the Geostationary Operational Environmental Satellite (GOES), and leaf area index (LAI) from the 8-day *Terra* MODIS product (MOD15A2), used to estimate vegetation cover fraction f_c. Instantaneous latent heat fluxes retrieved from ALEXI are upscaled to daytime-integrated ET estimates, assuming a self-preservation of the ratio of latent heat flux and incoming shortwave radiation f_{SUN} during daytime hours (Cammalleri et al. 2014). Incoming shortwave radiation inputs are taken from the NCEP Climate Forecast System Reanalysis (CFSR; Saha et al. 2010). Currently, ALEXI is not executed over snow-covered surfaces. These periods are instead gap filled with a linear interpolation of f_{SUN}.

While based on very different fundamental principles, ALEXI and Noah share some common inputs. Therefore, in order to minimize commonality in inputs between $\overline{ET}_{Noah}$ and $\overline{ET}_{ALEXI}$, every effort was made to ensure that these inputs did not induce cross-correlated error in ET predictions. For instance, while both Noah and ALEXI require incoming solar radiation as a forcing, ALEXI simulations were based by radiation products generated by CFSR while Noah simulations were instead forced by radiation fields from the NARR. Likewise, while both ALEXI and Noah require f_c, Noah uses a fixed monthly climatology acquired from a retrospective analysis of Advanced Very High Resolution Radiometer observations while ALEXI uses actual 8-day MODIS LAI composites to estimate f_c.

d. *MTE*–ET

For an additional analysis, $\overline{ET}$ estimates were also acquired from the flux tower observations and the MTE machine-learning algorithm introduced by Jung et al. (2009). The MTE upscales in situ ET measurements from a network of regional networks (FLUXNET) using the remotely sensed fraction of photosynthetically active radiation and gridded meteorological data to produce monthly gridded ET estimates at a 0.50° spatial resolution. These estimates were temporally averaged to an annual scale (within calendar years 2002–11) and spatially averaged within the five major basins listed in Table 2.

e. *GRACE*–ΔTWS

Monthly GRACE ΔTWS data were obtained by applying the rescaling coefficients of Landerer and Swenson (2012) to gridded 0.25° GRACE ΔTWS products provided by the GeoForschungsZentrum (GFZ) and the University of Texas Center for Space Research (CSR) and averaging the resulting two fields together. December and January GRACE ΔTWS values from this unified product were averaged together to estimate 1 January ΔTWS. The difference between successive 1 January ΔTWS values was then used to obtain $\overline{\Delta TWS}$ for calendar years 2002–12. Finally, the resulting 0.25° annual $\overline{\Delta TWS}$ fields were spatially averaged within the five major river basins listed in Table 2.

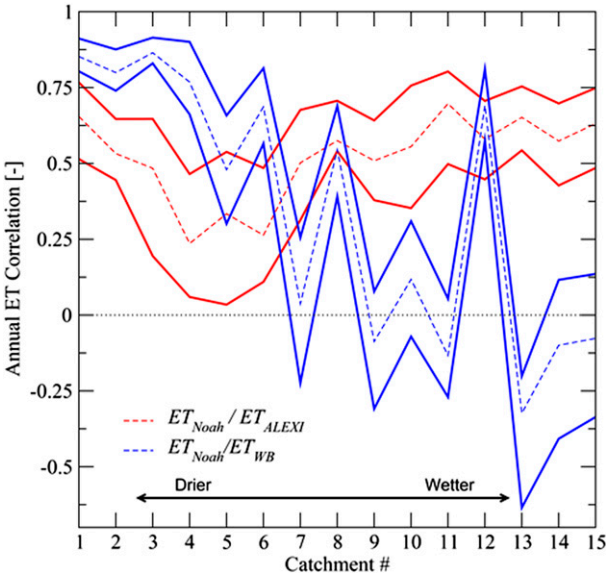

FIG. 2. The $\overline{ET}_{Noah}$–$\overline{ET}_{ALEXI}$ (red lines) and $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ (blue lines) correlations for the 15 unregulated, medium-scale basins listed in Table 1 (ordered from driest to wettest). Dashed lines indicate sampled correlations and solid lines indicate the interquartile spread of sampled correlations derived using a 5000-member bootstrapping approach.

4. Results

Our analysis focused on calculating the (lag zero) Pearson correlation coefficient between normalized anomalies of interannual, basin-scale $\overline{ET}_{Noah}$ variations and interannual variation found in other $\overline{ET}$ products. Given the west-to-east increase in P within the SGP region, mean $\overline{P}$ within the 15 medium-scale basins (Table 1, Fig. 1) ranges from 500 to 900 mm yr^{-1} (Table 1). Figure 2 plots the correlation between $\overline{ET}_{Noah}$ and $\overline{ET}_{WB}$ sampled in each medium-scale basin, where basins are sorted according to mean $\overline{P}$. Within the driest basins, $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlations are uniformly high. However, since $\overline{ET}_{Noah}$ and $\overline{ET}_{WB}$ estimates are based on the same (uncertain) precipitation product, some of this correlation may be spurious because of positively correlated errors. In contrast, $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlations become highly erratic (and frequently negligible)

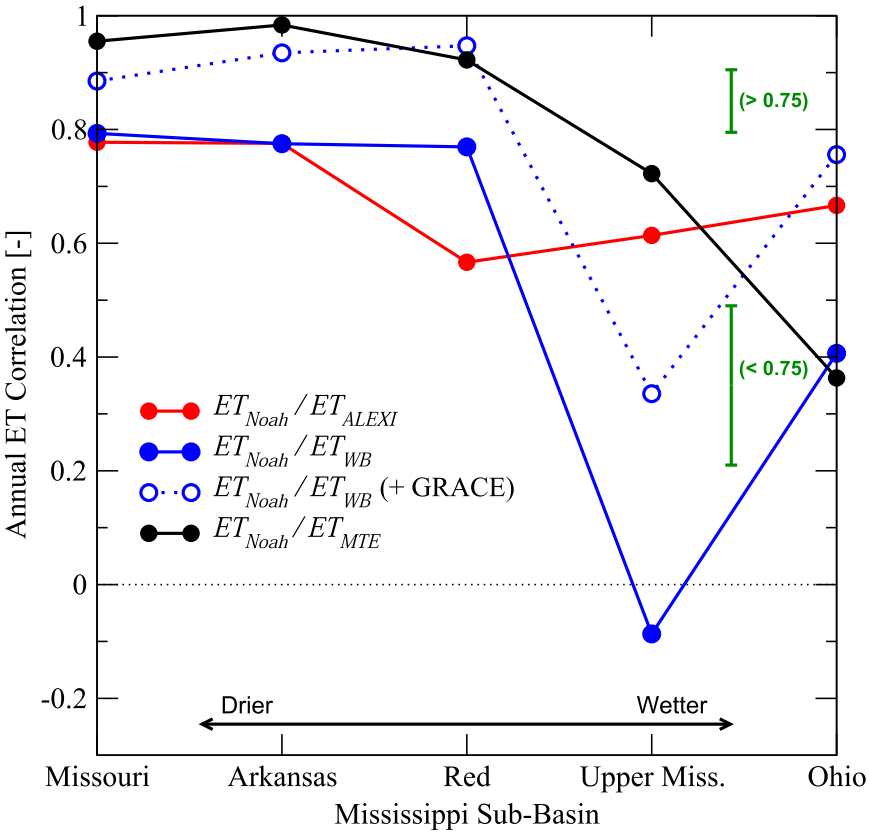

FIG. 3. The $\overline{ET}_{Noah}$–$\overline{ET}_{ALEXI}$ (red line) and $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ (solid blue line) correlations for the five major Mississippi River subbasins listed in Table 2 (ordered from driest to wettest). Also shown are correlations between $\overline{ET}_{Noah}$ and 1) GRACE-corrected $\overline{ET}_{WB}$ (dotted blue line) and 2) $\overline{ET}_{MTE}$ (black line).

within wetter basins (Fig. 2). This tendency has been attributed to the inability of land models to accurately capture interannual ET variability in humid climates (Zhang et al. 2012).

However, a different interpretation emerges when also considering the correlation between $\overline{ET}_{Noah}$ and $\overline{ET}_{ALEXI}$. In particular, $\overline{ET}_{Noah}$–$\overline{ET}_{ALEXI}$ correlations are uniformly positive for all medium-scale basins, even the wettest basins, which exhibit very low $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlations (Fig. 2). In addition, all sampled $\overline{ET}_{Noah}$–$\overline{ET}_{ALEXI}$ correlations have interquartile sampling ranges (derived from a boot-strapping approach) that do not include zero. Such robust positive correlations occur despite the fact that $\overline{ET}_{Noah}$ and $\overline{ET}_{ALEXI}$ are obtained via wholly independent means and cold season $\overline{ET}_{ALEXI}$ is based on a simplistic temporal interpolation technique (section 2c). Therefore, Fig. 2 strongly implies that the aforementioned reduction in $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlation within humid basins is attributable to error in $\overline{ET}_{WB}$ and not uncertainty in $\overline{ET}_{Noah}$.

An obvious error source for $\overline{ET}_{WB}$ is the neglect of $\overline{\Delta TWS}$. Within the larger-scale major basins listed in Table 2, the impact of $\overline{\Delta TWS}$ can be directly examined using GRACE ΔTWS observations. Figure 3 is analogous to Fig. 2, except applied to much larger basins within the Mississippi River system (Fig. 1, Table 2). As in Fig. 2, $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlations are relatively high for the drier major basins (i.e., the Missouri, the Red, and the Arkansas) but fall sharply for the wetter major basins (i.e., the Ohio and the upper Mississippi). However, $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlations are uniformly improved by avoiding the problematic assumption that $\overline{\Delta TWS} = 0$ and instead estimating large-scale $\overline{\Delta TWS}$ directly from GRACE (Fig. 3). The consistent improvement implies that the neglect of $\overline{\Delta TWS}$ is playing a significant role in reducing sampled $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlations.

As in Fig. 2, $\overline{ET}_{Noah}$–$\overline{ET}_{ALEXI}$ correlations in Fig. 3 are relatively high and remain stable across all five major basins. Sampled $\overline{ET}_{Noah}$–$\overline{ET}_{MTE}$ correlations are even higher for all basins except the Ohio River basin. In particular, note that the exceptionally low (negative) correlation in (non-GRACE corrected) $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlations within the upper Mississippi basin is not

reflected in either $\overline{ET}_{Noah}$–$\overline{ET}_{ALEXI}$, $\overline{ET}_{Noah}$–$\overline{ET}_{MTE}$, or GRACE-corrected $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlations. Therefore, the observed shortcoming in (uncorrected) $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ appears linked to relatively large interannual variability in surface water and snow storage within the upper Mississippi River basin. As a result, Fig. 3 supports Fig. 2 by suggesting that the decline in $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlations within humid basins is attributable to problems with the accuracy of the $\overline{ET}_{WB}$ benchmark and not the ability of Noah to accurately capture interannual ET variability.

The low $\overline{ET}_{Noah}$–$\overline{ET}_{MTE}$ correlation sampled within the (humid) Ohio River basin (Fig. 3) runs somewhat counter to this interpretation. However, the lack of a comparable reduction in either $\overline{ET}_{Noah}$–$\overline{ET}_{ALEXI}$ or GRACE-corrected $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlation results within the Ohio River basin implies that the reduction is due to increased error in $\overline{ET}_{MTE}$ and not $\overline{ET}_{Noah}$.

5. Discussion and conclusions

Here, we examine the correlation in interannual ET variations observed via a variety of independent means. When transitioning from a dry to a wet climate within the SGP, a large reduction is seen in the correlation between $\overline{ET}_{Noah}$ and $\overline{ET}_{WB}$. This trend holds along a transect of both medium-scale unregulated basins in the SGP (Fig. 2) and among five large-scale major Mississippi River subbasins (Fig. 3). However, an analogous reduction with wetter climate is not observed in the correlation between $\overline{ET}_{Noah}$ and $\overline{ET}_{ALEXI}$ (Figs. 2 and 3). Therefore, the reduction in the $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlation for wet climates appears to be a consequence of neglecting $\overline{\Delta TWS}$ in water balance calculations and not reflective of any shortcoming in $\overline{ET}_{Noah}$.

In addition, within the major Mississippi River system basins, the introduction of GRACE-based $\overline{\Delta TWS}$ uniformly improves the $\overline{ET}_{Noah}$–$\overline{ET}_{WB}$ correlation (Fig. 3). Therefore, taken as a whole, results imply that WB-based calculations with the neglect of interannual ΔTWS do not represent a robust benchmark for the validation of interannual ET variations in relatively humid climates. Instead, a more robust verification approach appears to be the examination of correlations between model-based predictions and independently generated ET datasets and/or the use of GRACE $\overline{\Delta TWS}$ data to refine annual water balance calculations (within sufficiently large basins). This may suggest the need to reevaluate previous work (Zhang et al. 2012) that utilized water balance approaches to conclude that model-based ET products contain little skill in capturing interannual ET variability.

The impact of $\overline{\Delta TWS}$ on (annual) $\overline{ET}_{WB}$ calculations has been previously noted (Rodell et al. 2007; Syed et al. 2008; Zeng et al. 2012); however, this analysis leads to several novel insights. First, by using $\overline{ET}$ correlations (and not GRACE observations) to infer the presence of significant $\overline{\Delta TWS}$ variations, these results provide an independent source of verification for earlier studies based only on GRACE $\overline{\Delta TWS}$ retrievals. One consequence of this is our ability to extend the observation-based analysis of $\overline{\Delta TWS}$ down to small-scale catchments (1000–8000 km^2 in size) that cannot be resolved by GRACE (Table 1, Fig. 2). Despite these small-scale basins being free of any major anthropogenic impoundment (and generally clear of major snowpack storage), $\overline{\Delta TWS}$ still appears to play a major role in any attempt to estimate $\overline{ET}$ via water balance considerations and the neglect of terrestrial water storage variations. In addition, results suggest a relatively larger impact for $\overline{\Delta TWS}$ on $\overline{ET}_{WB}$ variations within relatively wet climates. This tendency is at odds with earlier GRACE-based studies (which suggested greater impacts in arid climates; Zeng et al. 2012) and provides an alternative explanation for the conclusion of Zhang et al. (2012) that land surface models cannot match $\overline{ET}_{WB}$ trends within relatively wet climates.

Nevertheless, several important caveats should be considered. For example, Hain et al. (2015) identified large relative biases in $\overline{ET}_{Noah}$ for areas with extensive irrigation and/or direct groundwater extraction by plant roots. In such areas, correlations between $\overline{ET}_{Noah}$ and $\overline{ET}_{ALEXI}$ may be degraded and consequently unsuitable as a verification tool. Finally, all results are based on relatively short (9 or 10 years) data records because of limitations in the length of available satellite data records. Care should therefore be taken to avoid the overinterpretation of small—and potentially nonsignificant—differences in correlations. Finally, a number of obvious follow-on research topics can be defined based on initial results presented here. Such topics include 1) examining the impact of snow water storage on $\overline{\Delta TWS}$ by modifying the start/stop times used to define an annual average, 2) replicating the analysis for multiple land surface models, and 3) evaluating water balance calculations on subannual time scales.

Acknowledgments. Research was funded by a grant to Wade T. Crow (PI) from the Water Resources Group within the NASA Applied Sciences Program.

REFERENCES

Anderson, M. C., and Coauthors, 2011: Mapping daily evapotranspiration at field to continental scales using geostationary and polar orbiting satellite imagery. *Hydrol. Earth Syst. Sci.,* **15,** 223–239, doi:10.5194/hess-15-223-2011.

Cammalleri, C., M. Anderson, and W. Kustas, 2014: Upscaling of evapotranspiration fluxes from instantaneous to daytime

scales for thermal remote sensing applications. *Hydrol. Earth Syst. Sci.,* **18,** 1885–1894, doi:10.5194/hess-18-1885-2014.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.,* **129,** 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

——, and Coauthors, 1996: Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *J. Geophys. Res.,* **101,** 7251–7268, doi:10.1029/95JD02165.

Ek, M. B., and Coauthors, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.,* **108,** 8851, doi:10.1029/2002JD003296.

Hain, C. R., W. T. Crow, M. C. Anderson, and M. T. Yilmaz, 2015: Diagnosing neglected soil moisture source/sink processes via a thermal infrared-based two-source energy balance model. *J. Hydrometeor.,* doi:10.1175/JHM-D-14-0017.1, in press.

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. *Biogeosciences,* **6,** 2001–2013, doi:10.5194/bg-6-2001-2009.

——, and Coauthors, 2010: Recent decline in the global land evapotranspiration trend due to limited moisture supply. *Nature,* **467,** 951–954, doi:10.1038/nature09396.

Landerer, F., and S. Swenson, 2012: Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resour. Res.,* **48,** W04531, doi:10.1029/2011WR011453.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.,* **87,** 343–360, doi:10.1175/BAMS-87-3-343.

Mitchell, K. E., and Coauthors, 2004: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J. Geophys. Res.,* **109,** D07S90, doi:10.1029/2003JD003823.

Mueller, B., and Coauthors, 2011: Evaluation of global observations-based evapotranspiration datasets and IPCC AR4 simulations. *Geophys. Res. Lett.,* **38,** L06402, doi:10.1029/2010GL046230.

Rodell, M., J. Chen, H. Kato, J. S. Famiglietti, J. Nigro, and C. R. Wilson, 2007: Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeol. J.,* **15,** 159–166, doi:10.1007/s10040-006-0103-7.

Saha, S., and Coauthors, 2010: The NCEP Climate Forecast System Reanalysis. *Bull. Amer. Meteor. Soc.,* **91,** 1015–1057, doi:10.1175/2010BAMS3001.1.

Syed, T. H., J. S. Famiglietti, M. Rodell, J. Chen, and C. R. Wilson, 2008: Analysis of terrestrial water storage changes from GRACE and GLDAS. *Water Resour. Res.,* **44,** W02433, doi:10.1029/2006WR005779.

Xia, Y., and Coauthors, 2012: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.,* **117,** D03109, doi:10.1029/2011JD016048.

Zeng, Z., S. Piao, X. Lin, G. Yin, S. Peng, P. Ciais, and R. B. Myneni, 2012: Global evapotranspiration over the past three decades: Estimation based on the water balance equation combined with empirical models. *Environ. Res. Lett.,* **7,** 014026, doi:10.1088/1748-9326/7/1/014026.

——, T. Wang, F. Zhou, P. Ciais, J. Mao, X. Shi, and S. Piao, 2014: A worldwide analysis of spatiotemporal changes in water balance–based evapotranspiration from 1982 to 2009. *J. Geophys. Res. Atmos.,* **119,** 1186–1202, doi:10.1002/2013JD020941.

Zhang, Y., and Coauthors, 2012: Decadal trends in evaporation from global energy and water balances. *J. Hydrometeor.,* **13,** 379–391, doi:10.1175/JHM-D-11-012.1.