2-2015

Distributed Hydrologic Modeling Using Satellite-Derived Potential Evapotranspiration

Ryan Randall Spies
*Iowa State University*, ryspi@iastate.edu

Kristie J. Franz
*Iowa State University*, kfranz@iastate.edu

Terri S. Hogue
*Colorado School of Mines*

Angela L. Bowman
*Iowa State University*, albowman@iastate.edu

Follow this and additional works at: [http://lib.dr.iastate.edu/ge_at_pubs](http://lib.dr.iastate.edu/ge_at_pubs)

Part of the *Atmospheric Sciences Commons*, *Civil and Environmental Engineering Commons*, and the *Geology Commons*

The complete bibliographic information for this item can be found at [http://lib.dr.iastate.edu/ge_at_pubs/73](http://lib.dr.iastate.edu/ge_at_pubs/73). For information on how to cite this item, please visit [http://lib.dr.iastate.edu/howtocite.html](http://lib.dr.iastate.edu/howtocite.html).

This Article is brought to you for free and open access by the Geological and Atmospheric Sciences at Iowa State University Digital Repository. It has been accepted for inclusion in Geological and Atmospheric Sciences Publications by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
Distributed Hydrologic Modeling Using Satellite-Derived Potential Evapotranspiration

Abstract
Satellite-derived potential evapotranspiration (PET) estimates computed from Moderate Resolution Imaging Spectroradiometer (MODIS) observations and the Priestley-Taylor formula (M-PET) are evaluated as input to the Hydrology Laboratory Research Distributed Hydrologic Model (HL-RDHM). The HL-RDHM is run at a 4-km spatial and 6-h temporal resolution for 13 watersheds in the upper Mississippi and Red River basins for 2003-10. Simulated discharge using inputs of daily M-PET is evaluated for all watersheds, and simulated evapotranspiration (ET) is evaluated at two watersheds using nearby latent heat flux observations. M-PET-derived model simulations are compared to output using the long-term average PET values (default-PET) provided as part of the HL-RDHM application. In addition, uncalibrated and calibrated simulations are evaluated for both PET data sources. Calibrating select model parameters is found to substantially improve simulated discharge for both datasets. Overall average percent bias (PBias) and Nash-Sutcliffe efficiency (NSE) values for simulated discharge are better from the default-PET than the M-PET for the calibrated models during the verification period, indicating that the time-varying M-PET input did not improve the discharge simulation in the HL-RDHM. M-PET tends to produce higher NSE values than the default-PET for the Wisconsin and Minnesota basins, but lower NSE values for the Iowa basins. M-PET-simulated ET matches the range and variability of observed ET better than the default-PET at two sites studied and may provide potential model improvements in that regard.

Keywords
Evapotranspiration, forecasting, hydrologic models, hydrology, satellite observations

Disciplines
Atmospheric Sciences | Civil and Environmental Engineering | Geology

Comments
This article is from Journal of Hydrometeorology 16 (2015): 129, doi: 10.1175/JHM-D-14-0047.1. Posted with permission.

Rights
© Copyright 2015 American Meteorological Society (AMS). Permission to use figures, tables, and brief excerpts from this work in scientific and educational works is hereby granted provided that the source is acknowledged. Any use of material in this work that is determined to be “fair use” under Section 107 of the U.S. Copyright Act September 2010 Page 2 or that satisfies the conditions specified in Section 108 of the U.S. Copyright Act (17 USC §108, as revised by P.L. 94-553) does not require the AMS’s permission. Republication, systematic reproduction, posting in electronic form, such as on a web site or in a searchable database, or other uses of this material, except as exempted by the above statement, requires written permission or a license from the AMS. Additional details are provided in the AMS Copyright Policy, available on the AMS Web site located at (http://www.ametsoc.org/) or from the AMS at 617-227-2425 or copyrights@ametsoc.org.

This article is available at Iowa State University Digital Repository: http://lib.dr.iastate.edu/ge_at_pubs/73
Distributed Hydrologic Modeling Using Satellite-Derived Potential Evapotranspiration

RYAN R. SPIES AND KRISTIE J. FRANZ
Department of Geological and Atmospheric Sciences, Iowa State University, Ames, Iowa

TERRI S. HOGUE
Department of Civil and Environmental Engineering, Colorado School of Mines, Golden, Colorado

ANGELA L. BOWMAN
Department of Geological and Atmospheric Sciences, Iowa State University, Ames, Iowa

(Manuscript received 27 February 2014, in final form 13 August 2014)

ABSTRACT

Satellite-derived potential evapotranspiration (PET) estimates computed from Moderate Resolution Imaging Spectroradiometer (MODIS) observations and the Priestley–Taylor formula (M-PET) are evaluated as input to the Hydrology Laboratory Research Distributed Hydrologic Model (HL-RDHM). The HL-RDHM is run at a 4-km spatial and 6-h temporal resolution for 13 watersheds in the upper Mississippi and Red River basins for 2003–10. Simulated discharge using inputs of daily M-PET is evaluated for all watersheds, and simulated evapotranspiration (ET) is evaluated at two watersheds using nearby latent heat flux observations. M-PET–derived model simulations are compared to output using the long-term average PET values (default-PET) provided as part of the HL-RDHM application. In addition, uncalibrated and calibrated simulations are evaluated for both PET data sources. Calibrating select model parameters is found to substantially improve simulated discharge for both datasets. Overall average percent bias (PBias) and Nash–Sutcliffe efficiency (NSE) values for simulated discharge are better from the default-PET than the M-PET for the calibrated models during the verification period, indicating that the time-varying M-PET input did not improve the discharge simulation in the HL-RDHM. M-PET tends to produce higher NSE values than the default-PET for the Wisconsin and Minnesota basins, but lower NSE values for the Iowa basins. M-PET–simulated ET matches the range and variability of observed ET better than the default-PET at two sites studied and may provide potential model improvements in that regard.

1. Introduction

Remotely sensed satellite data streams have seen a significant rise in hydrologic modeling applications. The use of satellite-based data within data assimilation schemes is one of the most common applications, with snow data assimilation being among the most well studied (e.g., Andreadis and Lettenmaier 2006; Slater and Clark 2006; Tang and Lettenmaier 2010; Thirel et al. 2013; Liu et al. 2013). Exploration of satellite-based data for model calibration (Parajka and Blöschl 2008; Šorman et al. 2009; Franz and Karsten 2013) and as model input (Jacobs et al. 2009; Behrangi et al. 2011; Gokmen et al. 2013; Xue et al. 2013) has been more limited. In particular, there remains a need to identify high spatial and temporal resolution datasets for driving and validating spatially distributed hydrologic models.

Research has shown that distributed models have the potential to perform as well as, or better than, well-calibrated lumped models (Koren et al. 2004; Smith et al. 2012a), but expanding the use of these models requires identifying sources of reliable and robust input data (Reed et al. 2004; Koren et al. 2004; Smith et al. 2004; Smith et al. 2012a; Smith et al. 2012b). The National Weather Service (NWS) Office of Hydrologic Development (OHD) developed the Hydrology Laboratory Research Distributed Hydrologic Model (HL-RDHM)
as a tool to investigate the use of distributed models for hydrologic forecasting and to support related research, such as the application of gridded model inputs (Koren et al. 2004; NWS 2011).

The HL-RDHM is grid based and employs a conceptual rainfall–runoff model to perform the water balance functions for each grid, including meeting potential evapotranspiration (PET) demands (NWS 2011). The PET data implemented in the HL-RDHM are based on climatological potential evaporation (PE) values estimated from the seasonal and annual free water surface maps and mean monthly station data from National Oceanic and Atmospheric Administration (NOAA) Technical Reports 33 and 34. PE adjustment factors are used within the model to account for the seasonal variation in vegetation throughout the year (NWS 2008). The PE data and adjustment factors (hereinafter referred to as default-PET) are monthly gridded values from which daily values are interpolated. These data are static with no annual variation. As a result, the influence of daily weather, interannual climatic variability, and land cover change on PET is not reflected in the current simulation structure.

The use of regionalized climatological PET is common in hydrology because the sporadic coverage of meteorological stations limits the ability to estimate daily PET with ground-based measurements (Farnsworth and Thompson 1982). For this reason, researchers have looked to satellite data as a means to better estimate the spatial and temporal PET variability within watersheds (e.g., Kim and Hogue 2008; Jacobs et al. 2009). Satellite-based PET data have the ability to capture changes in the hydrologic cycle, such as those due to climate (Pechlivanidis et al. 2011; Irmak et al. 2012) or land cover change, potentially reducing model uncertainty associated with using long-term average PET input. Although models are able to adapt to biases in PET input data through the calibration process, model results are sensitive to PET inputs (Andréassian et al. 2004).

Several spatial PET datasets are available, the majority of which combine satellite remote sensing and ground-based meteorological observations. Examples include the North American Land Data Assimilation System, phase 2 (NLDAS-2), data forcing and a semiclimirical Moderate Resolution Imaging Spectroradiometer (MODIS) product (Jin et al. 2011). The MOD16 product is a land surface evapotranspiration (ET) product that also uses ground-based meteorological observations to estimate 1-km, 8-day ET. The focus of the current study is the Kim and Hogue (2008, 2013) method that uses nine products from the MODIS Terra and Aqua platforms and the Priestley–Taylor formula to estimate a daily PET (M-PET). M-PET differs from the other datasets in that it is generally a “stand alone” satellite-based product and does not incorporate a physical model or ground-based observational data. Kim and Hogue (2008) evaluated M-PET at the point scale against four flux-tower sites in the contiguous United States (CONUS) and found good correlation between variables used to derive a daily PET. Net radiation, the primary variable in the formulation, has shown relative low errors in several studies, with biases of less than 4% when compared with ground-based estimates (Kim and Hogue 2008, 2013).

Most model sensitivity studies have focused on precipitation inputs, while few have investigated the sensitivity of rainfall–runoff models to PET inputs (Andréassian et al. 2004; Oudin et al. 2005; Jacobs et al. 2009). Through a series of studies, we are evaluating the satellite-derived M-PET as potential input to operational hydrologic prediction models, including testing M-PET as input to the lumped NWS modeling framework (Bowman et al. 2013, manuscript submitted to J. Hydrol. Eng.; Barik 2014) and incorporating satellite data with higher temporal resolution to improve the PET estimate (Barik 2014). The primary objective of the current study is to test the M-PET data product as a potential forcing for the HL-RDHM. M-PET is used to model watersheds in the north-central United States for the evaluation period of 2003–10 and is compared to application of the NWS default-PET. The HL-RDHM is calibrated for each PET dataset and the model calibration and simulations of discharge and evapotranspiration are evaluated. The impact of a time-varying PET input on the performance of the rainfall–runoff model, which has traditionally been applied and developed using climatological estimates of PET, is explored. We test the simple assumption that data with higher spatial and temporal variability will lead to improved hydrologic modeling.

2. Methodology

a. Study area

The study area includes 13 headwater basins distributed throughout the states of Iowa, Minnesota, and Wisconsin, in the north-central United States (Fig. 1). All basins are within the upper Mississippi and Red River of the north basins and are established forecast points of the NWS North Central River Forecast Center (NCRFC). Study basins were chosen in consultation with NCRFC forecasters to represent a range of watershed sizes (from 530 to 6242 km²; Table 1), land use, and forecasting conditions. Daily discharge observations at the outlet of each basin are collected by the U.S. Geological Survey (USGS).

The study region is characterized by minimal topographic relief. All study basins have less than 200 m of
elevation change and the average slope is less than 4% in the study area, except for a small region of southwestern Wisconsin, where slope increases to 6% (Fig. 1). Land use in southern Minnesota, southern Wisconsin, and Iowa is largely agricultural, with extensive areas of corn, soybeans, and pasture. Basins in northern Minnesota and northern Wisconsin contain large forested and wetland regions (Fry et al. 2011). The abundant precipitation during the growing season negates the need for continual irrigation in much of the area, with some irrigation occurring mostly in Minnesota and Wisconsin. Several basins in Iowa and southern Minnesota have artificial subsurface drainage networks that promote agricultural activity by lowering the water table. Artificial drainage has resulted in observed changes in basin hydrology, such as increased base flow (Schilling and Libra 2004; Schilling and Helmers 2008).

The climate of the north-central United States has significant seasonal variability. Summers are characterized by warm, humid conditions and winters are cool and dry. Convective thunderstorm systems, enhanced by moisture transport from the Gulf of Mexico, produce a large percentage of the annual precipitation from early spring into late summer.

b. HL-RDHM

The HL-RDHM, version 3.2.0, is used in the current study. The framework includes the SNOW-17 model (Anderson 1973), the Sacramento Soil Moisture Accounting Heat Transfer model (SAC-HT; Koren et al. 2007, 2014), and the physically based kinematic hillslope and channel routing model. SNOW-17 is an empirically based snow accumulation and ablation model that uses air temperature as an index to the energy exchanges between the snowpack and the atmosphere, snow accumulation, and snowmelt. SAC-HT is the conceptual rainfall–runoff Sacramento Soil Moisture Accounting model (SAC-SMA; Burnash et al. 1973) incorporated with a physically based frozen ground model. In the SAC-SMA, flow of water through the subsurface occurs through two soil zones, a faster upper zone and slower lower zone. Each zone has free water storages that represent water that is drained by gravitational forces and tension water storages that represent water that can only be removed by evaporation and transpiration. Surface runoff occurs when upper-zone storages become full. ET occurs in five regions of the model: upper-zone tension water, upper-zone free water, lower-zone tension water, impervious areas, and channel inflows. Total ET is a function of the available moisture within each zone and the PET input at each time step, which cannot be exceeded.

Routing of surface and subsurface runoff within grid cells occurs over conceptual hillslopes and channels using drainage density, surface slope, and hillslope roughness properties (Koren et al. 2004). Within a basin, water

![Fig. 1. Map of study basins and location of the Brookings and Brooks Field Site 11–Ames AmeriFlux towers. The shaded area indicates the average percent slope for the NCRFC region studied.](image-url)
moves from upstream cells to downstream cells using the topographically defined cell-to-cell connectivity sequence developed by OHD based on digital elevation model (DEM) data (Reed 2003; NWS 2008).

Inputs to the HL-RDHM are 6-h precipitation, 6-h air temperature, and daily PE or PET. During the model simulation, the daily PET is uniformly interpolated into 6-h time steps. HL-RDHM is structured on the Next Generation Weather Radar (NEXRAD) Hydrologic Rainfall Analysis Project (HRAP) grid coordinate system (Fulton 1998), and all input, including parameters and data, must be in this coordinate system. The grid cell resolution of 1 HRAP pixel is approximately 4 km × 4 km and data are mapped by a polar stereographic projection.

Model parameter grids (a priori parameters) are obtained from OHD for almost all major parameters implemented in the HL-RDHM. Readers are referred to Koren et al. (2000, 2003) for more information regarding the method used by OHD to derive the SAC-SMA parameters. Where gridded data were not available, a single value was used for the entire basin. In such cases, the basin specific value used by the NCRFC in their lumped operational modeling system is applied. All model simulations are run in the default spatial resolution of 1 HRAP and at 6-h temporal resolution to allow the use of the OHD parameter grids, which are developed for those spatial and temporal resolutions. All model simulations are continuous starting with a spinup period of calendar year 2002, with 2003–10 used for model evaluation.

c. HL-RDHM calibration

This study follows the current practices for HL-RDHM calibration using observed discharge at the basin outlet and an automated stepwise line search (SLS) procedure (Kuzmin et al. 2008; NWS 2008). The SLS technique steps through each parameter successively, minimizing the objective function with respect to each parameter. If the parameter value remains the same for three consecutive loops, it is eliminated in subsequent optimization loops.

Parameters are optimized at multiple time scales (we use 24-, 240-, and 720-h intervals) using the following objective function \( J \) (NWS 2008):

\[
J = \sqrt{\sum_{k=1}^{n} \left( \frac{\sigma_k}{\sigma} \right)^2 \sum_{i=1}^{m_k} \left[ q_{o,k,i} - q_{s,k,i}(X) \right]^2},
\]

where \( q_{o,k,i} \) and \( q_{s,k,i} \) are the observed and simulated streamflows averaged over the time interval \( k \), \( \sigma_k \) is the standard deviation of the observed streamflow, \( n \) is the total number of time scales used, and \( m_k \) is the number of ordinates for time scale \( k \). In Eq. (1), the weight associated with each term is given by the inverse of the standard deviation of the flow at the respective time scales. This weighting scheme assumes that the uncertainty in modeled streamflow for each time scale is proportional to the variability of the observed flow at that scale. An important motivation for using the multiscale objective function is that it creates a smooth objective function surface, thereby reducing the likelihood of the search getting stuck in false minimums (NWS 2008).

Rather than calibrating the value of the parameter itself, the HL-RDHM program calibrates multipliers for each parameter that are applied to the a priori values. For a given basin and parameter, a single multiplier value is identified. The same multiplier value is applied to each grid cell associated with that basin.

Ten SAC-SMA and two SNOW-17 parameters are selected for calibration and associated parameter ranges

<table>
<thead>
<tr>
<th>Table 1. Study sites and locations.</th>
</tr>
</thead>
<tbody>
<tr>
<td>River basin</td>
</tr>
<tr>
<td>East Branch</td>
</tr>
<tr>
<td>Pecatonica River</td>
</tr>
<tr>
<td>Fox River</td>
</tr>
<tr>
<td>Pecatonica River</td>
</tr>
<tr>
<td>Crawfish River</td>
</tr>
<tr>
<td>South Skunk River</td>
</tr>
<tr>
<td>Squaw Creek</td>
</tr>
<tr>
<td>Beaver Creek</td>
</tr>
<tr>
<td>North Raccoon River</td>
</tr>
<tr>
<td>High Island Creek</td>
</tr>
<tr>
<td>Redwood River</td>
</tr>
<tr>
<td>Clearwater River</td>
</tr>
<tr>
<td>Blue Earth River</td>
</tr>
<tr>
<td>Kettle River</td>
</tr>
</tbody>
</table>

132 JOURNAL OF HYDROMETEOROLOGY VOLUME 16
are identified (Table 2) based on previous studies and NWS documentation (Hogue et al. 2000; Anderson 2002; Tang et al. 2007; Steffens and Franz 2012). Initial hydrograph analysis indicates that the timing of the discharge peaks is quite accurate relative to the magnitude. Calibration of routing parameters in select basins results in minor changes to the simulated hydrograph; therefore, to reduce computing time during the calibration process, routing parameters are not included.

To identify the allowable range of the parameter multipliers, basin-specific maximum and minimum multipliers are computed for each parameter according to the following equations:

\[ x_{\text{min}} = \alpha / p_{\text{min}} \]  
\[ x_{\text{max}} = \alpha / p_{\text{max}} \]

where \( x_{\text{min}} \) and \( x_{\text{max}} \) are the minimum and maximum multipliers, \( \alpha \) is the basin mean a priori parameter value, and \( p_{\text{min}} \) and \( p_{\text{max}} \) are the maximum and minimum values for the parameter (Table 2). Not all basin mean a priori parameter values fall within the range indicated in Table 2. For these instances, \( p_{\text{min}} \) and \( p_{\text{max}} \) are set to the a priori mean value ± 10%.

Tang et al. (2007) found significant parameter sensitivity in the HL-RDHM when comparing calibrations to wet versus dry conditions. In general, it has been recommended that calibration data should contain several very wet periods (Yapo et al. 1996). Therefore, we split the 8-yr data record into the calibration period, from 1 January 2007 to 31 December 2010, and the verification period, from 1 January 2003 to 31 December 2006. The calibration period was chosen to coincide with two significantly wetter-than-average years for most basins.

Results of the calibration process are summarized as the percent change from the a priori parameter value relative to the range of the calibrated parameter multipliers:

\[ (\Delta) \Delta = \frac{(x_{\text{calib}} - x_{\text{apriori}}) / (x_{\text{max}} - x_{\text{min}})}{100}, \quad (4) \]

where

\[ x_{\text{calib}} = a \times x_{\text{apriori}} \]

and \( x_{\text{calib}} \) is the basin-averaged parameter value after calibration; \( x_{\text{apriori}} \) is the basin-averaged a priori parameter value, \( x_{\text{max}} \) and \( x_{\text{min}} \) are the maximum and minimum parameter range (Table 2), respectively; and \( a \) is the calibrated parameter multiplier.

### d. Evaluation statistics

Three standard summary statistics are used to evaluate data and simulations in this study: bias, percent bias (Pbias), coefficient of determination (\( R^2 \)), and Nash–Sutcliffe efficiency (NSE; Nash and Sutcliffe 1970):

\[ \text{Bias} = \frac{1}{n} \left[ \sum_{i=1}^{n} (x_{\text{model},i} - x_{\text{obs},i}) \right], \quad (6) \]

\[ \text{Pbias} = \left[ \frac{\sum_{i=1}^{n} (x_{\text{model},i} - x_{\text{obs},i})}{\sum_{i=1}^{n} (x_{\text{obs},i})} \right] \times 100, \quad (7) \]

\[ R^2 = \left[ \frac{\sum_{i=1}^{n} (x_{\text{model},i} - \bar{x}_{\text{model}}) (x_{\text{obs},i} - \bar{x}_{\text{obs}})}{\sqrt{\sum_{i=1}^{n} (x_{\text{model},i} - \bar{x}_{\text{model}})^2} \sqrt{\sum_{i=1}^{n} (x_{\text{obs},i} - \bar{x}_{\text{obs}})^2}} \right]^2, \quad (8) \]
\[
NSE = 1 - \frac{\sum_{i=1}^{n} (x_{\text{obs},i} - x_{\text{model},i})^2}{\sum_{i=1}^{n} (x_{\text{obs},i} - \bar{x}_{\text{obs}})^2},
\]

(9)

where \( n \) is the number of days analyzed; \( x_{\text{obs},i} \) and \( x_{\text{model},i} \) are the observed and modeled data at day \( i \), respectively; and \( \bar{x}_{\text{obs}} \) and \( \bar{x}_{\text{model}} \) are the mean of the observed and modeled data, respectively. NSE values can range from \(-\infty \) to 1. A value of 1 indicates a perfect match of simulated to observed, a value of 0 indicates the model simulation is as accurate as the mean of the observed data, and negative values indicate the observed mean is a better predictor than the model. All error statistics are calculated for the warm season months from May through September.

3. Data

a. Default-PET

Default-PET data were developed by OHD for the HL-RDHM and are the product of the climatological PE and PE adjustment factors (Anderson 2002; NWS 2008). The climatological PE is based primarily on an unpublished method that combines NOAA data from seasonal and annual free water surface maps and mean monthly station data to predict climatological mean daily PE variability (Farnsworth and Thompson 1982; Farnsworth et al. 1982). PE adjustment factors for each month are developed using an empirical function that links PE adjustment factors to satellite observations of green vegetation fraction (NWS 2008).

Both the climatological PE and the PE adjustment factors are daily average values and are provided for each month at 4-km spatial resolution. They are assumed to be midmonth values from which daily values are interpolated (Fig. 2). Because it is climatology, the default-PET is the same for every year of the simulations.

b. M-PET

Instantaneous M-PET values are initially computed as instantaneous values at the satellite overpass time using estimates of net radiation \( R_n \), air temperature \( T_a \), and soil heat flux \( G \) derived from nine MODIS products. Longwave radiation is estimated using normalized difference vegetation index (NDVI), air temperature, dewpoint temperature, geopotential height, surface temperature, and emissivity. Soil heat flux is estimated using a relationship with the NDVI, albedo, and land surface temperature values. Shortwave radiation is estimated by two methods depending on the amount of cloud cover at the time of the satellite overpass. For clear days (cloud fractional coverage <20%), the solar zenith angle, precipitable water, ozone, and albedo are used to estimate shortwave radiation at the surface. On cloudy days (cloud fraction coverage >20%), the algorithm derives a theoretical clear-day shortwave radiation by interpolating between adjacent clear days. The daily mean cloud fraction and daily cloud optical depth products are then incorporated into an empirical formula to estimate the theoretical clear-sky shortwave radiation product. The shortwave radiation is used to estimate net radiation by applying a simple linear regression. Only positive net radiation values are applied. Total daily PET is estimated from a sinusoidal model applied to the instantaneous PET using the day length (difference between time of sunrise and sunset; Kim and Hogue 2008). The final result is an all-sky daily PET in millimeters per day.

M-PET is computed at 500-m resolution (Fig. 3a). Geographic coordinates of each 500-m pixel are converted to the HRAP coordinate system and grid cells are aggregated to 4-km resolution for HL-RDHM input (Fig. 3b). Daily M-PET is generated for the warm season months from May through September 2003–10, to coincide with the period when ET rates have the largest impact on the water balance in the study region. Default-PET values are used for days with missing M-PET data and for the cool season months (1 October to 30 April). On average, 25% of the days from May through September are missing. PE adjustment factors are not applied in the HL-RDHM when using the M-PET as model input.
c. **Flux-tower observations**

Latent heat flux (LE), net radiation, ground heat flux, and air temperature data are collected from two AmeriFlux flux-tower sites (https://public.ornl.gov/ameriflux) located within 5 km of the Squaw Creek, Iowa, basin (Brooks Field site 11–Ames, Ames, Iowa) and within 20 km of the Redwood River basin, Minnesota (Brookings site, Brookings, South Dakota; Fig. 1). Both sites use sonic anemometer and an open path CO$_2$/H$_2$O gas analyzer to calculate LE values. The Webb et al. (1980) and coordinate rotation corrections are applied. Land cover at the Ames site is rotations of corn and soybean, and land cover at the Brookings site is pasture and grassland. The flux-tower sites are similar to land cover conditions in the associated watersheds.

Observations at the AmeriFlux sites are every 30 min. The observations are averaged to daily mean values for the months of May–September for the available period of record (2005–10 at the Ames site and 2004–09 at the Brookings site). Positive daily mean LE values are converted to an equivalent depth of liquid water ET and used to evaluate the simulated ET for the two nearby basins. The mean daily ET is 3.0 mm day$^{-1}$ at the Ames flux-tower site and 3.9 mm day$^{-1}$ at the Brookings site.

d. **Temperature, precipitation, and observed discharge**

Ground-based station data from nearby Automated Surface Observing System (ASOS) and Automated Weather Observing System (AWOS) networks are used to develop air temperature grids for SNOW-17 model input. Station data recorded every 20 min are obtained from the Iowa Environmental Mesonet (http://mesonet.agron.iastate.edu/) and converted to a 6-h mean value. A 4-km resolution grid is interpolated for each region through the inverse distance weighting (IDW) method for the years 2002–10.

Precipitation data from the National Centers for Environmental Prediction (NCEP) Environmental Modeling Center Climatology-Calibrated Precipitation Analysis (CCPA) are used. CCPA is a 6-h precipitation product for the CONUS at the HRAP resolution available from 2002 to present. CCPA combines the high climatological reliability of the Climate Prediction Center (CPC) Unified Global Daily Gauge Analysis (24-h accumulation at 1/8° resolution) and the high temporal and spatial resolution of the NCEP stage IV analysis (6-h accumulation at 4-km resolution) (Hou et al. 2014).

During initial testing, observed discharge was considerably underestimated when using the CCPA as model input. This occurred when using both a priori and calibrated parameters. Initial calibrated simulations with default-PET for the North Raccoon River, Redwood River, Blue Earth River, East Branch Pecatonica River, and Pecatonica River resulted in simulated discharge with −46.2%, −71.0%, −56.0%, −55.0%, and −72.7% PBias, respectively. Hou et al. (2014) noted that CCPA precipitation estimates are better for lower and medium daily precipitation amounts compared to heavy precipitation events. Basin-averaged CCPA precipitation data are found to be consistently lower compared to mean areal precipitation (MAP) data obtained from the NCRFC for
each study basin (Table 3). To correct for the apparent bias, a precipitation adjustment factor is developed for each basin (Table 3, column 4) by calculating the average difference between total 1 March–30 September precipitation for the period of 2003–08 from the CCPA and MAP. The period of March–September is used because that is when the majority of the annual precipitation occurs and heavy events are most common. After applying the precipitation adjustment factors, annual basin-averaged CCPA values more closely match the annual average MAP (Table 3). Precipitation adjustment factors are specified in the HL-RDHM and applied to each grid cell in the corresponding basin.

### Table 3. Annual basin-averaged precipitation for 2003–10 for data from the NCRFC and the CCPA, precipitation adjustment factors for the CCPA data, and basin-averaged CCPA precipitation after application of the adjustment factor for all study sites. Note NCRFC data do not include 2010 calendar year.

<table>
<thead>
<tr>
<th>River basin</th>
<th>NCRFC precipitation (mm)</th>
<th>CCPA precipitation (mm)</th>
<th>Precipitation adjustment factor (%)</th>
<th>Adjusted CCPA precipitation (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>East Branch</td>
<td>935</td>
<td>870</td>
<td>7</td>
<td>937</td>
</tr>
<tr>
<td>Pecatonica River</td>
<td>824</td>
<td>809</td>
<td>0</td>
<td>809</td>
</tr>
<tr>
<td>Fox River</td>
<td>981</td>
<td>732</td>
<td>22</td>
<td>939</td>
</tr>
<tr>
<td>Crawfish River</td>
<td>925</td>
<td>901</td>
<td>1</td>
<td>939</td>
</tr>
<tr>
<td>South Skunk River</td>
<td>991</td>
<td>908</td>
<td>11</td>
<td>1008</td>
</tr>
<tr>
<td>Squaw Creek</td>
<td>1011</td>
<td>839</td>
<td>23</td>
<td>1032</td>
</tr>
<tr>
<td>Beaver Creek</td>
<td>966</td>
<td>942</td>
<td>3</td>
<td>970</td>
</tr>
<tr>
<td>North Raccoon River</td>
<td>818</td>
<td>736</td>
<td>9</td>
<td>809</td>
</tr>
<tr>
<td>High Island Creek</td>
<td>718</td>
<td>519</td>
<td>31</td>
<td>680</td>
</tr>
<tr>
<td>Redwood River</td>
<td>706</td>
<td>587</td>
<td>19</td>
<td>725</td>
</tr>
<tr>
<td>Clearwater River</td>
<td>602</td>
<td>436</td>
<td>25</td>
<td>545</td>
</tr>
<tr>
<td>Blue Earth River</td>
<td>814</td>
<td>654</td>
<td>18</td>
<td>798</td>
</tr>
<tr>
<td>Kettle River</td>
<td>725</td>
<td>560</td>
<td>24</td>
<td>694</td>
</tr>
</tbody>
</table>

4. Results

#### a. Assessment of M-PET and default-PET

The basin-averaged M-PET and default-PET are evaluated against observed daily PET from flux-tower sites near two study basins, Squaw Creek and Redwood River (Fig. 1). Observed daily PET at the flux towers is computed using the Priestley–Taylor formula and daily mean net radiation, air temperature, and ground heat flux measurements. To be consistent with the M-PET techniques, only positive values of net radiation are applied to calculate daily net radiation. The basin-averaged default-PET has low correlation to the flux-tower PET (average $R^2 = 0.13$) and a negative bias (average $-0.5$ mm day$^{-1}$; Table 4). By comparison, the basin-averaged M-PET data for Squaw Creek and Redwood River have good correlation to flux-tower PET values (average $R^2 = 0.66$) and an average positive bias of 1.2 mm day$^{-1}$ (Table 4).

Given the lower bias, the default-PET likely better represents the long-term average PET. But the higher $R^2$ suggests that the M-PET data better represents the full range and day-to-day variability of the observed PET (Fig. 4). The range of the default-PET is significantly smaller than the observed for both basins (Figs. 4a,b), whereas M-PET is more similar to observations with some overestimation of the higher values (Figs. 4c,d). The M-PET data in Fig. 3b range from 5.9 to 7.2 mm, illustrating that the M-PET values are larger and more variable than the default-PET which range from 5.3 to 5.7 mm (Fig. 2) for the day depicted.

Bowman et al. (2013, manuscript submitted to J. Hydrol. Eng.) also found the range of climatological PET values used operationally by the NCRFC to be less than that of the observed PET for these same basins and flux-tower sites.

Kim and Hogue (2008) found better correlation ($R^2 = 0.89$) and lower bias ($-0.34$ mm day$^{-1}$) in point-to-grid comparisons of daily M-PET and observed PET for humid sites. Our correlations are slightly lower most likely because 1) we are comparing to basin-scale PET as opposed to point-to-grid comparisons and 2) we include only May–September whereas Kim and Hoge

### Table 4. Bias and $R^2$ in daily basin-averaged PET compared to flux-tower PET for and default-PET at Squaw Creek and Redwood River basins.

<table>
<thead>
<tr>
<th>Bias (mm day$^{-1}$)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Default-PET</td>
</tr>
<tr>
<td>Squaw Creek</td>
<td>$-0.55$</td>
</tr>
<tr>
<td>Redwood River</td>
<td>$-0.53$</td>
</tr>
<tr>
<td>Mean</td>
<td>$-0.54$</td>
</tr>
</tbody>
</table>
include all months—correlations to observed PET have been found to be highest in the colder months. More recent studies (Barik 2014) confirm that M-PET slightly overestimates PET as compared to ground-based measurements, likely because of the observed biases in MODIS surface temperature products (Wan et al. 2002; Bisht et al. 2005).

### b. Model calibration

Compared to the a priori value, the upper-zone tension water maximum storage (UZTWM) and lower-zone tension water maximum storage (LZTWM) parameters have the largest average change of all parameters calibrated (Tables 2 and 5). These two parameters define the amount of stored moisture available to meet PET demands. The decrease in the LZTWM and UZTWM is a result of the calibration algorithm attempting to distribute more water to streamflow to correct under-simulated discharge that was produced by both PET datasets (Table 6; Figs. 5a,b). Associated with this, UZFWM, the parameter that controls quick-moving water storages in the upper zone (i.e., runoff and interflow), is significantly increased on average (Table 5).

For some basins, ZPERC (which controls the flow of water from the upper zone to the lower zone) is also decreased to reduce flow to the lower soil storages, effectively keeping the upper zone wetter and producing higher and more rapidly occurring peak flows (Fig. 5b). For example, for the East Branch Pecatonica River, overestimation of peaks and underestimation of base flow (Figs. 5c,d) are corrected during the calibration with an increase in LZFSM and LZFPM—the lower-zone free water parameters that store water for postpeak and baseflow recessions—and a decrease in the UZTWM. No consistent increase or decrease in the parameters related to the recession coefficients (LZPK, LZSK, and UZK; see Table 2) are observed (Table 5); therefore, the significant improvement in baseflow simulations after calibration appear to be primarily due to changes in the water storage parameters.

When comparing the average change in parameters between the default-PET and the M-PET calibrations,
the greatest differences are seen in parameters that control lower-zone storages and percolation from the upper to the lower zone (LZFSM, LZFPM, LZSK, ZPERC, and REXP; Table 5). These lower-zone parameters influence the lower flows in the hydrograph, although as their values decrease, more water moves through the upper zone, increasing the magnitude of higher flows (Yilmaz et al. 2008). The smaller increase in LZFSM and larger reduction in LZFPM for the M-PET calibrations indicate that, on average, lower-zone storages are smaller in the M-PET calibrations compared to the default-PET calibrations. The larger values of the M-PET produce more simulated ET and less water retention in the soil zone; therefore, less lower-zone storage is needed. The reason for differences in the percolation parameters is difficult to diagnose because of compensatory interactions between REXP and ZPERC (Gupta and Sorooshian 1983; Anderson 2002).

Simulated discharge is substantially improved after calibration for all watersheds and both PET datasets based on visual inspection of the hydrographs and the summary statistics. The average PBias for the verification period improved from −40.3% to 12.0% for default-PET simulations and from −63.2% to −14.4% for M-PET simulations (Table 6). The average NSE improved from −0.05 to 0.46 for the default-PET simulations and from −0.19 to 0.42 for M-PET simulations (Table 7). Bias in daily simulated ET as compared to the flux-tower observations is improved for Squaw Creek but degraded for Redwood River for both default-PET and M-PET (Table 8). Correlation for simulated ET was worse or showed no improvement after calibration for both basins.

c. Evaluation of simulated discharge

The remainder of this section focuses on calibrated model results only. On average, the default-PET simulations have a lower PBias for both the calibration and verification periods compared to the M-PET (Table 6). Because M-PET values are biased high, the M-PET model simulations show primarily negative PBias values. The default-PET simulations have slightly better mean NSE values than the M-PET simulations, although most values show little change between the two datasets (Table 7). NSE scores for the verification period tend to be better for the M-PET in the Wisconsin basins and better for the default-PET in the Iowa and Minnesota basins. Correlations (not shown) range from 0.48 to 0.94 and are slightly better for the default-PET, with an average across all basins of 0.73 for the default-PET compared to 0.67 for the M-PET.

The degree to which the simulations match the observed hydrograph varies between basins and from year-to-year for individual basins. Hydrographs from all years and all basins can be found in Spies (2013). Results depicted in Fig. 6 are typical in that the M-PET often provides a good match to the hydrograph but underestimates the streamflow to a slightly greater degree than the default-PET. Sites that tended to perform well using the default-PET, also tended to perform well when using the M-PET (i.e., using the M-PET neither substantially improved nor degraded the quality of the simulation for most study basins). Hydrographs produced using the default-PET and M-PET are most similar for the East Branch Pecatonica River (Wisconsin; Fig. 6c), Pecatonica River (Wisconsin), Clearwater River (Minnesota), North Raccoon River (Iowa; Fig. 6d), and Squaw Creek (Iowa; Fig. 6f). The NSE scores (Table 7) are among the highest for these same sites.

Redwood River (Minnesota) was one of the worst performing sites; NSE scores for the verification period were −0.15 for M-PET and −0.34 for default-PET (Table 7). Late summer discharge is consistently overestimated at this site (Fig. 6e). The Blue Earth River (Minnesota) shows the most improvement from the application of the M-PET data: NSE = 0.57 and PBias = −4.45% for M-PET for the verification period compared to NSE = 0.12 and PBias = 46.73% for default-PET (Tables 6, 7). However, both default-PET and M-PET simulations frequently overestimate peaks for Blue Earth River (Fig. 6a).

The higher M-PET estimates result in discharge simulations with lower peaks (Figs. 6a,b,e), and more commonly lower base flow (Figs. 6a,b,e,f) compared to the default-PET. The higher M-PET values do not lead to consistent improvements in the discharge
Table 6. PBias (%) in simulated discharge for all study sites and the mean of all sites for a priori and calibrated parameter sets for the period of 2007–10 (calibration period) and 2003–06 (verification period). PBias values of ±20% are in boldface.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>East Branch</td>
<td>-44.05</td>
<td>-43.73</td>
<td>-51.20</td>
<td>-48.90</td>
<td>26.15</td>
<td>30.78</td>
<td>16.35</td>
<td>22.25</td>
</tr>
<tr>
<td>Pecatonica River</td>
<td>-44.40</td>
<td>-48.83</td>
<td>-63.90</td>
<td>-64.18</td>
<td>-20.00</td>
<td>-23.00</td>
<td>-36.80</td>
<td>-38.88</td>
</tr>
<tr>
<td>Fox River</td>
<td>-46.25</td>
<td>-42.23</td>
<td>-53.55</td>
<td>-48.83</td>
<td>-12.23</td>
<td>-4.05</td>
<td>-24.05</td>
<td>-16.93</td>
</tr>
<tr>
<td>Pecatonica River</td>
<td>-2.25</td>
<td>-14.10</td>
<td>-28.10</td>
<td>-30.05</td>
<td>15.50</td>
<td>12.65</td>
<td>-1.78</td>
<td>-2.43</td>
</tr>
<tr>
<td>Crawfish River</td>
<td>8.57</td>
<td>-19.13</td>
<td>-51.38</td>
<td>-64.90</td>
<td>19.03</td>
<td>36.13</td>
<td>-20.08</td>
<td>-6.48</td>
</tr>
<tr>
<td>South Skunk</td>
<td>-29.90</td>
<td>-49.30</td>
<td>-58.50</td>
<td>-75.45</td>
<td>0.03</td>
<td>1.35</td>
<td>-38.08</td>
<td>-41.10</td>
</tr>
<tr>
<td>Squaw Creek</td>
<td>-22.83</td>
<td>-12.78</td>
<td>-57.83</td>
<td>-56.00</td>
<td>15.33</td>
<td>53.08</td>
<td>-20.42</td>
<td>14.45</td>
</tr>
<tr>
<td>Beaver Creek</td>
<td>-23.15</td>
<td>-40.05</td>
<td>-65.38</td>
<td>-74.40</td>
<td>15.43</td>
<td>6.68</td>
<td>1.20</td>
<td>7.68</td>
</tr>
<tr>
<td>North Raccoon</td>
<td>-41.65</td>
<td>-49.80</td>
<td>-64.80</td>
<td>-69.05</td>
<td>25.67</td>
<td>33.35</td>
<td>-17.88</td>
<td>-19.70</td>
</tr>
<tr>
<td>High Island</td>
<td>-54.83</td>
<td>-47.05</td>
<td>-69.47</td>
<td>-61.20</td>
<td>17.13</td>
<td>58.63</td>
<td>4.65</td>
<td>43.18</td>
</tr>
<tr>
<td>Creek</td>
<td>-77.32</td>
<td>-83.90</td>
<td>-84.23</td>
<td>-87.95</td>
<td>-49.45</td>
<td>-52.78</td>
<td>-61.13</td>
<td>-60.08</td>
</tr>
<tr>
<td>Clearwater River</td>
<td>-9.25</td>
<td>-18.15</td>
<td>-52.95</td>
<td>-59.55</td>
<td>38.23</td>
<td>46.73</td>
<td>-2.45</td>
<td>-4.45</td>
</tr>
<tr>
<td>Blue Earth River</td>
<td>-59.95</td>
<td>-53.08</td>
<td>-86.08</td>
<td>-81.63</td>
<td>-48.93</td>
<td>-44.25</td>
<td>-71.30</td>
<td>-69.45</td>
</tr>
<tr>
<td>Kettle River</td>
<td>-35.72</td>
<td>-40.32</td>
<td>-60.58</td>
<td>-63.24</td>
<td>3.22</td>
<td>11.95</td>
<td>-20.91</td>
<td>-14.41</td>
</tr>
<tr>
<td>Mean</td>
<td>-20.91</td>
<td>-24.13</td>
<td>-48.90</td>
<td>-54.83</td>
<td>-20.00</td>
<td>-23.00</td>
<td>-36.80</td>
<td>-38.88</td>
</tr>
</tbody>
</table>

Simulations (Tables 6 and 7), and in some cases the discharge is obviously less accurate (e.g., Fig. 6b). There are examples where M-PET produces improved estimates of discharge (e.g., May–June; Figs. 6a,e). Overestimation of discharge peaks during August and September is common for the Minnesota and Iowa sites and occurs for both the M-PET and default-PET (Figs. 6d–f).

Simulated hydrographs are observed to be mostly identical during the wettest years and wetter time period. This occurs despite differences in simulated ET (e.g., Figs. 6c,d). In these instances, sufficient moisture is available to continually meet PET demands and soil stores are mostly full. Precipitation will predominantly go to runoff, and minor variations in the depletion rate of the soil stores are not noticed in the hydrographs.

d. Evaluation of simulated ET

In most cases, the M-PET produces simulated ET that is larger and more variable than the default-PET (Fig. 6). Simulated ET is higher for M-PET than the default-PET, ranging from 0.7% more ET for Clearwater River to 8.2% more ET for Squaw Creek (Fig. 7). The Redwood River is the only study site for which the default-PET simulation has more ET than the M-PET, but only by 0.2%. The largest differences in simulated ET tend to occur from May to mid-July (Fig. 6), which is the wetter time period. With enough moisture available to meet the higher demand of the M-PET inputs, the simulated ET is larger. From mid-July through September, there is less water available and the ET simulations are more similar.

When compared to the flux-tower observations, the basin-averaged simulated daily ET from the M-PET have slightly better correlation and bias for Redwood River, but slightly worse for Squaw Creek compared to the default-PET (Table 8; see calibrated results). For the Squaw Creek basin, the lower values of the default-PET result in lower simulated ET values and, subsequently, less overestimation of ET (Fig. 8a) compared to the M-PET simulation (Fig. 8c). The larger negative PBias in discharge (Table 6) and tendency to underestimate base flow compared to the default-PET (Fig. 6f) is further evidence that the M-PET is overestimated at this site, particularly from May to mid-July.

For the Redwood River basin, the M-PET produces a daily ET output that matches both the magnitude and range of the observed well (Fig. 8d), particularly from June to mid-August (Fig. 6e). The default-PET on the other hand, results in a minimally varying ET simulation. The low values of the default-PET place an erroneously low upper limit on the simulated ET, resulting in an inability to model the largest observed values (Fig. 8b). Despite the difference in PET inputs, in most years total May–September ET from both sets of simulations is nearly identical and underestimated as compared to the flux-tower data for the Redwood River (Spies 2013). A simple sensitivity analysis is undertaken to test the possibility that the PET is underestimated at this site and the underestimated discharge (Fig. 6e) could be corrected with higher PET inputs. M-PET input was increased by 50% during July, August, and September, but little to no
change in the overestimated discharge is observed (not shown here).

5. Discussion

Discharge simulations are improved after calibration from the a priori parameters, similar to findings by others (Reed et al. 2004; Wagener et al. 2009; Smith et al. 2012b). In general, model calibration increases the amount of free water storage (moisture available for base flow, interflow, and runoff) while slowing the percolation of water from the upper zone to the lower zone (more upper-zone moisture available for interflow and storm runoff). The model calibration results in a decrease in simulated ET for all basins and both tested PET datasets. There is evidence that the calibration may improve the discharge simulation at the expense of the accuracy of the simulated ET. For instance, bias in ET increases after calibration at Redwood River for the default-PET and M-PET simulations.

At some sites, in particular the Clearwater, East Branch of the Pecatonica, and Pecatonica Rivers, the simulated discharge from the calibrated model is similar between the default-PET and M-PET despite differences in the magnitude, range, and variability of the PET inputs. These results support others (Fowler 2002; Oudin et al. 2005) who found that temporally varying potential evaporation inputs do not necessarily produce better streamflow simulations than long-term average potential evaporation inputs. However, these previous studies do not consider the accuracy of the simulated ET.

Using the flux-tower observations we find that the daily varying M-PET may lead to improved simulations of overall ET dynamics while still maintaining, and sometimes improving, the accuracy in the simulated discharge. In the example of the Redwood River for 2005, the M-PET–simulated ET matches the observed ET quite well, particularly from May to July, and corresponds to a high accuracy in simulated discharge (Fig. 6e). This suggests that improvement in the simulated daily ET can contribute to more accurately simulated peaks and base flow for some conditions in this basin. The influence of the ET is likely to be most important for event-based analysis and short-term forecasting, when accurate soil moisture conditions are most critical.

Uncertainties in accuracy of the latent heat observations and the use of point-scale observations as representations of the basin-averaged ET limit the ability to draw more definitive conclusions from the ET analysis. However, flux towers are currently the best available ground-based source of evapotranspiration data, and we demonstrate that these records are useful for understanding the ET output. As stated by others (Beven 2001; Kirchener 2006), use of multiple datasets in hydrologic model validation should be encouraged.

Through parameterization, the SAC-SMA is able to adjust to different PET inputs such that the model efficiency is similar for the M-PET and default-PET in most basins. Andréassian et al. (2004) came to this same conclusion when they tested different estimates of long-term average PET for input to the TOPMODEL for watersheds in the Massif Central highlands of France. Similar to Andréassian et al. (2004), we advocate that we do not see significant improvements in simulated discharge when “improved” estimates of daily PET are used given either data errors or model limitations (parameters and structure).

A potential limitation of the M-PET product is the uncertainty associated with the Priestley–Taylor formulation (Kim and Hogue 2008). Meteorological variables that are not included in the Priestley–Taylor equation are accounted for by the alpha parameter. Alpha, which
is treated as a constant in this study, has been found to vary seasonally (Jin et al. 2011). Uncertainties also exist in the MODIS variables used to derive M-PET. Air temperature values are biased low when compared to AmeriFlux sites associated with both the Squaw Creek and Redwood River basins (not shown here). Net radiation values are biased high for Squaw Creek but low for Redwood River. Kim and Hogue (2008) found errors in MODIS variables, such as air and surface temperature and albedo, to be dependent on viewing angle, spatial and temporal resolution, and atmospheric corrections, among others. The impact of PET errors may propagate through the model over time, leading to consistent underperformance. For instance, water budget deficits that occur in late summer periods can cause reductions in lower-zone reserves that last throughout the fall and winter and ultimately impact the next season’s simulations (Tang et al. 2007).

The choice of PET input may be more critical for dry periods and at sites with less annual precipitation. Hydrographs in the wetter Iowa and Wisconsin sites were often very similar between the two PET inputs, while at all basins the simulated hydrographs varied mostly with respect to recessions and base flow. For the Redwood River basin, M-PET tends to produce larger daily ET, resulting in lower soil stores and leading to lower, more accurate low flows in some years. During peak flows and when basin conditions are wet, the hydrographs are dominated by direct runoff. Although ET is at its maximum during wet periods and the depletion rate in soil stores will be slightly different between the M-PET and default-PET simulations, the minor variations do not appear to impact the hydrograph significantly.

The relatively poor correlation to observed ET for both PET inputs ($R^2$ values range from 0.13 to 0.35; Table 8) suggests that the HL-RDHM has difficulties reproducing daily ET for the two basins evaluated. Koren et al. (2010) state that the SAC-SMA and SAC-HT are limited by an inability to explicitly account for changes in the watershed due to vegetation and canopy controls on evapotranspiration. The SAC-HT evapotranspiration formulation model (SAC-HTET; Koren et al. 2010) has been developed as an option within the HL-RDHM to resolve this limitation by incorporating advanced canopy resistance parameterizations developed for land surface models. A comparison of our results to SAC-HTET simulations for these basins may help further clarify the current state of modeling ET in the HL-RDHM.

The tendency to overestimate discharge peaks during late summer for the Minnesota and Iowa basins and the inability to alter modeled low flows during the late summer in the Redwood River by increasing M-PET by

### Table 7. NSE for simulated discharge for all study sites and the mean of all sites for a priori and calibrated parameter sets for the period of 2007–10 (calibration period) and 2003–06 (verification period). Values greater than 0.60 are in boldface.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>East Branch Pecatonica River</td>
<td>–0.57</td>
<td>–3.55</td>
<td>0.28</td>
<td>–2.69</td>
<td>0.29</td>
<td>0.50</td>
<td>0.41</td>
<td>0.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fox River</td>
<td>0.10</td>
<td>–0.02</td>
<td>0.12</td>
<td>–0.16</td>
<td>0.82</td>
<td>0.78</td>
<td>0.72</td>
<td>0.61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pecatonica River</td>
<td>0.31</td>
<td>0.06</td>
<td>0.40</td>
<td>0.15</td>
<td>0.79</td>
<td>0.65</td>
<td>0.68</td>
<td>0.76</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crawfish River</td>
<td>0.07</td>
<td>0.44</td>
<td>0.49</td>
<td>0.51</td>
<td>0.90</td>
<td>0.86</td>
<td>0.92</td>
<td>0.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Skunk River</td>
<td>0.80</td>
<td>0.70</td>
<td>0.51</td>
<td>0.23</td>
<td>0.80</td>
<td>0.57</td>
<td>0.68</td>
<td>0.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squaw Creek</td>
<td>0.70</td>
<td>0.30</td>
<td>0.46</td>
<td>–0.02</td>
<td>0.84</td>
<td>0.69</td>
<td>0.72</td>
<td>0.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beaver Creek</td>
<td>0.66</td>
<td>0.72</td>
<td>0.45</td>
<td>0.46</td>
<td>0.64</td>
<td>0.23</td>
<td>0.56</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>North Raccoon River</td>
<td>0.66</td>
<td>0.43</td>
<td>0.19</td>
<td>–0.04</td>
<td>0.73</td>
<td>0.66</td>
<td>0.63</td>
<td>0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Island Creek</td>
<td>0.32</td>
<td>0.25</td>
<td>0.04</td>
<td>–0.03</td>
<td>0.65</td>
<td>0.58</td>
<td>0.52</td>
<td>0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Redwood River</td>
<td>0.22</td>
<td>0.22</td>
<td>0.04</td>
<td>0.07</td>
<td>0.53</td>
<td>–0.34</td>
<td>0.54</td>
<td>–0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clearwater River</td>
<td>–0.16</td>
<td>–0.58</td>
<td>–0.27</td>
<td>–0.69</td>
<td>0.35</td>
<td>0.19</td>
<td>0.17</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blue Earth River</td>
<td>0.42</td>
<td>0.26</td>
<td>0.21</td>
<td>0.06</td>
<td>0.34</td>
<td>0.12</td>
<td>0.46</td>
<td>0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kettle River</td>
<td>0.19</td>
<td>0.17</td>
<td>–0.27</td>
<td>–0.34</td>
<td>0.45</td>
<td>0.44</td>
<td>0.11</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.29</td>
<td>–0.05</td>
<td>0.16</td>
<td>–0.19</td>
<td>0.63</td>
<td>0.46</td>
<td>0.55</td>
<td>0.42</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 8. Bias and $R^2$ in simulated basin-averaged ET compared to flux-tower ET for M-PET and default-PET simulations for a priori and calibrated parameters.

<table>
<thead>
<tr>
<th>Basin</th>
<th>Bias (mm day$^{-1}$)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Default-PET</td>
<td>M-PET</td>
</tr>
<tr>
<td>A priori</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squaw Creek</td>
<td>0.9</td>
<td>2.2</td>
</tr>
<tr>
<td>Redwood River</td>
<td>–0.4</td>
<td>0.0</td>
</tr>
<tr>
<td>Calibrated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Squaw Creek</td>
<td>0.3</td>
<td>1.3</td>
</tr>
<tr>
<td>Redwood River</td>
<td>–0.7</td>
<td>–0.4</td>
</tr>
</tbody>
</table>
50% are further evidence of model structural problems. We suspect that the HL-RDHM is unable to properly capture the full range of hydrologic conditions in these basins as vegetation activity increases and precipitation decreases later in the summer, leading to lower soil moisture, higher surface roughness, and a slower watershed response. The SAC-SMA is designed for flood prediction and therefore may not do well in drier periods or less flashy watersheds when the watershed can receive a significant amount of rainfall without producing a response. Jacobs et al. (2009) also noted difficulties in using the SAC-SMA during low flows.

Results demonstrate that the model accuracy is heavily influenced by the physical and climatic characteristics for which it is applied and is not always a function of whether or not the model was calibrated to the time period for which it is evaluated. For both PET datasets tested, the model performs better for basins with higher annual precipitation—that is, basins in southern Wisconsin (Tables 4 and 5, rows 1–4) typically performed the best, followed by the central Iowa basins (Tables 4 and 5, rows 5–8), with Minnesota (Tables 4 and 5, rows 9–13) basins performing the worst. Even prior to calibration, the model performs better for the generally wetter 2007–10 calibration period compared to 2003–06 validation period for both default-PET and M-PET (see a priori parameter results on Table 6).

There are no obvious trends in model performance based on watershed size, land cover, or latitude. Model performance appears to be correlated with average basin slope. The East Branch Pecatonica and Pecatonica Rivers have among the highest model performances and

![Figure 6](image-url)

**Fig. 6.** Daily simulated and observed ET and discharge from 1 May to 30 Sep 2005 at select sites using calibrated parameters. The periods of zero observed ET in (e) and (f) are due to missing data.
6. Conclusions

We examine the potential for spatially and temporally variable satellite-based PET estimates (M-PET) to be implemented into distributed hydrologic models, using the HL-RDHM as a case study. Results show that the M-PET is able to produce discharge simulations of similar quality as the default-PET at several basins after calibration. While the evaluation statistics are better for the M-PET in some cases, for the 13 watersheds studied, use of the time-varying M-PET input did not produce better overall discharge results than the climatological default-PET. A potential benefit of M-PET data is the ability to more closely reproduce the range and variability of ET as illustrated for the Redwood River compared to default-PET.

While the M-PET product is found to be a suitable alternative to the historical PET data used in the HL-RDHM, additional work to understand the impact of data and model errors is needed. Specifically, the source of potential biases in the M-PET under different conditions and the impact of the precipitation uncertainty on the model results require further exploration. Investigations to test model improvement for dry conditions, including base flow and rainfall following extended dry periods, are also needed. We conducted our model discharge analysis in a very traditional manner by using summary statistics that were averaged across all time steps and looking at long-term hydrographs. Evaluating individual events or testing PET datasets in forecast mode would likely lend more insight into the importance of daily PET dynamics for simulation of discharge in the region. Work is underway to test the M-PET in the SAC-SMA for different watersheds with different climatic and geologic conditions from those explored here, and to further assess the role of watershed characteristics and climate on model performance.
In conclusion, it is often assumed that the application of new hydrologic methods, including data products, will necessarily give better results. In our work, using PET inputs with higher spatial and temporal variability led to mixed results. Our findings emphasize the need to carefully conduct application-specific evaluations of the benefits and limitations of new approaches and datasets.

Acknowledgments. Financial support for this work was provided by the NASA New Investigator Program Grant NNX10AQ77G. Special thanks to Mike DeWeese at the NCRFC and Mike Smith at the OHD for their assistance with the model logistics and data. We also thank the AmeriFlux Research Network PIs who provided observation data used in this analysis.

REFERENCES


FIG. 8. Comparison of mean daily observed ET to the modeled basin-averaged ET from the (a),(b) default-PET and the (c),(d) M-PET simulations for (left) Squaw Creek and (right) Redwood River. All simulations use the calibrated parameters. The sample size $n$ is indicated in each subplot.
FEBRUARY 2015 S P I E S E T A L . 145


