Using Reanalysis and Remotely Sensed Temperature and Precipitation Data for Hydrological Modeling in Monsoon Climate: Mekong River Case Study

H. LAURI
EIA Finland Ltd., Espoo, Finland

T. A. RÄSÄNEN AND M. KUMMU
Water & Development Research Group, Aalto University, Espoo, Finland

(Manuscript received 15 May 2013, in final form 8 April 2014)

ABSTRACT

Many large basins in the “Monsoon Asia” region have sparse surface observation networks of the hydro-meteorological parameters needed for hydrological modeling. These models are often used in water resources-related planning, impact assessments, and flood forecasting, which sets strict requirements for model accuracy and reliability. The aim of this study was to assess the performance of several publicly available reanalyses and remotely sensed datasets when used in modeling of discharges in the Mekong River basin. Tested precipitations were extracted from Tropical Rainfall Measuring Mission (TRMM) 3B42, versions 6 and 7; Asian Precipitation–Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE); Climate Forecast System Reanalysis (CFSR); and Interim ECMWF Re-Analysis (ERA-Interim) datasets. Temperature data were extracted from CFSR and ERA-Interim datasets. The model results obtained using these datasets were compared to measured discharges and modeled values based on daily surface observations. It was found that using TRMM, version 7, and APHRODITE precipitation datasets together with CFSR temperature data resulted in similar accuracy of computed discharges in the Mekong main stem as using surface observation data. This indicates that these gridded datasets might support well the modeling efforts in monsoon-driven large river basins in Monsoon Asia.

1. Introduction

Hydrological modeling of a large river basin requires a substantial amount of meteorological data to drive the model. In some regions available surface observations may be sparse, and often the quality of the historical measurement data can also be questioned. Many large river basins in the “Monsoon Asia” region (South, Southeast, and East Asia), such as the Ganges–Brahmaputra, Irrawaddy, Salween, and Mekong (excluding the Thai part of the basin), suffer from rather poor data coverage. At the same time, water resources–related development is rapid in the region (e.g., Grumbine and Pandit 2013; Johnston and Kummu 2012), mainly due to the large hydropower potential in the Himalayas and the Tibetan Plateau, from where many of the rivers originate, combined with the rapidly growing energy needs of the region. Moreover, a large part (around 25%) of the global population lives in these monsoon-driven large river basins (Varis et al. 2012).

Currently, the publicly available regional or global gridded meteorological datasets based on remote sensing or reanalysis data provide promising, convenient, and relatively easy-to-use input data for modeling in remote areas where no measurements exist or where measurement network data is scarce (Su et al. 2008). However, the suitability and accuracy of such data for hydrological modeling of the Mekong catchment has not been extensively tested. The Mekong catchment is also interesting from a precipitation modeling point of view, as precipitation modeling (including reanalysis) in tropical monsoon climates is considered to be particularly difficult and prone to errors (e.g., Turner et al. 2011).

This study aims to aid in filling this research gap by assessing the performance of two temperature and five precipitation datasets.
precipitation gridded datasets in a distributed hydrological model for the Mekong basin in Southeast Asia. The Mekong was selected as a case study because a hydrological model application using observed data was available for the area (Lauri et al. 2012). The two temperature datasets selected for this study were the Climate Forecast System Reanalysis (CFSR) dataset (Saha et al. 2010) and the Interim European Centre for Medium-Range Weather Forecasts (ECMWF) Re-Analysis (ERA-Interim) dataset (Berrisford et al. 2011). The temperature data were used to compute potential evapotranspiration (PET) in the hydrological model using the Hargreaves–Samani method (Hargreaves and Samani 1982) that estimates daily potential evapotranspiration from daily minimum, maximum, and average temperatures. The Penman–Monteith method could not be used because of lack of sufficient surface observation data for the baseline case. The five selected precipitation datasets were remotely sensed Tropical Rainfall Measuring Mission (TRMM) 3B42, versions 6 and 7 (Huffman et al. 2007); the Asian Precipitation–Highly-Resolved Observational Data Integration Towards Evaluation of Water Resources (APHRODITE) dataset, derived primarily from ground-based rain gauge observations (Yatagai et al. 2009, 2012); and the CFSR (Saha et al. 2010) and ERA-Interim (Berrisford et al. 2011) datasets. Datasets that could be used for hydrological modeling, but which are not involved in this study, include, for example, the remote sensing–based precipitation datasets Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN; Hsu et al. 1997) and Climate Prediction Center (CPC) morphing technique (CMORPH; Joyce et al. 2004) and the Japanese 25-year Reanalysis Project (JRA-25) dataset (Onogi et al. 2007; recently updated to JRA-55). Compared to TRMM, the PERSIANN and CMORPH datasets have a shorter range of available data, which limits the time period available for analysis (PERSIANN starting from the year 2000 and CMORPH starting from 2002). The JRA-25 dataset, on the other hand, had a resolution of 1.25°, which was considered somewhat too sparse for this study.

CFSR data have been applied to the computation of potential evapotranspiration, runoff, and discharge on the global scale by Sperna Weiland et al. (2012). In their study, several potential evapotranspiration methods were tested using the CFSR data, including the Hargreaves–Samani method. Their results indicate that the recalibrated Hargreaves–Samani method with CFSR data gave consistent PET results comparable to observation-derived PET values in multiple climate conditions. ERA-Interim temperature data have been compared to temperature observations, for example, in the European Alpine region by Gao et al. (2012), where the correlation between the reanalysis data and gridded observation data was between 0.947 and 0.992. The suitability of these datasets to the Mekong region for computation of potential evaporation remains, however, an open question.

The TRMM v6 precipitation dataset has been used for hydrological modeling in other parts of the world with variable success. Hydrological model applications using TRMM v6 data in large watersheds include work by Collischonn et al. (2008) and Gu et al. (2010), who used the TRMM data in the Amazon and Yangtze basins, respectively. These studies found that in the context of modeling a large catchment with sparse observation network, the TRMM data produced comparable results to those obtained using rain gauge data, at least on a monthly scale. According to Stisen and Sandholt (2010), who used TRMM v6 data in the Senegal basin, the data still had problems reproducing the interannual variation and seasonal dynamics even after a bias correction for the data and recalibration of the hydrological model.

Applications of TRMM data to smaller watersheds in Asia include studies by Li et al. (2012) and Vu et al. (2012), who used TRMM v6 data in Xinjiang (part of the Yangtze basin) and Dak Bla catchments (part of the Mekong basin), respectively. Both studies found that the TRMM v6 data did not produce accurate results for daily streamflow simulations [Nash–Sutcliffe efficiency (NSE) (Nash and Sutcliffe 1970) of 0.3 and <0.74]. Monthly results were good in the Xinjiang catchment (NSE = 0.86), but inaccurate in the Dak Bla catchment (NSE = 0.27).

A comparison of measured precipitation to TRMM v6 data in Thailand with 1° grid resolution was conducted by Chokngamwong and Chiu (2008). The results obtained show that the TRMM v6 data in Thailand tended to underestimate heavy precipitation rates and overestimate low rain rates. Daily correlation to gridded precipitation data derived from ground-based observations was not good, but the correlation improved when a longer accumulation time was used. Accumulated precipitations for periods longer than 5 days were considered appropriate for hydrological applications.

The APHRODITE precipitation data in Asia have been compared to the Global Precipitation Climatology Center (GPCC) rain gauge–based data by Yatagai et al. (2012). Generally, the pattern of precipitation was similar in both datasets in several test areas. In the Mekong region, the APHRODITE dataset showed somewhat less precipitation compared to the GPCC data. APHRODITE and other gridded observation-based precipitation datasets were also used by Vu et al. (2012) in modeling discharges from a 2560 km² catchment in Vietnam. The APHRODITE data had the best
fit to measured discharges among the gridded datasets on a daily scale, even though local gauge data provided still better results.

The performance of state-of-the-art reanalysis precipitations (such as CFSR and ERA-Interim used in this study) for hydrological modeling is location dependent and may have rather large bias in tropical regions, as noted by Lorenz and Kunstmann (2012). Nevertheless, the CFSR and ERA-Interim precipitation datasets were included in the comparison in order to assess how these two precipitation datasets would perform in the Mekong region.

2. Data and methods

a. The VMod hydrological model

The discharges of the Mekong basin were modeled using the VMod distributed hydrological model (Koponen et al. 2010; Lauri et al. 2012). The model is based on square grid cells, the side length of which may be set from a few hundred meters up to several kilometers. Each model grid cell has a separate submodel that computes soil surface energy balance and water balance for the cell. Outflow from the grid cells is then routed using a one-dimensional river network model. Parameters for each grid cell are estimated from elevation, land use, and soil data, as well as from calibration against measured discharges. A detailed description of the model equations and computation methods can be found in the VMod model manual (Koponen et al. 2010). The model has been successfully applied in the Mekong by, for example, Räsänen et al. (2012), Lauri et al. (2012), and Darby et al. (2013).

b. The Mekong basin VMod model application

The Mekong basin is the largest river basin in Southeast Asia. It covers an area of 795 000 km$^2$ and has an average discharge of 15 000 m$^3$ s$^{-1}$ (475 km$^3$ yr$^{-1}$) (Mekong River Commission 2005). The northernmost part of the basin lies in the Tibetan Plateau in China, while the southernmost point, the Mekong Delta, resides in southern Vietnam. The basin is located between 8° and 34°N and contains mountains over 5000 m in elevation in the northern part of the basin and large tropical floodplains in the southern part of the basin. The annual hydrological cycle is driven mainly by a monsoon climate, resulting in a wet summer season from approximately July until September and a dry season for the rest of the year (Mekong River Commission 2010).

The VMod application used in this study had a grid resolution of 5 × 5 km$^2$ and daily time resolution. A more detailed description of the model setup can be found in Lauri et al. (2012); only a brief summary is given here. The baseline computation was conducted using daily meteorological surface observation data obtained from the Mekong River Commission (MRC; Mekong River Commission 2011) and the National Climatic Data Center (NCDC) Global Surface Summary of Day (GSOD) database (NCDC 2010). At some locations with very sparse temperature, data reanalysis temperatures from the National Centers for Environmental Prediction (NCEP; NOAA 2011) were used (Fig. 1b). Because of data availability and data quality reasons, the model was set up to use precipitation and daily average, minimum, and maximum temperatures for the soil surface water and energy balance computations. Meteorological data were collected for the period 1981–2005 from 151 precipitation and 61 temperature stations (see locations in Figs. 1a and 1b, respectively).

Five Mekong main stem discharge gauging stations were selected to compare the model results to measurements: Chiang Saen, Nong Khai, Mukdahan, Pakse, and Stung Treng (see locations in Fig. 1c). The discharge data were acquired from the MRC database (Mekong River Commission 2011). The modeling takes into account the basin upstream from the Stung Treng gauging station (area of about 637 000 km$^2$). The lower part of the basin contains large floodplains and a complex delta system that are not properly described in the current version of the VMod model.

The computation period used in the performance assessment of reanalysis and remotely sensed datasets consisted of the years 1999–2005, for which all datasets had data. The TRMM dataset limits the starting time (1998 was used for model initialization), and the ending time is limited by the baseline dataset. The baseline model was calibrated against the observed discharges using the whole time period (1999–2005). Model performance using longer calibration and validation periods was tested earlier in Lauri et al. (2012). The 10-yr calibration and 7-yr validation performances, when measured with NSE, were 0.922 and 0.941, respectively (Lauri et al. 2012).

c. Datasets

The performance of two gridded temperature and five gridded precipitation datasets were assessed. A summary of the resolutions and available time periods of the tested datasets are shown in Table 1. All datasets are well documented and are publicly available free of charge. Data for this study were downloaded and converted for model use during February–May 2012.

The TRMM 3B42 data (Huffman et al. 2007) are based on remotely sensed radar and infrared data and are corrected using monthly GSOD rain gauge data. In contrast with gauge data, which measures cumulative precipitation,
the TRMM 3B42 dataset is a combination of satellite snapshots from different sources; that is, it is based on measurements of instantaneous precipitation. The data are updated every few months and are available starting from the year 1998. Version 7 has been available since May 2012. The accuracy of precipitation estimates in Southeast Asia has improved considerably in version 7 when compared to version 6 (as shown later). Most of the existing studies have used version 6 data, which is the reason for using both TRMM versions in this study. The TRMM datasets have a spatial resolution of 0.25° and a 3-h temporal resolution. For modeling purposes, the 3-h values were converted to daily values.

The APHRODITE dataset (Yatagai et al. 2009, 2012) is based on daily or higher-resolution ground observations, which are interpolated to a 0.25° grid having daily temporal resolution. The dataset is available for the period 1951–2007. This dataset is very similar to the baseline, as both datasets are based on interpolated station observations. The APHRODITE dataset uses a comparatively dense precipitation observation dataset; according to Yatagai et al. (2009), the amount of precipitation data

<table>
<thead>
<tr>
<th>Dataset name</th>
<th>Avg distance between points (km)</th>
<th>Spatial resolution</th>
<th>Temporal resolution</th>
<th>n</th>
<th>Data years</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>113</td>
<td>variable</td>
<td>daily</td>
<td>61</td>
<td>1981–2005</td>
<td>See Fig. 1</td>
</tr>
<tr>
<td>CFSR</td>
<td>32</td>
<td>0.375°</td>
<td>6 h</td>
<td>793</td>
<td>1979–2010</td>
<td>Saha et al. (2010)</td>
</tr>
<tr>
<td>ERA-Interim</td>
<td>77</td>
<td>0.75°</td>
<td>6 h</td>
<td>188</td>
<td>1979–2012</td>
<td>Berrisford et al. (2011)</td>
</tr>
<tr>
<td>Precipitation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>46</td>
<td>variable</td>
<td>daily</td>
<td>151</td>
<td>1981–2005</td>
<td>See Fig. 1</td>
</tr>
<tr>
<td>TRMM 3B42 v6</td>
<td>25</td>
<td>0.25°</td>
<td>3 h</td>
<td>1329</td>
<td>1998–2012</td>
<td>Huffman et al. (2007)</td>
</tr>
<tr>
<td>TRMM 3B42 v7</td>
<td>25</td>
<td>0.25°</td>
<td>3 h</td>
<td>1329</td>
<td>1998–2013</td>
<td>Huffman et al. (2007)</td>
</tr>
<tr>
<td>CFSR</td>
<td>32</td>
<td>0.3125°</td>
<td>6 h</td>
<td>793</td>
<td>1979–2010</td>
<td>Saha et al. (2010)</td>
</tr>
<tr>
<td>ERA-Interim</td>
<td>77</td>
<td>0.75°</td>
<td>6 h</td>
<td>188</td>
<td>1979–2012</td>
<td>Berrisford et al. (2011)</td>
</tr>
</tbody>
</table>
The CFSR temperature and precipitation data (Saha et al. 2010) are based on meteorological model reanalysis, which is a combination of a meteorological model field and surface observations and remote sensing data. At the time this study was conducted (April 2013) CFSR data were available for the years 1979–2010. The data from the ds093.1 dataset with a spatial resolution of 0.3125° and a temporal resolution of 6 h were used. For modeling, the 6-h values were converted to daily values. The CFSR land surface analysis precipitation is compiled using CPC daily rain gauge analysis and CPC merged 5-day precipitation analysis utilizing remote sensed data (Saha et al. 2010).

The ERA-Interim temperature and precipitation data are also based on a reanalysis of precipitation fields generated with a meteorological model (Berrisford et al. 2011). The ERA-Interim data have a spatial resolution of 0.75° and a 3-h temporal resolution. For modeling purposes, the 3-h values were converted to daily values. The ERA-Interim data are available for years 1979–2012. The ERA-Interim precipitation data are not rescaled using observation data.

For all datasets, the gridded precipitation data were set in the hydrological model to represent the precipitation in the middle of the data grid location. The precipitation value for each model grid cell was computed using the inverse distance squared weighted interpolation from three nearest precipitation data points. For the baseline case, the precipitation estimate was corrected for the elevation difference between the model grid location and the data grid location. The correction method was multiplicative, that is, the measured precipitation was multiplied with a coefficient of $1 + 0.2d$, where $d$ is the elevation difference between the data point and the model grid point in kilometers. The coefficient 0.2 used in the correction was obtained using calibration [see more in Lauri et al. (2012)]. For the gridded precipitation data, no elevation correction was used, as the gridded data value represents areal precipitation (instead of a point precipitation).

Gridded temperature was interpolated in the VMod model similarly as precipitation data (see above). For the baseline case, additive elevation correction with a constant lapse rate of 6 K km$^{-1}$ was used for the whole catchment area. For the gridded temperature datasets, similar elevation correction was used, where the elevation reference level in each grid cell was also selected so that the elevation correction did not change the average temperature within the grid cell.

d. Dataset performance assessment

In the dataset performance assessment, the hydrological model was run over the study period (1999–2005) using the data from the datasets described above. The results were then compared to measured discharges and results from the baseline computation. To obtain comparable results with the baseline, both gridded temperature datasets were bias corrected prior to the model runs, and each precipitation dataset was corrected to have similar weighted average discharge for five main stem stations. All other model parameters were kept the same as in the baseline run (see section 2b).

1) TEMPERATURE BIAS CORRECTION AND PRECIPITATION ADJUSTMENT

A bias correction has often been considered to be necessary when climate model or regional meteorological model data is used for hydrological modeling (see, e.g., Terink et al. 2010). A bias correction was required also in this case to correct for lower-than-measured surface temperatures in the reanalysis datasets. A spatially and temporally constant bias correction was conducted by adding a constant value to all temperatures of a dataset in question. The value was selected so that the average temperature of the whole catchment over the whole computation period was same as in the baseline model run.

After the temperature bias correction, the PET multiplier coefficient (PET mult.coeff.) was used to correct the difference in minimum and maximum temperatures, which the model uses in evaporation computation. The PET mult.coeff. was determined so that the simulated average weighted discharge of five discharge stations along the Mekong main stem was similar to measurement data. In the procedure, the discharges were weighted using average annual discharge of each station; for example, the lowermost discharge station had the largest weight. The discharge stations are shown in Fig. 1c, and average annual discharges for each station can be found in Table 3, found below.

After initial test runs with the hydrological model with different precipitation datasets, it was noted that some adjustment to the precipitation levels was also necessary to obtain comparable results for different datasets. The adjustment was done by multiplying the precipitation values for the whole catchment by a constant coefficient, determined using the same criteria as for the PET multiplier.

The method described above was used instead of a full model calibration in order to compare solely the dataset performance in modeling. Another possibility would have been to use full model calibration of all model parameters.
for each dataset. This would have resulted in better model performance results for each dataset, but part of the differences in the performances would have then been hidden in the changes in model parameterization.

2) PERFORMANCE ASSESSMENT AT BASIN SCALE

After the bias correction and the precipitation level adjustment, the performance of the datasets were evaluated using the NSE coefficients and the relative difference in average annual discharge between modeled and measured data for the five selected discharge stations in the Mekong main stem (see Fig. 1c) computed from daily discharge values. Additionally, the results conducted by gridded datasets were compared to baseline simulation results modeled with observed precipitation and temperature datasets.

For temperature dataset assessment, the observed precipitation dataset was used with the two gridded temperature datasets (CFSR and ERA-Interim). For precipitation dataset assessment, the CFSR temperature dataset was used with the TRMM v6, TRMM v7, CFSR, and APHRODITE precipitation datasets, while the ERA-Interim temperature dataset was used with the ERA-Interim precipitation dataset. The CFSR temperature dataset was used for the above-mentioned precipitation datasets as it was considered to be the better of the two gridded temperature datasets (as shown in section 3). The ERA-Interim temperature was thought to be most compatible with the ERA-Interim precipitation, which is the reason why they were used together.

3) PERFORMANCE ASSESSMENT AT SUBBASIN SCALE

At the Mekong basin scale, the hydrological model integrates precipitation data from a very large area, which dampens the effect of spatial variations in the source data on the catchment outflow. Therefore, the model performance was also assessed at the subbasin level. To evaluate the performance of the baseline and the two best-performing precipitation datasets (i.e., APHRODITE and TRMM v7; see Table 3 below for more details) at the subbasin scale, a difference of two adjacent discharge stations was computed from the observed discharge data and from the model results. These were then compared to each other using daily and monthly resolutions. The difference represents the lateral inflows in the river section between the two discharge measurement locations. A constant time lag (full days) was used in computing the differences in both measured and simulated discharges, to account for the travel time between the upstream and downstream stations. The lag was selected so that the resulting discharge difference time series had a minimum number of negative values.

3. Results

a. Temperature

The average temperature map of the baseline model run and the difference of that to the CFSR and ERA-Interim model runs are shown in Fig. 2. The difference maps (Figs. 2b,c) were computed by subtracting the interpolated average baseline temperature from the average temperature of CFSR or ERA-Interim model run results. Compared to the baseline (Fig. 2a), both datasets have regional differences in average temperatures that are mostly no further than 2°C from the baseline (Figs. 2b,c). The most notable temperature differences in both reanalysis datasets are in the southern part of the catchment, where both reanalysis datasets show temperatures that are higher than in the baseline model run. The mean temperature averaged over the whole basin was 22°C in the baseline (Table 2) and also for the CFSR and ERA-Interim datasets (as a result of bias correction).

The average temperature difference (daily maximum minus daily minimum) that is used in potential evaporation computation differs substantially between the three datasets (Fig. 3). In the baseline dataset, the temperature difference is smallest in the southern and middle parts of the basin, excluding the Thailand upland area (Fig. 3a). The temperature differences increase when moving northward and are largest in the northernmost part of the basin. The temperature difference computed using the CFSR dataset is more evenly distributed in the north–south direction (Fig. 3b). In the ERA-Interim data, the variation of temperature difference has rather a similar spatial pattern compared to the baseline dataset (Fig. 3c), except for the Thailand upland area, where the temperature difference is much smaller than in the baseline. The average temperature difference over the whole basin of the ERA-Interim dataset is over 2°C smaller than in the baseline, whereas for the CFSR it is rather close to the baseline value.

The evaporation computation parameterization was modified for each temperature dataset by adjusting the PET multiplier coefficient to obtain a similar level of cumulative discharge in the model. Resulting NSE values for discharge results computed as an average from the five main stem discharge stations are near each other (range 0.915–0.926) for the three temperature datasets (Table 2). The ERA-Interim dataset, however, required a rather large PET adjustment, which is due to the lower maximum–minimum temperature difference in the ERA-Interim dataset when compared to the baseline dataset.

Of the ERA-Interim and CFSR datasets, the CFSR dataset was selected as the more suitable data for modeling purposes. The main reason for that is the problematic
daily minimum and maximum temperature difference in the ERA-Interim data, which requires an overly large evaporation multiplier coefficient (see Table 2). Below, the CFSR dataset was used in further computations with gridded precipitation datasets, except for the ERA-Interim precipitation, with which the ERA-Interim temperature data was used.

b. Precipitation

Precipitation distributions of the reanalysis datasets (i.e., CFSR and ERA-Interim; see Figs. 4b and 4c, respectively) clearly differ from the baseline distribution (Fig. 4a). The CFSR dataset has an area of high annual precipitation in the lower-middle part of the catchment that is not present in the baseline data. In the case of ERA-Interim precipitation data, the area of high annual precipitation is concentrated in the middle part of the basin. Such an area of increased annual precipitation is not present in the baseline data. The precipitation distributions of the APHRODITE dataset based on ground measurements and the remotely sensed datasets TRMM v6 and TRMM v7 (Figs. 4d, 4e, and 4f, respectively) resemble that of the baseline dataset much better than the precipitation distributions of the reanalysis datasets. APHRODITE and TRMM v7 datasets are most similar to the baseline, while the TRMM v6 dataset deviates somewhat more, mostly in the southern part of the basin.

Interannual precipitation variation of each dataset after precipitation adjustment was investigated by computing the annual average precipitation for the catchment above Stung Treng (see location in Fig. 1c). The annual average precipitation values (Fig. 5) are remarkably similar for all datasets up to 2004, except for the CFSR dataset, which has a higher average precipitation. After 2005 the TRMM v6 and CFSR datasets indicate very low precipitation values, whereas TRMM v7, APHRODITE, and ERA-Interim dataset precipitations are at the same level as before 2005.

The similarity of annual values in the baseline, APHRODITE, and TRMM datasets are partly explained by the fact that all datasets utilize the same GTS

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Avg NSE</th>
<th>PET mult.</th>
<th>Avg temp (°C)</th>
<th>Avg dT (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.926</td>
<td>1.02</td>
<td>22.0</td>
<td>9.5</td>
</tr>
<tr>
<td>CFSR</td>
<td>0.923</td>
<td>1.00</td>
<td>22.0</td>
<td>10.3</td>
</tr>
<tr>
<td>ERA-Interim</td>
<td>0.915</td>
<td>1.35</td>
<td>22.0</td>
<td>6.9</td>
</tr>
</tbody>
</table>

![Fig. 2. Areal distribution of average annual temperature over the study period 1999–2005 for (a) baseline (i.e., based on surface observations), (b) difference of bias-corrected CFSR and baseline results, and (c) difference of bias-corrected ERA-Interim and baseline results.](image-url)
ground station data: baseline and APHRODITE directly and TRMM to adjust remotely sensed observations. However, compared to GTS data, the baseline and APHRODITE datasets contain additional ground station data points and should therefore be more accurate than the GSOD data used in the TRMM dataset.

c. Discharge at Mekong main stem stations

1) BASELINE

The performance of different datasets in discharge computation was assessed by comparing the modeled discharges to observed ones at five discharge stations along the main stem of the Mekong (Table 3). For the baseline model run, based on surface observations, the NSE coefficients are over 0.86 for all discharge stations. The modeled average flow exceeded the measured one by 16% at Nong Khai, and at Stung Treng it was 5% smaller than the measured discharge. In other stations, the computed average annual discharge was within 4% of the measured value (Table 3). The baseline run discharges at Chiang Saen and at Stung Treng are shown in Figs. 6a and 6b, respectively. Computed average monthly discharges over the calibration period, shown in Figs. 6c and 6d, were close to the measured values, though there was a consistent underestimation of discharges during the beginning of the wet season at Chiang Saen (May–July) and some underestimation of the peak flows (August–September) at Stung Treng.

2) GRIDDED DATASETS

When using the CFSR datasets, the results are acceptable with the NSE coefficients over 0.85 except for the Chiang Saen station (NSE = 0.661). The overall model efficiency is clearly lower than in the baseline model run (Table 3). Model results using ERA-Interim datasets show acceptable agreement to measurements only in the Mukdahan and Pakse stations (Table 3). This is most probably due to the problems in the ERA-Interim precipitation distribution (see Fig. 4b). The model run carried out with the APHRODITE precipitation dataset and the CFSR temperature data performs almost as well as the baseline, having the NSE values exceeding 0.86 for all assessed stations (Table 3). The ratio between observed and computed annual average flow is within the range from −5% to +13%, when for baseline the range is from −5% to +16%. Both TRMM model versions (with CFSR temperature dataset) perform well. TRMM v6 has NSE values exceeding or equal to 0.86 in all discharge stations with an average of 0.879, whereas the TRMM v7 NSE values exceed 0.88 in all stations and have an average of 0.932. The cumulative differences in discharge range from

![Areal distribution of average difference between daily maximum and daily minimum temperature (T_max - T_min) for (a) baseline, (b) bias-corrected CFSR, and (c) bias-corrected ERA-Interim datasets.](image-url)
−7% to +10% for the TRMM v6 data and from −7% to +5% for the TRMM v7 data.

d. Discharge difference between Mekong main stem stations

To assess the performance of the baseline and two best-performed datasets (i.e., APHRODITE and TRMM v7) at the subbasin scale, a difference of two adjacent discharge stations from the observed discharge data and from the model results was computed. The NSE coefficients for daily and monthly averaged discharge differences for computed and measured values are listed in Table 4. The NSE coefficients for the differences are lower than the NSE coefficients for the individual station discharges (Table 3). At the daily time step, the NSE coefficients and also the average discharge differences are rather similar for all datasets, and no single dataset clearly outperforms the others. The discharges computed using the APHRODITE dataset have slightly better average NSE values than discharges computed using either the baseline or the TRMM v7 datasets (Table 4).

The use of monthly average discharge values for assessing the performance improves the NSE values for all model runs (Table 4). The largest improvement is obtained in the TRMM v7 dataset model run. Monthly data are, however, most consistent in the model run using the
The APHRODITE and TRMM v7 datasets have better model efficiencies than the baseline dataset, especially for monthly data. It seems that both datasets bring additional information (e.g., more accurate precipitation distribution) to the computation when compared to the baseline dataset, thus improving the results, especially in areas with sparse surface observation network.

The river section with highest model efficiency is the section above Chiang Saen, whereas the lowest efficiency is obtained for the section between Chiang Saen and Nong Khai. The monthly average discharges/discharge differences of these highest and lowest performing sections are shown in Figs. 7a and 7b, respectively.

Relative differences of computed discharges to measured data are shown in Fig. 8 and Table 4 for the river sections between the five main stem stations. These differences seem to be correlated between the different datasets; for example, all datasets have either positive or negative difference in a given river section. The river sections Nong Khai–Mukdahan and Pakse–Stung Treng have lower simulated cumulative discharges, whereas the Chiang Saen–Non Khai and Mukdahan–Pakse sections have higher simulated discharges when compared to measurements. Both sections with lower simulated discharges obtain most of the runoff from areas of high annual precipitation and relatively sparse observation network near the eastern border of the Mekong catchment.

### 4. Discussion and conclusions

In this study, we assessed the performance of various global and regional gridded temperature and precipitation datasets in discharge modeling in a monsoon-driven large river basin, the Mekong basin. This study aims to facilitate the use of open datasets for hydrological modeling not only in the Mekong basin but also further in the Monsoon Asia region, thereby contributing to improvement of water resources–related planning and impact assessment work in the area.

The assessed data included reanalysis temperature and precipitation from CFSR and ERA-Interim datasets, rain gauge–observed precipitation from the APHRODITE dataset, and remotely sensed precipitation from the TRMM datasets (3B42, versions 6 and 7). A constant bias correction of temperature and adjustments for potential evapotranspiration and precipitation levels were used to obtain comparable results for the different datasets.

We found that the two best-performing combinations of the tested datasets comprising local daily ground observations of temperature and precipitation (APHRODITE + CFSR; TRMM v7 + CFSR) resulted in a similar accuracy
of computed daily discharges at the most downstream discharge point (NSE values of 0.948 and 0.938 in Stung Treng, respectively) than the baseline data (NSE value of 0.964). The best-performing precipitation datasets also had similar spatial distribution of annual precipitation than the baseline dataset (Fig. 4).

Reanalysis temperature data were found to be suitable for potential evapotranspiration computation after a bias correction. Average temperatures in both reanalysis datasets (CFSR and ERA-Interim) were accurate enough for modeling purposes; however, the difference of daily minimum and maximum temperatures, used in the VMod model for evapotranspiration computation, was more realistic in the CFSR dataset (compared to ground observations). Therefore, the CFSR data were considered as the better alternative in this case. Both reanalysis temperature datasets resulted in slightly too high temperatures in the Mekong floodplain area in Cambodia. There are, however, only a few temperature measurement points in that area (see Fig. 1b), and it is therefore likely that the measurement data might also be somewhat inaccurate. Our results about the suitability of CFSR temperature dataset for potential evapotranspiration computation using the Hargreaves–Samani method agree with those obtained by Sperna Weiland et al. (2012).

Existing studies show that for large catchments with sparse observation networks, the performance of TRMM precipitation in discharge modeling, when compared to rain gauge data, varies from case to case (Collischonn et al. 2008; Gu et al. 2010; Sisen and Sandholt 2010). For the Mekong catchment, the TRMM v7 dataset seems to yield equally good modeled discharge results as by using ground observed precipitation data. The model fit for daily discharge values along the main stem of the river had NSE coefficients exceeding 0.8 for all assessed monitoring stations. Results for the subbasin evaluation using the

---

Fig. 6. Comparison of (a),(b) daily and (c),(d) monthly average computed baseline flow results (i.e., using daily ground observations) and measured flows at Chiang Saen and Stung Treng stations during 1999–2005. See location of the stations in Fig. 1c.
difference between two main stem discharge measurement stations did not reach a similar level of accuracy (NSE coefficients were in the range 0.6–0.9). This indicates that the TRMM might not perform that well in smaller catchments, but it is also likely to be related to limitations of the hydrological model. This needs to be assessed in more detail in future studies.

The APHRODITE dataset has been used with good results by Vu et al. (2012) for discharge modeling in Vietnam. Our findings agree with their results. Similarly to the results by Vu et al., our findings indicate that the average precipitation in the APHRODITE dataset seems to be somewhat lower than that obtained from ground observations.

Precipitation data obtained from CFSR and ERA-Interim datasets had problems with spatial precipitation distribution and consequently did not give accurate results when used for discharge modeling. Moreover, the ERA-Interim data did not produce the precipitation pattern properly. The CFSR data were better in this respect, but they had several areas of heavy rainfall not present in the data observed on the ground. The heavy precipitation areas in the CFSR data were associated with orographic changes. Also, Lorenz and Kunstmann (2012) noted

<table>
<thead>
<tr>
<th></th>
<th>Daily</th>
<th>Monthly avg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>APHRODITE</td>
</tr>
<tr>
<td>Above Chiang Saen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qm (m³ s⁻¹)</td>
<td>2701</td>
<td>0.895</td>
</tr>
<tr>
<td>ΔQ/Qc</td>
<td>1.04</td>
<td>1.12</td>
</tr>
<tr>
<td>Chiang Saen–Nong Khai</td>
<td>2030</td>
<td>0.590</td>
</tr>
<tr>
<td>Qm/Qc</td>
<td>1.32</td>
<td>1.15</td>
</tr>
<tr>
<td>Nong Khai–Mukdahan</td>
<td>4265</td>
<td>0.859</td>
</tr>
<tr>
<td>Qm/Qc</td>
<td>0.82</td>
<td>0.77</td>
</tr>
<tr>
<td>Mukdahan–Pakse</td>
<td>1925</td>
<td>0.794</td>
</tr>
<tr>
<td>Qm/Qc</td>
<td>1.14</td>
<td>1.29</td>
</tr>
<tr>
<td>Pakse–Stung Treng</td>
<td>4260</td>
<td>0.823</td>
</tr>
<tr>
<td>Qm/Qc</td>
<td>0.75</td>
<td>0.78</td>
</tr>
<tr>
<td>Avg NSE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg [Qm/Qc – 1]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4. Efficiency coefficients and average annual discharge for discharge difference of adjacent discharge stations from three model runs using different precipitation datasets.

Fig. 7. (a) Monthly average discharges at Chiang Saen discharge station and (b) monthly average of discharge difference between Chiang Saen and Nong Khai discharge stations for model runs using baseline, APHRODITE, and TRMM v7 datasets. See also Table 4.
that the performance of state-of-the-art reanalysis precipitation products for hydrological modeling in tropical regions is not very good, and our findings support this statement. The field of reanalysis is, however, constantly improving, so new reanalysis datasets will hopefully improve the situation and should be tested as they become available.

Possible sources of errors in this study are related to errors in source data, limitations in the data interpolation, and limitations of the hydrological model. A somewhat unexpected result is that the errors in discharge for different datasets for subcatchments (Fig. 8) seem to be correlated. There are at least two possible explanations: 1) the datasets presented in Fig. 8 are not independent and 2) the hydrological model has errors in the distribution of precipitation or in forwarding the runoff to correct direction. The first point is obviously correct, as both the baseline and APHRODITE datasets are based on surface observations, and the TRMM 3B42 dataset is corrected using monthly average precipitation data from ground observations (GSOD dataset). Improvement in this respect is possible by utilizing better interpolation and data correction algorithms that take into account orographic effects, strongly variable precipitation, and long distances from precipitation observation to the interpolation points. The second point is also relevant; for example, the relatively large grid size of the model introduces errors in subcatchment boundaries as well as in land use, soil, and slope information in the model. The reasons behind these correlated errors are left as a subject for future studies.

Our findings suggest that easily accessible temperature and precipitation products covering large regions can yield a solid performance in hydrological modeling in the Mekong basin. According to our results, we recommend the use of APHRODITE or TRMM v7 precipitation datasets and the CFSR temperature data. The two most suitable datasets have their own particular strengths. The APHRODITE data are suitable for hydrological studies extending over longer time periods, as they are available from year 1951 onward, whereas the TRMM v7 data are suitable for near-real-time assessments or even short-term forecasting. When used with CFSR temperature data, the temporal extents of the datasets limit the use of the APHRODITE data to the period of 1979–2007 and TRMM v7 to 1998–2013. The CFSR dataset is complemented with near-real-time NCEP Climate Forecast System version 2 (CFSv2; Saha et al. 2013), and thus, TRMM v7 together with a combined use of CFSR and CFSv2 can support near-real-time hydrological modeling and hydrological forecasting within the limits defined by the above-mentioned data. ERA-Interim data are also updated regularly, but at the time of writing (April 2012), it was not available in real time.

Acknowledgments. This work was funded by Maa- ja vesiteknikan tuki ry. In addition to this, T.A.R. received funding from RYM-TO doctoral school and M.K. received funding from Aalto University postdoc funds and the Academy of Finland project SCART (Grant 267463). We thank our colleagues at Aalto University Water & Development Research Group. We are grateful to Teemu Kokkonen from Aalto University for his support and helpful comments and to EIA’s personnel Markku Virtanen and Jorma Koponen for their support. The comments and suggestions of the editor and the four reviewers are very much acknowledged.

REFERENCES


