Predictions in Ungauged Basins: Promise and Progress (Proceedings of symposium S7 held during the Seventh IAHS Scientific Assembly at Foz do Iguaçu, Brazil, April 2005). IAHS Publ. 303, 2006.

Evaluation of a grid-based distributed hydrological model over a large area

VICTOR KOREN, FEKADU MOREDA, SEANN REED, MICHAEL SMITH & ZIYA ZHANG

Office of Hydrologic Development, National Weather Service, 1325 East-West Highway, Silver Spring, Maryland 20910, USA

Victor.Koren@noaa.gov

Abstract In this paper we present an on-going effort to utilize available soil moisture data. This study evaluates the performance of a distributed hydrological model using runoff and soil moisture over 75 basins with watershed areas varying from 20 km^2 to 15 000 km^2. These basins are selected in a region where unique soil moisture data of the Oklahoma Mesonet are available. While simulated runoff is compared to measured streamflow at a basin outlet, simulated soil moisture is compared to basin average soil moisture derived from Oklahoma Mesonet observations. Our results show that the modified Sacramento model driven by *a priori* parameters performs reasonably well and allows explicit estimation of soil moisture at desired layers. Annual, monthly, and 10-day runoff volumes are found in good agreement with observed data for a range of spatial scales. Simulated and observed soil moisture of the 0–25 cm layer agrees well with a slight (9%) negative bias. However, 25–75 cm layer soil moisture shows a significant (26%) negative bias for most watersheds located in a dry region with $P/PE < 0.8$.

Key words distributed model; prediction; runoff; soil moisture; space-time averaging

INTRODUCTION

One of the biggest challenges of distributed modelling is the prediction of streamflow over a range of spatial scales, e.g. at basin interior locations. To address this challenge, a distributed model should reasonably well represent the heterogeneities of watershed properties through its model structure and parameters. Unfortunately, spatial data limitations reduce model evaluation to a simple comparison of modelled and observed streamflow at the gauged outlet (Reed *et al*., 2004) and greatly impede an evaluation of the spatial correctness of model parameters.

In addition to the scarcity of spatial data, many hydrological models do not represent watershed states such as soil moisture states but rather soil water storages/indices which also limit comparison of model simulation to available data. In this regard for example, Robock *et al*. (2004) compared observed soil moisture storage to conceptual model storage in the North American Land Data Assimilation System (NLDAS) project. Experiencing the problem of comparing model outputs at specific points such as Mesonet sites, they averaged soil moisture observations of 72 Oklahoma Mesonet stations and compared these values to regional model outputs. Similarly, Schaake *et al*. (2004) also merged 17 soil moisture stations in the state of Illinois, USA to compare observed total 2-m water storage and simulated total water storage of several models. While such merging of point measurements reduces data errors, it does not allow full utilization of the soil moisture measurements and evaluation of model performance over a range of spatial scales.

In this paper we present an on-going effort to fully utilize available spatially variable soil moisture data for distributed model development. This study evaluates the performance of a distributed hydrological model using runoff and soil moisture over 75 basins with watershed areas varying from 20 km^2 to 15 000 km^2. These basins are selected in a region where unique soil moisture observations are available from the Oklahoma Mesonet (Brock *et al.*, 1995). While simulated basin average runoff is compared to measured streamflow at basin outlets, basin averages of simulated soil moisture and soil moisture derived from Oklahoma Mesonet observations are compared. An extended version of the Sacramento Soil Moisture Accounting Model with an added Heat Transfer component (referred to as SAC-SMAHT) is used. This version allows linking a conceptual storage-type model states to actual soil moisture data, and so helping in finding connection between physically based and conceptual models (Robinson & Sivapalan, 1995). Furthermore, SAC-SMAHT has a number of practical advantages such as data assimilation using soil moisture and temperature measurements, and water balance simulations under frozen ground conditions.

DISTRIBUTED MODELLING SYSTEM

System structure

The Hydrology Laboratory Research Modelling System (HL-RMS) (Koren *et al.*, 2003, 2004) is used in this study. The HL-RMS is defined on a regular rectangular grid. The system consists of a water balance, hillslope routing and channel routing components. A number of conceptual hillslopes at each grid cell are defined to make overland flow distances physically realistic for the relatively large cell size (~16 km^2). A drainage density parameter is used to subdivide a cell into equally sized overland flow planes. Conceptual hillslopes drain water to a conceptual channel within the same grid cell. A conceptual channel usually represents the highest order stream of a selected grid cell. It is assumed that all hillslopes have the same properties inside each grid cell but they may be different from cell to cell. Cell-to-cell channel routing is done using a flow direction grid. To facilitate efficient routing calculations, the drainage network is translated into a computational sequence of grid cells in an upstream to downstream order. Fast response runoff from the water balance model is routed over conceptual hillslopes within each cell to a conceptual channel. Slow response runoff is assumed to enter the channel system directly from the soil and therefore bypass the hillslope routing. There is no physical connection between soil moisture states in adjacent grid cells. The conceptual channel is the only source of water exchange between neighboring pixels.

Water balance and routing models

The Sacramento Soil Moisture Accounting Model (SAC-SMA) (the water balance component of HL-RMS) was modified to transform conceptual soil moisture storage into soil moisture states of a soil profile. Koren *et al.* (2002) developed a set of physical relationships that link the SAC-SMA storages (parameters) and soil properties such as porosity, field capacity, wilting point, and hydraulic conductivity. They assume that tension water storages of the SAC-SMA model are related to available soil water, and that free water storage is related to gravitational soil water. These relationships allow recalculation of the upper and lower soil moisture capacities into soil moisture contents at a number of soil layers. Five layer depths are defined *a priori* to cover a 2 m soil profile with thinner layers closer to the soil surface. However, an actual number of soil layers and their thicknesses are automatically adjusted using SAC-SMA parameter values. To make this adjustment, the upper and lower zone depths are estimated first to be sure that the upper and lower SAC-SMA capacities are preserved. *A priori* defined layer depths are then adjusted to be consistent with these estimates. Because of this, the number of soil layers may be less than five, and can be different for different pixels. For more detail on this procedure see Koren *et al.* (2002) and Koren (this issue). At each time step, SAC-SMA liquid water storage changes due to rainfall/snowmelt are computed, and then transformed into soil moisture states of the heat transfer model. The heat transfer model (Koren *et al.*, 1999) calculates the temperature of each soil layer. Consequently, based on the simulated soil temperature profile, the total water content is split into frozen and liquid water portions. Estimated new soil moisture states are then converted back into SAC-SMA model storages. This new version (SAC-SMAHT) also accounts for the frozen ground effect on runoff (Koren, 2006).

Hillslope and channel routing uses the kinematic wave model. A fairly general numerical scheme that provides the unconditional stability is used (Koren *et al.*, 2004). We note that truncation errors of the scheme increase independently of the space–time increment ratio allowing a flexible selection of space–time increments to compensate for some accuracy reduction.

TEST REGION AND DATA

The tests were performed on 75 watersheds (with areas ranging from 20 km^2 to 15 000 km^2) within the Arkansas-Red River basin in Oklahoma, as shown in Fig. 1. With a total drainage area of 409 300 km^2, the basin encompasses a wide variety of climatic conditions, ranging from an arid region in the western part to a humid region in the eastern part. The test region has a relatively good data set to evaluate our distributed model. This region has the longest archive of 4-km NEXRAD-based multi-sensor precipitation grids, and these rainfall estimates have been evaluated thoroughly. United States Geological Survey (USGS) streamflow measurements are available at each selected basin outlet. Also, the test region has a unique soil moisture data collection network, the Oklahoma Mesonet.

The Oklahoma Mesonet provides real-time data including soil moisture measurements at four depths (5, 25, 60, and 75 cm) from more than 100 sites since 1997. However, only 64 sites provide measurements at all four depths. All sites are equipped with heat dissipation soil moisture sensors which measure the temperature change of a heat pulse (Brock *et al.*, 1995). In this study, the 30 min volumetric soil moisture data are resampled at the top of the hour. Then, the daily mean values of soil

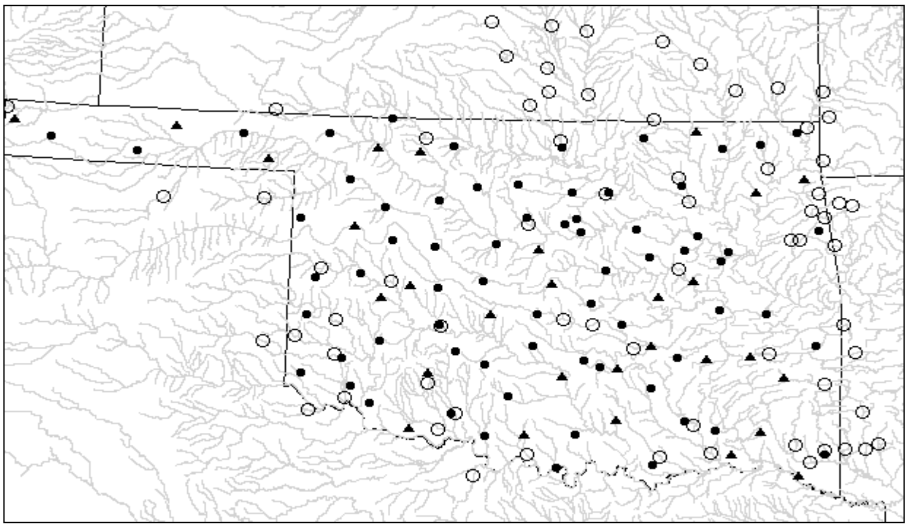

Fig. 1 A map of the Oklahoma Mesonet soil moisture sites (shared circles for four layer measurements, and triangles for only two top layer measurements) and test basin outlets (open circles).

moisture at each station are computed for the time period 1 January 1997 to 31 December 2002.

There are two issues to consider while using the volumetric soil moisture data from the Mesonet sites. First, the instantaneous absolute soil moisture measurement at a station is related to the soil type and the physiographic properties of the location in addition to the availability of moisture supply i.e. precipitation in the area. This hampers comparisons of stations located in different areas even during similar weather conditions. Secondly, hydrological model states and volumetric soil moisture measurements may not have a one-to-one correspondence; therefore one may not be able to compare these two quantities objectively. To reduce the impacts of these issues, the volumetric soil moisture is converted into a saturation ratio (*SR*) by using the soil properties at each station. *SR* is defined as:

$$SR = \frac{\theta - \theta_r}{\theta_s - \theta_r}$$

where θ is a volumetric water content ($m^3\ m^{-3}$), θ_s is the saturation volumetric water content ($m^3\ m^{-3}$), and θ_r is a residual volumetric water content ($m^3\ m^{-3}$). $SR = 0$ corresponds to dry soil conditions while $SR = 1$ corresponds to saturation or wet soil conditions. The saturation ratio attempts to reduce the effects of the individual soil property variation for intercomparison as well as generating soil moisture maps.

Our analyses are performed for weighted averages of soil moisture over two soil layers: the top 0–25 cm layer, and the deeper layer (25–75 cm). For each layer, point saturation ratio values are interpolated to 4 km grid cells for the entire Oklahoma state using an inverse distance weighting method. Weights are computed on a daily basis depending on station locations with available data at a given day. Later, the gridded daily maps of SR have been used to generate time series of basin average soil moisture. The range of the soil property variability over the Oklahoma Mesonet is shown in Table 1.

Table 1 The range of soil properties at four depths for the Oklahoma Mesonet.

Depth, cm	Residual moisture, θ_r	Saturation, θ_s
5	0.0–0.250	0.399–0.660
25	0.0–0.268	0.240–0.493
60	0.0–0.263	0.377–0.467
75	0.022–0.281	0.390–0.513

A continuous HL-RMS run for a 7-year period was performed at a 4 × 4 km grid using NEXRAD precipitation estimates. Because of minor effect of frozen ground in this region (for analysis of frozen ground effects see accompanying paper Koren (2006, this issue), the heat transfer component was not included meaning that water balance simulation results would be the same as from the original SAC-SMA model. However, the original version does not allow a direct comparison to measured soil moisture profiles. *A priori* soil-based SAC-SMA parameter grids over the coterminous USA (Koren *et al*., 2004) were used without any calibration. Rough estimates of channel and hillslope routing parameters from Koren *et al*. (2004) were applied to generate hydrographs at the selected watershed outlets. Both Oklahoma-region grids and time series of basin averages of water balance components were generated. To match to the observation soil layer thickness, SAC-SMAHT soil moisture contents at variable layers were recalculated into soil moisture at measurement soil layers. It is worth mentioning that Oklahoma Mesonet soil properties were not used in derivation of *a priori* parameters.

RESULTS AND DISCUSSION

Runoff analysis

First, we perform long-term (7 years) water balance tests with the model. Figure 2(a) shows that there is good agreement between simulated and observed annual average runoff. Another test is performed on the dependency of observed and simulated runoff on the climate. For this test, a climate index expressed as a ratio of mean annual precipitation to potential evaporation (P/PE) is calculated for all basins. A lower value of P/PE for a basin indicates the basin is dry while a higher value denotes the basin is wet. In this analysis, both simulated and observed annual runoffs display similar dependency on the climate index (Fig. 2(b)).

Since the routing model parameters were not verified, daily or shorter time step comparisons were not performed. Instead, to reduce the effect of routing uncertainties, runoff statistics for all watersheds were calculated using 10-day averaging time series. One can observe a high correlation of simulated and observed 10-day average time series for most watersheds with a few outliers which are located in the very dry western part of the region (Fig. 3(a)). A large range of simulation errors is observed for very small basins, Fig. 3(b). There is a tendency of decreasing simulation errors with increasing of basin size. However, measurable error reductions are observed for only large basins over 5000 km^2. Although the root mean square errors are usually lower for dryer basins, relative errors are much higher for these basins due to near zero

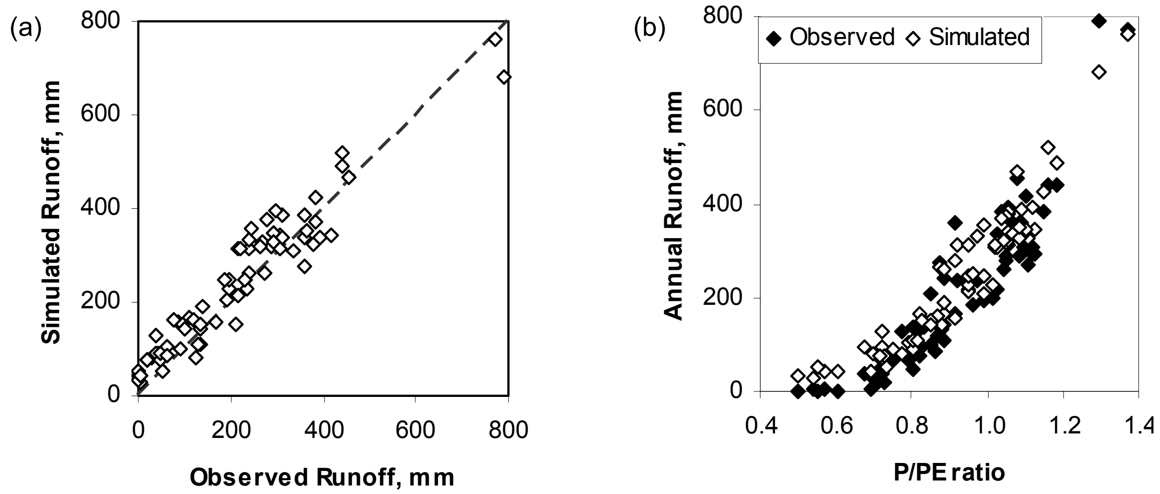

Fig. 2 Comparison of simulated and observed annual runoff for test basins: (a) simulated *vs* observed, (b) simulated and observed *vs* P/PE ratio.

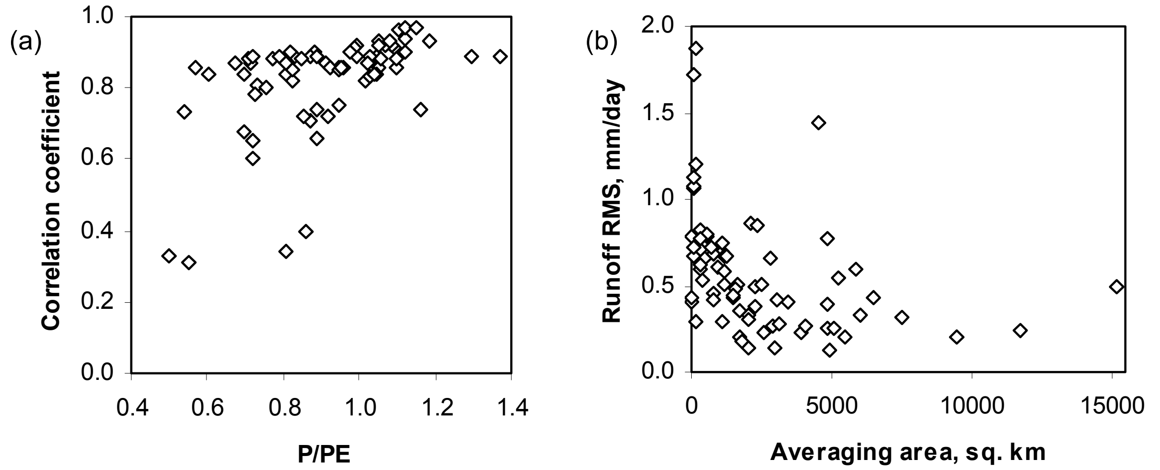

Fig. 3 Comparison of 10-day average simulated and observed runoff statistics for test basins: (a) correlation coefficients, (b) root mean square errors (RMS).

observed runoff. The model usually overestimates runoff for such basins. Perhaps one reason that these dry basins experience extremely low runoff at the outlet is that there are significant channel losses to the groundwater. Channel losses are not accounted by the model used in this study. Uncertainties in *a priori* parameter estimates may have been another cause of the positive bias of simulated runoff.

Soil moisture analysis

Daily averages of observed and measured saturation ratios of 0–25 cm and 25–75 cm soil layers are compared for all test basins. The correlation coefficients between the daily observed and simulated soil moisture are shown in Fig. 4(a) and (b). High correlation between simulated and measured saturation ratios for the both soil layers can be seen in these figures. The 0–25 cm layer correlation does not display dependency on the climate index but there is a clear reduction in correlation for the 25–75 cm layer in dryer watersheds. We also computed the biases of the simulated and observed soil moisture. As shown in Fig. 4(c), the 0–25 cm layer biases do not show dependency on

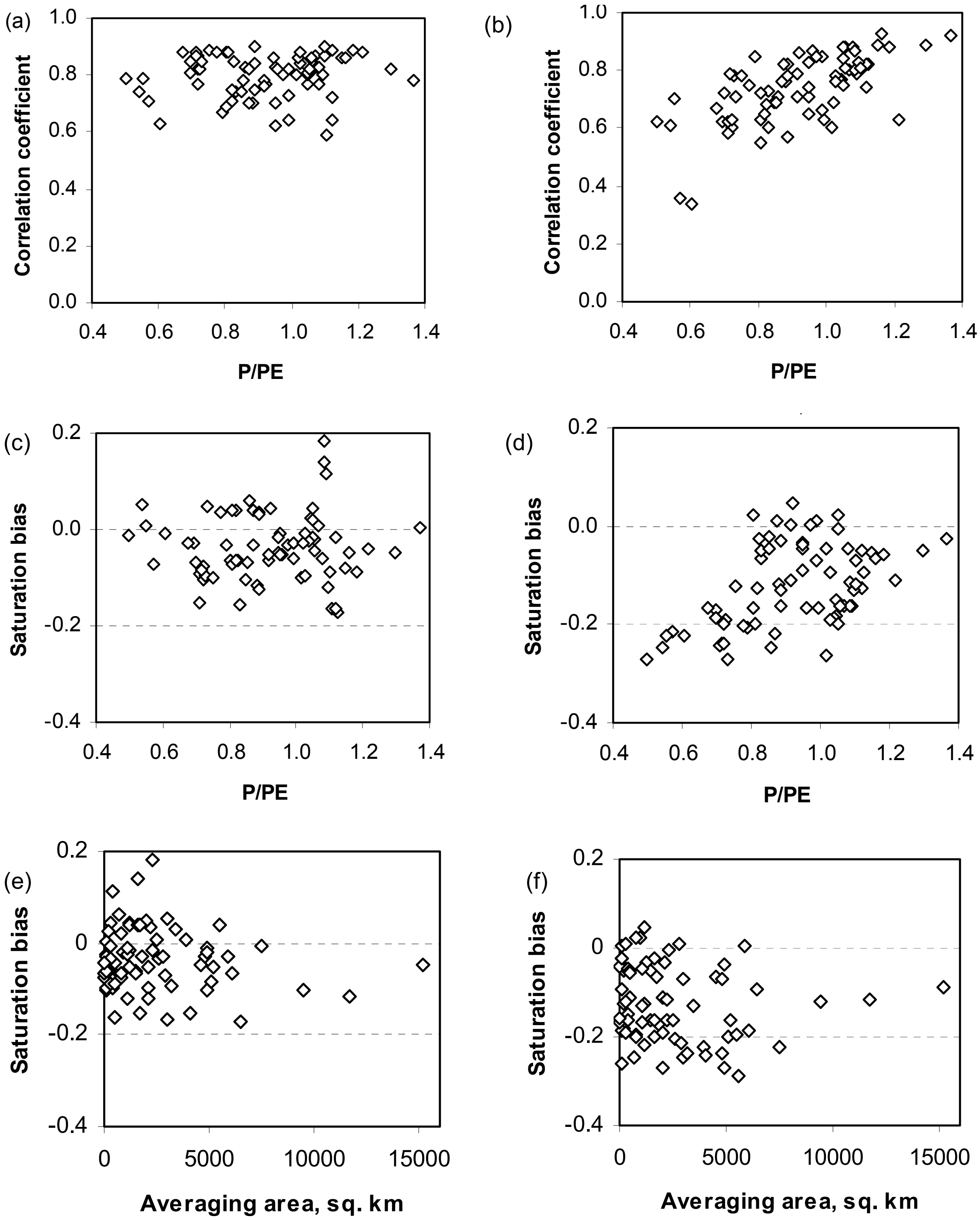

Fig. 4 Comparison of simulated and observed soil moisture statistics at the upper (left panels) and lower (right panels) soil layers for test basins: (a), (b) correlation coefficient, (c), (d) saturation ratio bias *vs* climate index, and (e), (f) saturation ratio bias *vs* averaging area.

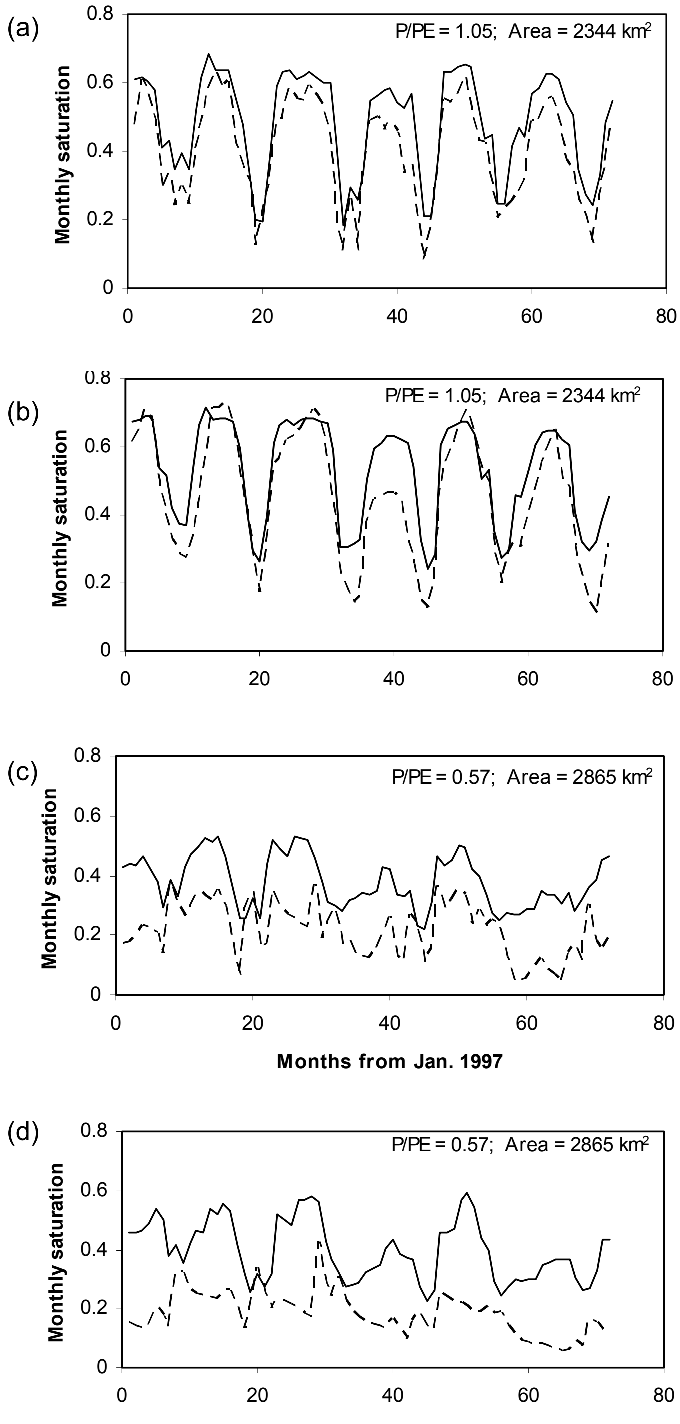

Fig. 5 Comparison of simulated (dashed lines) and observed (solid lines) monthly saturation ratios at 0–25 cm ((a), (b)) and 25–75 cm ((c), (d)) soil layers for selected wet ((a), (b)) and dry ((c), (d)) basins.

the climate index. However, the 25–75 cm layer biases show some dependency on the climate index similar to the correlation coefficient dependency seen in Fig. 4(b). The results also show little dependency of soil moisture bias on an averaging area, as can be seen from Fig. 4(e) and (f). Overall, there is a slight (9%) negative bias for the 0–25 layer. The 25–75 cm layer soil moisture shows up to 26% negative bias for watersheds located in the very dry western region but only slight bias similar to the 0–25 cm layer for wetter watersheds with the climate index above 0.8. Uncertainty of soil moisture measurements may contribute to this bias. Illston *et al.* (2004) compared soil moisture measurements at Mesonet sites with soil core samples at 5 cm and 25 cm during the enhanced drying phase. They observed a 2-fold decrease in the range of soil water content values at 25 cm from Mesonet data (0.23 $m^3 m^{-3}$) compared to collected soil cores (0.44 $m^3 m^{-3}$). As a result, significant positive bias of Mesonet sensors at the dry end of the spectrum was experienced. Simulation results from the Sacramento model are consistent with these findings. As can be seen from Fig. 5, simulated soil moisture dynamics agree well with Mesonet measurements, however, simulation accuracy decreases significantly for very dry watershed (P/PE = 0.57). Another cause of simulation errors may be discrepancies in the definition of soil properties from measurement sites and STATSGO data at modelled pixels (Robock *et al.*, 2003). In addition, there was no use of climate characteristics in the derivation of *a priori* parameters.

SUMMARY

The modified Sacramento model driven by *a priori* parameters performs reasonably well and allows explicit estimation of soil moisture at desired layers. Annual, monthly, and 10-day runoff volumes are in good agreement with observed data for a range of spatial scales. Soil moisture dynamics are consistent with measurements for all soil layers with correlation coefficients above 0.6. Simulated and observed soil moisture of the top layer (0–25 cm) agrees well with a slight (9%) negative bias. However, deeper layer (25–75 cm) soil moisture has a significant (26%) negative bias for most watersheds located in the dry western region with P/PE < 0.8. The spatial averaging scale is not the dominant contributor to simulated soil moisture and runoff errors; the major factor is climate represented by the ratio of precipitation to potential evapotranspiration.

A priori parameter uncertainties, soil moisture measurement and interpolation errors, and channel losses may cause soil moisture biases. Further analyses to evaluate and improve model performance include: (a) establishing relationships between model parameters and climate characteristics by calibrating the model for selected watersheds in different climates; (b) testing the effects of the model structure and physics, specifically the evaporation component from the lower zone, and effects of channel losses on overall balance and soil moisture states; and (c) using more reliable sources of soil moisture measurements.

REFERENCES

Brock, F. V., Crawford, K. C., Elliott, R. L., Cuperus, G. W., Stadler, S. J., Johnson, H. & Eillts, M. D. (1995) The Oklahoma Mesonet: A technical overview. *J. Atmos. Oceanic Technol.* **12**, 5–19.

Illstone, B. G., Caldwell, J. & Bodnar, S. G. (2004) Representativeness of soil moisture conditions in central Oklahoma during the enhanced drying phase. In: *Applied Climatology* (Proc. of the 14th Conf., AMS, Seattle, Washington, USA).

Koren, V., Schaake, J., Mitchell, K., Duan, Q.-Y., Chen, F. & Baker, J. M. (1999) A parameterization of snowpack and frozen ground intended for NCEP weather and climate models. *J. Geophys. Res.* **104**(D16), 19569–19585.

Koren, V., Smith, M. & Duan, Q. (2002) Use of *a priori* parameter estimates in the derivation of spatially consistent parameter sets of rainfall-runoff models. In: *Calibration of Watershed Models, Water Science and Application 6* (ed. by Q. Duan, S. Sorooshian, H. Gupta, A. Rosseau & R. Turcotte), 239–254. AGU, Washington DC, USA.

Koren, V., Reed, S., Smith, M. & Zhang, Z. (2003) Combining physically-based and conceptual approaches in the development and parameterization of a distributed system. In: *Weather Radar Information and Distributed Hydrological Modelling* (Proc. Symposium HS03 at Sapporo, July 2003), 101–108. IAHS Publ. 282. IAHS Press, Wallingford, UK.

Koren, V., Reed, S., Smith, M., Zhang, Z. & Seo, D.-J. (2004) Hydrology laboratory research modelling system (HL-RMS) of the US national weather service. *J. Hydrol.* **291**, 297–318.

Reed, S., Koren, V., Smith, M., Zhang, Z., Moreda, F. & Seo, D.-J. (2004) Overall distributed model intercomparison project results. *J. Hydrol.* **298**, 27–60.

Robinson, J. S. & Sivapalan, M. (1995) Catchment scale runoff generation model by aggregation and similarity analysis. In: *Scale Issues in Hydrological Modelling* (ed. by J. D. Kalma & M. Sivapalan), 311–330. Wiley, New York, USA.

Robock, A., Lue, L., Wood, E. F., Wen, F., Mitchell, K. E., Houser, P. R., Schaake, J. C., Lohmann, D., Cosgrove, B., Sheffield, J., Duan, Q., Higgins, R. W., Pinker, R. T., Tarpley, D., Basara, J. B. & Crawford, K. C. (2004) Evaluation of the North American Land Data Assimilation System over the Southern Great Plains during the warm season. *J. Geophys. Res.* **108**(D22), 8846, GCP 7.1–7.21.

Schaake, J. C., Duan, Q., Koren, V., Mitchell, K. E., Houser, P. R., Wood, E. F., Robock, A., Lettenmaier, D. P., Lohmann, D., Cosgrove, B., Sheffield, J., Luo, L., Higgins, R. W., Pinker, R. T. & Tarpley, J. D. (2004) An Intercomparison of soil moisture fields in the North American Land Data Assimilation System (NLDAS). *J. Geophys. Res*. **109**(D1), D01S90 1–16.