A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions

Ben Livneh

Eric A. Rosenberg

Chiyu Lin

Bart Nijssen

Vimal Mishra

See next page for additional authors

Authors

Ben Livneh, Eric A. Rosenberg, Chiyu Lin, Bart Nijssen, Vimal Mishra, Kostas M. Andreadis, Edwin P. Maurer, and Dennis P. Lettenmaier

A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions*,#

BEN LIVNEH

Cooperative Institute for Research in Environmental Science, University of Colorado Boulder, Boulder, Colorado

ERIC A. ROSENBERG, CHIYU LIN, BART NIJSSEN, AND VIMAL MISHRA[+]

Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington

KOSTAS M. ANDREADIS

Jet Propulsion Laboratory, Pasadena, California

EDWIN P. MAURER

Department of Civil Engineering, Santa Clara University, Santa Clara, California

DENNIS P. LETTENMAIER

Cooperative Institute for Research in Environmental Science, University of Colorado Boulder, Boulder, Colorado

ABSTRACT

This paper describes a publicly available, long-term (1915–2011), hydrologically consistent dataset for the conterminous United States, intended to aid in studies of water and energy exchanges at the land surface. These data are gridded at a spatial resolution of $1/16°$ latitude/longitude and are derived from daily temperature and precipitation observations from approximately 20 000 NOAA Cooperative Observer (COOP) stations. The available meteorological data include temperature, precipitation, and wind, as well as derived humidity and downwelling solar and infrared radiation estimated via algorithms that index these quantities to the daily mean temperature, temperature range, and precipitation, and disaggregate them to 3-hourly time steps. Furthermore, the authors employ the variable infiltration capacity (VIC) model to produce 3-hourly estimates of soil moisture, snow water equivalent, discharge, and surface heat fluxes. Relative to an earlier similar dataset by Maurer and others, the improved dataset has 1) extended the period of analysis (1915–2011 versus 1950–2000), 2) increased the spatial resolution from $1/8°$ to $1/16°$, and 3) used an updated version of VIC. The previous dataset has been widely used in water and energy budget studies, climate change assessments, drought reconstructions, and for many other purposes. It is anticipated that the spatial refinement and temporal extension will be of interest to a wide cross section of the scientific community.

---

1. Introduction

Increased computational capabilities, the availability of new data sources such as remote sensing, and better understanding of the Earth system have resulted in considerable improvements in the ability to represent long-term variations in land surface water and energy fluxes and state variables. Earth system models require, among other things, consistent observational datasets for model testing and diagnosis. Furthermore, predictions of alternative future scenarios of land surface conditions resulting from changes in climate and/or land cover

require benchmark historical data against which to evaluate.

We describe an observational dataset (herein called L13) that provides a means to analyze and verify hydroclimatic predictions as well as to drive land surface models, along with fluxes and state variables from the variable infiltration capacity (VIC) model. The L13 dataset is based on the methods of Maurer et al. (2002; hereafter M02), who developed a set of publicly available gridded meteorological data from ground-based measurements, together with model-derived hydrologically consistent surface fluxes and states. M02 spanned the period 1 January 1949–31 July 2000 (~51.5 yr) at a ⅛° latitude/longitude spatial resolution. The L13 dataset described here refines the spatial resolution to ¹⁄₁₆°, extends the period of record backward to 1 January 1915 and forward to 31 December 2011 (97 yr), and provides fluxes and states from an updated version of the VIC land surface model. The significance of each of these aspects is described below, followed by a summary of evaluations of the new dataset relative to M02.

An examination of the most widely cited studies that reference and/or use the M02 data (in total, over 300 Web of Science citations; http://wokinfo.com/) suggests that the applications can be grouped into three general areas: 1) studies that use the meteorological and hydrological data directly to characterize the state or variability of a specific hydroclimatic variable (e.g., temperature, precipitation, and snowpack), 2) studies that use the data as a spatially and temporally complete observational baseline for downscaling climate model output (especially for bias correction) to generate future climate scenarios, and 3) water and energy balance studies, for which the model forcings and derived fluxes are of particular interest because the derived surface water and energy budgets close at all grid cells at each time step by construct. Examples of each of these applications are discussed below.

Westerling et al. (2006) used the M02 gridded meteorological data along with other sources to isolate the signal of climatic variability on wildfire frequency in the western United States. Hayhoe et al. (2007) used the archived hydrological fluxes and states to represent historical hydrologic conditions from which future meteorological scenarios were assessed via hydrologic simulations in the northeastern United States. Soil moisture data were used by Castro et al. (2007) to initialize a regional climate model to simulate U.S. climatology. Sheffield et al. (2004) used M02 soil moisture to derive a hydrologically based drought index, which showed good agreement with time series of U.S. drought from two Palmer drought severity index (PDSI) datasets.

The second group of studies typically follows procedures that include correcting downscaled climate model output with the M02 forcing data, and then using the bias-corrected model to produce future climatic scenarios. Cayan et al. (2008) and Hayhoe et al. (2004) both exploited the downscaled climate model outputs to assess climatic implications of future greenhouse gas emission scenarios in California using this general approach. Loarie et al. (2008) used downscaled outputs to examine impacts on the diversity of flora in California. Salathé (2003) downscaled climate model outputs to simulate streamflow over a river basin in Washington State and also evaluated performance of several climate models after bias correction. Wood et al. (2004) used the forcing and hydrological datasets to evaluate six bias correction methods for downscaling climate model outputs over the continental United States.

Among the third type of applications, Stewart et al. (2004) utilized the forcing data to simulate streamflow over large river basins in the Pacific Northwest. Smith et al. (2004) used the meteorological data to force a suite of land surface models and compared their performance. Christensen et al. (2004) applied the forcing data over the Colorado River basin to search for robust VIC model parameters over small river basins that were then used to assess climatic impacts under future forcing scenarios. Carpenter and Georgakakos (2004) utilized the energy forcing data to compute potential evapotranspiration (ET) in a radar–rainfall uncertainty study. Maurer et al. (2004) used climate data and the archived VIC-derived soil moisture, snow, and runoff data to examine predictability of runoff across the United States. Andreadis and Lettenmaier (2006) used the forcing data to evaluate a data assimilation scheme using satellite-based snow water equivalent (SWE) information.

Climate model outputs, remote sensing, and land cover data continue to become available at finer spatial resolutions, making the spatial refinement of L13 a significant improvement (from ⅛° to ¹⁄₁₆°). The extended period of record (from 50 to 97 yr) will help to improve the statistical strength of computed trends from hydroclimatic analyses and model corrections for downscaled climate model outputs, and it captures important historic extremes such as the 1930s drought that were outside the period of M02. Also, the climate model simulations for historic periods conducted as part of phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) extend through 2005, for which an extended observational dataset is useful for various purposes discussed below. Finally, an updated VIC model version (described below) includes refinements that will be of interest for some applications of the model-derived variables.
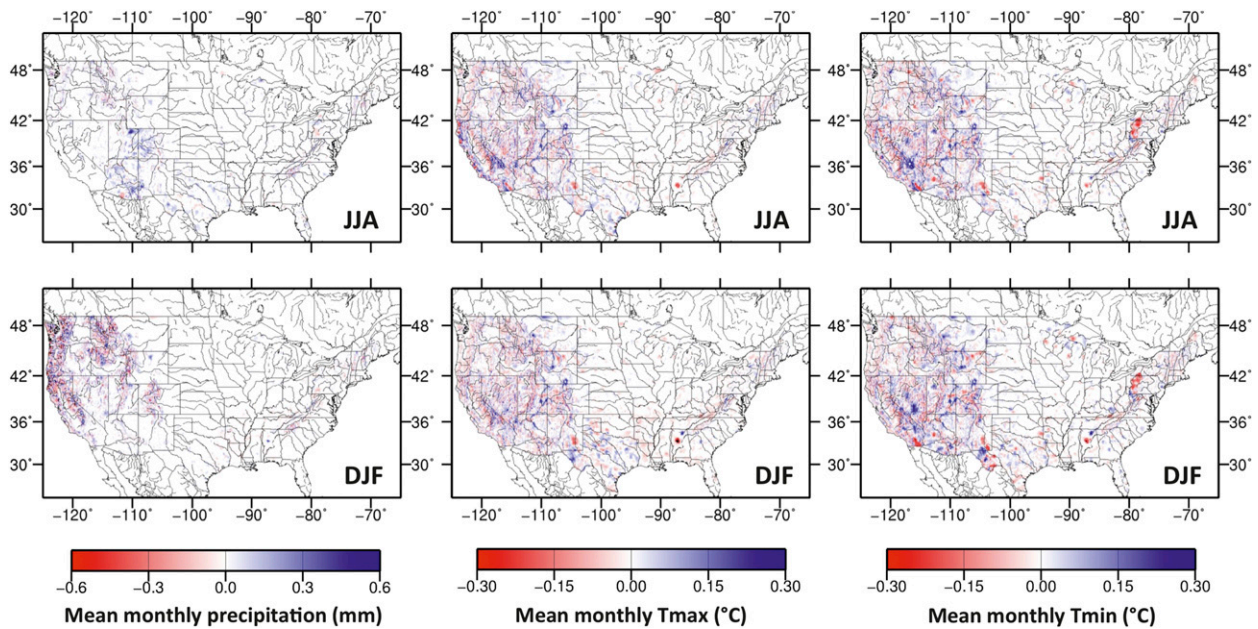
FIG. 1. L13 minus M02 for (top) summer [June–August (JJA)] and (bottom) winter [December–February (DJF)] mean monthly (left) precipitation, (center) max temperature, and (right) min temperature for the concurrent period (1 Jan 1950–31 Jul 2000).

As in the original M02 dataset, we produced three types of data, all of which are publicly available (http://www.hydro.washington.edu/SurfaceWaterGroup/Data/gridded/index.html): 1) station-based daily precipitation and temperature data, and wind fields from the National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) reanalysis (Kalnay et al. 1996); 2) derived subdaily (3 hourly) land surface model forcing data, including precipitation and temperature, as well as downwelling solar and longwave radiation, humidity, and surface air pressure; and 3) model-based hydrological states and fluxes. We attempted as much as possible to follow the methods of M02, so that studies using those data can apply L13 to extend or refine their previous analyses. Below, we briefly describe the spatial gridding methodology and updates to the hydrologic model and provide comparisons between M02 and L13. We also compare model-based hydrologic outputs with observations of streamflow, soil moisture, and surface heat and radiative fluxes, presented here in a format consistent with M02.

2. Gridding methodology

The gridding methods of M02 were closely followed in L13, and the reader may find complete details in that publication. The L13 dataset is derived from observations of precipitation and minimum and maximum daily temperature at National Climatic Data Center (NCDC) Cooperative Observer (COOP) stations across the conterminous United States (DSI-3200). Although the cumulative total number of stations used is about 20 000, the number at any time varies, with a peak of approximately 12 000 stations in 1970. As in M02, we used only stations with at least 20 years of valid data. L13 uses the same relationships as in M02 to estimate those variables (downward solar and longwave radiation and humidity) that are not observed directly using algorithms described in the next paragraph. Both temperature and precipitation were gridded to $1/16°$ using the synergraphic mapping system (SYMAP) algorithm. Precipitation was linearly apportioned among days based on the time of observation. Daily maximum and minimum temperature were assumed to occur in the day of record, because the times of observation were not consistently recorded. Station metadata were incorporated into the gridding process through use of the quality control (QC) flags; however, issues beyond those that qualified for flagging (e.g., instrument error or upgrade) were not explicitly accounted for given the lack of documentation, aside from a few obvious inconsistencies in precipitation data noted in the supplemental material (available online at http://dx.doi.org/10.1175/JCLI-D-12-00508.s1). Gridded precipitation values were subsequently scaled on a monthly basis so as to match the long-term mean from the Parameter-Elevation Regressions on Independent Slopes Model (PRISM; Daly et al. 1994); for consistency with M02, a 1961–90 PRISM climatology was used. Wind data were linearly interpolated from a larger (approximately 1.9° grid) NCEP–NCAR reanalysis grid
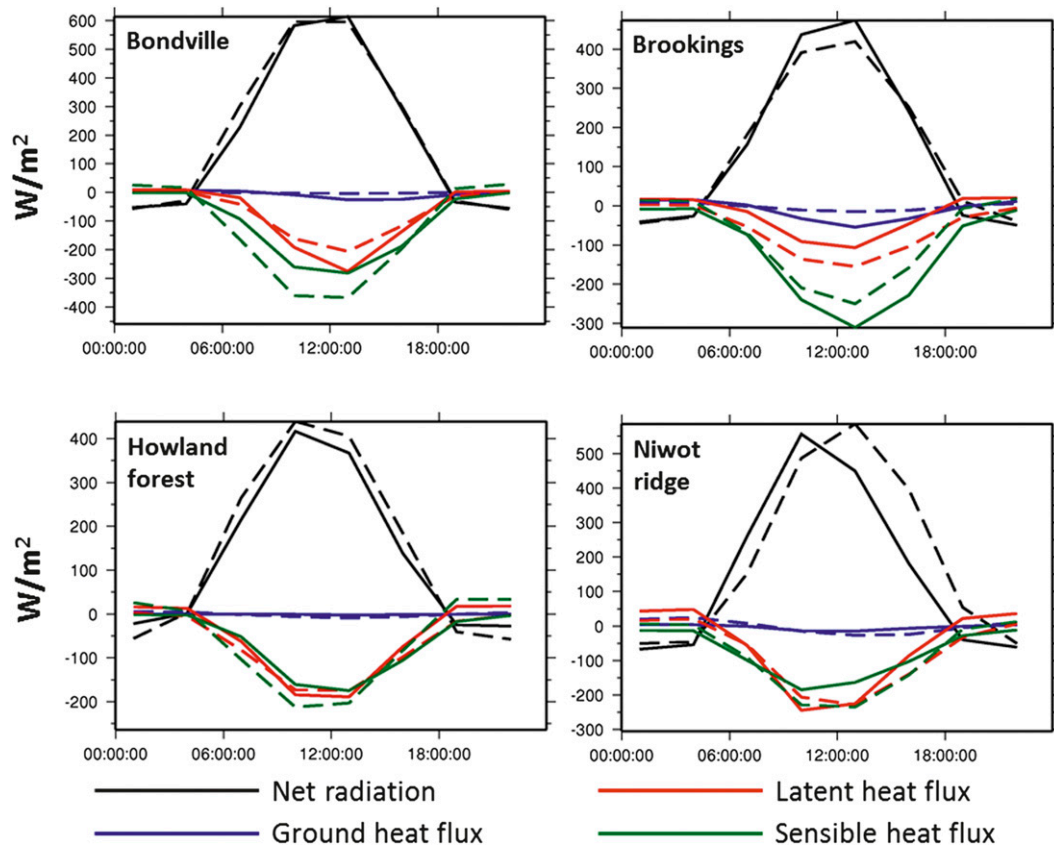
FIG. 2. L13 (dashed lines) modeled energy budget components compared with observations (solid lines) for a single summer for each Ameriflux site (JJA), specifically Blodgett Forest, CA (2004); Niwot Ridge, CO (2006); Brookings, SD (2005); and Howland Forest, ME (2001).

(Kalnay et al. 1996). Because the reanalysis data are only available from 1948 onward, a daily wind climatology for 1948–2011 was used for years prior to 1948.

Vapor pressure, humidity, and incoming shortwave and longwave radiation were derived using algorithms from mountain microclimate simulator (MTCLIM; Kimball et al. 1997; Thornton and Running 1999; Thornton et al. 2001) as described in M02, with several updates outlined in Bohn et al. (2013) The major difference between the version of MTCLIM used in L13 and M02 is a change in the estimate of longwave radiation from Tennessee Valley Authority (1972) to Prata (1996). To provide subdaily (3 hourly) temperature, a spline was applied to daily minimum and maximum temperatures to estimate the diurnal cycle (see Bohn et al. 2013).

Hydrologic model

As in M02, hydrologic states and fluxes were simulated using the VIC model (Liang et al. 1994). VIC is a grid-based hydrologic model that balances surface energy and water budgets at typical spatial resolutions ranging from a few to hundreds of kilometers. VIC

represents subgrid variability of vegetation and runoff generation, while also accounting for subgrid topography through elevation bands. Land-cover input data are the same as in M02, with static vegetation (Hansen et al. 2000), and soil information (Miller and White 1998) aggregated from a 1-km database for the effective years of 2000 and 1998, respectively. The VIC model used in L13, version 4.1.2, was run in energy balance mode, and has undergone a number of upgrades since the M02 data were published (using version 4.0.3). Readers are referred to the VIC website (http://www.hydro.washington.edu/Lettenmaier/Models/VIC/) for a complete description of these upgrades, the most important of which are related to the snow accumulation and ablation model, which now performs a separate energy balance for canopy snowpack and snow on the ground (Andreadis et al. 2009).

3. Evaluation

We first compared gridded station data from L13 and M02 (Fig. 1) for summer and winter over the concurrent

TABLE 1. Details of observational Ameriflux data used for comparison with simulated fluxes and states.

Tower name	Climate	Elevation (m)	Lat (N)	Lon (W)
Blodgett Forest	Mediterranean	1315	38.89°	120.63°
Niwot Ridge	Subalpine mixed coniferous	3050	40.03°	105.55°
Brookings	Temperate grassland	510	44.35°	96.84°
Howland Forest	Temperate continental	60	45.20°	68.74°

TABLE 2. Details of observational Global Soil Moisture Data Bank data used for comparison with simulated fluxes and states.

Number of stations	Elevation range (m)	Lat range (N)	Lon range (W)
19	130–265	38.13°–42.28°	88.10°–90.83°

period from 1 January 1950 to 31 July 2000. The two datasets are largely consistent, with differences mainly over topographically complex regions in the western United States. The major source of discrepancies are 1) intragrid variability from the four $1/16°$ L13 grid cells that weight the station data slightly differently than the single $1/8°$ M02 cell and 2) the 20-yr constraint on valid stations, which leads to L13 having slightly more valid stations at the beginning and end of the concurrent period (i.e., 1950s and 1990s) than in M02.

Figure 2 compares derived surface energy budget components with observations from four Ameriflux towers during summer (see Table 1). Simulated fluxes track observations fairly well at each site with several exceptions. First, derived fluxes tend to underpredict sensible heat fluxes and overpredict latent heat fluxes. The average difference in latent heat flux across sites and time intervals is $-19.5 \, W \, m^{-2}$, or 17%, which is equivalent to an overestimation of $0.69 \, mm \, day^{-1}$ of evaporation during summer. Second, at Niwot Ridge (a high-alpine site) there is a timing lag in the simulated peak radiation. Downward solar radiation is a derived quantity based on minimum and maximum daily air temperatures, which suggests that the assigned timing for the $1/16°$ grid cell is not representative of the Ameriflux site, which is situated on a ridge. In the solar radiation derivation algorithm, minimum daily temperature is assumed to occur at sunrise, while maximum daily temperature is assumed to occur at two-thirds of the duration between sunrise and sunset. The reader is referred to Bohn et al. (2013) for further explanation of the radiation algorithm. For details on model-derived cold-season fluxes and their evaluation, the reader is referred to Andreadis et al. (2009) and Cherkauer and Lettenmaier (2003). Feng et al. (2008) evaluated the VIC snow model in comparison with other state-of-the-art models using data collected as part of the Cold Land Processes Field Experiment (CLPX; Elder et al. 2009), and found that VIC predictions agreed well with higher complexity snow models in realistically capturing the duration of snow cover and snow density.

Soil moisture plays a central role in hydrologic processes such as runoff generation and ET, and is a key indicator of drought. Simulated soil moisture from VIC, driven as described above, was compared with observations from 19 sensors in Illinois retrieved from the Global Soil Moisture Data Bank (Robock et al. 2000) summarized in Table 2. Figure 3 shows mean monthly soil moisture values as well as the autocorrelations. VIC climatological soil moisture values are consistently lower than the observations for the 19-sensor average. However, the VIC-simulated intermonthly variability tracks very closely with observations, indicating that the model realistically simulates moisture storage changes and water budget dynamics for this part of the domain. The monthly autocorrelation is a measure of persistence of soil moisture anomalies in time, important for seasonal runoff forecasting and characterizing drought evolution. Figure 3b demonstrates that the temporal structure of model response effectively captures observed persistence for the first 3 months, becoming slightly less
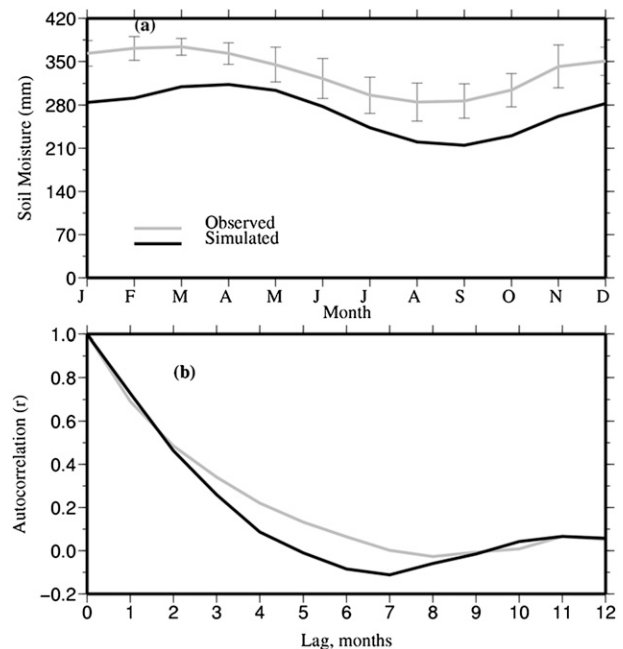

FIG. 3. Comparison of L13 (a) mean monthly soil moisture and (b) autocorrelations with 19 sensors in Illinois retrieved from the Global Soil Moisture Data Bank (1981–2004) for the top 1 m of soil (observed and modeled soil columns may extend slightly deeper). Note: Bars in (a) indicate monthly std dev.
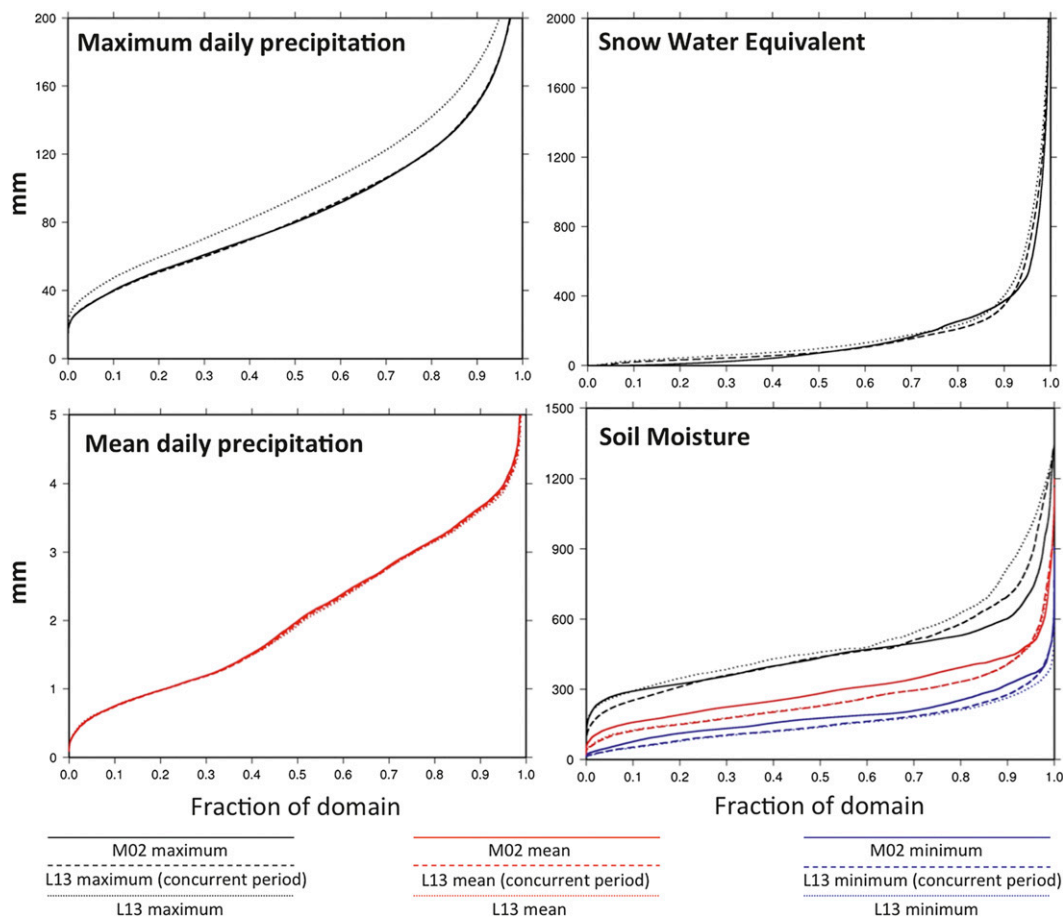
FIG. 4. Cumulative density functions (CDFs) of (left) inputs max and mean daily precipitation and (right) state variables soil moisture and SWE, comparing the historic M02 dataset (solid lines) with the L13 dataset over the concurrent period (1 Jan 1950–31 Jul 2000; dashed lines) and the entire L13 record (1 Jan 1916–31 Dec 2011; dotted lines). Total soil depth is variable across the domain ranging from roughly 1.5 to 2.7 m. Mean and max precipitation are separated for ease of viewing.

persistent thereafter, while autocorrelations become almost negligible beyond 5 months. The L13 magnitude and autocorrelations track those from observed soil moisture comparably to the original M02 data.

In addition to soil moisture, SWE is a key hydrologic state variable. Figure 4 shows histograms of the dynamic soil moisture range, mean and maximum SWE, and precipitation for M02 and L13 over the concurrent time period, as well as L13 for the extended period (1 January 1915–31 December 2011). SWE values were frequently larger for the finer spatial domain ($\frac{1}{16}°$) than the coarser ($\frac{1}{8}°$) during the concurrent period, corresponding to an increased meteorological variability, while the extended period had maximum values that were still larger. The dynamic soil moisture range was accordingly greatest for the extended period (at $\frac{1}{16}°$). Maximum daily precipitation was comparable between the two datasets over the concurrent

period; however, larger daily values were frequently recorded for the extended period corresponding to a wet period before 1925, as well as over topographically complex regions. Conversely, the mean daily precipitation values were stable across both periods and resolutions.

Simulated streamflows are compared with observations in Fig. 5 from major river basins covering large portions of the domain. For several basins, particularly in the western United States, naturalized streamflow data were obtained that have been adjusted for anthropogenic impacts, including upstream regulation, water withdrawals, and evaporation from upstream reservoirs (see Table 3). Limited VIC parameter estimation was performed to match surface and subsurface runoff from the previously calibrated VIC (version 4.0.3) used in M02. We employed a technique similar to Troy (2008) with the objective of matching the runoff

Runoff Comparison of major basins
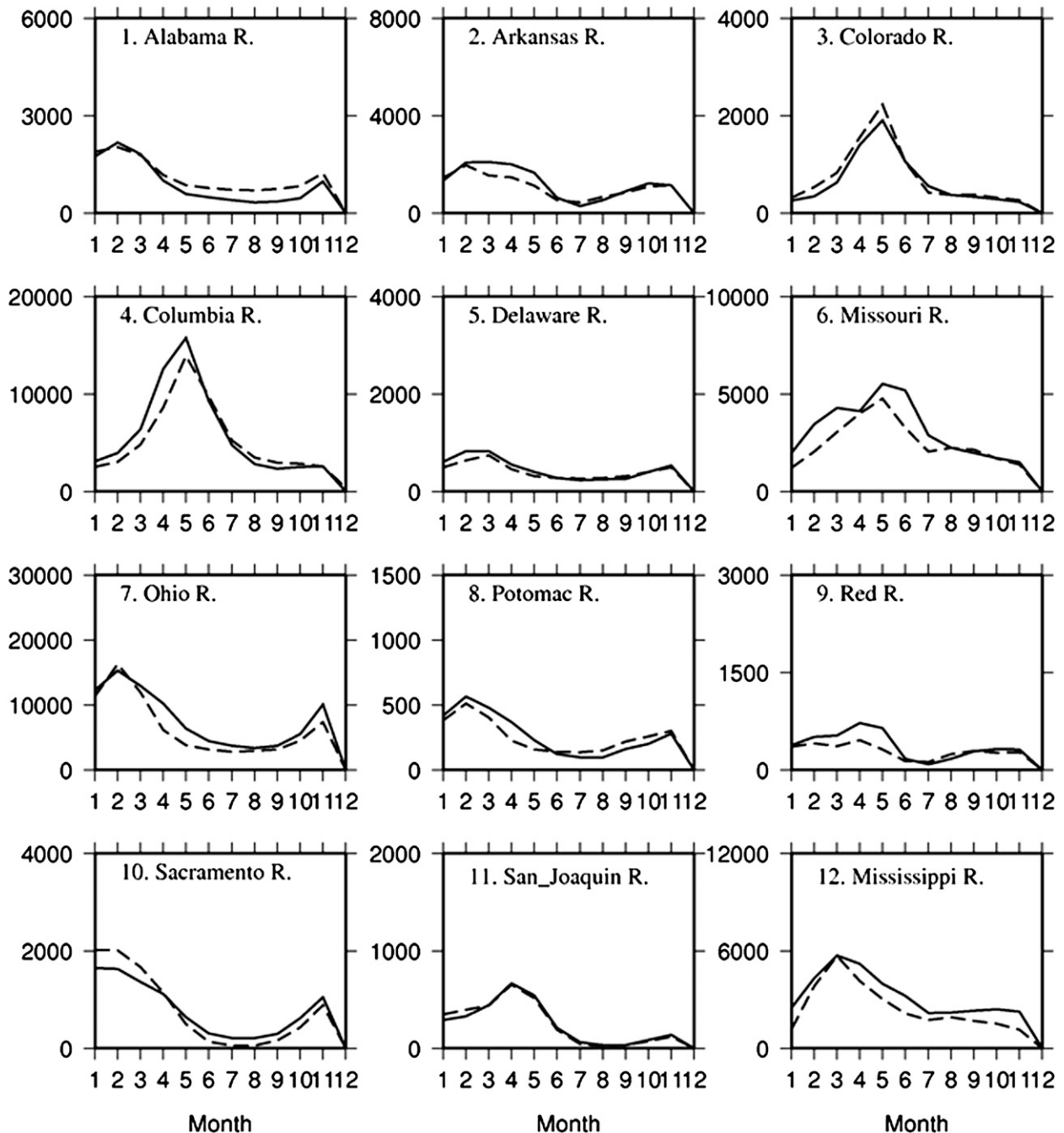
FIG. 5. Mean monthly hydrographs from L13 ($m^3 s^{-1}$) over the period 1961–90. Simulated flows are denoted by dashed lines, while observed or naturalized flows are solid lines.

ratio (in this case between model versions 4.0.3 and 4.1.2) at regularly spaced intervals of 1°. A Monte Carlo search consisting of 200 iterations was applied, which varied three VIC soil parameters—the variable

infiltration curve parameter b, the maximum velocity of baseflow parameter D_{smax}, and the depth of the bottom soil layer D_3—within a narrow range ($\pm 10\%$) of their previous values.

TABLE 3. Details of observational streamflow data used for comparison with simulated fluxes and states.

River name	Station name	Area (km^2)
Alabama	Clairborne	56 900
Arkansas*	Ralston	121 340
Colorado*	Lees Ferry	278 070
Columbia*	Dalles	613 280
Delaware	Memorial Bridge	28 500
Missouri*	Hermann	1 357 670
Ohio	Metropolis	525 760
Potomac	Point of Rocks	25 000
Red*	Index	124 390
Sacramento*	Bend Bridge	23 050
San Joaquin*	Mokelumne Hill	1860
Upper Mississippi	Grafton	443 660

* Naturalized streamflow were obtained.

Offline simulations were conducted to evaluate the impact of using climatological winds prior to 1948 (see the supplemental material). These comparisons showed that with few exceptions use of the climatological winds slightly reduce the temporal and spatial variability of hydrologic fluxes but have small relative impacts on long-term mean values. Relative impacts on short-term (3-hourly and daily) values are greatest, and are less at monthly time steps. Given the uncertainty in using static vegetation and soil (from 2000 and 1998, respectively), the derived model outputs for the earlier part of the simulation period serve as a reference scenario (rather than a reconstruction), while providing the necessary meteorological inputs for users who might desire to produce more detailed dynamic reconstructions [as a point of reference, Matheussen et al. (2000) found maximum changes in runoff and ET of less than 10% for reconstructed 1900 versus 1990 vegetation in the Columbia River basin]. Additional uncertainty arises from using a constant lapse rate in regions of topographical complexity (i.e., western United States), with the potential to bias daily temperature range in certain cases, which may impact derived downwelling shortwave radiation based on the MTCLIM algorithm. It follows that undocumented or incomplete QC of instrument change error may hinder the robustness of trends in these data, as pointed out by Menne et al. (2009). Menne et al. (2010) subsequently showed that there has been no successful correction of the biases associated with the change from liquid-in-glass (LiG) thermometers to Maximum–Minimum Temperature System (MMTS) in the daily NCDC station dataset. This is estimated to result in an artificial negative bias in maximum temperatures less than approximately 0.5°C, which would translate into a slight underestimation of diurnal temperature range

and reduction in derived shortwave radiation for L13. This instrument bias could have further implications for simulated cold-season processes (snowpack evolution) and surface heat fluxes, as well as for drought assessment and climate-oriented analyses such as downscaling.

4. Data format and availability

The data are available in Network Common Data Form (NetCDF) format, conforming to the Assistance for Land-Surface Modelling Activities (ALMA) convention of Polcher et al. (2000). This means that moisture fluxes are expressed as kilograms per square meter per second, energy fluxes as watts per square meter, and moisture states as kilograms per square meter. (The data are freely accessible from ftp://ftp.hydro.washington. edu/pub/blivneh/CONUS/, where we also provide plots comparing a range of other states and fluxes between M02 and L13.)

5. Conclusions

We have described an observation-based hydrologically consistent dataset for the period 1915–2011 at a $\frac{1}{16}°$ spatial resolution. Gridded station data for precipitation and temperature, surface wind from an atmospheric reanalysis, and derived downward solar and longwave radiation and vapor pressure were used to force a hydrologic model that was shown to reproduce, on average, observed surface heat fluxes, soil moisture, and runoff. These data have potential uses for model evaluation and diagnosis in energy and water balance studies and climate change impact studies. We expect that these data will complement studies that have used the M02 dataset, given the wider range of conditions that are included in a longer time period and at finer spatial resolution.

REFERENCES

Andreadis, K. M., and D. P. Lettenmaier, 2006: Assimilating remotely sensed snow observations into a macroscale hydrology model. *Adv. Water Resour.,* **29,** 872–886.

——, P. Storck, and D. P. Lettenmaier, 2009: Modeling snow accumulation and ablation processes in forested environments. *Water Resour. Res.,* **45,** W05429, doi:10.1029/2008WR007042.

Bohn, T. J., B. Livneh, J. Oyler, S. W. Running, B. Nijssen, and D. P. Lettenmaier, 2013: Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. *Agric. For. Meteor.,* **176,** 38–49, doi:10.1016/j.agrformet.2013.03.003.

Carpenter, T. M., and K. P. Georgakakos, 2004: Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model. *J. Hydrol.,* **298,** 202–221.

Castro, C. L., R. A. Pielke Sr., and J. O. Adegoke, 2007: In-
   vestigation of the summer climate of the contiguous United
   States and Mexico using the Regional Atmospheric Model-
   ing System (RAMS). Part I: Model climatology (1950–2002).
   J. Climate, **20,** 3844–3865.

Cayan, D. R., E. P. Maurer, M. D. Dettinger, M. Tyree, and
   K. Hayhoe, 2008: Climate change scenarios for the California
   region. *Climatic Change,* **87** (Suppl. 1), 21–42, doi:10.1007/
   s10584-007-9377-6.

Cherkauer, K. A., and D. P. Lettenmaier, 2003: Simulation of
   spatial variability in snow and frozen soil. *J. Geophys. Res.,*
   108, 8858, doi:10.1029/2003JD003575.

Christensen, N. S., A. W. Wood, N. Voisin, D. P. Lettenmaier, and
   R. N. Palmer, 2004: The effects of climate change on the hy-
   drology and water resources of the Colorado River basin.
   Climatic Change, **62,** 337–363.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-
   topographic model for mapping climatological precipitation
   over mountainous terrain. *J. Appl. Meteor.,* **33,** 140–158.

Elder, K., D. Cline, A. Goodbody, P. Houser, G. E. Liston,
   L. Mahrt, and N. Rutter, 2009: NASA Cold Land Processes
   Experiment (CLPX 2002/03): Ground-based and near-surface
   meteorological observations. *J. Hydrometeor.,* **10,** 330–337.

Feng, X., A. Sahoo, K. Arsenault, P. Houser, Y. Luo, and T. J.
   Troy, 2008: The impact of snow model complexity at three
   CLPX sites. *J. Hydrometeor.,* **9,** 1464–1481.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, and R. Sohlberg,
   2000: Global land cover classification at 1 km spatial resolution
   using a classification tree approach. *Int. J. Remote Sens.,* **21,**
   1331–1364.

Hayhoe, K., and Coauthors, 2004: Emissions pathways, climate
   change, and impacts on California. *Proc. Natl. Acad. Sci. USA,*
   101, 12 422–12 427.

——, and Coauthors, 2007: Past and future changes in climate and
   hydrological indicators in the US Northeast. *Climate Dyn.,* **28,**
   381–407.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Re-
   analysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kimball, J. S., S. W. Running, and R. R. Nemani, 1997: An im-
   proved method for estimating surface humidity from daily
   minimum temperature. *Agric. For. Meteor.,* **85,** 87–98.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A
   simple hydrologically based model of land surface water and
   energy fluxes for GSMs. *J. Geophys. Res.,* **99** (D7), 14 415–14 428.

Loarie, S. R., B. E. Carter, K. Hayhoe, S. McMahon, R. Moe, C. A.
   Knight, and D. D. Ackerly, 2008: Climate change and the fu-
   ture of California's endemic flora. *PLoS ONE,* **3,** 2502,
   doi:10.1371/journal.pone.0002502.

Matheussen, B., R. L. Kirschbaum, I. A. Goodman, G. M.
   O'Donnell, and D. P. Lettenmaier, 2000: Effects of land cover
   change on streamflow in the interior Columbia River Basin
   (USA and Canada). *Hydrol. Processes,* **14,** 867–885.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and
   B. Nijssen, 2002: A long-term hydrologically based dataset of
   land surface fluxes and states for the conterminous United
   States. *J. Climate,* **15,** 3237–3251.

——, D. P. Lettenmaier, and N. J. Mantua, 2004: Variability and
   predictability of North American runoff. *Water Resour. Res.,*
   40, W09306, doi:10.1029/2003WR002789.

Menne, M. J., C. N. Williams, and R. S. Vose, 2009: The U.S.
   Historical Climatology Network monthly temperature data,
   version 2. *Bull. Amer. Meteor. Soc.,* **90,** 993–1007.

——, ——, and M. A. Palecki, 2010: On the reliability of the U.S.
   surface temperature record. *J. Geophys. Res.,* **115,** D11108,
   doi:10.1029/2009JD013094.

Miller, D. A., and R. A. White, 1998: A conterminous United
   States multilayer soil characteristics dataset for regional
   climate and hydrology modeling. *Earth Interact.,* **2,** doi:10.1175/
   1087-3562(1998)002<0001:ACUSMS>2.3.CO;2.

Polcher, J., and Coauthors, 2000: GLASS: Global Land–Atmosphere
   System Study. *GEWEX News,* No. 10 (2), International
   GEWEX Project Office, Silver Spring, MD, 3–5.

Prata, A. J., 1996: A new long-wave formula for estimating down-
   ward clear-sky radiation at the surface. *Quart. J. Roy. Meteor.
   Soc.,* **122,** 1127–1151.

Robock, A., K. Y. Vinnikov, G. Srinivasan, J. K. Entin, S. E.
   Hollinger, N. A. Speranskaya, S. Liu, and A. Namkhai, 2000:
   The Global Soil Moisture Data Bank. *Bull. Amer. Meteor.
   Soc.,* **81,** 1281–1299.

Salathé, E. P., 2003: Comparison of various precipitation down-
   scaling methods for the simulation of streamflow in a rain-
   shadow river basin. *Int. J. Climatol.,* **23,** 887–901.

Sheffield, J., G. Goteti, F. Wen, and E. F. Wood, 2004: A simulated
   soil moisture based drought analysis for the United States.
   J. Geophys. Res., **109,** D24108, doi:10.1029/2004JD005182.

Smith, M. B., D. Seo, V. I. Koren, S. M. Reed, Z. Zhang, Q. Duan,
   F. Moreda, and S. Cong, 2004: The distributed model inter-
   comparison project (DMIP): Motivation and experiment de-
   sign. *J. Hydrol.,* **298,** 4–26.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger, 2004: Changes in
   snowmelt runoff timing in western North America under a
   "business as usual" climate change scenario. *Climatic Change,*
   62, 217–232.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of
   CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.,*
   93, 485–498.

Tennessee Valley Authority, 1972: Heat and mass transfer between
   a water surface and the atmosphere. Laboratory Rep. 14,
   Tennessee Valley Authority, Norris, TN, 270 pp.

Thornton, P. E., and S. W. Running, 1999: An improved algorithm
   for estimating incident daily solar radiation from measure-
   ments of temperature, humidity, and precipitation. *Agric. For.
   Meteor.,* **93,** 211–228.

——, H. Hasenauer, and M. A. White, 2001: Simultaneous esti-
   mation of daily solar radiation and humidity from observed
   temperature and precipitation: An application over com-
   plex terrain in Austria. *Agric. For. Meteor.,* **104,** 255–271,
   doi:10.1016/S0168-1923(00)00170-2.

Troy, T. J., 2008: An efficient calibration method for continental-
   scale land surface modeling. *Water Resour. Res.,* **44,** W09411,
   doi:10.1029/2007WR006513.

Westerling, A. L., H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam,
   2006: Warming and earlier spring increase western U.S. forest
   wildfire activity. *Science,* **18,** 940–943.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004:
   Hydrologic implications of dynamical and statistical ap-
   proaches to downscaling climate model outputs. *Climatic
   Change,* **62,** 189–216.

CORRIGENDUM

BEN LIVNEH

Cooperative Institute for Research in Environmental Science, University of Colorado Boulder, Boulder, Colorado

ERIC A. ROSENBERG, CHIYU LIN, BART NIJSSEN, AND VIMAL MISHRA[+]

Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington

KOSTAS M. ANDREADIS

Jet Propulsion Laboratory, Pasadena, California

EDWIN P. MAURER

Department of Civil Engineering, Santa Clara University, Santa Clara, California

DENNIS P. LETTENMAIER

Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington

(Manuscript received 20 November 2013, in final form 4 December 2013)

Due to a production error, Figs. 1 and 5 were published incorrectly in Livneh et al. (2013). In addition, the third row of Table 3 contained an error and the final coauthor's affiliation and corresponding author address was not accurate. To correct this, the following pages contain the full article as it should have appeared, with all of the above corrections implemented.

The staff of the *Journal of Climate* regrets any inconvenience these errors may have caused.

REFERENCE

Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. M. Andreadis, E. P. Maurer, and D. P. Lettenmaier, 2013: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *J. Climate,* **26,** 9384–9392.

A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States: Update and Extensions*

Ben Livneh

Cooperative Institute for Research in Environmental Science, University of Colorado Boulder, Boulder, Colorado

Eric A. Rosenberg, Chiyu Lin, Bart Nijssen, and Vimal Mishra[+]

Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington

Kostas M. Andreadis

Jet Propulsion Laboratory, Pasadena, California

Edwin P. Maurer

Department of Civil Engineering, Santa Clara University, Santa Clara, California

Dennis P. Lettenmaier

Department of Civil and Environmental Engineering, University of Washington, Seattle, Washington

ABSTRACT

This paper describes a publicly available, long-term (1915–2011), hydrologically consistent dataset for the conterminous United States, intended to aid in studies of water and energy exchanges at the land surface. These data are gridded at a spatial resolution of $1/16°$ latitude/longitude and are derived from daily temperature and precipitation observations from approximately 20 000 NOAA Cooperative Observer (COOP) stations. The available meteorological data include temperature, precipitation, and wind, as well as derived humidity and downwelling solar and infrared radiation estimated via algorithms that index these quantities to the daily mean temperature, temperature range, and precipitation, and disaggregate them to 3-hourly time steps. Furthermore, the authors employ the variable infiltration capacity (VIC) model to produce 3-hourly estimates of soil moisture, snow water equivalent, discharge, and surface heat fluxes. Relative to an earlier similar dataset by Maurer and others, the improved dataset has 1) extended the period of analysis (1915–2011 versus 1950–2000), 2) increased the spatial resolution from $1/8°$ to $1/16°$, and 3) used an updated version of VIC. The previous dataset has been widely used in water and energy budget studies, climate change assessments, drought reconstructions, and for many other purposes. It is anticipated that the spatial refinement and temporal extension will be of interest to a wide cross section of the scientific community.

Corresponding author address: Dennis P. Lettenmaier, Department of Civil and Environmental Engineering, Box 352700, University of Washington, Seattle, WA 98195-2700.
E-mail: dennisl@u.washington.edu

1. Introduction

Increased computational capabilities, the availability of new data sources such as remote sensing, and better understanding of the Earth system have resulted in considerable improvements in the ability to represent long-term variations in land surface water and energy fluxes and state variables. Earth system models require, among other things, consistent observational datasets for model testing and diagnosis. Furthermore, predictions of alternative

future scenarios of land surface conditions resulting from changes in climate and/or land cover require benchmark historical data against which to evaluate.

We describe an observational dataset (herein called L13) that provides a means to analyze and verify hydroclimatic predictions as well as to drive land surface models, along with fluxes and state variables from the variable infiltration capacity (VIC) model. The L13 dataset is based on the methods of Maurer et al. (2002; hereafter M02), who developed a set of publicly available gridded meteorological data from ground-based measurements, together with model-derived hydrologically consistent surface fluxes and states. M02 spanned the period 1 January 1949–31 July 2000 (~51.5 yr) at a $\frac{1}{8}°$ latitude/longitude spatial resolution. The L13 dataset described here refines the spatial resolution to $\frac{1}{16}°$, extends the period of record backward to 1 January 1915 and forward to 31 December 2011 (97 yr), and provides fluxes and states from an updated version of the VIC land surface model. The significance of each of these aspects is described below, followed by a summary of evaluations of the new dataset relative to M02.

An examination of the most widely cited studies that reference and/or use the M02 data (in total, over 300 Web of Science citations; http://wokinfo.com/) suggests that the applications can be grouped into three general areas: 1) studies that use the meteorological and hydrological data directly to characterize the state or variability of a specific hydroclimatic variable (e.g., temperature, precipitation, and snowpack), 2) studies that use the data as a spatially and temporally complete observational baseline for downscaling climate model output (especially for bias correction) to generate future climate scenarios, and 3) water and energy balance studies, for which the model forcings and derived fluxes are of particular interest because the derived surface water and energy budgets close at all grid cells at each time step by construct. Examples of each of these applications are discussed below.

Westerling et al. (2006) used the M02 gridded meteorological data along with other sources to isolate the signal of climatic variability on wildfire frequency in the western United States. Hayhoe et al. (2007) used the archived hydrological fluxes and states to represent historical hydrologic conditions from which future meteorological scenarios were assessed via hydrologic simulations in the northeastern United States. Soil moisture data were used by Castro et al. (2007) to initialize a regional climate model to simulate U.S. climatology. Sheffield et al. (2004) used M02 soil moisture to derive a hydrologically based drought index, which showed good agreement with time series of U.S. drought from two Palmer drought severity index (PDSI) datasets.

The second group of studies typically follows procedures that include correcting downscaled climate model output with the M02 forcing data, and then using the bias-corrected model to produce future climatic scenarios. Cayan et al. (2008) and Hayhoe et al. (2004) both exploited the downscaled climate model outputs to assess climatic implications of future greenhouse gas emission scenarios in California using this general approach. Loarie et al. (2008) used downscaled outputs to examine impacts on the diversity of flora in California. Salathé (2003) downscaled climate model outputs to simulate streamflow over a river basin in Washington State and also evaluated performance of several climate models after bias correction. Wood et al. (2004) used the forcing and hydrological datasets to evaluate six bias correction methods for downscaling climate model outputs over the continental United States.

Among the third type of applications, Stewart et al. (2004) utilized the forcing data to simulate streamflow over large river basins in the Pacific Northwest. Smith et al. (2004) used the meteorological data to force a suite of land surface models and compared their performance. Christensen et al. (2004) applied the forcing data over the Colorado River basin to search for robust VIC model parameters over small river basins that were then used to assess climatic impacts under future forcing scenarios. Carpenter and Georgakakos (2004) utilized the energy forcing data to compute potential evapotranspiration (ET) in a radar–rainfall uncertainty study. Maurer et al. (2004) used climate data and the archived VIC-derived soil moisture, snow, and runoff data to examine predictability of runoff across the United States. Andreadis and Lettenmaier (2006) used the forcing data to evaluate a data assimilation scheme using satellite-based snow water equivalent (SWE) information.

Climate model outputs, remote sensing, and land cover data continue to become available at finer spatial resolutions, making the spatial refinement of L13 a significant improvement (from $\frac{1}{8}°$ to $\frac{1}{16}°$). The extended period of record (from 50 to 97 yr) will help to improve the statistical strength of computed trends from hydroclimatic analyses and model corrections for downscaled climate model outputs, and it captures important historic extremes such as the 1930s drought that were outside the period of M02. Also, the climate model simulations for historic periods conducted as part of phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) extend through 2005, for which an extended observational dataset is useful for various purposes discussed below. Finally, an updated VIC model version (described below) includes refinements that will be of interest for some applications of the model-derived variables.
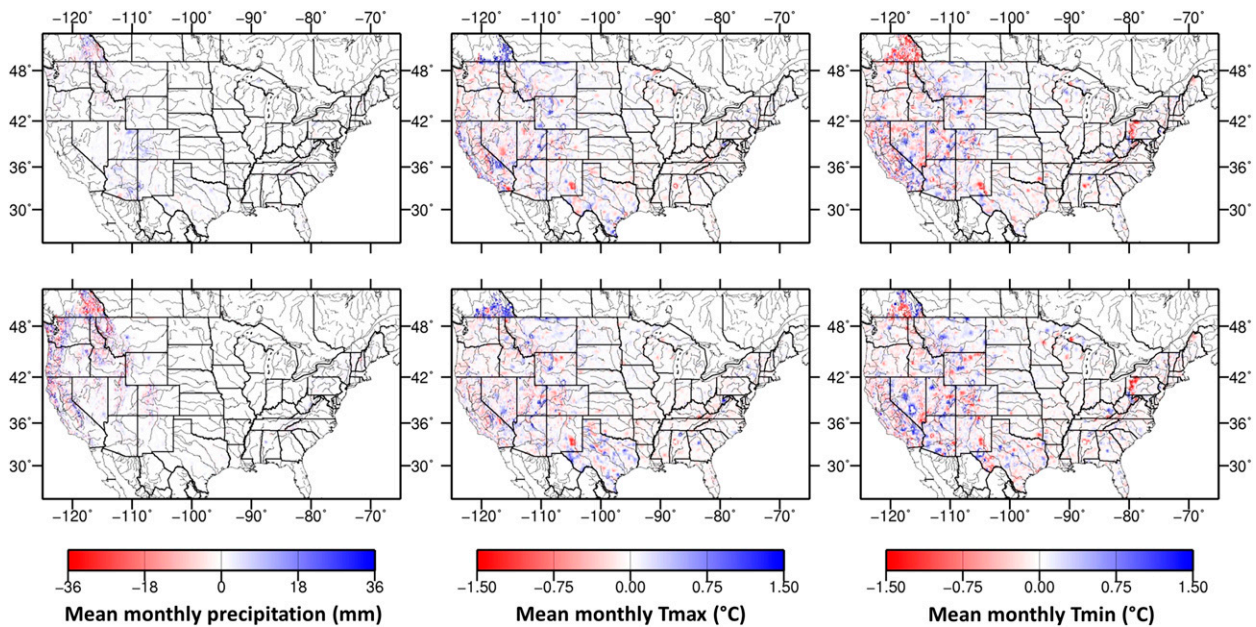
FIG. 1. L13 minus M02 for (top) summer [June–August (JJA)] and (bottom) winter [December–February (DJF)] mean monthly (left) precipitation, (center) max temperature, and (right) min temperature for the concurrent period (1 Jan 1950–31 Jul 2000).

As in the original M02 dataset, we produced three types of data, all of which are publicly available (http://www.hydro.washington.edu/SurfaceWaterGroup/Data/gridded/index.html): 1) station-based daily precipitation and temperature data, and wind fields from the National Centers for Environmental Prediction (NCEP)–National Center for Atmospheric Research (NCAR) reanalysis (Kalnay et al. 1996); 2) derived subdaily (3 hourly) land surface model forcing data, including precipitation and temperature, as well as downwelling solar and longwave radiation, humidity, and surface air pressure; and 3) model-based hydrological states and fluxes. We attempted as much as possible to follow the methods of M02, so that studies using those data can apply L13 to extend or refine their previous analyses. Below, we briefly describe the spatial gridding methodology and updates to the hydrologic model and provide comparisons between M02 and L13. We also compare model-based hydrologic outputs with observations of streamflow, soil moisture, and surface heat and radiative fluxes, presented here in a format consistent with M02.

2. Gridding methodology

The gridding methods of M02 were closely followed in L13, and the reader may find complete details in that publication. The L13 dataset is derived from observations of precipitation and minimum and maximum daily temperature at National Climatic Data Center (NCDC) Cooperative Observer (COOP) stations across the conterminous United States (DSI-3200). Although the cumulative total number of stations used is about 20 000, the number at any time varies, with a peak of approximately 12 000 stations in 1970. As in M02, we used only stations with at least 20 years of valid data. L13 uses the same relationships as in M02 to estimate those variables (downward solar and longwave radiation and humidity) that are not observed directly using algorithms described in the next paragraph. Both temperature and precipitation were gridded to $1/16°$ using the synergraphic mapping system (SYMAP) algorithm. Precipitation was linearly apportioned among days based on the time of observation. Daily maximum and minimum temperature were assumed to occur in the day of record, because the times of observation were not consistently recorded. Station metadata were incorporated into the gridding process through use of the quality control (QC) flags; however, issues beyond those that qualified for flagging (e.g., instrument error or upgrade) were not explicitly accounted for given the lack of documentation, aside from a few obvious inconsistencies in precipitation data noted in the supplemental material (available online at http://dx.doi.org/10.1175/JCLI-D-12-00508.s1). Gridded precipitation values were subsequently scaled on a monthly basis so as to match the long-term mean from the Parameter-Elevation Regressions on Independent Slopes Model (PRISM; Daly et al. 1994); for consistency with M02, a 1961–90 PRISM climatology was used. Wind data were linearly interpolated from a larger (approximately 1.9° grid) NCEP–NCAR reanalysis grid
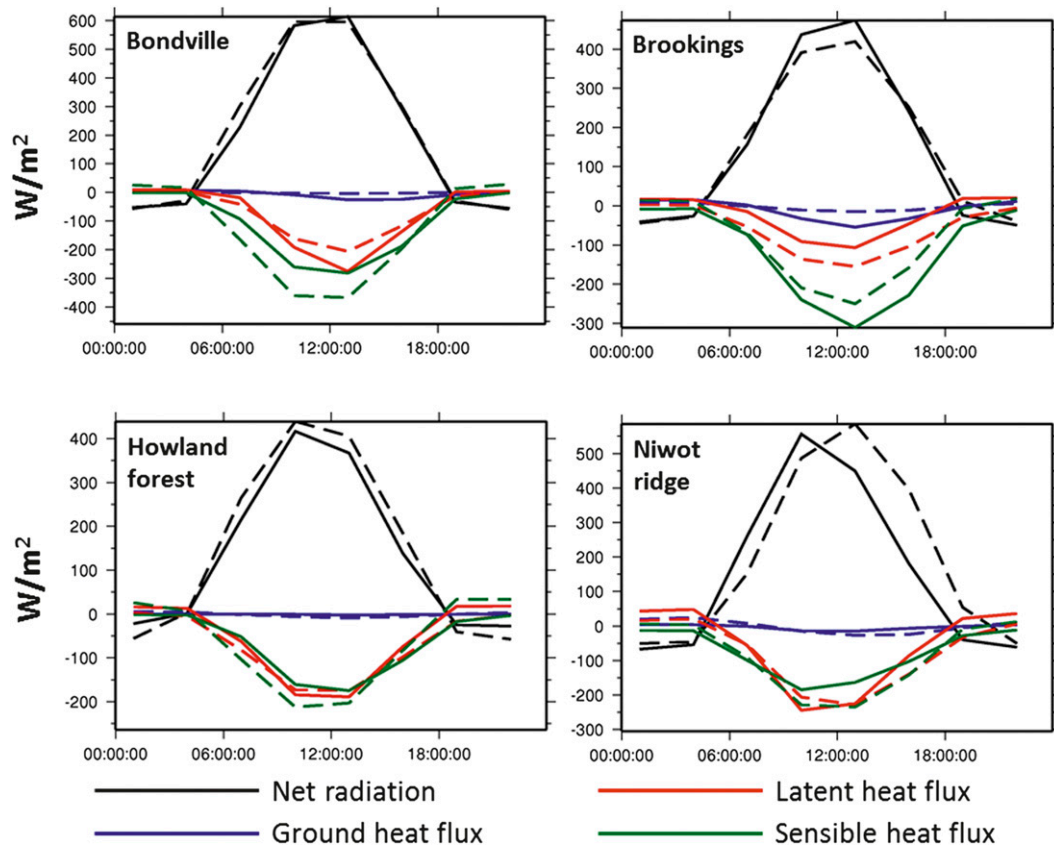
FIG. 2. L13 (dashed lines) modeled energy budget components compared with observations (solid lines) for a single summer for each Ameriflux site (JJA), specifically Blodgett Forest, CA (2004); Niwot Ridge, CO (2006); Brookings, SD (2005); and Howland Forest, ME (2001).

(Kalnay et al. 1996). Because the reanalysis data are only available from 1948 onward, a daily wind climatology for 1948–2011 was used for years prior to 1948.

Vapor pressure, humidity, and incoming shortwave and longwave radiation were derived using algorithms from mountain microclimate simulator (MTCLIM; Kimball et al. 1997; Thornton and Running 1999; Thornton et al. 2001) as described in M02, with several updates outlined in Bohn et al. (2013) The major difference between the version of MTCLIM used in L13 and M02 is a change in the estimate of longwave radiation from Tennessee Valley Authority (1972) to Prata (1996). To provide subdaily (3 hourly) temperature, a spline was applied to daily minimum and maximum temperatures to estimate the diurnal cycle (see Bohn et al. 2013).

Hydrologic model

As in M02, hydrologic states and fluxes were simulated using the VIC model (Liang et al. 1994). VIC is a grid-based hydrologic model that balances surface energy and water budgets at typical spatial resolutions ranging from a few to hundreds of kilometers. VIC

represents subgrid variability of vegetation and runoff generation, while also accounting for subgrid topography through elevation bands. Land-cover input data are the same as in M02, with static vegetation (Hansen et al. 2000), and soil information (Miller and White 1998) aggregated from a 1-km database for the effective years of 2000 and 1998, respectively. The VIC model used in L13, version 4.1.2, was run in energy balance mode, and has undergone a number of upgrades since the M02 data were published (using version 4.0.3). Readers are referred to the VIC website (http://www.hydro.washington. edu/Lettenmaier/Models/VIC/) for a complete description of these upgrades, the most important of which are related to the snow accumulation and ablation model, which now performs a separate energy balance for canopy snowpack and snow on the ground (Andreadis et al. 2009).

3. Evaluation

We first compared gridded station data from L13 and M02 (Fig. 1) for summer and winter over the concurrent

TABLE 1. Details of observational Ameriflux data used for comparison with simulated fluxes and states.

Tower name	Climate	Elevation (m)	Lat (N)	Lon (W)
Blodgett Forest	Mediterranean	1315	38.89°	120.63°
Niwot Ridge	Subalpine mixed coniferous	3050	40.03°	105.55°
Brookings	Temperate grassland	510	44.35°	96.84°
Howland Forest	Temperate continental	60	45.20°	68.74°

TABLE 2. Details of observational Global Soil Moisture Data Bank data used for comparison with simulated fluxes and states.

Number of stations	Elevation range (m)	Lat range (N)	Lon range (W)
19	130–265	38.13°–42.28°	88.10°–90.83°

period from 1 January 1950 to 31 July 2000. The two datasets are largely consistent, with differences mainly over topographically complex regions in the western United States. The major source of discrepancies are 1) intragrid variability from the four $1/16°$ L13 grid cells that weight the station data slightly differently than the single $1/8°$ M02 cell and 2) the 20-yr constraint on valid stations, which leads to L13 having slightly more valid stations at the beginning and end of the concurrent period (i.e., 1950s and 1990s) than in M02.

Figure 2 compares derived surface energy budget components with observations from four Ameriflux towers during summer (see Table 1). Simulated fluxes track observations fairly well at each site with several exceptions. First, derived fluxes tend to underpredict sensible heat fluxes and overpredict latent heat fluxes. The average difference in latent heat flux across sites and time intervals is $-19.5\,\mathrm{W\,m}^{-2}$, or 17%, which is equivalent to an overestimation of $0.69\,\mathrm{mm\,day}^{-1}$ of evaporation during summer. Second, at Niwot Ridge (a high-alpine site) there is a timing lag in the simulated peak radiation. Downward solar radiation is a derived quantity based on minimum and maximum daily air temperatures, which suggests that the assigned timing for the $1/16°$ grid cell is not representative of the Ameriflux site, which is situated on a ridge. In the solar radiation derivation algorithm, minimum daily temperature is assumed to occur at sunrise, while maximum daily temperature is assumed to occur at two-thirds of the duration between sunrise and sunset. The reader is referred to Bohn et al. (2013) for further explanation of the radiation algorithm. For details on model-derived cold-season fluxes and their evaluation, the reader is referred to Andreadis et al. (2009) and Cherkauer and Lettenmaier (2003). Feng et al. (2008) evaluated the VIC snow model in comparison with other state-of-the-art models using data collected as part of the Cold Land Processes Field Experiment (CLPX; Elder et al. 2009), and found that VIC predictions agreed well with higher complexity snow models in realistically capturing the duration of snow cover and snow density.

Soil moisture plays a central role in hydrologic processes such as runoff generation and ET, and is a key indicator of drought. Simulated soil moisture from VIC, driven as described above, was compared with observations from 19 sensors in Illinois retrieved from the Global Soil Moisture Data Bank (Robock et al. 2000) summarized in Table 2. Figure 3 shows mean monthly soil moisture values as well as the autocorrelations. VIC climatological soil moisture values are consistently lower than the observations for the 19-sensor average. However, the VIC-simulated intermonthly variability tracks very closely with observations, indicating that the model realistically simulates moisture storage changes and water budget dynamics for this part of the domain. The monthly autocorrelation is a measure of persistence of soil moisture anomalies in time, important for seasonal runoff forecasting and characterizing drought evolution. Figure 3b demonstrates that the temporal structure of model response effectively captures observed persistence for the first 3 months, becoming slightly less
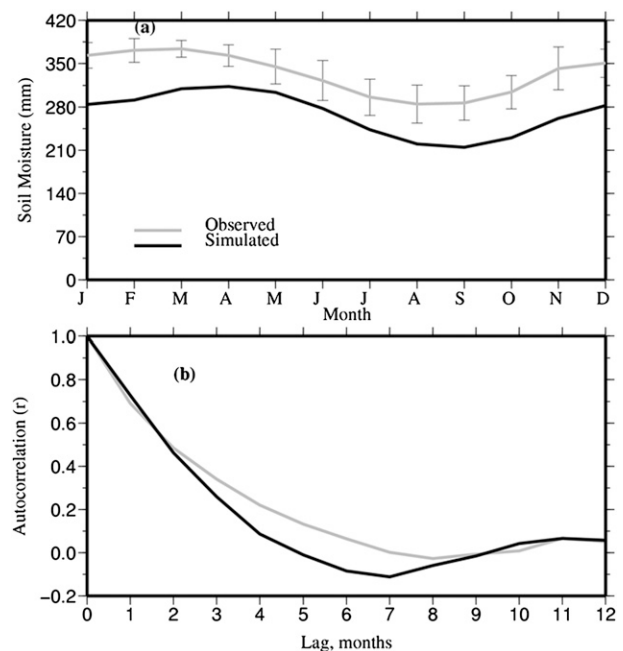

FIG. 3. Comparison of L13 (a) mean monthly soil moisture and (b) autocorrelations with 19 sensors in Illinois retrieved from the Global Soil Moisture Data Bank (1981–2004) for the top 1 m of soil (observed and modeled soil columns may extend slightly deeper). Note: Bars in (a) indicate monthly std dev.
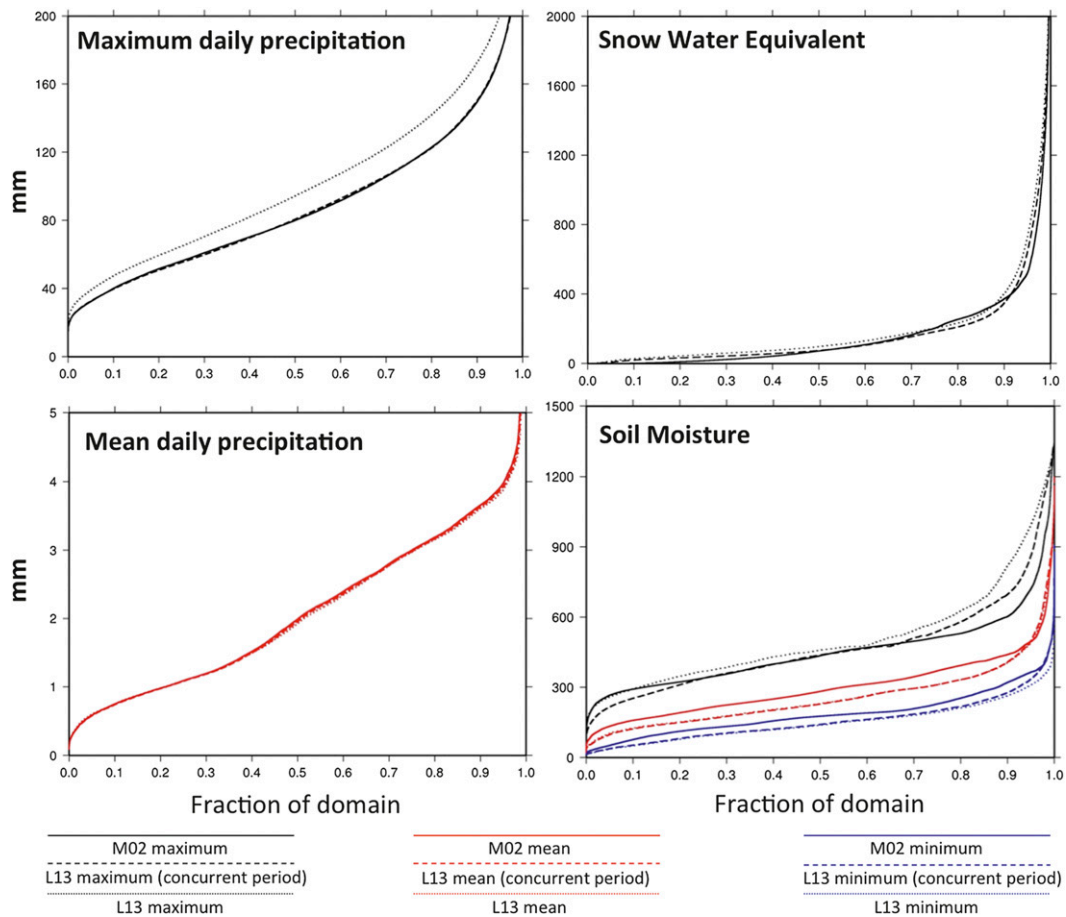
FIG. 4. Cumulative density functions (CDFs) of (left) inputs max and mean daily precipitation and (right) state variables soil moisture and SWE, comparing the historic M02 dataset (solid lines) with the L13 dataset over the concurrent period (1 Jan 1950–31 Jul 2000; dashed lines) and the entire L13 record (1 Jan 1916–31 Dec 2011; dotted lines). Total soil depth is variable across the domain ranging from roughly 1.5 to 2.7 m. Mean and max precipitation are separated for ease of viewing.

persistent thereafter, while autocorrelations become almost negligible beyond 5 months. The L13 magnitude and autocorrelations track those from observed soil moisture comparably to the original M02 data.

In addition to soil moisture, SWE is a key hydrologic state variable. Figure 4 shows histograms of the dynamic soil moisture range, mean and maximum SWE, and precipitation for M02 and L13 over the concurrent time period, as well as L13 for the extended period (1 January 1915–31 December 2011). SWE values were frequently larger for the finer spatial domain ($\frac{1}{16}°$) than the coarser ($\frac{1}{8}°$) during the concurrent period, corresponding to an increased meteorological variability, while the extended period had maximum values that were still larger. The dynamic soil moisture range was accordingly greatest for the extended period (at $\frac{1}{16}°$). Maximum daily precipitation was comparable between the two datasets over the concurrent

period; however, larger daily values were frequently recorded for the extended period corresponding to a wet period before 1925, as well as over topographically complex regions. Conversely, the mean daily precipitation values were stable across both periods and resolutions.

Simulated streamflows are compared with observations in Fig. 5 from major river basins covering large portions of the domain. For several basins, particularly in the western United States, naturalized streamflow data were obtained that have been adjusted for anthropogenic impacts, including upstream regulation, water withdrawals, and evaporation from upstream reservoirs (see Table 3). Limited VIC parameter estimation was performed to match surface and subsurface runoff from the previously calibrated VIC (version 4.0.3) used in M02. We employed a technique similar to Troy (2008) with the objective of matching the runoff
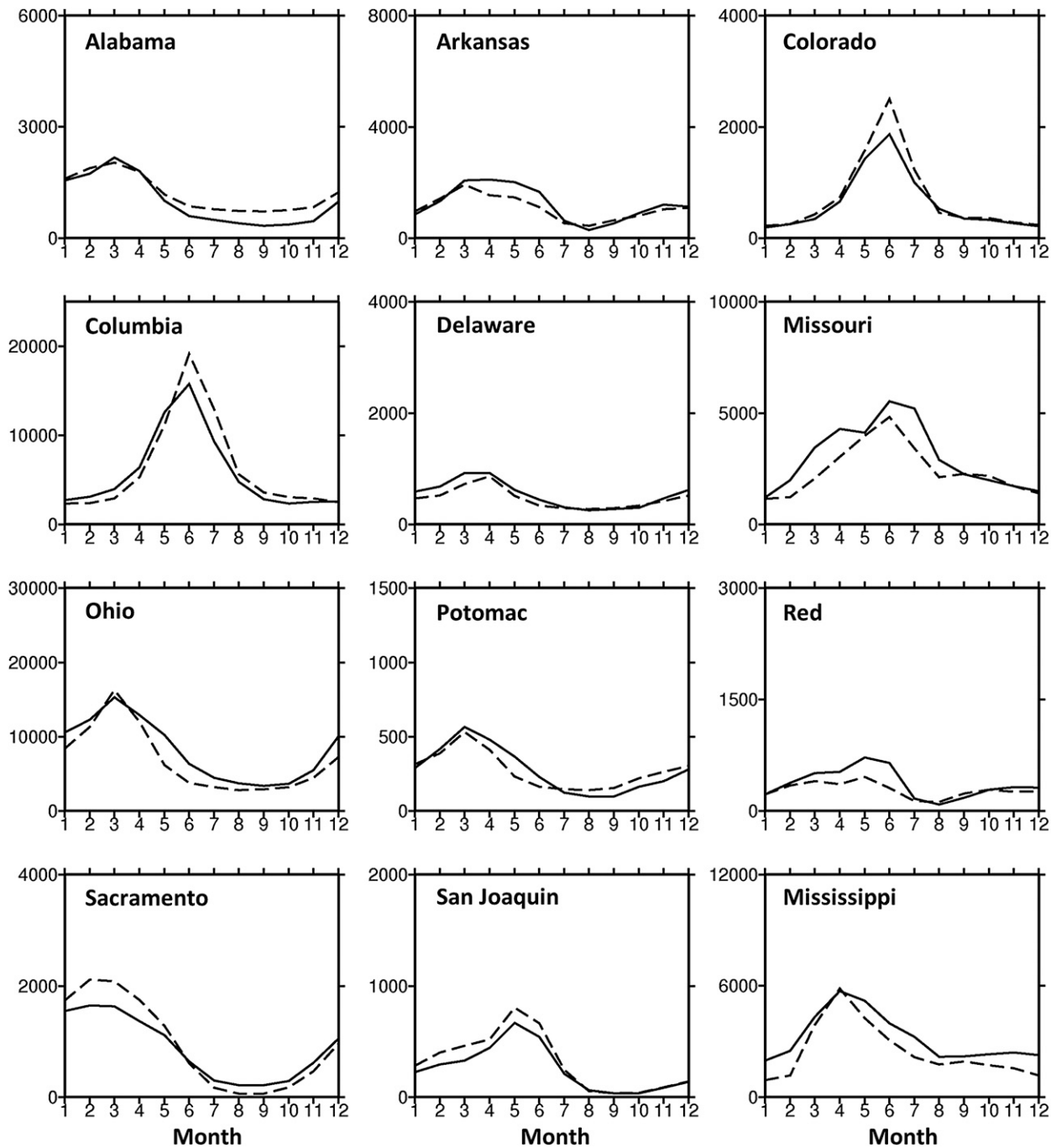
FIG. 5. Mean monthly hydrographs from L13 ($m^3 s^{-1}$) over the period 1961–90. Simulated flows are denoted by dashed lines, while observed or naturalized flows are solid lines.

ratio (in this case between model versions 4.0.3 and 4.1.2) at regularly spaced intervals of 1°. A Monte Carlo search consisting of 200 iterations was applied, which varied three VIC soil parameters—the variable infiltration curve parameter b, the maximum velocity of baseflow parameter D_{smax}, and the depth of the bottom

soil layer D_3—within a narrow range ($\pm 10\%$) of their previous values.

Offline simulations were conducted to evaluate the impact of using climatological winds prior to 1948 (see the supplemental material). These comparisons showed that with few exceptions use of the climatological winds

TABLE 3. Details of observational streamflow data used for comparison with simulated fluxes and states.

River name	Station name	Area (km^2)
Alabama	Clairborne	56 900
Arkansas*	Ralston	121 340
Colorado*	Imperial Dam	488 215
Columbia*	Dalles	613 280
Delaware	Memorial Bridge	28 500
Missouri*	Hermann	1 357 670
Ohio	Metropolis	525 760
Potomac	Point of Rocks	25 000
Red*	Index	124 390
Sacramento*	Bend Bridge	23 050
San Joaquin*	Mokelumne Hill	1860
Upper Mississippi	Grafton	443 660

* Naturalized streamflow were obtained.

slightly reduce the temporal and spatial variability of hydrologic fluxes but have small relative impacts on long-term mean values. Relative impacts on short-term (3-hourly and daily) values are greatest, and are less at monthly time steps. Given the uncertainty in using static vegetation and soil (from 2000 and 1998, respectively), the derived model outputs for the earlier part of the simulation period serve as a reference scenario (rather than a reconstruction), while providing the necessary meteorological inputs for users who might desire to produce more detailed dynamic reconstructions [as a point of reference, Matheussen et al. (2000) found maximum changes in runoff and ET of less than 10% for reconstructed 1900 versus 1990 vegetation in the Columbia River basin]. Additional uncertainty arises from using a constant lapse rate in regions of topographical complexity (i.e., western United States), with the potential to bias daily temperature range in certain cases, which may impact derived downwelling shortwave radiation based on the MTCLIM algorithm. It follows that undocumented or incomplete QC of instrument change error may hinder the robustness of trends in these data, as pointed out by Menne et al. (2009). Menne et al. (2010) subsequently showed that there has been no successful correction of the biases associated with the change from liquid-in-glass (LiG) thermometers to Maximum–Minimum Temperature System (MMTS) in the daily NCDC station dataset. This is estimated to result in an artificial negative bias in maximum temperatures less than approximately 0.5°C, which would translate into a slight underestimation of diurnal temperature range and reduction in derived shortwave radiation for L13. This instrument bias could have further implications for simulated cold-season processes (snowpack evolution) and surface heat fluxes, as well as for drought

assessment and climate-oriented analyses such as downscaling.

4. Data format and availability

The data are available in Network Common Data Form (NetCDF) format, conforming to the Assistance for Land-Surface Modelling Activities (ALMA) convention of Polcher et al. (2000). This means that moisture fluxes are expressed as kilograms per square meter per second, energy fluxes as watts per square meter, and moisture states as kilograms per square meter. (The data are freely accessible from ftp://ftp.hydro.washington.edu/pub/blivneh/CONUS/, where we also provide plots comparing a range of other states and fluxes between M02 and L13.)

5. Conclusions

We have described an observation-based hydrologically consistent dataset for the period 1915–2011 at a $\frac{1}{16}°$ spatial resolution. Gridded station data for precipitation and temperature, surface wind from an atmospheric reanalysis, and derived downward solar and longwave radiation and vapor pressure were used to force a hydrologic model that was shown to reproduce, on average, observed surface heat fluxes, soil moisture, and runoff. These data have potential uses for model evaluation and diagnosis in energy and water balance studies and climate change impact studies. We expect that these data will complement studies that have used the M02 dataset, given the wider range of conditions that are included in a longer time period and at finer spatial resolution.

REFERENCES

Andreadis, K. M., and D. P. Lettenmaier, 2006: Assimilating remotely sensed snow observations into a macroscale hydrology model. *Adv. Water Resour., 29,* 872–886.

——, P. Storck, and D. P. Lettenmaier, 2009: Modeling snow accumulation and ablation processes in forested environments. *Water Resour. Res., 45,* W05429, doi:10.1029/2008WR007042.

Bohn, T. J., B. Livneh, J. Oyler, S. W. Running, B. Nijssen, and D. P. Lettenmaier, 2013: Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models. *Agric. For. Meteor., 176,* 38–49, doi:10.1016/j.agrformet.2013.03.003.

Carpenter, T. M., and K. P. Georgakakos, 2004: Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed hydrologic model. *J. Hydrol., 298,* 202–221.

Castro, C. L., R. A. Pielke Sr., and J. O. Adegoke, 2007: Investigation of the summer climate of the contiguous United States and Mexico using the Regional Atmospheric Modeling System (RAMS). Part I: Model climatology (1950–2002). *J. Climate, 20,* 3844–3865.

Cayan, D. R., E. P. Maurer, M. D. Dettinger, M. Tyree, and K. Hayhoe, 2008: Climate change scenarios for the California region. *Climatic Change,* **87** (Suppl. 1), 21–42, doi:10.1007/s10584-007-9377-6.

Cherkauer, K. A., and D. P. Lettenmaier, 2003: Simulation of spatial variability in snow and frozen soil. *J. Geophys. Res.,* **108,** 8858, doi:10.1029/2003JD003575.

Christensen, N. S., A. W. Wood, N. Voisin, D. P. Lettenmaier, and R. N. Palmer, 2004: The effects of climate change on the hydrology and water resources of the Colorado River basin. *Climatic Change,* **62,** 337–363.

Daly, C., R. P. Neilson, and D. L. Phillips, 1994: A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *J. Appl. Meteor.,* **33,** 140–158.

Elder, K., D. Cline, A. Goodbody, P. Houser, G. E. Liston, L. Mahrt, and N. Rutter, 2009: NASA Cold Land Processes Experiment (CLPX 2002/03): Ground-based and near-surface meteorological observations. *J. Hydrometeor.,* **10,** 330–337.

Feng, X., A. Sahoo, K. Arsenault, P. Houser, Y. Luo, and T. J. Troy, 2008: The impact of snow model complexity at three CLPX sites. *J. Hydrometeor.,* **9,** 1464–1481.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, and R. Sohlberg, 2000: Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.,* **21,** 1331–1364.

Hayhoe, K., and Coauthors, 2004: Emissions pathways, climate change, and impacts on California. *Proc. Natl. Acad. Sci. USA,* **101,** 12 422–12 427.

——, and Coauthors, 2007: Past and future changes in climate and hydrological indicators in the US Northeast. *Climate Dyn.,* **28,** 381–407.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kimball, J. S., S. W. Running, and R. R. Nemani, 1997: An improved method for estimating surface humidity from daily minimum temperature. *Agric. For. Meteor.,* **85,** 87–98.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for GSMs. *J. Geophys. Res.,* **99** (D7), 14 415–14 428.

Loarie, S. R., B. E. Carter, K. Hayhoe, S. McMahon, R. Moe, C. A. Knight, and D. D. Ackerly, 2008: Climate change and the future of California's endemic flora. *PLoS ONE,* **3,** 2502, doi:10.1371/journal.pone.0002502.

Matheussen, B., R. L. Kirschbaum, I. A. Goodman, G. M. O'Donnell, and D. P. Lettenmaier, 2000: Effects of land cover change on streamflow in the interior Columbia River Basin (USA and Canada). *Hydrol. Processes,* **14,** 867–885.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *J. Climate,* **15,** 3237–3251.

——, D. P. Lettenmaier, and N. J. Mantua, 2004: Variability and predictability of North American runoff. *Water Resour. Res.,* **40,** W09306, doi:10.1029/2003WR002789.

Menne, M. J., C. N. Williams, and R. S. Vose, 2009: The U.S. Historical Climatology Network monthly temperature data, version 2. *Bull. Amer. Meteor. Soc.,* **90,** 993–1007.

——, ——, and M. A. Palecki, 2010: On the reliability of the U.S. surface temperature record. *J. Geophys. Res.,* **115,** D11108, doi:10.1029/2009JD013094.

Miller, D. A., and R. A. White, 1998: A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. *Earth Interact.,* **2,** doi:10.1175/1087-3562(1998)002<0001:ACUSMS>2.3.CO;2.

Polcher, J., and Coauthors, 2000: GLASS: Global Land–Atmosphere System Study. *GEWEX News,* No. 10 (2), International GEWEX Project Office, Silver Spring, MD, 3–5.

Prata, A. J., 1996: A new long-wave formula for estimating downward clear-sky radiation at the surface. *Quart. J. Roy. Meteor. Soc.,* **122,** 1127–1151.

Robock, A., K. Y. Vinnikov, G. Srinivasan, J. K. Entin, S. E. Hollinger, N. A. Speranskaya, S. Liu, and A. Namkhai, 2000: The Global Soil Moisture Data Bank. *Bull. Amer. Meteor. Soc.,* **81,** 1281–1299.

Salathé, E. P., 2003: Comparison of various precipitation downscaling methods for the simulation of streamflow in a rain-shadow river basin. *Int. J. Climatol.,* **23,** 887–901.

Sheffield, J., G. Goteti, F. Wen, and E. F. Wood, 2004: A simulated soil moisture based drought analysis for the United States. *J. Geophys. Res.,* **109,** D24108, doi:10.1029/2004JD005182.

Smith, M. B., D. Seo, V. I. Koren, S. M. Reed, Z. Zhang, Q. Duan, F. Moreda, and S. Cong, 2004: The distributed model intercomparison project (DMIP): Motivation and experiment design. *J. Hydrol.,* **298,** 4–26.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger, 2004: Changes in snowmelt runoff timing in western North America under a "business as usual" climate change scenario. *Climatic Change,* **62,** 217–232.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.,* **93,** 485–498.

Tennessee Valley Authority, 1972: Heat and mass transfer between a water surface and the atmosphere. Laboratory Rep. 14, Tennessee Valley Authority, Norris, TN, 270 pp.

Thornton, P. E., and S. W. Running, 1999: An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation. *Agric. For. Meteor.,* **93,** 211–228.

——, H. Hasenauer, and M. A. White, 2001: Simultaneous estimation of daily solar radiation and humidity from observed temperature and precipitation: An application over complex terrain in Austria. *Agric. For. Meteor.,* **104,** 255–271, doi:10.1016/S0168-1923(00)00170-2.

Troy, T. J., 2008: An efficient calibration method for continental-scale land surface modeling. *Water Resour. Res.,* **44,** W09411, doi:10.1029/2007WR006513.

Westerling, A. L., H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam, 2006: Warming and earlier spring increase western U.S. forest wildfire activity. *Science,* **18,** 940–943.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change,* **62,** 189–216.