5.2 Multi-Decade Analysis of Record for Hydrologic Model Calibration

David Kitzmiller[1], Wanru Wu[1], Yu Zhang[1], Dennis Miller[1], and Ziya Zhang[2]

[1]Office of Hydrologic Development
National Weather Service, NOAA
Silver Spring, Maryland

[2]University Corporation for Atmospheric Research
Boulder, Colorado

1. Introduction

Hydrologic prediction operations within the National Weather Service (NWS) require periodic recalibration of rainfall-runoff and routing models. The calibration process in turn requires high-quality time series estimates of precipitation, temperature, and river discharge covering multiple years (Smith et al. 2003). Adaptable parameters within the runoff and routing models are tuned so that errors in model output streamflow are optimized relative to observed discharge.

To date, NWS weather datasets for calibration have been based largely on station reports of temperature and precipitation, spatially interpolated to basin-average values, which then drive lumped-parameter river models. The basin average values are generally constrained to agree with some long-term climatology, and must be physically realistic relative to observed streamflow and assumptions about long-term evapotranspiration processes.

There is a need to adapt this general approach to gridded inputs. First, some hydrologic prediction operations are better served by spatially distributed, rather than lumped, river models. Such distributed models required gridded, rather than basin-average, inputs. However, basin averages can also be easily estimated from grids of values. Second, considerable effort must be expended in the collection and quality control of station inputs. It would be advantageous to make use of several existing product suites to develop gridded analyses of weather variables.

Our initial goal in this effort is to produce temperature and precipitation gridded time series using previously-published climatology and monthly and hourly time series, covering multiple decades starting in 1979. These analyses will later be enhanced with station observations such as have been used in calibration datasets in the past.

This extended abstract includes seven additional sections, as follows. Section 2 has a description of the overall approach to developing hourly data grids from inputs with multiple time scales. Section 3 describes the map projection used in the current work. Section 4 covers the compositing of grids of 30-year (1981-2010) monthly climatology of total precipitation and maximum and minimum temperature. Section 5 describes generation of monthly time series constrained to agree with this 30-year climatology, while Sections 6 and 7 show some results of verification of the grids

Corresponding author address: David Kitzmiller, W/OHD-12, 1325 East West Highway, Silver Spring, MD 20910; email <David.Kitzmiller@noaa.gov>

relative to Global Historical Climate Network and other station reports over the United States, Canada, and Mexico. Section 8 describes the next steps, which will result in hourly time series of temperature and precipitation.

2. Basic approach to developing climatic, monthly, and hourly gridded records

In the initial phases of this development effort, we wished to make optimum use of previously-published gridded climatic and weather information, particularly long-term climatologies and monthly and hourly time series. Climate grids of mean monthly precipitation, maximum and minimum temperature, and other variables, contain important information that can be used to downscale coarse-mesh inputs and constrain time series information toward a realistic long-term mean.

First, datasets of monthly, gridded climatologies for precipitation and for maximum and minimum temperature (T_{mx} and T_{mn} respectively) from previously published sources. These climatologies (i.e., one mean value per month) cover the period 1981-2010.

We next constructed monthly time series of total precipitation and mean T_{mx} and T_{mn}. These were taken from existing data sources that cover from the early 20th century to the present. The 1981-2010 mean values for these variables were calculated and compared with the climatic means. Bias adjustment factors for each grid box, for each of the twelve months, were derived from this comparison. The bias factors were applied to the monthly time series over the entire 1979-2012 period, so that the mean of the corrected time series matches the climatic values.

Finally, following the approaches of Abatzoglou (2011) and Zhang et al. (2012), an hourly time series of temperature and precipitation from the North American Land Data Assimilation System Version 2 (NLDAS2; Cosgrove et al. 2003; Xia et al. 2013) will be bias-corrected to agree with the climatologically-corrected monthly time series values. That is, the hourly precipitation and temperature will be adjusted so that, for each month, the total precipitation and the means of the T_{mx} and T_{mn}, match the monthly values derived in the previous step.

This approach insures some time stability and spatial consistency in the final hourly time series, since the monthly time series and 30-year mean grids generally incorporate more station inputs than do the hourly (Abatzoglou 2011).

3. Map grid projection

The gridded inputs described in the previous section are all distributed in multiple latitude-longitude projections, with different origins, and therefore require remapping. However, the spatial framework within NWS hydrologic software is heavily geared toward the Hydrologic Rainfall Analysis Projection (HRAP) grid (Reed and Maidment 1999). To facilitate analysis and testing of our weather records, we have mapped all data inputs to this projection, which is polar stereographic, oriented along the -105° E meridian, with a mesh length of 4,762.5 m at 60° N. The mesh length is ~4 km over most of the CONUS. Later, the grids will be further downscaled to 1/4th of the original HRAP grid, a fine-mesh projection also used in some NWS modeling. As necessary, the data can be remapped to other projections, given appropriate elevation and climatology background information.

4. Long-term climatology, 1981-2010

It is generally desirable to initially constrain long-term monthly means of the weather variables to some standard that has incorporated a large number of quality-controlled observations. There are a number of gridded climatology datasets covering our area of interest (summarized by Riverside Technology 2012). A commonly-used source for the United States is gridded monthly mean precipitation and maximum/minimum temperature derived by the Parameter-Elevation Relationships on Independent Slopes Method (PRISM; Daly et al. 2008). These have been most recently derived for the 1981-2010 period, and have been published for general use on a 30 arc-second grid mesh, corresponding to roughly 0.008° or approximately 1-km mesh length over the conterminous United States (PRISM Climate Group 2014). In the current study we have applied these climatology grids to the Conterminous United States (CONUS) portion of the domain. Because the grid mesh of the PRISM data is finer than the 4-km mesh initially used here, the 4-km HRAP grid values were taken as the mean of all PRISM values lying within each grid box.

A number of high-resolution monthly climatologies have also been published for the remainder of our project domain; these include WorldClim (Hijmans et al. 2005), which covers the global land surface, and the Atlas Climático Digital de México (Fernandez-Eguiarte et al. 2010), which covers Central America, Mexico, and the southern CONUS. However, these sources have been generated from a much longer data record than the most recent thirty years. We therefore chose to apply 0.04° grids generated by the NOAA National Climatic Data Center (NCDC), originally for the purpose of defining mean values for climatic zones in and near the CONUS (Russell Vose, personal communication 2012). These were derived from quality-controlled monthly station data from 1981-2010 through the ANUSPLIN method (Hutchinson 1991). This method executes geospatial interpolation based on local relationships between the meteorological variable of interest, and latitude, longitude, and elevation. Because the latitude-longitude grid mesh of the NCDC data is more coarse than that of HRAP in portions of the domain, values for HRAP grid boxes that had no NCDC data were estimated from the local average of nearby NCDC grid points. A spatial blending of these grids, hereafter referred to as NCDC, with the PRISM grids, thus covers our target domain with high-resolution information for a common, recent period.

For the full-domain climatology grids, we applied PRISM data wherever available over the CONUS, and the original NCDC data over Mexico and Canada in areas removed from national borders. We found generally good agreement between the PRISM and NCDC datasets wherever they overlapped. Some small discontinuities existed at CONUS borders. In a spatial belt of approximately 120 km over southern Canada and northern Mexico, the NCDC data were applied, after a bias correction based on collocated PRISM/NCDC data in the nearby CONUS regions. This correction is similar to that applied to coverage boundaries between different precipitation products in the NLDAS2 grids (Cosgrove et al. 2003).

Examples of the resulting HRAP grids of mean total precipitation, T_{mx}, and T_{mn}, for January and July, are shown in Fig. 1. Precipitation features related to upslope enhancement and the position of Gulf of Mexico water vapor in the

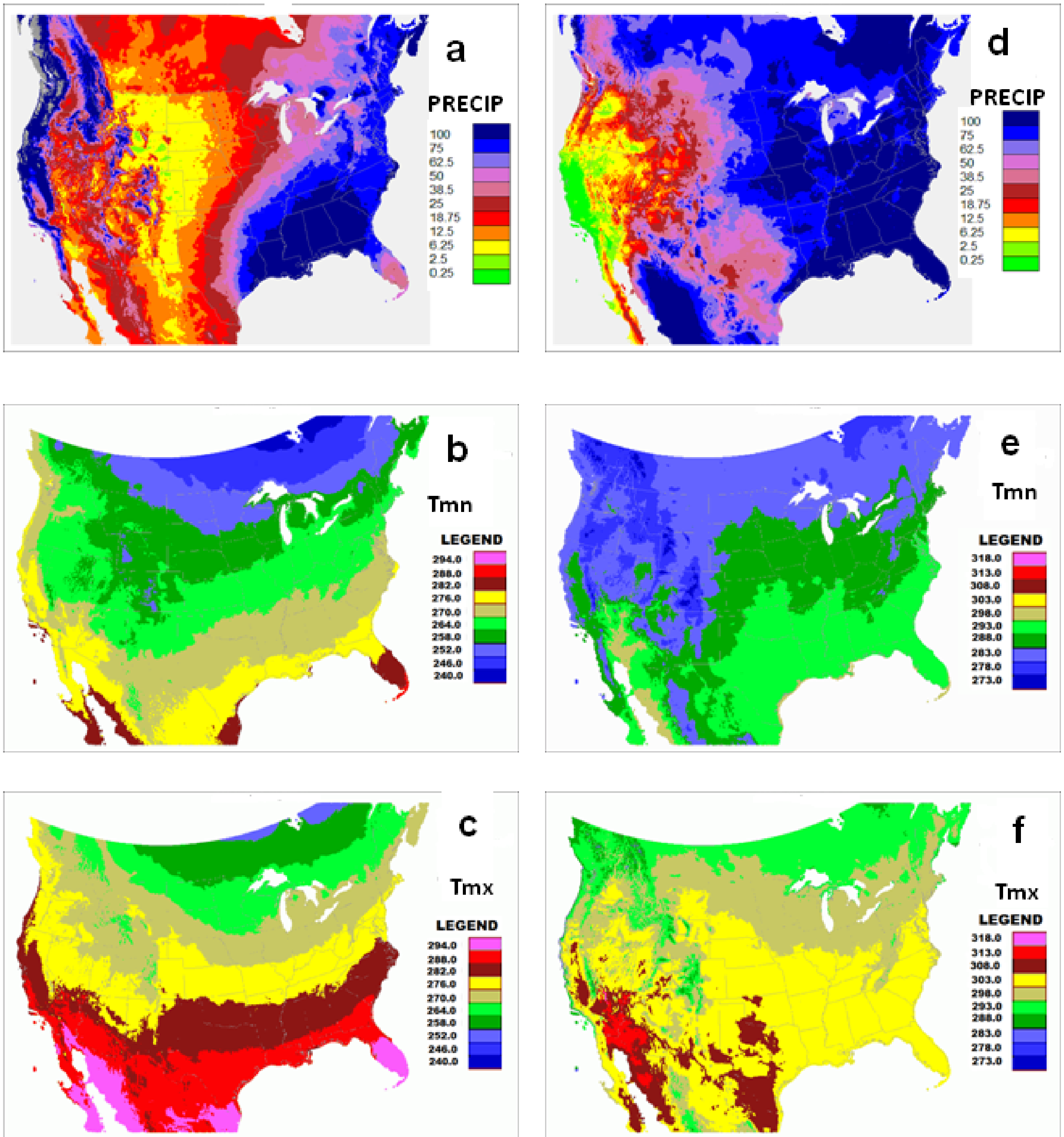

Figure 1. Spatially-blended PRISM-NCDC 1981-2010 climatic fields. January mean values of (a) precipitation, mm; (b) T_{mn} (K) and (c) T_{mx}; (K). July mean values for (d) precipitation, (e) T_{mn}, and (f) T_{mx}. Note that temperature scales change from b,c to e,f.

central CONUS (Fig. 1a,d) are clearly shown. The south-north increase in January precipitation in the area of the U.S.-Canadian Plains (Fig. 1a) is climatically realistic, and appears in the NCDC grids as well. Terrain-related temperature features (Fig. 1b,c and Fig. 1e,f) are also rendered with realistic detail across the national border regions.

5. Monthly time series of total precipitation, mean T_{mx}, and mean T_{mn}, 1979-2010

While it might be possible to apply these 30-year climatic means directly to bias correction of hourly NLDAS2, we found evidence of time variation in the bias of certain NLDAS2 records.

Therefore we identified monthly time series datasets that were developed independently of NLDAS2 and of the North American Regional Reanalysis (NARR; Mesinger et al. 2004) that contributed to NLDAS2 (Cosgrove et al. 2003). These monthly time series datasets will be used to adjust the NLDAS2 hourly data.

As noted by Abatzoglou (2011), some published monthly gridded climatic records generally feature more station inputs than do the hourly NLDAS2 records, as well as higher spatial resolution than do the $1/8^{th}$ ° NLDAS2 grids. He developed multi-decade hourly time series datasets of several variables by applying monthly bias correction factors to the NLDAS2 data. The monthly time series were those generated by the PRISM group. However, the datasets of Abatzoglou were limited to the CONUS, and they were created prior to the availability of the 1981-2010 climatology described above in section 4. We used the PRISM-NCDC climatology, and both the PRISM and Tyndall Research Centre monthly time series, to generate bias-corrected monthly datasets covering the larger domain.

The PRISM monthly time series that are publicly available cover the CONUS with a 120 arc-second (~4-km) grid mesh. These data are available as far back as 1900. It should be noted that because of variations in station networks and processing procedures, the means of these data are not insured to match those of the 1981-2010 climatology (PRISM Climate Group 2014). Therefore we calculated and applied bias corrections, based on differences between the 1981-2010 climatology and corresponding monthly time series means, to the original monthly time series. Bias correction factors are multiplicative for precipitation and additive for temperature. We have assumed that the same bias corrections would be applicable in 1979-1980 and 2011-2012. As with the 30-year data, we applied PRISM in the monthly time series wherever it was available.

To obtain monthly-scale information outside the CONUS, we chose time series data of the Tyndall Research Centre Climate Research Unit, henceforth referred to as CRU (Mitchell and Jones 2005). These data cover land surfaces with a 0.5° grid, and cover the period 1900-2012. The data were interpolated to the HRAP grid by applying an inverse-distance weighting analysis with a 1.5° radius of influence, consistent with the feature resolution of the original 0.5^{o} grid.

While the NARR dataset includes the same variables and has a higher spatial resolution than does the CRU, it extends only back to 1979, and has certain time varying precipitation biases that argue against its direct use. In particular, wintertime precipitation over Quebec and Ontario was biased very low prior to 1993, an artifact created by data scarcity and features of the precipitation assimilation algorithm within the reanalysis (Wesley Ebisuzaki, personal communication, 2012).

The process of bias correction of monthly gridded time series toward agreement with 30-year climatic grids is illustrated schematically in Fig. 2. It should be noted that the long-term bias correction does not affect time-dependent information within the monthly time series; the

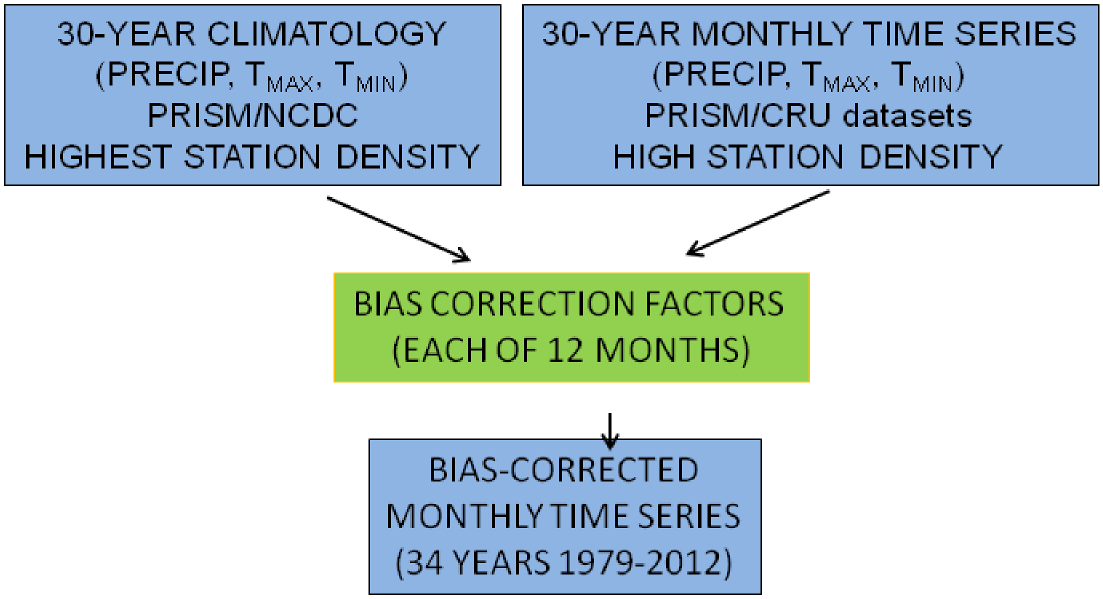

Figure 2. Schematic of data flow for development of bias-adjusted monthly time series of total precipitation and mean T_{mx} and T_{mn}. Climatic grids of 1981-2010 mean values (upper left) and compared with mean time series values (upper right) to derive grids of bias correction factors, which are applied to data within and near the 1981-2010 time period (bottom)

bias corrections are constant for each grid box, month of the year, and weather element. Though the precipitation climatology is not necessarily consistent with observed runoff, and must be adjusted in some locations during the hydrologic calibration process, prior experience has shown the PRISM datasets to be generally reliable.

Though the CRU data outside the CONUS are at much lower spatial resolution than the final 4-km grid mesh, bias correction of this monthly data to match the 1981-2010 climatology effectively downscales it to reflect higher-detail terrain-related features. The bias correction also insures spatial continuity across the PRISM-CRU coverage boundaries.

This downscaling effect is evident in the original spatial merging of January 1981 mean T_{mx}, and after downscaling using PRISM-NCDC climatology (Fig. 3a,b). Terrain-related features over Canada and Mexico are clearly delineated in the bias-corrected form (Fig. 3b).

It must be noted that we used PRISM monthly time series of precipitation and temperature that were distributed prior to the release of new data by the PRISM group in 2013. We will investigate the application of this new monthly time series, as well as daily time series (PRISM Climate Group 2014) in the near future.

6. Effectiveness of bias correction of monthly time series: Temperature

To verify that the bias corrections applied to the NLDAS2 monthly time series were effective in improving their absolute accuracy, we compared the PRISM-CRU values to those reported at observing sites in the Global Historical Climate Network (GHCN; Durre et al.

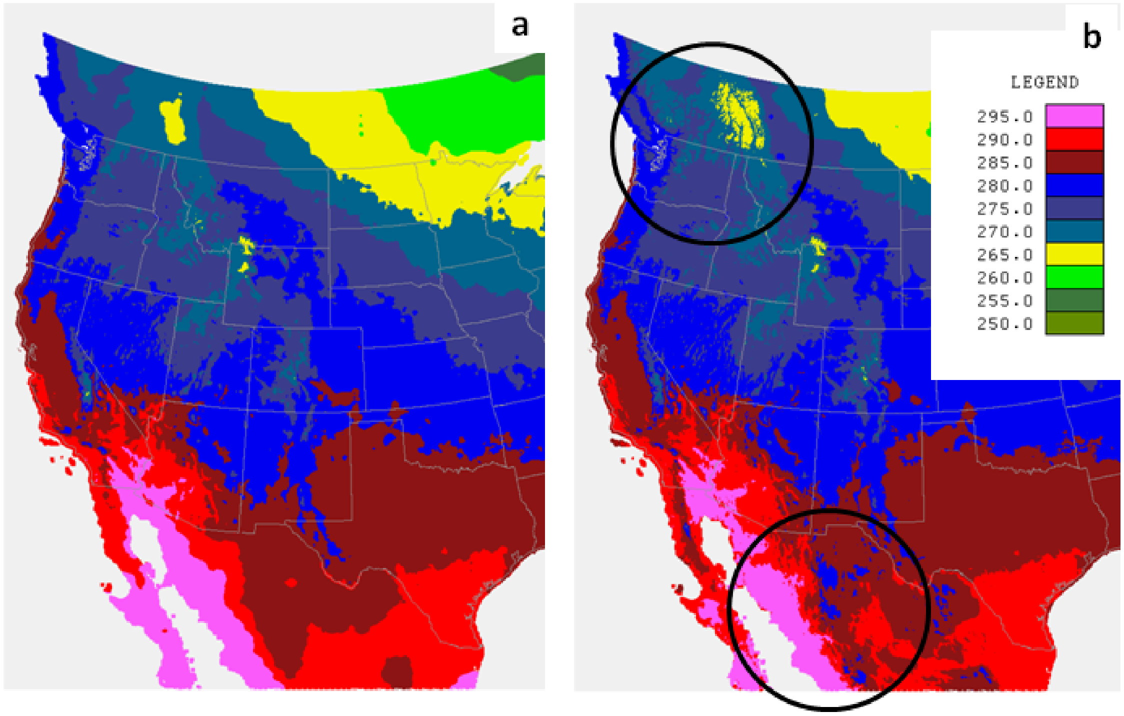

Figure 3. Effects of downscaling and bias adjustment on a temperature field. Original spatial blend of monthly PRISM and CRU T_{mx} (a) gains terrain- and climatology-related detail when adjusted for agreement to 30-year 4-km climatic fields (b); see especially circled regions. The field is for mean T_{mx} in January 1981, K.

2010). Monthly total precipitation and mean T_{mx}/T_{mn} values during the 1979-2011 period were taken from GHCN data, and compared with point values extracted from NLDAS2 and bias-corrected PRISM-CRU grids. We compared the multi-year means of data from individual months, and the year-to-year time series correlations at individual sites for those months. Samples from one month per season, namely January, April, July and October, were tested in this fashion. Results for the winter and spring (melt) seasons are shown below.

The number of sites meeting GHCN standards and reporting data for any significant portion of the 1979-2011 period is small, particularly for Mexico and southern Canada. We present results only from those sites reporting for at least 10 years within this 33-year window. This amounted to ~25 sites in the CONUS, 9-11 sites over Canada, and 6-10 sites in northern Mexico. The distribution of these sites for January, for those reporting T_{mx}/T_{mn} and those reporting precipitation, are shown in Fig. 4a,b, respectively.

Comparison of long-term means and time series correlations can best be explained through Fig. 5, which covers January temperatures. Figure 5a shows the multi-year means of the PRISM-CRU and NLDAS2 minimum temperatures as functions of the corresponding GHCN values. The NLDAS2 values are generally biased high by 1°C to 2°C, while the PRISM-CRU values were generally biased slightly low. The time-series correlations, that is, correlations between January temperatures over multiple years, are generally > 0.9 (R^2 > 0.8). As shown in

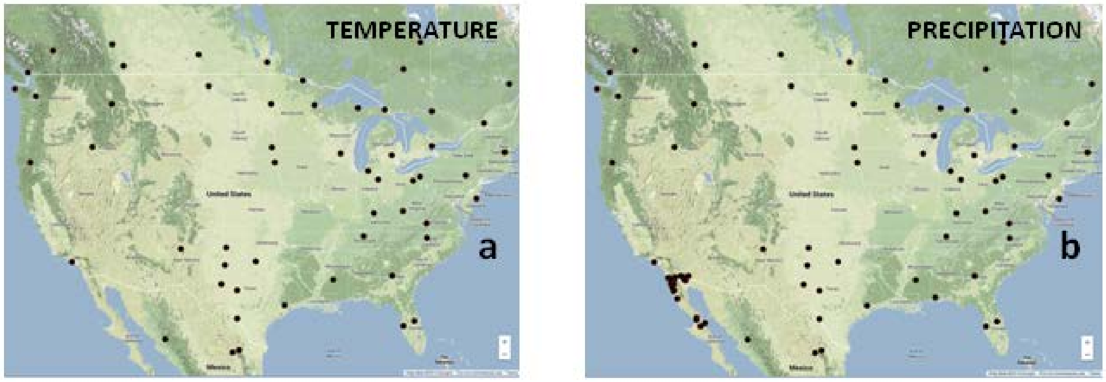

Figure 4. Distribution of GHCN sites used in verification of monthly means and time series, for (a) temperature and (b) precipitation.

Fig. 5b, where the PRISM-CRU correlation is plotted as a function of the NLDAS2 correlation, the correlations for PRISM-CRU were generally higher.

Similar results were realized for January T_{mx} (Fig. 5c,d), though neither NLDAS2 nor PRISM-CRU temperatures appeared to have a domain-wide bias (Fig. 5c). Long-term means were more accurate than for T_{mn} (smaller scatter about the diagonal reference line) and time-series correlations were slightly higher. This finding might because the models used to interpolate station point values to grids are most accurate for afternoon conditions, when solar heating and vertical mixing drive lower tropospheric temperature profiles toward the adiabatic, and make the terrain height a very effective predictor of horizontal temperature variations. Spatial differences in night time temperatures are more difficult to model, being strongly influenced by local radiative effects and recurring temperature inversions (Daly et al. 2008).

Comparisons for April, during the western North American melt season, yielded similar results. As in the January cases, the NLDAS2 mean values were biased slightly high for minimum temperatures, but neither time series had a systematic bias relative to maximum temperatures (Fig. 6a,c). Again, time-series correlations for T_{mx} were higher than those for T_{mn}, and the corrected PRISM-CRU estimates were more strongly correlated with station values than were the NLDAS2 (Fig. 6b,c). Correlations were lower for the April T_{mn} time series than for the January T_{mn}.

Very similar results, not shown here, were obtained for summer (July) and autumn (October). The findings indicate that the adjusted PRISM-CRU temperature time series had generally smaller biases, relative to GHCN station reports, than did the NLDAS2 time series; the PRISM-CRU time series also generally had a higher temporal correlation to GHCN. Based on these findings, we are confident that our approach to constructing gridded monthly time series for mean T_{mx} and T_{mn} is sound.

7. Effectiveness of bias correction of monthly time series: Precipitation

a. CONUS and Canada

For monthly precipitation time series, 13 GHCN sites in Canada and 40 sites in the CONUS were

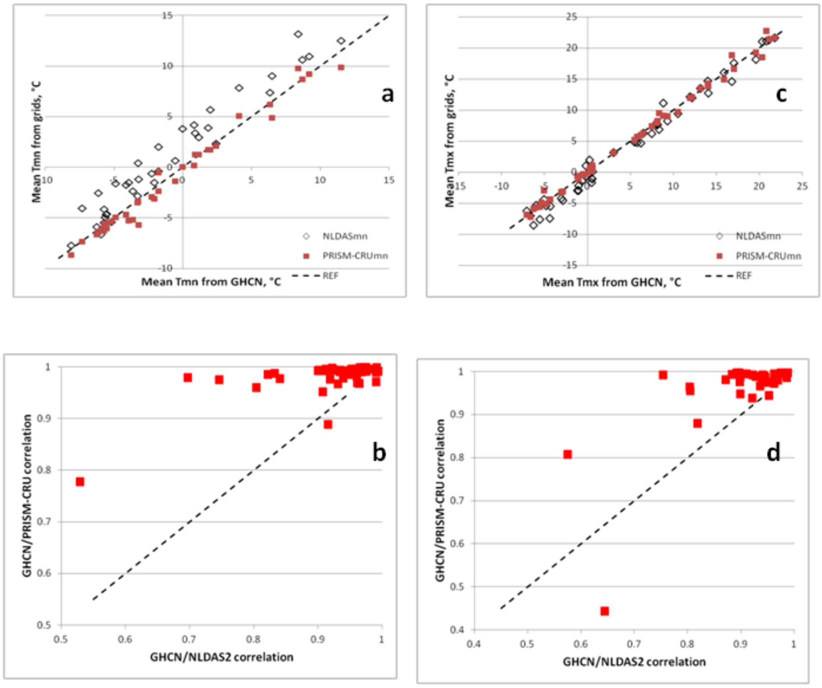

Figure 5. Verification of gridded monthly gridded Tmx and Tmn time series relative to GHCN station reports, January. Accuracy of long-term means is shown in (a) and (c) for Tmx and Tmn respectively; open diamonds for NLDAS2 and red squares for PRISM-CRU. Time series correlation of the PRISM-CRU grids is plotted as a function of NLDAS2 correlation for (b) Tmx and (d) Tmn, respectively,

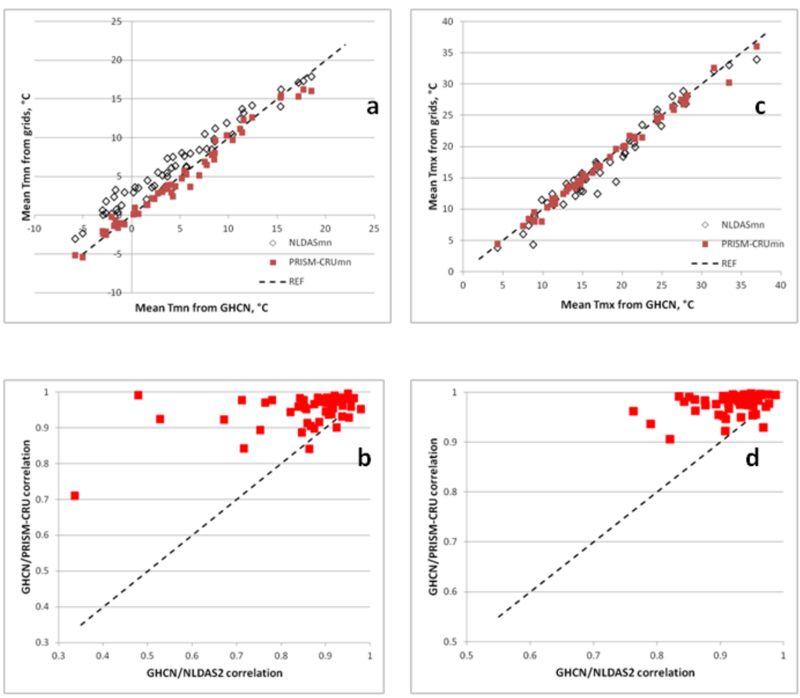

Figure 6. Temperature verification as in Fig. 5, except for April.

available for verification. The CONUS sites all had 33 years of coverage, while the Canadian sites had 13 to 25 years' coverage. As with temperature, we analyzed the months of January, April, July, and October as being characteristic of each season.

Verification of long-term average precipitation showed that both gridded datasets had no major systematic bias relative to the GHCN reference sites. The PRISM/CRU datasets had slightly smaller errors than did the NLDAS2. Errors were smaller for the January and April precipitation mean totals than for the July and October totals (Figs. 7b, 8b, 9b, and 10b).

We found that for all four months, both the NLDAS2 and PRISM/CRU datasets generally had monthly time series correlations > 0.8 (R^2 > 0.65) relative to the GHCN reference, and most of the PRISM/CRU datasets had correlations > 0.9 (R^2 > 0.8). As shown in Figs. 7a, 8a, 9a, and 10a, the PRISM/CRU time series generally had a higher correlation to GHCN than did the NLDAS2. Time series correlations were somewhat higher for January and April than for July and October. This might be due to the effect of locally-heavy convective precipitation in July, and generally light precipitation in October.

We found that while the monthly NLDAS2 time series had good accuracy relative to station reports from the Global Historical Climate Network (GHCN) dataset, application of the 30-year mean climatology and monthly gridded data improved existing biases and the monthly time series correlation.

b. Mexico

There was only a limited number of Mexican GHCN stations with reports over most of the 1979-2011 period. We found only two suitable GHCN precipitation stations in the area of the Rio Conchos basin, which is a major contributor to the Rio Grande, and the majority of precipitation stations were concentrated in the semi-arid northwestern part of the country.

In general, the adjusted CRU monthly precipitation records appeared poorly correlated with the GHCN reports (results not shown here). This was particularly true for the spring season (April). To further investigate this potential problem, we are examining a larger, alternative verification set of monthly totals from 157 stations over northern Mexico, published by the Mexican Servicio Meteorológico Nacional (SMN) (SMN 2013). These include a number of sites over the central and eastern parts of the Mexican NLDAS2 domain. To rectify any deficiencies confirmed by results from this larger dataset, we might choose to modify the 30-year climatology and/or monthly time series grids for the northern Mexico area.

It must be noted that additional quality-controlled hourly station information will be entered into the temperature and precipitation grids prior to their use in hydrologic model calibration. These additional data should mitigate systematic and random errors in the reanalysis first guess, such as might be present over Mexico.

8. Future work

Following final adjustments, the monthly time series datasets described in preceding sections will be applied to bias adjustment of hourly NLDAS2 time series, following the methods used by Abatzoglou (2011) and Zhang et al. (2011). Initial results show that the NLDAS2 contains considerable correct detail, based on comparisons with observed station hourly temperature and precipitation. Tests of the hourly time series of 4-km grids through hydrologic modeling will then be carried out.

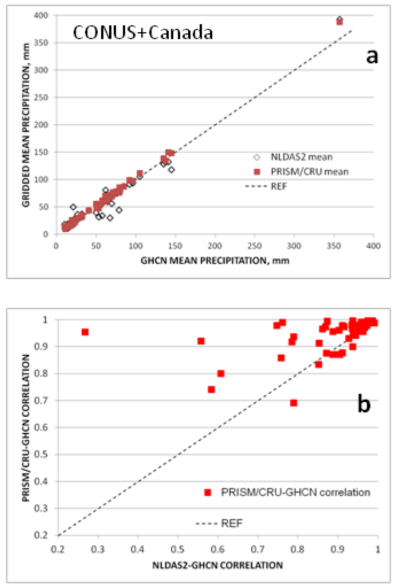

Figure 7. As in Fig. 5, except for monthly total precipitation verification for CONUS and Canada, January.

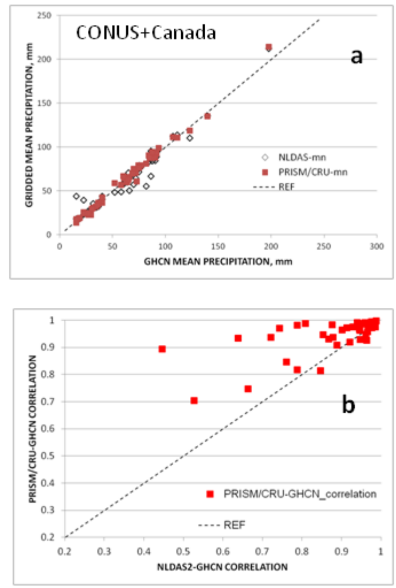

Figure 8. As in Fig. 7, except for April.

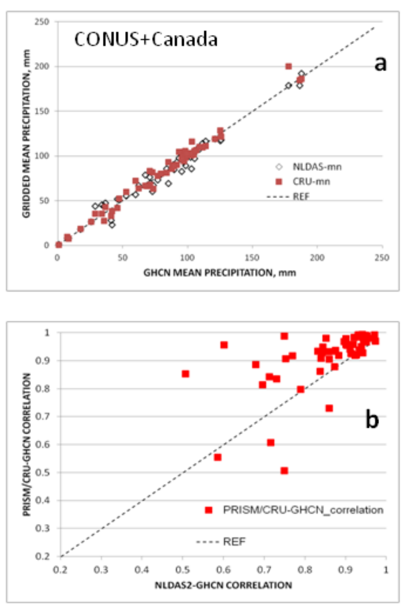

Figure 9. As in Fig. 7, except for July.

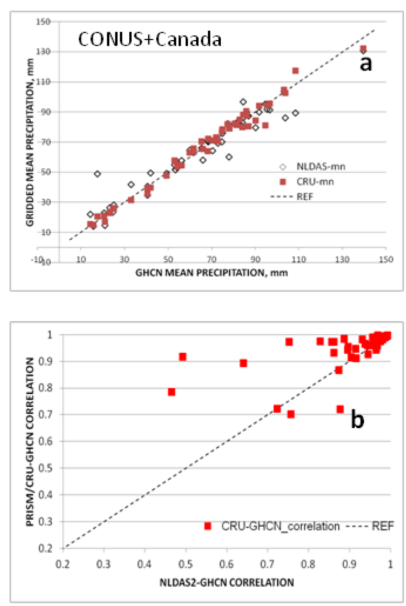

Figure 10. As in Fig. 7, except for October.

Acknowledgements

This work has been supported by the Office of Hydrologic Development through the Advanced Hydrologic Prediction System (AHPS initiative). We are indebted to Amy Volckens and Shaun Carney of Riverside Technology, Inc., and staff of several River Forecast Centers, who contributed valuable information to the development process. Russell Vose of NCDC provided North American climatic grids. NLDAS2 data were obtained from NASA

Goddard Space Flight Center (http://ldas.gsfc.nasa.gov/nldas/NLDAS2forcing.php).

References

Abatzoglou, J, 2011: Development of gridded surface meteorological data for ecological applications and modeling. *International Journal of Climatology.* (http://onlinelibrary.wiley.com/doi/10.1002/joc.3413/full).

Cosgrove, B. and Coauthors, 2003: Real-time and retrospective forcing in the NorthAmerican Land Data Assimilation System (NLDAS) project. *Journal of Geophysical Research*, **108**, doi:10.1029/2002JD003118.

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Doggett, M.K., Taylor, G.H., Curtis, J., and Pasteris, P.A. 2008. Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *International Journal of Climatology*, **28**, 2031-2064.

Durre, I. , M. Menne, B. Gleason, T. Houston, and R. Vose, 2010: Comprehensive automated quality assurance of daily surface observations. *Journal of Applied Meteorology and Climatology*, **49**, 1615-1633.

Fernandez-Eguiarte A., J. Zavala-Hidalgo., R. Romero-Centeno. 2010: Atlas Climático Digital de México. Centro de Ciencias de la Atmósfera. UNAM. [http://uniatmos.atmosfera.unam.mx/]

Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones and A. Jarvis, 2005: Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology,* **25**, 1965-1978.

Hutchinson M, 1991: The application of thin plate smoothing splines to continent-wide data assimilation. Jasper JD (ed.) *BMRC Research Report No.27, Data Assimilation Systems*. Melbourne: Bureau of Meteorology: 104-113.

Mesinger, F. and Coauthors 2006: North American regional reanalysis. *Bulletin of the American Meteorological Society*, **87**, 343-360.

Mitchell, T. and P. Jones, P.D., 2005: An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, **25**, 693-712.

PRISM Climate Group, 2014: http://www.prism.oregonstate.edu/

Reed, S., and Maidment, D., 1999: Coordinate transformations for using NEXRAD data in GIS-based hydrologic modeling. *Journal of Hydrologic Engineering*, **4**, 174–182.

Riverside Technology, Inc. 2012: Hydrometeorological Forcings for Hydrologic Prediction and Hydrologic Model Calibration. Technical Report for Office of Hydrologic Development, 132 pp. [Available at: http://amazon.nws.noaa.gov/articles/HRL_Pubs_PDF_May12_2009/New_Scans_June_2012/_Riverside_AOR_Final_Report_Rev1.pdf]

Servicio Meteorológico Nacional, México, 2013: http://smn.conagua.gob.mx/ [accessed 2 December 2013]

Smith, M., D. Laurine, V. Koren, S. Reed, and Z. Zhang, 2003: Hydrologic model calibration in the National Weather Service. In *Calibration of Watershed Models* (Q. Duan, H. Gupta, S. Sorooshian, A. Rousseau, and R. Turcotte, ed.), American Geophysical Union, Washington DC., 345 pp.

Xia, Y., B. Cosgrove, M. Ek, J. Sheffield, L. Luo, E. Wood, K. Mo, and NDLAS team 2013: Overview of the North American Land Data Assimilation System (NLDAS). *Land Surface Observation, Modeling and Data Assimilation*. S. Liang, X. Li, and X. Xie, eds. World Scientific Publishing Company, 492 pp.

Zhang, Y., S. Reed, and D. Kitzmiller, 2011: Effects of retrospective gauge-based readjustment of multisensor precipitation estimates on hydrologic simulations. *Journal of Hydrometeorology*, **12**, 429–443.