Earth Syst. Sci. Data Discuss., 8, 401–425, 2015
www.earth-syst-sci-data-discuss.net/8/401/2015/
doi:10.5194/essdd-8-401-2015

Open Access Earth System Science Data Discussions

A global satellite assisted precipitation climatology

C. Funk[1,3], A. Verdin[2], J. Michaelsen[3], P. Peterson[3], D. Pedreros[1,3], and G. Husak[3]

[1]U.S. Geological Survey, Earth Resources Observation and Science Center, Sioux Falls, SD, USA
[2]University of Colorado, Boulder, CO, USA
[3]University of California, Santa Barbara Climate Hazards Group, Santa Barbara, CA, USA

Received: 6 February 2015 – Accepted: 6 March 2015 – Published: 12 May 2015

Correspondence to: C. Funk (cfunk@usgs.gov)

Published by Copernicus Publications.

Abstract

Accurate representations of mean climate conditions, especially in areas of complex terrain, are an important part of environmental monitoring systems. As high-resolution satellite monitoring information accumulates with the passage of time, it can be increasingly useful in efforts to better characterize the earth's mean climatology. Current state-of-the-science products rely on complex and sometimes unreliable relationships between elevation and station-based precipitation records, which can result in poor performance in food and water insecure regions with sparse observation networks. These vulnerable areas (like Ethiopia, Afghanistan, or Haiti) are often the critical regions for humanitarian drought monitoring. Here, we show that long period of record geo-synchronous and polar-orbiting satellite observations provide a unique new resource for producing high resolution (0.05˚) global precipitation climatologies that perform reasonably well in data sparse regions.

Traditionally, global climatologies have been produced by combining station observations and physiographic predictors like latitude, longitude, elevation, and slope. While such approaches can work well, especially in areas with reasonably dense observation networks, the fundamental relationship between physiographic variables and the target climate variables can often be indirect and spatially complex. Infrared and microwave satellite observations, on the other hand, directly monitor the earth's energy emissions. These emissions often correspond physically with the location and intensity of precipitation. We show that these relationships provide a good basis for building global climatologies. We also introduce a new geospatial modeling approach based on moving window regressions and inverse distance weighting interpolation. This approach combines satellite fields, gridded physiographic indicators, and in situ climate normals. The resulting global 0.05˚ monthly precipitation climatology, the Climate Hazards Group's Precipitation Climatology version 1 (CHPclim v.1.0, http://dx.doi.org/10.15780/G2159X), is shown to compare favorably with similar global climatology products, especially in areas with complex terrain and low station densities.

ESSDD
8, 401–425, 2015

1 Introduction

Systematic spatial variations in climate have been studied since at least the first century AD, when Ptolemy's Geographia identified the earth's polar, temperate, and equatorial temperature zones. Analysis of these climatological surfaces continues to be an important aspect of environmental monitoring and modeling. In the 1960s, computers enabled the automatic interpolation of point data, and several important algorithms such as Shepard's modified inverse distance weighting function (Shepard, 1968) and optimal surface fitting via kriging (Krige, 1951; Matheron, 1963) were developed. The value of spatially continuous ancillary data, such as elevation, was soon recognized (Willmott and Robeson, 1995) and the current state-of-the-science climatologies all use background physiographic indicators combined with in situ observations. The most widely used current global climatologies, such as those produced by the University of East Anglia's Climatological Research Unit (CRU) (New et al., 1999), and the Worldclim (Hijmans et al., 2005) global climate layers, typically base their estimates on elevation, latitude, and longitude. Daly et al. (1994) used locally varying regressions fit to the topographic facets, while the CRU and Worldclim climatologies use thin-plate splines (Hutchinson, 1995) to minimize the roughness of the interpolated field, with the degree of smoothing determined by generalized cross validation.

While these approaches have proven to be very useful, as the depth of the satellite record approaches "climatological" lengths, the satellite estimates of precipitation and land surface temperatures also appear to be good candidates for auxiliary predictors. Such predictors may be especially helpful in data sparse regions with complex terrain, areas like East Africa or southwest Asia, where the climate is complicated and station density is low.

In Africa, Climate Hazards Group (CHG) scientists have demonstrated the utility of satellite fields as a source of ancillary data for climatological precipitation and air temperatures surfaces (Funk et al., 2012; Knapp et al., 2011). This new approach combines satellite fields, gridded physiographic indicators, and in situ climate normals us-

ESSDD
8, 401–425, 2015

ing local moving window regressions and inverse distance weighting interpolation. Expanding from our work in Africa, we have produced a global 0.05° monthly precipitation climatology, the Climate Hazards Group Precipitation Climatology version 1 (CHPclim v.1.0, http://dx.doi.org/10.15780/G2159X). This paper summarizes our statistical approach and modeling results, and presents a validation of the resulting dataset. The CHPclim, Worldclim, and CRU climatologies are compared with independent sets of station normals for Colombia, Afghanistan, Ethiopia, the Sahel, and Mexico.

2 Data

2.1 Precipitation normals

Two sets of monthly precipitation normals (long term averages) were used to create the CHPclim. The first set was a collection of 27 453 monthly station averages obtained from the Agromet Group of the Food and Agriculture Organization of the United Nations (FAO). This extensive collection has a fairly detailed level of representation in many typically data sparse regions, but suffers from a limitation. The FAO database does not provide the period of record used to calculate the long term averages, although most observations roughly correspond to averages over the 1950s through the 1980s. This data set, therefore, was augmented with 20 591 station climate normals taken from version two of the Global Historical Climate Network (GHCN) (Peterson and Vose, 1997). We compensated for the FAO database's varied coverage in time by supplementing it with averages from a less dense but more temporally consistent information source – the GHCN. The more extensive FAO normals were used to build the preliminary climate surfaces (as described below in Sect. 3). The differences between this surface and GHCN 1980–2009 averages were then estimated and interpolated, and then used to adjust the final monthly surfaces to a 1980–2009 time period.

ESSDD
8, 401–425, 2015

2.2 Satellite surfaces

Monthly means of five satellite products were evaluated as potential background climate surfaces: Tropical Rainfall Measuring Mission (TRMM) 2B31 microwave precipitation estimates (Huffman et al., 2007), the Climate Prediction Center morphing method (CMORPH) microwave-plus-infrared based precipitation estimates (Joyce et al., 2004), monthly mean geostationary infrared (IR) brightness temperatures (Janowiak et al., 2001), and Land Surface Temperature (LST) estimates (Wan, 2008). The TRMM and CMORPH precipitation estimates are based primarily on passive microwave observations from meteorological satellites in asynchronous orbits. The monthly mean infrared brightness temperatures, on the other hand, are derived from a combination of multiple geostationary weather satellites. The LST estimates are derived from multispectral observations from Moderate Resolution Imaging Spectrometers (MODIS) aboard the Terra and Aqua satellites. The LST fields are global, while the CMORPH, TRMM, and IR brightness temperatures span 60° N/S. For each month, for all available years (typically ~2001–2010), the satellite data were averaged. All four products were convolved to a common 0.05° grid. A fifth predictor was created based on the average of the CMORPH and TRMM precipitation fields.

2.3 Topographic and physiographic surfaces

Mean 0.05° elevation, slope, compound topographic index, flow accumulation, aspect, and slope were calculated from global 30 arc seconds GTOPO30 elevation grids following the methodology developed for the HYDRO1K (Verdin and Greenlee, 1996). While the utility of all the topographic fields was explored, only elevation and slope were used in the final analysis because they proved to be the most robust predictors. Latitude and longitude were also included as potential predictor variables.

ESSDD
8, 401–425, 2015

3 Methods – the CHG climatology modelling process

The modelling methodology involved three main steps that were repeated for each month for a set of 56 modelling regions. The extent of the regions were based on (a) station density, (b) homogeneity of predictor response, and (c) availability of the predictor fields. The first step used a series of moving window regressions (MWR) to create an initial prediction of a 0.05° precipitation grid. The second step calculated the at-station residuals from step 1 (station observations minus regression estimates), and then interpolated these values using a modified inverse-distance weighting (IDW) interpolation scheme to create grids of MWR model residuals. The gridded MWR estimates and gridded residuals were combined, to create an initial set of climatological surfaces based on the FAO normals. In the third step, these surfaces were then adjusted using the 1980–2009 GHCN station averages. The differences (ratios) from 1980–2009 GHCN climate normals were computed and used to produce final surfaces corresponding to a 1980–2009 baseline period.

3.1 Localized correlation estimates

Our process relies heavily on local regressions between our target variable and background field. We begin by explaining the bivariate standardized case of this process, which corresponds to a localized correlation. At a certain location we can sample a number of points and background variables that fall within a certain distance (d_{max}) and calculate their distance weighted (localized) correlation. The localized correlation process finds a set of n neighbouring points (within d_{max}), and estimates their weighted correlation. This study uses a cubic function of the distance (d) and a user-defined, regionally-variable, maximum distance (d_{max}).

$$w(d) = 0, d > d_{max}$$

$$w(d) = \left[1 - (d/d_{max})^3\right]^3, d \le d_{max} \quad (1)$$

ESSDD
8, 401–425, 2015

These weights are then used to estimate a localized correlation.

$$r_{x,y} = (n-1)^{-1}\left(\sum_{i=1}^{n} w(d_i)\right)^{-1} \sum_{i=1}^{n} w(d_i)\left[(x_i - \overline{x})\sigma_x^{-1}\right]\left[(y_i - \overline{y})\sigma_y^{-1}\right] \quad (2)$$

The localized correlation ($r_{x,y}$) at some location (x, y) corresponds with the standardized cross-product of the neighboring points, weighted by their distance. This process can be used to generate correlation maps (Fig. 1). Typically, the direct physical relationship between the station normals and a satellite field, such TRMM/CMORPH precipitation, results in a stronger correlation pattern than that which is produced by an indirect physiographic indicator such as elevation. Figure 1 provides an example of this by contrasting the local correlations between station precipitation, elevation and TRMM/CMORPH precipitation.

3.2 Localized moving window regressions

The core of the CHG climatology modeling process is based on a series of local regressions between in situ observations and spatially continuous predictor fields. For each location, a set of neighboring observations is obtained, and a regression model constructed using weighted least squares, with the weight of each observation determined by its distance from the regression centroid (Eq. 1).

For each region and month, a grid of center points is defined on a regular 1° grid over land-only locations. Figure 2 shows the modeling regions. At each center-point, station values within the radius (d_{max}) are collected, and a regression model is fit based on weights determined by Eq. (1). The d_{max} values are defined individually for each model region, varying from 650 km for the larger or data sparse regions (e.g. Australia, northwest Asia) to 300 km for Central America and the Galapagos.

ESSDD
8, 401–425, 2015

3.3 Model fitting

For each modeling region and month, regression models were determined through a combination of automated regression subset selection and visual inspection of the output. In some cases, visual inspection indicated that a combination of statistically powerful predictors produced obvious artifacts. In these cases, the selection pool was reduced by hand. Based on the boundaries of the interpolation window, certain predictors were omitted (TRMM, CMORPH, IR) because the satellite range did not extend northward or southward enough for these areas.

3.4 Interpolation of model residuals

Following the MWR modeling procedure, at-station anomalies (the arithmetic difference between the FAO station normals and the nearest 0.05° regression estimate) are calculated and interpolated using a modified IDW interpolation procedure. For each 0.05° grid cell, the cube of inverse distances is used to produce a weighted average of the surrounding station residuals, r. This value is then modified based on a local interpolation radius, d_{IDW} and the distance to the closest neighboring station (d_{min}).

$$r^* = \left(1 - \frac{d_{\mathrm{min}}}{d_{\mathrm{IDW}}}\right) r \tag{3}$$

This simple thresholding procedure forces the interpolated residual field to relax towards zero, based on the distance to the closest station. The d_{min} values were defined by modeling region, and ranged from 350 to 100 km, based on station density.

3.5 Rescaling by GHCN ratios

In the final stage, for each month, the regional tiles are composited on a global 0.05° grid and compared with 1980–2009 GHCN climate normals. The ratio of the GHCN and gridded climatology is calculated at each station location. These ratios are capped

ESSDD
8, 401–425, 2015

between 0.3 and 3.0, and interpolated to a 0.05° grid for each month. The values were capped to limit the potential influence of poor station data. A modified IDW procedure, similar to Eq. (3), is used, but instead of relaxing to zero, the interpolation is forced to a ratio of 1 (no change) as the distance to the minimum neighbour reaches d_{IDW}. This ratio grid is multiplied against the sum of the MWR and interpolated residuals, producing the final CHG Climatology field.

3.6 Cross-validation

Selection bias can inflate the estimated accuracy of statistical estimation procedures, producing artificial skill (Michaelsen, 1987). To limit such inflation, this study uses cross-validation. This technique removes 10 % of the station data, fits the model using the remaining 90 % of the values, and evaluates the accuracy for the withheld locations. This process is repeated ten times, eventually withholding all of the data, to produce a robust estimate of the model accuracy.

3.7 Independent validation studies

As additional validation, high quality climatology data sets were obtained for five focus regions: Afghanistan, Colombia, Ethiopia, Mexico, and the Sahel region of western Africa (Senegal, Burkina Faso, Mali, Niger and Chad). Means, spatial R^2 values, mean bias errors (MBE [mm]), mean absolute errors (MAE [mm]), percent MBE, and percent MAE statistics were evaluated. These regions (as opposed to the continental United States or Europe) were chosen to represent challenging estimation domains.

ESSDD
8, 401–425, 2015

4 Results

4.1 Model fitting results

Figure 2 shows the best predictor for each individual modeling region and the FAO station locations. For regions between 60° N and 60° S, the combined CMORPH and TRMM field tended to be the most useful predictor. The TRMM-only precipitation was selected, however, for southern Africa. Regions beyond 60° N and 60° S could not be modeled with the TRMM or CMORPH means. These regions were generally best fit with LST, slope, or elevations from a digital elevation model (DEM).

Figures 3 and 4 show the proportion of modeled cross-validated variance for the MWR and interpolated residuals components for each of the modeling regions. These results are averaged across the twelve months. For most regions, the MWR accounted for over 80 % the total variance. The interpolated residuals typically accounted for another 10–25 %. Most regions of the globe had average monthly percent errors of between 15 and 25 % (Fig. 5).

4.2 Validation studies

We next present results from our validation studies for Afghanistan, Colombia, Ethiopia, Mexico, and the Sahel (Senegal, Burkina Faso, Mali, Niger, and Chad). In each case, additional high quality gauge data were obtained from national meteorological agencies (Table 1). These data were screened, and only values not in the FAO or GHCN archive were retained. Table 1 summarizes the number of independent stations and presents the monthly validation statistics, averaged across all twelve months. For each validation station, the closest CHPclim, CRU, or Worldclim grid cell was extracted. The CHPclim percent biases were substantially smaller in magnitude than the CRU or Worldclim biases, ranging between −2 to +5 %, as compared to −28 to +16 % (CRU) or −16 to 0 % (Worldclim). While all the climatologies did well in regions with a large number of stations (e.g. Mexico and Colombia), CHPclim's performance was substantially better

ESSDD
8, 401–425, 2015

in data sparse areas like the Sahel, Ethiopia, and Afghanistan. Averaged across the study regions, the CHPclim/CRU/Worldclim datasets had overall mean absolute error (MAE) values of 16, 26 and 20 mm month^{-1}, respectively. The average spatial R^2 values for the three climatologies were 0.77 (CHPclim), 0.58 (CRU), and 0.67 (Worldclim). Overall, the CHPclim compared favorably to both the CRU and Worldclim data sets.

Plotting the monthly validation statistics provides more temporal information. Figure 6 shows monthly time series of the MAE values for each region and for each set of climatological estimates. In Afghanistan, data was only obtained for the rainy season. The low spatial correlations with the CRU and Worldclim estimates (Table 1) translate into high MAE scores (Fig. 6). In Colombia, the spatial R^2 (Table 1) and MAE time series of the CHPclim and Worldclim are similar – both perform well. In Ethiopia, the Worldclim and CRU MAE peak in concert with the seasonal rainfall maxima, while the CHPclim values remain substantially lower. This pattern is recreated for the Sahel and, to a lesser extent, for Mexico. We postulate that the CHPclim performance benefits from the fact that satellite precipitation estimates do a good job of representing heavy convection in these countries during the heart of the precipitation season. Conversely, the thin plate spline fitting procedure, combined with low gauge density in Ethiopia and the Sahel, may make it difficult to statistically represent precipitation gradients in these countries, degrading the performance of the CRU and Worldclim climatologies.

Figure 7 shows similar time series for the spatial R^2 statistics. In Afghanistan, Ethiopia, and the Sahel, the CHPclim appears substantially better at representing spatial gradient information. In Colombia and Mexico, CHPclim and Worldclim performance is similar. This may relate to the number of climate normals available in each region (cf. Fig. 2). In Colombia and Mexico, relatively dense gauge networks result in similar Worldclim and CHPclim performance. In regions with fewer stations, the correlation structure of the satellite precipitation data (Fig. 1) probably helps boost the relative performance of CHPclim.

ESSDD
8, 401–425, 2015

5 Discussion

This paper has introduced a new climatology modeling process developed by the CHG to support international drought early warning and hydrologic modeling. While this process has been applied to African rainfall and temperatures (Funk et al., 2012; Knapp et al., 2011), we report here for the first time global results, and evaluate the relative accuracy of the CHPclim v1.0 (http://dx.doi.org/10.15780/G2159X). The CHPclim is one part of the CHG's overall strategy to provide improved drought early warning information (Fig. 8). Working closely with early warning scientists from the U.S. Geological Survey's Center for Earth Resources Observation and Science (EROS), the CHG develops improved earth science tools to support food security and disaster relief for the US Agency for International Development's Famine Early Warning System Network (FEWS NET).

These activities fall into two main categories: analytic studies focused on understanding the relationship between local climate variations and large scale climate drivers (Funk et al., 2008, 2014; Hoell and Funk, 2013a, b; Liebmann et al., 2014), and the development of integrated datasets and tools supporting agro-climatic monitoring in the developing world. While early precipitation efforts focused on the use of a model (Funk and Michaelsen, 2004) to represent orographic precipitation (Funk et al., 2003), the potential issues produced by spurious model-based trends led us to focus on the use of high resolution climatologies as proxies for orographic precipitation enhancement (Funk et al., 2007). The global 0.05° CHPclim presented here is the global expansion of that work.

CHPclim provides the first component of our global precipitation monitoring system, which is built on the Climate Hazard Group Infrared Precipitation with Stations (CHIRPS, Fig. 8). The monthly CHPclim fields, described and evaluated here, have been temporally disaggregated to pentadal (five-day) means. These pentadal mean fields are then combined with 1981-near present 0.05° 60° S–60° N IR brightness (Janowiak et al., 2001; Knapp et al., 2011) precipitation estimates to produce the Cli-

ESSDD
8, 401–425, 2015

mate Hazards Group Infrared Precipitation fields (CHIRP). A modified inverse distance-weighting procedure is then used to blend these fields with global precipitation gauge station data to produce the CHIRPS (Funk et al., 2014b). These data, which benefit from the high resolution CHPclim climatology, can be used to drive a gridded crop Water Requirement Satisfaction Index model (WRSI) (Verdin and Klaver, 2002), force a special Land Data Assimilation System developed for the U.S. Agency for International Development's FEWS NET (the FLDAS), or populate interactive early warning displays like the Early Warning eXplorer (EWX, http://earlywarning.usgs.gov/fews/ewxindex.php). Improved background climatologies can enhance the efficacy of crop models, increasing their drought monitoring capacity.

Ongoing efforts are being directed towards linking seasonal forecast information with historical CHIRPS archives (Shukla et al., 2014a, b). In East Africa, for example, daily 0.05° rainfall values are used to force a hydrologic model. These results can then be combined with precipitation forecasts that translate large-scale climate conditions into region-specific predictions of CHIRPS rainfall. These rainfall forecasts can be used to drive crop and hydrologic models. In this way, for some high priority regions like East Africa, CHG scientists hope to combine the climatological constraints described by high resolution climatologies like the CHPclim, historic precipitation distributions (Husak et al., 2013), the latent information contained in the land surface state as represented by land surface models (Shukla et al., 2014b, 2013), and the foreshadowing of future weather provided by climate forecasts (Funk et al., 2014a; Shukla et al., 2014a, b). The CHPclim, described here, has been designed to provide a good foundation for this, and similar, hydrologic modeling and monitoring systems. The CHPclim and CHIRPS data sets are available at http://dx.doi.org/10.15780/G2RP4Q and http://chg.geog.ucsb.edu.

Acknowledgements. This research was supported by U.S. Geological Survey (USGS) cooperative agreement #G09AC000001 "Monitoring and Forecasting Climate, Water and Land Use for Food Production in the Developing World" with funding from the U.S. Agency for International Development Office of Food for Peace award #AID-FFP-P-10-00002 for "Famine Early Warning Systems Network Support" and NOAA Award NA11OAR4310151, "A Global Standardized

ESSDD
8, 401–425, 2015

Precipitation Index supporting the US Drought Portal and the Famine Early Warning System Network".

References

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

Funk, C. and Michaelsen, J.: A simplified diagnostic model of orographic rainfall for enhancing satellite-based rainfall estimates in data-poor regions, J. Appl. Meteorol., 43, 1366–1378, 2004.

Funk, C., Michaelsen, J., Verdin, J., Artan, G., Husak, G., Senay, G., Gadain, H., and Magadazire, T.: The collaborative historical African rainfall model: description and evaluation, Int. J. Climatol., 23, 47–66, 2003.

Funk, C., Husak, G., Michaelsen, J., Love, T., and Pedreros, D.: Third Generation Rainfall Climatologies: Satellite Rainfall and Topography Provide a Basis for Smart Interpolation, in: Proceedings of the JRC – FAO Workshop, 27–29 March, Nairobi, Kenya, 283–296, 2007.

Funk, C., Dettinger, M. D., Michaelsen, J. C., Verdin, J. P., Brown, M. E., Barlow, M., and Hoell, A.: Warming of the Indian Ocean threatens eastern and southern African food security but could be mitigated by agricultural development, P. Natl. Acad. Sci. USA, 105, 11081–11086, 2008.

Funk, C., Michaelsen, J., and Marshall, M.: Mapping recent decadal climate variations in precipitation and temperature across eastern Africa and the Sahel, in: Remote Sensing of Drought: Innovative Monitoring Approaches, edited by: Wardlow, B., Anderson, M., and Verdin, J., CRC Press, Boca Raton, Florida, USA, 25 p., 332–356, 2012.

Funk, C., Hoell, A., Shukla, S., Bladé, I., Liebmann, B., Roberts, J. B., Robertson, F. R., and Husak, G.: Predicting East African spring droughts using Pacific and Indian Ocean sea surface temperature indices, Hydrol. Earth Syst. Sci., 18, 4965–4978, doi:10.5194/hess-18-4965-2014, 2014a.

Funk, C., Peterson P., Landsfeld, M., Pedreros, D., Verdin, J., Rowland, J., Romero, B., Husak, G., Michaelsen, J., and Verdin A.: A Quasi-global Precipitation Time Series for Drought Monitoring, US Geological Survey, Reston VA USA, 2014b.

ESSDD
8, 401–425, 2015

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A.: Very high resolution interpolated climate surfaces for global land areas, Int. J. Climatol., 25, 1965–1978, 2005.
Hoell, A. and Funk, C.: The ENSO-related West Pacific sea surface temperature gradient, J. Climate, 26, 9545–9562, 2013a.
Hoell, A. and Funk, C.: Indo–Pacific sea surface temperature influences on failed consecutive rainy seasons over eastern Africa, Clim. Dynam., 1–16, doi:10.1007/s00382-013-1991-6, 2013b.
Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B.: The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, J. Hydrometeorol., 8, 38–55, 2007.
Husak, G. J., Funk, C. C., Michaelsen, J., Magadzire, T., and Goldsberry, K. P.: Developing seasonal rainfall scenarios for food security early warning, Theor. Appl. Climatol., 114, 291–302, 2013.
Hutchinson, M. F.: Interpolating mean rainfall using thin plate smoothing splines, Int. J. Geogr. Inf. Syst., 9, 385–403, doi:10.1080/02693799508902045, 1995.
Janowiak, J. E., Joyce, R. J., and Yarosh, Y.: A real-time global half-hourly pixel-resolution infrared dataset and its applications, B. Am. Meteorol. Soc., 82, 205–217, 2001.
Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P.: CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution, J. Hydrometeorol., 5, 487–503, 2004.
Knapp, K. R., Ansari, S., Bain, C., L., Bourassa, M. A., Dickinson, M. J., Funk, C., Helms, C. N., Hennon, C. C., Holmes, C., Huffman, G. J., Kossin, J. P., Lee, H.-T., A. Loew, A., and Magnusdottir, G.: Globally gridded satellite (GriSat) observations for climate studies, B. Am. Meteorol. Soc., 92, 893–907, 2011.
Krige, D. G.: A Statistical Approach to Some Mine Valuations and Allied Problems at the Witwatersrand, University of Witwatersrand, Johannesburg, South Africa, 1951.
Liebmann, B., Hoerling, M. P., Funk, C., Bladé, I., Dole, R. M., Allured, D., Quan, X., Pegion, P., and Eischeid, J. K.: Understanding recent eastern Horn of Africa rainfall variability and change, J. Climate, 27, 8630–8645, 2014.
Matheron, G.: Principles of geostatistics, Economic Geology, 58, 1246–1266, 1963.
Michaelsen, J.: Cross-validation in statistical climate forecast models, J. Clim. Appl. Meteorol., 26, 1589–1600, 1987.

ESSDD
8, 401–425, 2015

New, M., Hulme, M., and Jones, P.: Representing twentieth-century space–time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology, J. Climate, 12, 829–856, 1999.
Peterson, T. C. and Vose, R. S.: An overview of the Global Historical Climatology Network temperature database, B. Am. Meteorol. Soc., 78, 2837–2849, 1997.
Shepard, D.: A Two-Dimensional Interpolation Function for Irregularly-Spaced Data, in: Proc. 23rd National Conference Association for Computing Machinery, 517–524, 1968.
Shukla, S., Sheffield, J., Wood, E. F., and Lettenmaier, D. P.: On the sources of global land surface hydrologic predictability, Hydrol. Earth Syst. Sci., 17, 2781–2796, doi:10.5194/hess-17-2781-2013, 2013.
Shukla, S., Funk, C., and Hoell, A.: Using constructed analogs to improve the skill of March-April-May precipitation forecasts in equatorial East Africa, Environ. Res. Lett., 9, 094009, 2014a.
Shukla, S., McNally, A., Husak, G., and Funk, C.: A seasonal agricultural drought forecast system for food-insecure regions of East Africa, Hydrol. Earth Syst. Sci., 18, 3907–3921, doi:10.5194/hess-18-3907-2014, 2014b.
Verdin, J. and Klaver, R.: Grid-cell-based crop water accounting for the famine early warning system, Hydrol. Process., 16, 1617–1630, 2002.
Verdin, K. L. and Greenlee, S. K.: Development of Continental Scale Digital Elevation Models and Extraction of Hydrographic Features, in: Third International Conference/Workshop on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, 1996.
Wan, Z.: New refinements and validation of the MODIS Land-Surface Temperature/Emissivity products, Remote Sens. Environ., 112, 59–74, 2008.
Willmott, C. J. and Robseon, S. M.: Climatologically Aided Interpolation (CAI) of terrestrial air temperature, Int. J. Climatol., 15, 221–229, 1995.

ESSDD
8, 401–425, 2015

Table 1. CHPclim validation results.

	Region	N-stns	Station Mean	Climatology Mean	MBE	MAE	Pct MBE	Pct MAE	R^2
CHPclim	Colombia	194	168	159	8	30	5	18	0.84
	Afghanistan	22	35	34	1	9	3	25	0.53
	Ethiopia	76	97	94	3	10	4	10	0.91
	Sahel	28	55	53	0	6	0	10	0.93
	Mexico	1814	77	78	−1	23	−2	30	0.65
CRU	Colombia	194	168	174	−6	47	−4	28	0.59
	Afghanistan	22	35	45	−10	20	−28	57	0.18
	Ethiopia	76	97	101	−4	23	−4	24	0.68
	Sahel	91	55	65	−11	14	16	21	0.87
	Mexico	1814	77	75	2	24	2	31	0.60
Worldclim	Colombia	194	168	178	−11	31	−6	19	0.82
	Afghanistan	22	35	41	−6	18	−17	52	0.18
	Ethiopia	76	97	97	0	20	0	21	0.72
	Sahel	28	55	65	−10	14	−16	22	0.86
	Mexico	1814	77	79	−2	18	−2	23	0.78

ESSDD
8, 401–425, 2015

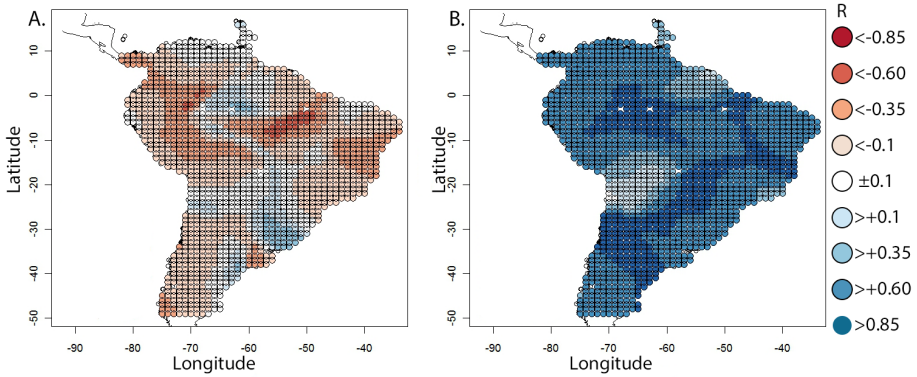

Figure 1. Local correlations with July station means. **(a)** Elevation. **(b)** Combined TRMM/CMORPH precipitation.

ESSDD
8, 401–425, 2015

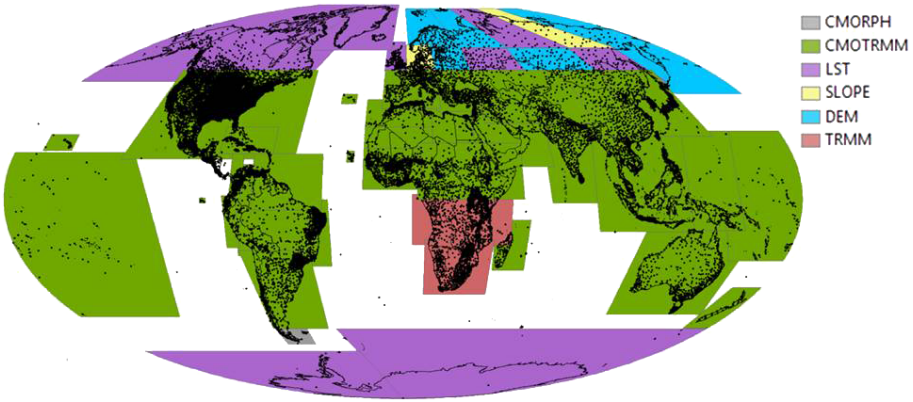

Figure 2. Best predictor, by model region, with station locations.

ESSDD
8, 401–425, 2015

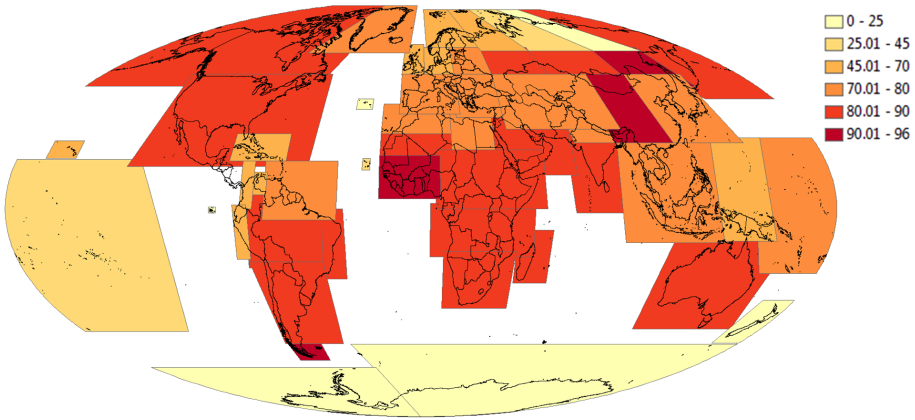

Figure 3. Percent of variance explained by cross-validated moving window regression.

ESSDD
8, 401–425, 2015

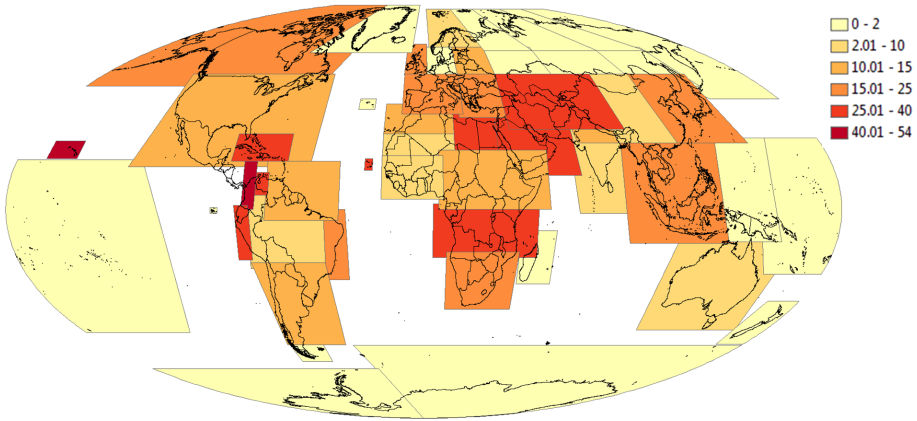

Figure 4. Percent of variance explained by cross-validated inverse distance weighting.

ESSDD
8, 401–425, 2015

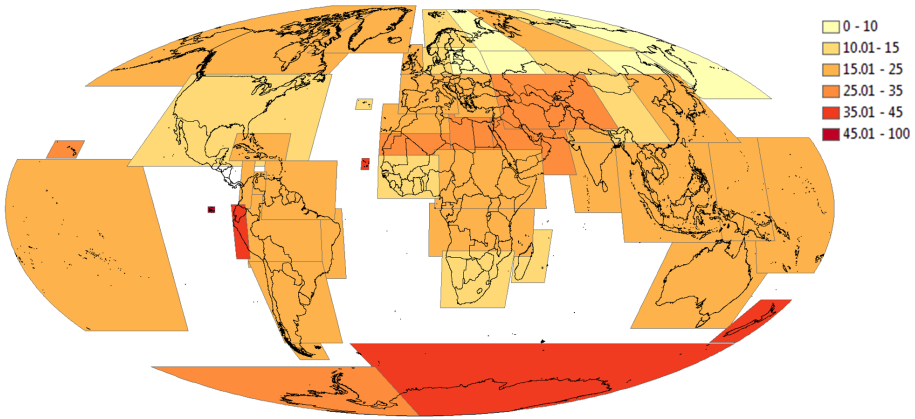

Figure 5. Percent standard error explained by cross-validation.

ESSDD
8, 401–425, 2015

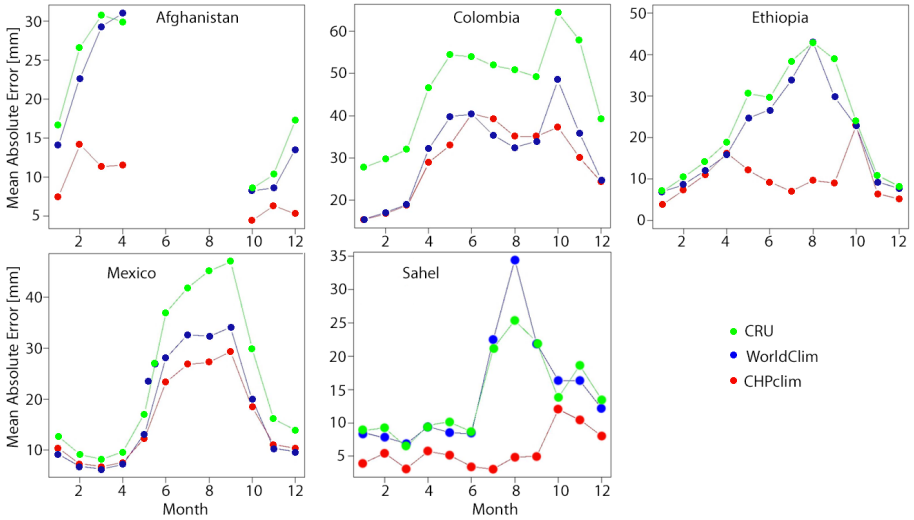

Figure 6. Mean absolute error time series [mm month^{-1}].

ESSDD
8, 401–425, 2015

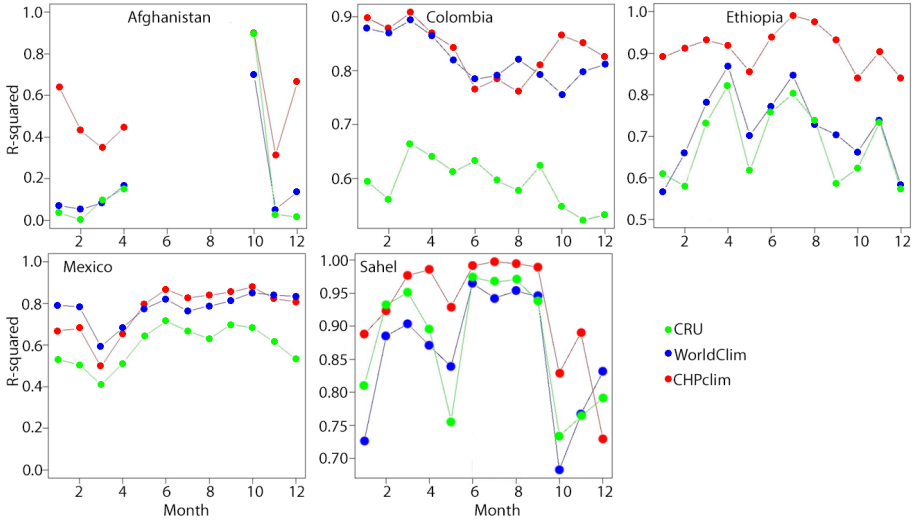

Figure 7. Spatial R^2 time series.

ESSDD
8, 401–425, 2015

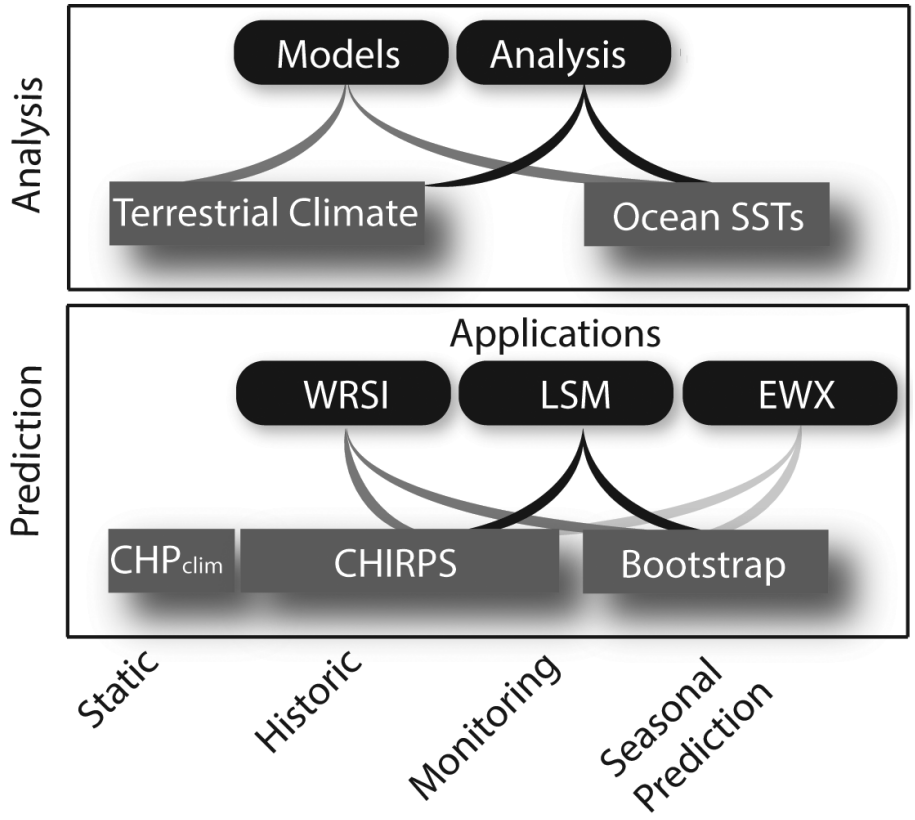

Figure 8. Schema of CHG analysis and prediction activities.

ESSDD
8, 401–425, 2015