A New Precipitation Dataset for Bias Correction and Downscaling

Purpose:

- Establish a precipitation dataset that is reliable and quality controlled, with a high spatial and temporal resolution.

Uses:

- Bayesian Processor of Ensemble at NCEP
- Downscaling
- Verification

Available Datasets:

Dataset	Available Since	Spatial Res.	Temporal Res.	Coverage
CPC Analysis	Jan, 2002	⅛°	1 day	CONUS
	Jan, 1948	¼°	1 day	CONUS/Mexico
RFC QPE	Oct, 2000	4-5 km	6 hrs	CONUS
CMORPH	Dec, 2002	~8 km	30 min	Semi-global (60 °N to 60 °S)
Future Products (available in the next 6 months)				
CPC Unified Gauge-based Analysis of Precip.	Jan, 1948	⅛°	1 day	CONUS/Mexico
RMORPH	1998	?	?	?

New Precipitation Dataset:

Summary:

The objective is to combine the established reliability of the ¼° CPC daily precipitation analysis with the higher spatial and temporal resolution of the River Forecast Center's Quantitative Precipitation Estimates (RFC QPE). The following steps will describe how the RFC QPE data is scaled so that its climatology matches that of the CPC data.

Objective 1 - RFC adjustment

1. *Prepare data*
   a. Sum the 7 years of 6-hourly RFC precip grids into 24-hourly grids (12-12 UTC)
   b. Interpolate these 24-hourly grids to ¼° using the budget option in copygb
2. *Obtain scaling terms α and β at each gridpoint*
   a. Select a 30-day window surrounding each day of the year (Fig. 1)
   b. Linear regression to obtain α and β on ¼° grid
      i. Collect all CPC and RFC grids within 30-day window, and insert timeseries of precip at each gridpoint into the following equation:

$$W_{CPC} = \alpha \cdot W_{RFC} + \beta$$

      ii. Do not include gridpoints with zero precip (becomes non-linear)
      iii. These scaling terms are calculated once
3. *Adjust RFC precipitation*
   a. Adjust all 7 years of RFC precip grids (daily, ¼°) by α and β
   b. Disaggregate in time to 6-hourly (algorithm from Roman)
   c. Disaggregate in space to ~5 km (algorithm from Roman)
   d. In real time, the RFC precip grids will be interpolated to ¼°, summed into 24-hour grids, adjusted by the scaling factors α and β and disaggregated in time and space.

Objective 2 - Pseudo-precipitation prior

1. *Generate negative pseudo-precipitation (Paul Schultz and Huiling Yuan)*
   a. Prepare data
      i. The negative pseudo-precipitation will be calculated using the 2.5° NCEP-NCAR reanalysis, back to 1959
      ii. Interpolate to 1° (for bias correction) and 5 km (for downscaling)
   b. Calculate pseudo-precipitation
      i. The negative pseudo-precipitation describes the moisture deficit in a column, and is defined at a gridpoint (i,j) as:

$$PP(i,j) = \frac{1}{g}\int (q_v - q_{vsat}(T))\,dp$$

   c. *Obtain transform coefficients at each gridpoint*
      i. Fit censored distribution to negative pseudo-precipitation, independently of positive, for each day of the year (Fig. 3, solid brown line)
      ii. Fit censored distribution to positive precipitation, independently of negative, for each day of the year (Fig. 3, solid green line)
      iii. Solve for transform between the distributions of positive and negative precipitation (Fig. 3, black arrows)
      iv. Either:
         - Save unique 'mapping function' for each day of the year, for each gridpoint
         - Find a transform function that is uniform over all days and/or gridpoints

- Most likely use unique transforms

v. Fourier smooth transform coefficients over 365 days

vi. These coefficients will be calculated only once

2. *Adjust negative pseudo-precipitation at each gridpoint*
   a. Transform negative pseudo-precipitation with the saved transform coefficients for that day, for all 7 years of RFC data (Fig. 3, dashed red line)
   b. End up with continuous pseudo-precipitation at 5 km (1°) for downscaling (bias-correction)
   c. Now treat adjusted pseudo-precipitation prior as any other continuous variable, and insert into normal Bayesian Processor
3. *Adjust operational pseudo-precipitation*
   - Since the negative portion of the pseudo-precipitation prior originates from 2.5° reanalysis variables, its distribution is different than the distribution one would find if the variables came from a higher resolution grid (such as the operational analysis).
   - The operationally forecasted pseudo-precipitation will be on a 1° grid, and will therefore have an inherently different distribution.
   - Using the mapping function found in 1c would not be appropriate here, because it was designed to map 2.5° reanalysis-based pseudo-precipitation to the positive precipitation distribution.
   - Before using the forecast (1°) pseudo-precipitation in the Bayesian Processor, we must:

   a. Calculate the difference between the reanalysis-based pseudo-precipitation and the operational analysis-based pseudo-precipitation
      i. The difference between the reanalysis and operational reanalysis is already calculated recursively for several variables. For pseudo-precipitation, either:
         1. Recursively estimate the difference between analyses for all component variables of pseudo-precipitation
         2. Recursively estimate the difference between analyses for pseudo-precipitation itself
         3. Recursively estimate the linear regression coefficients relating the 2.5° reanlysis-based pseudo-precipitation to the 1° operational analysis-based pseudo-precipitation
      ii. Scale the forecast pseudo-precipitation by this difference to make the forecast pseudo-precipitation look like the 2.5° reanalysis-based pseudo-precipitation
      iii. Now apply the mapping function found in 1c

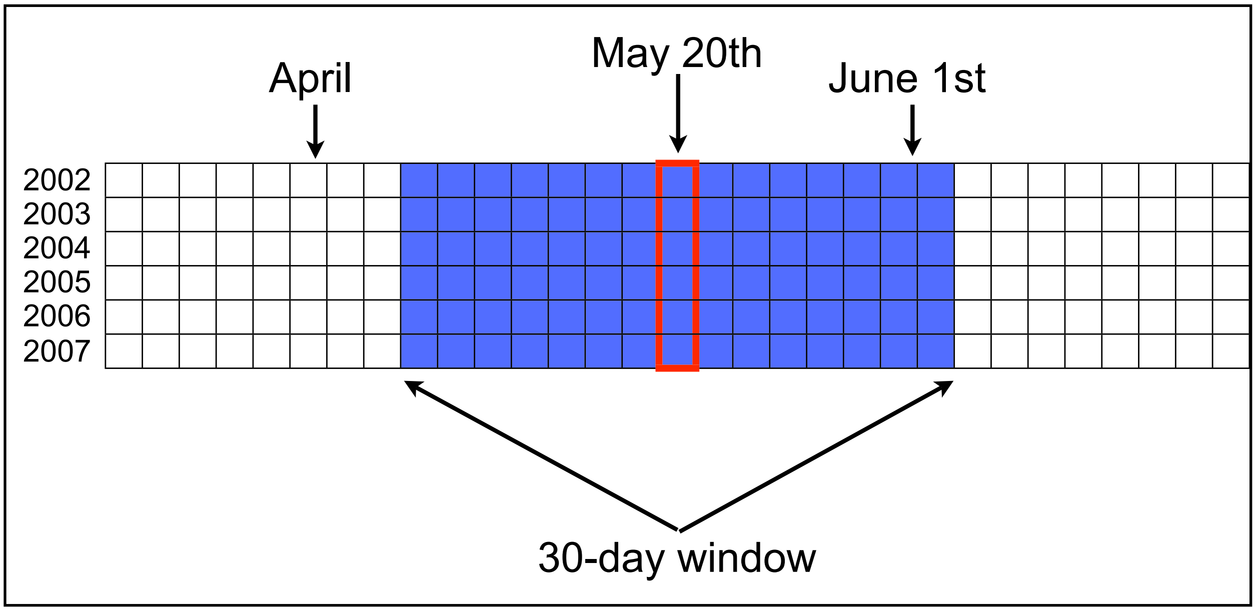

Figure 1. Example of average precipitation over a 30-day window for each day of the year. Each row represents a year of data, and each column represents two days of the year. In this example, the daily average precipitation is being calculated for May 20th. The blue boxes are all within the 30-day window surrounding the day in question, and represent daily precipitation grids that will be box-averaged, resulting in an average daily precipitation grid for May 20th.

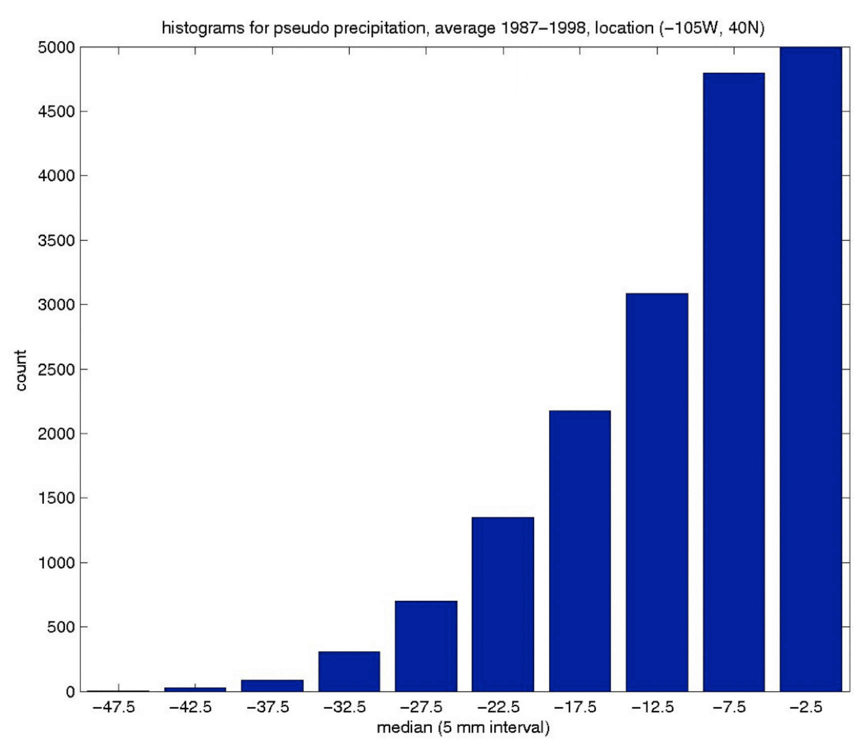

Figure 2. Histogram of pseudo precipitation (mm), from 1987 to 1998, for Denver, CO.

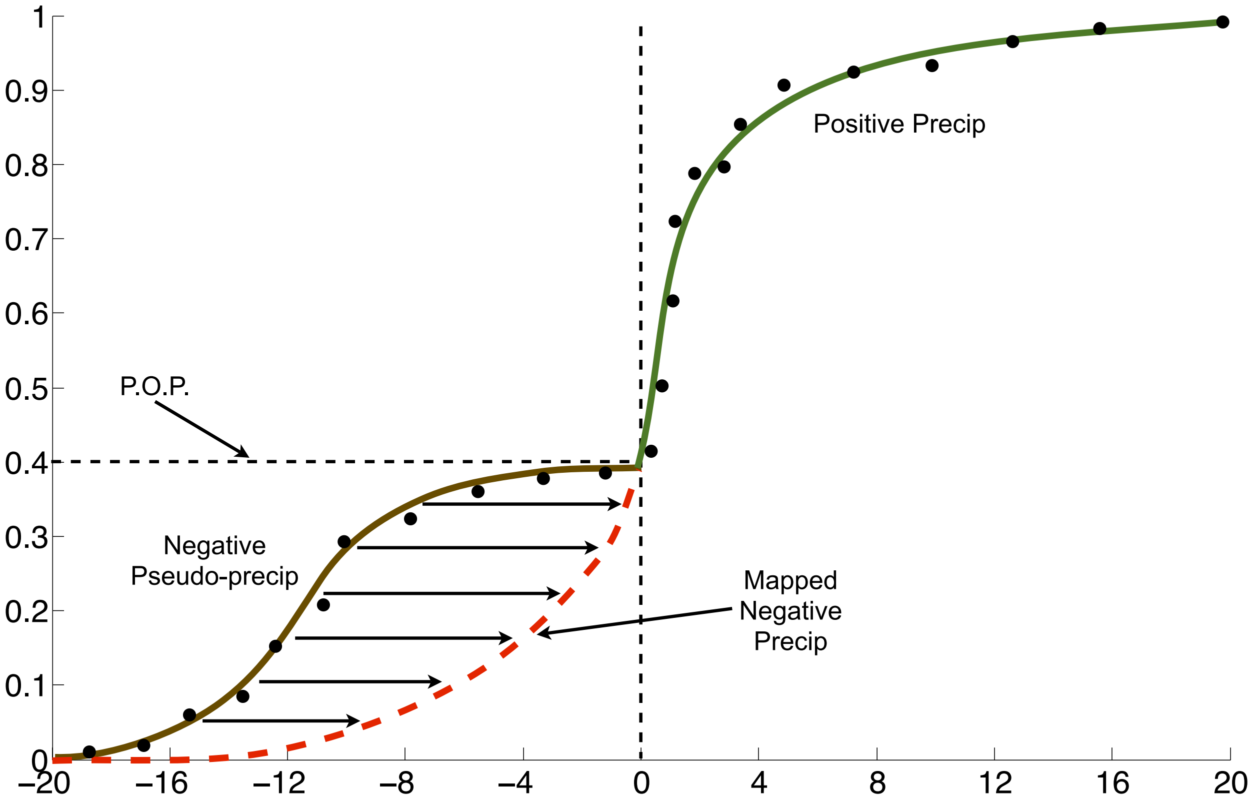

Figure 3. Example of mapping of pseudo-precipitation CDF to make it continuous with the positive precipitation CDF at a single gridpoint. In this example, the probability of precipitation (P.O.P.) is 0.4. The brown line indicates the curve fit to the CDF of negative pseudo-precipitation (raw values of CDF indicated by black dots). The green line indicates the CDF of positive precipitations. The dashed red line is the negative pseudo-precipitation mapped (black arrows) so that its CDF is consistent with the positive precipitation.