

A High Resolution Precipitation Dataset over CONUS: Climatology-Calibrated Precipitation Analysis (CCPA)

Dingchen Hou¹, Mike Charles^{1,}, Zoltan Toth^{1,#}, Yuejian Zhu¹,
Roman Krzysztofowicz², Yan Luo¹, Ying Lin¹, Pingping Xie³,
Dong-Jun Seo⁴, Malaquias Pena¹ and Bo Cui¹*

¹ Environmental Modeling Center/NCEP/NWS/NOAA, Camp Springs, MD

² University of Virginia, Charlottesville, VA

³ Climate Prediction Center/NCEP/NWS/NOAA, Camp Springs, MD

⁴ Office of Hydrology Development/NWS/NOAA, Silver Springs, MD

* Current affiliation: Climate Prediction Center/NCEP/NWS/NOAA, Camp Springs, MD

Current affiliation: Global System Division/ESRL/OAR/NOAA, Boulder, CO

Acknowledgements

Steve Lord, EMC/NCEP/NWS/NOAA

Letitia Soulliard, HPC/NCEP/NWS/NOAA

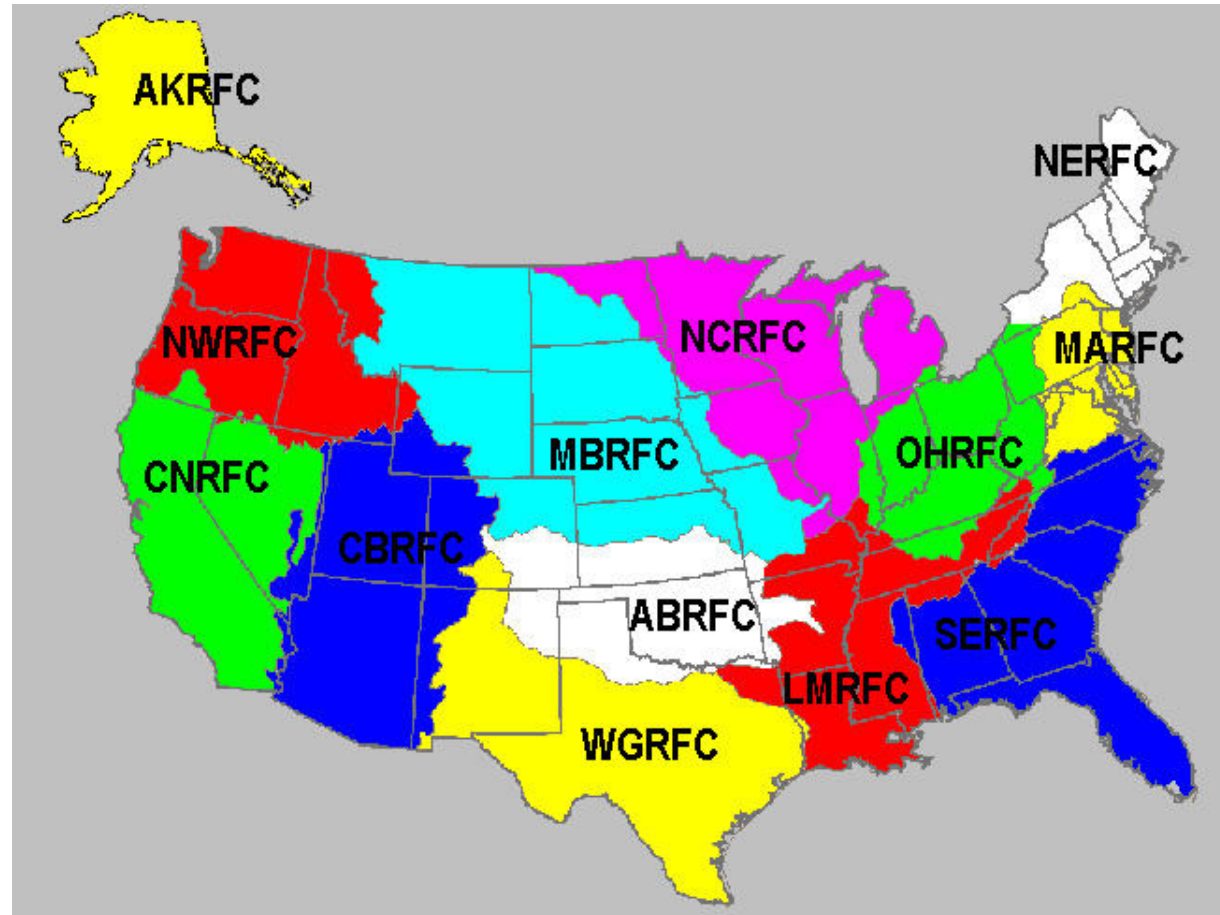
Sep. 17 2010 OHD

Why do we need another dataset?

- Want an accurate, 5x5 km (NDFD), 6-hourly precip grid
 1. Downscale NAEFS precipitation forecasts to NDFD
 2. Verify NAEFS precipitation forecasts
 3. Input to Bayesian Processor of Ensemble (BPE)
- What do we Have?
 - Stage IV
 - CPC
- **Note: This effort has limitations, as it was developed to simply combine existing datasets. Much more work will be needed for a more comprehensive approach, but this is out of the scope of this work**

What is Stage IV

- National Stage IV QPE product
 - Mosaicked from Individual RFC's Multi-sensor Precipitation Analyses ([RMPAs](#))
 - Available within 1h of receiving any new hourly/6-hourly data from one or more RFCs.
 - 12 RFCs over CONUS
 - Some manual QC at RFCs
 - 4km HRAP grid



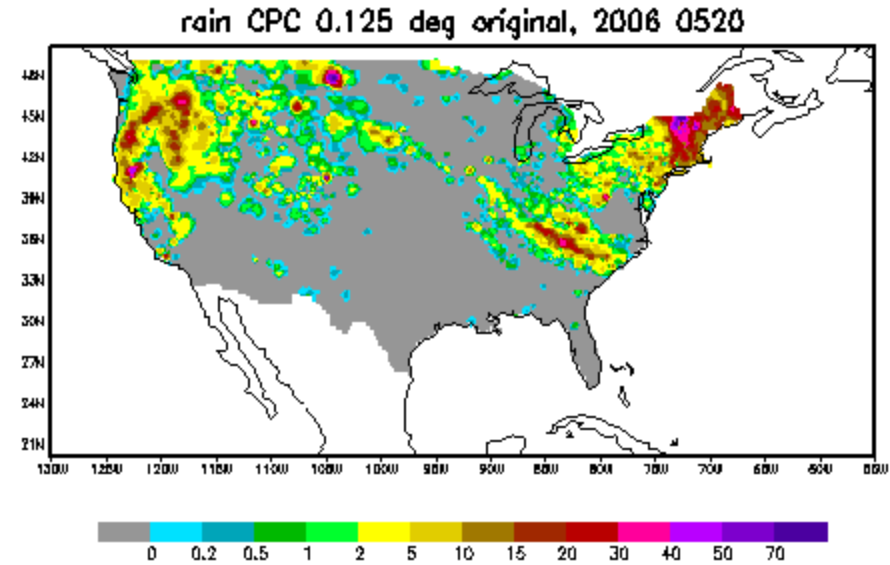
Why Calibration ?

(following Charles et. al)

- Want an accurate, 5x5km (NDFD) 6-hourly precipitation for
 - Down scaling NAEFS precip forecast to NDFD
 - Verifying NAEFS precipitation forecast
- Stage IV, a good candidate but ...
 - ✓ High resolution (close to NDFD) → better representation of fine scale temporal and spatial variability
 - Non-uniform QC (different RFCs have different methods)
 - Each RFC may make their own adjustments before mosaicking
- CPC Unified Precipitation Analysis
 - Back to 2000 (eventually back to 1979, then 1948)
 - $\frac{1}{8}^{\circ}$ spatial resolution
 - Daily
 - Global land
 - ✓ More confidence in long term statistics of CPC dataset
 - a. Uniform QC across entire domain
 - b. Gauge-based
 - Too low resolution for downscaling

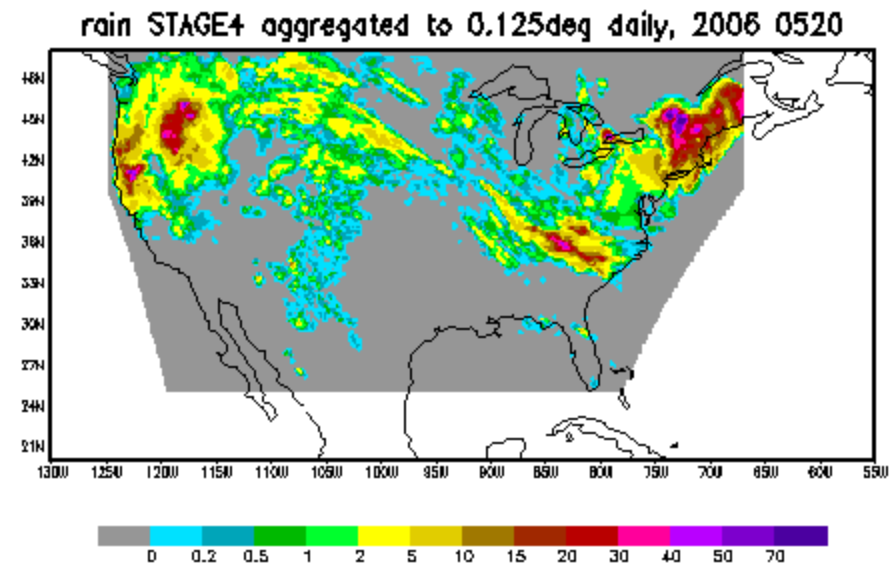
Comparison of CPC and Stage IV

**CPC:
CONUS
LAND
ONLY**



Units: mm/day

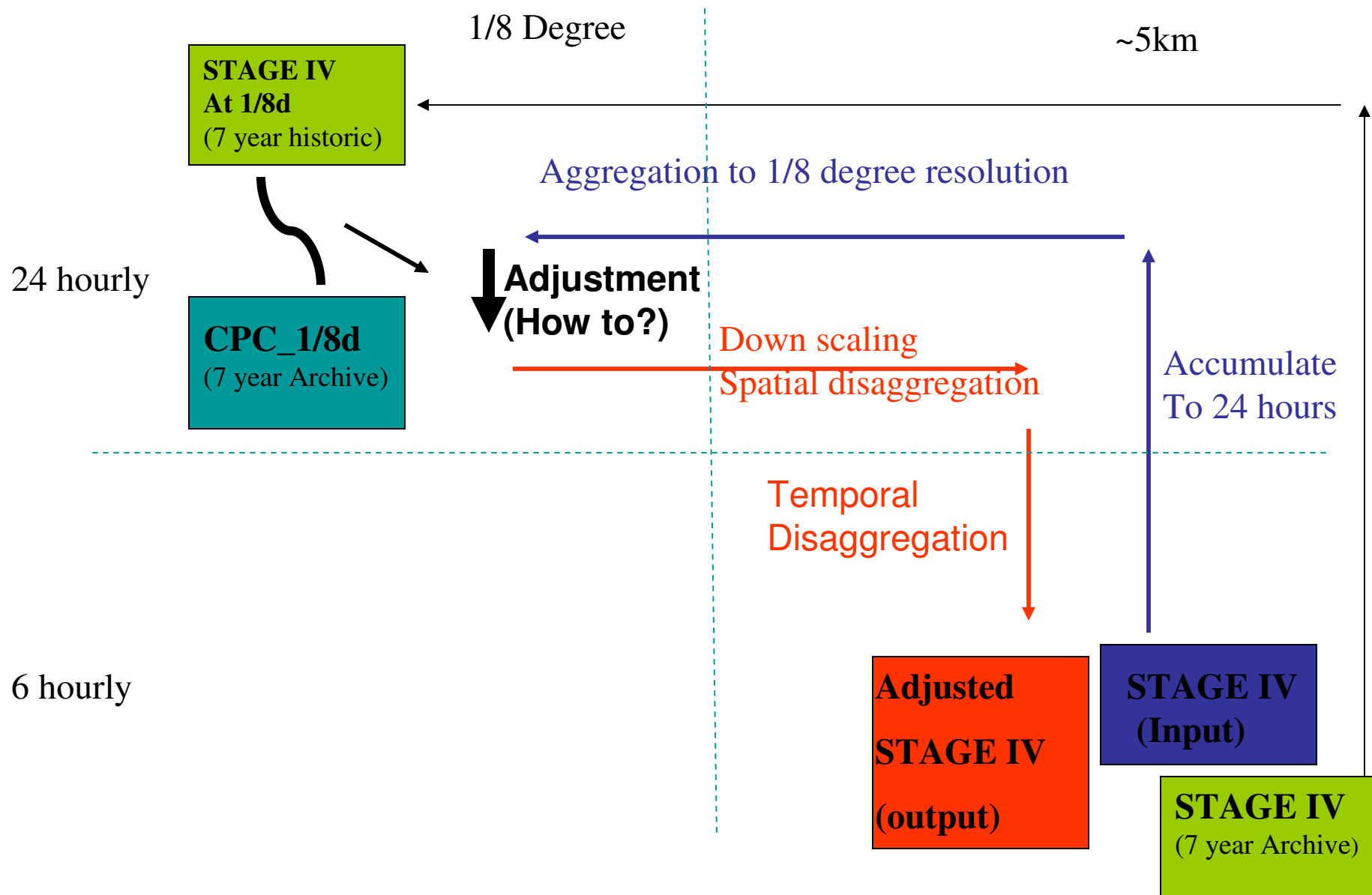
**Stage IV:
Geometric
Boundary**



How to Calibrate

- *Solution:* adjust Stage IV grids so their climatology is consistent with the CPC dataset
 - ✓ Have the reliability of the CPC dataset, with the high spatial and temporal resolution of the RFC dataset
- **Note: This effort has limitations, as it was developed to simply combine existing datasets. Much more work will be needed for a more comprehensive approach, but this is out of the scope of this work**

Flow Chart



Establish Statistical Relationship

1. Historical data sets

June 1 2002 to July 31 2009 For CPC and STAGE IV

2. Match resolutions

- a. Accumulate RFC over 24 hours
- b. Interpolate to $\frac{1}{8}^\circ$ (copygb w/ volume preservation)

3. Collect precip samples

- a. For each day of the year and at each grid point, collect all precip within 60 day window centered around that day, over all 7 years (max ~427 data points)
- b. Use only data points with ST4 > 0

4. Linear regression

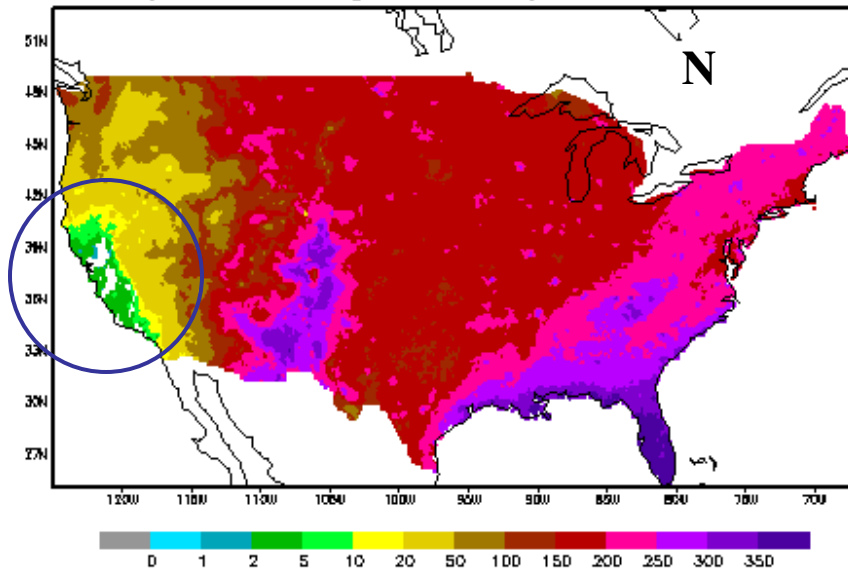
- a. $\text{CPC} = a \cdot \text{ST4} + b$

• End Result

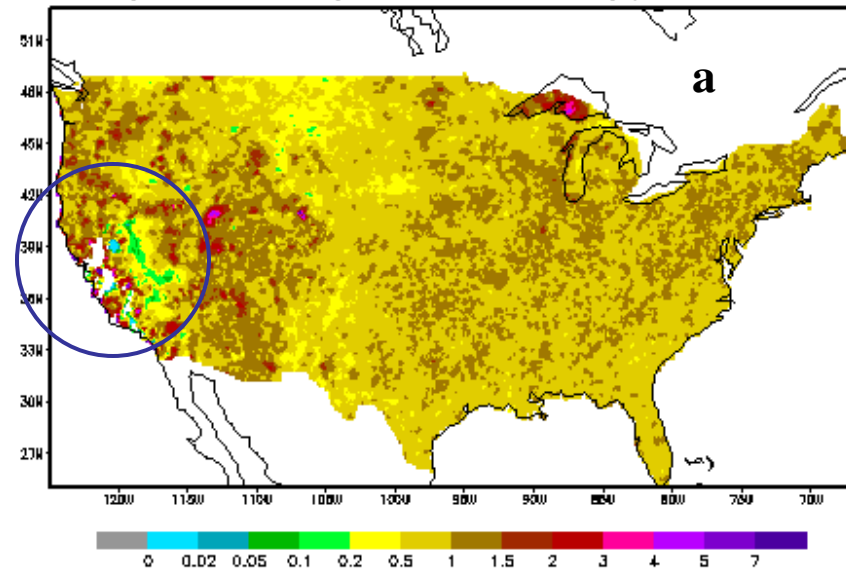
- Linear relationship (a & b) on $\frac{1}{8}^\circ$ grid for each day of the year

Regression Aug. 1st (SW US, Summer Gaps, maximum 369 points)

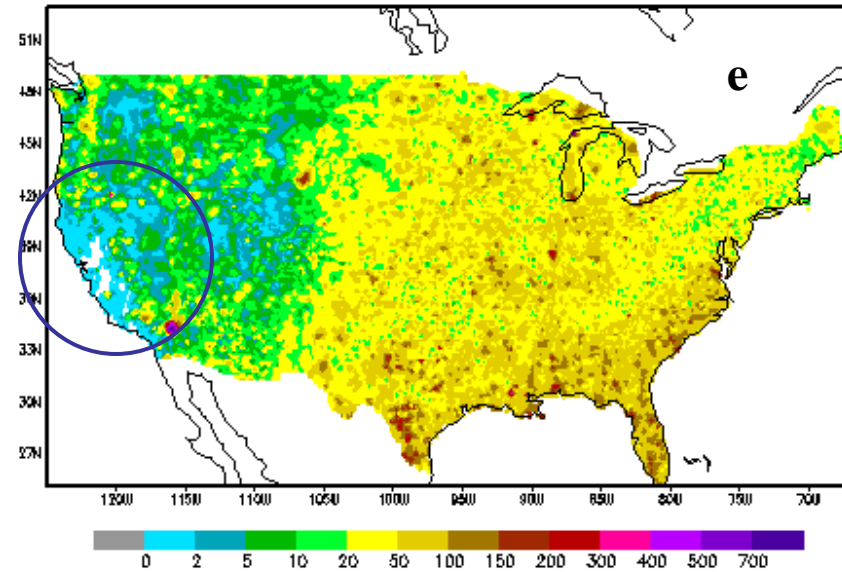
Prcp CPC-ST4 Regression Sample size, 2000 0801



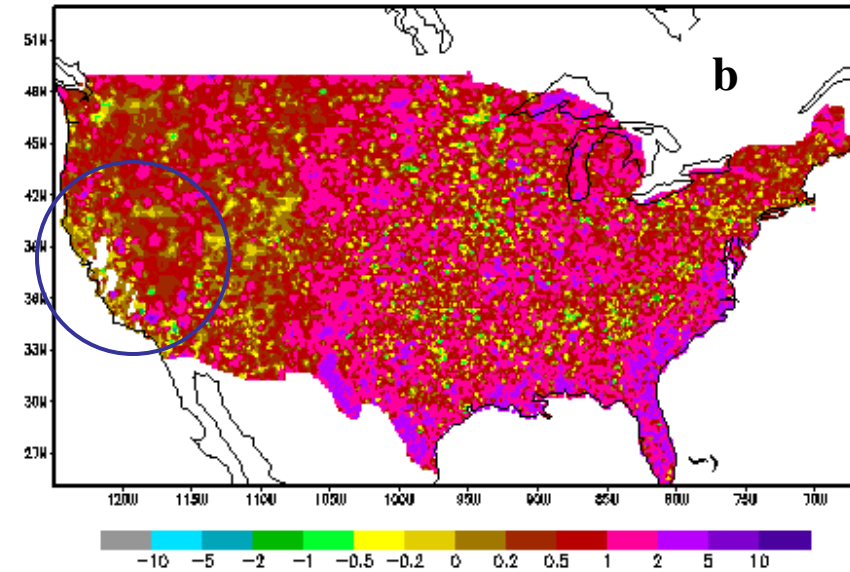
Prcp CPC-ST4 Regression Coefficient (a), 2000 0801



Prcp CPC-ST4 Regression Res. Square (e), 2000 0801

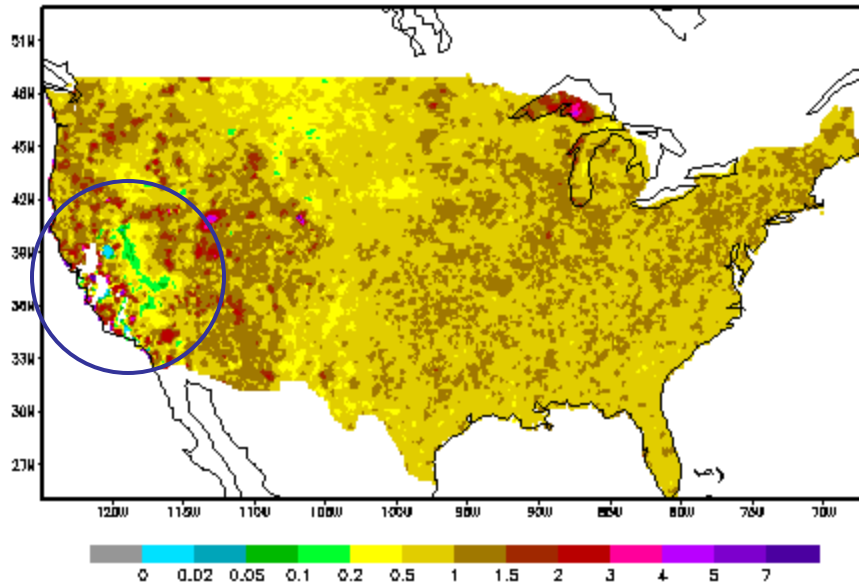


Prcp CPC-ST4 Regression Intercept (b), 2000 0801

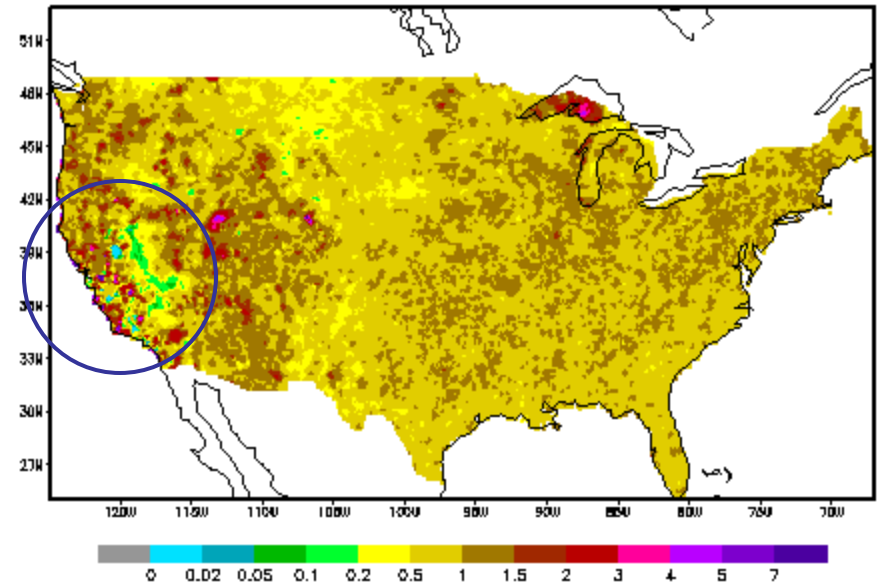


Filling the gap in Space (linear interpolation)

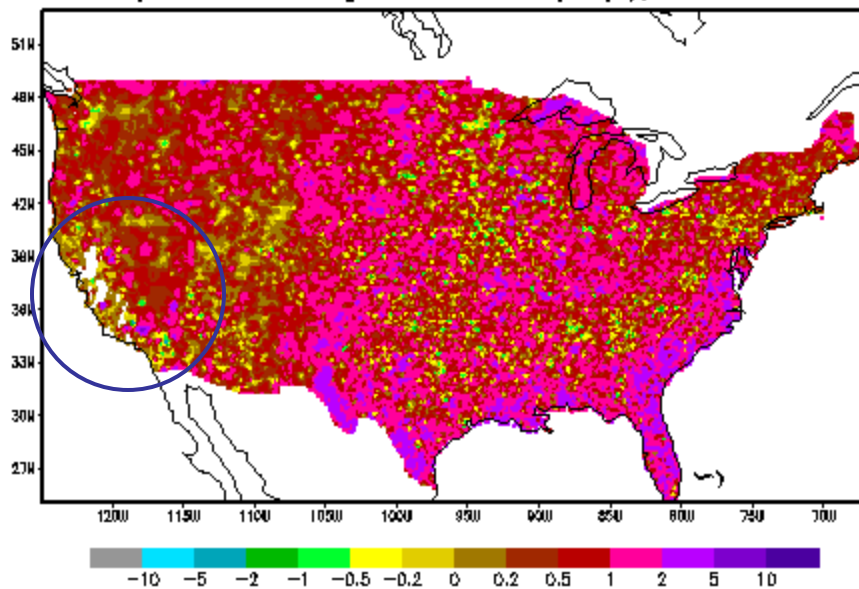
Prcp CPC-ST4 Regression Coefficient (a), 2000 0801



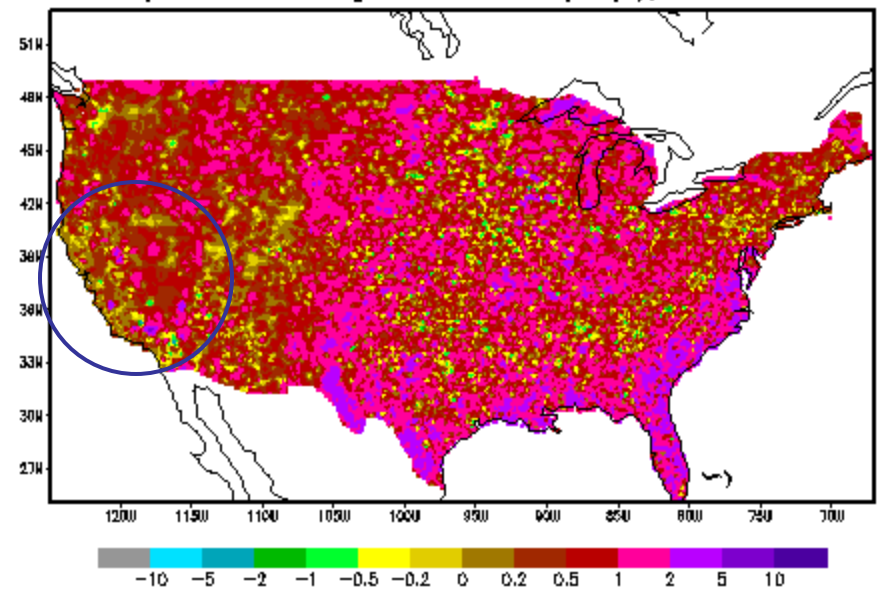
Prcp CPC-ST4 Regression Coefficient (a), 2000 0801



Prcp CPC-ST4 Regression Intercept (b), 2000 0801

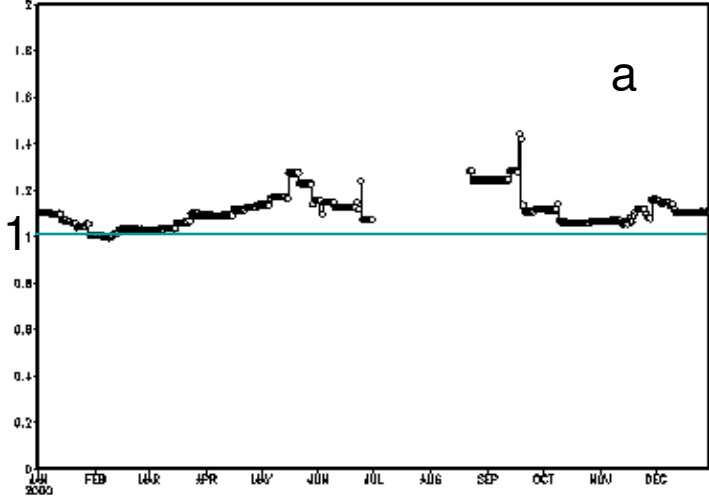


Prcp CPC-ST4 Regression Intercept (b), 2000 0801



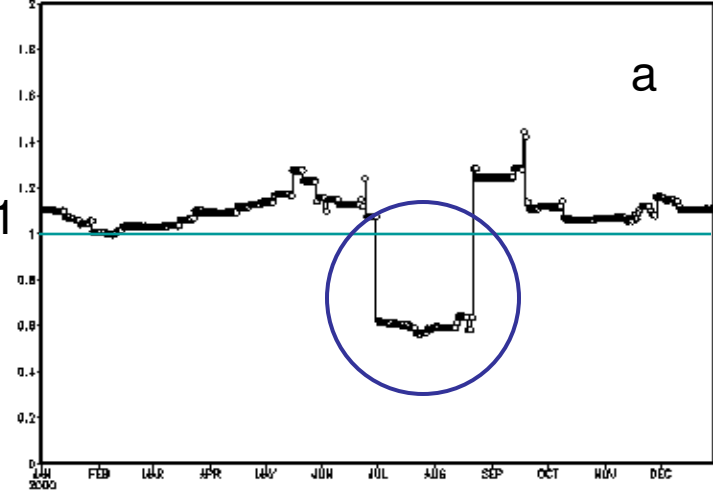
Time Series; Gap is filled by Spatial Interpolation

Prpc CPC-ST4 Regression Coefficient (a), (Point 54,72)



1/15/00

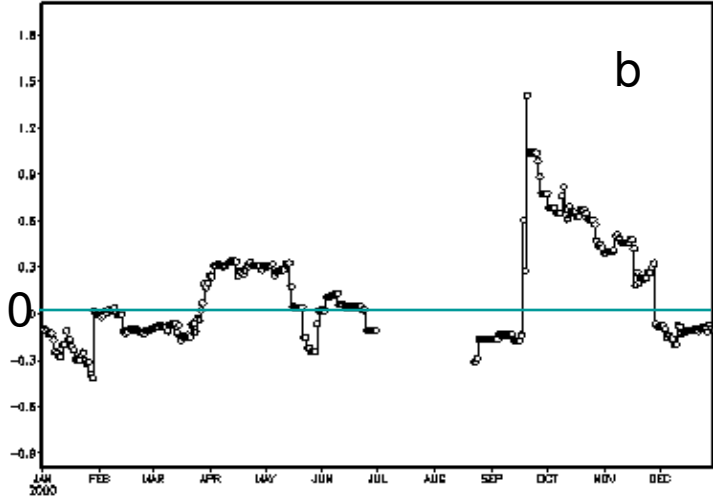
Prpc CPC-ST4 Regression Coefficient (a), (Point 54,72)



1/15/00

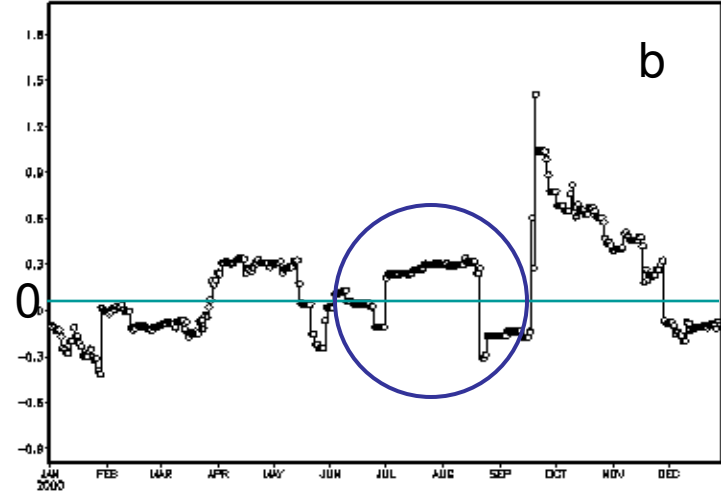


Prpc CPC-ST4 Regression Intercept (b), (Point 54,72)



1/15/00

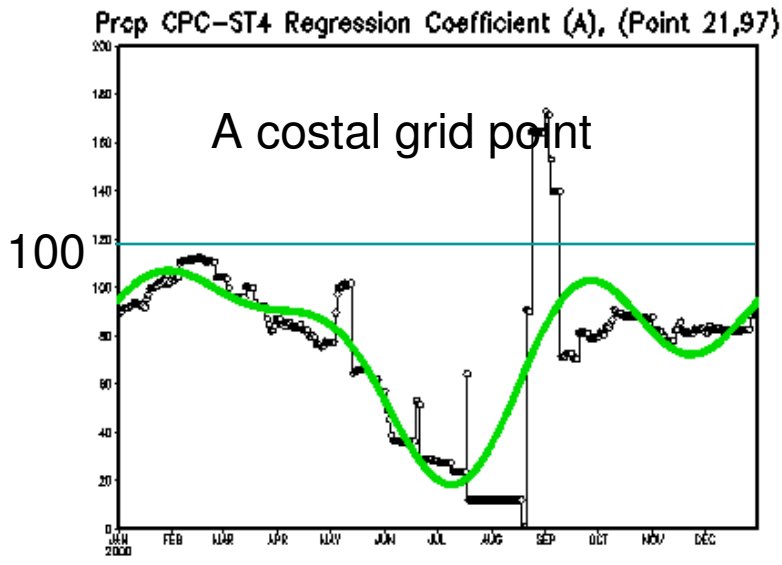
Prpc CPC-ST4 Regression Intercept (b), (Point 54,72)



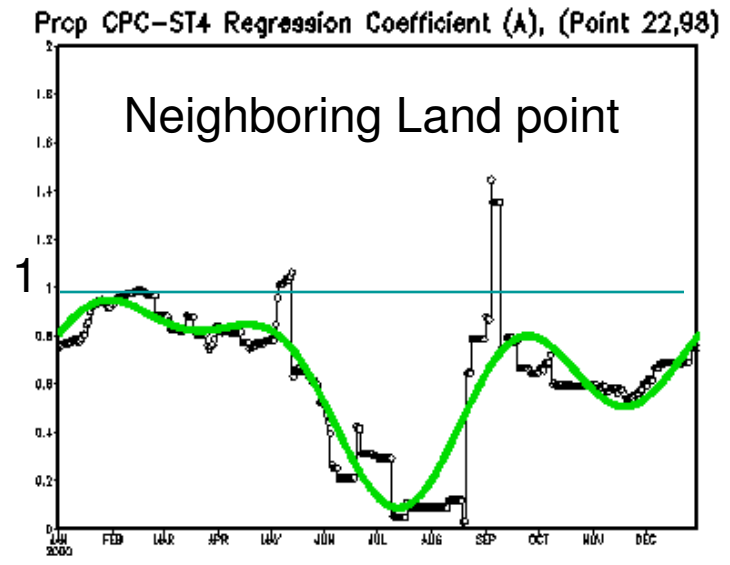
1/15/00



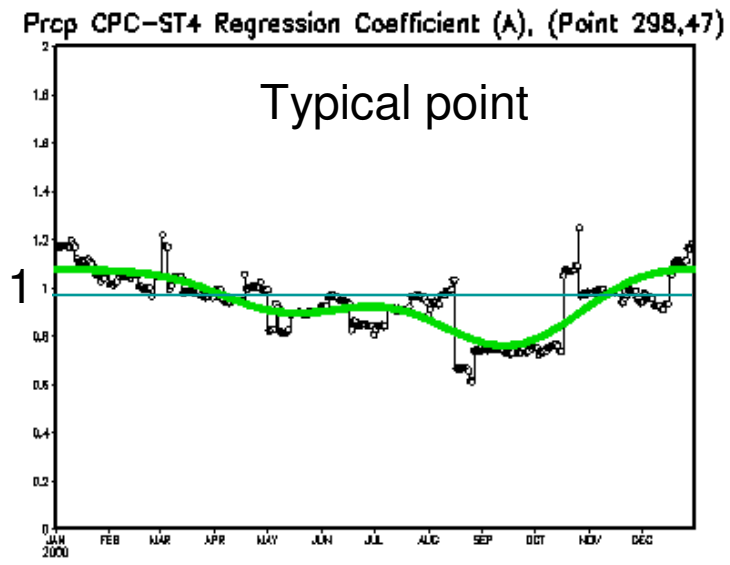
Temporal Smoothing (3 harmonics) of a



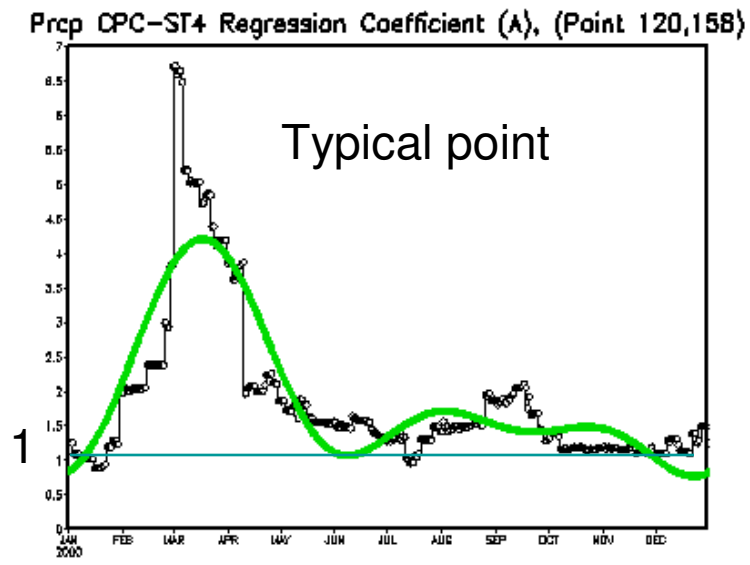
3



0489: COLA/NCES



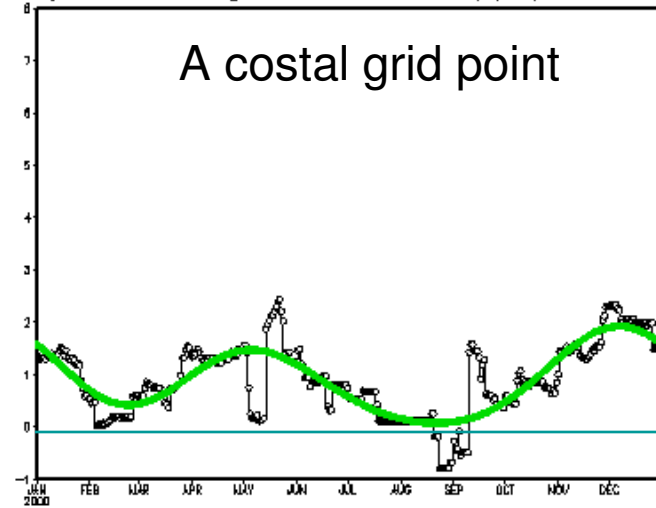
3



0489: COLA/NCES

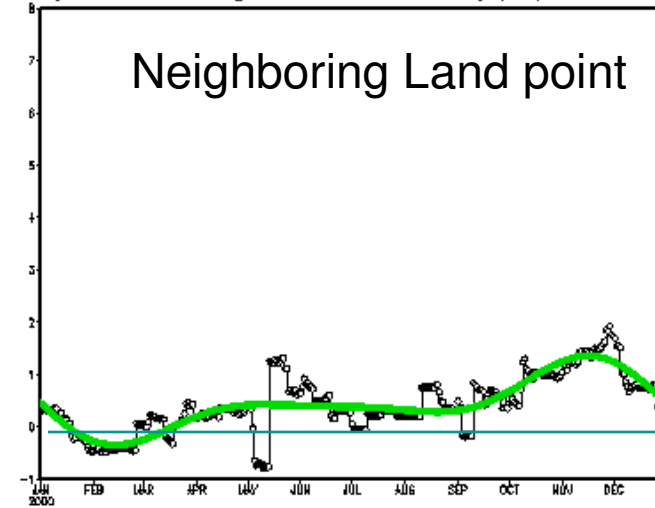
Temporal Smoothing (3 harmonics) of b

Prpc CPC-ST4 Regression Coefficient (B), (Point 21,97)



A coastal grid point

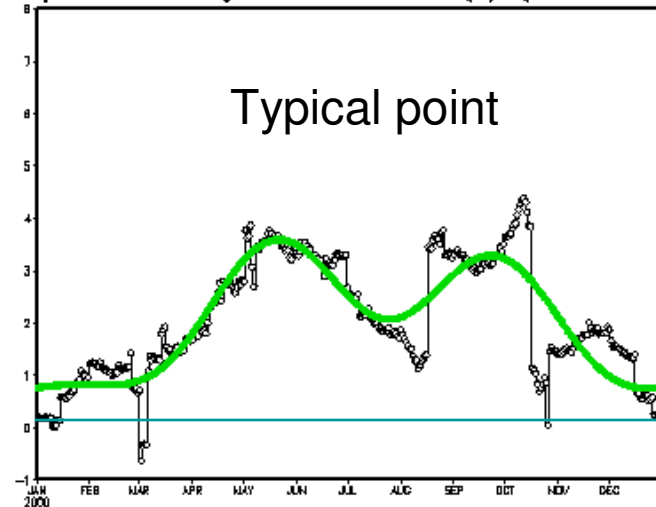
Prpc CPC-ST4 Regression Coefficient (B), (Point 22,98)



Neighboring Land point

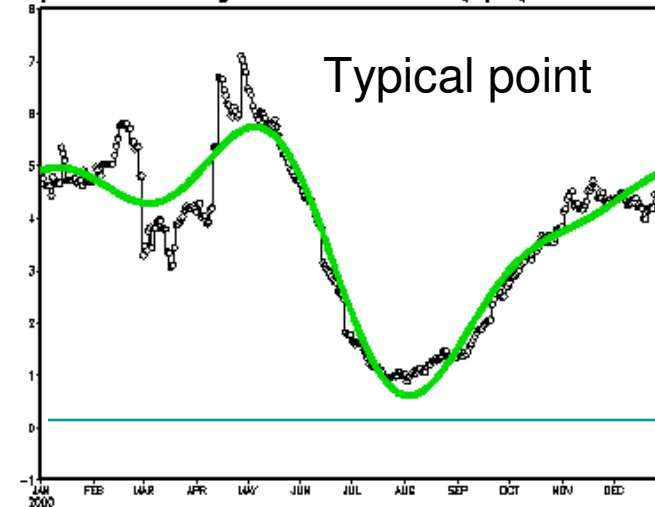
GRID: CELA/RES

Prpc CPC-ST4 Regression Coefficient (B), (Point 298,47)



Typical point

Prpc CPC-ST4 Regression Coefficient (B), (Point 120,158)

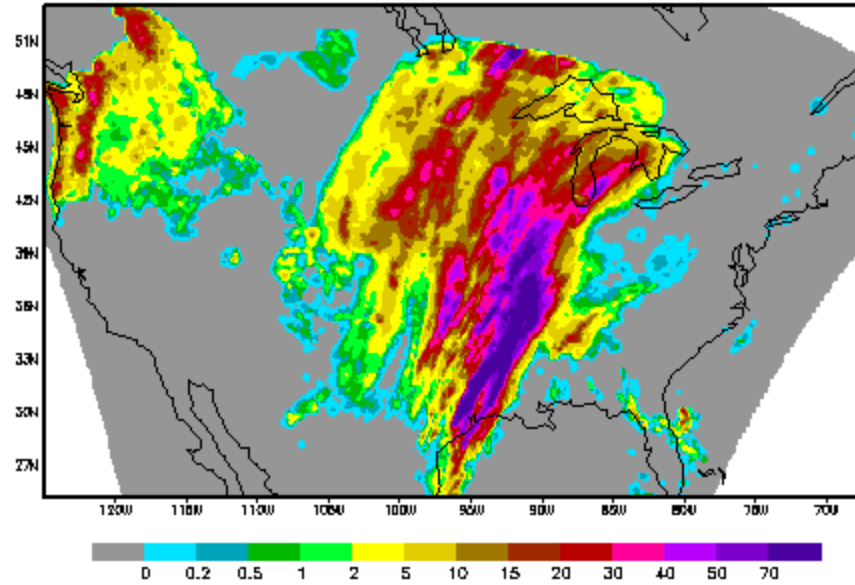


Typical point

GRID: CELA/RES

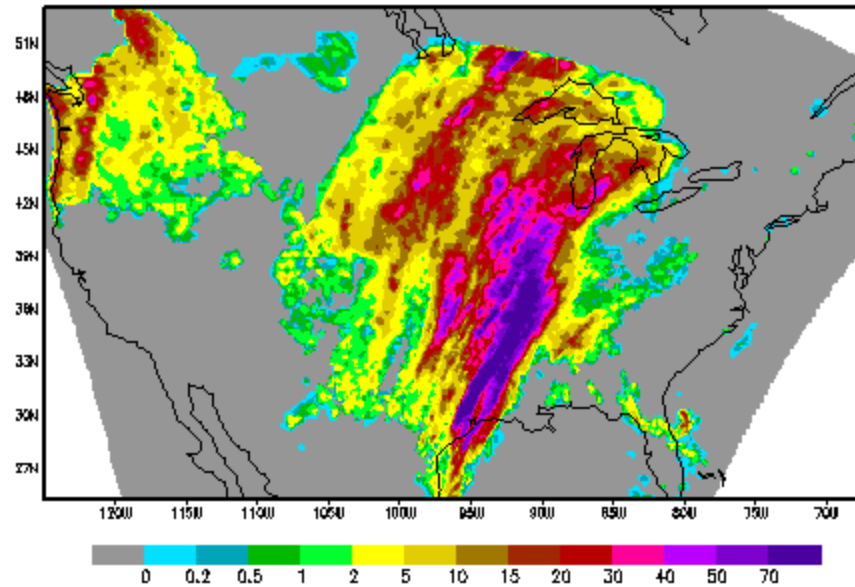
Adjustment with raw and filled a&b: $ST4^* = a \cdot ST4 + b$

Prcp STAGE4 0.125 deg original, 2009 1030



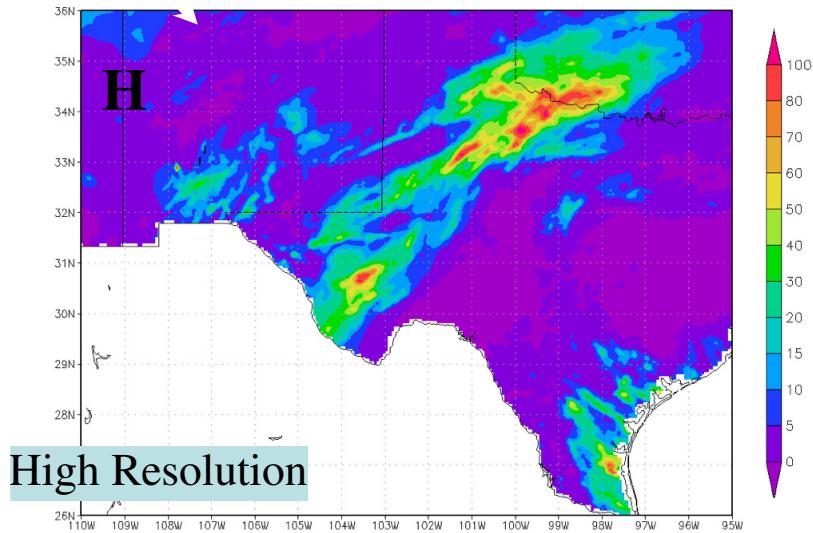
BASE: DOLA/DES

Prcp, STAGE4 0.125 deg adjust_02, 2009 1030

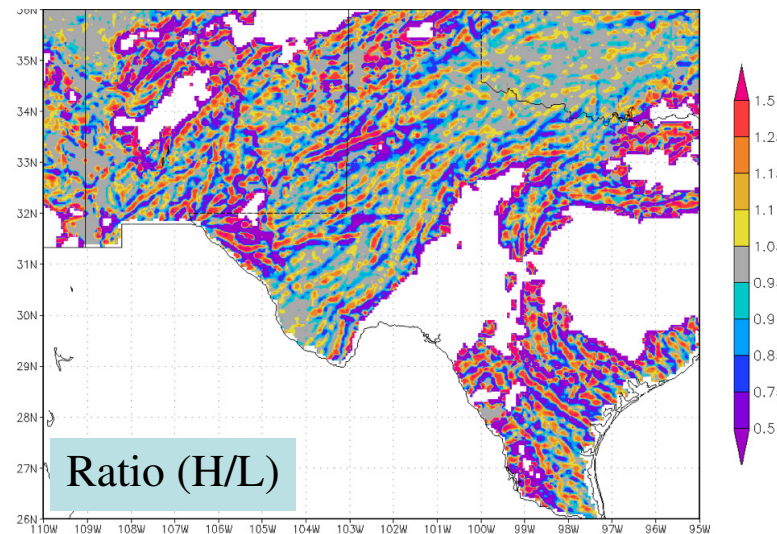
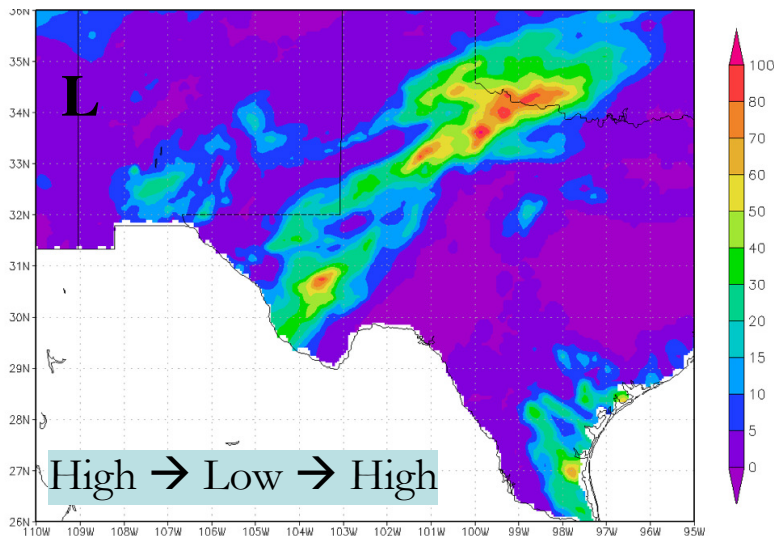


Recovering Original RFC Resolutions

Spatial Disaggregation



- Information is lost in ST4* (from H to L res.)
 - What does lost information look like?
 - ST4,Take ratio H/L (below) from the original ST4
 - This ratio can be used to put high resolution information back into ST4*
1. Interpolate ST4* to HRAP
 2. Multiply by H/L
- End with ST4* at HRAP resolution.
 - Spatial information recovered from ST4_{orig}



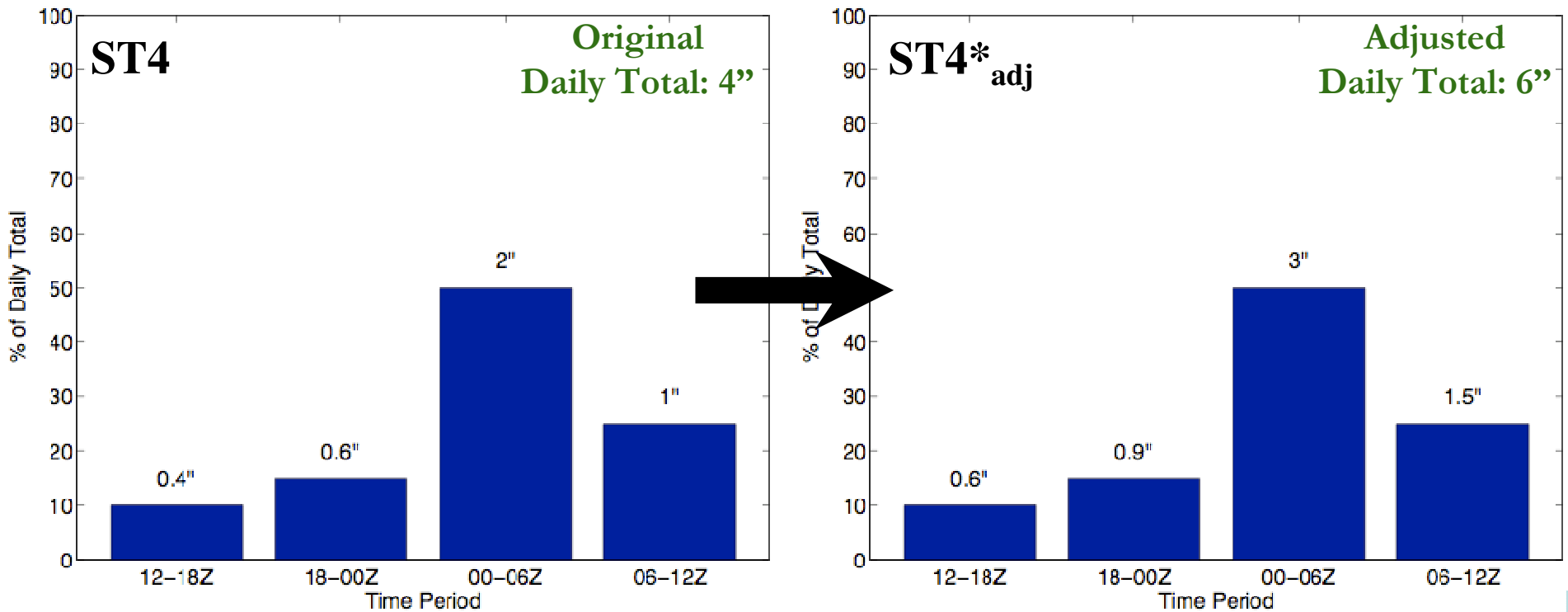
Recovering Original RFC Resolutions

Temporal Disaggregation

1. Determine percentage of daily total precipitation in each 6-hour period in original ST4

2. Divide 24 hour ST4* into four 6-hour precip amounts using the percentages from original ST4

Percent of daily total in each 6-hourly period

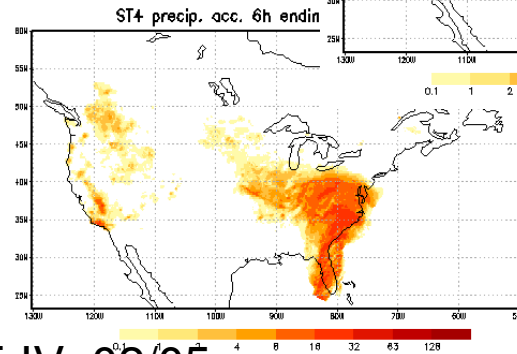
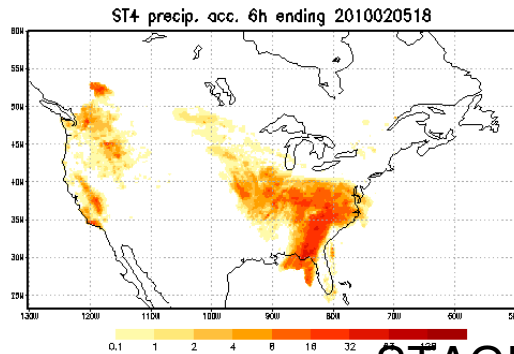
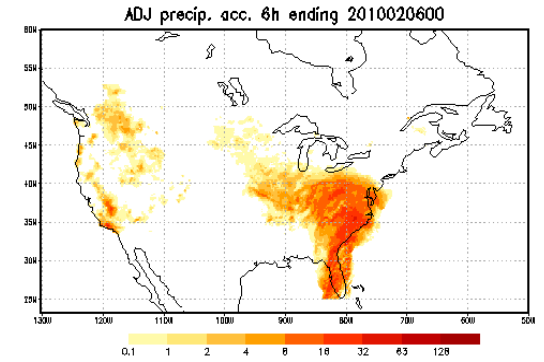
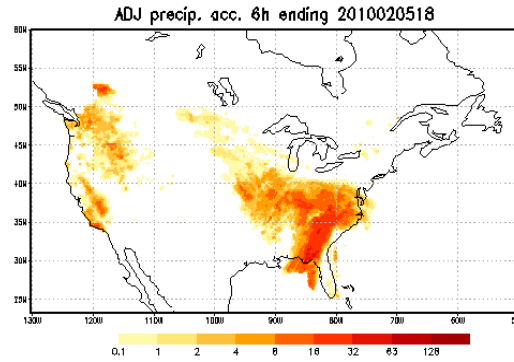


Implementation Details

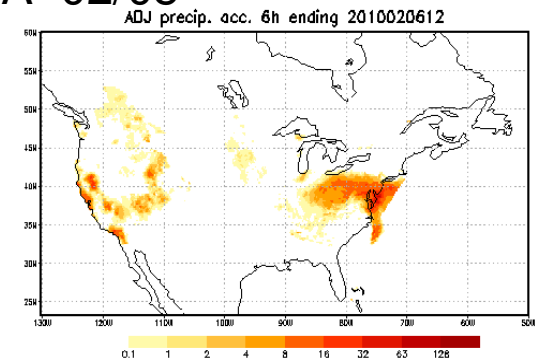
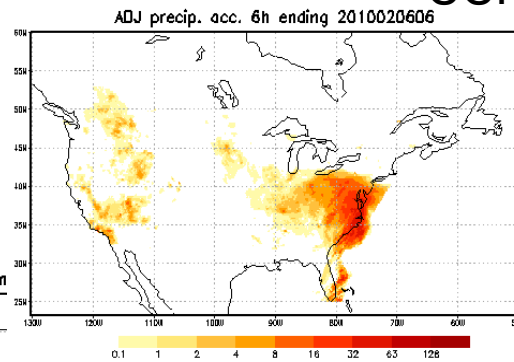
- Rules
 - Only Non-Zero Stage IV is adjusted
 - Zero values remains zero
 - Adjustment is applied over CONUS LAND only
- Leap Year
 - 366 day convention is adapted in regression calculations
 - Feb 29 has its own regression coefficients a and b
- Spatial Continuity
 - US Boundaries
 - Land/Ocean Boundary
 - Zero/Non-Zero Boundary
- Rare cases of abnormal regression coefficients
 - Temporal smoothing of a and b reduces abnormal values
 - Discard the regression coefficients a and b, if too large
 - Set an upper limit to the adjusted St4 value

An Example Feb. 5 2010

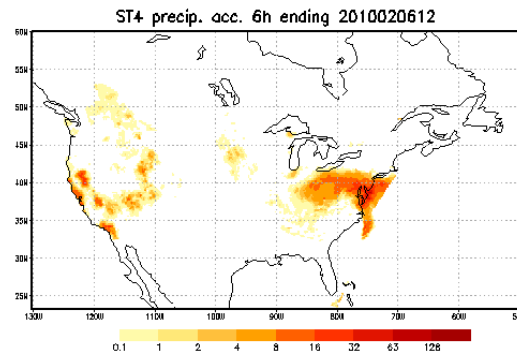
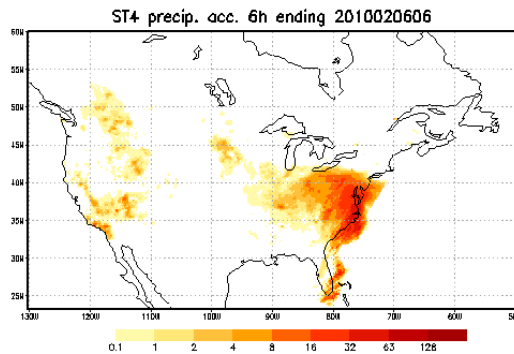
East Coast Snowstorm 2010



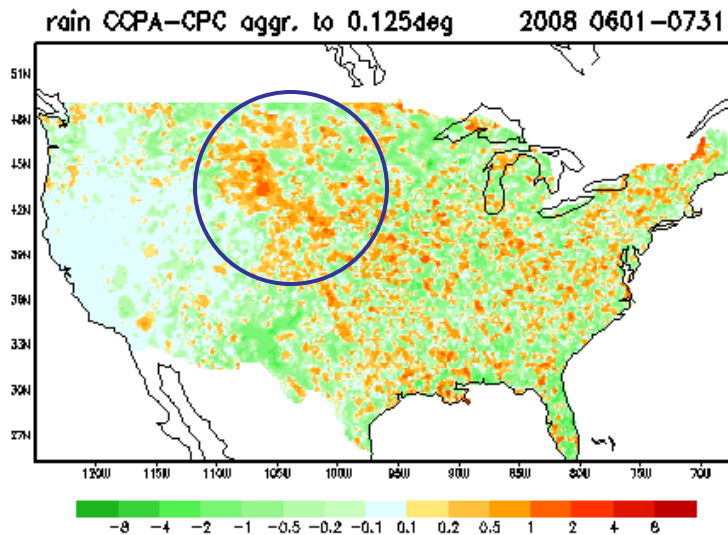
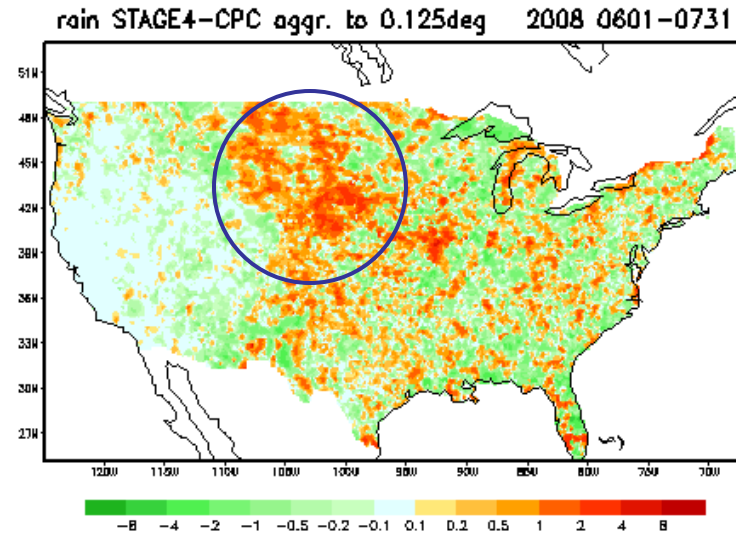
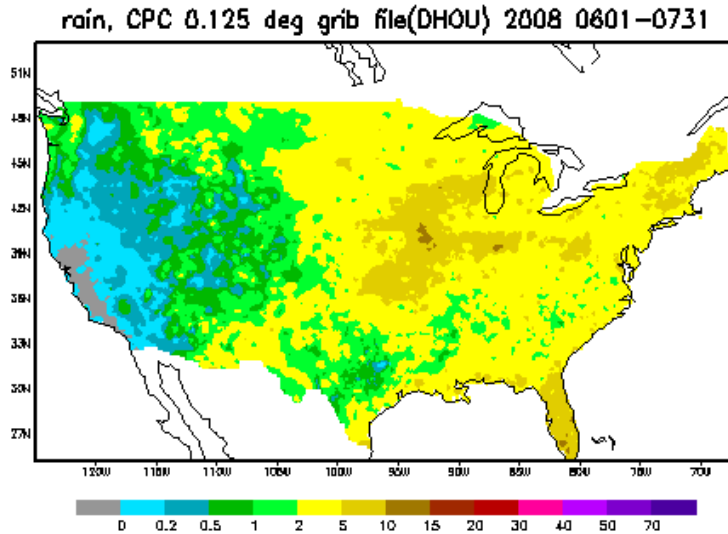
CCPA 02/05



STAGE IV 02/05



Comparison of Stage IV and CCPA Wrt. CPC

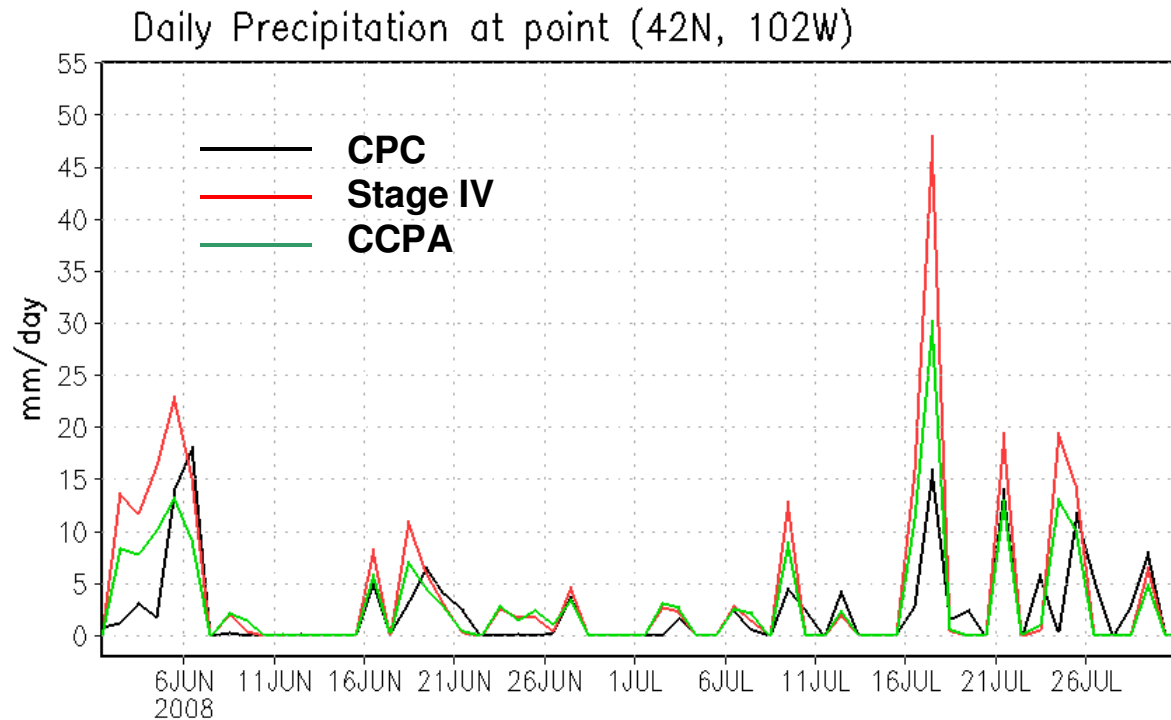


-Two Month Mean (June 1 – July 31, 2008)

For Stage IV and CCPA

- Aggregated from HRAP to 0.125 deg
- Aggregated from 6-hourly to daily

Comparison of time series of CPC, Stage IV and CCPA



- Example: A Point (42N, 102W) near Ashby, NE
- Selected from 0.125 deg datasets for June 1 – July 31 2008

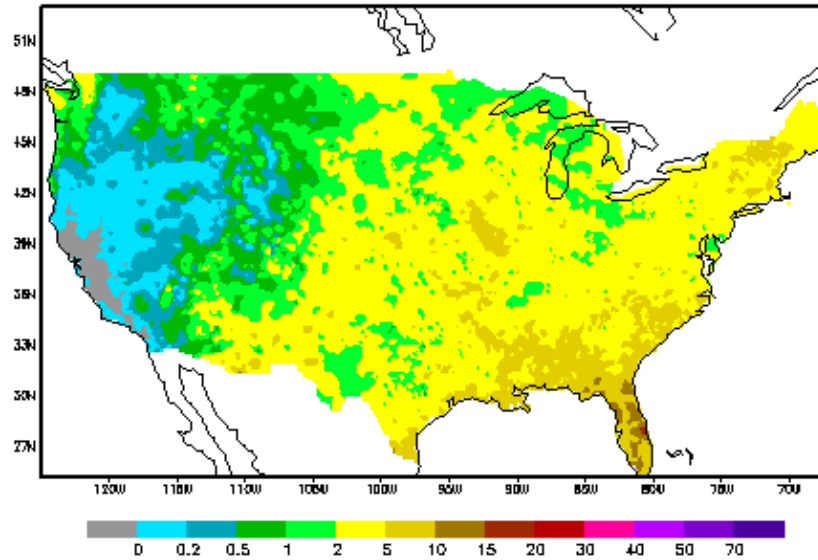
Cross Validation

- Hold historical data for a particular year (July to June next year).
- Re-estimate the regression coefficients with reduced historical data set.
- Apply the re-estimated regression coefficients to the Stage-IV data set of the held year to generate a data set similar to CCPA (call it CCPA_CV).
- Compare CCPA_CV (instead of CCPA) to CPC

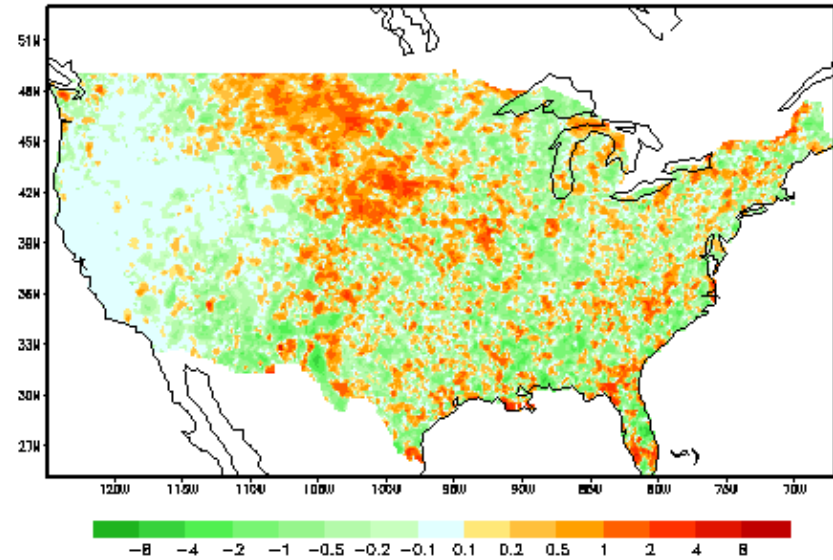
Cross Validation for the last year of the training period (July 2008 - Jun 2009)

Average July-Aug. 2008

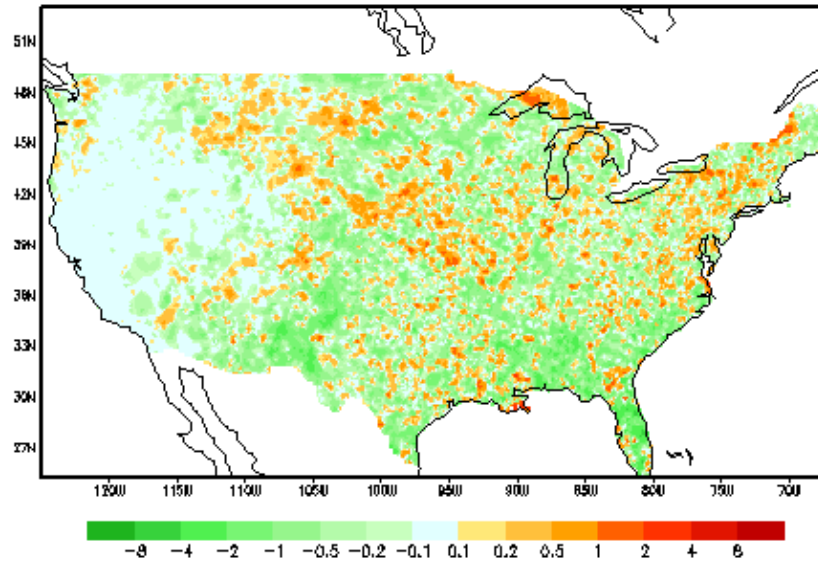
rain, CPC 0.125 deg grib file(DHOU) 2008 0701-0831



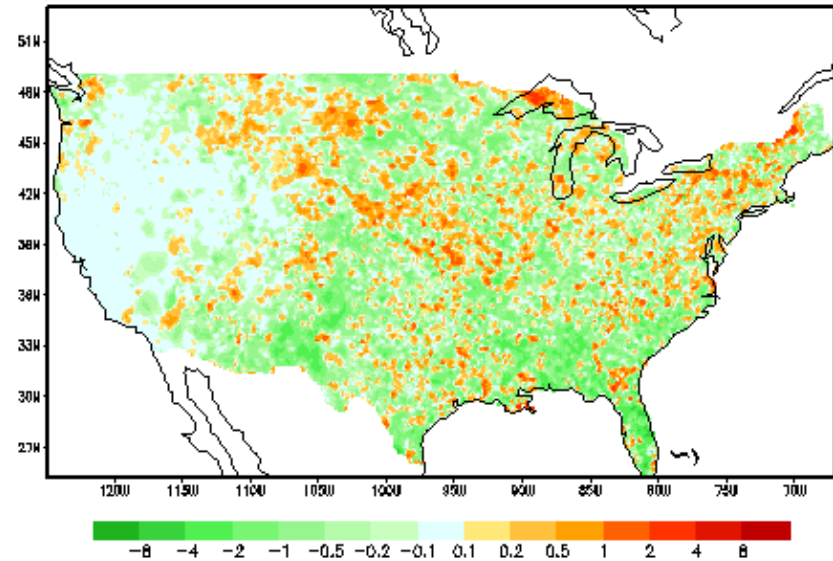
rain STAGE4 - CPC aggr. to 0.125deg 2008 0701-0831



rain ST4_ADJ - CPC aggr. to 0.125deg 2008 0701-0831

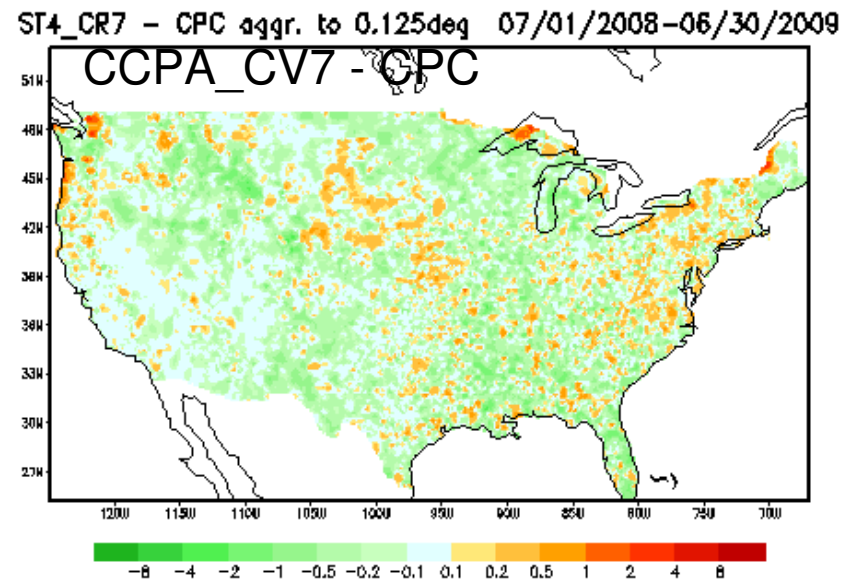
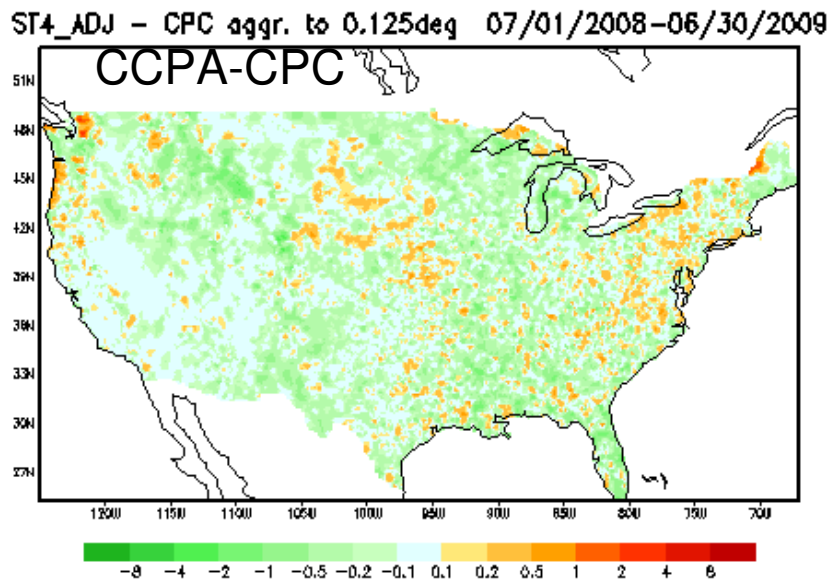
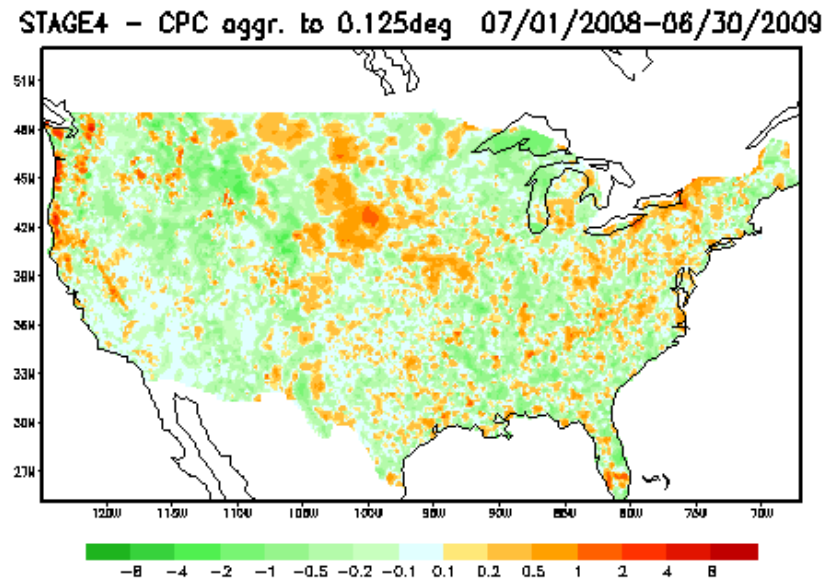
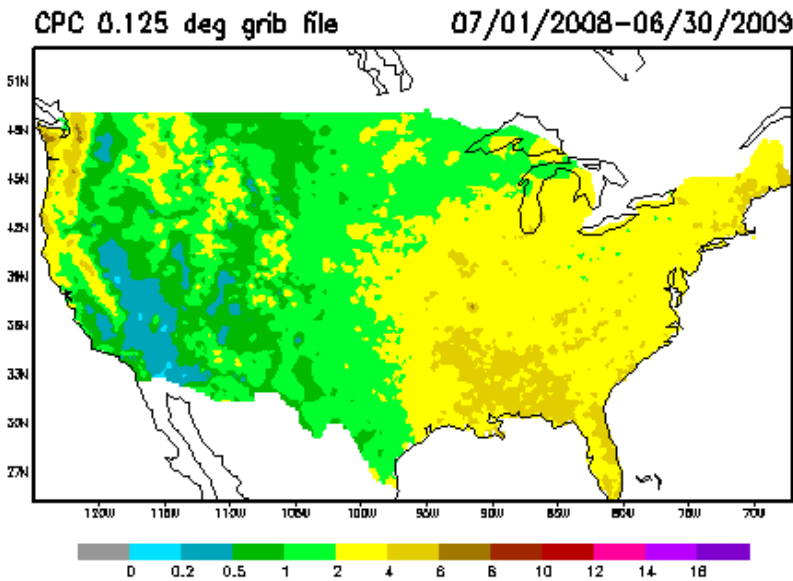


rain ST4_CR7 - CPC aggr. to 0.125deg 2008 0701-0831



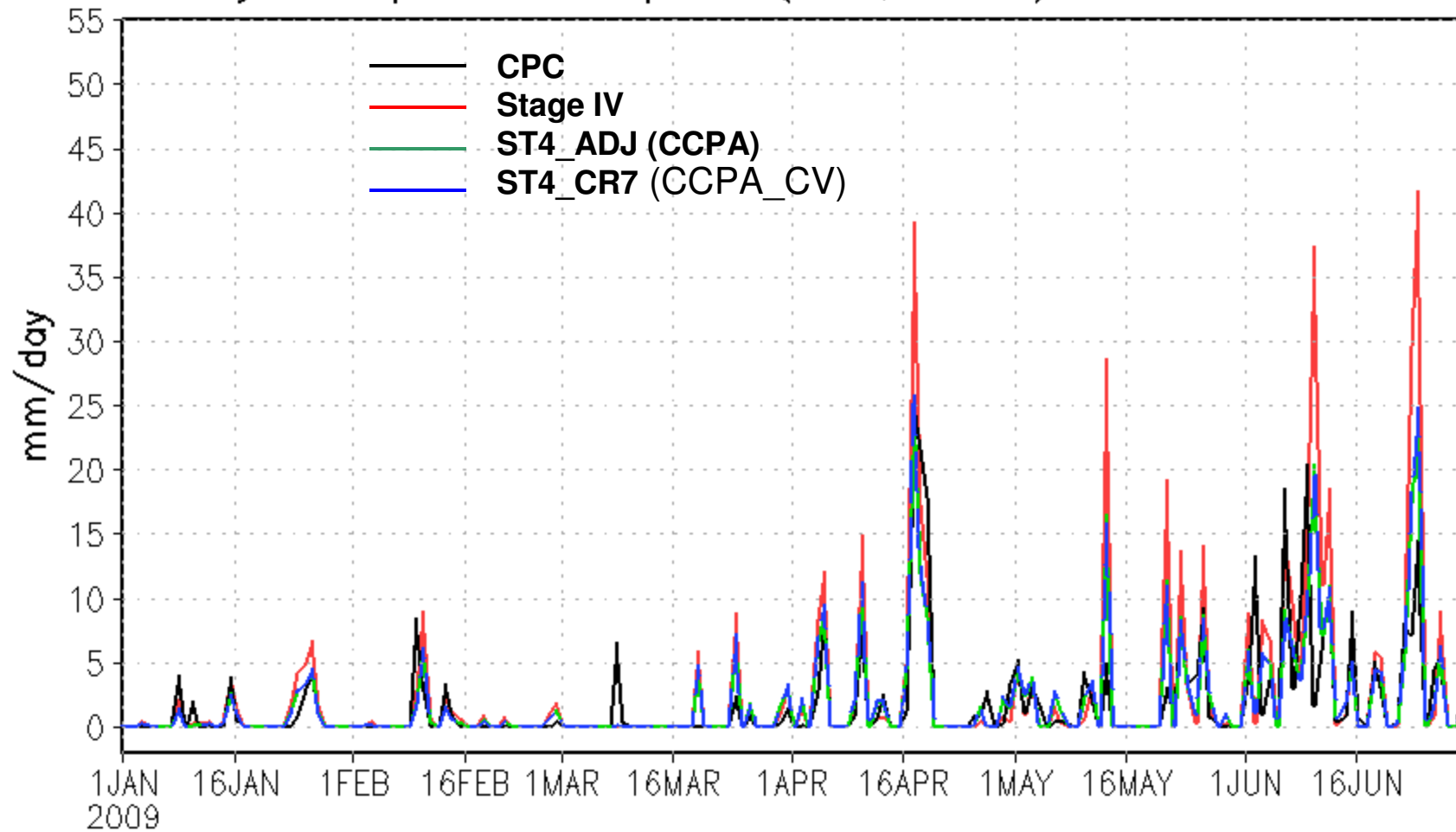
Cross Validation for the last year of the training period (July 2008 - Jun 2009)

Average July-Aug. 2008



Cross Validation: Time Series

Daily Precipitation at point (42N, 102W)



Concluding Remarks

- A new dataset of precipitation analysis, over CONUS at 6h, 4km resolution
- Statistical adjustment of Stage IV data with same resolution toward CPC analysis
- A combination of the Stage IV and CPC Unified Precipitation Analysis
- Simple linear regression at 0.125 degree and 24h accumulation
- Spatial interpolation and temporal smoothing to regression coefficients
- Keep the fine scale structures of Stage IV
- Closer to CPC Unified Precipitation Analysis, in the sense of climatology
- Provide a proxy of truth for precipitation forecast calibration and downscaling
- Operational implementation at NCEP, on July 13, 2010
 - Generate the historical data set of CCPA for 2002-2010

What is next?

- Cross validation for all years
- Quantitative evaluation
 1. Make comparisons of STAGE4, CCPA, CCPA_CV and CPC gauge analysis against ground measurements from RFC rain-gauge network.

QPE:

 - 24h (12UTC-12UTC) STAGE4, CCPA and CPC gauge analysis at 1/8 deg (or 1deg?), over CONUS

OBS (the Truth):

 - 24h (12UTC-12UTC) RFC rain-gauge data (in-situ measurements, point data) over CONUS
 - Raw gauge data box averaged to same analysis grid at 1/8 deg (or 1deg?)
 2. Verification statistics
 - Compute 24h (12UTC-12UTC) ETS, TSS, bias scores and RMSE over CONUS for STAGE4, CCPA and CPC gauge analysis
 - Generate scores, RMSE plots
- Periodic (annual) upgrading regression coefficients with increasing sample size
 - Updating coefficients a and b for real time CCPA
 - Re-generate the CCPA historical data set ?
- Improving the methodology