CCPA Precipitation Analysis: Data Set, Cross Validation and Evaluation

Dingchen Hou[1], Yan Luo[1] ,Yuejian Zhu[1],
Pingping Xie[2], and Ying Lin[1]

[1] Environmental Modeling Center/NCEP/NWS/NOAA
2 Climate Prediction Center/NCEP/NWS/NOAA

Acknowledgements

Mike Charles, Zoltan Toth, Roman Krzysztofowicz, Dong-Jun Seo,
Malaquias Pena, Bo Cui, Steve Lord, Letitia Soulliard, Julie Demargne, John Schaake

25th Conference on Hydrology
23-27 January, 2011, Seattle, WA

What is CCPA?

(Climatology-Calibrated Precipitation Analysis)

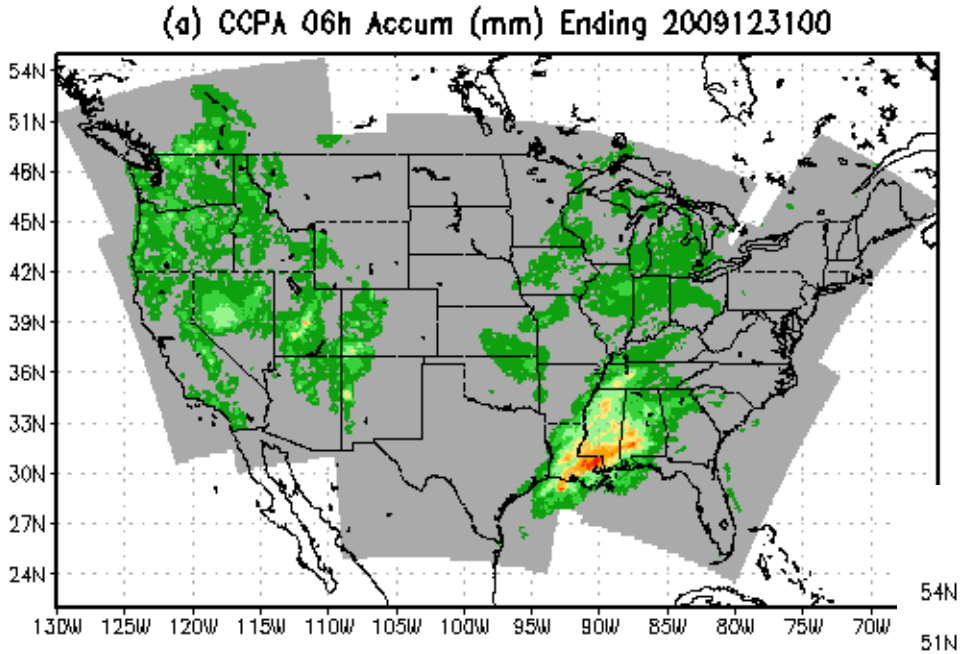
- A new dataset of precipitation analysis, over CONUS at 6h, ~4km resolution
- Statistical adjustment of Stage IV data toward CPC analysis
- Simple linear regression at 0.125 degree and 24h accumulation
- Spatial interpolation and temporal smoothing to regression coefficients
- Keep the fine scale structures 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

Status and Availability of CCPA data sets

- Operational implementation at NCEP on July 13, 2010
  - Real time generation of CCPA after STAGE IV
  - Generate at noon and update in the evening
- Generate the historical data set of CCPA for 2002-2010
- Product grids:
  - HRAP (primary)
  - NDGD, 0.125, 0.5 and 1.0 degree resolutions (byproducts)
- Contact information: Yan.Luo@noaa.gov
- CCPA website:

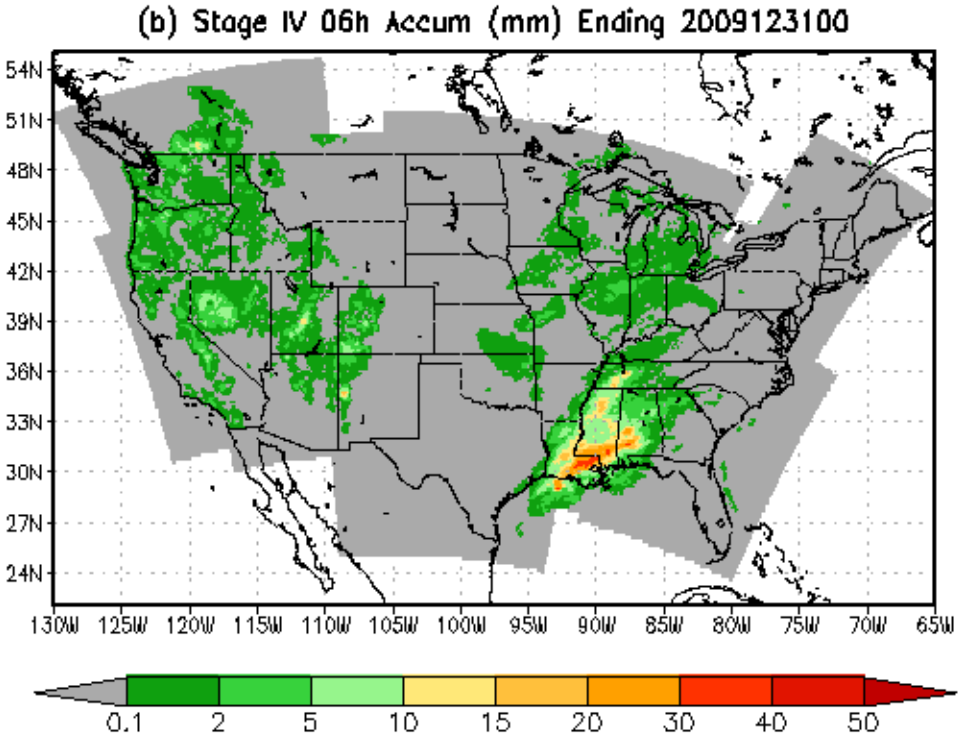
  http://www.emc.ncep.noaa.gov/gmb/yzhu/html/imp/201007_imp.html

Comparison of CCPA and Stage IV



6-h accumulation
(18Z , 30th to 00Z 31st, December 2009)
~4km HRAP

Spatial pattern correlation coefficient
= 0.990016



CCPA Evaluation Study

Goal

•Examine the impact and robustness of the CCPA methodology and evaluate the quality of CCPA data set

Data availability and processing

CPC Unified Precipitation Analysis :

- 1/8 deg, daily(12UTC-12UTC), 24 hr accumulation

RFC Rain Gauge Analysis:

- Point data, daily(12UTC-12UTC), 24 hr accumulation
- Box averaged to 1/8deg

Stage IV and CCPA:

- Aggregated from HRAP to 1/8 deg
- Aggregated from 6-hourly to daily

CVA (Cross Validation Analysis):

- An alternative data set of CCPA
- Cross validation method (see next slide)

Cross Validation Method

(Data holding technique, similar to Xie et al, 2007)

Linear Regression:
CPC= a * ST4 + b

Estimate a &b for CCPA from data pool
6/1/2002 – 7/31/2009 (7yr)

CCPA=**a** * ST4+**b**
(1/1/2002 – 6/30/2002)
(7/1/2002 – 6/30/2003)
(……)
(7/1/2008 – 6/30/2009)
(after 6/30/2009)
Same a&b for all years

Estimate a &b for CVA from data pool
6/1/2002 – 7/31/2003
6/1/2003 – 7/31/2004
6/1/2004 – 7/31/2005
6/1/2005 – 7/31/2006
6/1/2006 – 7/31/2007
6/1/2007 – 7/31/2008
6/1/2008 – 7/31/2009 6yr

CVA=**a** * ST4+**b** (7/1/2008 – 6/30/2009)

Estimate a &b for CVA from data pool
6/1/2002 – 7/31/2003
6/1/2003 – 7/31/2004
6/1/2004 – 7/31/2005
6/1/2005 – 7/31/2006
6/1/2006 – 7/31/2007
6/1/2007 – 7/31/2008
6/1/2008 – 7/31/2009 6yr

CVA=**a** * ST4+**b** (7/1/2007 – 6/30/2008)

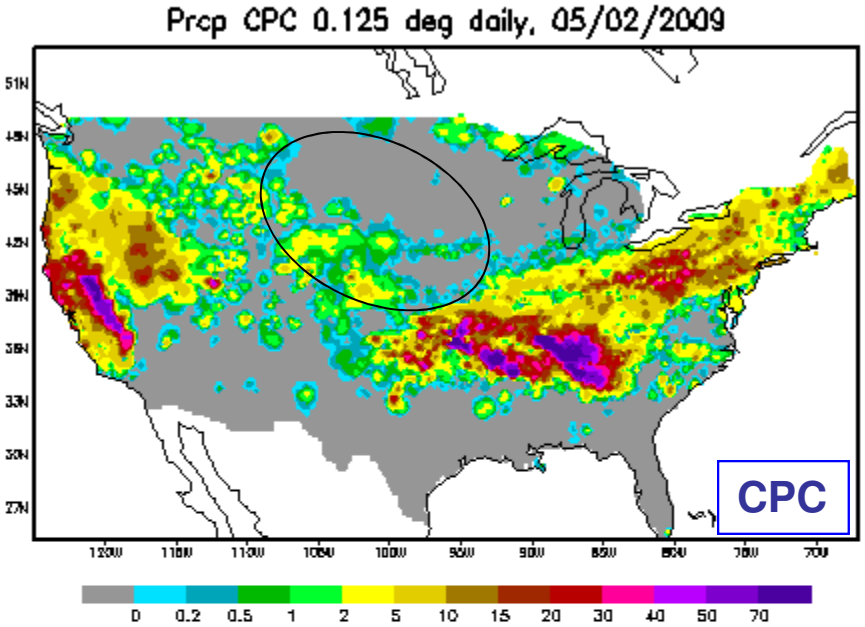
and so on, a&b vary year by year ……

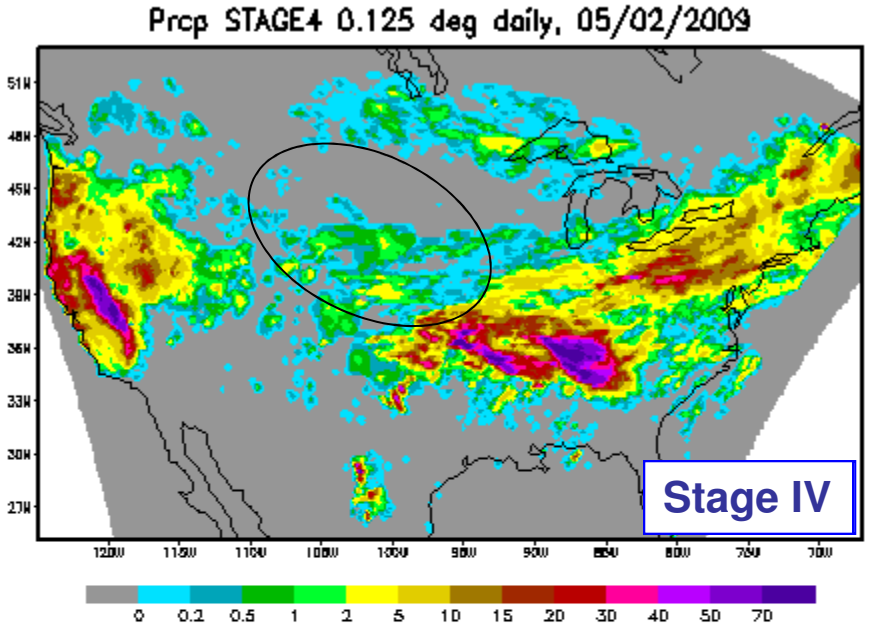
Evaluation method

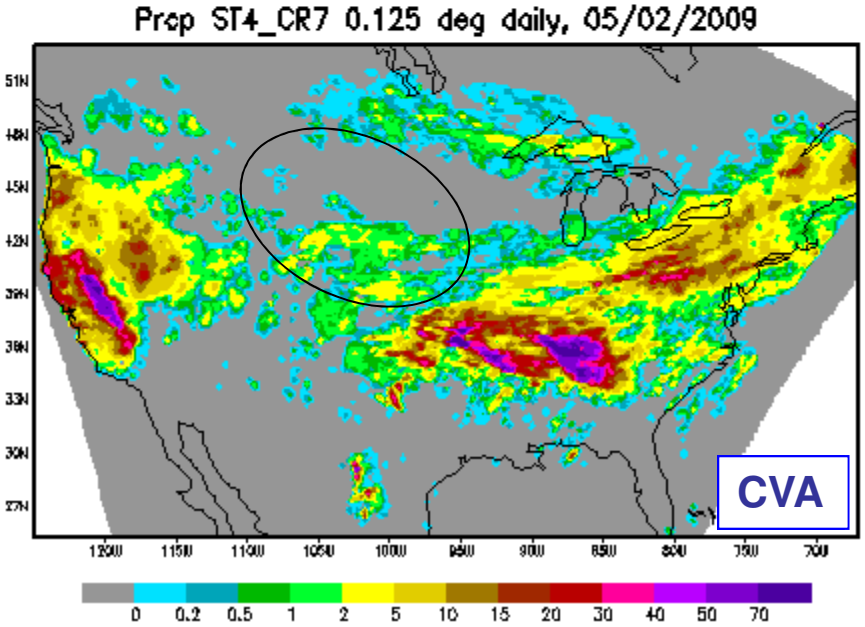
- Comparisons of ST4, CVA and CCPA against CPC
  - Daily based (12UTC-12UTC, 24 hr accumulation)
  - Daily cases
  - Annual Average
  - Time Series

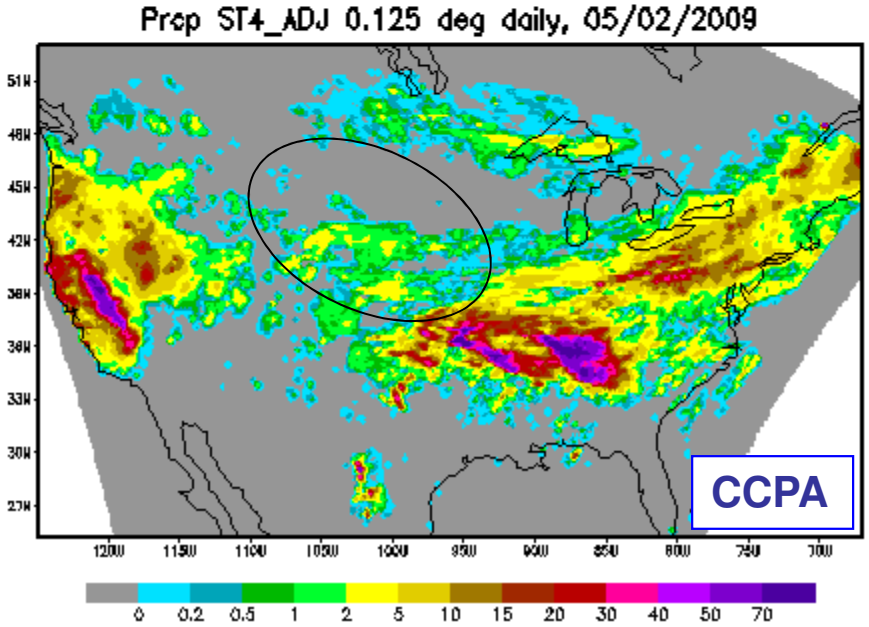
- Verifications of ST4,CVA and CCPA against RFC rain gauge observations
  - Daily based (12UTC-12UTC, 24 hr accumulation)
  - 1/8 deg over CONUS domain
  - Annual statistics (7/1/2008 – 6/30/2009 shown)
  - Verification Metrics: RMSE, ABSE, ETS and TSS scores
  - For various thresholds

Comparison of Stage IV, CVA and CCPA against CPC

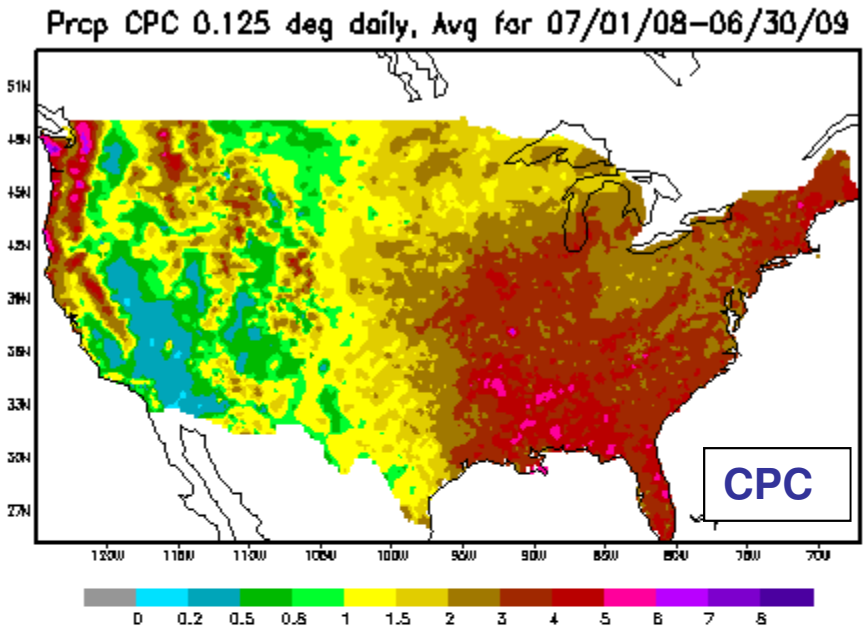


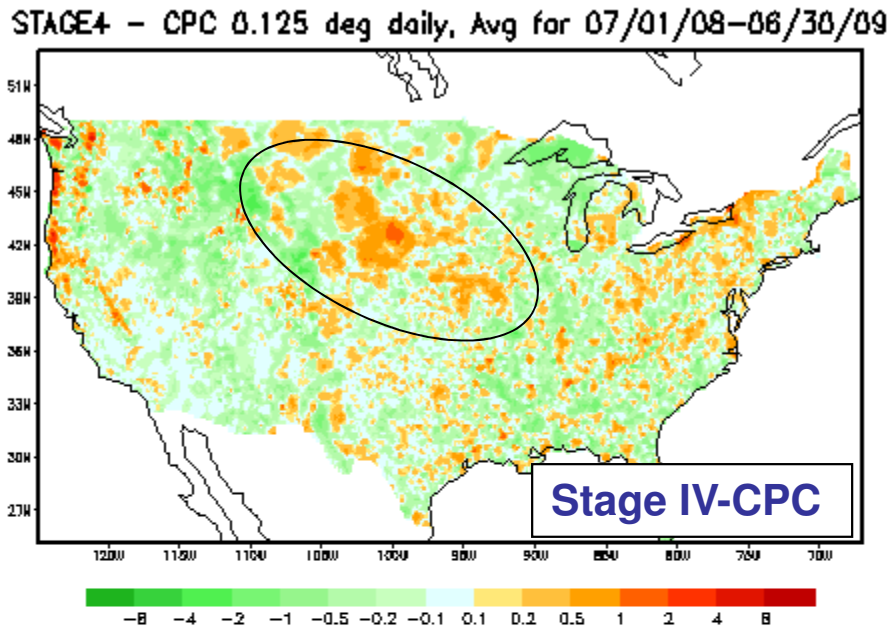


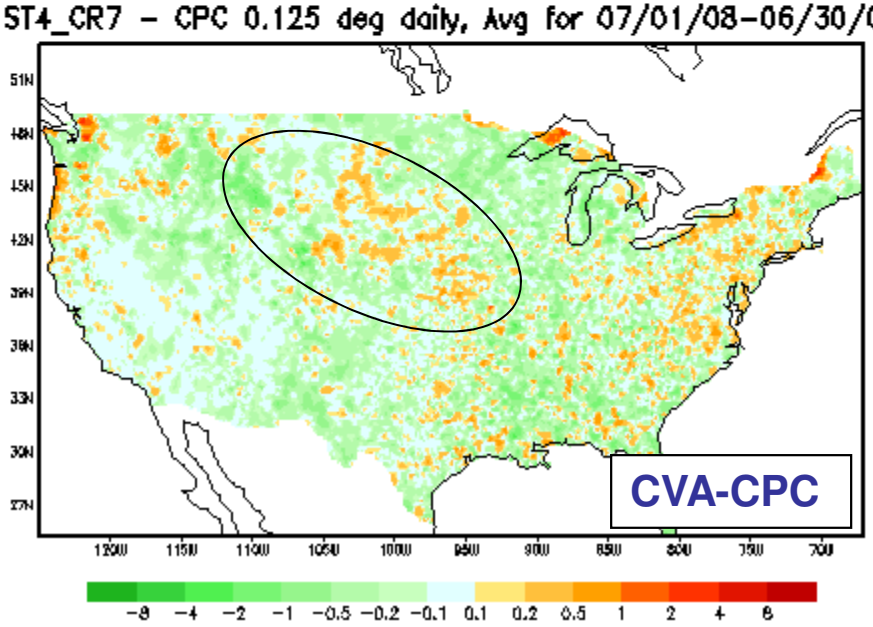


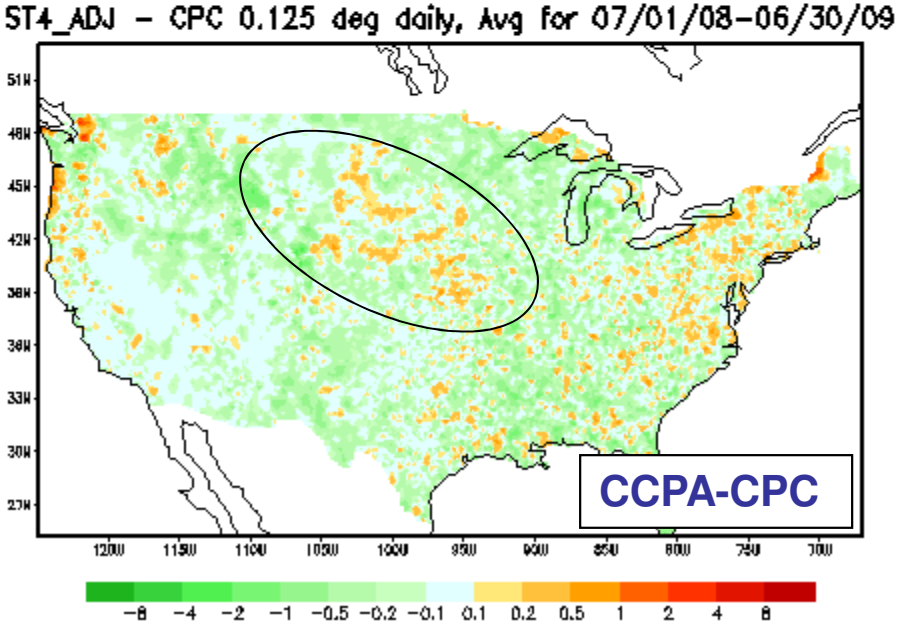


Comparison of Stage IV, CVA and CCPA against CPC

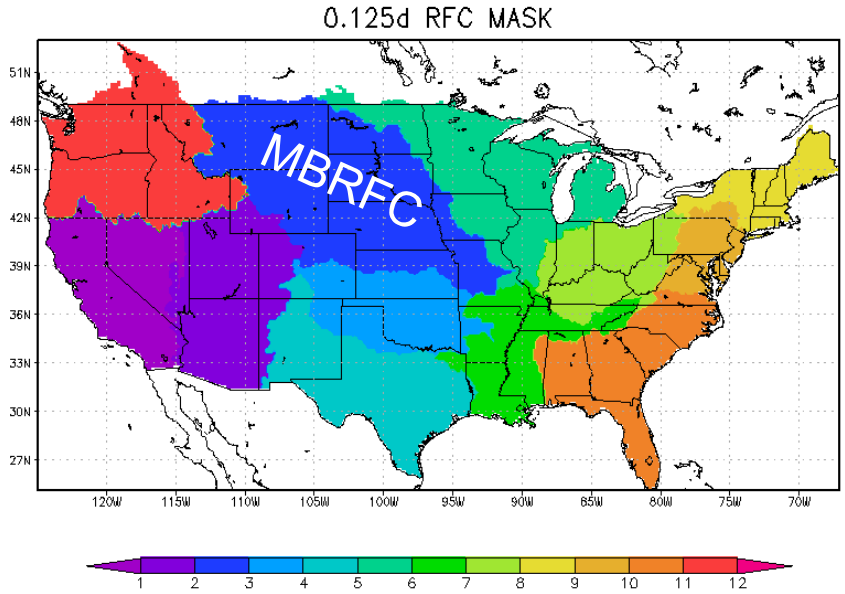


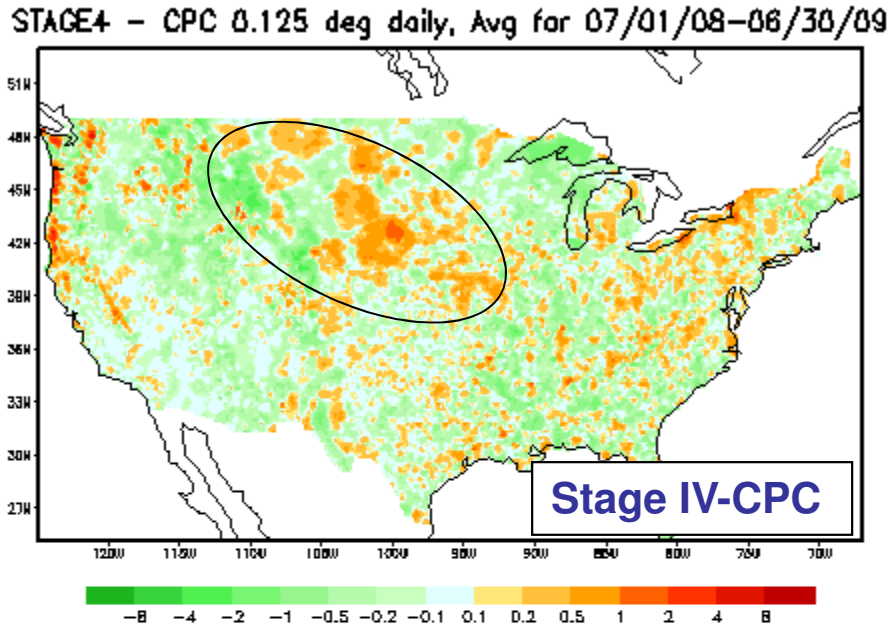


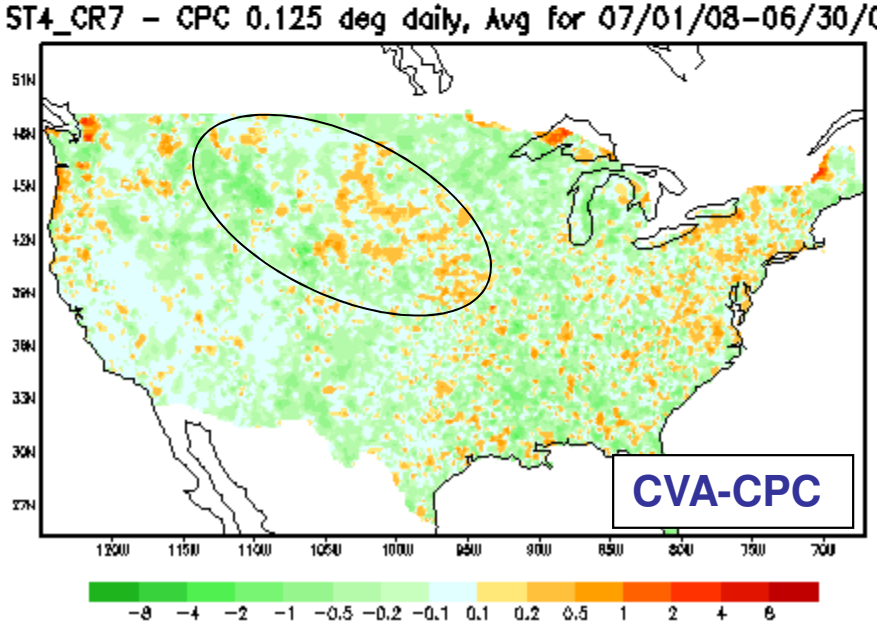


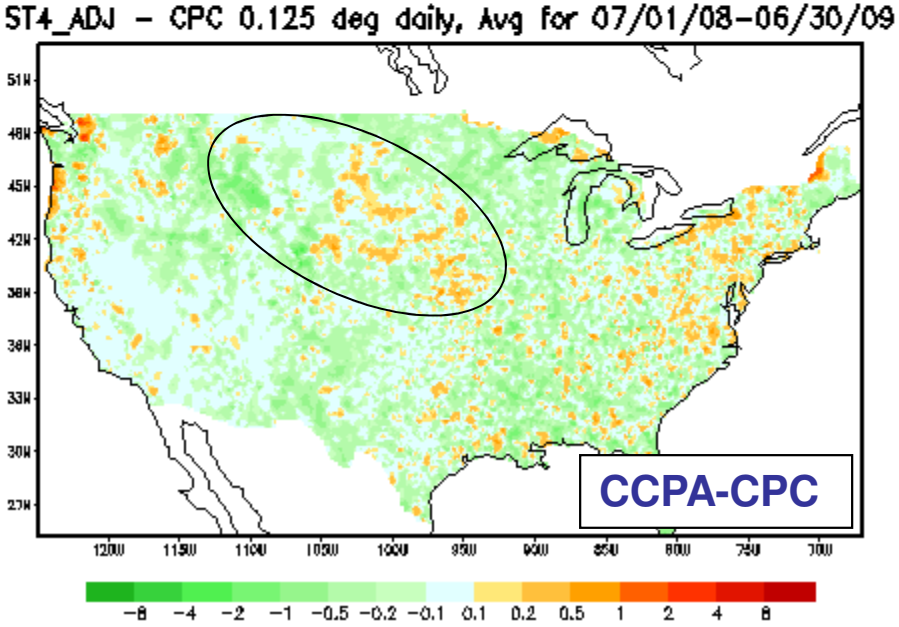


Comparison of Stage IV, CVA and CCPA against CPC

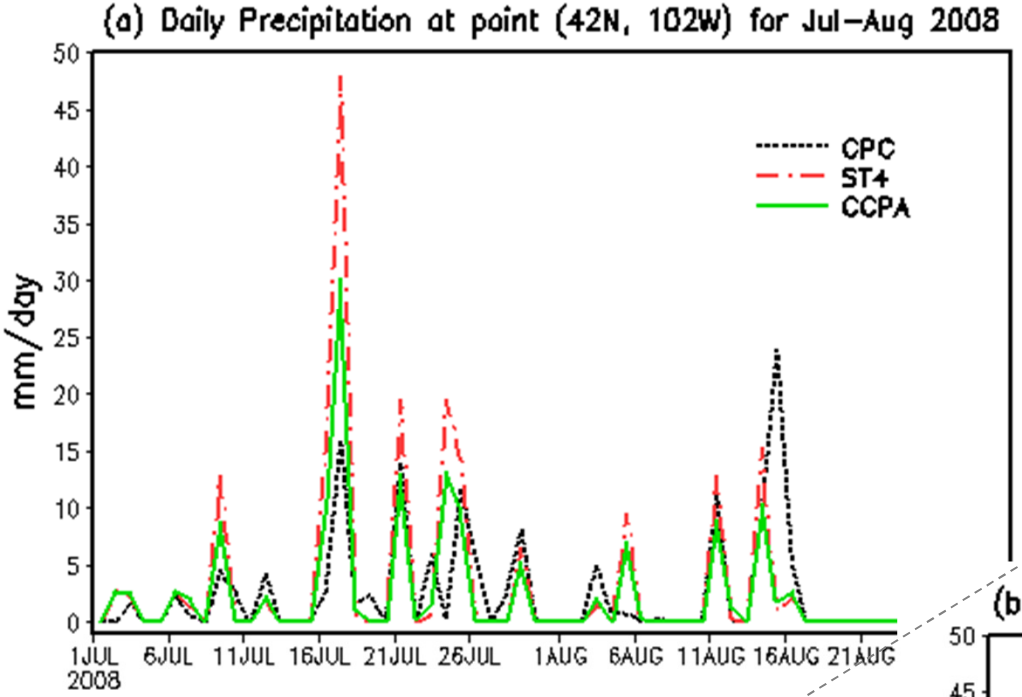


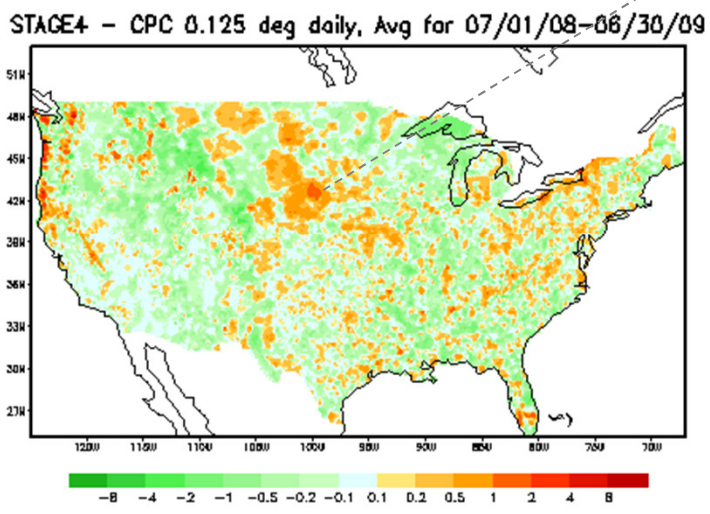




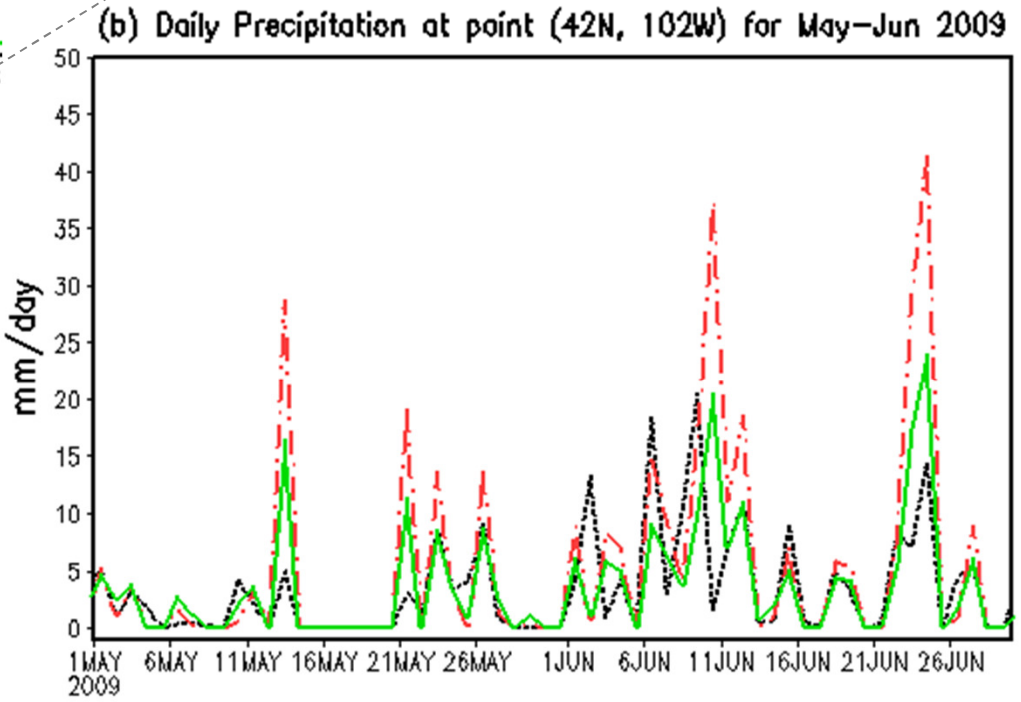


Comparison of CCPA and Stage IV against CPC



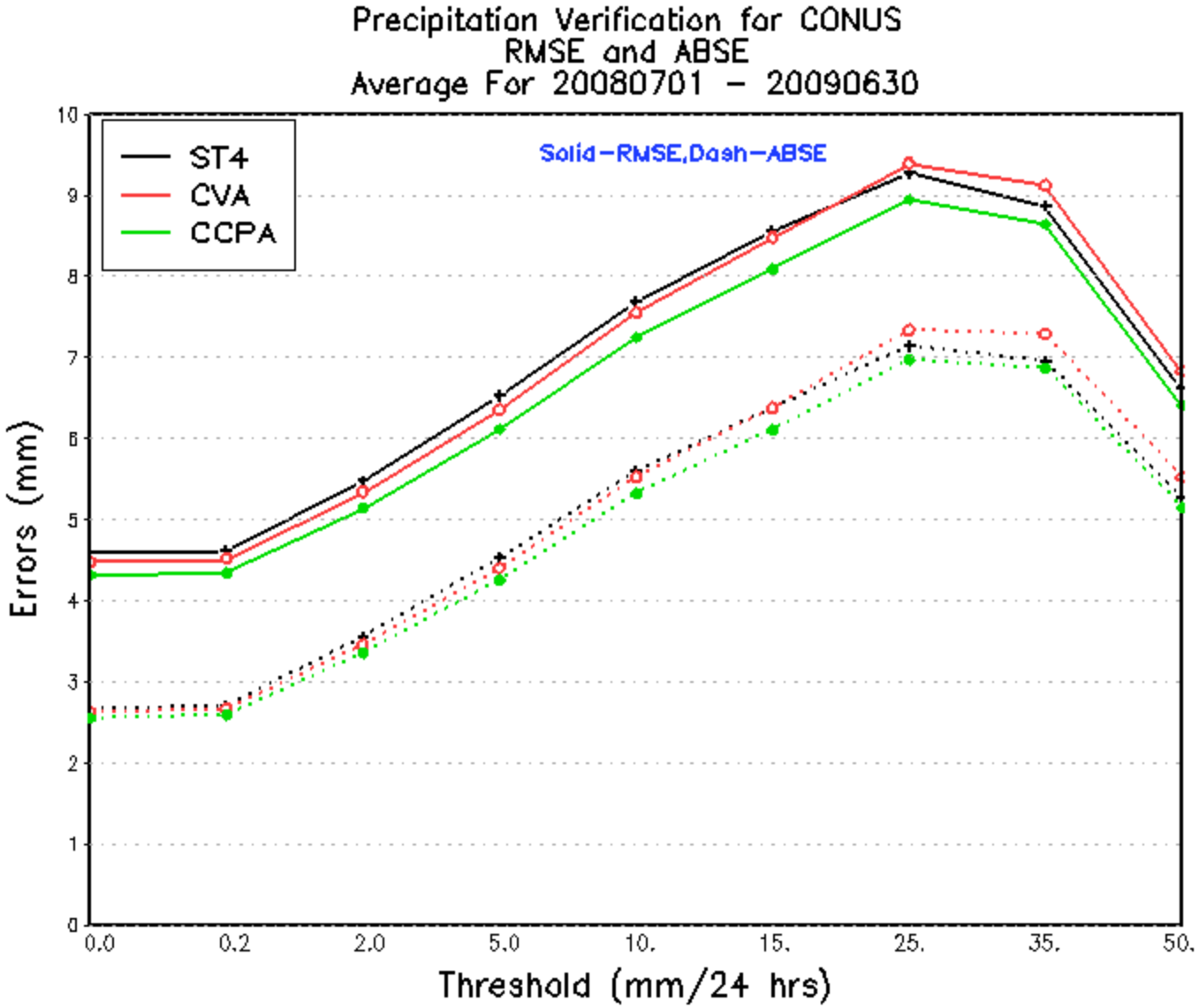


- Example: A Point (42N, 102W) near Ashby, NE in MBRFC
- Selected from 0.125 deg datasets for two warm and wet seasons



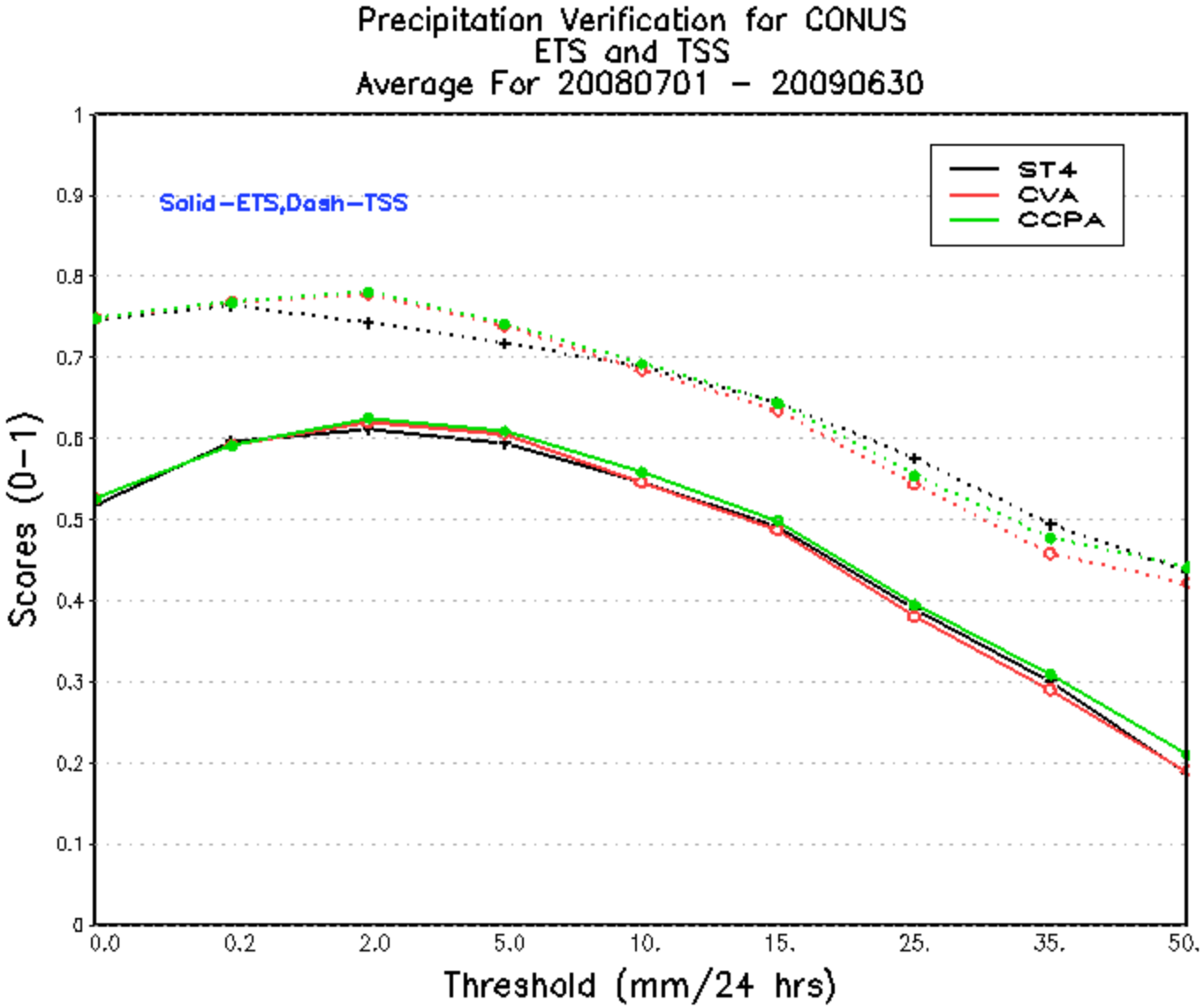
Verification against RFC-gauge network

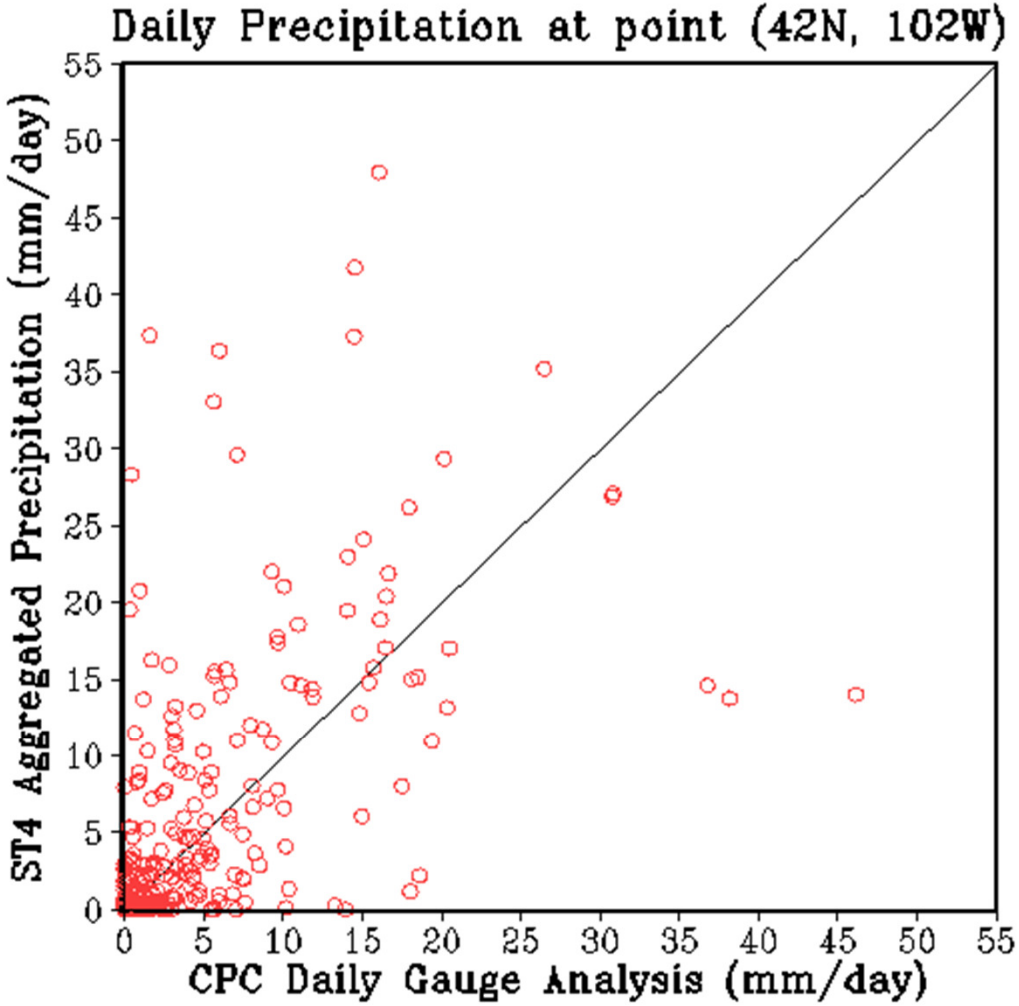
Results – RMSE and ABSE



Verification against RFC-gauge network

Results – ETS and TSS





Scatter plot of Stage IV against CPC. All data pairs here are sampled to estimate regression coefficients at point (42N,102W) for day July 1st (Julian day 182).

- Different sample size for the lower and higher precipitation ranges
- Small size for heavy precipitation
- A "linear" regression likely dominated by the lower precipitation points.

Conclusion

- CCPA methodology is robust; this is supported by the fact that cross validation analysis is fairly close to CCPA.
- Non-uniform quality control as one shortcoming of Stage IV is (at least partially) corrected.
- CCPA retains spatial and temporal patterns of Stage IV data set.
- CCPA long term average is closer to that of CPC analysis than Stage IV.
- The improvement is more significant with low and medium daily precipitation amounts.

Limitations and Future Work

- Limitations
  - Inadequate sample of high amount precipitation
  - Validity of the simple linear regression model

- Future Work
  - Perform annual updating of the regression coefficients with increased sample size
  - Employ more realistic non-linear regression models
  - Other calibration methods