# Investigation of Model Precipitation Biases over the Indian Subcontinent in the CORDEX South Asia Hindcast

J. Kim[1], D. Waliser[2], C. Mattmann[2], P. Ramirez[2], H. Lee[2], P. Loikith[2], M. Bounstani[2], C. Goodale[2], A. Hart[2], J. Sanjay[3], M.V.S. Rama Rao[3], R. Krishnan[3], M. Mujumdar[3], S. Ingle[3]

[1]University of California Los Angeles

[2]Jet Propulsion Laboratory/NASA

[3]Center for Climate Change Research, Indian Institute of Tropical Meteorology

# Introduction

- Precipitation is among the most important climate variables for the Indian subcontinent region.
  - Water resources, natural disasters, agriculture, ecosystems.
- Evaluation of models' capability in simulating regional precipitation is a crucial part of the application of climate projections to local impact assessments.
- Sparse observation network in the mountainous regions can introduce significant uncertainties in the "*observational*" data against which model performance is measured.
- This study examines the observed and simulated precipitation over the Indian subcontinent, fundamental in applying the model results to climate change impact assessment.
  - Uncertainties in the observational datasets
  - Model performance in various parts of the Indian subcontinent.

# Data

- Multiple observations are used for model evaluation
  - CRU3.1: Raingauge-based gridded analysis, land surfaces, 0.5deg resolutions
  - UDEL: Raingauge-based gridded analysis, land surfaces, 0.5deg resolutions
  - TRMM: Satellite- and raingauge-based gridded analysis, 0.25deg resolutions
  - GPCP: Satellite- and raingauge-based gridded analysis, 2.5deg resolutions

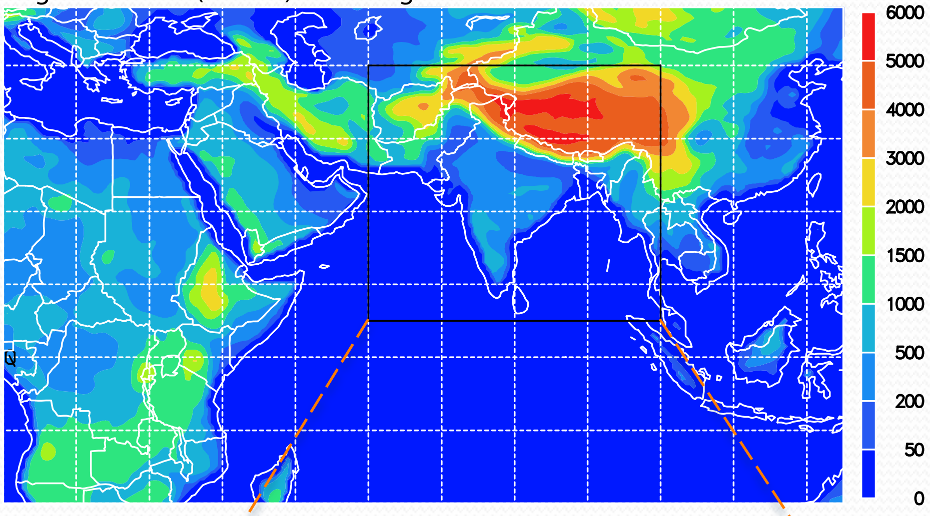
- Models
  - Model simulations were performed by the IITM authors
  - RCM3 with Emanuel or Grell convection scheme
  - WRF3 with Betts-Miller-Janjic or Kain-Frisch scheme

- Evaluation period
  - January 1998 – December 2006 (9 years)
  - The period of evaluation is limited by the TRMM data (beginning) and the simulation period.
  - E*valuation for a longer period using three observations* (*excluding TRMM*) *yields qualitatively similar results as the shorter one presented here.*

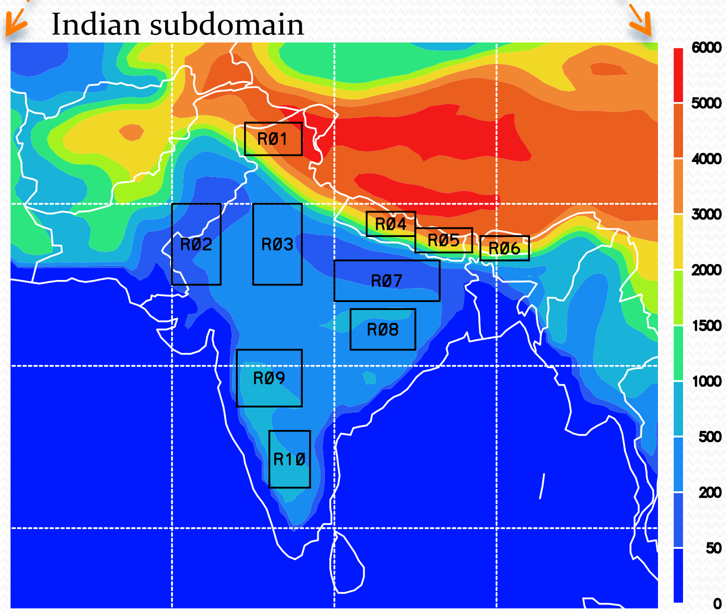
# Domain and the Model Terrain:

- Model precipitation is evaluated over the Indian subcontinent part of the CORDEX-South Asia simulation domain using Regional Climate Model Evaluation System (RCMES).

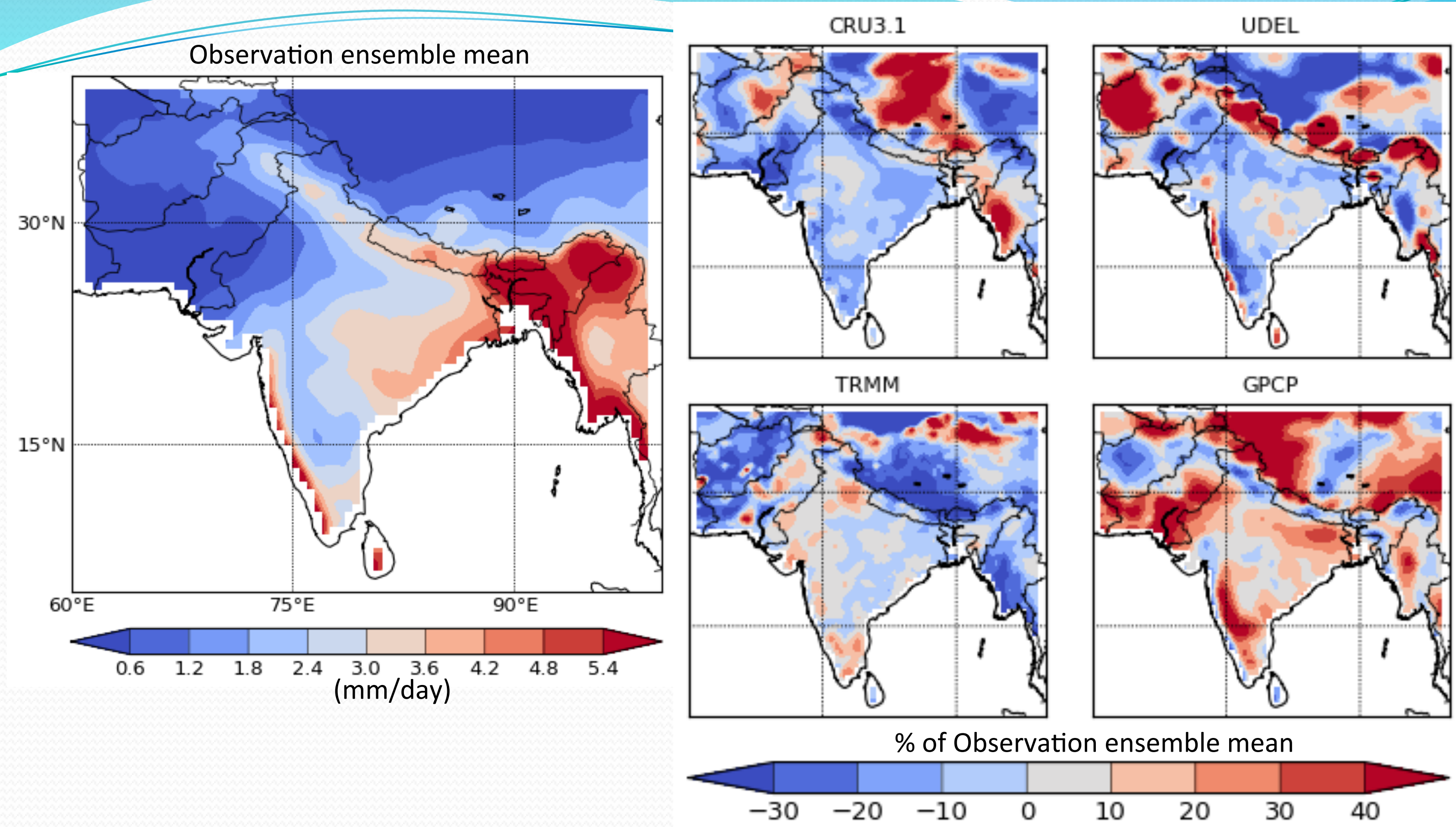
RegCM4 Terrain (meters) at 0.5 deg horizontal resolution




Indian subdomain

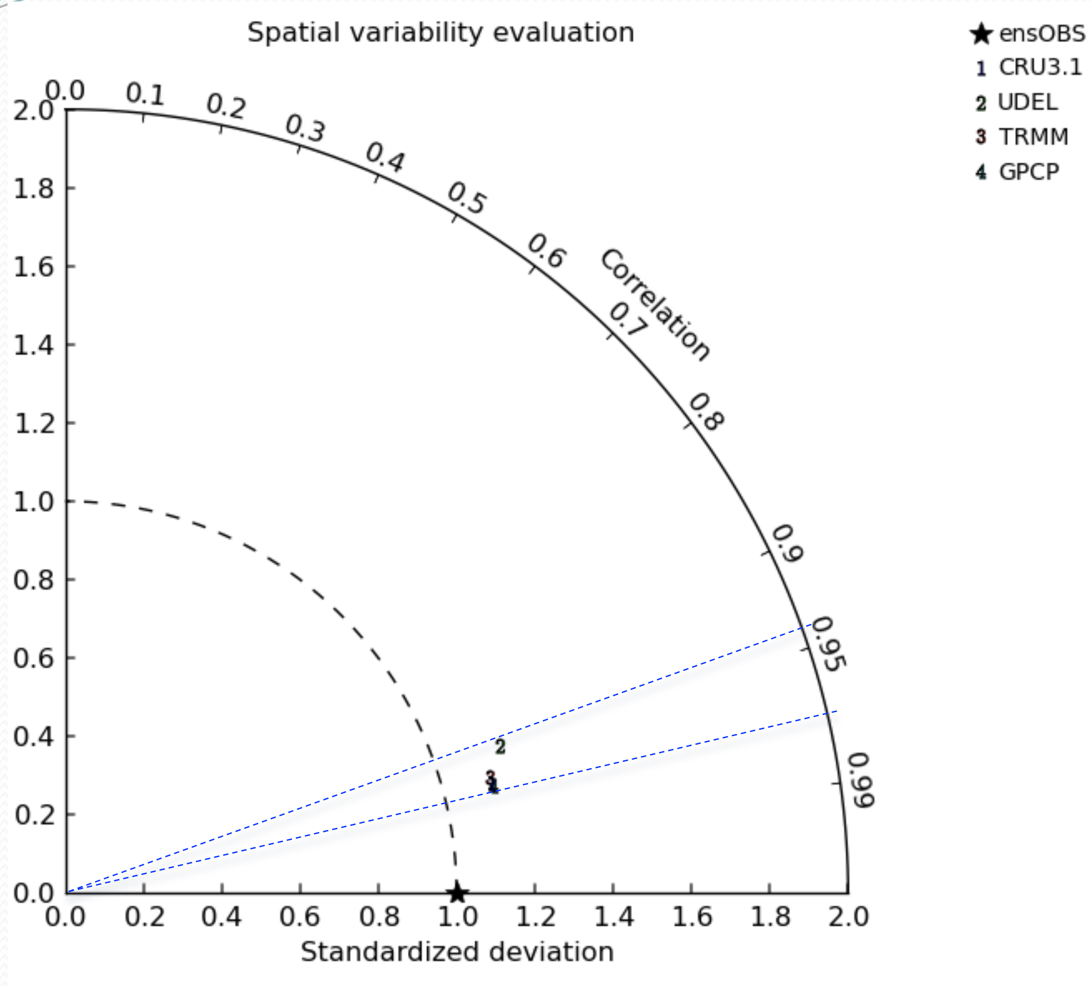


## Observed Annual-mean Precipitation



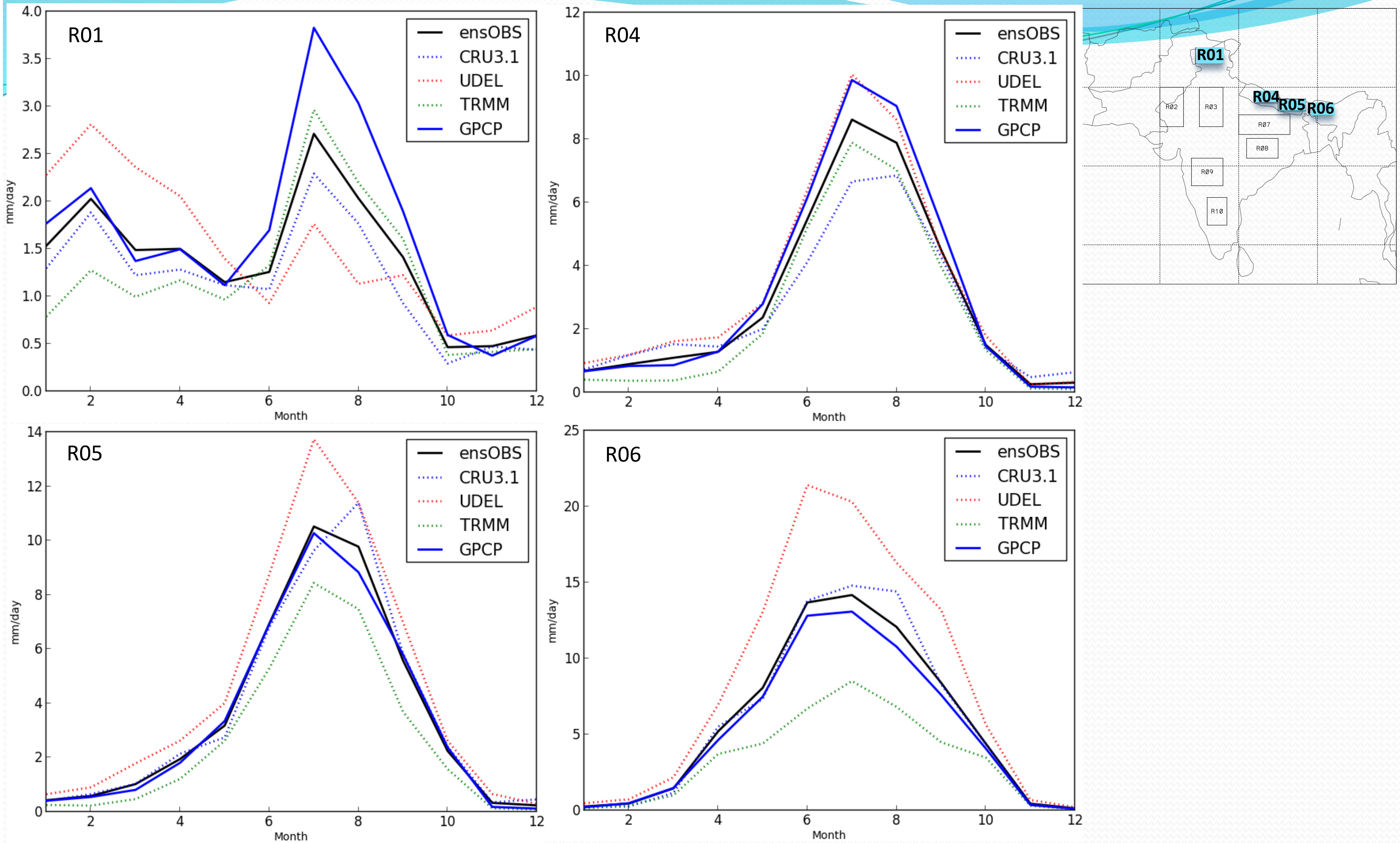

- Observational data are the key information for model evaluation.
- Wide variations amongst datasets, *especially in the regions of complex terrain*:
  - northern India, Nepal, Bhutan, Tibetan Plateau, west coast of India (facing the Arabian Sea).
- The large disagreement amongst these observation data may originate from:
  - the lack of observation stations in the northern mountain region
  - coarse horizontal resolution (GPCP)

# Spatial Variability of the Observed Annual-mean Precipitation



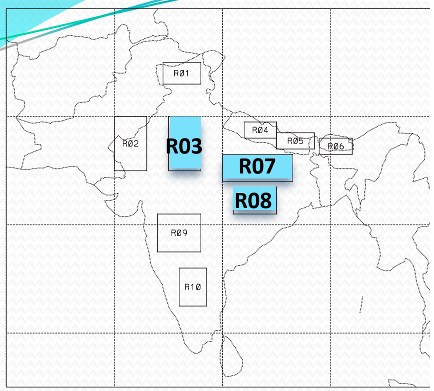

- Uncertainties in the observed spatial variability of the annual-mean precipitation due to different observation datasets are examined using the Taylor diagram.
- Compared against the ensemble of the four observation datasets:
  - Magnitude of the spatial variability of individual datasets is similar (1.15-1.2)
  - The pattern correlation with the observation ensemble ranges 0.95 - 0.97
  - UDEL data are the most noticeable outlier among the four observations.

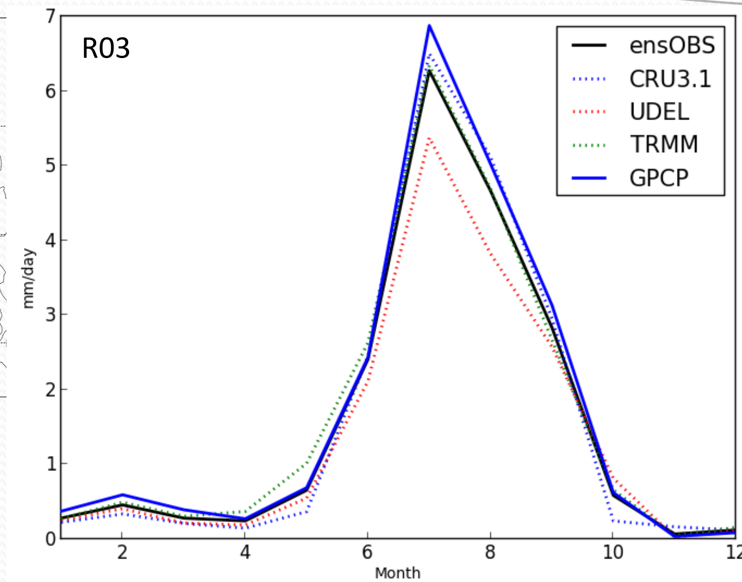
# Observed Precipitation Annual Cycle: Northern Mountains

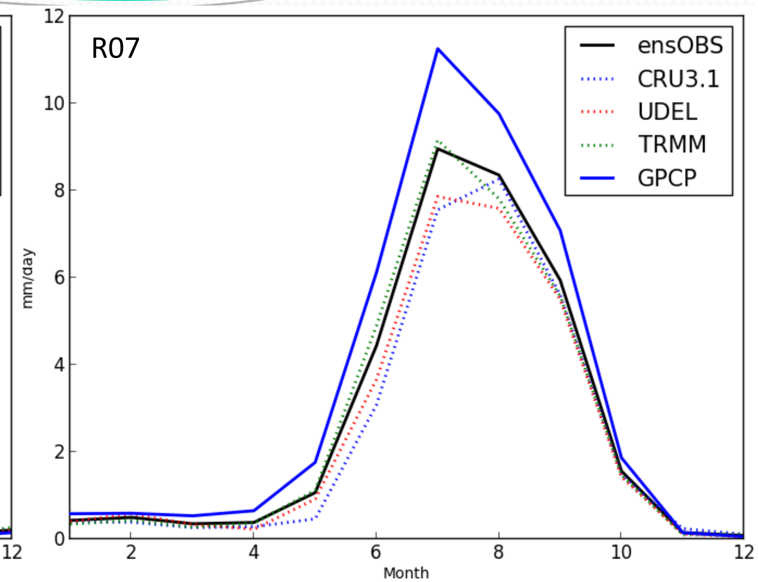



- Huge inter-dataset variations in the observed precipitation annual cycle over the northern mountain regions, especially the southern slope of the Hindu Kush Mountains.

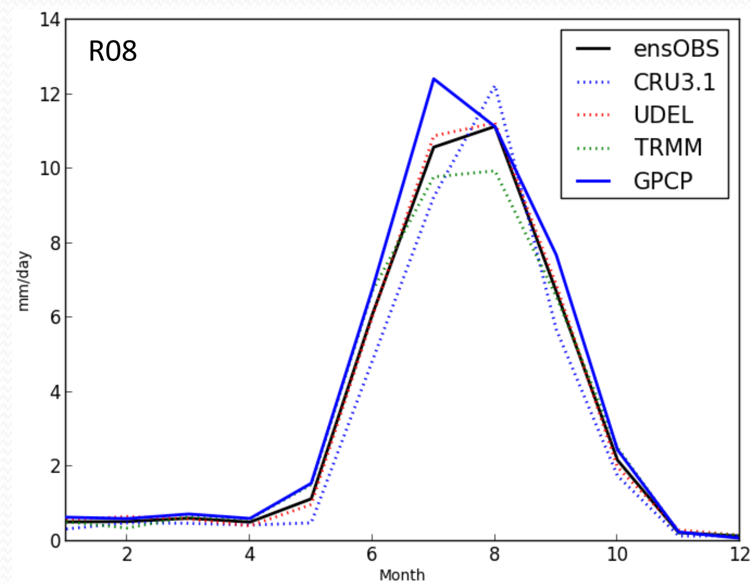
# Precipitation Annual Cycles: Northern Plains



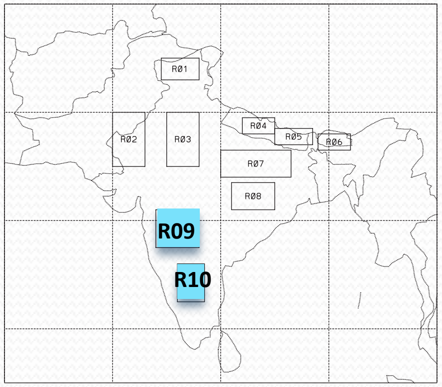



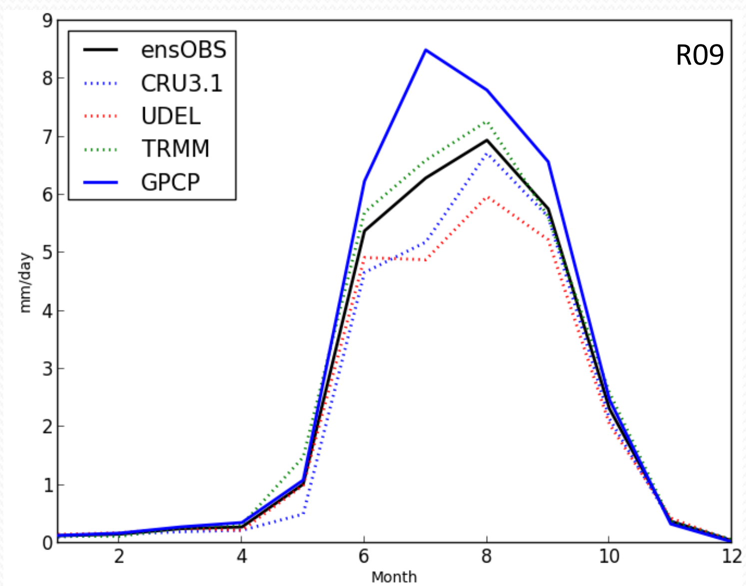

- Observations are much more consistent over the northern plains regions compared to the northern mountain regions.
- Still, their exist noticeable differences in the observed precipitation amongst datasets.

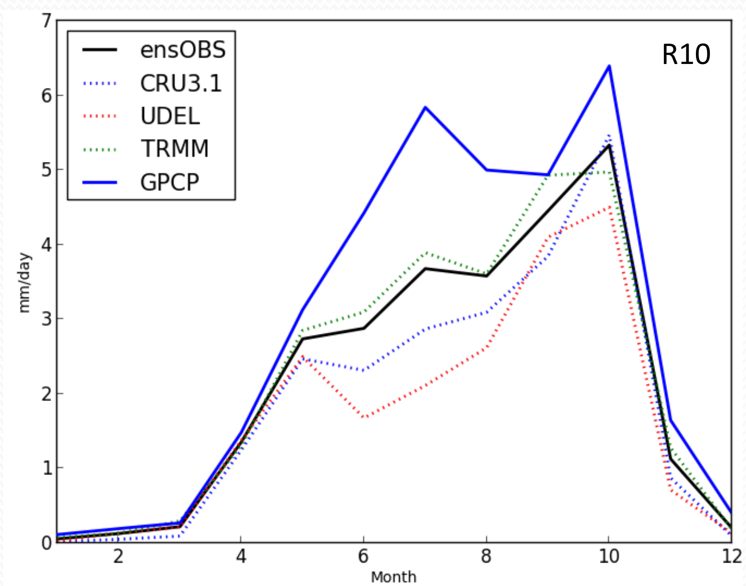


# Observed Precipitation Annual Cycle: Southern India




- Inter-dataset variations in the observed precipitation annual cycle is also large, although not as large as over the northern mountain regions.
- The most noticeable inter-observation variation occurs in the early (June - Aug) monsoon season over the southwestern coastal mountain regions (R10).
  - Similar problem also occurs for R09.
- For both regions, GPCP is the most serious outlier and overestimates precipitation throughout the year.
  - May be a problem related with horizontal resolution



# Observed Precipitation Data over the Indian Subcontinent Region

- There exist noticeable differences amongst '*observed*' precipitation datasets that are popularly used in climate research.

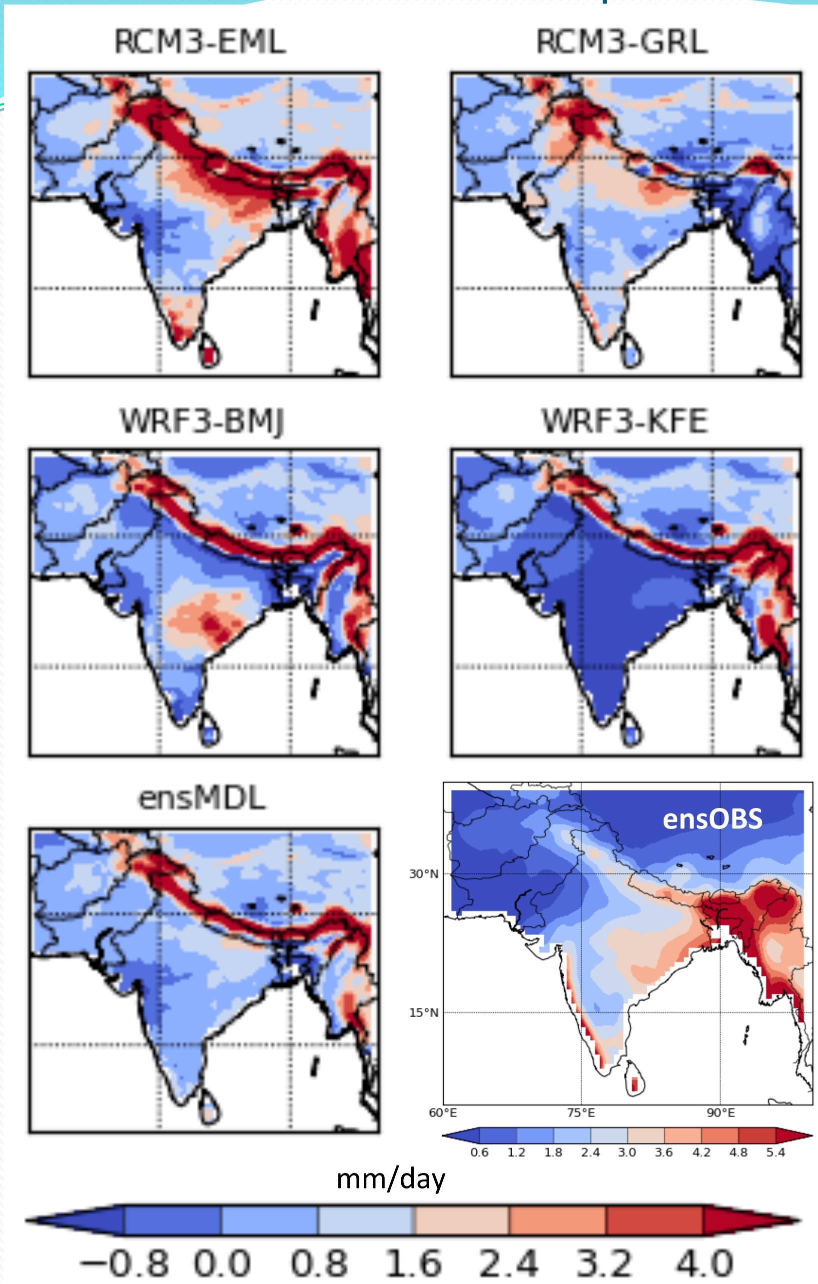
- Observational uncertainties vary strongly according to regions and terrains. The largest observational uncertainties occur in the mountain regions
  - The Himalayas and the Hindu Kush Mountains, Tibetan Plateau.
  - Southwestern India with steep coastal terrain.

- Not shown, but observational uncertainties are larger for the Sept – February period than the March – August period

- Thus, observational uncertainties must be kept in mind when evaluating model precipitation in this region.

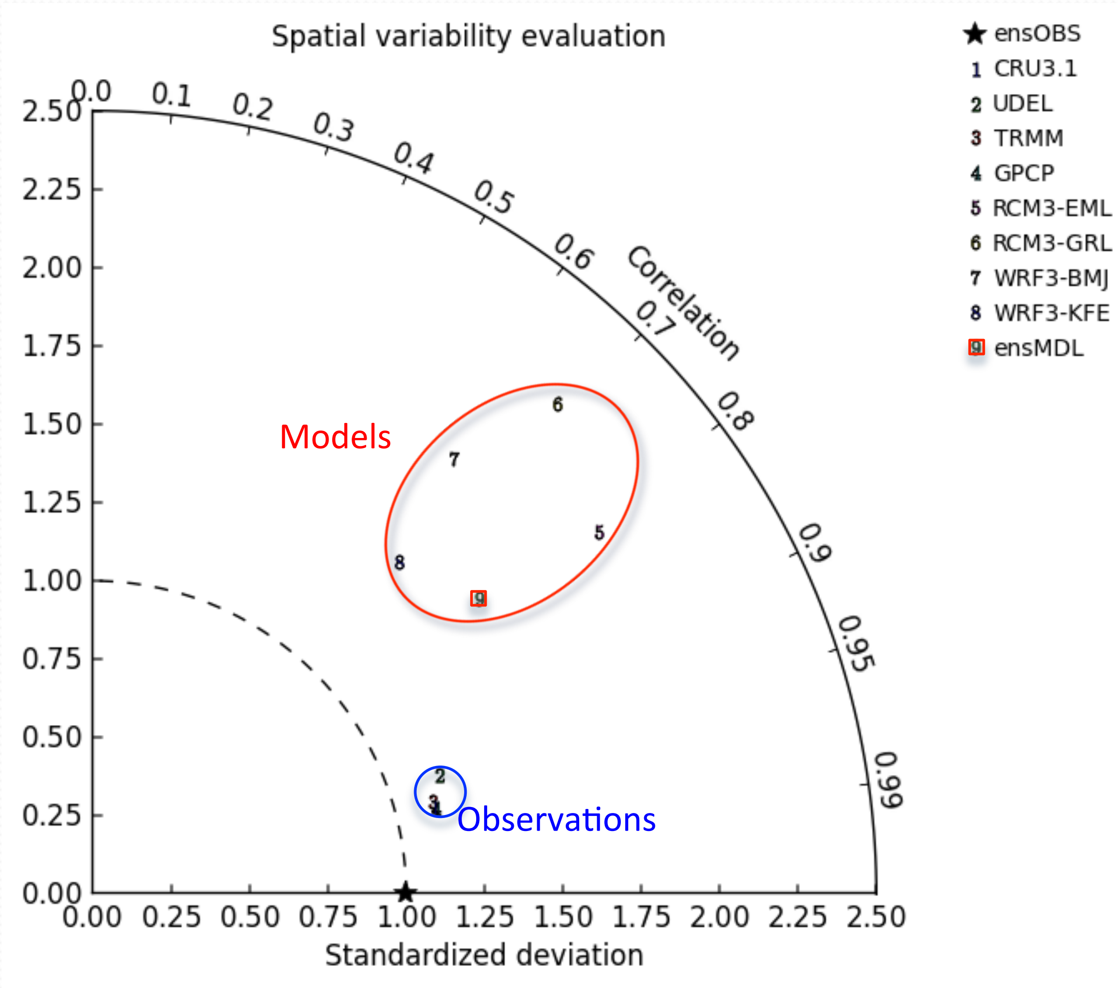
- Now, evaluate the RCM-simulated precipitation …

# Annual-mean Precipitation Biases from the Observation Ensemble



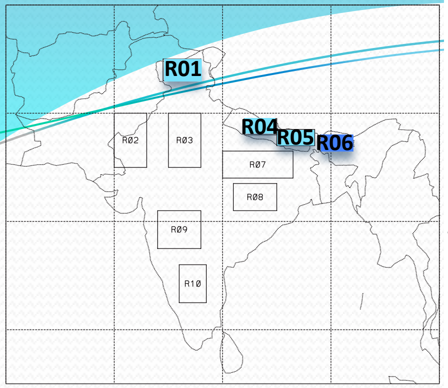

- The simulated annual-mean precipitation is compared against the four-observation ensemble.
- The most noticeable model errors, both in terms of absolute amounts and relative to the climatological mean, occur in the region of steep terrain over the southern slope of the Himalayas and the Hindu Kush Mountains.
  - The large errors in the northern mountain regions occur regardless of RCMs and convection schemes used in this study.
- Models generate large wet/dry biases over the Myanmar region.
  - These model biases are not systematic (2 wet and 2 dry).
  - Not clear if model performance vary systematically according to models and/or convection schemes.

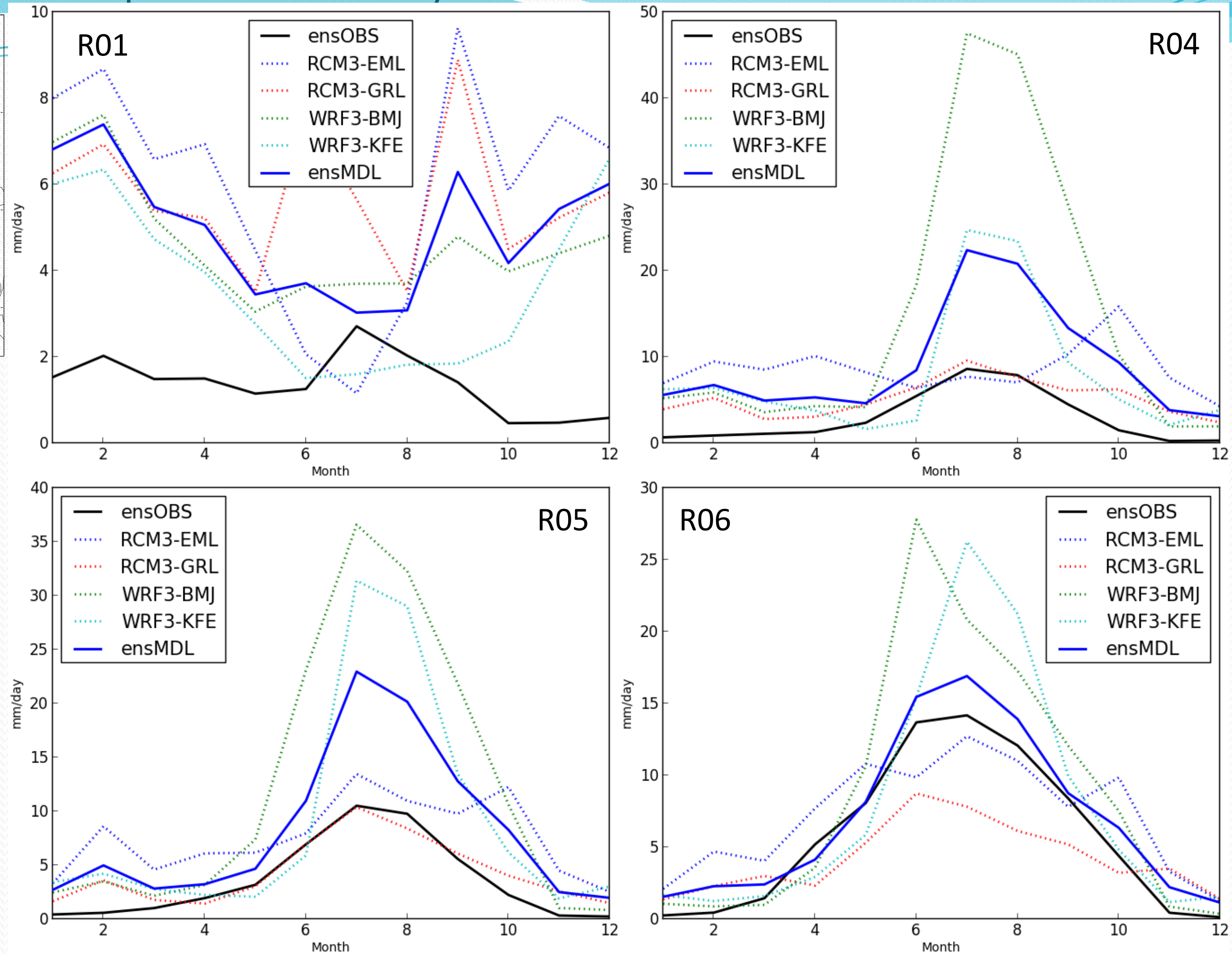
# Spatial Variability of the Simulated Annual-mean Precipitation




- All RCMs used in this study overestimate the observed spatial variability by 50% - 100% of the observation ensemble.
- Spatial pattern correlation between the simulated annual-mean precipitation and observation ensemble ranges from 0.6 to 0.8.
- Model ensemble outperforms individual models measured in terms of RMSE.
- Models and observations are clearly separated
  - *Observational uncertainties are large, but may not be crucial in model evaluation.*

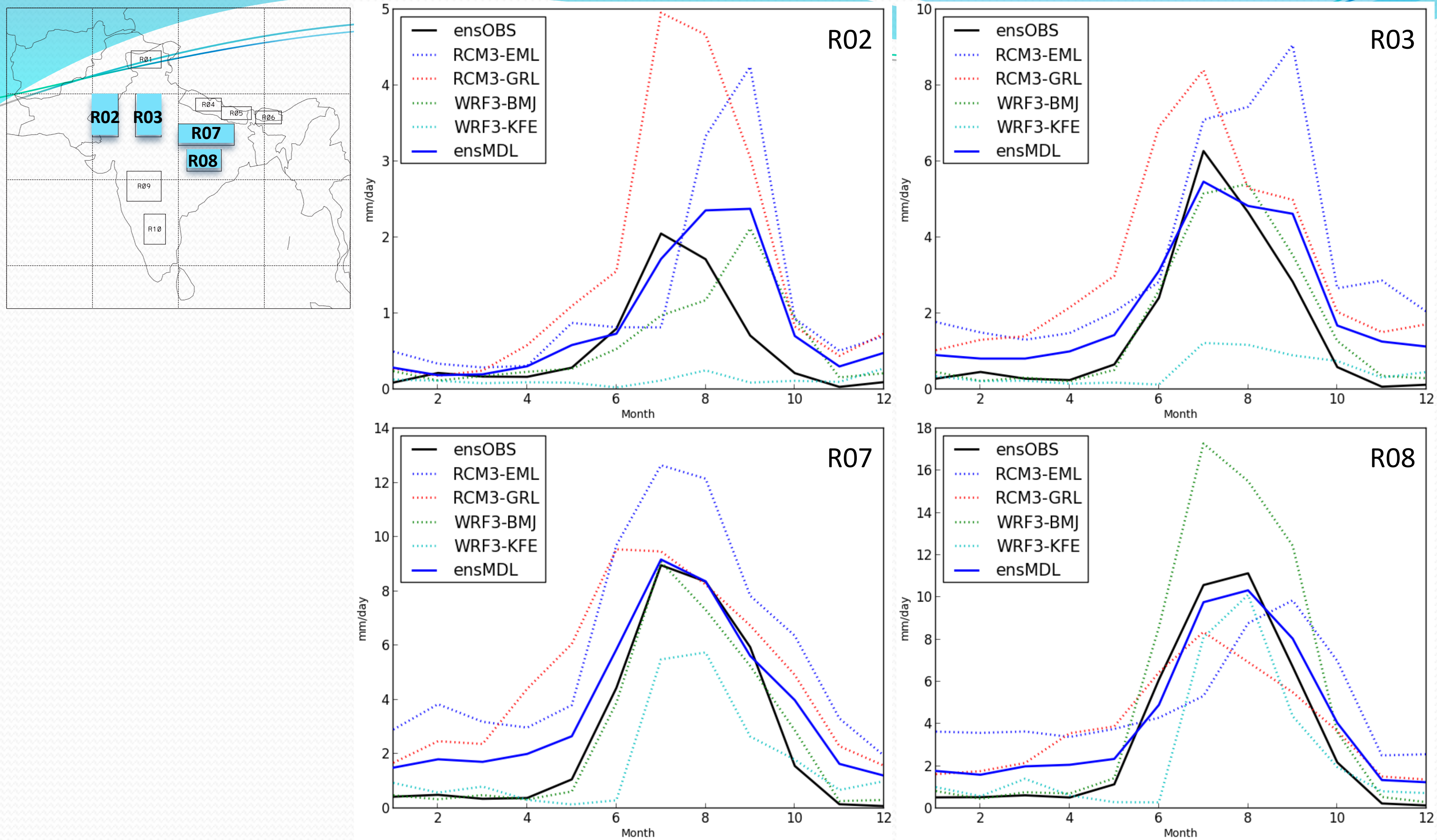
# Precipitation Annual Cycle: Northern Mountains



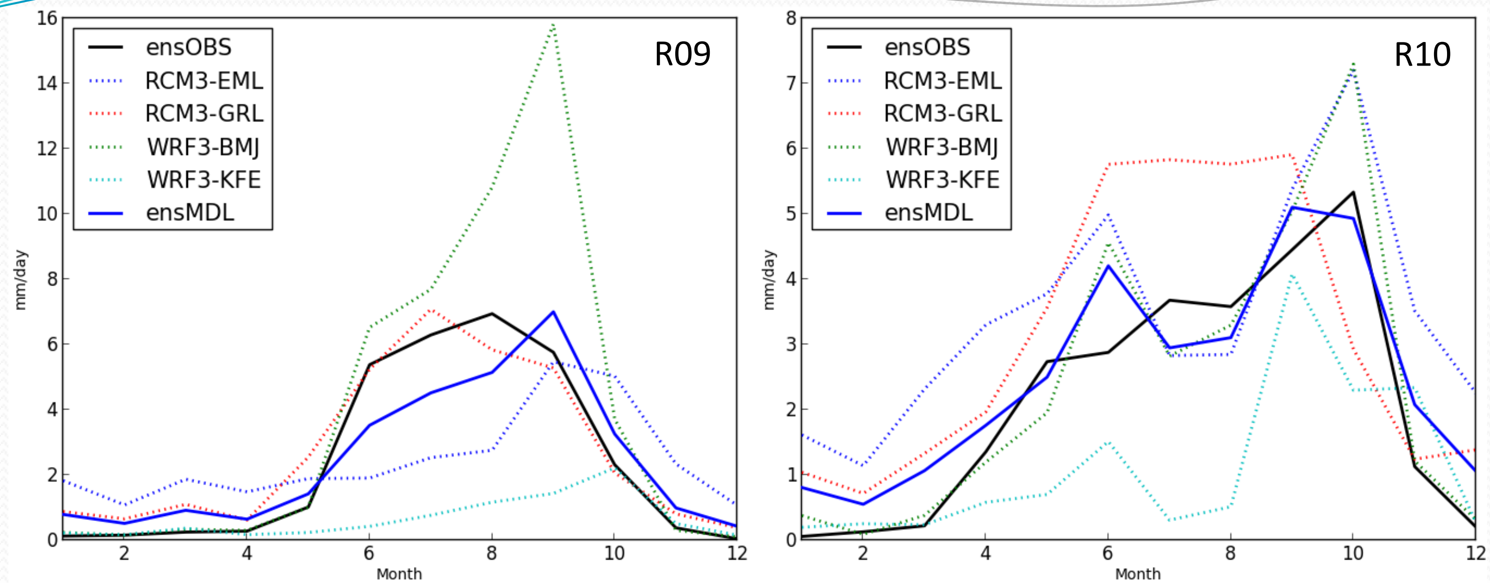

- No RCM nor the model ensemble show any skill in simulating precipitation annual cycle over the southern Hindu Kush regions.
- For other regions, a majority of models and the model ensemble generate correct phase of the observed annual cycle.
- The model ensemble performs better in the eastern Himalayas than in the western part.

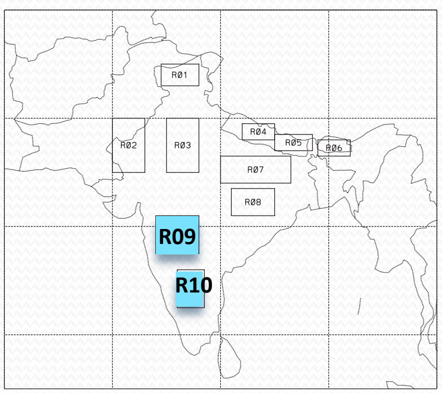
# Precipitation Annual Cycles: Northern Plains




- RCMs perform more consistently for these regions than for the northern mountain regions.
  - Inter-model variability is still large
  - Model ensemble performs well in the regions 3, 7, and 8.
- RCMs generally show better performance in wetter regions than drier regions.

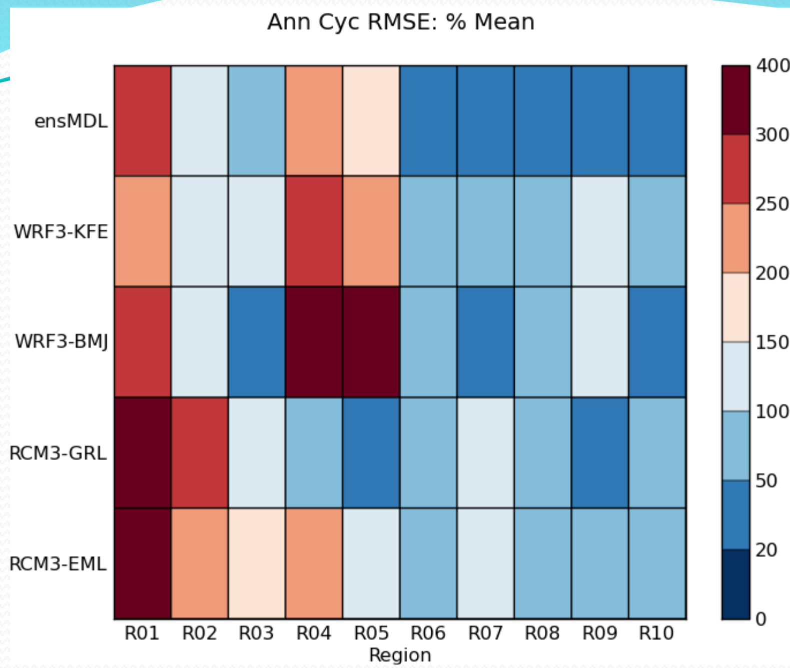
# Precipitation Annual Cycles: South-Central India

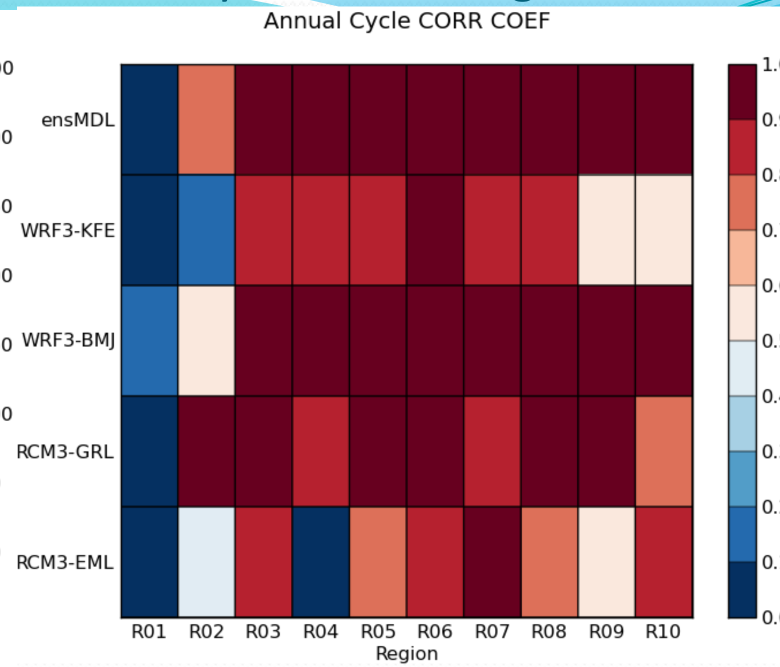


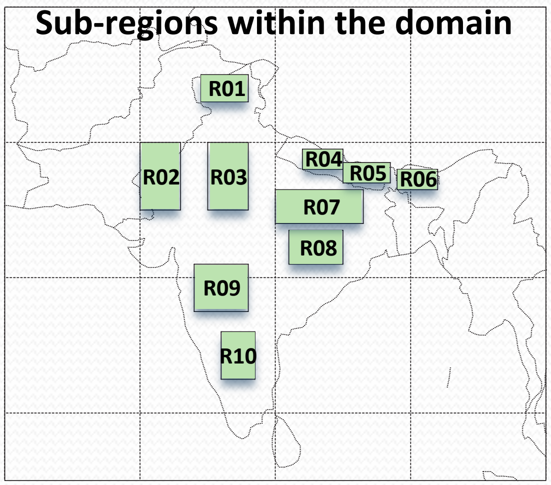

- Similarly as other regions, inter-RCM variations are large in these regions.
- The model ensemble is also similar to the observation ensemble in both the phase and magnitude of the annual cycle.

# Annual Cycle Evaluation – Summary for all subregions




- RCM performance varies widely among models and regions.
- Not very pronounced, but some regional structures exist:
  - No models/model ensemble performs well for the arid NW India/ SE Pakistan and the southern Hindu Kush Mountains (Kashmir).
  - In terms of RMSE, models generally performs better for the NE (R06, R07, R08) and the SW (R09, R10) parts of the Indian subcontinent than for the NW part of the region.
- Overall, it is difficult to stratify model performance.
  - The model ensemble shows better skill than individual models.



## Summary

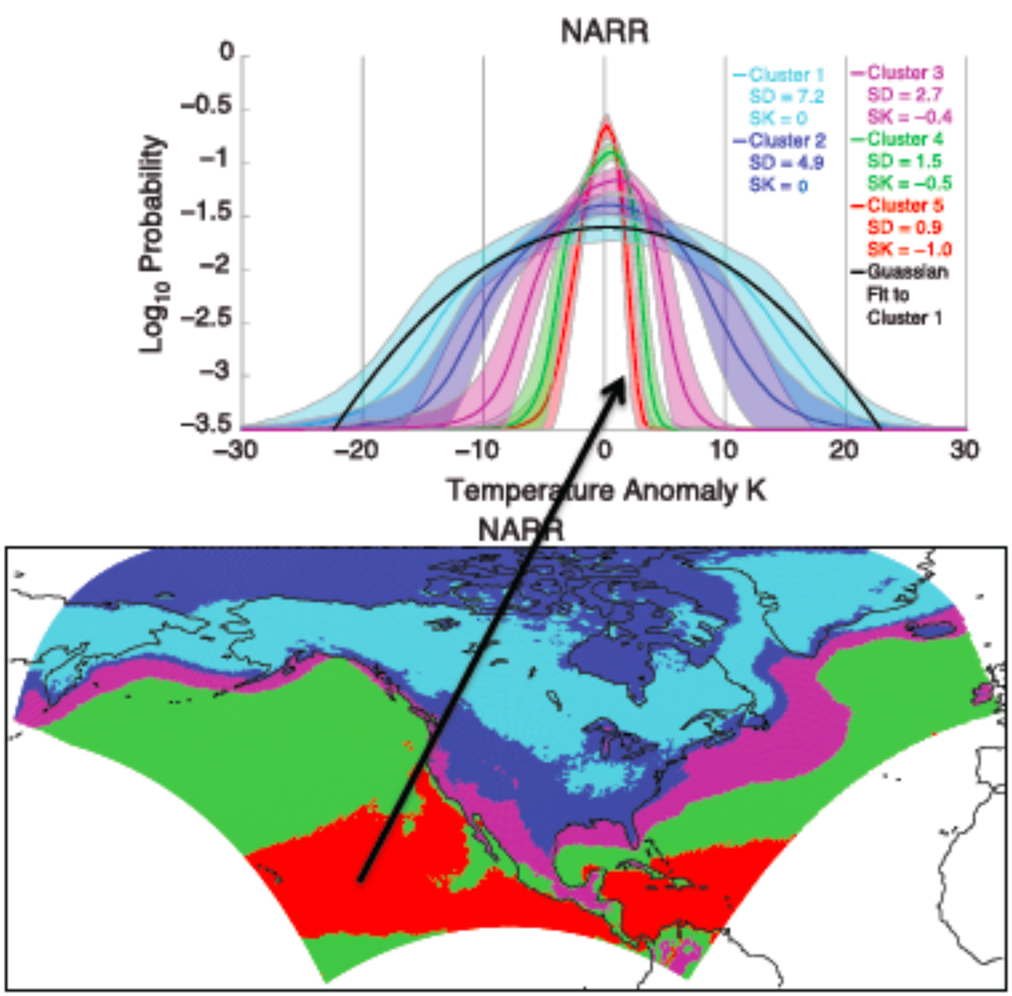
- RCMs' performance in simulating the annual-mean precipitation over the Indian subcontinent region has been examined using multiple observation datasets for January 1998 – December 2006 (9 years).
  - TRMM, GPCP, CRU, UDEL
  - 2 RCM3 runs using two different convection schemes (Emanuel, Grell)
  - 2 WRF3 runs using two different convection schemes (Betts-Miller Janjic, Kain-Frisch)
- Uncertainties related with observation data can be large
  - Spread among the observational data are large, especially in the mountainous regions, arid regions in the northwestern India/eastern Pakistan regions, and Tibetan Plateau.
- Model performance vary widely amongst regions:
  - May not be able to expect useful skill in simulating the annual-mean precipitation in the southern Hindu Kush Mountains, west-central Himalayas (Nepal), and the northwestern India/eastern Pakistan regions.
  - All RCMs also show substantial spread in all seasons for all regions. But the model ensemble may be useful for the plains and west coastal regions.
- The performance of RCM3 and WRF3 for the Indian subcontinent region is similar.

# Works to Follow – Climate science component

- Interannual variability of monsoon precipitation:
  - Onsets & terminations, and wet/dry monsoons (daily observation and model data).
- Radiation budget
  - Biases in OLR and the warm-season convective precipitation, TOA radiation budget
- Hydrology cycle
  - Precipitable water; Water vapor fluxes during the monsoon season, condensed hydrometeors.
- Extreme rainfall and heat episodes
  - Daily model data, TRMM; Local gauge data(?)
- Drought conditions
  - Identify and use the most suitable drought index for the region (soil moisture, precipitation)
- Surface fluxes
- Additional metrics

# Characterization of temperature variability using Cluster analysis

## An example of future metrics under development




- K-means clustering used to group the January surface temperature PDFs over the NARCCAP domain into 5 categories.
- (Top) Mean PDFs ofr each cluster. Shading indicates ±1 $\sigma$.
- Cluster assignments
  - The red curve is the average of all PDFs shaded in red on map, etc.
  - Cluster assignments primarily reflect variance, with some skewness
- Cluster analysis can provide a basis for identifying regions of common PDF morphology

***More details are available from***:
Loikith et al., 2013, Geophysical Research Letters.
Dr. Paul Loikith, JPL (Paul.C.Loikith@jpl.nasa.gov)

# Demonstration of Regional Climate Model Evaluation System (RCMES) for the CORDEX South Asia Domain

P. Ramirez[1], D. Waliser[1], C. Mattmann[1], J. Kim[2], H. Lee[1], P. Loikith[1], M. Bounstani[1], C. Goodale[1], A. Hart[1], M. Joyce[1], S. Khudikyan[1], J. Sanjay[3], M.V.S. Rama Rao[3], R. Krishnan[3], M. Mujumdar[3], S. Ingle[3]

[1]Jet Propulsion Laboratory/NASA

[2]University of California Los Angeles

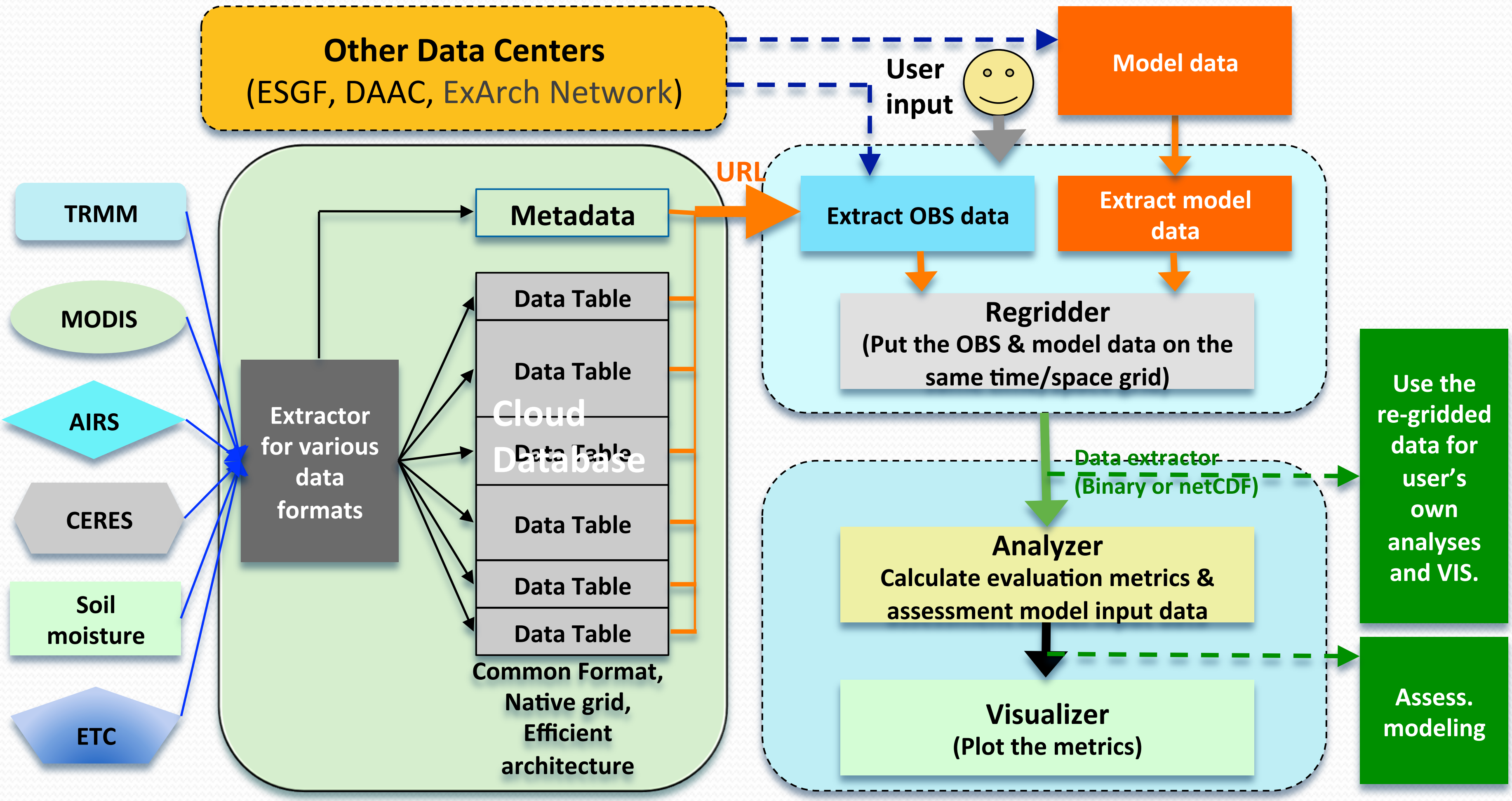
[3]Center for Climate Change Research, Indian Institute of Tropical Meteorology

# RCMES Goals

- Deliver a software suite for facilitating the evaluation of regional climate models
- Bring NASA observational datasets to aid in RCM evaluation
- Free scientists to focus on their science, not data management
- Focus and enhance the evaluation workflow
  - Download (Models, Observations), Preprocessing (Subsetting, Regridding), Evaluation (Statistical metrics), Visualization (Plot Results), Analysis and Interpretation (Scientist)
- Reduce the time and effort required to get to analysis and interpretation of results
- Collaborate with as many efforts as possible to improve adoption and reuse
- Support CORDEX community through collaborative science, software, and publications

# RCMES Architecture

**(http://rcmes.jpl.nasa.gov; Powered by *Apache Open Climate Workbench*)**




**Raw Data:**
Various sources, formats, Resolutions, Coverage

**RCMED**
(Regional Climate Model Evaluation Database)
A large scalable database to store data from variety of sources in a common format

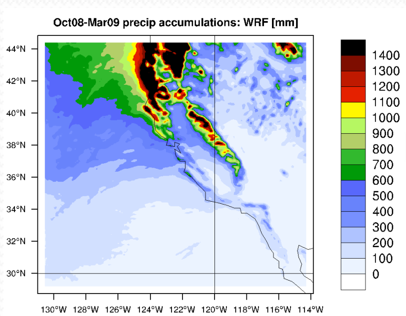
**RCMET**
(Regional Climate Model Evaluation Tool)
A library of code for extracting data from RCMED and model and for calculating evaluation metrics

# RCMES Workflow

| *Climate Model* | *Observational Data* | *Metrics* | *Results* |
|---|---|---|---|

## *Climate Model*

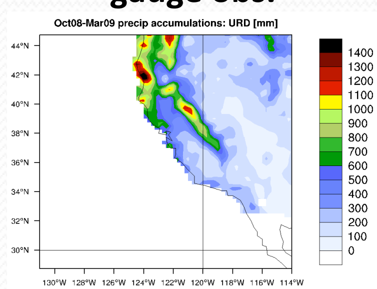
**WRF model**

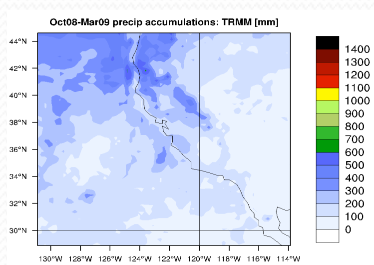



## *Observational Data*

(RCMED)

**URD gridded rain gauge obs.**




**TRMM observations**

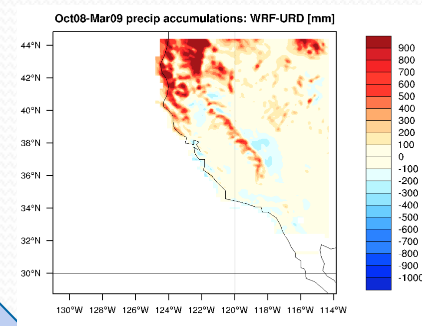
## *Metrics*

(RCMET)

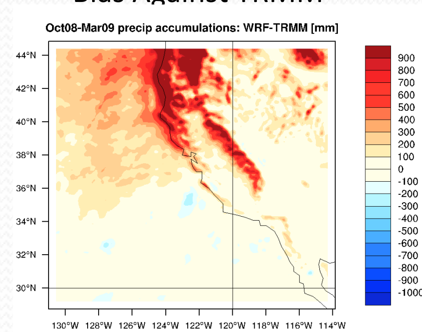
- **Spatial and temporal regridding for model to obs. Alignment**
- **Subsetting, masking, and averaging**
- **Metrics calculation**
- **Plotting/visualization**

## *Results*

Bias Against URD




Bias Against TRMM



# Demonstration Overview

- Command Line
  - Walk through configuration and running of toolkit
  - Show how to create all the plots you've seen
    - Contour Map, Portrait Diagram, Taylor Diagram, XY Line Plot (Time Series)
  - Demonstrate features and how to tailor an evaluation
  - Show the different metrics available through the toolkit
  - Create your own version by building on the code we have available
  - Version shown today available at https://svn.apache.org/repos/asf/incubator/climate/branches/rcmet-2.1.1
- Browser Based (Web) UI
  - Provide visual queues for model and observation overlap
  - Assume as many defaults as possible
  - History of evaluations to visit previous results and build new evaluations from
  - Shared configuration with command line to provide seamless integration
  - Screencast available for later viewing http://youtu.be/GZ2RindNmdI

# Live Demonstration

# Ongoing and Future Efforts

- Link to Earth System Grid Federation (ESGF)
  - Import model data directly from ESGF for processing
- Made-to-order System
  - Distribution of RCMES in a virtual machine packaged with python code and observation data as specified by users
- Data Processing and Metrics
  - Multiple re-gridding routines for irregularly distributed data
  - Improve handling of very large data sets by moving processing towards server
  - Develop metrics for evaluating PDF characteristics (e.g. variance, skewness)
  - Cluster analysis
- Reproducible Evaluations
  - Expand configuration to capture evaluation information and metadata
  - Allow import and export of configuration
  - Expand on history functionality to support building off of an existing evaluation
- Collaborative Development through Open Source
  - Leverage a well known open source software community and team expertise at Apache Software Foundation
  - Apache Open Climate Workbench effort built on meritocracy (invite collaborators and retain relationship with CORDEX communities and students)
  - Bring in other open source climate tools and link to other open source efforts across CORDEX and other climate communities (e.g. OpenClimateGIS, ESMPy)

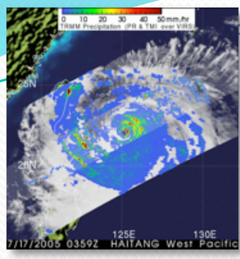
## Summary

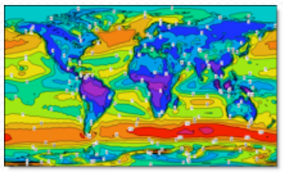
- The Regional Climate Model Evaluation System (RCMES) was developed to facilitate model evaluation via easy access to key observational datasets especially from satellite and remote sensing data sets
- RCMES is being used in evaluation studies for several climate experiments including CORDEX SA (seen today), NARCCAP, and other CORDEX domains
- RCMES includes a growing number of metrics and plots useful for model evaluation and is continuously being improved
- The RCMES team is building collaborations by attending and working with the climate community through workshop and conferences, jointly developing and publishing evaluations, and openly developing climate tools
- RCMES is working to make material used in publications easy to produce and have fellow scientist reproduce

# Contact Information

- RCMES Team
  - Website – http://rcmes.jpl.nasa.gov
  - Mailing List – rcmes-general@jpl.nasa.gov
  - Latest Demo – http://rcmes.jpl.nasa.gov/dev/ui/next/
- Apache Open Climate Workbench
  - Website – http://climate.incubator.apache.org
  - Mailing List – dev@climate.incubator.apache.org
  - Project Proposal – http://wiki.apache.org/incubator/ClimateProposal
  - Source Code – https://svn.apache.org/repos/asf/incubator/climate/trunk
  - Issue Tracking – https://issues.apache.org/jira/browse/climate
  - Documentation – https://cwiki.apache.org/confluence/display/CLIMATE/Home
  - Apache Software Foundation – http://www.apache.org/

# Backup Slides



**RCMED Datasets:** Satellite retrievals, Surface analysis, Reanalysis, Assimilations

- MODIS (satellite cloud fraction): [daily 2000 – 2010]
- TRMM (satellite precipitation): 3B42 [daily 1998– 2010]
- AIRS (satellite surface + T & q profiles) [daily 2002 – 2010]
- CERES and GEWEX-SRB radiation – Surface and Top of the atmosphere
- NCEP CPC Rain Gauge analysis (gridded precipitation): [daily 1948 – 2010]
- CRU TS 3.1: precipitation, Tavg, Tmax, Tmin [monthly means, 1901 – 2006]
- University of Delaware precipitation and temperature analysis
- Snow Water Equivalent over Sierra Nevada Mts [monthly 2000-2010]
- NASA MERRA Land Surface Assimilation & pressure-level data [daily, 1979-2011]
- ERA-Interim (reanalysis): [daily 1989 – 2010]
- AVISO sea-level height [1992-2010]
- (*In progress*) CloudSat atmospheric ice and liquid, Satellite-based snow (Himalayas), ISCCP cloud fraction, Fine-scale SST, etc.

**RCMET Metrics:**

- Bias (e.g. seasonal means or variance)
- RMS error (e.g. interannual variability)
- Anomaly Correlation (spatial patterns of variability)
- PDFs (likelihoods, extremes and their changes)
- Taylor Plots & Portrait Diagrams (overall model performance)
- Statistical Tests
- User-defined regions (e.g. water shed, desert, sea, political)

- Datasets and metrics are continuously updated.