

Training session on Regional Climate Model Evaluation System (RCMES)

Kyo Lee (huikyo.lee@jpl.nasa.gov)

(Jet Propulsion Laboratory, California Institute of Technology)

June 27, 2018

<http://rcmes.jpl.nasa.gov>

<http://climate.apache.org>

Acknowledgement

- My special thanks of gratitude to the workshop organizers.
- Regional Climate Model Evaluation System (RCMES) team
Duane Waliser (PI), Huikyo Lee (co-I), Alexander Goodman, Peter Gibson, Elias Massoud, Brian Wilson, Paul Loikith², and Antonio Monge³
¹JPL/Caltech, ²California State U. LA, ³Portland State U.
- **Virtual Information-Fabric Infrastructure (VIFI)** team led by Prof. William Tolone at U. of North Carolina, Charlotte
- Parallelized BCSD codes from Dr. TJ Vandal
- NASA Earth eXchange (NEX) team at NASA Ames center

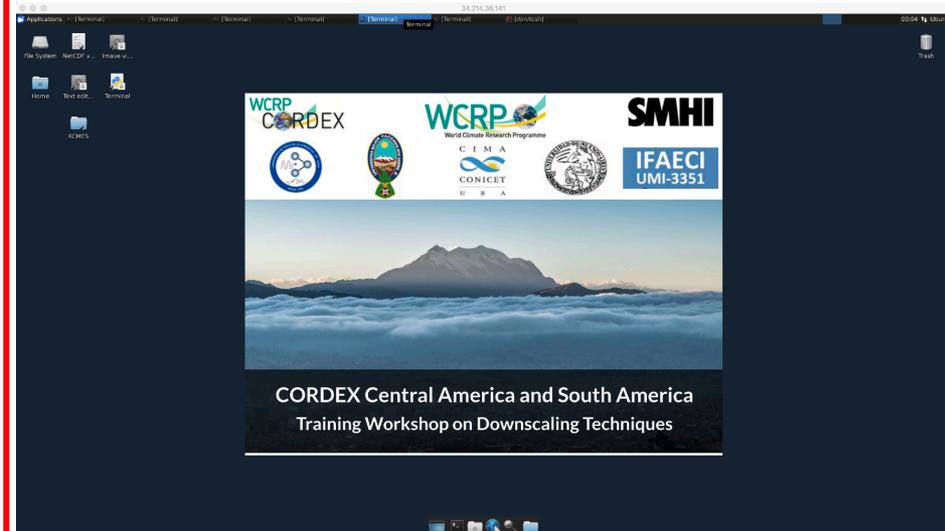
RCMES Training Outline (08:45-10:45)

Time	Agenda Item	Process/presentations/materials
08:45-09:00	Welcome and connect to Amazon Elastic Compute Cloud (EC2)	<ul style="list-style-type: none">• Check the IP address of the assigned server• Connect to the server using Microsoft Remote Desktop (or terminal software)
09:00-09:20	Activity #1 : Correct biases in CORDEX RCM simulations	<ul style="list-style-type: none">• Quantile-based bias correction of CORDEX CAM/SAM simulations using satellite-based precipitation observation data• Presentation: Systematic evaluation of CORDEX RCMs using RCMES
09:20-09:50	Activity #2 : Pointwise Statistical downscaling using RCMES	<ul style="list-style-type: none">• CMIP5 temperature and precipitation datasets for present and future climate• Compare the IPCC climate change scenarios (RCP 4.5 vs. RCP 8.5)
09:50-10:20	Activity #3 : Download and visualize the NEX-GDDP data	<ul style="list-style-type: none">• NASA Earth Exchange Globally Daily Downscaled Projections (NEX-GDDP) in Amazon Simple Storage Service (S3)• Presentation: Toward the future of Big climate data analysis in the cloud
10:20-10:45	Activity #4 : Analyze the bias corrected RCM output	<ul style="list-style-type: none">• Presentation: What powers RCMES and how to get involved with development

Two different ways to connect to the virtual Linux machine on Amazon Web Service

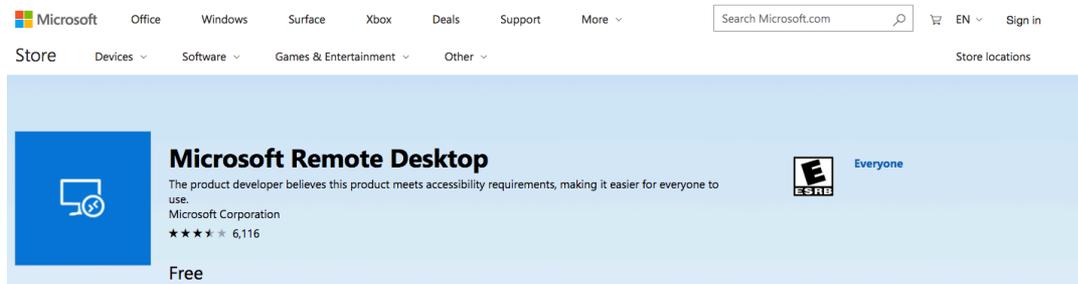
- SSH connection using your terminal application
- Prerequisite software
 - terminal: putty, xshell, xterm)
 - X Server: Xming, XQuartz)
 - NetCDF/HDF viewer: Panoply
 - (Optional) sftp client: xftp, FileZilla
- `ssh -Y ubuntu@xx.xxx.xx.xxx`
(your machine's IP address)
- password: **cordex**

- Remote desktop



Prerequisite software to run remote desktop

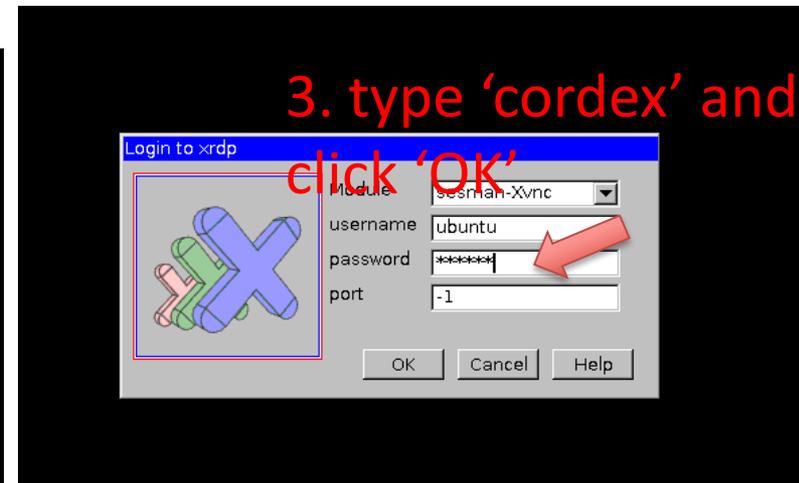
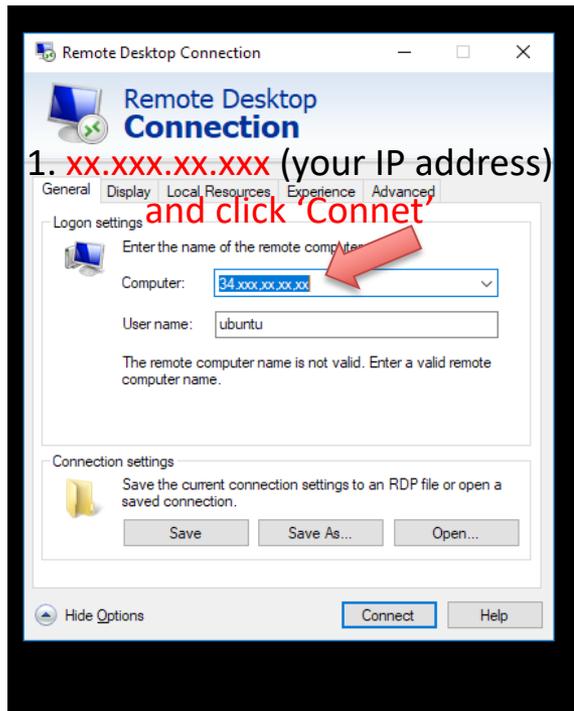
- ~~Linux based system~~
- Windows laptops: Microsoft Remote Desktop



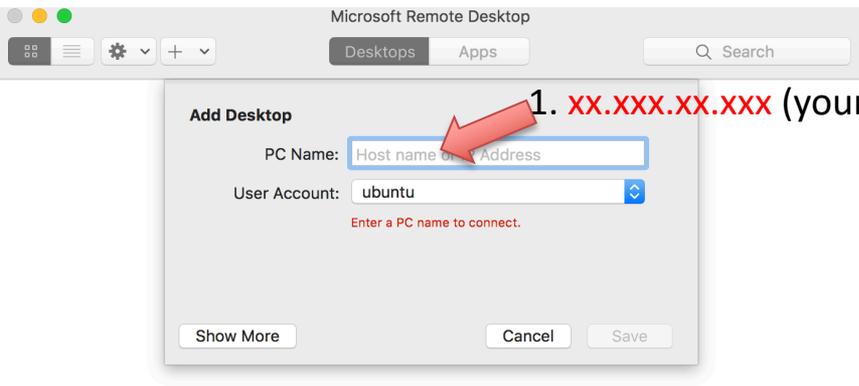
- Macbooks: Microsoft Remote Desktop **10**
(do not use version 8)



Set up your remote desktop (Windows)

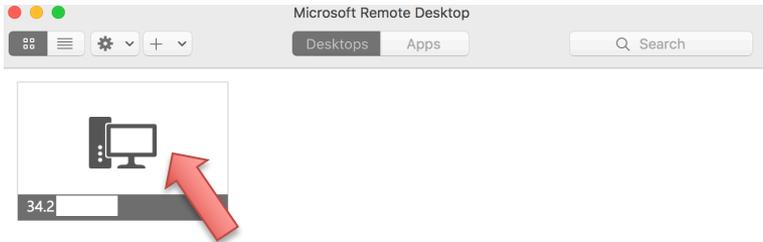


Set up your remote desktop (Mac)



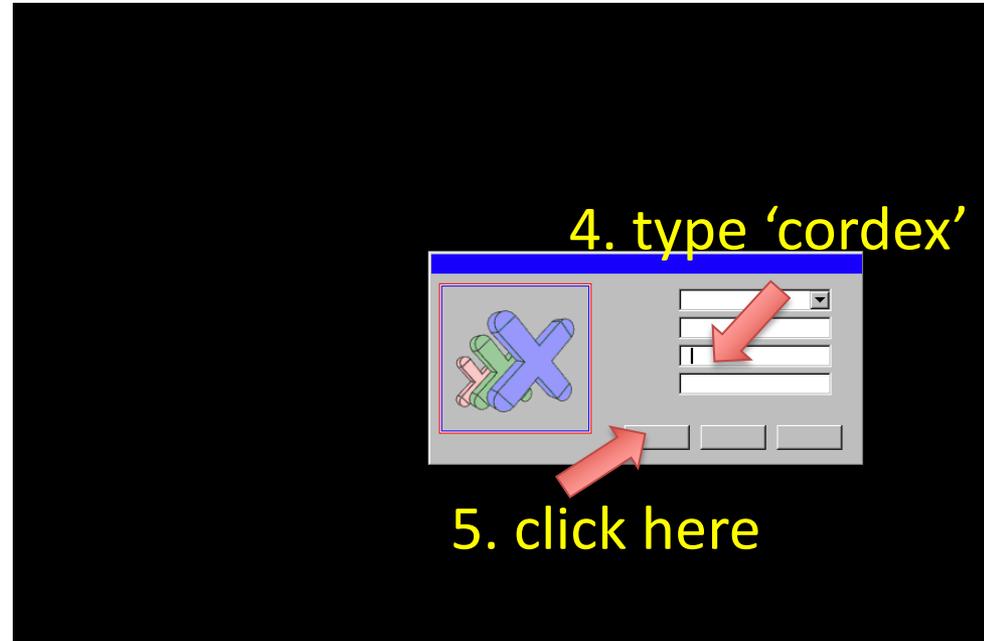
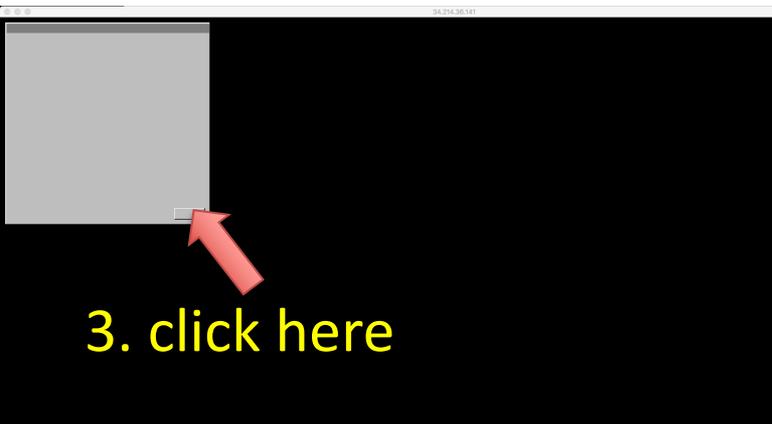
Add your first desktop connection to get started.

Add desktop



2. Double click!!!

Bugs in the Mac version



- **Activity #1**
: Correct biases in CORDEX RCM simulations
- Activity #2
: Pointwise Statistical downscaling using RCMES
- Activity #3
: Download and visualize the NEX-GDDP data
- Activity #4
: Analyze the bias corrected RCM output

Running the bias correction script

(courtesy of Dr. TJ Vandal at NASA Ames, <https://github.com/tjvandal/bcsd-python>)

1. Open Terminal and type `cd RCMES`

(Two options: please choose one of them)

2-1. To correct biases in a RegCM4 simulation for the CORDEX Central America,

```
python CORDEX_CAM-SAM_TRMM_BC_example.py CORDEX_CAM.yaml
```

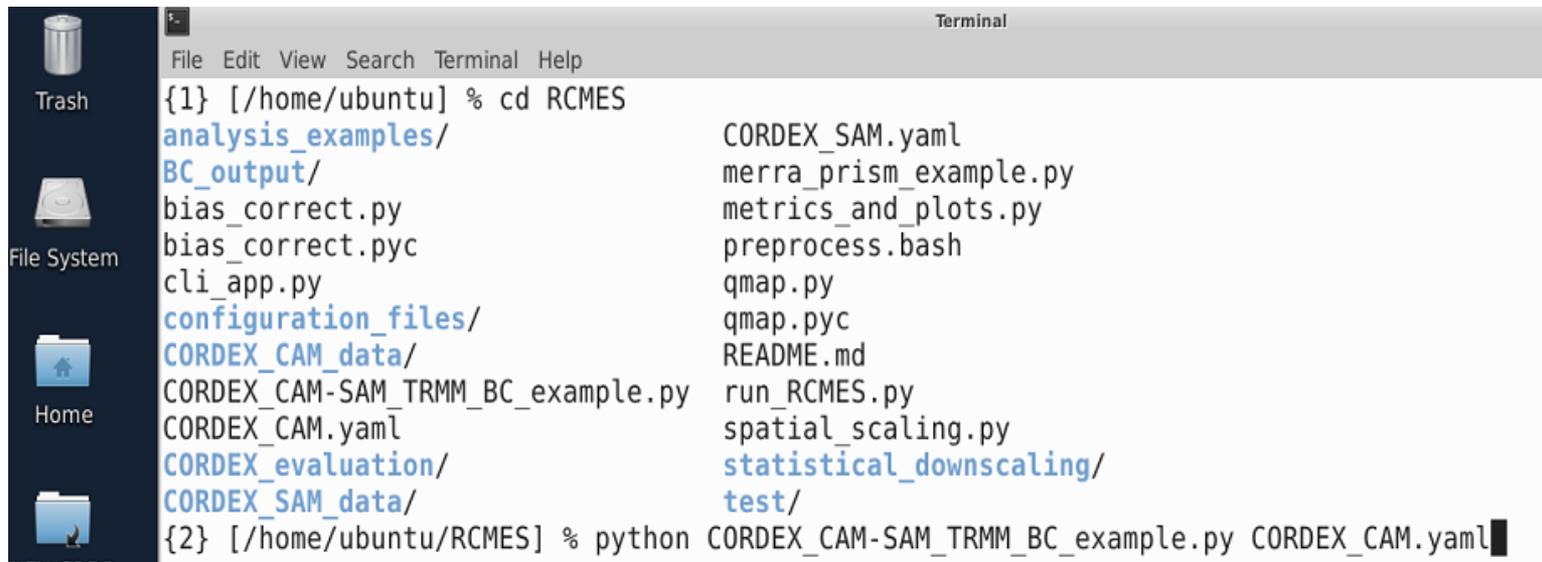
Python script

Configuration file

2-2. To correct biases in a RCA4 simulation for the CORDEX South America,

```
python CORDEX_CAM-SAM_TRMM_BC_example.py CORDEX_SAM.yaml
```

(Running 2-1 or 2-2 uses 15 CPUs and takes about 45-60 minutes.)



The screenshot shows a terminal window with a file manager sidebar on the left. The sidebar includes icons for Trash, File System, Home, and another folder. The terminal window title is "Terminal" and it has a menu bar with "File", "Edit", "View", "Search", "Terminal", and "Help". The terminal output shows the following commands and directory listing:

```
{1} [/home/ubuntu] % cd RCMES
analysis_examples/
BC_output/
bias_correct.py
bias_correct.pyc
cli_app.py
configuration_files/
CORDEX_CAM_data/
CORDEX_CAM-SAM_TRMM_BC_example.py
CORDEX_CAM.yaml
CORDEX_evaluation/
CORDEX_SAM_data/
CORDEX_SAM.yaml
merra_prism_example.py
metrics_and_plots.py
preprocess.bash
qmap.py
qmap.pyc
README.md
run_RCMES.py
spatial_scaling.py
statistical_downscaling/
test/
{2} [/home/ubuntu/RCMES] % python CORDEX_CAM-SAM_TRMM_BC_example.py CORDEX_CAM.yaml
```

Bias Correction of CORDEX simulations

- The two CORDEX RCM simulations have high spatial resolution (~44 km) relative to CMIP GCMs.
- ~~BCSD~~ => BC : spatial disaggregation (SD) may not be necessary thanks to the high resolution of CORDEX simulations.

Quantile mapping to correct simulated precipitation using TRMM observations (1)

- Inside the configuration file (CORDEX_CAM.yaml)

fobserved: TRMM_regridded_RegCM4-3_CAM-44.nc

observed_varname: TRMM_daily_pr

fmodeled_present: pr_CAM-44_MPI-M-MPI-ESM-MR_historical_r1i1p1_ICTP-RegCM4-3_v4_day_19980101-20131231.nc

fmodeled_future: pr_CAM-44_MPI-M-MPI-ESM-MR_rcp85_r1i1p1_ICTP-RegCM4-3_v4_day_20830101-20991231.nc

modeled_varname: pr

(Observation)

Read TRMM_daily_pr from
TRMM_regridded_RegCM4-3_CAM-44.nc

(Simulation for the present climate)

Read pr from pr_CAM-
44_***_19980101-20131231.nc

(Simulation for the future climate)

Read pr from pr_CAM-
44_***_20830101-20991231.nc

Quantile mapping to correct simulated precipitation using TRMM observations (2)

- At each RCM grid point, biases in simulated precipitation are corrected for each quantile (0.5-99.5%) by comparing two cumulative distributions from TRMM and the RCM (± 15 days).

(Observation)

TRMM_daily_pr for 19980101-20131231

(Simulation for the future climate)

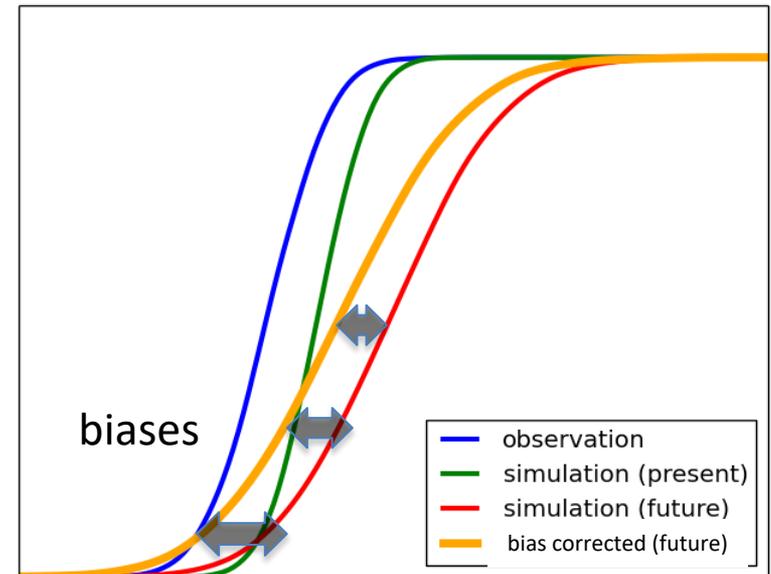
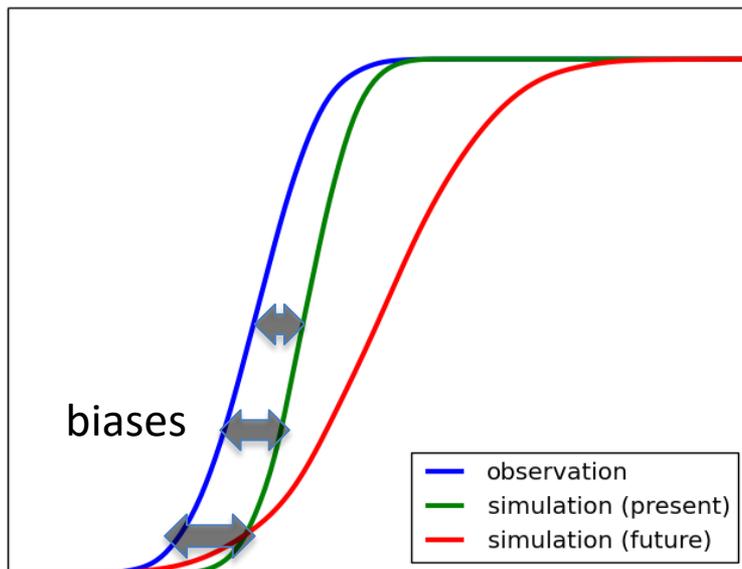
pr_CAM-44_***_20830101-20991231.nc

(Simulation for the present climate)

pr_CAM-44_***_19980101-20131231.nc

(Bias corrected future simulation)

BC_pr_CAM-44_***_20830101-20991231.nc



- Activity #1
: Correct biases in CORDEX RCM simulations
- **Activity #2**
: **Pointwise Statistical downscaling using RCMES**
- Activity #3
: Download and visualize the NEX-GDDP data
- Activity #4
: Analyze the bias corrected RCM output

The Regional Climate Model Evaluation System (RCMES, <https://rcmes.jpl.nasa.gov>)

- Lee et al. (2018) in GMDD under review.
- Python-based open source software powered by the Apache Open Climate Workbench (OCW)
- Main components
 - 1) Database of observations
 - 2) Toolkit for facilitating systematic evaluation of CORDEX RCMs using satellite observations
 - 3) Statistical downscaling of coarse-resolution GCM output
 - 4) Stand-alone scripts for data processing and visualization based on OCW

Regional Climate Model Evaluation System

High-Level Architecture

Observation for Evaluation

obs4MIPs
Over 30 Satellite
variables on ESGF

Other
Data
Centers

Local
Disk

RCMES
Observational
Database
(e.g., TRMM, CRU,
UDEL)

User Input

Spatial Boundaries
Temporal Boundaries & Resolution

Extract OBS data

Extract model data

Regridding
Put the OBS & model data on the same spatial grid

Metrics Calculator
(Calculate evaluation metrics)

Visualizer
(Plot the metrics)

Model data for Evaluation

ESGF
Several Models
(e.g. CMIP5, CORDEX)

Other
Data
Centers

Local
Disk

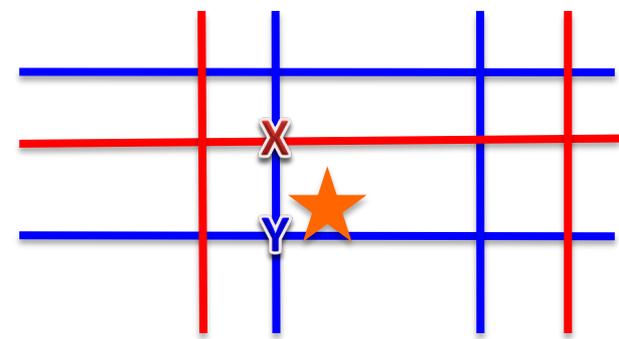
Data extractor to netCDF

Use the re-gridded data for
user's own analyses and
visualization.

RCMES captures the
entire workflow.

Another user can reproduce the same
results using the captured workflow.

Statistical downscaling using RCMES



- To statistically downscale CMIP5 variables at a specific location (star marker), RCMES uses statistical relationship between the nearest model grid point data (X) and observation grid point data (Y)

: simultaneous correction of both bias and collocation

$$Y = f(X)$$

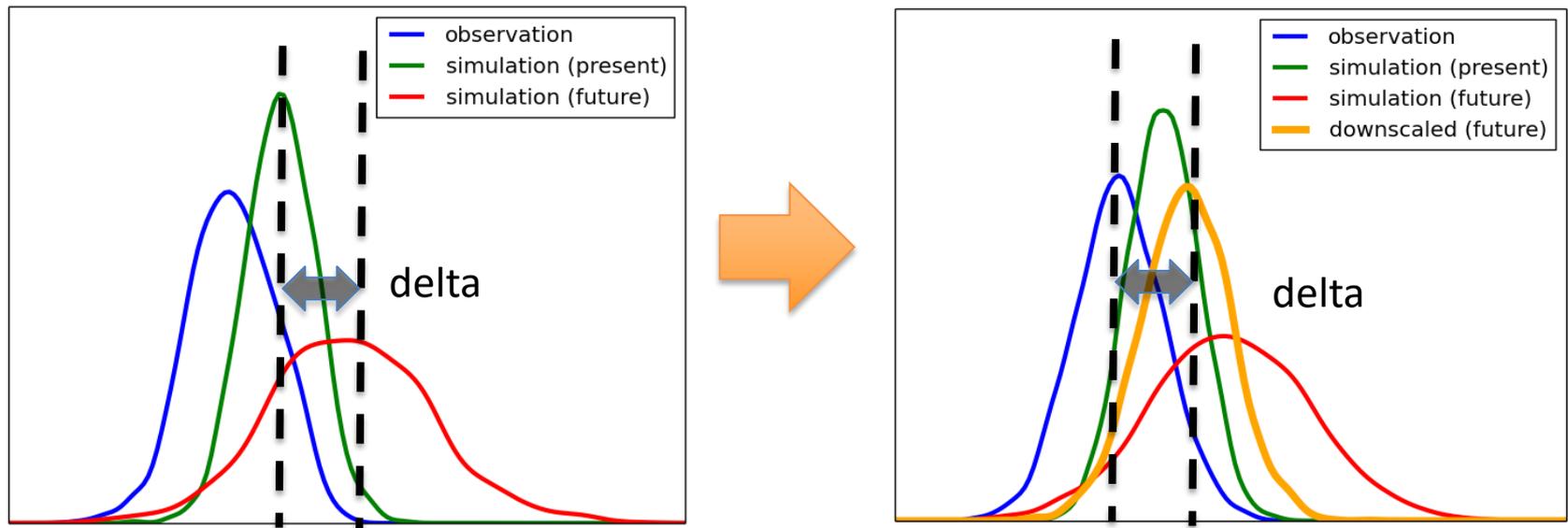
- Four different methods for model calibration (Stoner et al., 2013)
 - Delta method (addition)
 - Delta method (bias correction)
 - Quantile mapping
 - Asynchronous linear regression
- The observational datasets in RCMES database can be used to determine the observation-model relationship.

Delta method (Delta addition)

Y_0 : present observation, X_0 : present simulation, X_1 : future simulation

$$Y_1 = Y_0 + \bar{X}_1 - \bar{X}_0$$

- (future climate) = (present observation) + (mean difference between X_0 and X_1)

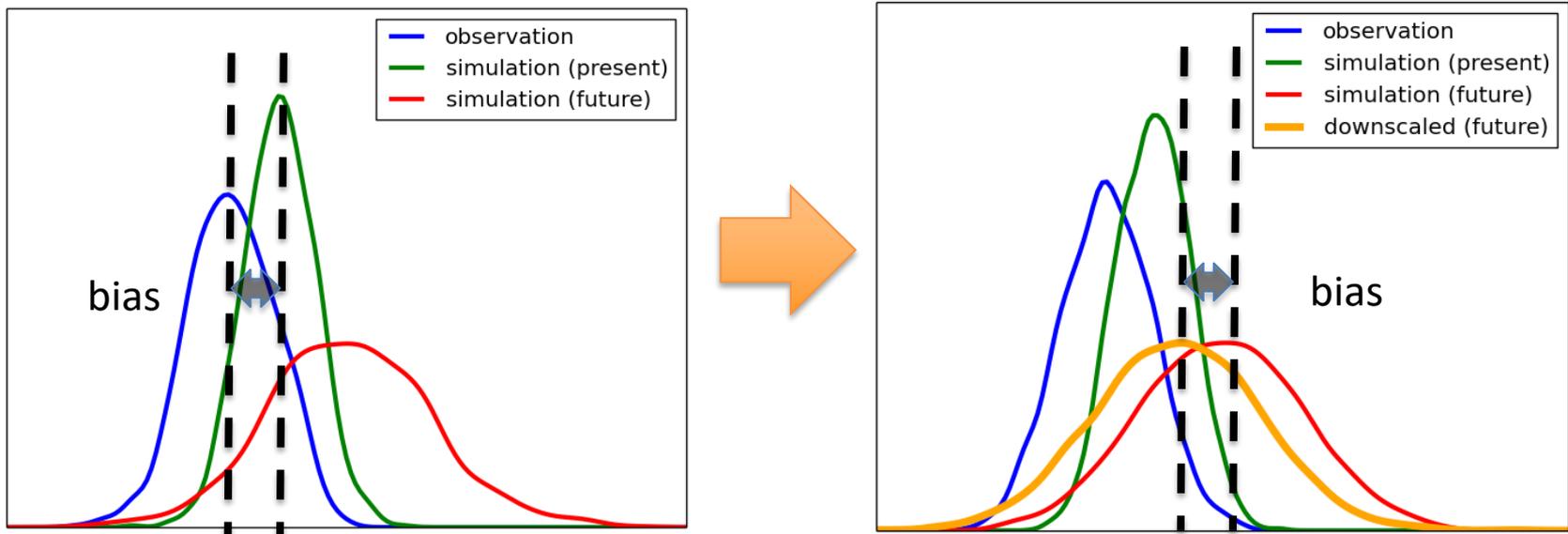


Delta method (Bias correction)

Y_0 : present observation, X_0 : present simulation, X_1 : future simulation

$$Y_1 = X_1 + \bar{Y}_0 - \bar{X}_0$$

- (future climate) = (future simulation) + (mean bias)

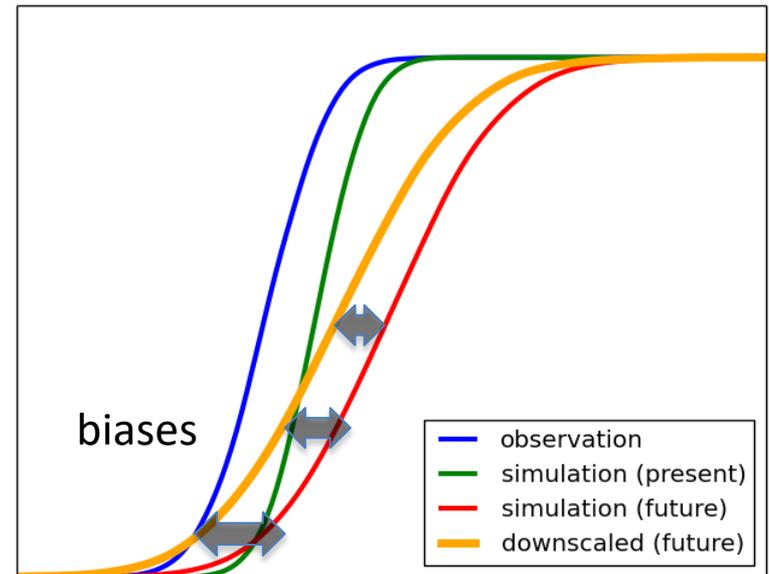
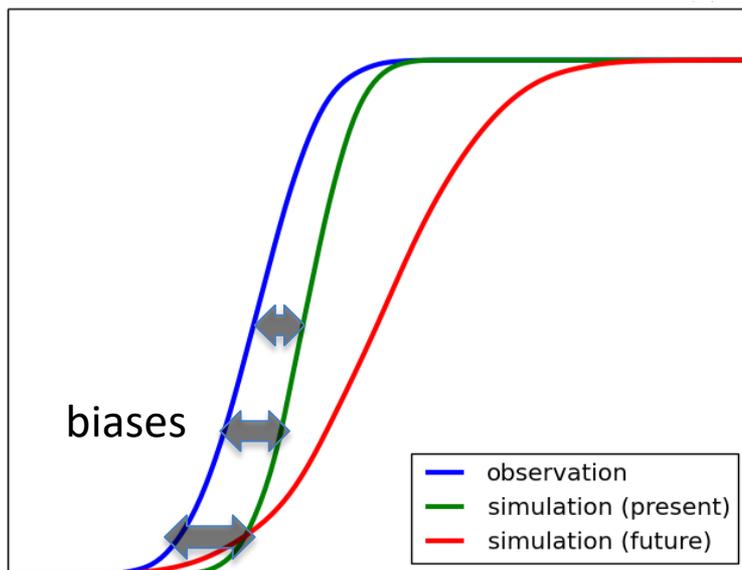


Quantile mapping

Y_0 : present observation, X_0 : present simulation, X_1 : future simulation

$Y_1 = f(X_1)$ where f is bias correction function for each quantile ($Y_0 = f(X_0)$).

- (future climate) = (bias corrected future simulation)
- Bias is corrected for each quantile.



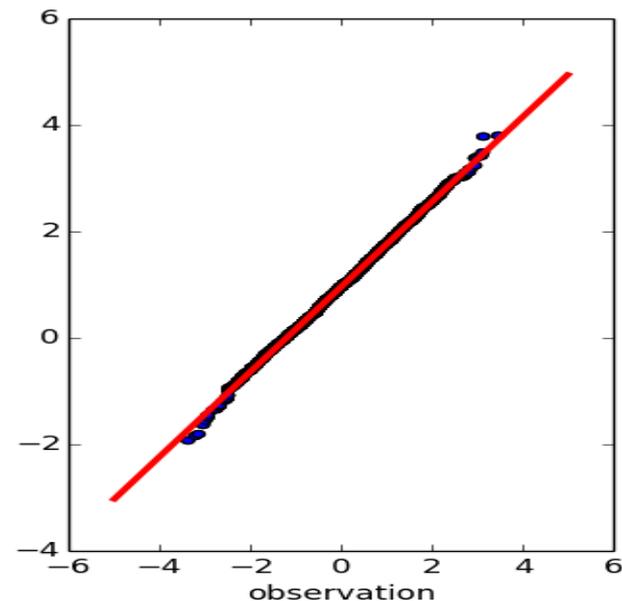
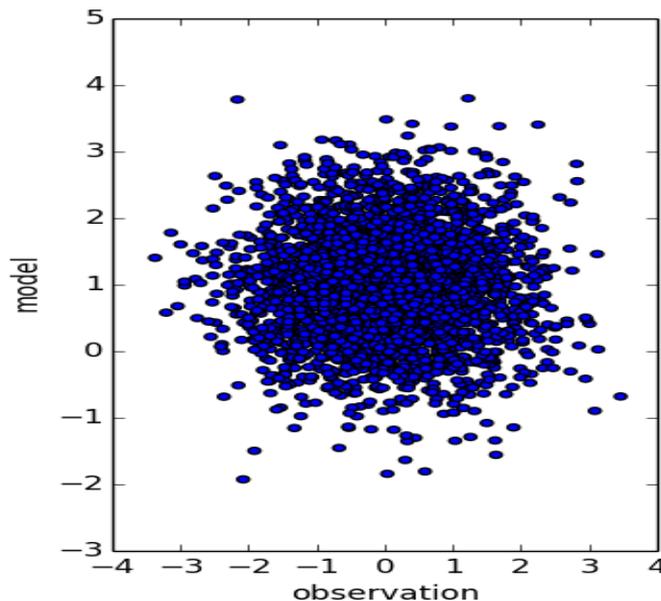
Asynchronous linear regression

Y_0 : present observation, X_0 : present simulation, X_1 : future simulation

Y'_0, X'_0, X'_1 : sorted in ascending order

$Y'_1 = aX'_1 + b$ where $Y'_0 = aX'_0 + b$. a and b are the slope and intercept for the least square regression line.

- The linear relationship between observation and present simulation is determined after sorting them in ascending order.



Statistical Downscaling using RCMES

1. Open Terminal and `cd RCMES/statistical_downscaling/`

2. To run the statistical downscaling script, type

```
python run_statistical_downscaling.py MPI_tas_JJA.yaml
```

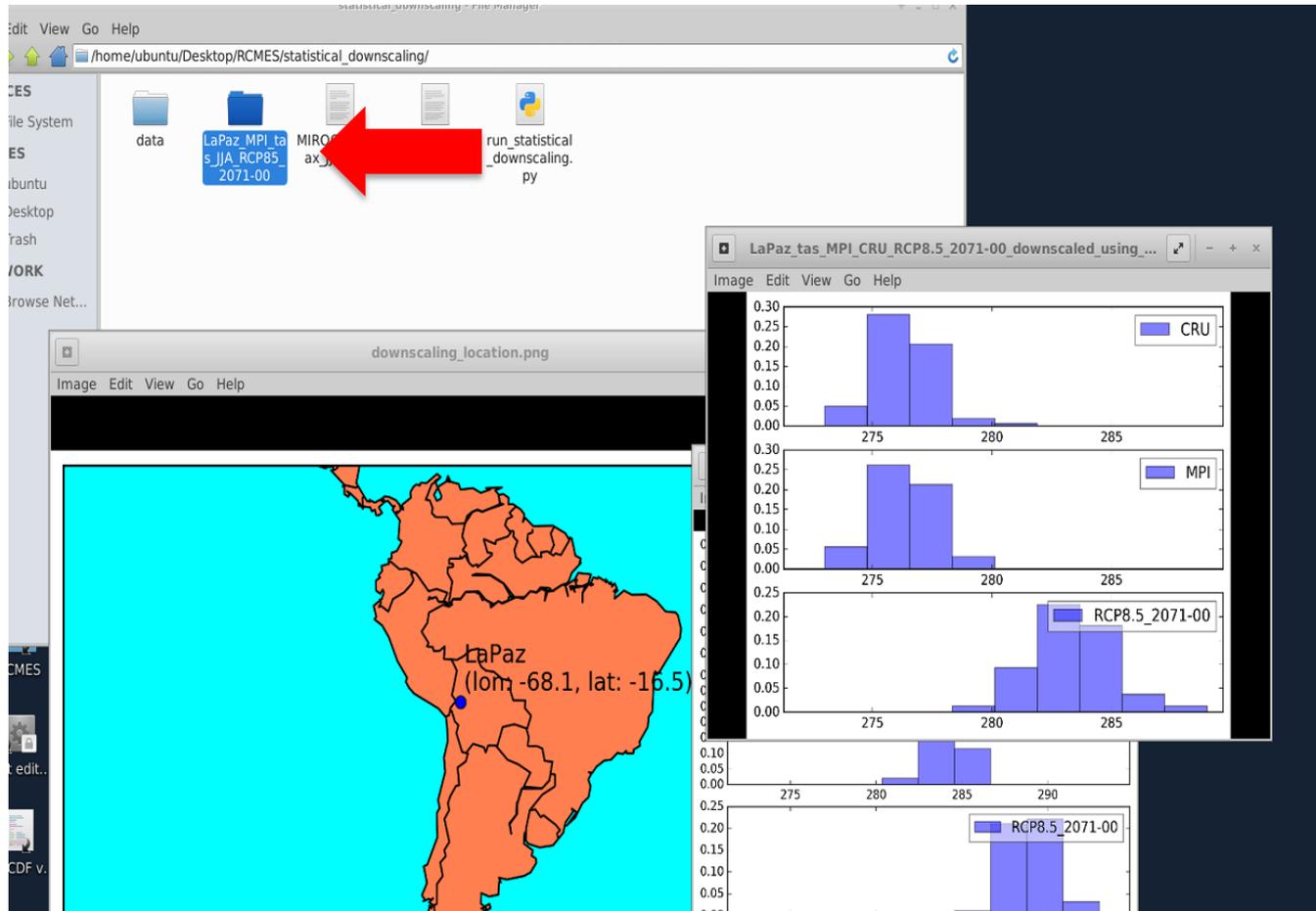
Python script

Configuration file

```
{13} [/home/ubuntu] % cd RCMES/statistical_downscaling/  
data/ LaPaz_MPI_tas_JJA_RCP85_2071-00/ MIROC5_tasmax_JJA.yaml MPI_tas_JJA.yaml run_statistical_downscaling.py  
{14} [/home/ubuntu/RCMES/statistical_downscaling] % python run_statistical_downscaling.py MPI_tas_JJA.yaml  
Reading the configuration file MPI_tas_JJA.yaml  
Processing CRU data  
Loading ./data/tas_cru_monthly_1981-2010.nc into an OCW Dataset Object  
CRU values shape: (times, lats, lons) - (360, 360, 720)  
  
Loading ./data/tas_Amon_MPI_decadal1980_198101-201012.nc into an OCW Dataset Object  
MPI values shape: (times, lats, lons) - (360, 96, 192)  
  
RCP8.5_2071-00:MPI values shape: (times, lats, lons) - (360, 96, 192)  
  
Temporal subsetting for the selected month(s)  
Spatial aggregation of observational data near latitude 10.75 and longitude 106.67  
Creating a statistical downscaling object  
asynchronous_regression: Downscaling model output  
Plotting results  
Generating spreadsheet  
{15} [/home/ubuntu/RCMES/statistical_downscaling] % █
```

View the statistically downscaled tas results

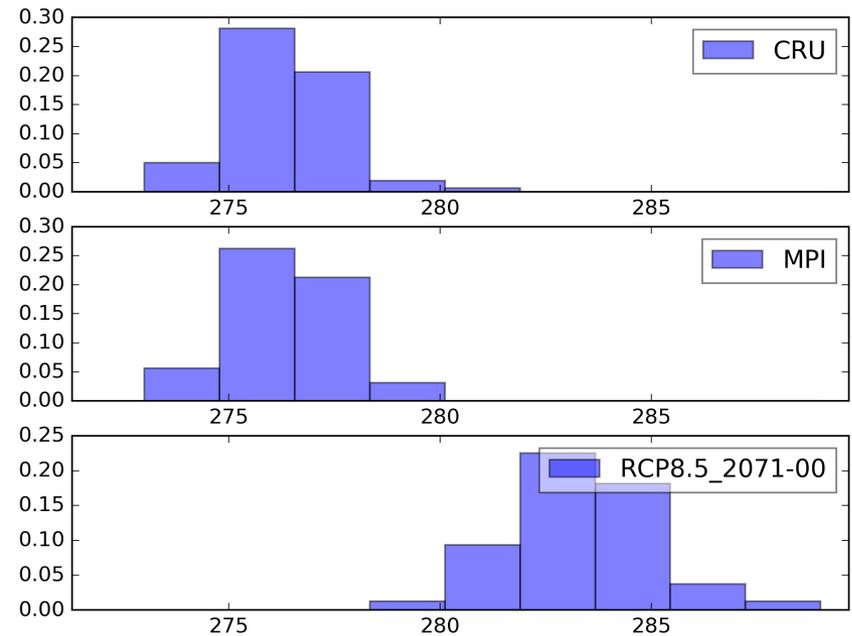
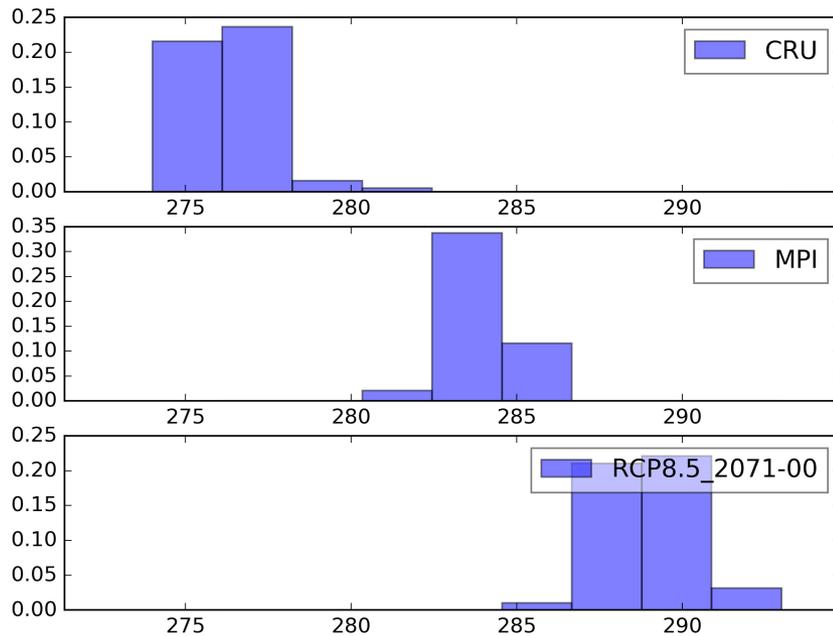
- The results can be found in statistical_downscaling/LaPaz_MPI_tas_JJA_RCP85_2071-00 folder



Quantile mapping of the near-surface air temperature for La Paz in JJA

Original model output

Statistically downscaled model output



Run another example: taxmax in Buenos Aires

```
python run_statistical_downscaling.py MIROC5_taxmax_JJA.yaml
```

Python script

Configuration file

Make your own example by editing the yaml file

```
case_name: BuenosAires_MIROC5_tasmax_JJA_RCP85_2071-00
```

Output folder name

```
downscaling_option: 3
```

```
location:
```

```
  name: BuenosAires
```

```
  grid_lat: -34.60
```

Search Google with the keyword 'latitude and longitude of XXX'

```
  grid_lon: -58.38
```

```
month_index: !!python/tuple [6,7,8]
```

Season

```
reference:
```

```
  data_source: local
```

```
  data_name: CRUs
```

```
  path: ./data/tasmax_cru_monthly_1981-2010.nc
```

```
  variable: tasmax
```

```
model:
```

```
  data_name: MIROC5
```

```
  variable: tasmax
```

```
present:
```

```
  path: ./data/tasmax_Amon_MIROC5_decadal1980_198101-201012.nc
```

```
future:
```

```
  scenario_name: RCP8.5_2071-00
```

```
  path: ./data/tasmax_Amon_MIROC5_rcp85_207101-210012.nc
```

(Options)

1. **tas**, **tasmin**, and **tasmax**
2. **RCP 4.5** and **8.5**
3. **(2041-2070)** and **(2071-2100)**

- Activity #1
: Correct biases in CORDEX RCM simulations
- Activity #2
: Pointwise Statistical downscaling using RCMES
- **Activity #3**
: **Download and visualize the NEX-GDDP data**
- Activity #4
: Analyze the bias corrected RCM output

NASA's Earth Exchange

(NEX, <https://nex.nasa.gov>)



- NEX is a platform for scientific collaboration, knowledge sharing and research for the Earth science community.
- The new project, Open NEX, is aimed at making a number of important datasets more accessible.

NASA Earth Exchange
Global Daily Downscaled Projections (NEX-GDDP)

CMIP5 historical and RCP
4.5/8.5 simulations
(from 21 models, 1950-2100)

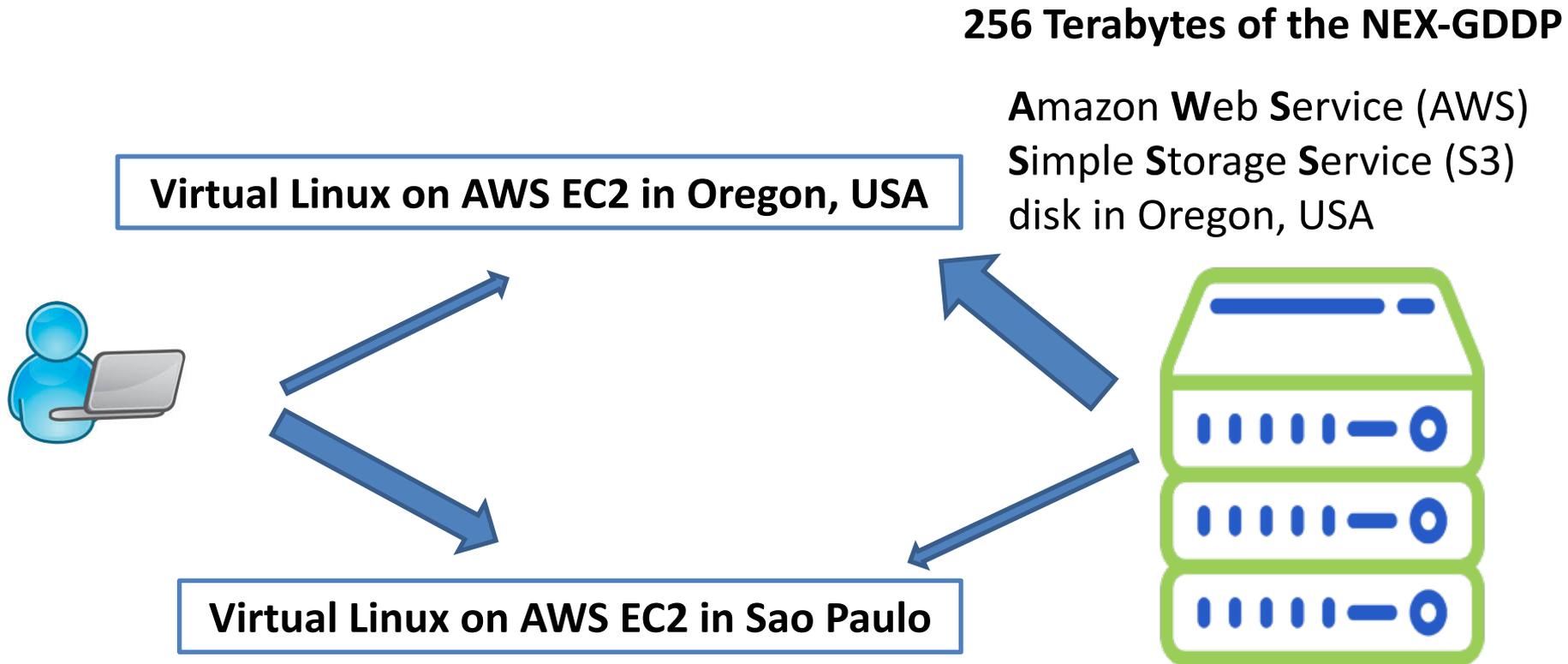
Bias-Correction Spatial
Disaggregation (BCSD)



Global Meteorological
Forcing Dataset
(observation, 1950-2005)

NEX-GDDP
: tasmax, tasmin, precipitation

Access to the statistically downscaled NEX-GDDP



- The NEX S3 is mounted in your linux EC2.
- Open terminal and type `df -h`

What are inside s3://nasanex?

```
[/home/ubuntu] % aws s3 ls s3://nasanex
```

```
PRE AVHRR/
```

```
PRE CMIP5/
```

```
PRE LOCA/
```

```
PRE Landsat/
```

```
PRE MAIAC/
```

```
PRE MODIS/
```

```
PRE NAIP/
```

```
PRE NEX-DCP30/
```

```
PRE NEX-GDDP/
```

List, download and visualize NEX-GDDP

1. Open terminal and `cd NEX-GDDP`

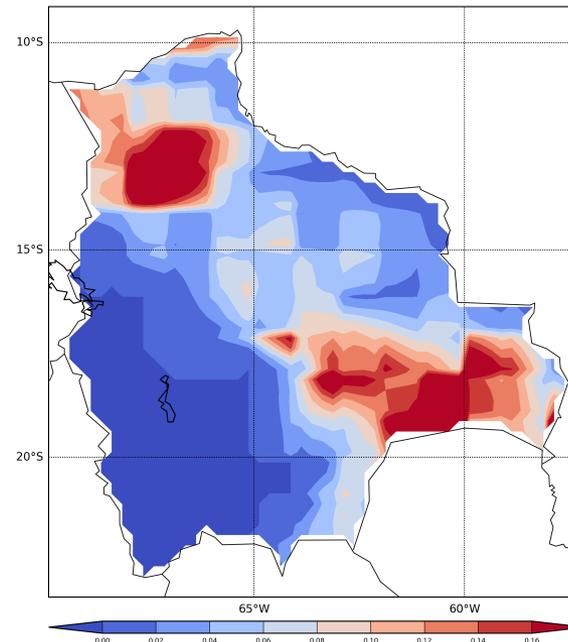
2. `./list`

3. `./download` **How fast!**

This script is an example of **Open Climate Workbench**, an open-source Python library that comprise **RCMES**.

4. `python plot_NEX-GDDP_example.py`

Statistically downscaled precipitation from NorESM1-M model for June 2100



Apache Open Climate Workbench (OCW)

<https://climate.apache.org/>

Apache Open Climate Workbench

Downloads

Development ▾

Documentation ▾

Community ▾

ASF ▾



Apache Open Climate Workbench

Apache Open Climate Workbench is an effort to develop software that performs climate model evaluation using model outputs from a variety of different sources the [Earth System Grid Federation](#), the [Coordinated Regional Climate Downscaling Experiment](#), the [U.S. National Climate Assessment](#) and the [North American Regional Climate Change Assessment Program](#) and temporal/spatial scales with remote sensing data from [NASA](#), [NOAA](#) and other agencies. The toolkit includes capabilities for rebinning, metrics computation and visualization.

Apache Open Climate Workbench 1.0.0 Released

September 24, 2015

The Apache Open Climate Workbench team is pleased to announce the 1.0.0 release! This release addresses no less than 52 issues, bugs, and improvements. For a full breakdown of the work packaged into this release please see the [release report](#).

Some important features this release packs include statistical downscaling capabilities such as Delta Method, Quantile Mapping and Quantile Regression, configuration driven evaluation improvements, better plot support to config based evaluations and a brand new module to calculate area mean and standard deviation with given subregion information.

Download

We urge all users to upgrade to this version immediately. Please let us know how you are using OCW over on the [community mailing lists](#).

Finally, please see our [1.1 Roadmap](#) for an idea of the next line of development.

Source at github.com/apache/climate

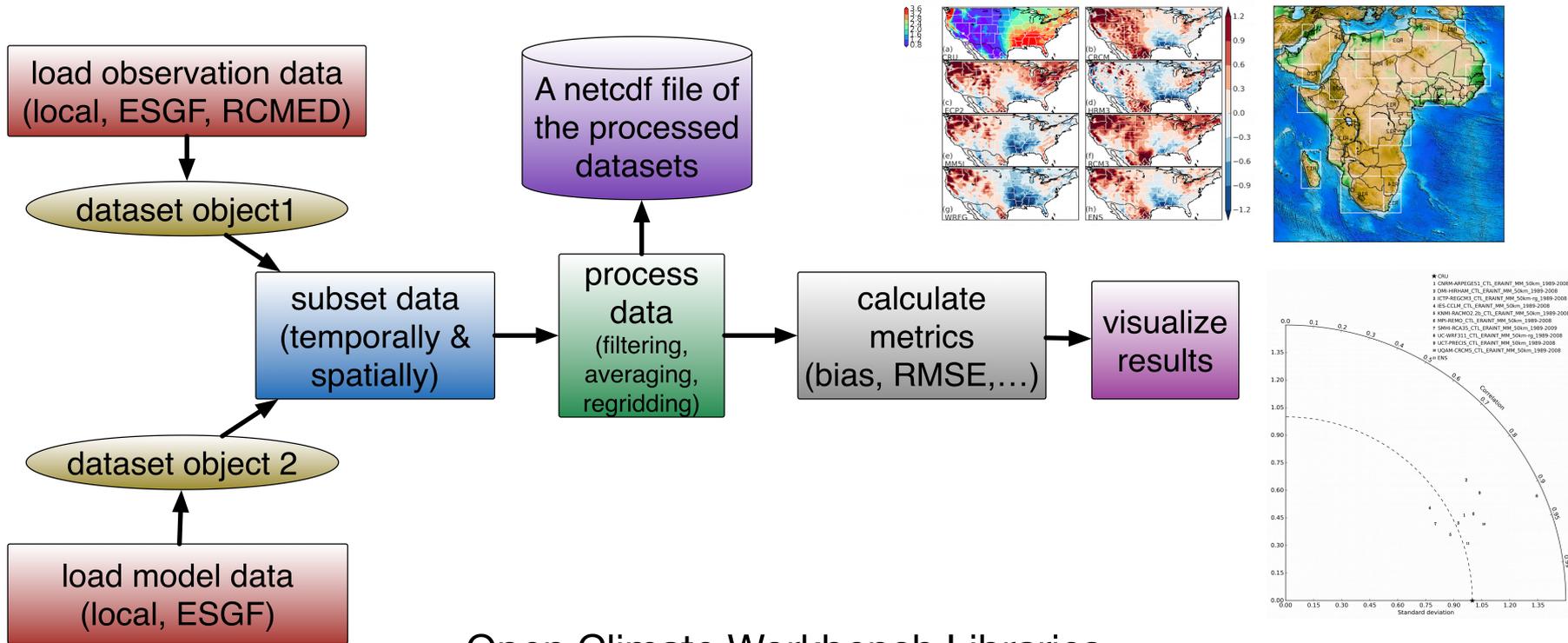
GitHub repository header for `apache / climate`. It includes a search bar, navigation links for Pull requests, Issues, and Gist, and social interaction buttons for Watch (8), Star (17), and Fork (36). The repository is noted as mirrored from `git://git.apache.org/climate.git`. Navigation tabs for Code, Pull requests (4), Pulse, and Graphs are also present.

Mirror of Apache Open Climate Workbench

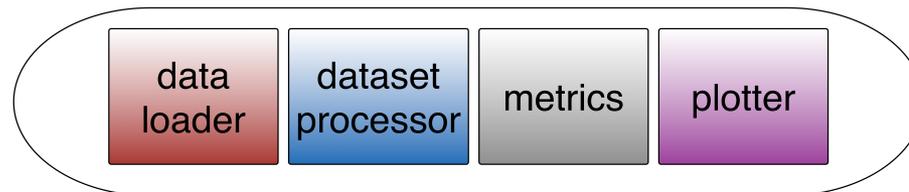
Repository statistics and navigation: 1,791 commits, 16 branches, 7 releases, and 13 contributors. The current branch is `master`. A `New pull request` button is available. The repository URL is `https://github.com/apache/climate`. A `Download ZIP` button is also present.

Commit	Description	Time
huikyole CLIMATE-771	Critical bugs in LAT_NAMES and LON_NAMES in local.py	Latest commit fd6deb 10 days ago
RCMES	CLIMATE-770 - Make boundary checking optional in spatial_regrid	12 days ago
docs	[RELEASE PREPARE] Prep for 1.0.0 release candidate	6 months ago
easy-ocw	Resolve CLIMATE-560 : Does not assume installation directory within e...	3 months ago
examples	Examples that use dataset_processor.temporal_rebin have been updated	a month ago
mccsearch	Resolve CLIMATE-559. Merge PR #142.	a year ago
obs4MIPs	Add Resources sheet into excel spreadsheet. It is no longer necessary...	2 years ago
ocw-ui	CLIMATE-572 Address deprecation and WARN's in ocw-ui/frontend npm ins...	12 days ago
ocw-vm	CLIMATE-712 - Update VM build to use conda install	4 months ago
ocw	Merge branch 'master' of https://github.com/apache/climate into CLIMA...	10 days ago
ocw_config_runner	adding init python file	6 months ago
.gitignore	Update gitignore so setup.py develop artifacts are ignored	a year ago
.mailmap	CLIMATE-608 - Add mailmap file to repo	a year ago
.pylintrc	CLIMATE-600 - Add basic .pylintrc with some sane defaults	a year ago
CHANGES.txt	[RELEASE PREPARE] Prep for 1.0.0 release candidate	6 months ago
KEYS	[RELEASE PREPARE] Prep for 1.0.0 release candidate	6 months ago
LICENSE.txt	add README to provide information on how to retrieve TRMM data from G...	2 years ago
MANIFEST.in	CLIMATE-725 Ensure that OCW 1.1 Test PyPi Works as Expected	2 months ago
NOTICE.txt	CLIMATE-342 - Update NOTICE with public domain note for TaylorDiagram	2 years ago
README.md	CLIMATE-684 - Add link to Python API to README	6 months ago

Running RCMES using configuration files: a complete start-to-finish workflow to evaluate multi-scale climate models using observational data



Open Climate Workbench Libraries



- Activity #1
: Correct biases in CORDEX RCM simulations
- Activity #2
: Pointwise Statistical downscaling using RCMES
- Activity #3
: Download and visualize the NEX-GDDP data
- **Activity #4**
: Analyze the bias corrected RCM output from Activity #1

List, download and visualize NEX-GDDP

1. Open terminal and `cd NEX-GDDP`

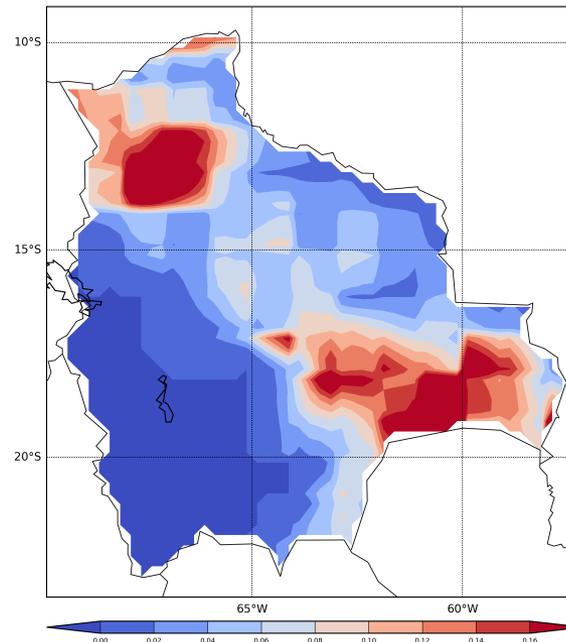
2. `./list`

3. `./download` **How fast!**

This script is an example of **Open Climate Workbench**, an open-source Python library that comprise **RCMES**.

4. `python plot_NEX-GDDP_example.py`

Statistically downscaled precipitation from NorESM1-M model for June 2100



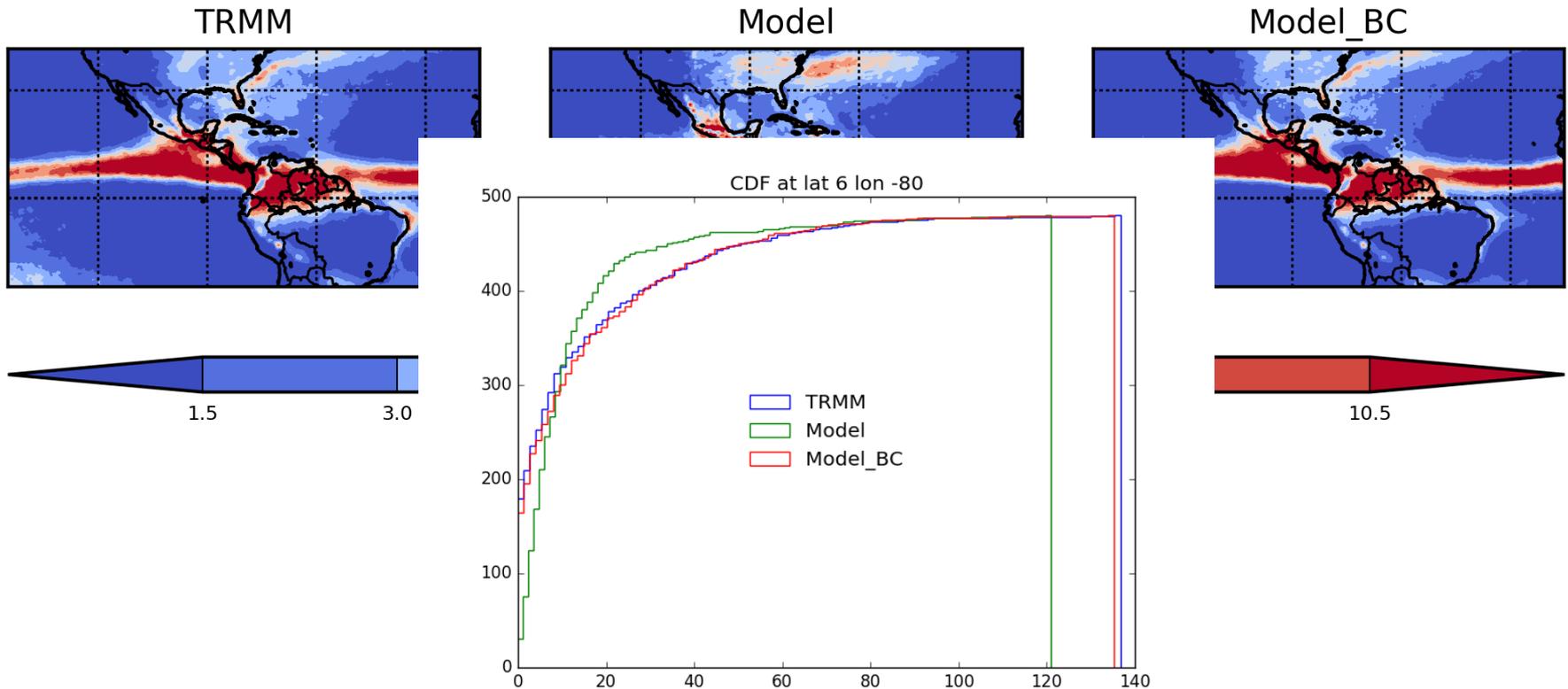
Compare TRMM, original RCM, and bias corrected RCM

1. Open terminal and `cd RCMES/analysis_examples`

2. `python check_bias_correction_CAM.py`

`python check_bias_correction_SAM.py`

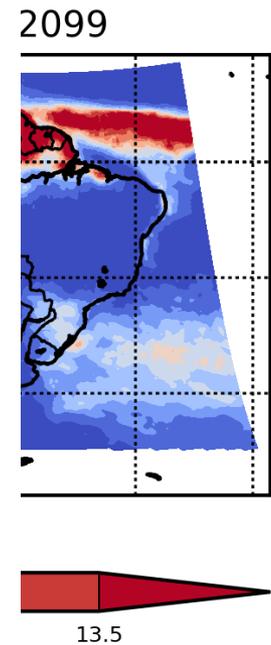
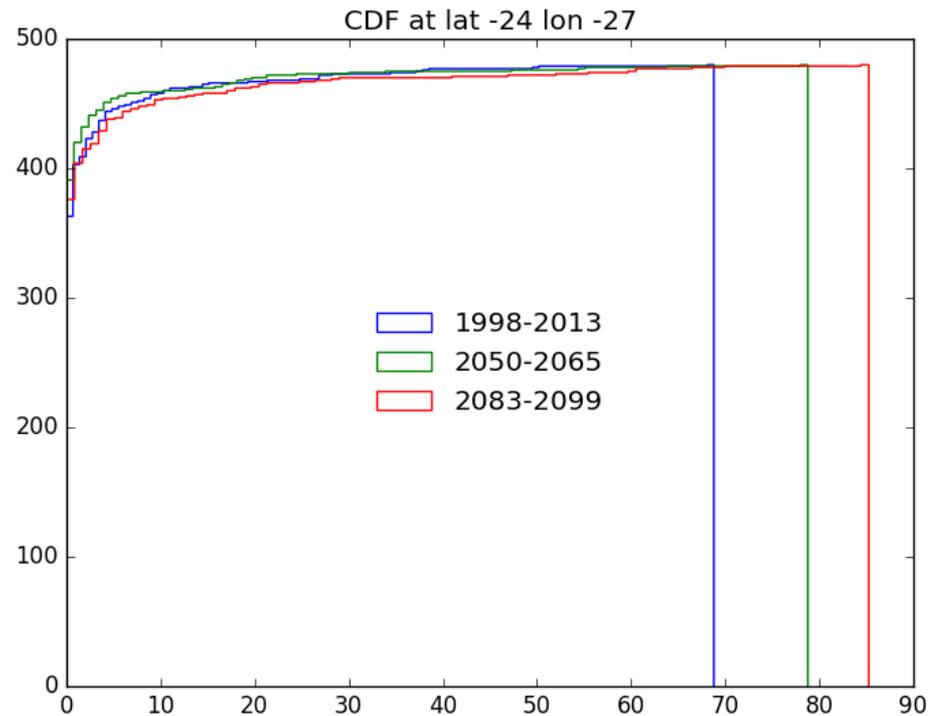
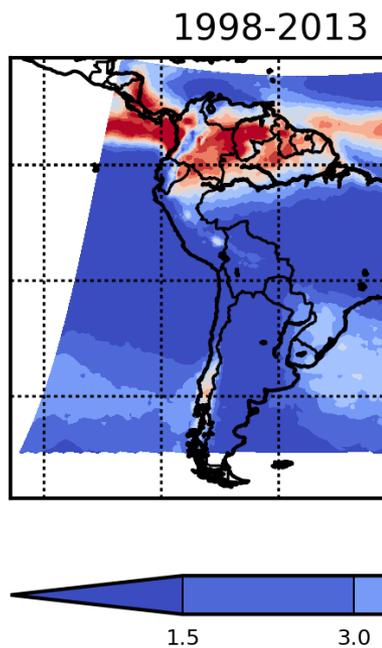
OCW-based script



How does the bias-corrected precipitation look like in the future (in June, 1998-2013 vs. 2050-2065 vs, 2083-2099)?

```
python compare_present_and_future_CAM.py
```

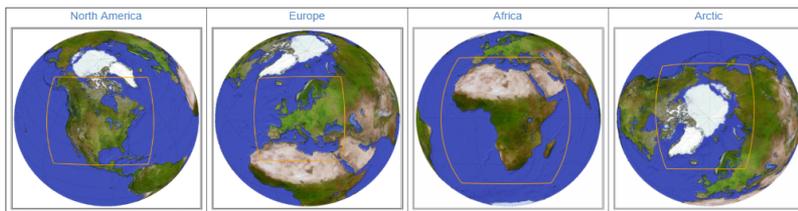
```
python compare_present_and_future_SAM.py
```



Future Direction

- Development is ongoing...
 - Adding more metrics to assure traceability and reproducibility of model evaluation results.
 - Growing user and developer base by utilizing AWS and OpenNEX datasets.
- Develop a comprehensive model evaluation system for the United States National Climate Assessment and CORDEX.

Quick Navigation - Jump to Evaluations



Results

North America Evaluations

Reference Dataset	Variables	Results Page by Seasons
CERES-EBAF	Downwelling Longwave Radiation (Surface) Upwelling Longwave Radiation (Surface) Upwelling Longwave Radiation (TOA) Downwelling Shortwave Radiation (Surface) Downwelling Shortwave Radiation (TOA) Upwelling Shortwave Radiation (Surface) Upwelling Shortwave Radiation (TOA)	Annual Summer Winter Annual Summer Winter Annual Summer Winter Annual Summer Winter Annual Summer Winter Annual Summer Winter Annual Summer Winter

<https://rcmes.jpl.nasa.gov/content/cordex-evaluation>

Where to find more information:

- <http://rcmes.jpl.nasa.gov>
- <http://climate.apache.org/>
- Email team members or dev@climate.apache.org
- <https://nex.nasa.gov>

[Lee et al. \(2018\)](#), Regional Climate Model Evaluation System powered by Apache Open Climate Workbench v1.3.0: an enabling tool for facilitating regional climate studies, GMDD under review.

Contact

Kyo Lee: huikyo.lee@jpl.nasa.gov