Training session on Regional Climate Model Evaluation System (RCMES)

Kyo Lee (huikyo.lee@jpl.nasa.gov)
(Jet Propulsion Laboratory, California Institute of Technology)
December 24, 2018

http://rcmes.jpl.nasa.gov
http://climate.apache.org

Acknowledgement

• My special thanks of gratitude to Dr. Sanjay, Dr. Krishnan, Dr. Neena, and Dr. Suhas.

• Regional Climate Model Evaluation System (RCMES) team
  Duane Waliser (PI), Huikyo Lee (co-I), Alexander Goodman, Peter Gibson, Elias Massoud, Brian Wilson, Paul Loikith\(^2\), and Antonio Monge\(^3\)
  \(^1\)JPL/Caltech, \(^2\)California State U. LA, \(^3\)Portland 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 and NASA Earth eXchange (NEX) team at NASA Ames center
# RCMES Training Outline (10:00-12:30)

<table>
<thead>
<tr>
<th>Time</th>
<th>Agenda Item</th>
<th>Process/presentations/materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00-10:15</td>
<td>Welcome and connect to Amazon Elastic Compute Cloud (EC2)</td>
<td>• Check the IP address of the assigned server  &lt;br&gt;• Connect to the server using Microsoft Remote Desktop (or terminal software)</td>
</tr>
<tr>
<td>10:15-10:30</td>
<td>Activity #1: Correct biases in CORDEX RCM simulations</td>
<td>• Quantile-based bias correction of the CORDEX WAS simulations using satellite-based precipitation observation data</td>
</tr>
<tr>
<td>10:30-10:50</td>
<td>Activity #2: Evaluate CORDEX RCM simulations</td>
<td>• Systematic evaluation of CORDEX RCMs against obs4mips using RCMES</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Break (10:50-11:00)</td>
</tr>
<tr>
<td>11:00-11:30</td>
<td>Activity #3: Pointwise Statistical downscaling using RCMES</td>
<td>• CMIP5 temperature and precipitation datasets for present and future climate  &lt;br&gt;• Compare the IPCC climate change scenarios  &lt;br&gt;• (RCP 4.5 vs. RCP 8.5)</td>
</tr>
<tr>
<td>11:30-11:45</td>
<td>Activity #4: Download and visualize the NEX-GDDP data</td>
<td>• NASA Earth Exchange Globally Daily Downscaled Projections (NEX-GDDP) in Amazon Simple Storage Service (S3)</td>
</tr>
<tr>
<td>11:45-</td>
<td>Activity #5</td>
<td>• Analyze the bias corrected RCM output and check the evaluation result</td>
</tr>
</tbody>
</table>
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 user1@xx.xxx.xxx.xxx`
- 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)

1. xx.xxx.xxx.xxx (your server’s IP address) and click ‘Connect’
2. Click ‘OK’
3. Type ‘cordex’ and click ‘OK’

user1
Set up your remote desktop (Mac)

1. **xx.xxx.xx.xxx** (your server’s IP address)

2. **Double click!!!**
Bugs in the Mac version

3. click here

4. type ‘cordex’

5. click here
• Activity #1
  : Correct biases in CORDEX RCM simulations

• Activity #2
  : Evaluate CORDEX RCM simulations

• Activity #3
  : Pointwise Statistical downscaling using RCMES

• Activity #4
  : Download and visualize the NEX-GDDP data
Running the bias correction script
(courtesy of Dr. TJ Vandal at NASA Ames, https://github.com/tjvandal/bcsd-python)

1. Open Terminal and type \texttt{cd RCMES}

(Five options: please choose one of yml files)

ex) To correct biases in the CSIRO-MK3-6-0_IITM-RegCM4-4_v5 for the CORDEX South Asia domain,

Python script
\texttt{python CORDEX_TRMM_BC_example.py}

One of the five configuration files
\texttt{CORDEX_WAS_CSIRO-QCCCE-CSIRO-Mk3-6-0_IITM-RegCM4-4_v5.yaml}

(Running this parallelized script takes more than an hour.)
Bias Correction of CORDEX simulations

• The IITM-RegCM simulations have high spatial resolution (~44 km) relative to CMIP GCMs.

• BCSD => BC : spatial disaggregation (SD) may not be necessary thanks to the resolution of CORDEX simulations.
Quantile mapping to correct simulated precipitation using TRMM observations (1)

- Inside the configuration file (CORDEX_WAS_CCCma-CanESM2_IITM-RegCM4-4_v5.vaml)
  
  **fobserved**: TRMM_regridded_RegCM4-4_v5_day_19980101-20131201_WAS-44.nc
  
  **observed_varname**: TRMM_daily_pr
  
  **fmodeled_present**: pr_WAS-44_CCCma-CanESM2_historical_r1i1p1_IITM-RegCM4-4_v5_day_19900101-20051231.nc
  
  **fmodeled_future**: pr_WAS-44_CCCma-CanESM2_rcp85_r1i1p1_IITM-RegCM4-4_v5_day_20840101-20991231.nc
  
  **modeled_varname**: pr

(Observation)
Read TRMM_daily_pr from TRMM_regridded_RegCM4-4_v5_day_19980101-20131201_WAS-44.nc

(Simulation for the present climate)
Read pr from pr_WAS-44_***_19900101-20051231.nc

(Simulation for the future climate)
Read pr from pr_WAS-44_***_20840101-20991231.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 present climate)
pr_WAS-44_***_19900101-20051231.nc

(Simulation for the future climate)
pr_WAS-44_***_20840101-20991231.nc

(Bias corrected future simulation)
BC_pr_WAS-44_***_20840101-20991231.nc
• Activity #1
  : Correct biases in CORDEX RCM simulations

• Activity #2
  : Evaluate CORDEX RCM simulations

• Activity #3
  : Pointwise Statistical downscaling using RCMES

• Activity #4
  : Download and visualize the NEX-GDDP data

- 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, Geoscientific Model Development.
- 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 (Activity #2)
  3) Statistical downscaling of coarse-resolution GCM output or bias correction of high-resolution RCM output (Activity #1 & 3)
  4) Stand-alone scripts for data processing and visualization based on OCW
Model evaluation using RCMES

1. **User Input**
   - a configuration file (.yaml)

2. **Observation for Evaluation**
   - obs4MIPs
   - Over 50 Satellite variables on ESGF
   - Other Data Centers
   - RCMES Observational Database (e.g., TRMM, CRU, UDEL)

3. **Local Disk**
   - Extract OBS data
   - Extract model data
   - Spatial Boundaries
   - Temporal Boundaries & Resolution
   - ESGF climate models (e.g., CORDEX)

4. **Regridder**
   - Put the OBS & model data on the same spatial grid

5. **Metrics Calculator**
   - (Calculate evaluation metrics)

6. **Visualizer**
   - (Plot the metrics)

7. **Data extractor to netCDF**
   - Use the re-gridded data for user’s own analyses and visualization.

8. **RCMES captures the entire workflow.**

9. **Another user can reproduce the same results using the captured workflow.**
Apache Open Climate Workbench (OCW)
https://climate.apache.org/

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

This repository is a mirror of Apache Open Climate Workbench.

**Repository Details**
- 1,791 commits
- 16 branches
- 7 releases
- 13 contributors

**Branches**
- `master`

**Recent Pull Requests**
- **huilicye** CLIMATE-771 - Critical bugs in LAT_NAMES and LON_NAMES in local.py
- RCMES
  - CLIMATE-770 - Make boundary checking optional in spatial_regrid
- docs
  - [RELEASE PREPARE] Prep for 1.0.0 release candidate
- easy-ocw
  - Resolve CLIMATE-580: Does not assume installation directory within e...
- examples
  - Examples that use dataset_processor.temporal_rebin have been updated
- mcsearch
  - Resolve CLIMATE-559. Merge PR #142.
- obs4MIPs
  - Add Resources sheet into excel spreadsheet. It is no longer necessary...
- ocw-ui
  - CLIMATE-572 Address deprecation and WARN's in ocw-ui/frontend npm ins...
- ocw-vm
  - CLIMATE-712 - Update VM build to use conda install
- ocw
  - Merge branch 'master' of https://github.com/apache/climate into CLIMA...
- ocw_config_runner
  - adding init pytton file
- .gitignore
  - Update .gitignore so setup.py develop artifacts are ignored
- .mailmap
  - CLIMATE-608 - Add mailmap file to repo
- .pylintrc
  - CLIMATE-600 - Add basic .pylintrc with some sane defaults
- CHANGES.txt
  - [RELEASE PREPARE] Prep for 1.0.0 release candidate
- KEYS
  - [RELEASE PREPARE] Prep for 1.0.0 release candidate
- LICENSE.txt
  - add README to provide information on how to retrieve TRMM data from G...
- MANIFEST.in
  - CLIMATE-725 Ensure that OCW 1.1 Test PyPi Works as Expected
- NOTICE.txt
  - CLIMATE-342 - Update NOTICE with public domain note for TaylorDiagram
- README.md
  - CLIMATE-341 - Add link to Python API in README

**Latest Commit**
- fd6debb 10 days ago
Running RCMES using configuration files: a complete start-to-finish workflow to evaluate multi-scale climate models using observational data.

Load observation data (local, ESGF, RCMED)

Dataset object 1

Subset data (temporally & spatially)

Process data (filtering, averaging, regridding)

Calculate metrics (bias, RMSE, ...)

Visualize results

Open Climate Workbench Libraries

Data loader

Dataset processor

Metrics

Plotter
Running the systematic evaluation of CORDEX WAS simulations

1. Open Terminal and type
   ```
   cd RCMES
   cd CORDEX_evaluation
   ```

2. Type
   ```
   ./evaluate_WAS-44
   ```

```
metrics 1/2: Map_plot_bias_of_multiyear_climatology
metrics 2/2: Taylor_diagram_spatial_pattern_of_multiyear_climatology
Reading the configuration file evaluation_result/WAS-44/MODIS/clt/summer/WAS-44_MODIS_clt_summer.yaml
Loading datasets:
[{{'file_path': '/mnt/CORDEX-efs/obs4mips/clt_MODIS_L3_C5_*.nc', 'name': 'MODIS', 'loader_name': 'local_split', 'variable_name': 'clt'}, {{'lat_name': 'lat', 'name': 'SMHI-RCA4', 'loader_name': 'local_split', 'variable_name': 'clt', 'file_path': '/mnt/CORDEX-efs/CORDEX/WAS-44/clt/clt_WAS-44_ECMWF-ERAINT_evaluation_r1i1p1_SMHI-RCA4_v1_mon_*.nc', 'lon_name': 'lon'}, {{'lat_name': 'lat', 'name': 'MOHC-HadRM3P', 'loader_name': 'local_split', 'variable_name': 'clt', 'file_path': '/mnt/CORDEX-efs/CORDEX/WAS-44/clt/clt_WAS-44_ECMWF-ERAINT_evaluation_r1i1p1_MOHC-HadRM3P_v1_mon_*.nc', 'lon_name': 'lon'}}]
Maximum overlap period
start time: 2001-01-01 00:00:00
end time: 2010-12-01 00:00:00
Dataset loading completed
Reference data: MODIS
Number of target datasets: 2
SMHI-RCA4
MOHC-HadRM3P
Regridding datasets: {'regrid_on_reference': True}
SMHI-RCA4 has been regridded
```
ERA-Interim Reanalysis Forced RCM simulations Available on ESGF

Schematic of Multi-Domain, Multi-Model and Multi-Variate CORDEX Model Evaluation with Obs4MIPs

Mount an EFS storage to an AWS EC2 instance

Elastic File System (EFS) storage containing obs4MIPs and CORDEX South Asia simulations
Why we need “Systematic Evaluation”

- This Config file (namelist file) is necessary to run each evaluation combination (CORDEX Domain, Season and Variable), forming a large “evaluation matrix”.
- 14 variables x 13 domains x 3 seasons x ~10 models > 5000 evaluations
- Writing that many Config files manually would be cumbersome/prohibitive.
Solution: Extract metadata from input filenames

Variable Domain

Model

User Input:
Dataset Locations
(obs4mips, CORDEX)

Evaluation Groups
(Season, Domain, Variable)

Config File

RCMES

We can group all models and obs datasets together by common attributes (domain and variable) to form a unique evaluation, and therefore automatically generate Config Files using only the dataset locations as user input.

./evaluate_WAS-44

python cordex.py WAS-44 evaluation_result /mnt/CORDEX-efs/obs4mips
/mnt/CORDEX-efs/CORDEX/WAS-44
• Activity #1
: Correct biases in CORDEX RCM simulations
• Activity #2
: Evaluate CORDEX RCM simulations
• Activity #3
: Pointwise Statistical downscaling using RCMES
• Activity #4
: Download and visualize the NEX-GDDP data
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
$Y_1', X_0', X_1'$: sorted in ascending order

$Y_1' = a\dot{X}_1' + b$ where $Y_0' = a\dot{X}_0' + b$. $a$ 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_tasmin_DJF.yaml`

Python script
Configuration file
View the statistically downscaled tasmin results

- The results can be found in statistical_downscaling/Pune_tasmin_MPI_CRU_RCP8.5_2071-00 folder
Quantile mapping of the daily minimum temperature for Pune in DJF

Original model output

Statistically downscaled model output
Run another example: taxmax in Pune

```python
python run_statistical_downscaling.py MIROC5_taxmax_DEC.yaml
```

Make your own example by editing the yaml file

```yaml
case_name: Pune_MIROC5_tasmax_DEC_RCP85_2071-00
downscaling_option: 3
location:
    name: Pune
    grid_lat: 18.5204
    grid_lon: 73.8567
month_index: !!python/tuple [12]
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_1981-2010.nc
future:
    scenario_name: RCP8.5_2071-00
    path: ./data/tasmax_Amon_MIROC5_rcp85_207101-210012.nc
```

Output folder name

- IPSL, MPI, and MIROC5
- tas, tasmin, and tasmax
- RCP 4.5 and 8.5
- (2041-2070) and (2071-2100)

Search Google with the keyword ‘latitude and longitude of XXX’

Season: December only in this case
• 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.

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

Global Meteorological Forcing Dataset (observation, 1950-2005)

Bias-Correction Spatial Disaggregation (BCSD)

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!
4. python plot_NEX-GDDP_example.py

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

Statistically downscaled precipitation from NorESM1-M model for December 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
Compare TRMM, original simulation, and bias corrected simulation

1. Open terminal and cd RCMES/analysis_examples
2. python check_bias_correction.py
   OCW-based script
How does the bias-corrected precipitation look like in the future (in July-August, TRMM (1998-2013) vs. two bias-corrected simulations (2084-2099))? 

```
python compare_present_and_future.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.

https://rcmes.jpl.nasa.gov/content/cordex-evaluation
Where to find more information:

- Email team members at [dev@climate.apache.org](mailto:dev@climate.apache.org)
- [https://nex.nasa.gov](https://nex.nasa.gov)

Contact

Kyo Lee: [huikyo.lee@jpl.nasa.gov](mailto:huikyo.lee@jpl.nasa.gov)