



# IRI's new Forecasting system and Introduction of PyCPT tools For regional Forecasting

Nachiketa Acharya

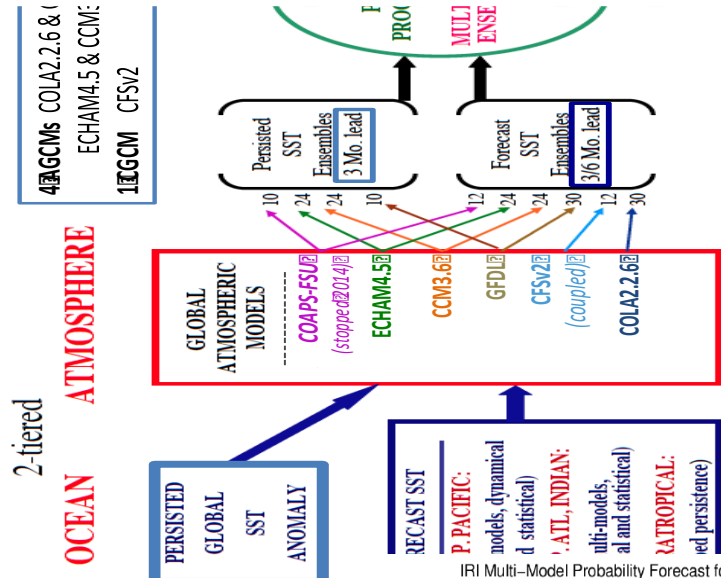
[nachiketa@iri.columbia.edu](mailto:nachiketa@iri.columbia.edu)

# Real-time IRI's Probabilistic Seasonal Forecasting

IRI began routinely providing calibrated user-oriented seasonal climate forecasts since the late 1990s based on a **2-tiered multi-model ensemble** dynamical prediction system.

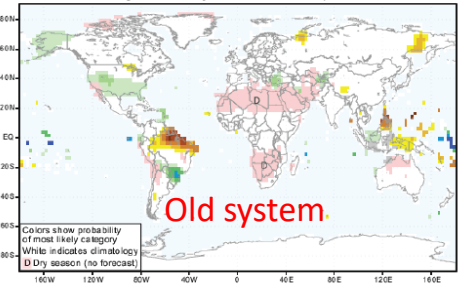
Multi-Model Ensemble Forecast

Seasonal rainfall temperature probabilistic forecast

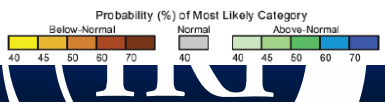
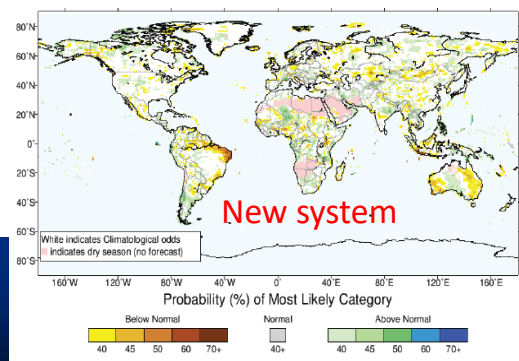


Latest predic  
<http://pred.ideo.columbia.edu/forec>

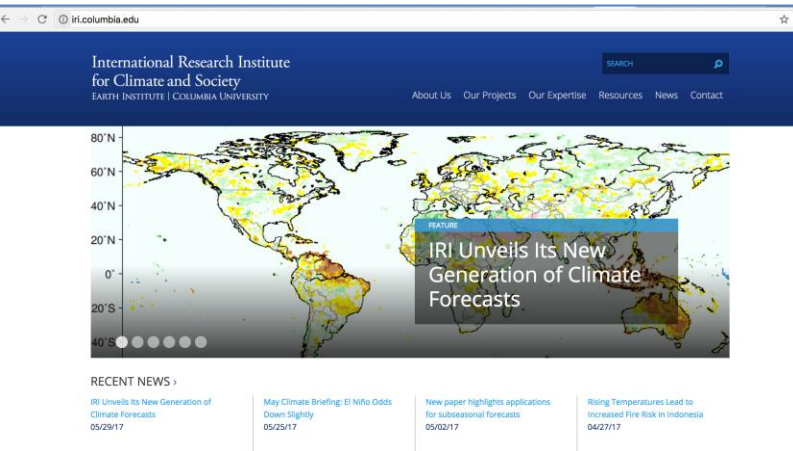
IRI Multi-Model Probability Forecast for Precipitation for May-June-July 2016, Issued April 2016



IRI Multi-Model Probability Forecast for Precipitation for May-June-July 2017, Issued April 2017



# Advances in Real-time IRI's Probabilistic Seasonal Forecasting



**New system of seasonal forecast has been operational from April, 2017**

	Old IRI forecast	New IRI forecast
GCM used (Predictors)	2-tier (uncoupled) ECHAM 4.5, CCM3.6, COLA, GFDL,CFSv2	1-tier (coupled) <b>NMME models</b>
Observed data used (Predictand)	Precip: CMAP Temp: CAMS	Precip: CPC-CMAP Temp: GCHN updated
Forecast Resolution	2.5 degree grid	1 degree grid
Calibration method	<ul style="list-style-type: none"> <li>• Pattern-based correction of ensemble means</li> <li>- PC Regression based on tropical precip EOFs</li> <li>- Spread estimate from historical forecasts with forecast SST</li> <li>• Equal weighting of corrected models</li> <li>• Parametric forecast probabilities (T - Gaussian, P - transformed Gaussian)</li> </ul>	<b>Extended Logistic Regression (Non-Gaussian) at grid point level.</b>
Dry mask	Forecast are only produced when the climatology being more than 30 mm precipitation in any given season	Forecast are only produced when the at least 10% of the training sample are non-zero.
<b>Making Flexible forecast</b>	Used mean and SD of the forecast, then use parametric approach	Integrated part of the ELR method



# NMME datasets



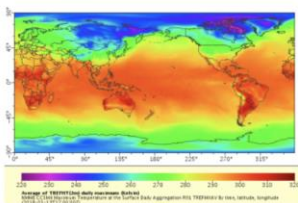
Home Climate Information Data Access Customer Support Contact About Search

## North American Multi-Model Ensemble

- Quick Links
- Land-Based Station
- Satellite
- Radar
- Model
  - Datasets
  - AEC
  - CFS
  - CM2.X
  - CMIP5
  - GDAS
  - GEFS
  - GFS
  - HYCOM
  - NAM
  - NARR
  - NCOM
  - NCOM-Regional
  - NDFD
  - NDGD
  - NGM
  - NMME
  - R1 / R2
  - RAP
  - RTOP5
  - SRR5
  - NOGAPS
  - Ocean Models
  - Reanalysis
  - Numerical Weather Prediction
  - Climate Prediction

The North American Multi-Model Ensemble (NMME) is a multi-model, seasonal forecasting system consisting of coupled models from North American modeling centers. This data contains global, 12-month forecasts of 13 key variables.

The NMME contributors are NOAA's National Centers for Environmental Prediction (NCEP), NOAA's Geophysical Fluid Dynamics Laboratory (GFDL), the National Center for Atmospheric Research (NCAR), NASA's Goddard Space Flight Center (GSFC), and Canada's Centre for Climate Modeling and Analysis (CCCma). The models include The Community Climate System Model version 4 (CCSM4), the Climate Forecast Model version 2 (CFV2-2011), the Canadian Coupled Climate Model versions 3 and 4 (CanCM3, CanCM4), the Global Environmental Multiscale/Nucleus for European Modeling of the Ocean (GEM/NEMO), and the Goddard Earth Observing System Model version 5 (GEOS-5).



NMME CCSM4 Daily Maximum Temperature

NMME data is daily or 6-hourly with a 1° by 1° spatial resolution. Most NMME datasets have 10 realizations for each variable. Variables include the following; by model and variable name ("N/A" denotes variables that are not available).

Variable	CCSM4	CESM1	CFV2-2011	CanCM3	CanCM4	FLORB-01	GEOS-5
Precipitation rate	pr	precip	pr	prlr	prlr	pr	pr
Daily Maximum Surface Air Temperature	TREFMXAV	Tasmax	tasmax	tasmax	tasmax	tasmax	tasmax
Daily Minimum Surface Air Temperature	TREFMNAV	Tasmin	tasmin	tasmin	tasmin	tasmin	tasmin
Zonal surface	STX	Stx	stx	stx	stx	N/A	stx

<https://iridl.ideo.columbia.edu/SOURCES/.Models/.NMME/>

**Models NMME**

Models NMME: North American Multi-Model Ensemble (NMME).

**Documents**

- [overview](#) an outline showing sub-datasets of this dataset
- [CTE home](#) Climate Test Bed NMME Page
- [NMME Description](#) North American Multi-Model Ensemble: Improving NOAA's Seasonal Prediction Capability
- [NMME Home](#) Information about the NMME project from CPC

**Semantic Documents**

[auxinfo owl](#)

**Datasets and Variables**

- [CanCM4](#) Models NMME CanCM4[HINDCAST FORECAST]
- [CanSISv2](#) Models NMME CanSISv2[HINDCAST FORECAST]
- [CMC1-CanCM3](#) Models NMME CMC1-CanCM3[HINDCAST FORECAST]
- [CMC2-CanCM4](#) Models NMME CMC2-CanCM4[HINDCAST FORECAST]
- [COLA-RSMAS-CCSM4](#) Models NMME COLA-RSMAS-CCSM4[ecb210 MONTHLY mc8210]
- [COLA-RSMAS-CCSM4](#) Models NMME COLA-RSMAS-CCSM4[ecb210 MONTHLY mc8210]
- [CPC-CMAP](#) Models NMME CPC-CMAP[grate]
- [CPC-CMAP-URD](#) Models NMME CPC-CMAP-URD[grate]
- [CPC-PRECIP](#) Models NMME CPC-PRECIP[grate]
- [GEM-NEMO](#) Models NMME GEM-NEMO[HINDCAST FORECAST]
- [GFDL-CM2p1](#) Models NMME GFDL-CM2p1[MONTHLY]
- [GFDL-CM2p1-ecb04](#) Models NMME GFDL-CM2p1-ecb04[ecb210 MONTHLY mc8210]
- [GFDL-CM2p5-FLOR-A06](#) Models NMME GFDL-CM2p5-FLOR-A06[MONTHLY mc8110 sc8110]
- [GFDL-CM2p5-FLOR-B01](#) Models NMME GFDL-CM2p5-FLOR-B01[MONTHLY mc8110 sc8110]
- [GEM-CAMS](#) Models NMME GEM-CAMS[temp updated]
- [IRI-ECHAM4q5-AnomalyCoupled](#) Models NMME IRI-ECHAM4q5-AnomalyCoupled[MONTHLY]
- [IRI-ECHAM4q5-DirectCoupled](#) Models NMME IRI-ECHAM4q5-DirectCoupled[MONTHLY]
- [LSMASX](#) Models NMME LSMASX[land]
- [NASA-GEOSS25](#) Models NMME NASA-GEOSS25[HINDCAST FORECAST]
- [NASA-GMAO](#) Models NMME NASA-GMAO[MONTHLY]
- [NASA-GMAO-062012](#) Models NMME NASA-GMAO-062012[MONTHLY mc8110 sc8110]
- [NCAR-CESM1](#) Models NMME NCAR-CESM1[HINDCAST FORECAST]
- [NCDC-OISST](#) Models NMME NCDC-OISST[stx]
- [NCEP-CFSv1](#) Models NMME NCEP-CFSv1[MONTHLY]
- [NCEP-CFSv2](#) Models NMME NCEP-CFSv2[HINDCAST FORECAST]
- [Owv\\_SST](#) Models NMME Owv\_SST[stx]

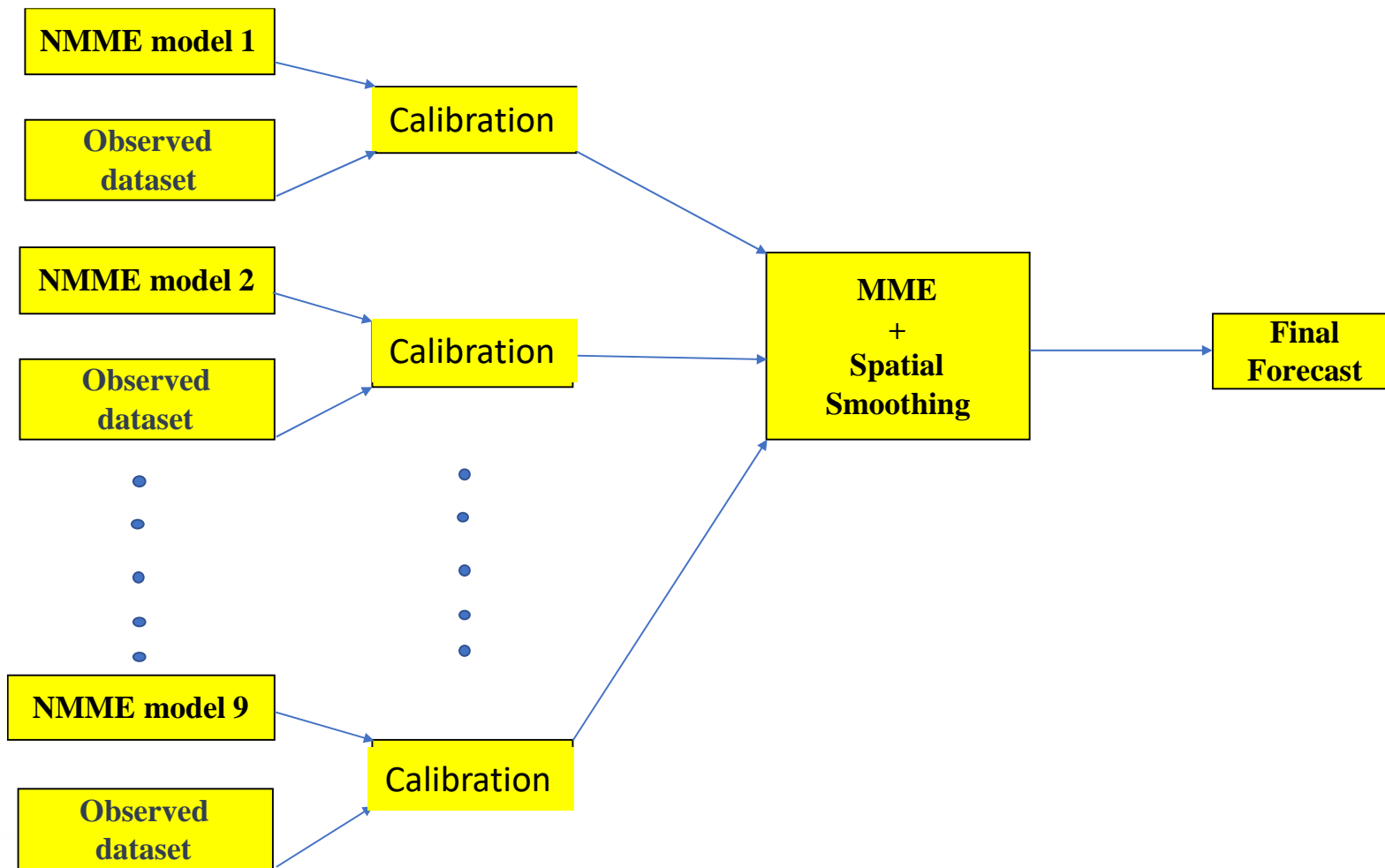
**Other Info**

**ACKNOWLEDGEMENTS**

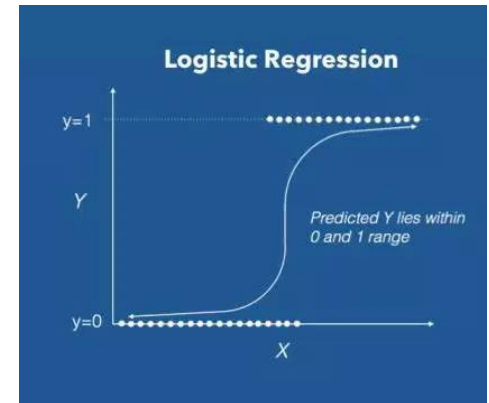
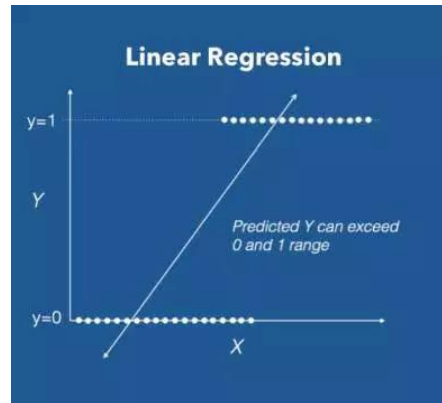
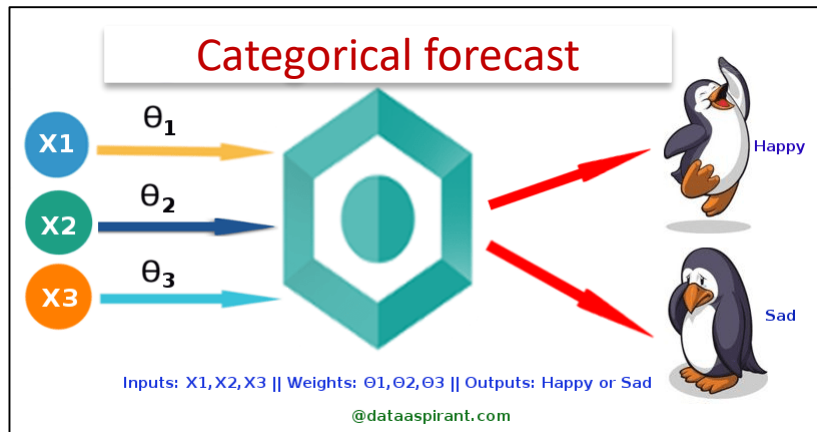
In order to document NMME-Phase II data impact and enable continuing support, users of NMME data are expected to acknowledge NMME data and the participating modeling groups. The NMME model output should be <https://www.earthsystemgrid.org/search/?project=Nmme> in publications, users should include a table (referred to below as Table XX) listing the models and institutions that provided model output used in the NMME documenting the models, where "Table XX" in the paper should list the models and modeling groups that provided the NMME data. In addition, an acknowledgement similar to the following should be included in any public system and we thank the climate modeling groups (Environment Canada, NASA, NCAR, NOAA/GFDL, NOAA/NCEP and University of Miami) for producing and making available their model output. NOAA/NCEP, NOAA/



# Flow chart of new forecast methodology



# IRI's New Calibration Method



## Logistic Regression

Logistic regression is well famous method to make probability forecast

$$\ln \left[ \frac{p}{1-p} \right] = f(x)$$

Where  $p$  is the (cumulative) probability of not exceeding the quantile  $q$

$$p = Pr\{V \leq q\}$$

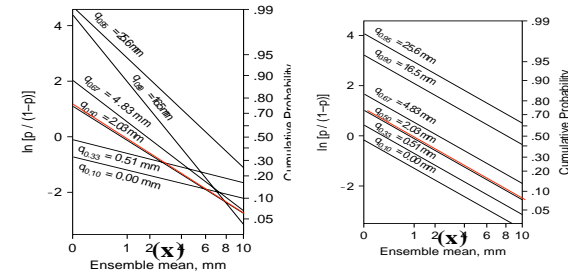
- ❖ Logistic regression (LR), a nonlinear regression method where probability itself can be considered as the predictand rather than a measurable physical quantity, is an **alternative model for Gaussian approach**.
- ❖ Logistic Regression is a **Machine Learning algorithm** which is used for the classification problems, it is a predictive analysis algorithm and based on the concept of probability
- ❖ Unlike linear regression, **no need to fulfill assumptions of linearity, normality and homoscedasticity**.

# Modification of LR Method

$$\ln \left[ \frac{p}{1-p} \right] = f(x) + g(q) \quad \text{Where} \quad \begin{aligned} f(x) &= b_0 + b_1 \overline{x}_{ens} \\ g(q) &= b_2 q \end{aligned}$$

## Limitations:

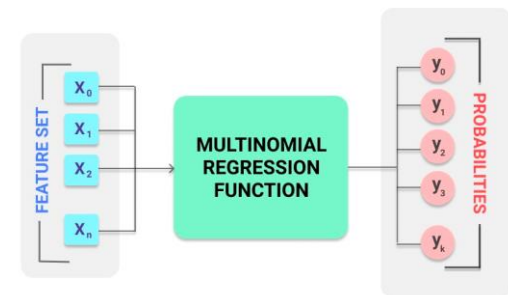
- ❖ Probabilities of different categories estimated by fitting separate equations for selected predictand quantile thresholds ( $q$ ), yielding a collection of **threshold probabilities rather than full forecast probability distributions**.
- ❖ However, the most problematic consequence of separate equations for different predictand thresholds is that forecasts derived from the different equations **are not constrained to be mutually consistent**.



## Extending Logistic Regression:

❖ Extending LR (ELR) by including the predictand threshold as an additional predictor (link function  $g$  itself function of the quantile  $q$ ), allows **the derivation of full predictive distributions** to avoid the problem of potentially incoherent forecast probabilities (Wilks, 2009).

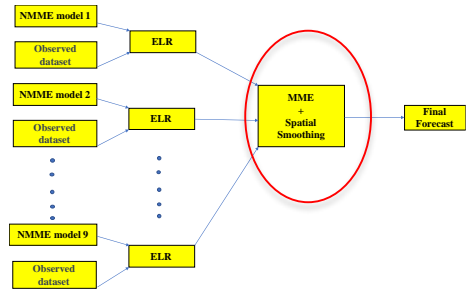
❖ Cumulative probability for a **smaller predictand threshold cannot be larger than the probability for a larger threshold**.



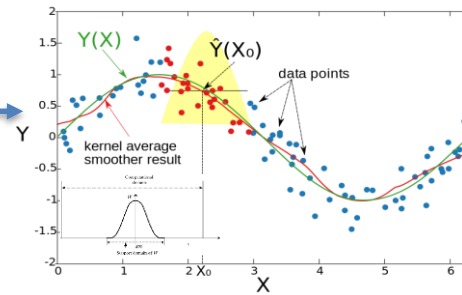
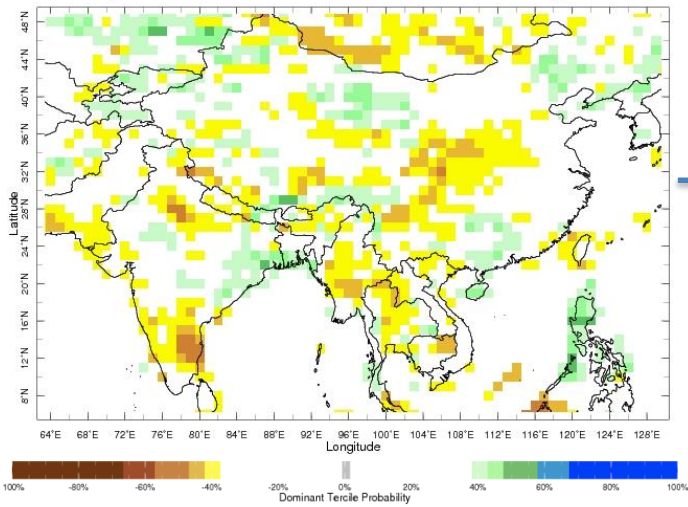
Wilks, D., 2009: Extending logistic regression to provide full-probability-distribution MOS fore- casts. *Meteor. Appl.*, 16, 361–368.



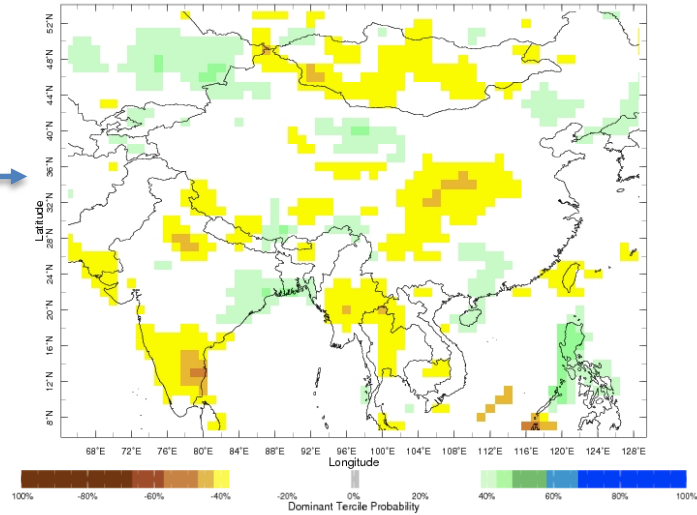
# Spatial Smoothing of final forecast: why & how?



Original Forecast



Smoothed Forecast



Smoothing with Kernel function (Gaussian) with a rectangle of size 9 by 9.



# Hindcast Skill Map (1982-2010)

<http://iridl.ideo.columbia.edu/maproom/Global/Forecasts/index.html#tabs-2>

Rank probability Skill Score:

$$RPSS = 1 - \frac{RPS_{forecast}}{RPS_{reference}}$$

$$RPS = \sum_{cat=1}^{Ncat} (Pcum_{F(cat)} - Pcum_{O(cat)})^2$$

**Precipitation Seasonal Hindcast Skill**

Seasonal skill score based on the historical performance of each calibrated NMME model and their multimodel ensemble (1982-2010).

Skill is mapped by calendar month for seasonal lead times. Lead 1 = months 2-4, Lead 2 = months 3-5, Lead 3 = months 4-6, Lead 4 = months 5-7 after the forecast is issued. Forecasts skill scores combine start times by calendar month and across years 1982 to 2010. The observational reference datasets are CMAP-URD for precipitation and GHCN-CAMS for temperature. The models included in the assessment are: the Center for Ocean-Land-Atmosphere Studies/University of Miami (COLA-RSMAS-COSMA), one from the National Aeronautics and Space Administration (NASA-GMAO-060712), three from the Geophysical Fluid Dynamics Laboratory (GFDL-CM2p1-a04, GFDL-CM2p5-FIOR-A06, GFDL-CM2p5-FIOR-B01), two from the Canadian Meteorological Center (CMC1-CanCM3, CMC2-CanCM4), one from NOAA's Centers for Environmental Prediction (NCEP-CFSv2) and one from the National Center for Atmospheric Research (NCAR-CESM1).

These skill scores diagnostics maps give a sense of where and when (issued which months of the year and for which seasonal lead times) the probabilistic seasonal forecasts have the potential to provide useful information, based on hindcasting.

**Skill scores definitions:**

- RPSS: Ranked Probability Skill Scores (RPSS; Epstein (1969); Murphy (1969, 1971); Weigel et al. (2007)) are used to quantify the extent to which the calibrated tercile-category predictions are improved compared to climatological frequencies. RPSS values tend to be small, even for skillful forecasts. The approximate relationship between RPSS and correlation being such that a RPSS value of 0.1 corresponds to a correlation of about 0.44 (Tippett et al. 2010).

**References:**

- Epstein, E.S., 1969: A Scoring System for Probability Forecasts of Ranked Categories. *J. Appl. Meteor.*, 8, 985-987
- Murphy, A.H., 1969: On the "Ranked Probability Score". *J. Appl. Meteor.*, 8, 988-989
- Murphy, A.H., 1971: A Note on the Ranked Probability Score. *J. Appl. Meteor.*, 10, 155-156
- Tippett, M.K., A.G. Barnston, and T. DelSole, 2010: Comments on "Trends, Sameness and Uncertainty Estimates for Skill Measures for Seasonal Prediction". *Mon. Wea. Rev.*, 138, 1487-1493
- Weigel, A.P., M.A. Liniger, and C. Appenzeller, 2007: The Discrete Brier and Ranked Probability Skill Scores. *Mon. Wea. Rev.*, 135, 118-124

**Temperature Seasonal Hindcast Skill**

Seasonal skill score based on the historical performance of each calibrated NMME model and their multimodel ensemble (1982-2010).

Skill is mapped by calendar month for seasonal lead times. Lead 1 = months 2-4, Lead 2 = months 3-5, Lead 3 = months 4-6, Lead 4 = months 5-7 after the forecast is issued. Forecasts skill scores combine start times by calendar month and across years 1982 to 2010. The observational reference datasets are CMAP-URD for precipitation and GHCN-CAMS for temperature. The models included in the assessment are: the Center for Ocean-Land-Atmosphere Studies/University of Miami (COLA-RSMAS-COSMA), one from the National Aeronautics and Space Administration (NASA-GMAO-060712), three from the Geophysical Fluid Dynamics Laboratory (GFDL-CM2p1-a04, GFDL-CM2p5-FIOR-A06, GFDL-CM2p5-FIOR-B01), two from the Canadian Meteorological Center (CMC1-CanCM3, CMC2-CanCM4), one from NOAA's Centers for Environmental Prediction (NCEP-CFSv2) and one from the National Center for Atmospheric Research (NCAR-CESM1).

These skill scores diagnostics maps give a sense of where and when (issued which months of the year and for which seasonal lead times) the probabilistic seasonal forecasts have the potential to provide useful information, based on hindcasting.

**Skill scores definitions:**

- RPSS: Ranked Probability Skill Scores (RPSS; Epstein (1969); Murphy (1969, 1971); Weigel et al. (2007)) are used to quantify the extent to which the calibrated tercile-category predictions are improved compared to climatological frequencies. RPSS values tend to be small, even for skillful forecasts. The approximate relationship between RPSS and correlation being such that a RPSS value of 0.1 corresponds to a correlation of about 0.44 (Tippett et al. 2010).

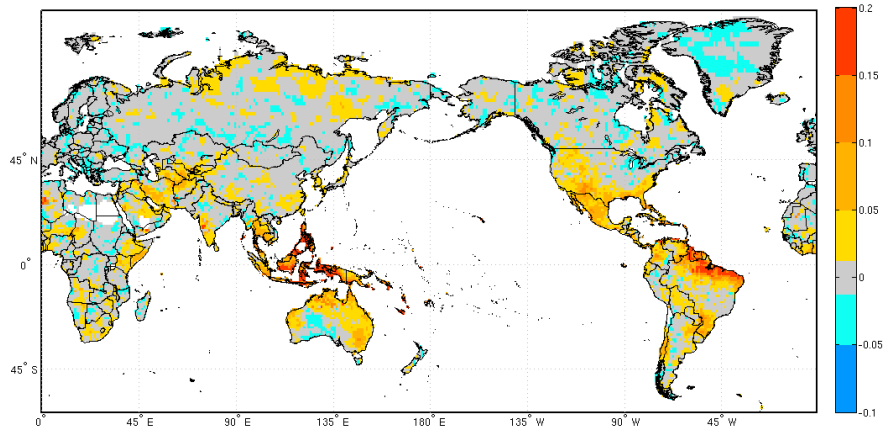
**References:**

- Epstein, E.S., 1969: A Scoring System for Probability Forecasts of Ranked Categories. *J. Appl. Meteor.*, 8, 985-987
- Murphy, A.H., 1969: On the "Ranked Probability Score". *J. Appl. Meteor.*, 8, 988-989
- Murphy, A.H., 1971: A Note on the Ranked Probability Score. *J. Appl. Meteor.*, 10, 155-156
- Tippett, M.K., A.G. Barnston, and T. DelSole, 2010: Comments on "Trends, Sameness and Uncertainty Estimates for Skill Measures for Seasonal Prediction". *Mon. Wea. Rev.*, 138, 1487-1493
- Weigel, A.P., M.A. Liniger, and C. Appenzeller, 2007: The Discrete Brier and Ranked Probability Skill Scores. *Mon. Wea. Rev.*, 135, 118-124

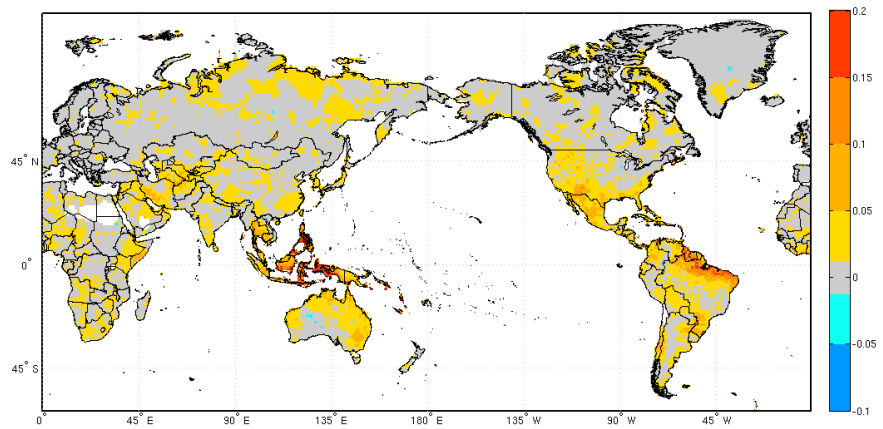


# Lead-1 skill scores for precipitation

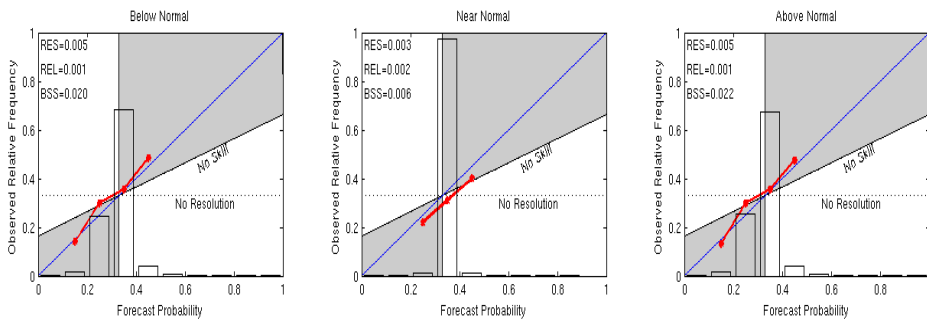
RPSS



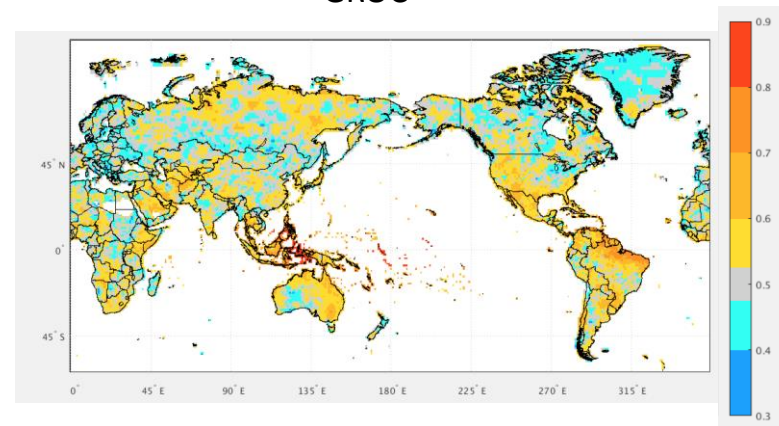
Logarithmic skill score (LSS)



Reliability



GROC



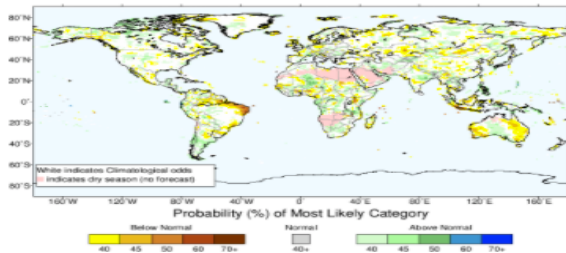
# Flexible forecast format

Please refer to our [licensing agreement](#) for permission to use any IRI forecast material.

Seasonal Climate Forecast

Region:  Type:  Issue Year:  Issue Month:  Leads:

## IRI Multi-Model Probability Forecast for Precipitation for May-June-July 2017, Issued April 2017



### Discussion

April 2017 Climate Forecast Discussion for May-Jul through Aug-Oct

The SST forecast shows El Niño development by Jul-Sep, reaching moderate strength by the fourth and final running forecast season of Aug-Oct. A positive Indian Ocean Dipole exists throughout the four forecast seasons, and becomes somewhat stronger by the fourth season. The tropical Atlantic maintains near to slightly above average SST during all four seasons.

Slightly enhanced probabilities for below-normal precipitation are predicted in far northern South America during May-Jul, progressively strengthening and expanding to include more of northern South America by Jul-Sep, and persisting into Aug-Oct. A similar scenario of strengthening chances for below-normal is predicted for Indonesia beginning in Jun-Aug, and for eastern Australia over the course of all four forecast periods. By Aug-Oct season, the region leaning toward below-normal in eastern Australia expands northward to join Indonesia.

A tendency toward above-normal temperature is predicted in a general manner over much of the globe for all four forecast seasons, with a few exceptions such as a small region in west-central interior North America, straddling the U.S./Canadian border, during the first two seasons. Probabilities for above-normal are strongest in Greenland and northeastern Canada during the second, third and fourth seasons, and in varying parts of Eurasia and South America during the course of the four seasons.

### Overview

Starting in April 2017, the IRI probabilistic seasonal climate forecast product is based on a re-calibration of model output from the U.S. National Oceanographic and Atmospheric Administration (NOAA's North American Multi-Model Ensemble Project (NMME)). This includes the ensemble seasonal prediction systems of NOAA's National Centers for Environmental Prediction, Environment and Climate Change Canada, NOAA/Geophysical Fluid Dynamics Laboratory, NASA, NCAR and COLA/University of Miami. The output from each NMME model is re-calibrated prior to multi-model ensembling to form reliable probability forecasts. The forecasts are now presented on a 1-degree latitude-longitude grid.

**Disclaimer:** The IRI seasonal forecast is a research product. Please see the NOAA CPC forecast for the official seasonal forecast over the U.S. Please consult your country's national meteorological service for the official forecast for your country.

Please see the "Discussion" item for an overview of the individual forecasts.

The climatological base period used is 1982-2010. [Details of the forecast system, post-processing, and recommended references for citation can be found here.](#) Forecasts from the individual NMME models are shown on NOAA CPC's website. Verifications of IRI's real-time forecasts issued since 1998 can be found on the [Seasonal Climate Verifications pages](#).

To aid in interpretation of the forecast probabilities, maps of the observed precipitation and temperature percentiles are plotted in physical units here: [Climatological Percentiles Maproom](#).

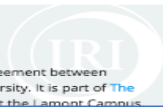
The IRI forecasts are also available as a flexible probabilistic format, providing the probability of exceedance (or non-exceedance) of a user-specified percentile of the climatological distribution: [Go to IRI Flexible Forecasts](#)

Temperature    Precipitation

Connect with IRI



The IRI was initially established as a cooperative agreement between NOAA's Climate Program Office and Columbia University. It is part of The Earth Institute, Columbia University, and is located at the Lamont Campus.



Staff Directory

Support our Work

Employment

Climate Forecasts    Flexible Forecasts    Region:    Target Time:    Probability:

Precipitation Flexible Seasonal Forecast    May-Jul 2018    exceeding    Percentile: 50.0    %ile

Description    Dataset Documentation    Instructions    Contact Us

## Precipitation Flexible Seasonal Forecast

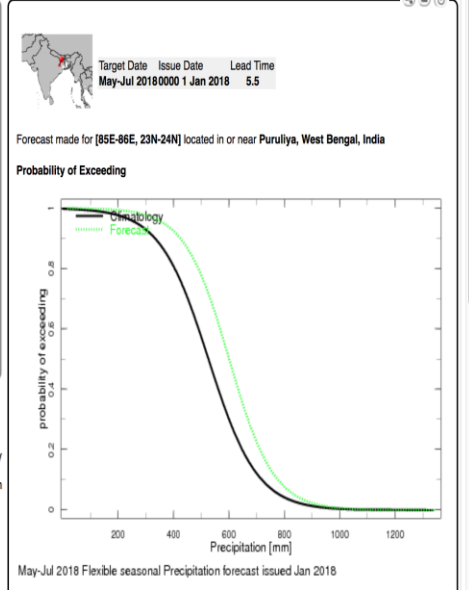
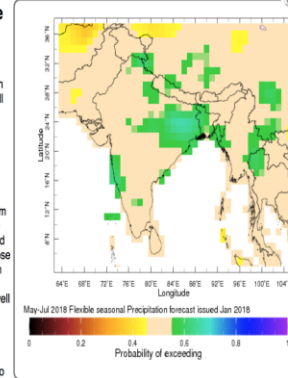
This seasonal forecasting system consists of probabilistic precipitation seasonal forecasts based on the full estimate of the probability distribution.

Please refer to our [licensing agreement](#) for permission to use any IRI forecast material.

Probabilistic seasonal forecasts from multi-model ensembles through the use of [statistical recalibration](#), based on the historical performance of those models, provide reliable information to a wide range of climate risk and decision making communities, as well as the forecast community. The flexibility of the full probability distributions allows to deliver interactive maps and point-wise distributions that become relevant to user-determined needs.

The default map shows globally the seasonal precipitation forecast probability (colors between 0 and 1) of exceeding the 50<sup>th</sup> percentile of the distribution from historical 1982-2010 climatology. The forecast shown is the latest forecast made (e.g. Dec 2017) for the next season to come (e.g. Jan-Mar 2018). Four different seasons are forecasted and it is also possible to consult forecasts made previously. The forecasts are directly computed from the [extended logistic regression](#) model as probabilities of exceeding (or non-exceeding) of every 5th percentile of the climatological distribution. The specific quantile (in steps of 5 percentile points) can then be selected. The user can also specify a quantitative value in physical units (here seasonal total precipitation in mm) for probability of exceeding or non-exceeding. The final probability maps are smoothed spatially with a 9x9 point Gaussian smoother.

Clicking on a point on the map will show the local probability of exceeding and probability distribution of the forecast (green) together with the climatological distribution (black).



The distribution of seasonal rainfall may often rightfully be approximated by a normal distribution, especially when considering only years when it actually rained. But there can be a significant number of years in the sample when rainfall is 0. This translates in situations where the probability of rainfall to be just above 0 is less than 100%, which are represented in by the vertical lines at 0mm in the distributions graphs.

### Colors Scales

Color scales are colors indicating that the distribution of the forecast tends towards drier (shades of brown) or wetter (shades of blue) conditions than normal (moccasin).

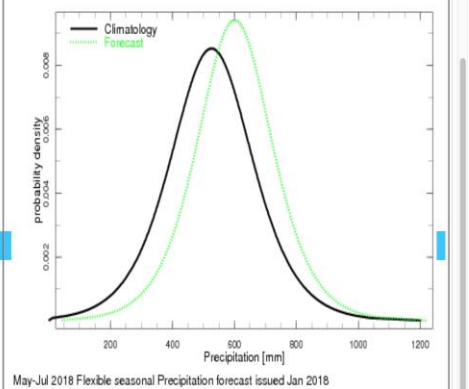
### Older Forecasts

Older forecasts can be found in the last tabbed entries of this section. Forecasts made from February 2017 use the exact same ELR method only the presentation of the local poe graph was discrete, and the local pdf was not shown. Forecasts made from July 2012 to March 2017 are discontinued and used a different methodology to obtain the full distribution of the forecasts.

### Release dates

Our seasonal forecasts are released on the 15th of each month. If the 15th falls on a weekend or holiday, they are released on the closest workday, earlier or later.

### Probability Distribution





# IRI's Experimental Precipitation Sub-seasonal Forecasts

<http://iridl.ldeo.columbia.edu/maproom/Global/ForecastsS2S/index.html>

← → ↻ Not Secure | iridl.ldeo.columbia.edu/maproom/Global/ForecastsS2S/index.html

Maproom Climate Subseasonal Forecasts Region Global

## Subseasonal Forecasts

Subseasonal forecasts of precipitation and temperature.

This section is dedicated to subseasonal forecasts, i.e. that bridge the gap between medium range weather forecasts (up to 10 days) and seasonal climate predictions (above 1 month). They are issued at different frequencies (from daily to once or twice a week) forecasting daily values with lead times from 1 to about 40 days, depending on the Global Producing Center (GPC). The availability of forecast products in the subseasonal-to-seasonal time range offers an unprecedented opportunity to develop intra-seasonal forecast information that other forecasts can't, in association with increased lead time compared to medium range weather forecasts, and with higher temporal resolution than seasonal forecasts that give an overview of an upcoming season (3 months). For instance, subseasonal forecasts may allow delivering relevant information about key climate characteristics such as the timing of the onset of a rainy season for agriculture, the risk of extreme rainfall events or heat waves in regards to public health.

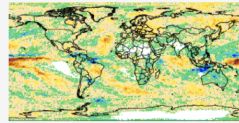
At present, these maprooms include experimental subseasonal forecasts of weekly and biweekly precipitation and temperature terciles and above median) based on the multi-model ensemble of individual forecasts issued every Saturdays through the SubX real-time database and every Thursday through the delayed S2S database.

SubX Forecasts SubX Hindcast Skill S2S Lagged Forecasts

### SubX Forecasts

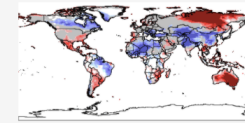
#### Precipitation Median Probability Forecast

Calibrated Subseasonal Two-category precipitation real-time forecasts



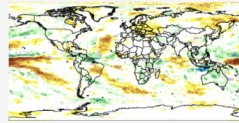
#### Temperature Weekly Probability Forecast

Calibrated subseasonal tercile categories temperature forecasts



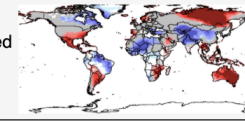
#### Precipitation Biweekly Probability Forecast

Calibrated Subseasonal tercile category biweekly-precipitation forecasts



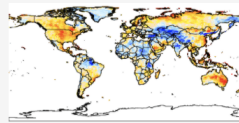
#### Temperature Weekly Probability Forecast (LELR)

Subseasonal tercile categories temperature forecasts with pattern-based calibration



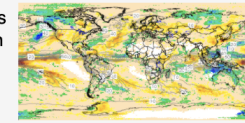
#### Temperature Median Probability Forecast

Calibrated Subseasonal Two-category temperature real-time forecasts



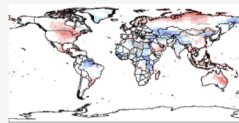
#### Precipitation Flexible Biweekly Forecast

This subseasonal forecasting system consists of probabilistic precipitation forecasts based on the full estimate of the probability distribution.



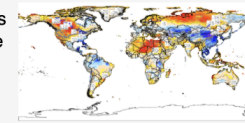
#### Temperature Biweekly Probability Forecast

Calibrated subseasonal tercile categories temperature forecasts



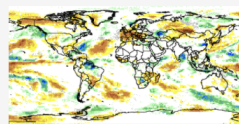
#### Temperature Flexible Biweekly Forecast

This subseasonal forecasting system consists of probabilistic temperature forecasts based on the full estimate of the probability distribution.



#### Precipitation Weekly Probability Forecast

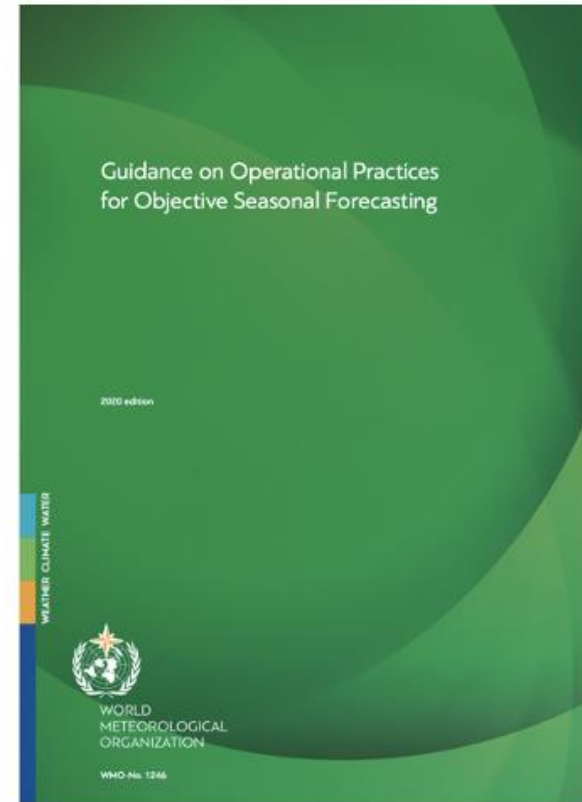
Calibrated Subseasonal Tercile categories precipitation real-time forecasts

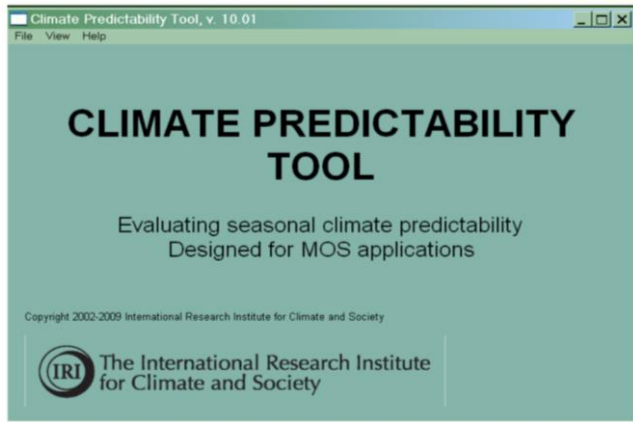


# Next Generation (NextGen) Regional Forecasting

The use of an objective seasonal forecast procedure which is defined as a traceable, reproducible, and well-documented set of steps that allows the quantification of forecast quality, are preferred and recommended by the **World Meteorological Organization in their recent seasonal forecast guidance.**

The Next Generation (NextGen) seasonal forecast system is a systematic and **objective approach.** It enables calibration, combination, and verification of objective climate forecasts from the state-of-the-art general circulation models (GCM) of the North American Multi-Model Ensemble project.





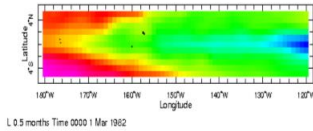
- Climate Predictability tool (CPT) is an easy-to-use software for making seasonal forecast using either empirical predictors, of the outputs from GCM.
- Developed and maintain by **Dr. Simon Mason**.
- CPT available for Windows 95+ and Linux Batch version.

Predictor (X)

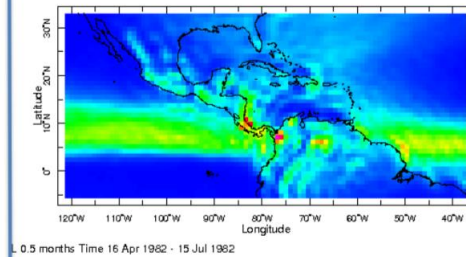
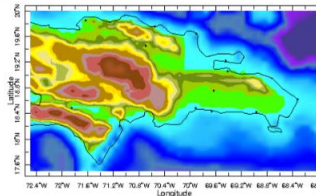
Predictand (Y)

Forecast

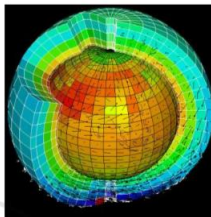
Empirical



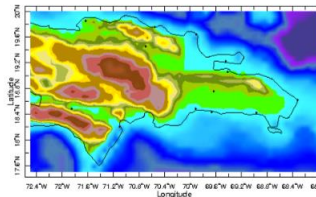
Observations  
Observations



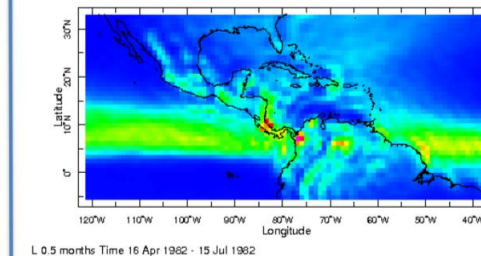
Dynamical



Model output

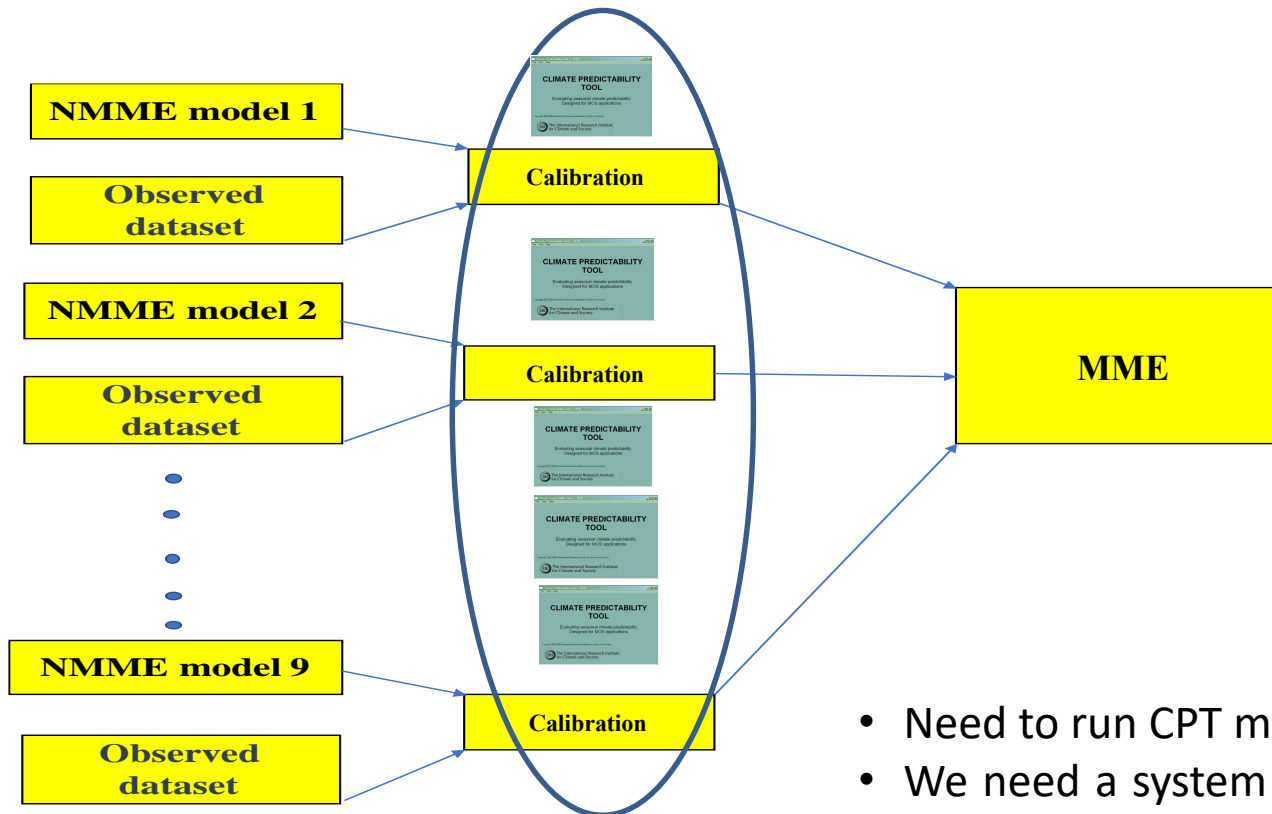


Observations





# NextGen Approach

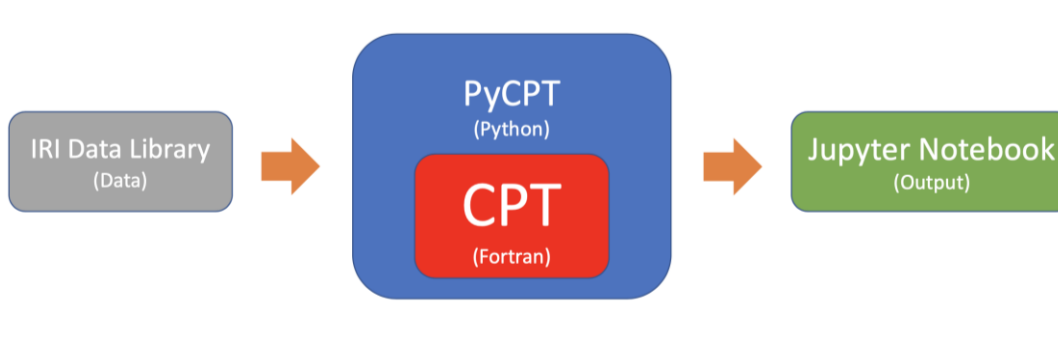


- Need to run CPT multiple times.
- We need a system which produce skill maps and forecasts for multiple models in a single run.

# PyCPT: Python script to run CPT batch version

PyCPT is a **Python library that provides an interface and extra functionalities** to IRI's Climate Predictability Tool (CPT), a widely used research and application Model Output Statistics/Prediction toolbox.

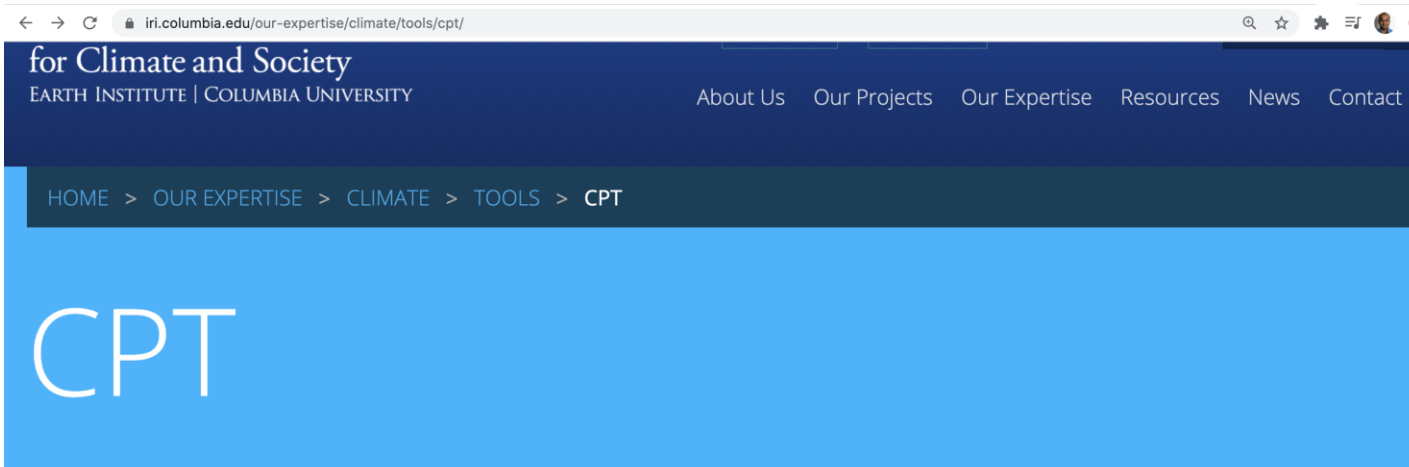
## PyCPT Layout



## PyCPT Structure



# PyCPT: Download



## The Climate Predictability Tool

The Climate Predictability Tool (CPT) is a software package for constructing a seasonal climate forecast model, performing model validation, and producing forecasts given updated data. Its design has been tailored for producing seasonal climate forecasts using model output statistics (MOS) corrections to climate



### Important Links

#### CPT Downloads

[Download Latest Version \(2020-09-08\)](#)

[Download Version 15.7.11 \(2019-07-26\)](#)

[PyCPT](#)

#### Release Notes

<https://bitbucket.org/py-iri/iri-pycpt/src/master/>



# PyCPT: Installation

## ➤ For Linux users:

- The user will need to install Anaconda (Python3), the Climate Predictability Tool (batch version) and the Python extension of CPT (PyCPT).

## ➤ For Window users:

- install a Virtual Machine with all needed packages and use PyCPT Ubuntu
- Any recent (< 2 years) Intel Processor should be able to run a Virtual Machine. The PC should have at least 4GB of RAM installed, but preferably more than 8GB You should have at least 20GB of free space to install the virtual machine and software.

For details: <https://bitbucket.org/py-iri/iri-pycpt/wiki/Home>



Live Demo!





# #NextGen Maproom: Example of Meteo Rwanda

The screenshot shows the homepage of the Rwanda Meteorology Agency website. The URL in the browser is [meteorwanda.gov.rw/index.php?id=2](http://meteorwanda.gov.rw/index.php?id=2). The page features a navigation menu with the following items: Home, About Us, Products & Services, Weather Forecast, Warning Services, **Maproom** (circled in red), Online services, and Feedback. Below the navigation menu, there are two main content areas: 'Today weather forecast' and 'Latest News'. The 'Today weather forecast' section includes a map of Rwanda with various districts labeled, such as Musanze, Burera, Nyagatare, Gicumbi, Gatsibo, Rubavu, Nyanza, and others. The 'Latest News' section features a map of East Africa with a legend for 'EA\_MAM contribution' and a title: 'Understanding the Evolution and Socio-Economic Impacts of the Extreme Rainfall Events in March-May 2017 to 2020 in East Africa'.

[http://maproom.meteorwanda.gov.rw/maproom/Climatology/Climate\\_Forecast/Forecast.html](http://maproom.meteorwanda.gov.rw/maproom/Climatology/Climate_Forecast/Forecast.html)

# #NextGen Maproom: Example of Meteorology in Rwanda

Not Secure | maproom.meteorwanda.gov.rw/maproom/Climatology/Climate\_Forecast/Forecast.html?region=bb%3A29.5%3A-2%3A29.75%3A-1.75%3Abb

**Precipitation Flexible Seasonal NextGen Forecast**

This Maproom shows the next 3 month forecast using a NextGen multi-model approach.

NextGen is a systematic general approach for designing, implementing, producing and verifying objective climate forecasts. It involves the identification of decision-relevant variables by the stakeholders; the analysis of the physical mechanisms, sources of predictability and suitable candidate predictors (in models and observations) for those key relevant variables. In those cases when prediction skill is high enough, NextGen helps select the best dynamical models for the region of interest through a process-based evaluation, and automatizes the generation and verification of tailored multi-model, statistically calibrated predictions at seasonal and sub-seasonal timescales.

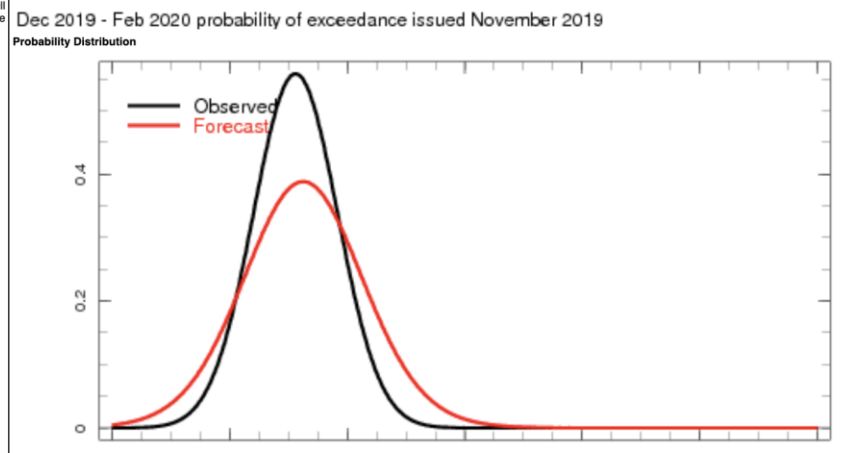
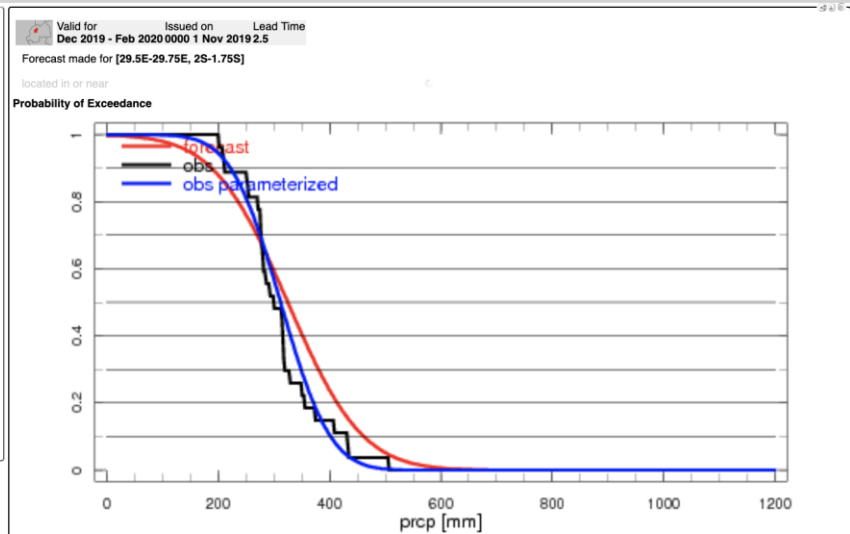
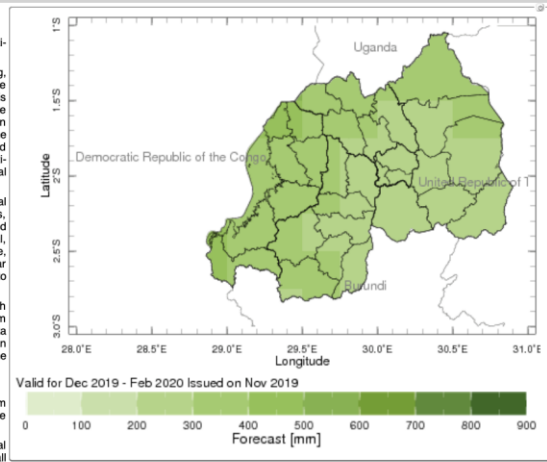
The system takes advantage of the expertise of forecasters and local scientists at the country's national meteorological service and universities, to maximize predictive skill and tailoring of the climate services generated by the process. Rather than focusing on probabilities of above normal, normal and below normal categories of total rainfall or mean temperature, NextGen also provides probabilities of exceeding (or not) particular thresholds of interest in the decision-making process, thus enabling users to forecast with the same system both mean and extreme values.

The models employed in this forecast are from the suite of the North American Multi-Model Ensemble (NMME) and the predictand is rainfall from the Climate Hazards Group InfraRed Precipitation with Station Data (CHIRPS). The default map shows, for the latest forecast made, the median value of the seasonal rainfall total forecast in the season. Users can use the Field menu to express the forecast in different ways, as follows:

- **Rainfall:** most likely seasonal total rainfall
- **Anomaly:** deviation in mm of the most likely seasonal total rainfall from yearly average of the most likely seasonal total rainfall predicted by the hindcast (1982-2009)
- **Percent of Median:** deviation in percentages of the most likely seasonal total rainfall from yearly median of the most likely seasonal total rainfall predicted by the hindcast (1982-2009)
- **Probability of non-exceeding a Percentile:** forecast probability of seasonal total rainfall to be below/above the historically observed (1982-2009) chosen percentile
- **Probability of non-exceeding a Precipitation amount:** forecast probability of seasonal total rainfall to be below/above the chosen rainfall amount

The Layers button, showing when mousing over the map, will reveal inactive layers on the map.

Clicking on the map will reveal information about the location clicked, as well as the full forecast distribution at that given location, compared with the historical distribution. Cumulative full distribution of the forecast (red) together with the climatological distribution (blue and black) for the forecast in view on the map shows under **Probability of Exceedance**, as well as the full probability distribution under **Probability Distribution**.



# Challenges in producing forecast

- Non-availability of GCMs in real time.
- GCM's version changes.
- Need update PyCPT scripts.

Models	hindcast	forecast
CMC1- CanCM3	1981-2010	Jan2011- current
CMC2- CanCM4	1981-2010	Jan2011- current
NCEP-CFSv2	1982-2010	March/Apr 2011- current
NCAR- CESM1	1980-2010	July-2016- April-2017
COLA- RSMAS- CCSM4	1982-2010	2011- current
NASA- GMAO-	1981-2010	2011- Jan2018
NASA- GEOSS2S	Feb 1981 - Jan 2017	Nov 2017- current
GFDL(3)	1982-2010	2011- current

# Concluding Remark

- ❖ ELR based non-Gaussian calibration method introduced in the real-time seasonal forecast at IRI.
- ❖ It is a more robust method compared to other calibration method based on the Gaussian assumption for precipitation.
- ❖ For regional forecast, NextGen system is introduced by IRI by the recommendation of WMO.
- ❖ PyCPT is the tool for NextGen where CPT can run multiples times in a single run.
- ❖ Very easy to use through Jupyter notebook.
- ❖ PyCPT keep on updating based on user's feedback and GCM availability.



**Thanks!**

[nachiketa@iri.Columbia.edu](mailto:nachiketa@iri.Columbia.edu)