

INTRODUCTION TO PYCPT

31ST AUGUST 2021



RESEARCH PROGRAM ON
**Climate Change,
Agriculture and
Food Security**



INTRODUCTION TO PYCPT

**Objective Climate Forecasts for Agriculture and Food Security
Sector in Eastern and Southern Africa Training Workshop,
Victoria Falls, Zimbabwe**

1st September 2021

WHAT IS CLIMATE PREDICTABILITY TOOL (CPT)?

- The Climate Predictability Tool (CPT) is a statistical tool developed by the IRI for;
 - constructing a seasonal climate forecast model (CCA, PCR, MLR)
 - performing model validation (retroactive and cross-validation method),
 - and producing forecasts given updated data (probabilistic and deterministic).
- It is designed for producing seasonal climate forecasts using model output statistic (MOS) corrections to climate predictions from general circulation model (GCM)
- It is also designed for producing forecasts using fields of sea-surface temperatures or similar predictors.

WHAT IS PYCPT?

- PyCPT is a Python library that provides interface and extra functionalities to IRI's Climate Predictability Tool (CPT),
- It is specially designed for mass-production of seasonal and sub-seasonal forecast, skill assessment maps, and probabilistic flexible forecasts.
- PyCPT can be installed in a local computer for both Windows and Linux OS.
- To install in a windows machine, Virtual Machine should be installed first

PYCPT ESSENTIAL SCRIPTS

- **Pycpt_functions_seasonal.py**: Main python code that contains all functions needed to process the forecasts among other outputs.
- **Pycpt_dictionary.py** : python script that holds the URL links to all your datasets: Model datasets are obtained from the IRI data library.
- The primary script **PyCPT_seav1.9.py**- this is the main script that you will use to run the PyCPT.

STRUCTURE OF PYCPT



Most of the configuration of your particular case happens in the namelist section, found at the beginning of the PyCPT code. This section enables the user to select predictor and predictand datasets, calibration methods and spatial and temporal domains among other parameters

The download and CPT execution section deals with preparing all the needed input datasets and running CPT to produce skill assessment and forecast files, conducting the calibration process selected by the user.

The skill assessment section visualizes predictive skill metrics as maps or text. Skill measures such as Pearson, ROC scores, Spearman, etc

This forecast section deals with the ensemble generation and production of forecast maps and other figures, with a special focus on flexible formats (the use of the entire forecast probability density function)

IMPORTANT DIRECTORIES UNDER 'WORK' DIRECTORY

Input: This is the directory that holds the downloaded data: model hindcast, forecast, and observation data. The data files have **.tsv** suffix

Output: This directory contains the output text files (used in objective forecast), **ctl** and **dat** files (output from CPT processing)

Figures: This directory contains all the output maps(forecast (probabilistic & deterministic), skill, flex format)

Scripts: This directory contains CPT log files...Helps trace where error occurred if CPT processing fails

➤ Also contains CPT job files.

LOG IN TO YOUR COUNTRY ACCOUNT

1. Username: ssh -X training@197.254.1.14

Password : *****

2. Change directory to SADC/country/pycpt/code/ then list;

```
[training@master code]$ ls -lrt
total 376
-rw-r--r-- 1 training modelers 192565 Aug 28 15:02 pycpt_dictionary.py
-rw-r--r-- 1 training modelers 151757 Aug 28 15:02 pycpt_functions_seasonal.py
drwxr-xr-x 2 training modelers   104 Aug 28 15:06 __pycache__
-rw-r--r-- 1 training modelers  13660 Aug 28 15:46 PyCPT_seav1.9.py
-rw-r--r-- 1 training modelers  13661 Aug 28 15:55 PyCPT_seav1.9monthly.py
```

Main script for
seasonal forecast

3. Vi PyCPT_seav1.9.py

STEP-BY-STEP PROCEDURE WITH PYCPT

Namelist

Step1: Create your working directory:

- work = './Aug2021'

Step2: select models (Available models have different hindcast periods, so process them separately); For this exercise, we will use

- Models: COLA-RSMAS-CCSM4, NCEP-CFSv2; NASA-GEOSS

Step3: Select Observation, MOS method, Predictand & Predictor

1. Obs (options are CPC-CMAP-URD, CHIRPS, TRMM, CPC, Chilestation). For this exercise, we will use;

Obs='CHIRPS'

Station=False

2. MOS method (None, PCR, CCA)

MOS='CCA'

3. Predictand (PRCP, RFREQ)

PREDICTAND='PRCP'

4. Predictor(GCM's PRCP,VQ,UQ)

PREDICTOR='PRCP'

Step4: Set target seasons and Related parameters

1. Set **initialization month** (for more than one seasons, add the list)

Mons= ['Aug' , 'Aug' , 'Aug']

2. Set the **Data Library index** for the first and last month of your target season

Tgtii = ['0.5', '1.5', '2.5'] - First month of target season

Tgtff = ['2.5', '3.5', '4.5'] – Last month of the target season

3. Indicate the **target seasons** corresponding to seasons in Step 3 (2)

tgts = ['Aug-Oct' , 'Sep-Nov' , 'Oct-Dec']

4. Set the training period (this will double up as the climatological period)

❖ `tini = 1982`

`tend = 2020`

5. Choose **number of Modes** (Do not change for now)

`xmodes_min = 1`

`xmodes_max = 10`

`ymodes_min = 1`

`ymodes_max = 10`

`ccamodes_min = 1`

`ccamodes_max = 10`

STEP5: SET FORECAST DATE (INIT. MONTH & YEAR)

`monf = 'Aug'`

`fyr = 2021`

STEP6: SWITCHES: DOWNLOAD DATA FILES EVEN WHEN ALREADY AVAILABLE?

Force download= True

STEP7: SET PREDICTOR & PREDICTAND DOMAIN (do not change for now)

1. Spatial domain for predictor

GHA	SADC
nla1=30	10 # Northernmost latitude
sla1=-30	-40 # Southernmost latitude
wlo1=0	-180 # Westernmost longitude
elo1=359	180 # Easternmost longitude

2. Spatial domain for predictand

GHA	SADC
nla2=23	-9 # Northernmost latitude
sla2=-12	-29 # Southernmost latitude
wlo2=20	11 # Westernmost longitude
elo2=52	41 # Easternmost longitude

STEP8: SET REQUIRED DIRECTORIES' PATH

1. Set the path of **working directory**

Workdir = '/home/training/**name**/pycpt/code/'

2. Set path to **CPT directory** (Don't change for now)

Cptdir='/home/forecast/modelLibs/LIBRARIES/CPT/16.5.9/'

STEP9: SKILL ANALYSIS

1. Select metrics of choice and season of interest

Met=('Pearson', 'RoCAbove')

Monss=['Aug-Oct', 'Sep-Nov', 'Oct-Dec']

STEP10: NEXTGEN: MULTI-MODEL ENSEMBLE & SKILL ASSESSMENT

models=['NCEP-CFSV2', 'COLA-RSMAS-CCSM4, NASA-GEOSS"] # Selected models should be part of, or all of those selected in **STEP 2**

Model= 'NextGen'

MOS= 'None'

STEP 11: SELECT SKILL METRICS & SEASON FOR NEXTGEN

```
Met=('Pearson','RocAbove')  
Monss=['Aug-Oct', 'Sep-Nov', 'Oct-Dec']  
Model= 'NextGen'
```

STEP 12: FLEXIBLE FORECAST (SPATIAL)

```
flexformat_threshold = 100 #threshold for the flexible format (only one target season)  
ffthres_ispctl= False      # #If True, threshold is a percentile: 0.001 - 0.999
```

STEP 13: FLEXIBLE FORECAST (POINT LOCATION)

```
flexformat_threshold = 160 #threshold for the flexible format (only one target season)  
ffthres_ispctl=False  
lon=37.  
lat=5.
```

STEP 14: RUN THE SCRIPT

To run the script, we need to activate the pycpt environment:
From your working directory, type

Conda activate pycpt_env

Now run the script:

```
ipython Pycpt_seasv1.9.py      #
```

The processing will take time depending on the number of models and seasons. Text outputs required for the objective forecast go to 'work'/ouputs/

THANK YOU VERY MUCH!

