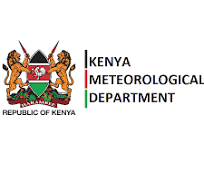
  ****

**Capacity Building Workshop for Kenya Meteorological Department Forecasters on Seasonal Forecasting using Next-Generation Forecasting System**

**21-25 MARCH 2022, MALINDI, KENYA**



Contents

[**List of ACRONYMS** 3](#_Toc163573112)

[**1.** **Introduction** 4](#_Toc163573113)

[**2.** **Opening Session** 4](#_Toc163573114)

[**3.** **Introduction to Python** 5](#_Toc163573115)

[**4.** **Installation of PyCPT** 5](#_Toc163573116)

[**5.** **Presentations and Discussions** 6](#_Toc163573117)

[**5.1** **Seasonal forecasting in KMD** 6](#_Toc163573118)

[**5.2** **Understanding Objective Forecasting and KMD Preparedness** 7](#_Toc163573119)

[**6.** **Introduction to Next Generation Forecasting System** 8](#_Toc163573120)

[**6.1** **Skill Metrics used in PyCPT** 9](#_Toc163573121)

[**7.** **Introduction to PYCPT** 9](#_Toc163573122)

[**7.1** **Generating a Seasonal Forecast using CCA** 10](#_Toc163573123)

[**7.2** **Assessing the Impact of Domain Size on Forecasts** 12](#_Toc163573124)

[**8.** **Generating a Seasonal Forecast using Principle Component Regression (PCR)** 17](#_Toc163573125)

[**9.** **Conclusion and Closing Remarks** 19](#_Toc163573126)

**List of ACRONYMS**

**IGAD** Inter-Governmental Authority on Development

**ICPAC** IGAD Climate Prediction and Applications Centre

**NMHSs** National Meteorological and Hydrological Services

**KMD** Kenya Meteorological Department

**CPT** Climate Predictability Tool

**PyCPT** Python based Climate Predictability Tool

**RCCs** Regional Climate Centres

**GHA** Greater Horn of Africa

**GHACOF** Greater Horn of Africa Climate Outlook Forum

# **Introduction**

IGAD Climate Prediction and Applications Centre (ICPAC) received funding from the European Union (EU) under the Intra-ACP Climate Services and Related Application (ClimSA) Project to improve climate services at regional and national levels for select countries (Kenya and Uganda) within the Greater Horn of Africa (GHA). One of the project’s outputs is to build the capacity of National Meteorological and Hydrological Services (NMHSs) in aspects related to climate services. To achieve this output for Kenya, a 1-week capacity building training of the Kenya Meteorological Department (KMD) forecasters was done at Diamonds Dream of Africa hotel in Malindi, Kenya.

The training involved 7 forecasters from the KMD drawn from Daily to seasonal range. The participants were 4 females and 3 males. The main objective of the Workshop was to provide both theoretical and practical skills towards generating high-skill seasonal climate forecasts, using the “NextGen” technique, which is one of the methods used in generating ICPAC’s objective operational forecast following recommendation by the World Meteorological organization (WMO) in its steering committee meeting in 2017. The specific objectives were;

* Installing PyCPT in personal computers
* Introduce the use of Python based Climate Predictability Tool in climate forecasting using Windows & Linux computers
* Explain the steps in configuring PyCPT for optimal forecasts, forecast verification and producing flexible forecasts
* Use PyCPT in forecasting sub-seasonal and seasonal climate over Kenya

# **Opening Session**

The training was officially opened by Mr. Paul Ombai, who represented the director. He gave a brief history on the evolution of seasonal forecasting at ICPAC from consensus to objective. He noted that the move to objective forecasting was a recommendation from the WMO to have the RCCs issue forecasts that are traceable, reproducible and verifiable. He re-iterated that ICPAC is committed to its mandate on capacity building of its member states through training workshops, infrastructure, attachments, and academic scholarships. He pointed out that currently, there are 22 students in different institutions of learning that are pursuing various climate related courses. He urged the participants to take the training seriously as it is a great opportunity to build their skills and knowledge in objective forecasting procedures developed by ICPAC. He then wished the participants a fruitful training throughout the week.

# **Introduction to Python**

PyCPT is a python-based version of the Climate Predictability Tool (CPT), and this requires that the participants got a basic understanding of the language. A simple survey showed that the participants had little or no skill on the language. To set the scene, the participants were taken through a practical session on how to manipulate and analyze various types of datasets including .tsv and netcdf using Python. Since most of them were using Windows laptops, the first activity was to install a Linux environment. They were guided on how to install the Anaconda Navigator which would allow then to utilize the Jupyter notebook. The choice of using Jupyter notebook was because the main PyCPT script is written on jupyter notebooks. They were introduced to the basic python libraries that are required to organize data and make simple line plots. Some of the libraries include pandas, statistics, numpy, matplotlib.pyplot. The specific tasks that they were able to perform include assigning missing numbers in a dataframe, calculating annual means, concatenating datasets, plotting and saving graphs. Sample temperature map that the participants were able to plot is shown below.

# **Installation of PyCPT**

PyCPT uses the Linux version of the CPT. The participants were directed to the IRI website where they were able to download the 16.5.8 version of CPT. They were then helped on how to compile it in Linux environment. They were also able to download and install Anaconda for Linux from the Anaconda website. Other essential python libraries that were not installed in the first session (introduction to python) such as xarray, scipy, cartopy, netcdf4 were installed in this session. The participants were also able to clone PYCPT from the IRI website into their local computers. The Step by Step installation instructions document customized for the training was available for participants to make reference.

# **Presentations and Discussions**

## **Seasonal forecasting in KMD**

To get an understanding on how seasonal forecasting is done at the Kenya Meteorological Department (KMD), a detailed presentation was done. It was evident that the seasonal forecasting starts during the capacity building and climate prediction development workshop (popularly known as Pre-GHACOF) that is hosted by the IGAD Climate Prediction and Applications Centre (ICPAC). During the workshop, a national forecast is produced using the objective techniques developed by ICPAC. After the workshop, the forecast is then subjected to a critical analysis by a team of climate scientists from the meteorologicaldepartment during pre-NCOF sessions. The process involves evaluation of the GHACOF statement, previous season’s review (temperature and rainfall), the team critics/improves upcoming seasonal forecast (including SPI)​. This is done through analysing analogue years chosen, querying daily data for the different homogeneous zones​, providing an indication of expected onset, distribution and cessation​ and then downscaling further to homogeneous zones.

The improved forecast is then released during NCOF, a national user interface platform, which bring together different stakeholders. The NCOF session involve a co-production workshop with sectors, generation of sector advisories, and dissemination of forecast/final product. Supplementary seasonal forecasts are issued to;

* Ministry of Agriculture​
* KenGen​
* Kenta Red Cross Society (KRCS)​
* National Drought Management Authority (NDMA)​
* Ministry of Environment​
* Ministry of Education

It was noted that after the forecast has been issued to the public, continuous monitoring is done and updates made when necessary.

## **Understanding Objective Forecasting and KMD Preparedness**

This was a group discussion. The participants were divided into three groups centered on data, technical capacity, Internet and computing. ICPAC staff joined the three groups and below are the summary from the three groups and general take away message

**Summary Notes**

* There’s data availability in terms of access and this can support objective forecasting. For more access of data to aid in capacity development at KMD with ICPAC, there is need to review the MoU on Data sharing between KMD-ICPAC
* KMD computing facilities (e.g HPC) is currently under-utilized, this if operationalized can provide huge resources to support staff in enhancing the understanding and full adoption of objective forecasting.
* The department is ready to adopt the objective forecasting in its operations.
* Boosting Internet connectivity is needed for better forecasting.
* Time allocated for the pre-COF capacity building is too short. It takes on average 3 consecutive seasons to fully master the process
* Confidence in forecasts needs to be raised by providing skill assessment of consensus and objective forecasting to support the basis for the shift.
* The institution has two experienced and trained forecasters and two semi-trained that needs more training.

**Key message**

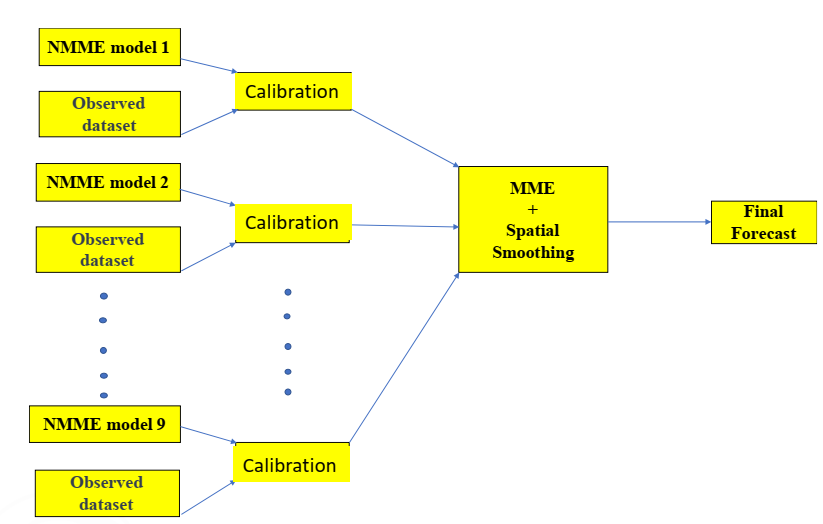
* Collaboration between ICPAC-KMD and other NHMSs on capacity development needs to continue; with a focus on basic computing skills during the first week of Foundational training
* Department needs to work on knowledge retention and transfer of its staff
* ICPAC to introduce an online self-training course on objective forecasting procedures, basic skills on data analysis and modelling and provide necessary materials (for individuals to learn at their own pace)
* ICPAC to institutionalize the Foundational training
* Model assessment and skill improvement should be based on Realtime forecasts and not over-reliance on hindcast skill assessment

# **Introduction to Next Generation Forecasting System**

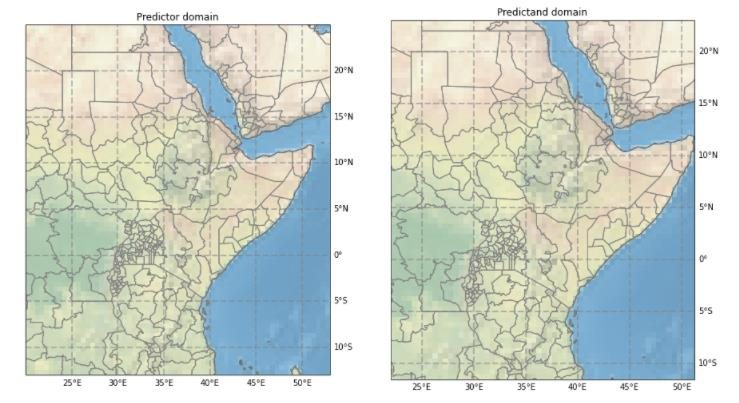
The participants were introduced to the NextGen forecasting system, which provides an opportunity for co-designing, implementing, producing and verifying objective forecasts at multiple timescales. It was pointed out that NextGen provides potential for improved forecasts at regional and national level due to availability of global ensemble forecast system products. It was also noted that the need to adopt the use of NextGen in forecasting is necessitated by the fact that many NMHSs in developing countries have not adopted the use of objective forecasting approaches and that they tend to use SSTs in the Pacific as the main predictor. Their use of empirical forecast approaches is limited by the fact that atmosphere-ocean coupling is normally not considered. A detailed explanation on how the NextGen technique using the CCA MOS works was done and the details are depicted by the diagram below.

**Figure 1: NextGen Forecasting System using CCA calibration**

**Figure 1: NextGen Forecasting System using CCA Calibration Method**



**Canonical Correlation Analysis**



Predictor

Domain

Predictand Domain

Calibration Procedure

 ​

## **Skill Metrics used in PyCPT**

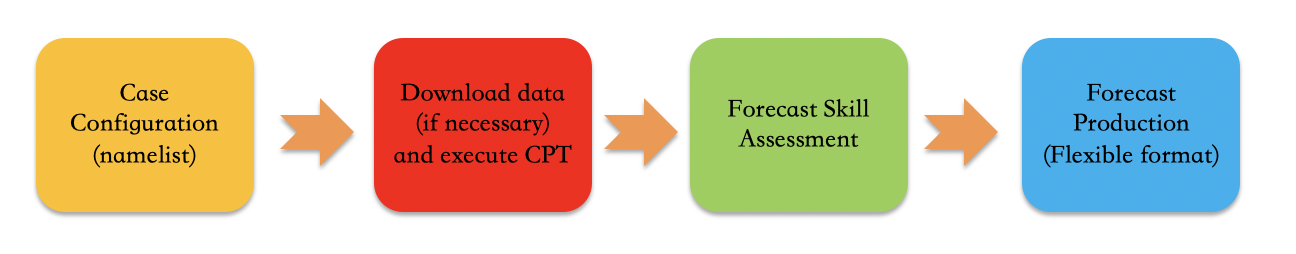
PyCPT automatically analyze model skills using metrics such as Pearson correlation, Spearman correlation, Ignorance, Ranked Probability Skill Score (RPSS), ROC (GROC, ROC above, ROC Below). Detailed explanation on Spearman and Pearson skill metrics was done. Operational model hindcast skill assessment at ICPAC is done using the Pearson correlation. Other skill scores that were discussed are Anomaly correlation coefficient and Root Mean Square Error.

# **Introduction to PYCPT**

Having gone through a session on the basics of Python and an introduction to NextGen forecasting systems, the participants were now introduced to the PyCPT. A brief explanation on the working principles of CPT was done and the advantages of PyCPT over the windows version of CPT was highlighted. These pros include;

* Automatic downloading of datasets
* Automatic calibration and assessment of model skill
* Automatic forecast generation
* Automatic model ensembling/NextGen

The essential scripts used in PyCPT (Pycpt\_functions\_seasonal.py, Pycpt\_dictionary.py, and PyCPT\_seav1.9.2.ipynb) were also explained and participants were able to understand the source of the data being used and how it is accessed. The structure of PyCPT as shown in the diagram below was also explained in detail.



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

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, spearman, RocAbove, RocBelow, RPSS etc

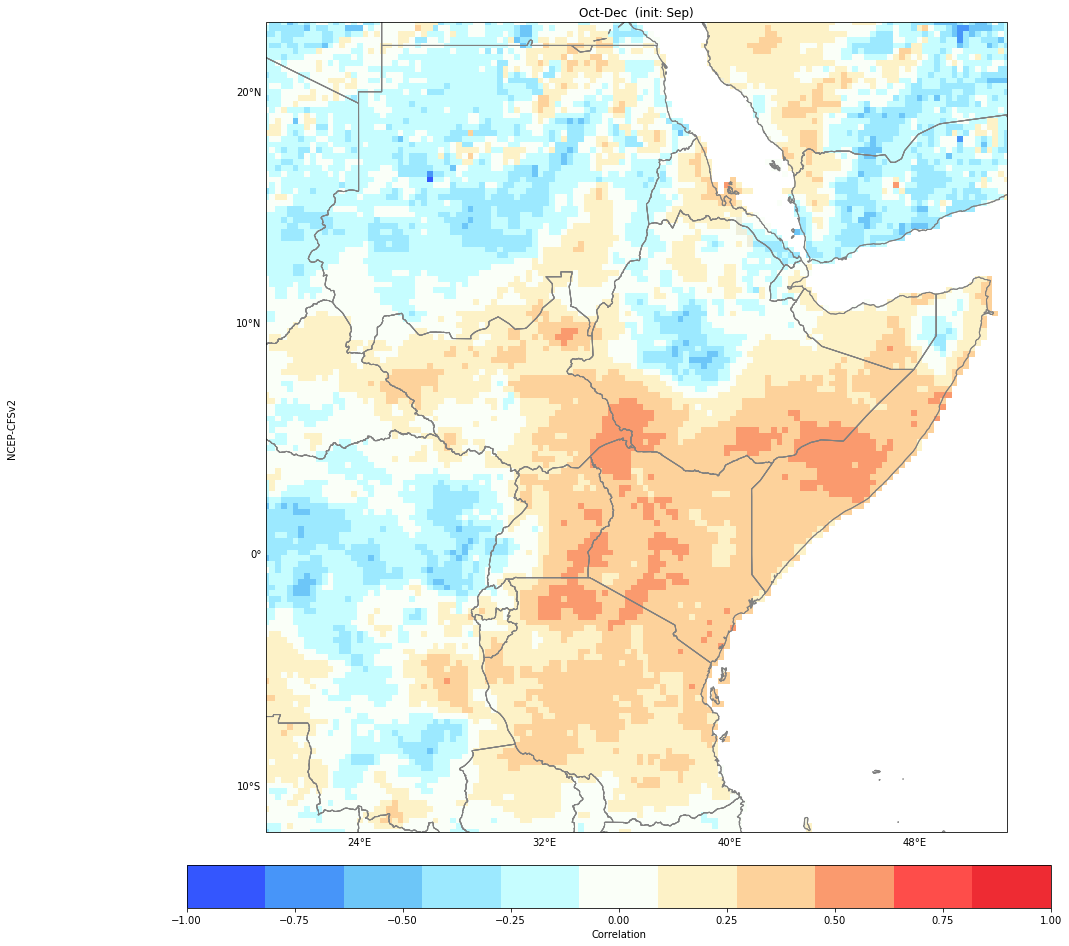
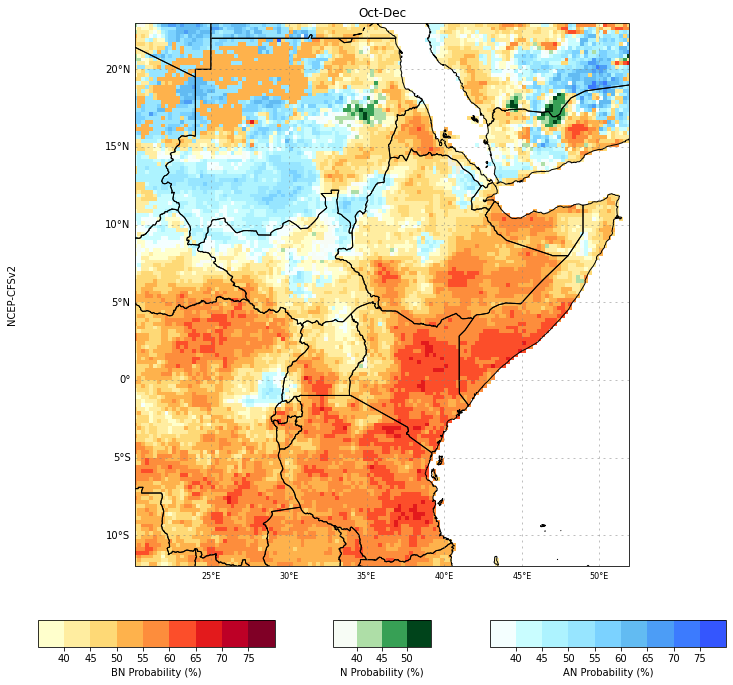
This forecast section deals with the ensemble generation and production of forecast maps and other figures, with a special focus on flexible formats (the use of the entire forecast probability density function).

## **Generating a Seasonal Forecast using CCA**

The participants were taken through the step-by-step process of generating a forecast using PyCPT. They were asked to launch the Jupyter notebook and open the PyCPT\_seav1.9.2.ipynb script.  At the namelist section, they were able to make case configurations such as creating their working directory, choosing calibration method (CCA, PCR or None), identify predictors and predictand variables and their domains, select season of interest and the respective initialization month, training/climatological period, and the models to be calibrated. All the participants were expected to use the global tropics predictor and the GHA predictand domains to make a forecast for GHA. In the second section of the PyCPT structure, the participants were able to download the model outputs data (hindcast and forecast) and the observation data (CHIRPS) and were able to locate the outputs of the calibrated models in the output folder. They were also able to plot and understand the EOFs, PCs, CCA modes. In the third section, visualization of various model skills (for 2 models; NCEP-CFSv2 and NASA-GEOSS2S) such as Pearson, Spearman, RoC Above/Below, GROC, RPSS was done. A discussion on how to interpret these skills was done. Lastly, the participants were able to plot individual model forecast, the NextGen forecasts, which refers to multi-model ensemble, multi-model skill, and the flexible format forecast at national and single point scales.  Sample results from the participants are shown below.

**Global tropics predictor domain with GHA predictand domain for OND 2021**





**Pearson Correlation**

**Model Forecast**

**Figure 2: Predictor Domain (top) and OND 2021 skill and forecasts using Global predictor Domain and GHA predictand domain**

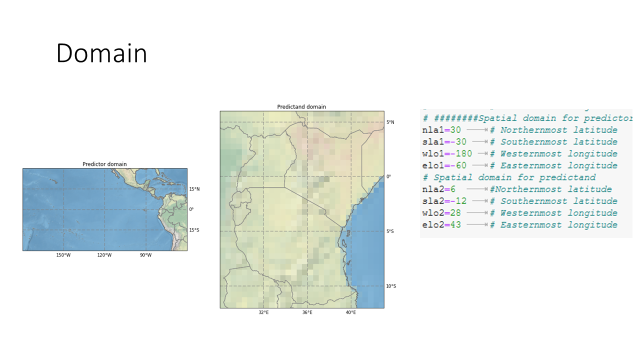
## **Assessing the Impact of Domain Size on Forecasts**

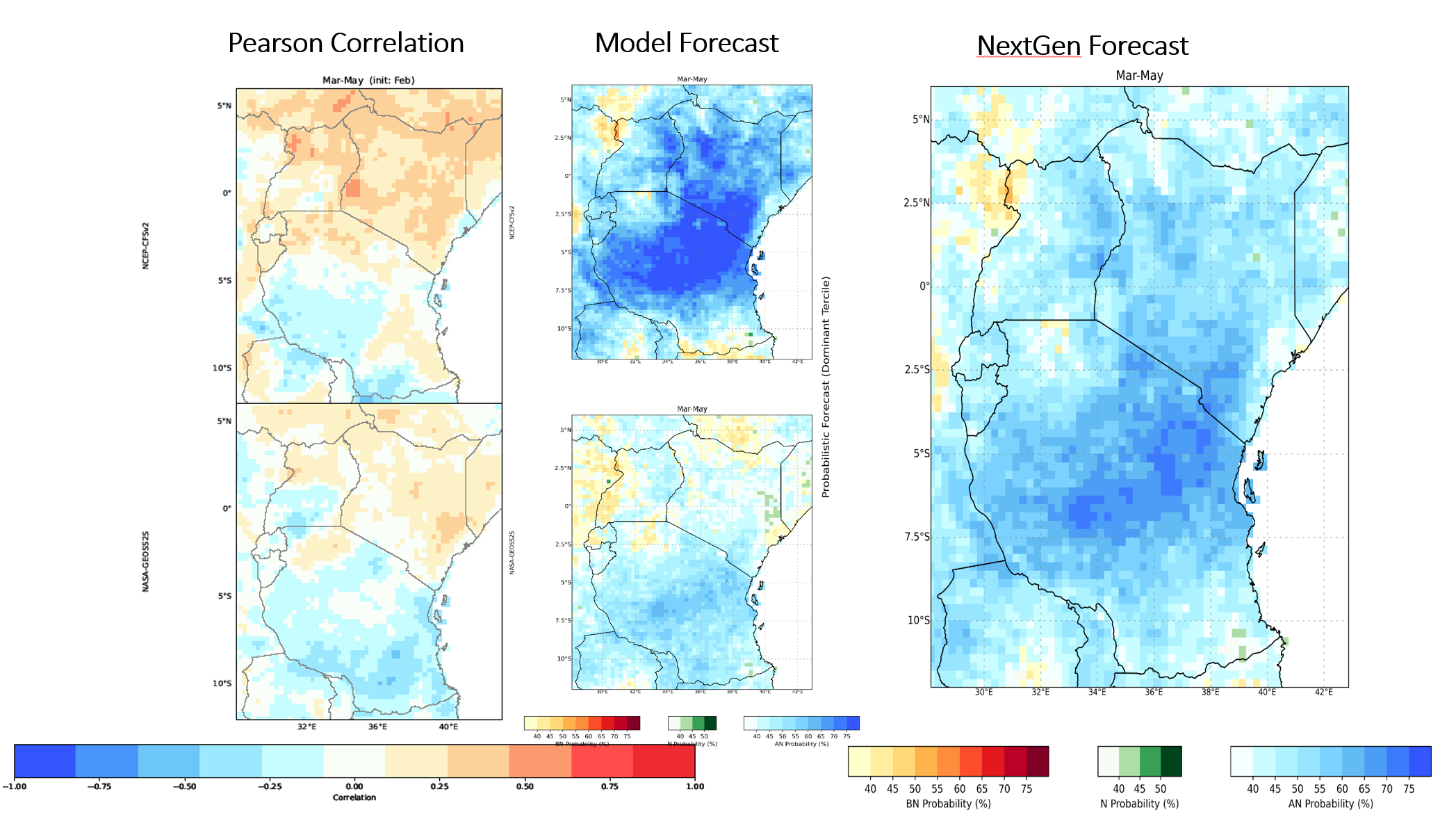
Canonical Correlation Analysis involves calculation of PCs, which makes the domain size an important factor in determining a forecast. To get more understanding on how different predictor domains would influence the forecast, the participants were divided into two groups. The first group were assigned to analyze the MAM 2021 season while the second group was assigned to analyze the OND 2021 season forecasts.

The main task involved using the CCA Model Output Statistics Method to calibrate NCEP-CFSv2 and NASA-GEOSS2S models. Three sets of experiments were to be done with predictor domains as the global tropics, Indian Ocean, Pacific, and any other 2 of their choice and the predictand domain set to Kenya, East Africa, and part of East Africa. The groups were expected to present their individual model skills and forecasts and the NextGen (multi-model ensemble) skill and forecast. ​A lengthy discussion on the observed differences was held and sample results are shown below.

**MAM 2021 Results**

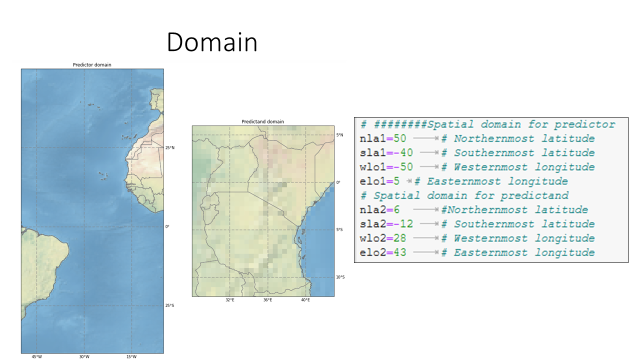
**Predictor Domain: Pacific Predictand Domain: East Africa**

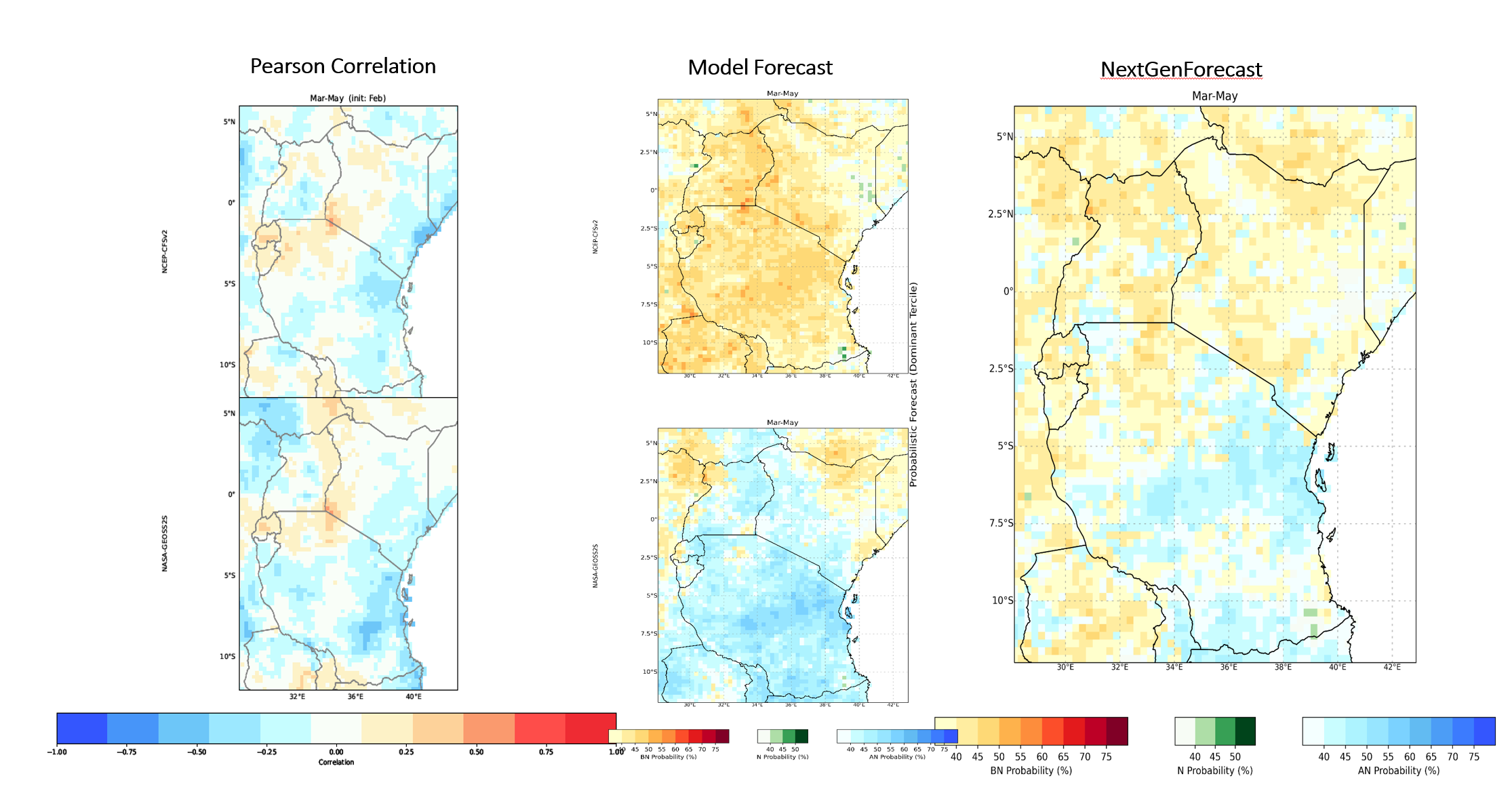
****

****

**Figure 3: Predictor Domain (top) and MAM 2021 skill and forecasts using Pacific predictor Domain**

**Predictor Domain: Atlantic Predictand Domain: East Africa**



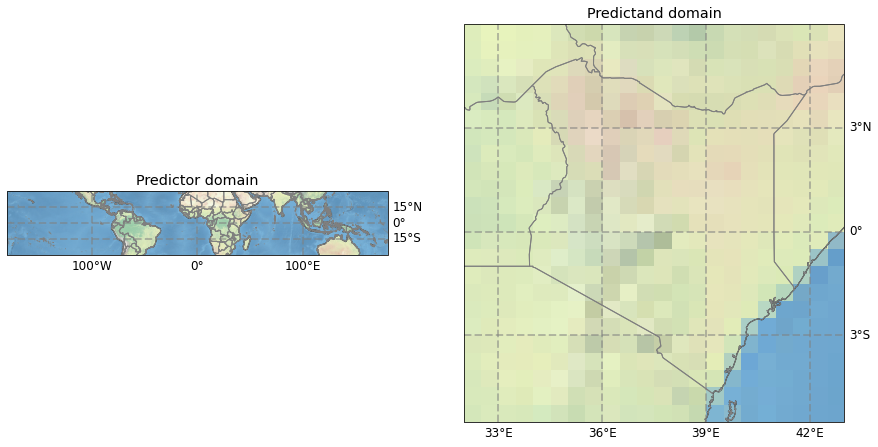


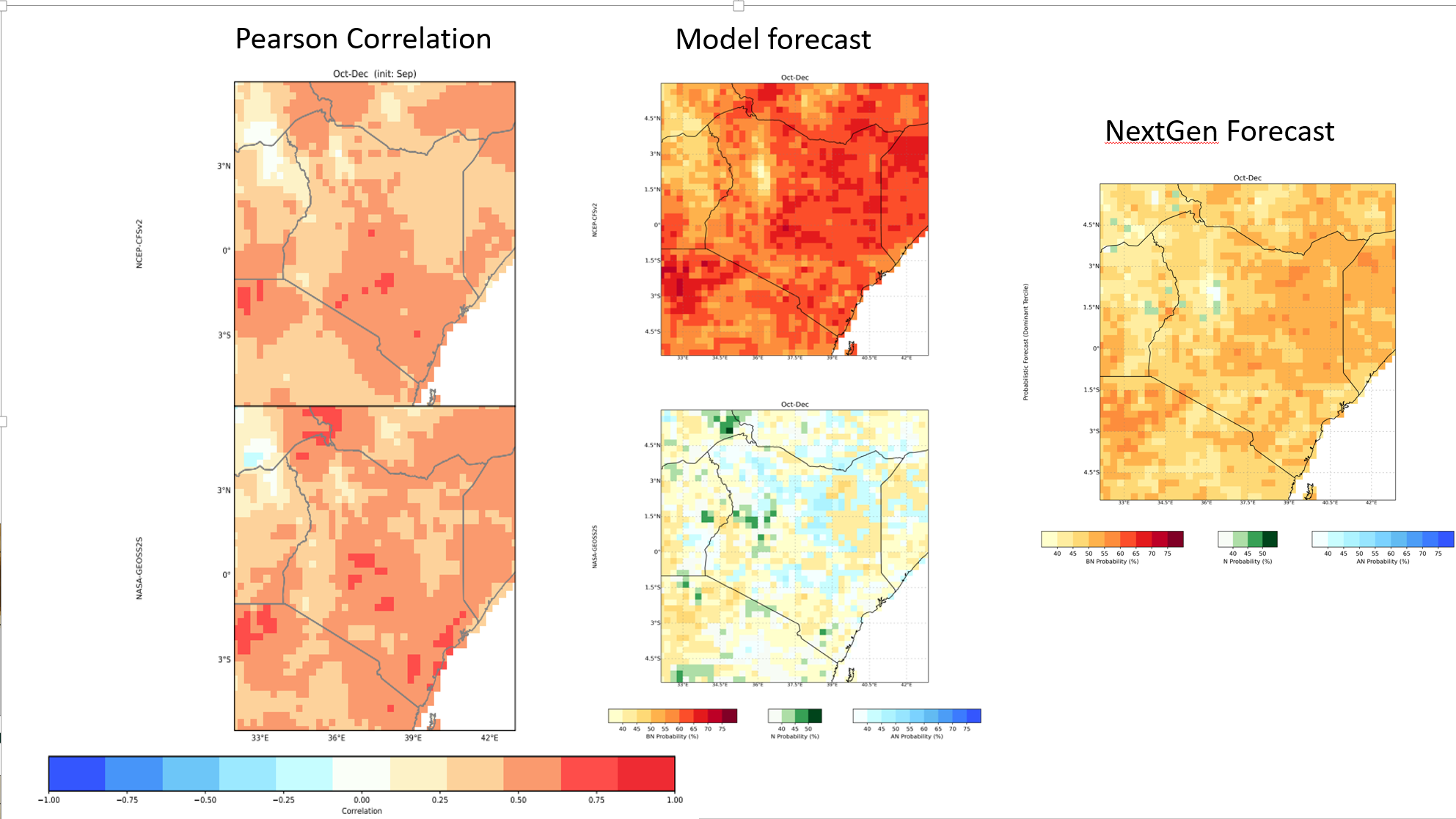
**Figure 4: MAM 2021 Pearson correlation (bottom left), individual model forecast for Atlantic predictor domain (middle) and NextGen forecast (right)**

Significant difference is observed on the forecast when the two different domains are used as shown in figures 3 and 4 above. The Atlantic gives a significantly dry season while the Pacific domains gives a significantly wet season.

**OND 2021 from September Initialization**

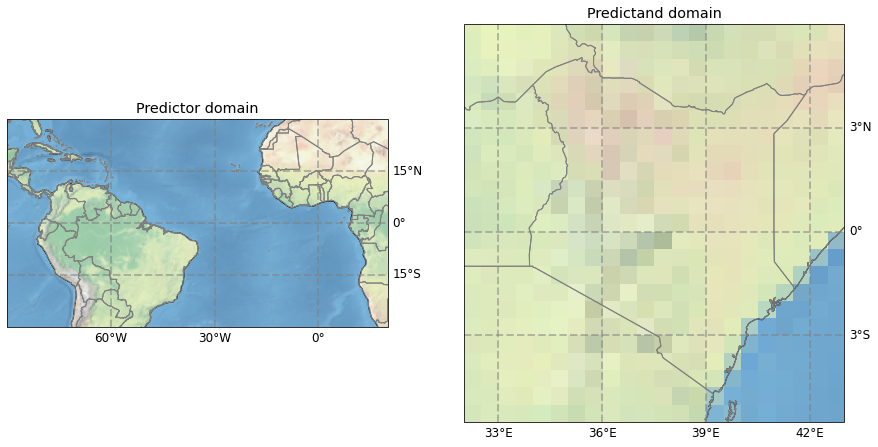
**Predictor Domain: Global Predictand Domain: Kenya**

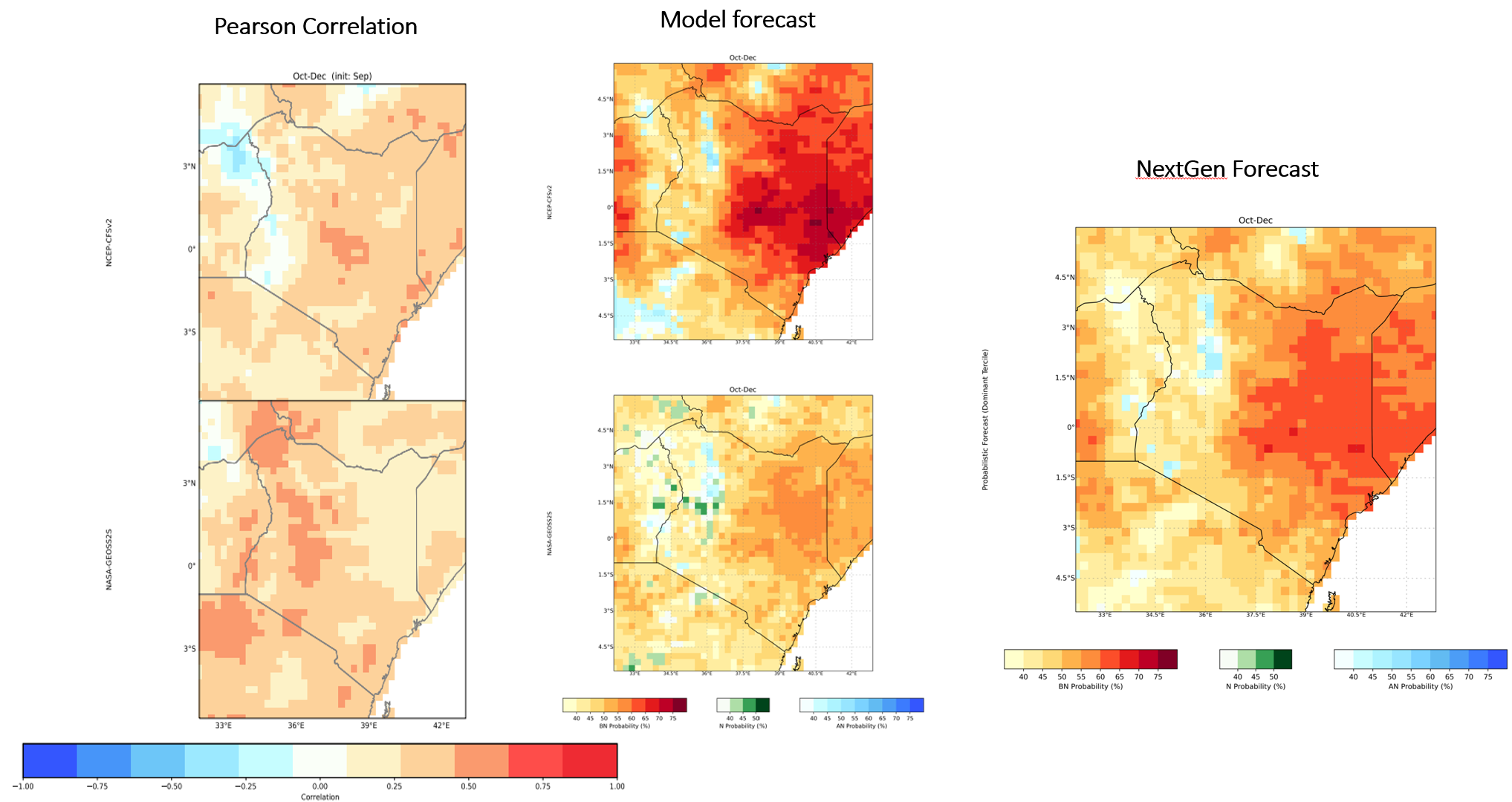
****

****

**Figure 5: OND 2021 skill (left), individual model forecast (middle) and NextGen forecast (right) from September initialization using global tropics predictor domain**

**Predictor Domain: Atlantic Predictand Domain Kenya**



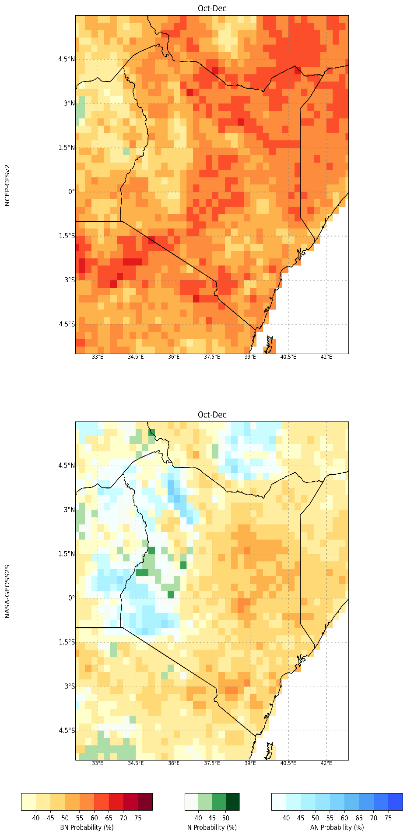
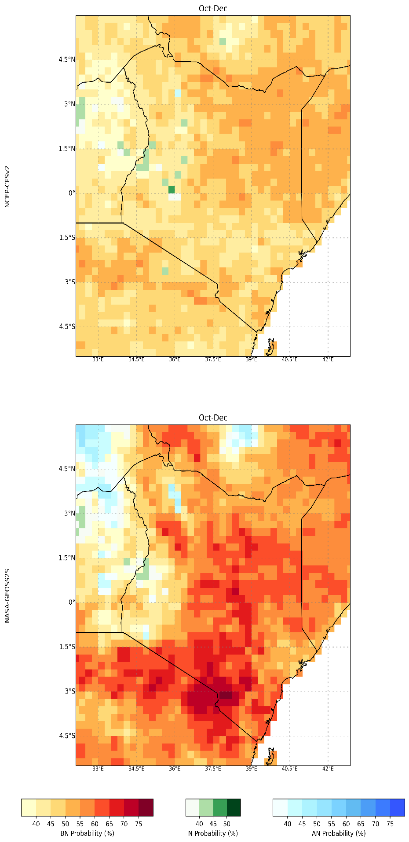
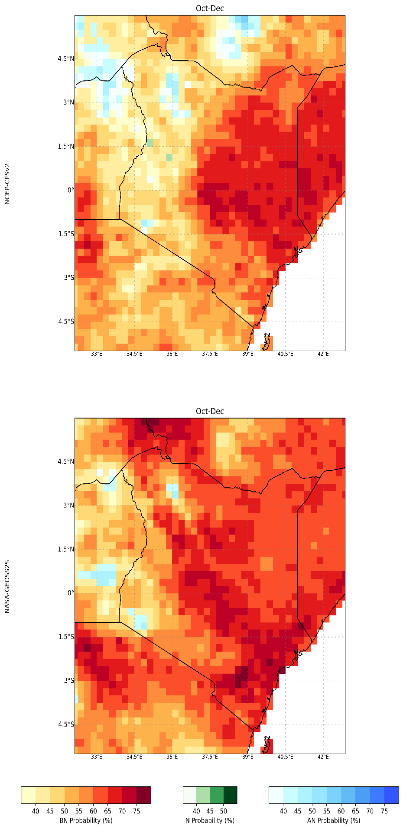
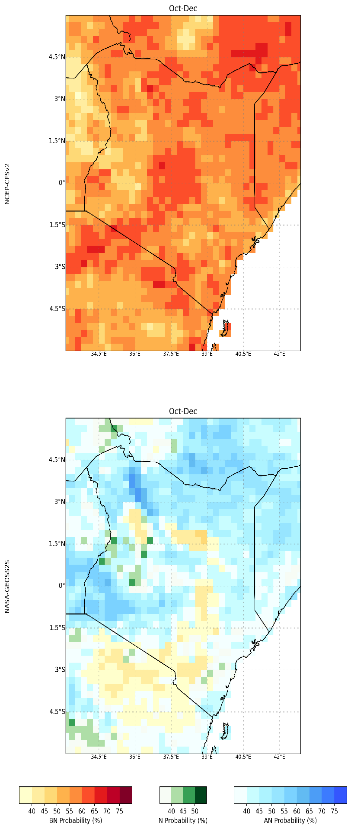


**Figure 6:OND 2021 skill (left), individual model forecast (middle) and NextGen forecast from September initialization using Atlantic as the predictor domain**

Theresults in figure 5 and 6 above uses the same predictand domain and different predictor domains, the impact on the forecast is evident as the probability of below normal when the Atlantic domain is used is enhanced compared to that of the global tropics predictor domain. On the other hand, the skill scores are enhanced when the global domain is used.​

# **Generating a Seasonal Forecast using Principle Component Regression (PCR)**

After understanding how to generate a seasonal forecast using CCA, the participants went ahead to generate the forecast using PCR. The two groups were assigned the seasons they previously worked on. However, the OND season was to be analyzed at both 1 and 2-months lead time. Sample individual model results for Global and Indian Ocean predictor domains for OND (1- and 2-months lead) are shown in figure 7 while Figure 8 show individual model results for Global predictor domains for MAM from February initialization. These results were presented and discussed in a plenary session.



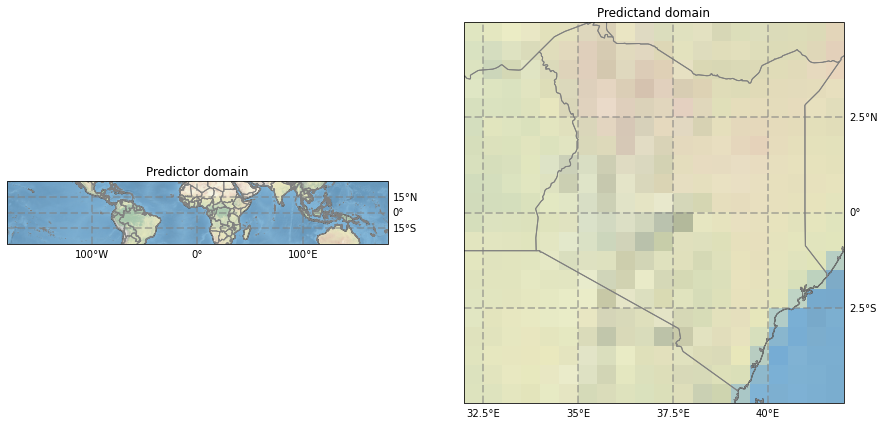
Global-Sep

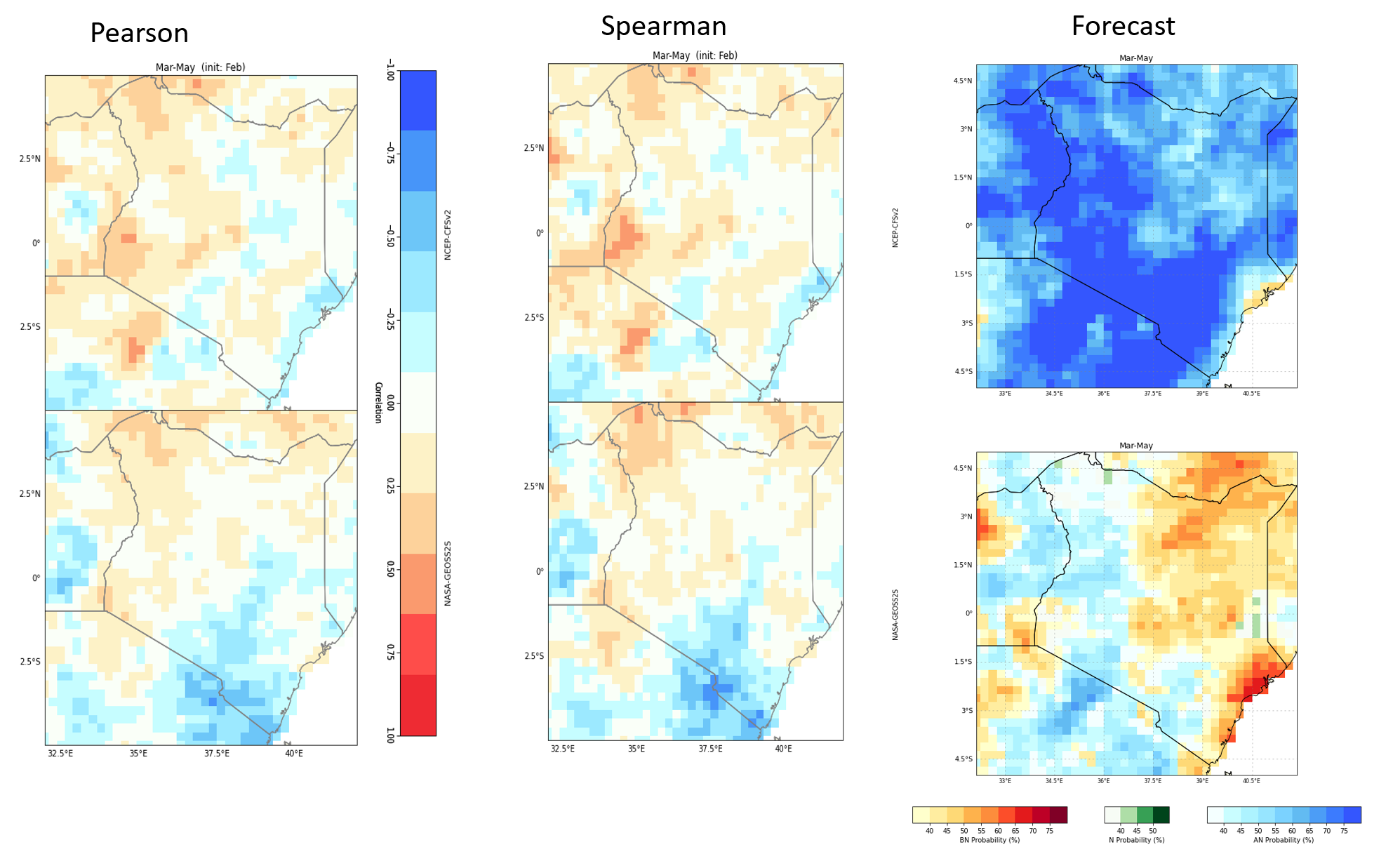
Global-Aug

IO-Aug

IO-Sep

**Figure 7: OND 2021 forecast from August and September initialization for Global and Indian Ocean predictor domains using NCEP-CFSv2 and NASA-GEOSS2S Models**





**Figure 8: MAM 2021 forecast from February initialization for Global predictor domain (top) using NCEP-CFSv2 and NASA-GEOSS2S Models**

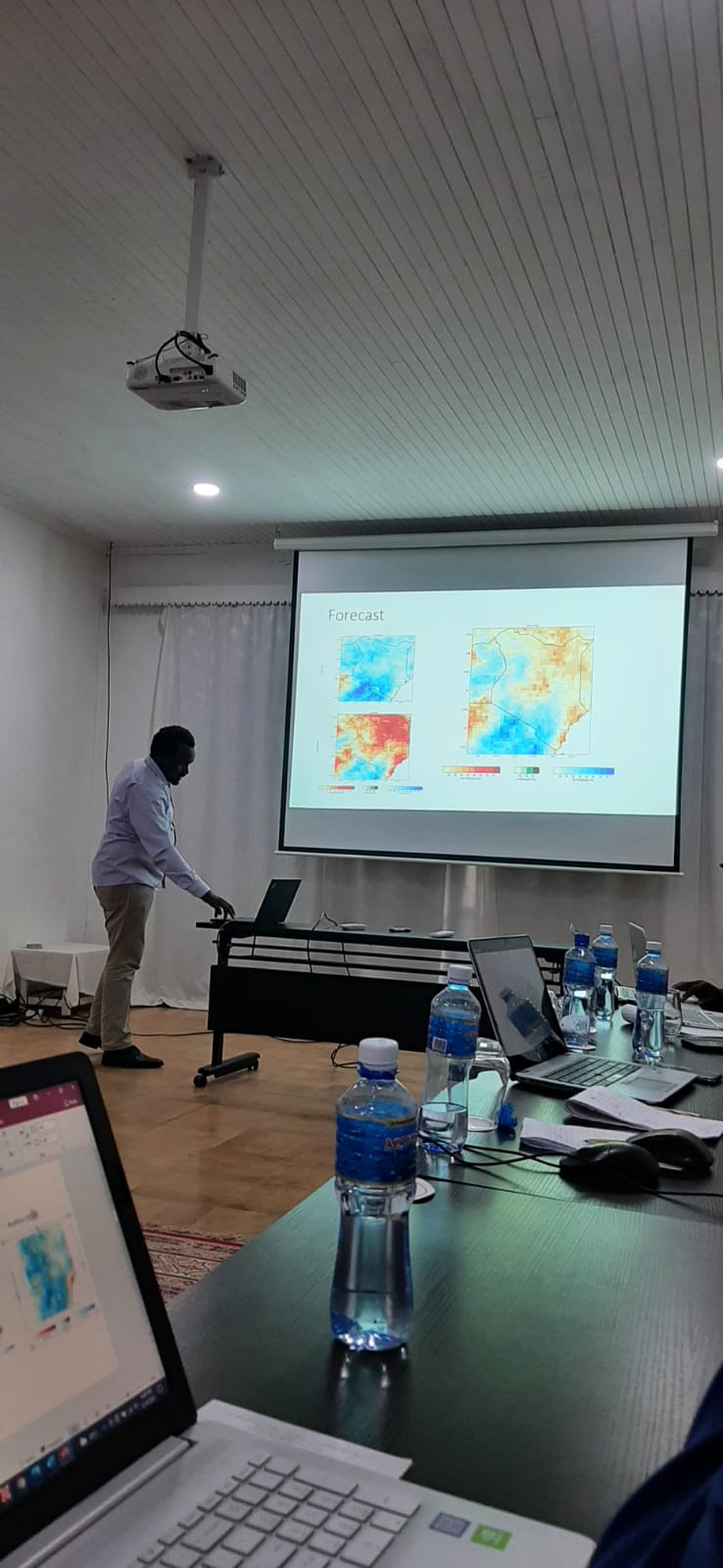
# **Conclusion and Closing Remarks**

The Next Generation technique for seasonal forecasting provides an objective way of combining individual model forecasts. Domain and lead time play a significant role in determining the forecast and this should be taken into consideration with prior knowledge on the domain with significant impact on the season. During the OND season, for instance, the Indian ocean and Pacific Ocean are significant and should be considered as the predictand domains. Other domains can also be considered depending on their influence during the season of interest.

The workshop was closed by Mr. Paul Ombai on behalf of the director ICPAC. He thanked the participants for the hard work they put for a whole week and also encouraged them to put the knowledge they have gained into practice to improve seasonal forecasting at national level. He pointed out that as climate continue to change, new methodologies should be adopted to ensure that forecasts accuracy is enhanced.

# **Annex: Workshop Photos and List of Participants**

****

****

****

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **First Name** | **Second Name** | **Institution** | **Email** |
| 1 | Herbert | Misiani | IGAD | [herbert.misiani@igad.int](mailto:herbert.misiani@igad.int) |
| 2 | Masilin | Gudoshava | IGAD | [masilin.gudoshava@igad.int](mailto:masilin.gudoshava@igad.int) |
| 3 | Barbara | Olwande | IGAD | [barbara.olwande@igad.int](mailto:barbara.olwande@igad.int) |
| 4 | Eunice | Koech | IGAD | [eunice.koech@igad.int](mailto:eunice.koech@igad.int) |
| 5 | George | Otieno | IGAD | [george.otieno@igad.int](mailto:george.otieno@igad.int) |
| 6 | Anthony | Mwanthi | IGAD | [anthony.mwanthi@igad.int](mailto:anthony.mwanthi@igad.int) |
| 7 | Titike | Bahaga | IGAD | [titike.bahaga@igad.int](mailto:titike.bahaga@igad.int%20) |
| 8 | Zewdu | Segele | IGAD | Zewdu.segele@igad.int |
| 9 | Mary | Kilavi | KMD | [marykilavi@yahoo.com](mailto:marykilavi@yahoo.com) |
| 10 | Hannah | Kimani | KMD | hwkimani1@gmail.com |
| 11 | Franklin | Komolkori | KMD | Frankom09@gmail.com |
| 12 | Jemima | Gacheru | KMD | gacherujeminah@gmail.com |
| 13 | Phinous | Okoth | KMD | Phyn.okoth@gmail.com |
| 14 | Chris | Kiptum | KMD | chriskiptum@gmail.com |
| 15 | Christine | Maswi | KMD | maswichristine@gmail.com |
| 16 | Paul | Ombai | IGAD | [Paul.ombai@igad.int](mailto:Paul.ombai@igad.int) |
| 17 | Joseph | Kiragu | IGAD | [Jeosph.kiragu@iga.int](mailto:Jeosph.kiragu@iga.int) |
| 18 | Leonard | Kimotho | IGAD | Leonard.kimotho@igad.int |