Seasonal Rainfall Forecasting with Multimodel Climate Ensembles

PREPARE Drought and Flood Early Warning for Pacific Islands Training Workshop
Nadi, Fiji, 15-20 July 2024

Dr. Katie Kowal

NOAA/CPC/International Desks

17 July 2024

All models are wrong, *but some are useful*

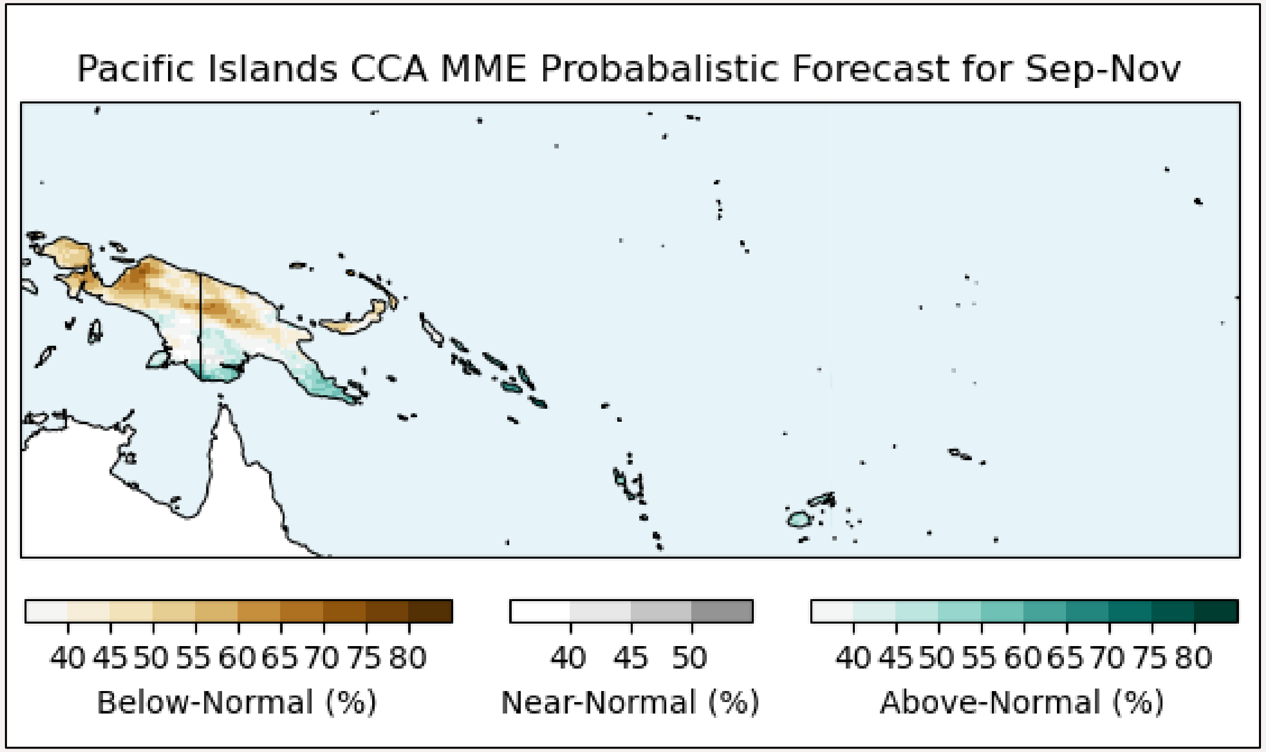
Earth system is complex and chaotic, no model will have a perfect representation of future conditions

We can **improve** the **usefulness of a prediction by**

(1) **Choosing a more attainable forecast type** for a model (e.g. tercile-based probabilities)

(2) **Combining forecast models** to aggregate different assumptions about the earth system

(3) **Bias-correcting** our raw model outputs using statistical relationships



Common Models / Rainfall Data Products Used at the CPC International Desk

Monitoring Products	Subseasonal Forecasts	Seasonal Forecasts
CMORPH 0.25' rainfall data	GFS (1 member, many variables)	**NMME - North American Multimodel Ensemble** -NOAA NCEP CFSv2 -NOAA GFDL SPEAR -NASA GEOSv5 -NCAR CSM4 and/or CESM1 -Canadian Models (GEM5-NEMO, CanCM4i)
CHIRPS 5km rainfall data	GEFS (31 members, a few variables)	C3S Multimodel Ensemble (European models) -ECMWF SEAS 51 -UK Met Office Glosea6 -Meteo France System 8 -German DWD System 21 -Italian CMCC System 35 -Canadian Models (GEM5-NEM, CanCM4i) -NOAA CFSv2
CMAP		
CPC Unified Gauge		
RFEv2		

Several techniques can correct raw forecasts

- **Canonical Correlation Analysis (CCA)** – spatial pattern matching using linear correlations
- Logistic Regression (LR) **–** probability of a binary outcome (probability of above or below normal) based on one or more predictor variables, applied separately for each tercile
- **Extended Logistic Regression (ELR)** – nonlinear ensemble calibration that extends Logistic Regression and is applied across all terciles simultaneously
- Probabilistic Output Extreme Learning Machine (POELM) – nonlinear advanced machine learning approach (neural network with a single hidden layer)
- **Extended Probability Output Extreme Learning Machine (EPOELM)** – combines ELR and POELM approaches, useful for non-exceedance forecasts for any threshold value after fitting

Considerations for applying these techniques

- Grid by grid vs. spatial pattern matching
- Need for examining nonlinear vs. linear relationships
- Choosing predictor variables, e.g. SSTs or Precip

slido

Please download and install the
Slido app on all computers you use

Join at slido.com
#3573517

ⓘ Start presenting to display the joining instructions on this slide.

slido

Please download and install the
Slido app on all computers you use

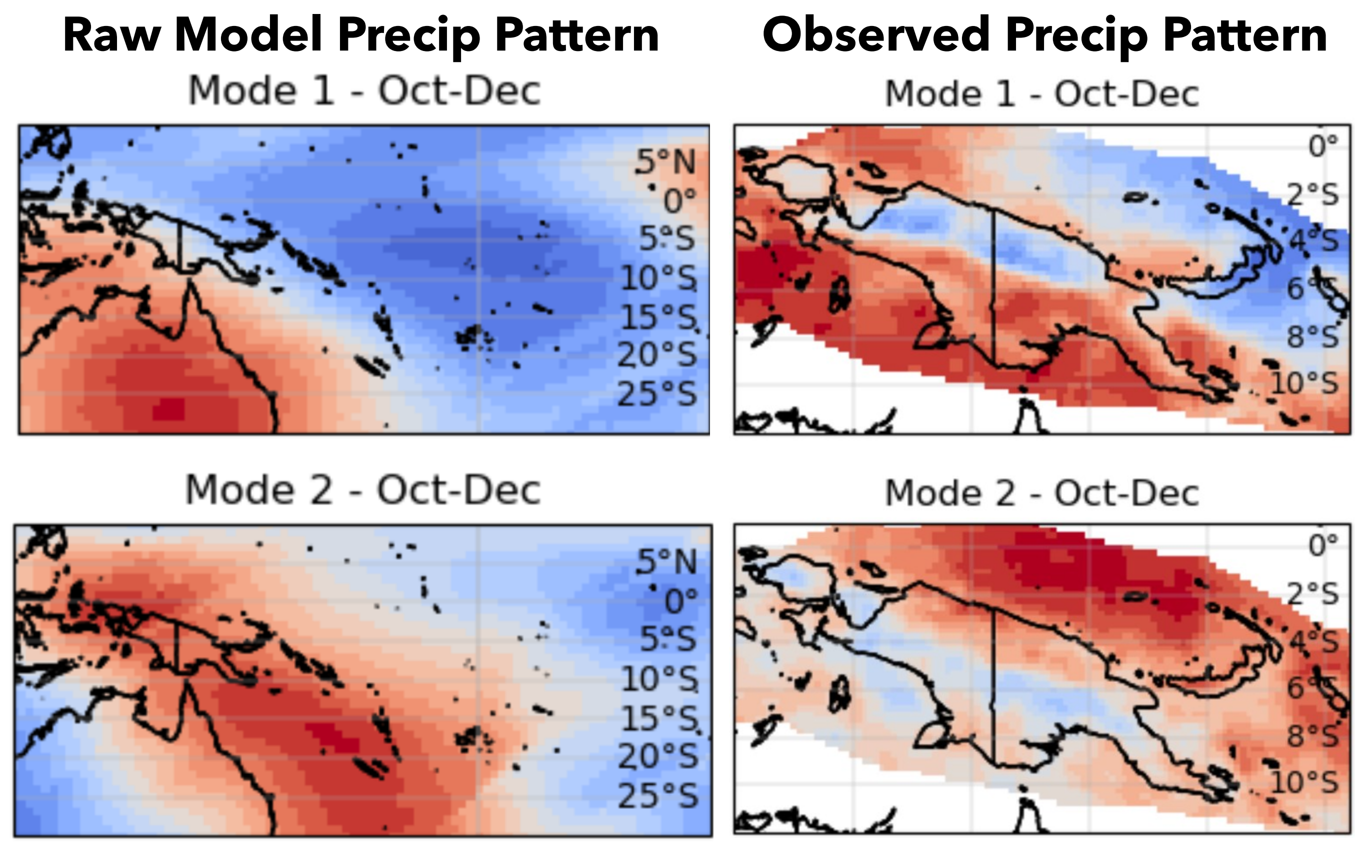
How comfortable do you feel with using CCA as a bias correction technique?

ⓘ Start presenting to display the poll results on this slide.

CCA corrects the pattern of the spatial bias and downscales raw model forecasts to the predictand resolution

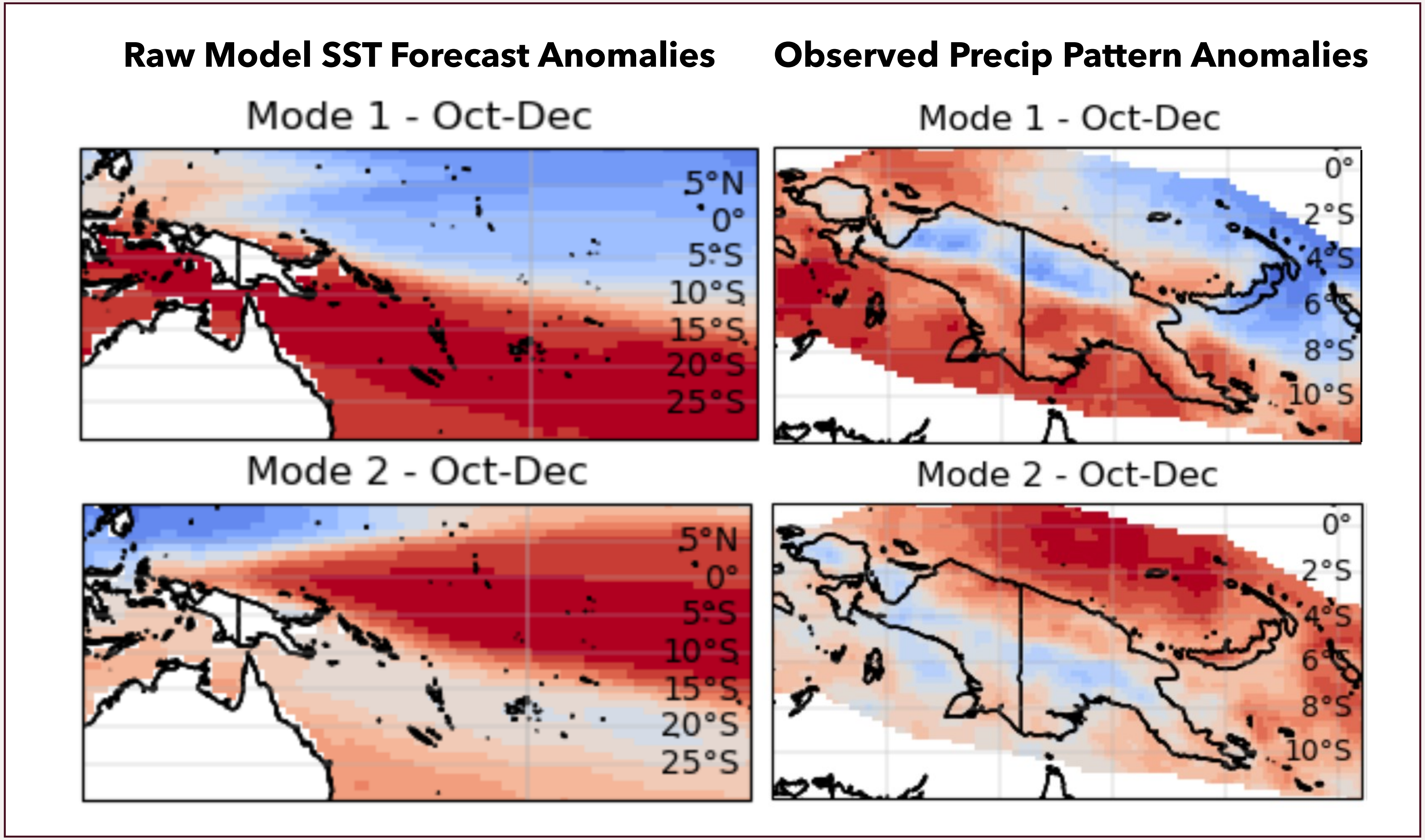
Key CCA Steps

1. **Estimate spatial pattern** with a principle component (PC) analysis
2. **Correlate PCs**
3. **Find combinations** between principle components **with maximum correlation**



Example: Comparing the Spatial Pattern of Rainfall Anomalies

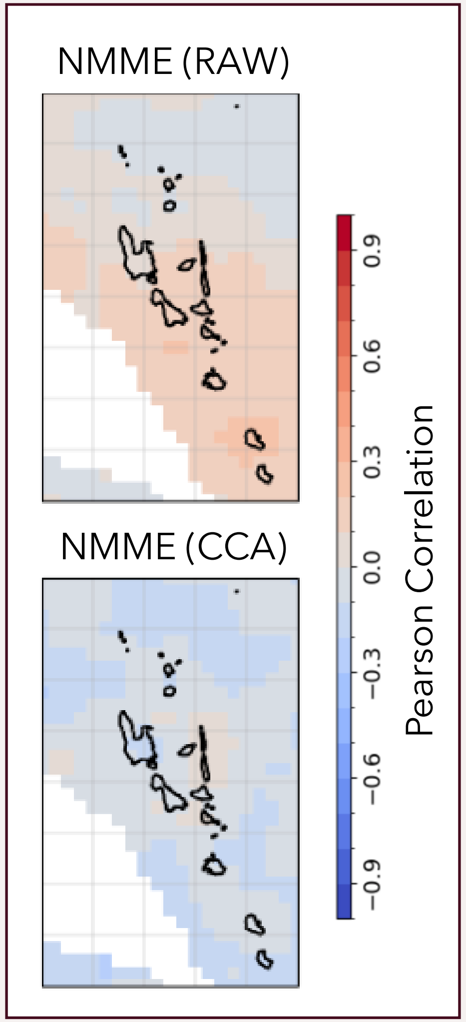
Diving More Deeply into CCA
Possibilities of Driving Precip Forecasts with SST



Verification is KEY - Skillfulness of Bias Correction Varies Spatially and Seasonally

- **CCA** is **often better over large zones but** can be **sensitive to** the **spatial extent** used to train a model **and** the **time of year**

- **Grid-based techniques** (ELR, EPOELM in this session) can **better account for nonlinear relationships**, **but require more data** to be useful **and often perform worse in small zones** where there are fewer grid cells to train a model (e.g. small islands)



Example of Bias Correction Not Helping Over Vanuatu
Model Initialized October, Predicting Dec-Feb

Step by Step Verification/Consolidated Ensemble Generation Process using CMORPH 0.25′ rainfall data

1. **We train and verify our model using hindcast data** – prior forecasts over a historical period
2. **We expand our dataset years** because CMORPH data is only available from 1998-2022 (25 years). To expand our dataset, we train our data over 3 seasons at a time

EXAMPLE: Model is initialized this month, 1st of July, 2024, for a one-month lead forecast (August-October, ASO), we pull hindcast data from July-Sep, Aug-Oct, and Sep-Nov for 25 available years (75 data points to train/test our model performance)

3. **We split training/testing years.** When we train the model using CCA we dump out X number of years (e.g. 5), but will randomly pull the test years 75 times.
4. **We apply a skill score** (e.g. Pearson or GROCS) to comparatively examine how the raw and bias-corrected forecasts compared
5. **We generate a consolidated ensemble** that combines different bias correction techniques, weighted by their skill (e.g. GROCS)

slido

Please download and install the
Slido app on all computers you use

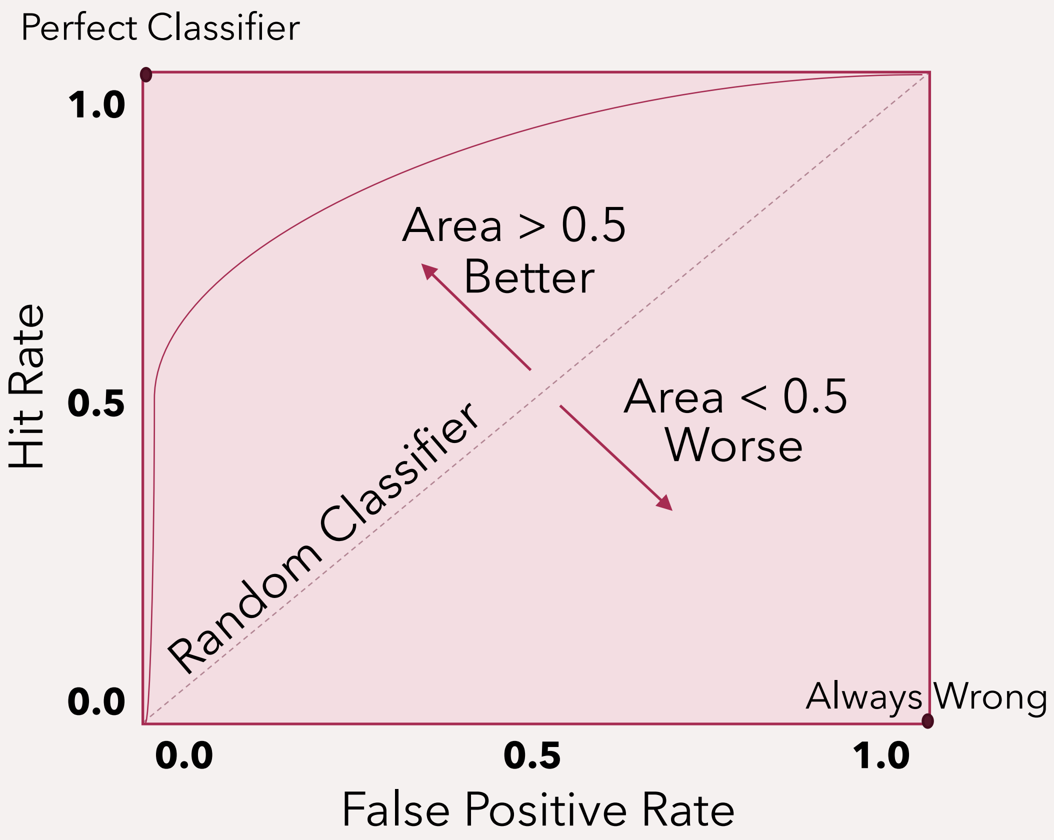
How comfortable are you with applying the Generalized Receiver Operating Characteristic Score (GROCS) as a verification score for forecasts?

ⓘ Start presenting to display the poll results on this slide.

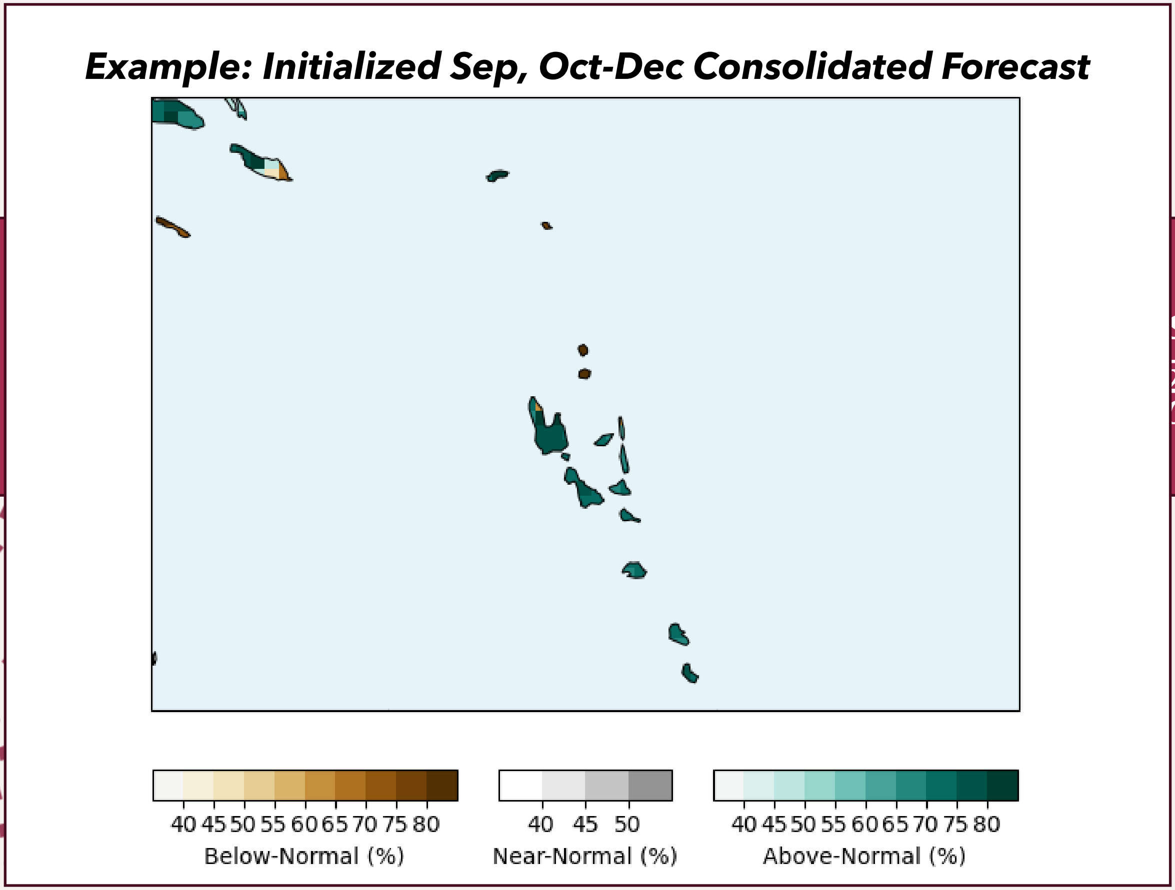
Verification Metric Used in Today's Session: Generalized Receiver Operating Characteristic Score (GROCS)

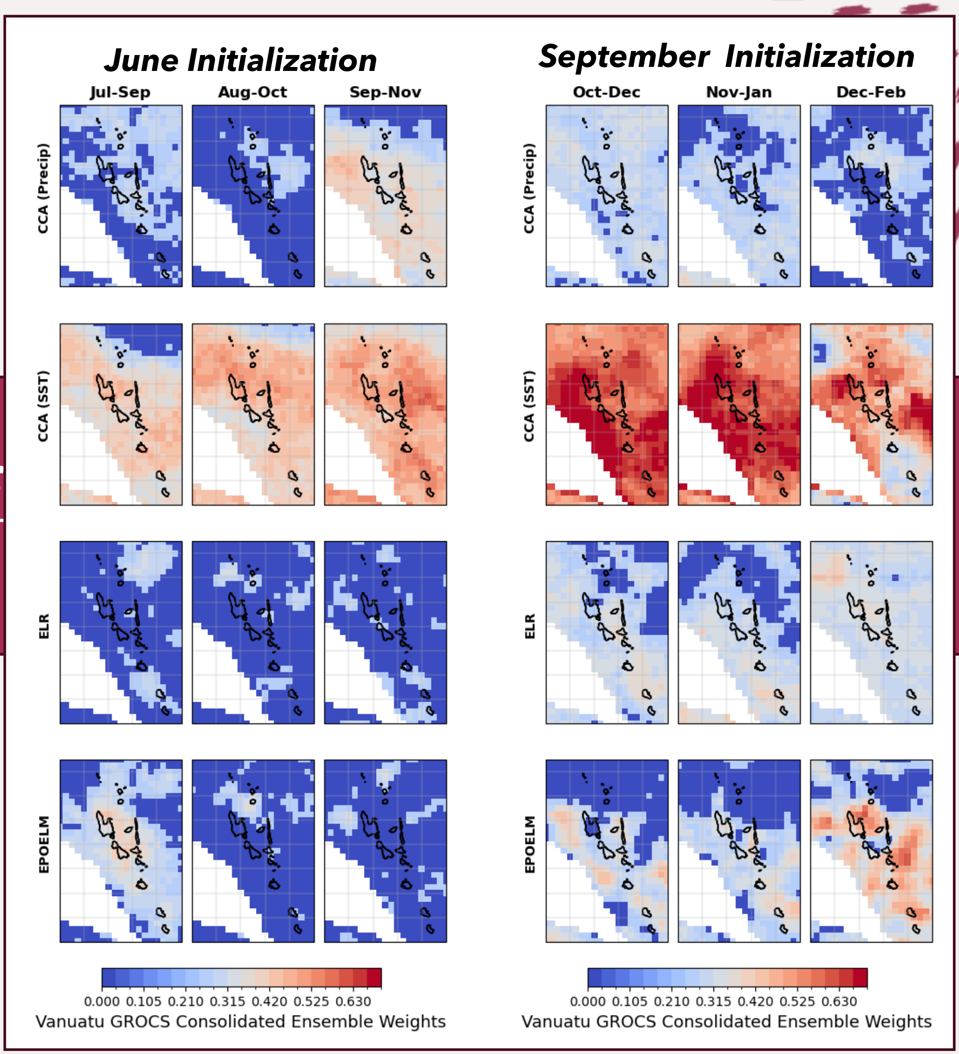
- **Receiving Operating Characteristic Curve – ROC:** visualization of the ability of a forecast to discriminate between binary classes (e.g. False Positive, Hit Rate, Misses)
- **Generalized ROC:** ROC applied to probabilistic forecasts, degree of correct probabilistic forecast discrimination between all forecast categories (e.g. below-normal, normal, and above normal)

<u>ROC Visualized</u>



Using GROCS to generated a consolidated weighted ensemble forecast



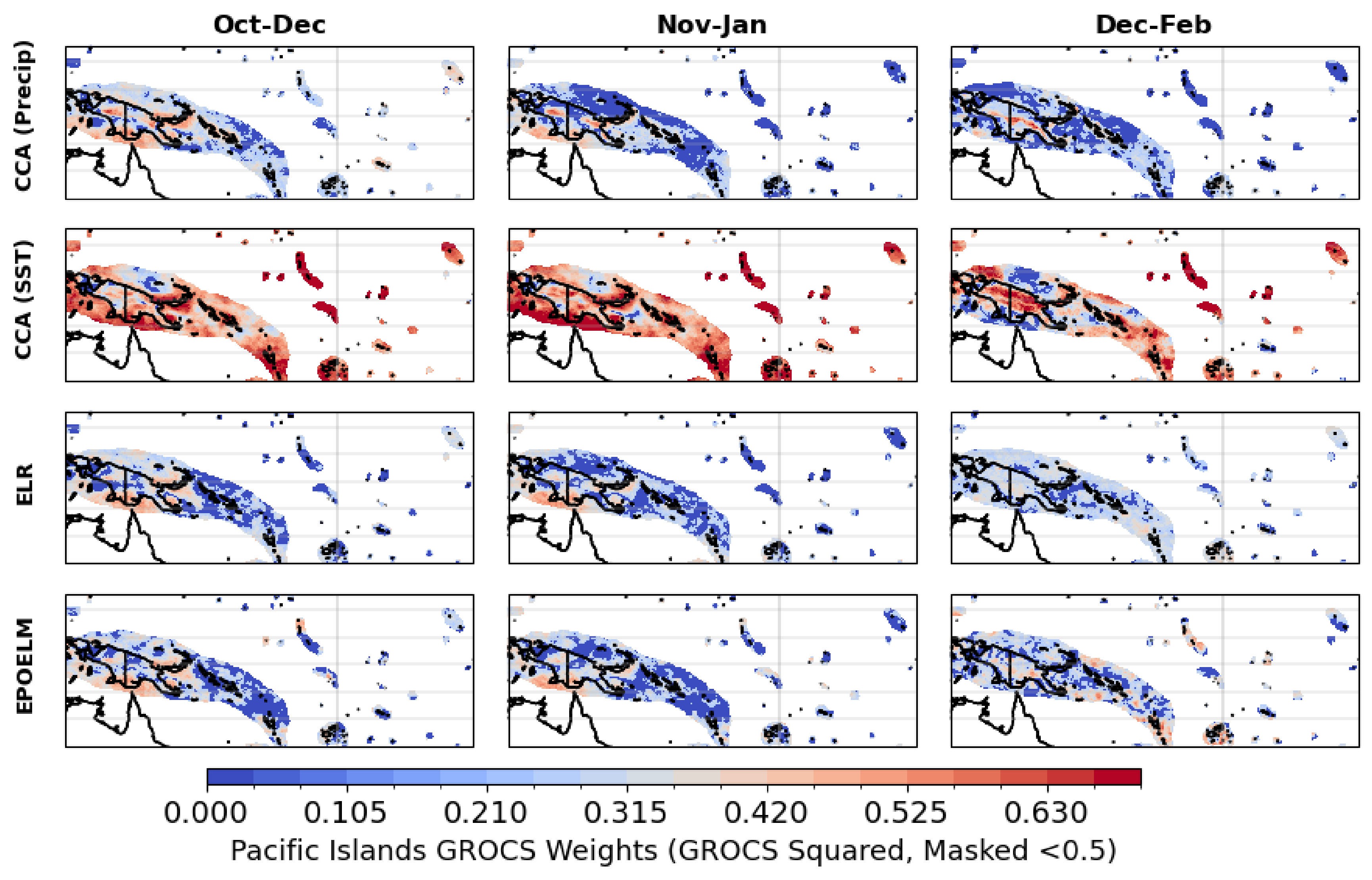


Consolidated Probabilistic Forecast (After masking GROCS < 0.5) =

$$\frac{prob(CCA{-}SST\ fcst) * GROCS^2_{CCA-SST} + prob\ (CCA{-}Precip\ fcst) * GROCS^2_{CCA-Precip} + prob(ELR\ fcst) * GROCS^2_{ELR} + prob(EPOELM\ fcst) * GROCS^2_{EPOELM}}{GROCS^2_{CCA-SST} + GROCS^2_{CCA-Precip} + GROCS^2_{ELR} + GROCS^2_{EPOELM}}$$

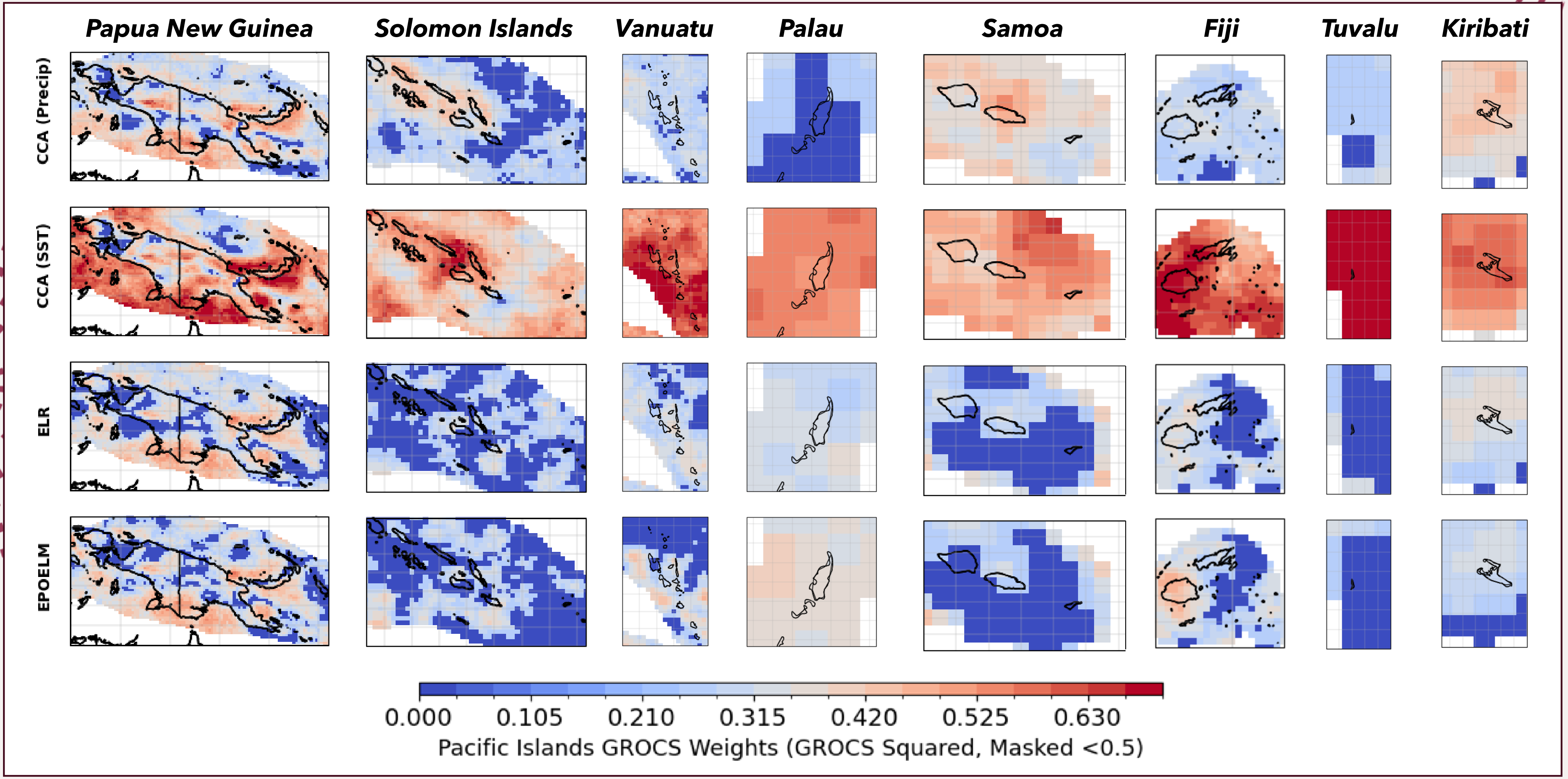
An Intercomparison of Methods

CCA, ELR, and EPOELM, initialized Sep for Oct-Dec Prediction



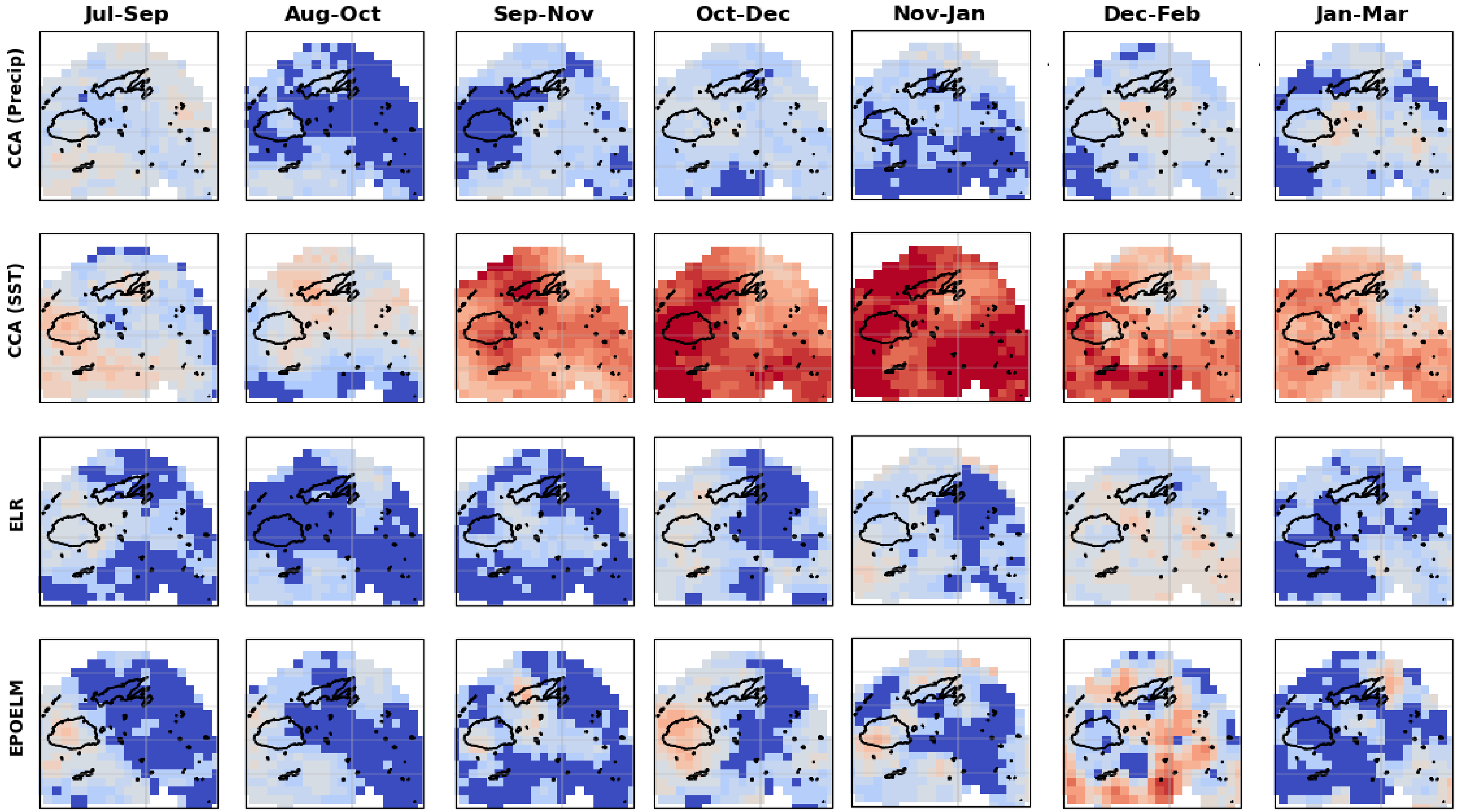
Spatial Intercomparison of Methods

CCA, ELR, and EPOELM … initialized Sep for Oct-Dec Prediction



Seasonal Intercomparison of Methods

CCA, ELR, and EPOELM … seasonal skill comparison in Fiji



Model Initialized One Month Ahead of Season (Same Scale as Prior Slide)

Goals of Today's Practical Session

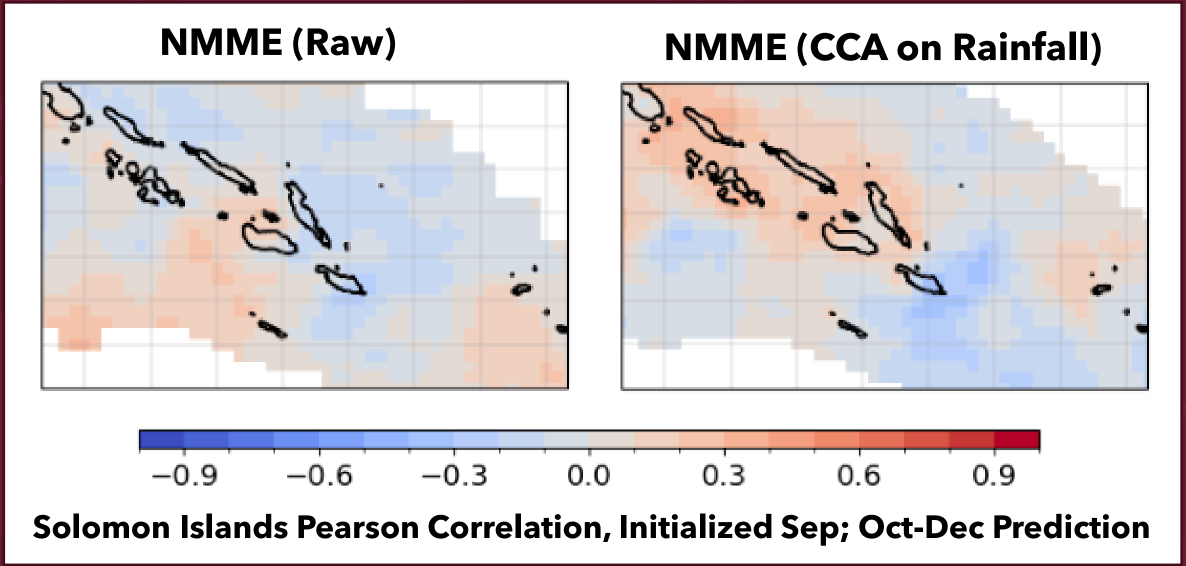
Use — **Use CCA to bias correct an NMME forecast**

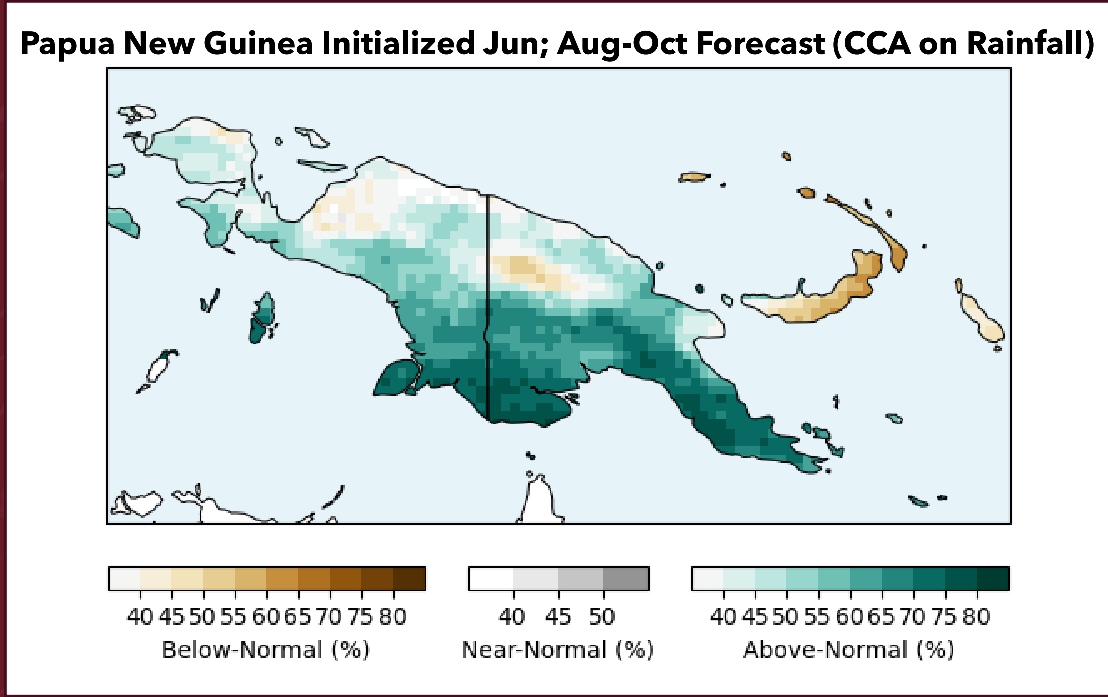
Compare — **Compare** the **CCA approaches** to see how **driving forecasts with SSTs vs rainfall** affects forecast performance

Consider — **Consider seasonal variability** of the correction techniques

Consider — **Consider spatial variability** of the above correction techniques

Predict — **Predict seasonal rainfall** by generating a probabilistic tercile forecast using a bias correction technique of choice





Example Plots