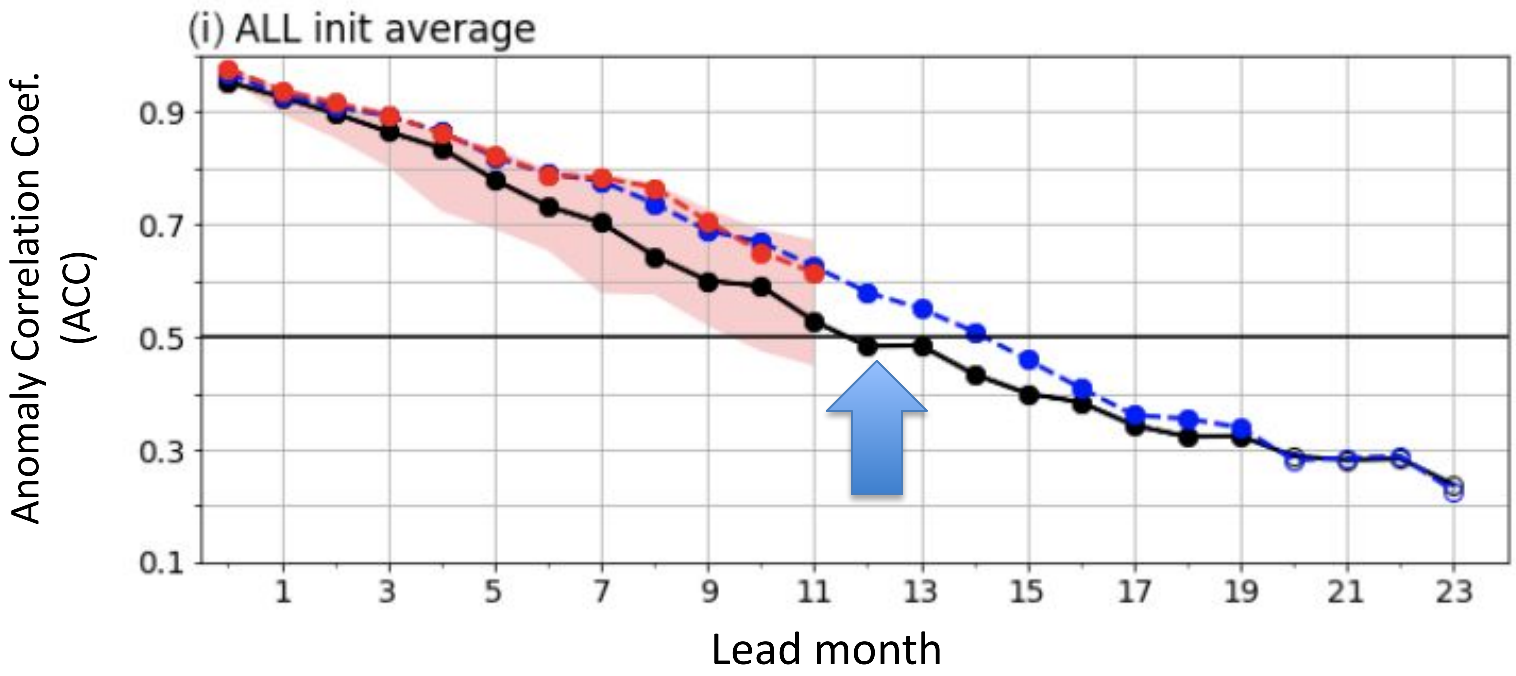
Advancing Long-Range ENSO Prediction through Large Ensemble Data and Transfer Learning with AI/ML Approaches

Sathish C. Akula
Department of Computer Science
Florida Polytechnic University, FL
Email: sakula@floridapoly.edu

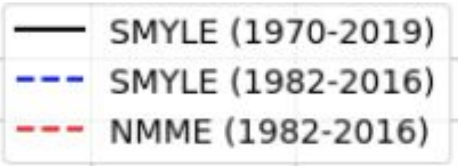
Sanjiv Kumar
College of Forestry, Wildlife and Environment
Auburn University, AL USA

Motivation

Improving long range (12-24 months) ENSO prediction using advance AI/ML methodologies



ENSO Prediction skill in Seasonal to Multiyear Large Ensemble (SMYLE) Prediction System



Yeager et al., 2022

Why – (1) Computational Efficiency, and (2) Range of applications – Agriculture, Water Resources, and Food Security

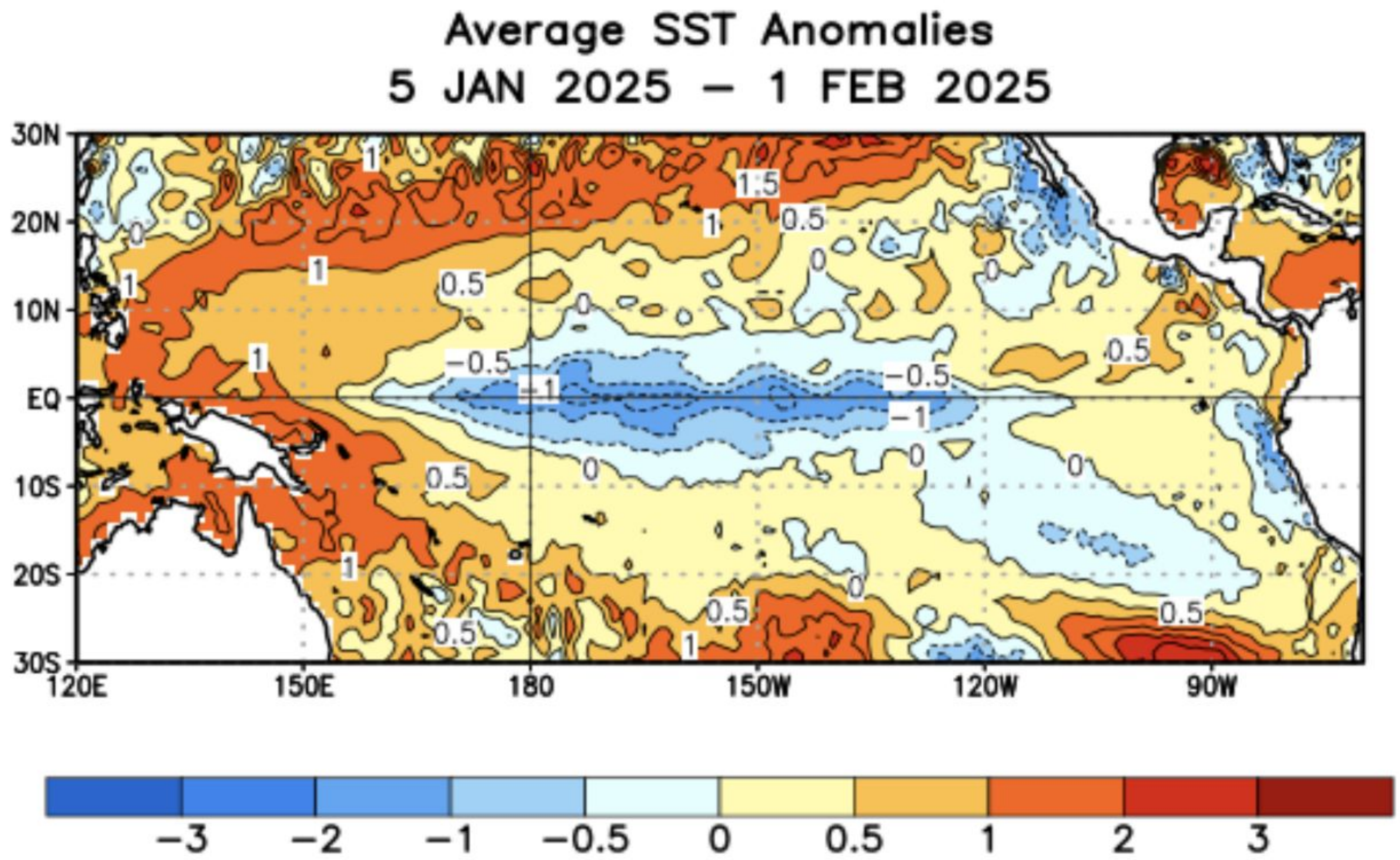
Literature Review

- **CNNs for ENSO Forecasting**: Studies by Ham et al., Kim et al., and Luo et al. demonstrated that Convolutional Neural Networks (CNNs) significantly improved ENSO prediction accuracy and extended lead times up to 24 months compared to traditional models.
- **LSTM Models' Performance**: Research by Kratzert et al. and Frame et al. showed that Long Short-Term Memory (LSTM) models achieved a 30% improvement in ENSO prediction accuracy over traditional methods while providing reliable forecasts up to 24 months.
- **ANNs for Climate Signal Extraction**: Barnes et al. utilized Artificial Neural Networks (ANNs) to enhance climate change signal detection, achieving a 10–15% improvement over traditional statistical approaches.

Research Objectives

Novel Contribution	**Tackles the "data-hungry" challenge in deep learning models using a large ensemble dataset (251 members × 100 years, CESM2-LE).**
	Addresses biases between climate model predictions and observations through Transfer Learning.
Specific Objectives	**Model Development**: Develop and evaluate deep learning models for long-term ENSO event predictions.
	Architecture Comparison: Compare the performance of CNNs, LSTMs, GRUs, Bi-directional LSTMs, and hybrid models.
	Lookback Analysis: Analyze the impact of different lookback periods on model accuracy and lead times.
	Transfer Learning: Implement transfer learning to enhance model generalizability and performance.

Data and Methods



SST anomaly in the equatorial Pacific

Current status (La Niña), NOAA

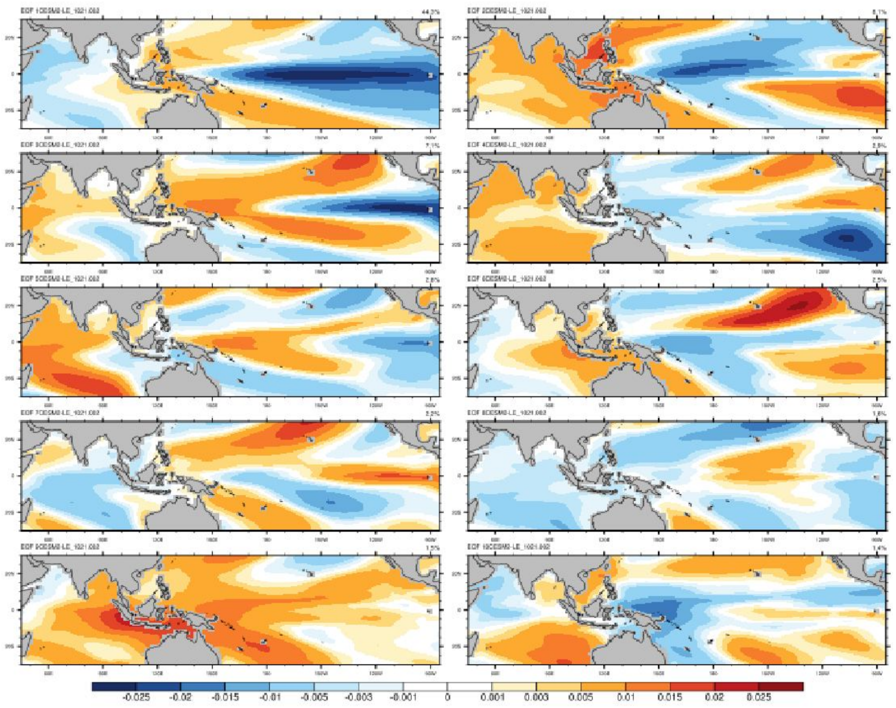
Aspect	CESM2-LE	Berkeley Earth
Model/Data Source	Community Earth System Model version 2 Large Ensemble (Rodger et al., 2021)	Meteorological stations, satellites, buoys, and ships
Duration	251 years (1850–2100) with 100 ensemble members	From 1950 to the present
Purpose	To understand long-term climate variability and change using CESM2-LE	To analyze recent climatic trends, validate climate models, and understand the impacts of climate change
Data Type	Modeled data	Observational data
Key Applications	Climate variability analysis, scenario projections	Model validation, trend analysis, and impact assessments

Data Preparation

1. A **12-month running mean** and **detrended anomalies** (grand ensemble mean removed) were analyzed.

2. **Principal Component Analysis (PCA)** was performed, retaining the **first 10 principal components (PCs)**.

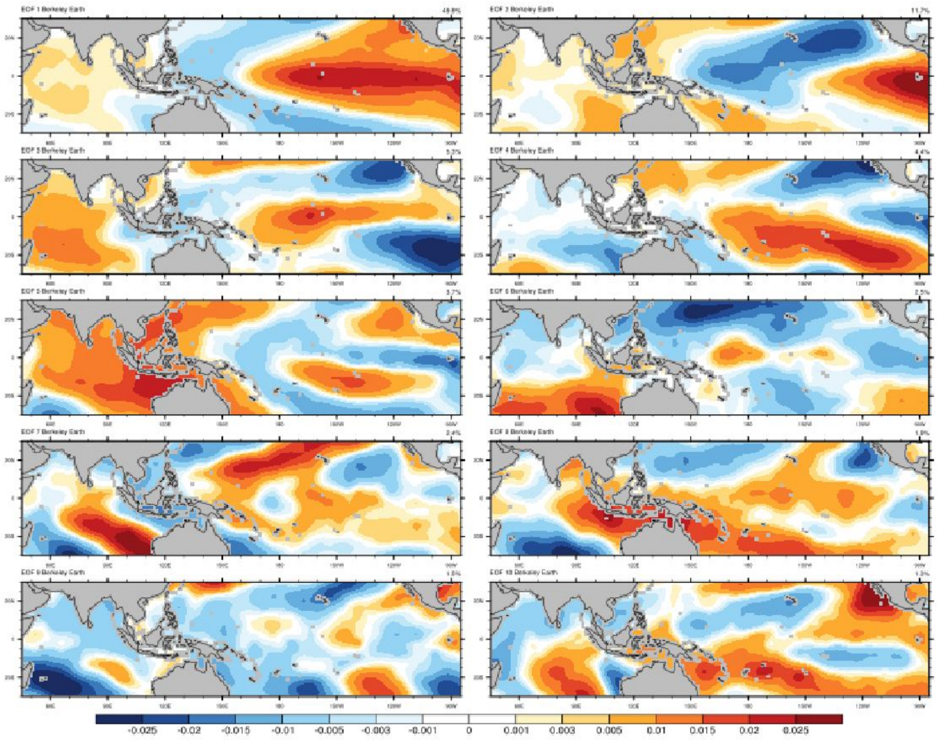
CESM2-LE



PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
44.3	8.1	7.1	2.9	2.8	2.5	2.2	1.6	1.5	1.4

Total variance explained: 74.4%

Berkeley Earth



PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
48.5	11.3	5.3	4.4	3.7	2.5	2.4	1.9	1.5	1.2

Total variance explained: 82.7%

Deep Learning Architectures

CNN1D Model:

Core Idea: Very effective at capturing local patterns and short-term dependencies through its convolutional filters.

Key Layers: Conv1D ⟶ Dropout ⟶ Flatten ⟶ Dense layers.

Standard LSTM Model:

Core Idea: They are designed to capture long-term dependencies in sequential data..

Key Layers: LSTM (with return_sequences) ⟶ Dropout ⟶ LSTM ⟶ Dense.

GRU Model:

Core Idea: Similar to LSTM but with a simpler gating mechanism. These are designed to capture long-term dependencies in sequential data.

Key Layers: GRU (return_sequences) ⟶ Dropout ⟶ GRU ⟶ Dense.

BiLSTM Model:

Core Idea: Processes the sequence in both forward and backward directions. which can be valuable if context from both past and future

Key Layers: Bidirectional LSTM layers with dropout followed by Dense.

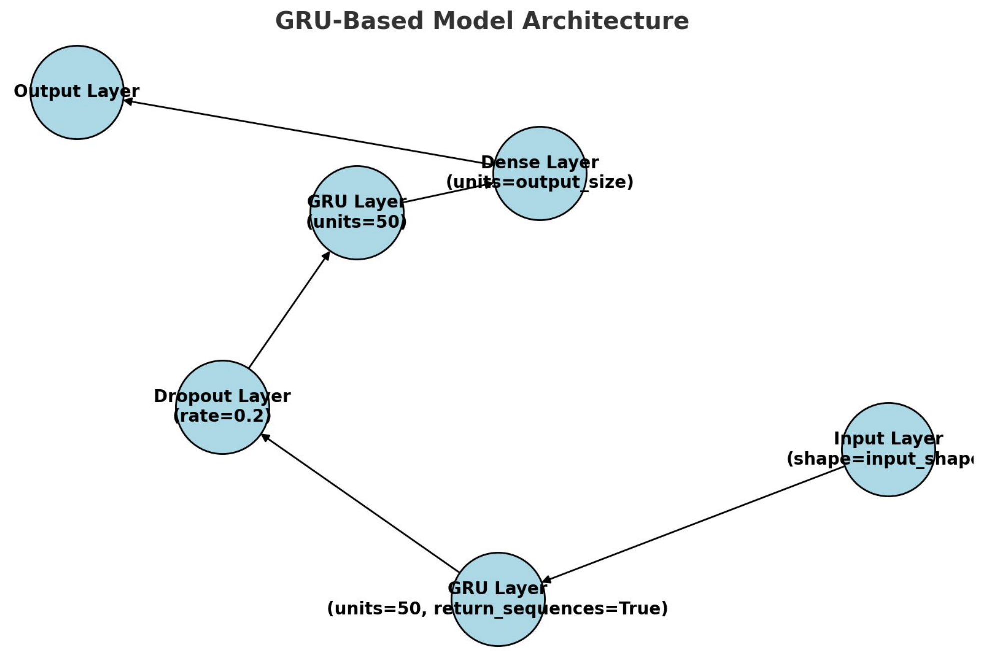
Hybrid Model (Conv1D + LSTM):

Core Idea: Leverage both local feature extraction and long-term sequential dependencies potentially capturing a richer representation of the data

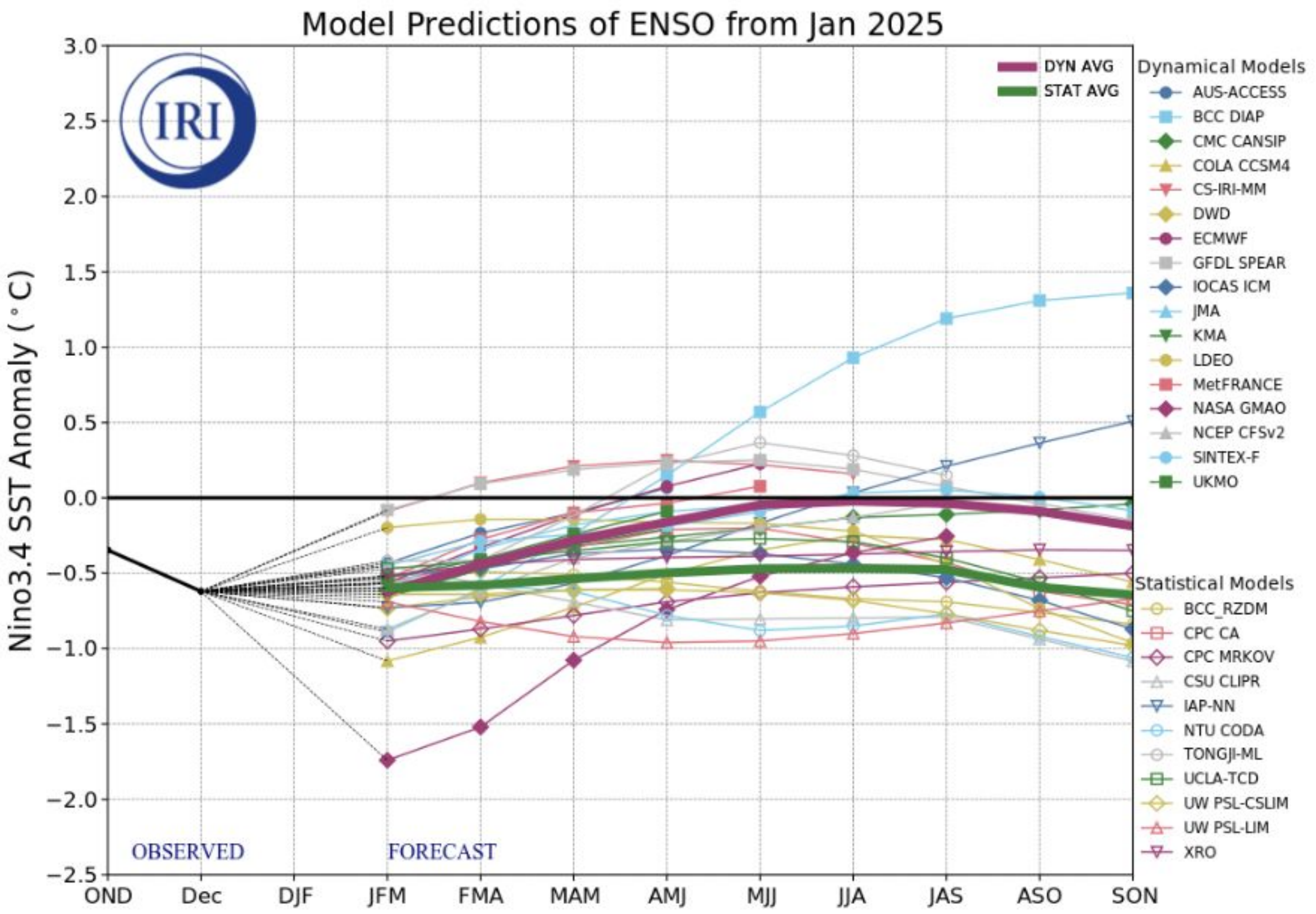
Key Layers: Conv1D ⟶ Dropout ⟶ LSTM ⟶ Dense.

Development of Deep Learning Models (GRU Architecture)

- **Input Data:** Processes sequences of SST data with lookback periods.
- **Model Type:** GRU-based recurrent architecture for capturing temporal dependencies.
- **Output:** Predicts SST anomalies for multiple lead times ahead.
- **Training:** Uses the Adam optimizer and MSE loss function, with early stopping to prevent overfitting.



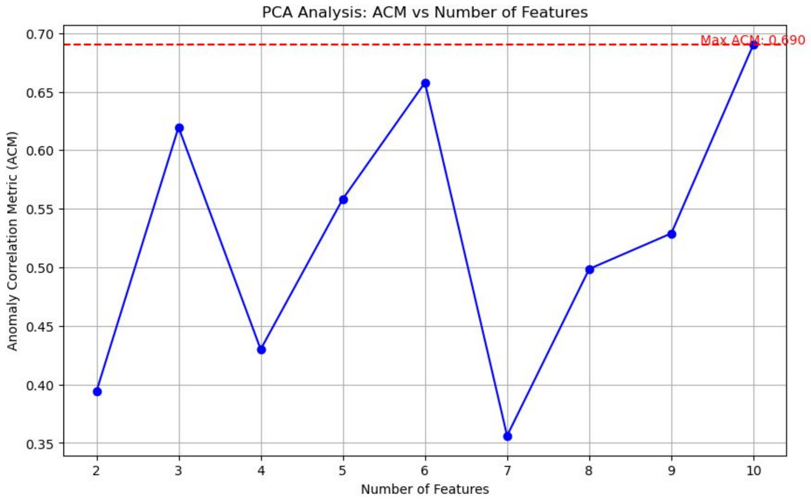
Results



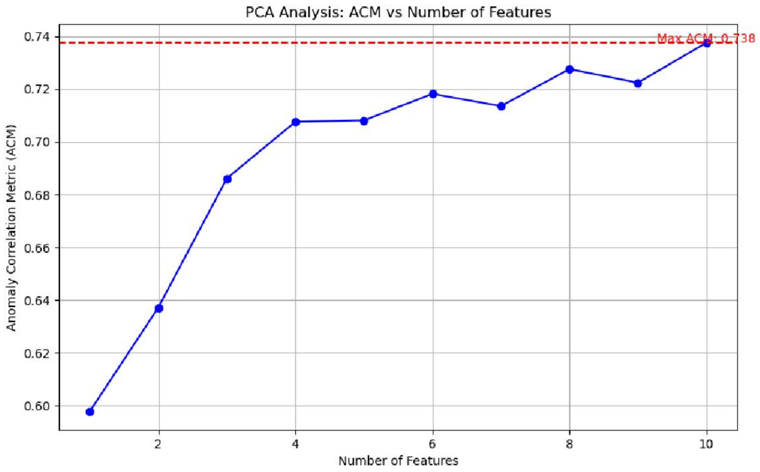
Model Predictions of ENSO from Jan 2025
Nino3.4 SST Anomaly (°C)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
−2.5
IRI
OBSERVED
FORECAST
OND
Dec
DJF
JFM
FMA
MAM
AMJ
MJJ
JJA
JAS
ASO
SON
DYN AVG
STAT AVG
Dynamical Models
AUS-ACCESS
BCC DIAP
CMC CANSIP
COLA CCSM4
CS-IRI-MM
DWD
ECMWF
GFDL SPEAR
IOCAS ICM
JMA
KMA
LDEO
MetFRANCE
NASA GMAO
NCEP CFSv2
SINTEX-F
UKMO
Statistical Models
BCC_RZDM
CPC CA
CPC MRKOV
CSU CLIPR
IAP-NN
NTU CODA
TONGJI-ML
UCLA-TCD
UW PSL-CSLIM
UW PSL-LIM
XRO

Feature Analysis

12-month Lead Anomaly correlation vs Number of features.

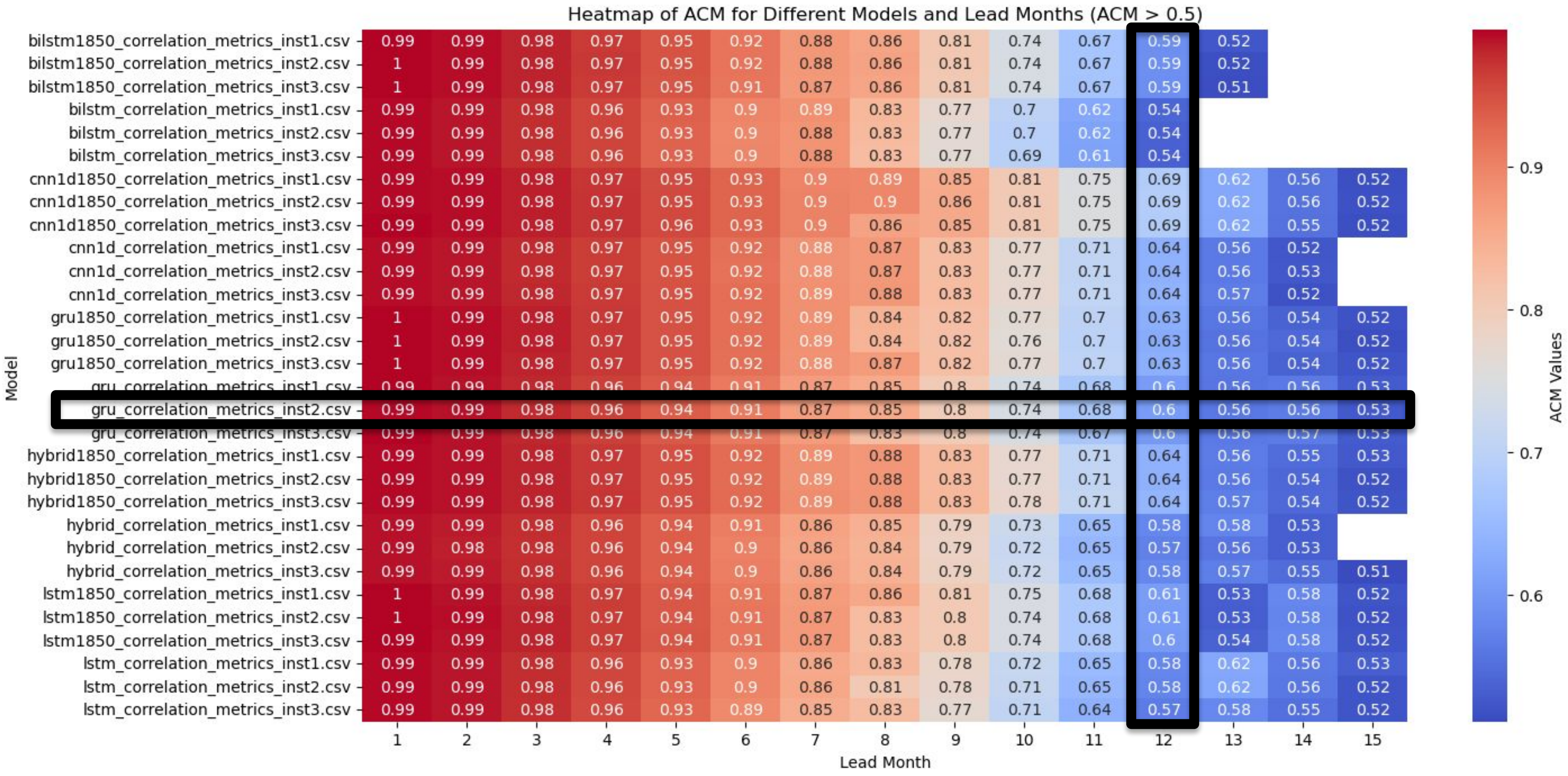


Berkeley Earth

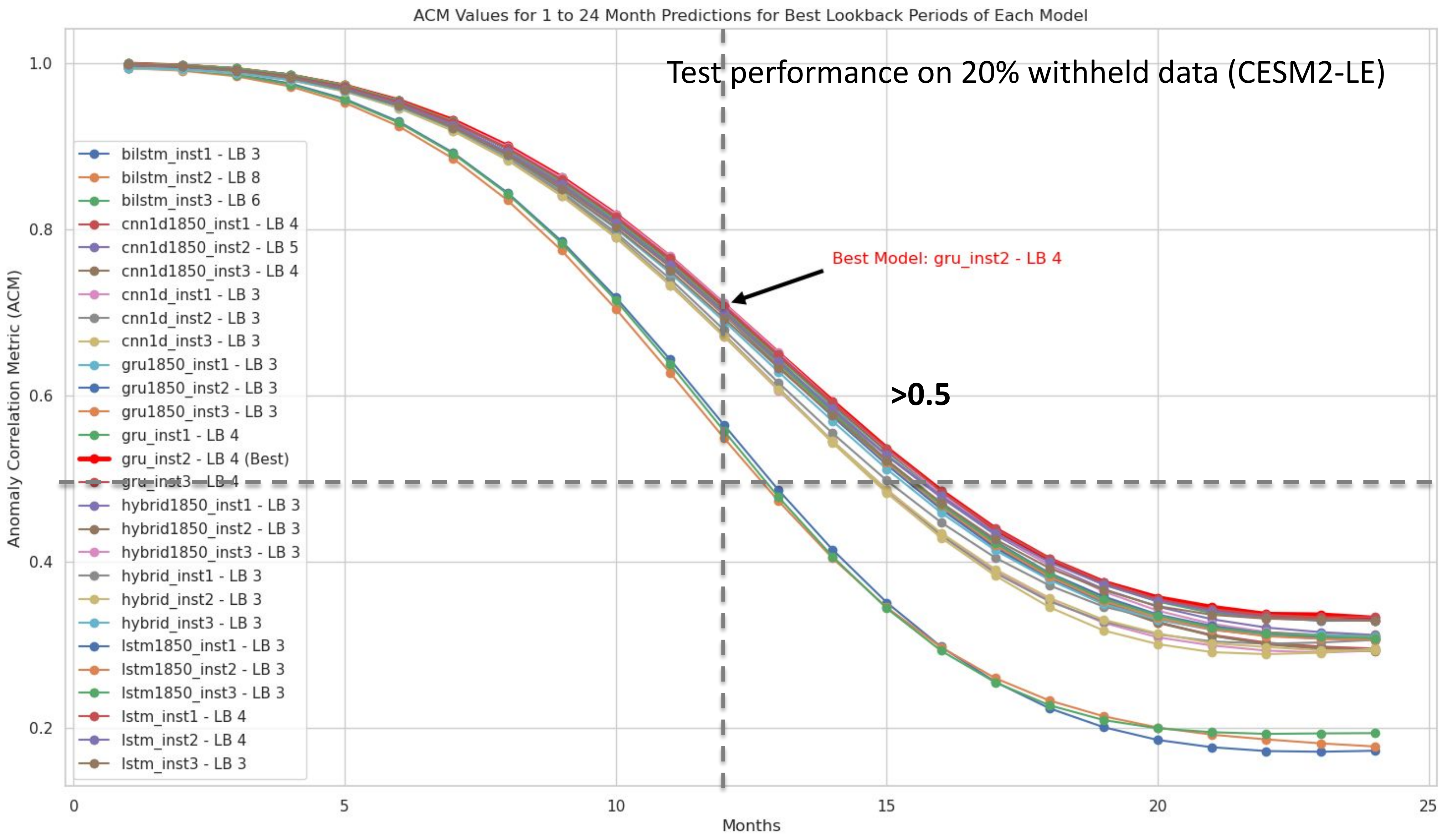


CESM2-LE

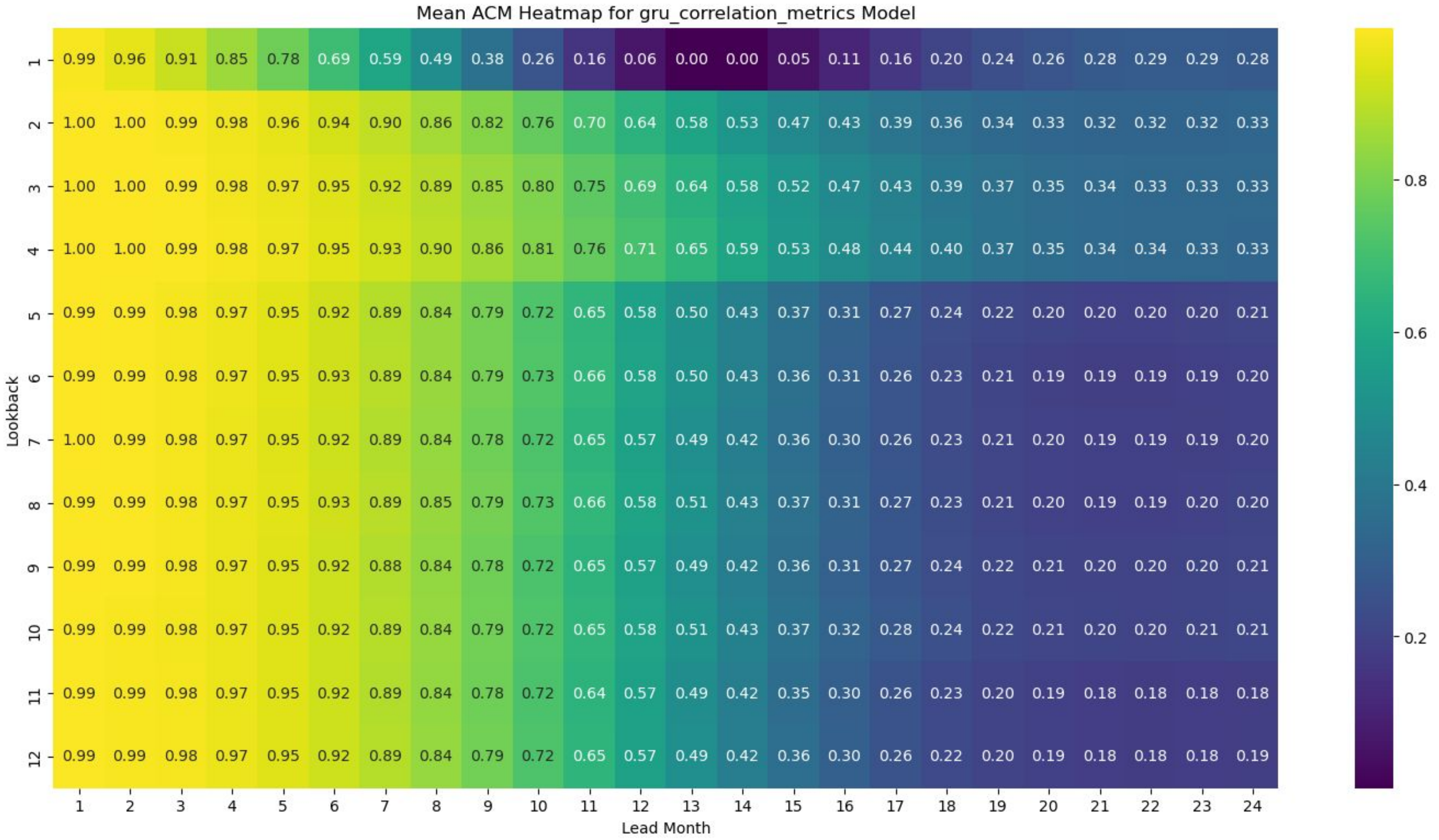
Model Selection Results –CESM 2 LE



DL model performances on CESM2-LE data

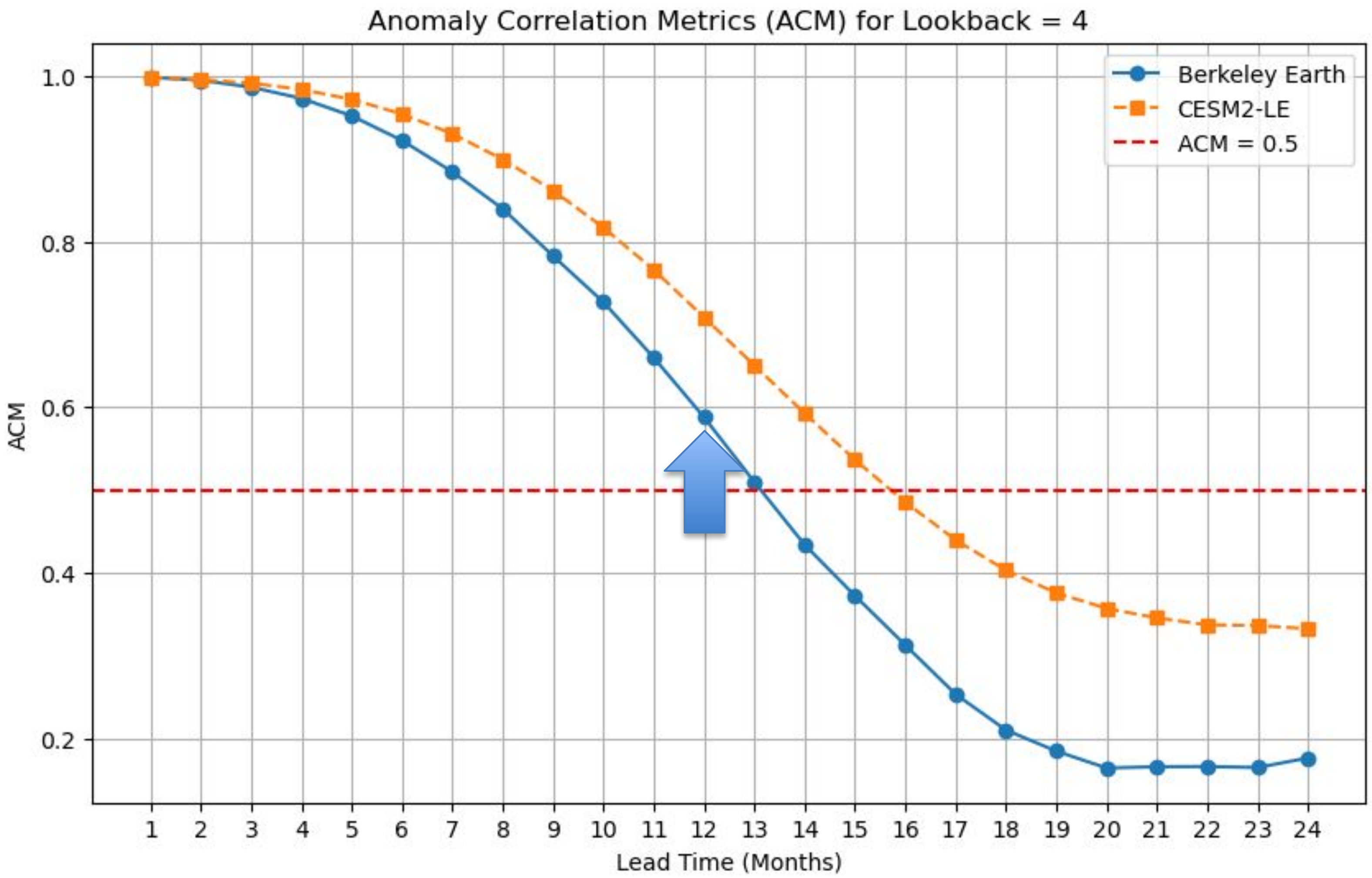


Effects of Lookback



DL model performances trained using CESM2-LE data, and verified using Berkely Earth (1950-2017)

GRU model Trained on CESM2, Applied on Berkeley Earth



ACC = 0.6 @ 12 months lead (comparable or better than SMYLE)

Transfer learning

01

Fine-Tuning a Pre-Trained Model :

- Load pre-trained weights and architecture.
- Remove the final Dense layer.
- Retain existing GRU layers for feature extraction.

02

Re-Train the Model:

- Add a new Dense layer for prediction.
  Train with early stopping to prevent overfitting.
- Use 1950 to 2017 to train and test

03

Evaluation
Use the Anomaly Correlation Metric (ACM) to assess performance.

04

Ungauged Dataset

Apply the fine-tuned model to a separate test dataset.
Measure ACM values for different lead months.

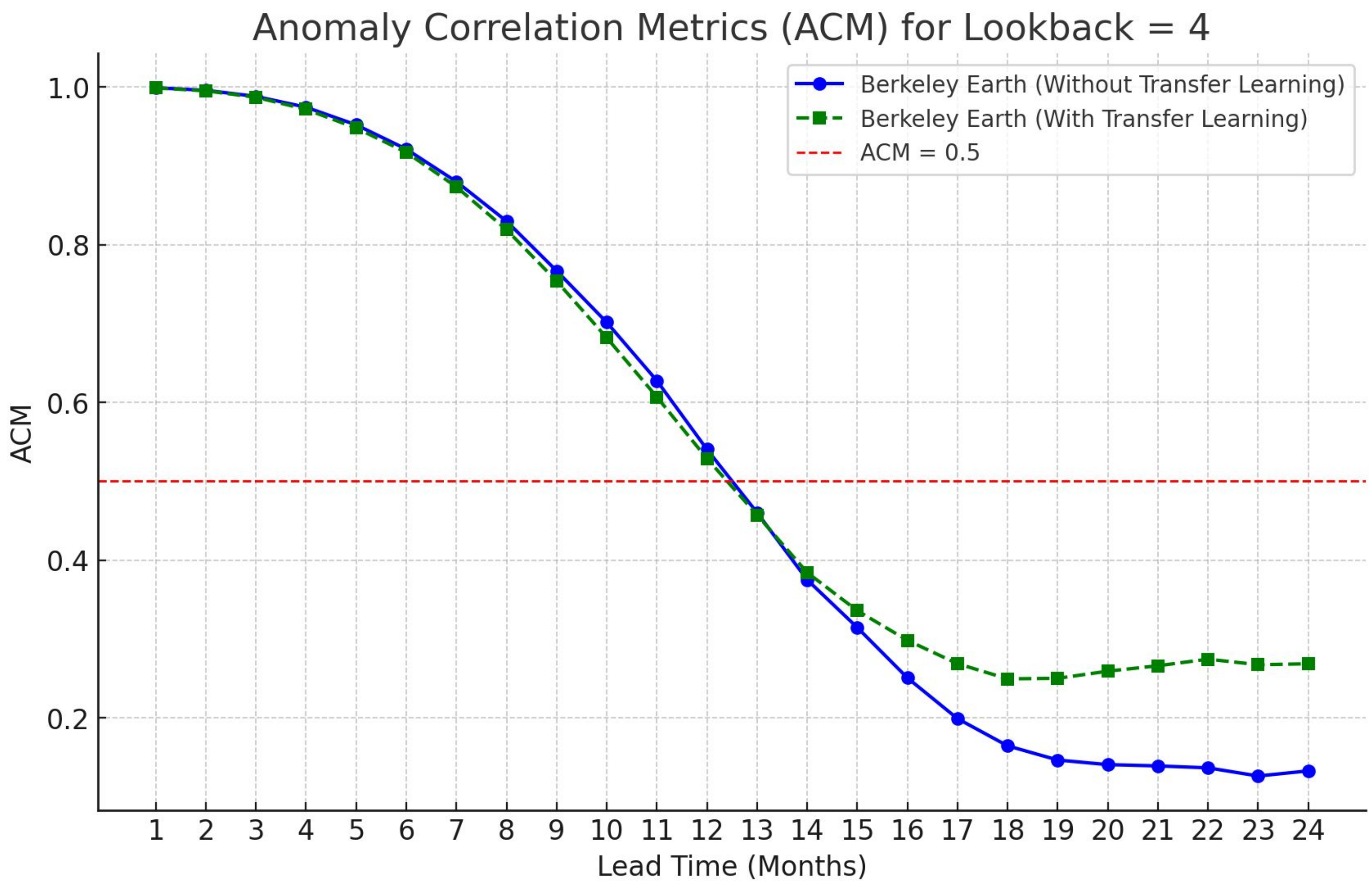
2017 to 2024 to predict

05

Save and visualize ACM results.
Store the fine-tuned model for future use.

DL model performances trained using CESM2-LE data, and verified using Berkely Earth (2004 to 2017)

Transfer Learning



Research Finidings

- **How does the length of the lookback period affect the performance and accuracy of DL models in ENSO prediction?**
  - Highest predictive performance is achieved with lookback periods of 1-6 months.
- **How does the number of principal components affect the performance of DL models in ENSO prediction?**
  - Increasing the number of principal components generally improved the performance of DL models by capturing more data variance. This was evident in the study as the Anomaly Correlation Metric (ACM) increased with the inclusion of more principal components.
- **Which model configuration is most effective for long-term ENSO predictions?**
  - The study found that GRU was the most effective configurations for accurate long-term ENSO predictions. These models leveraged their strengths in extracting spatial and temporal features.
- **Investigating the use of transfer learning to enhance model performance on low- availability datasets in ENSO forecasting, and determining which model is most effective**
  - Transfer learning significantly enhanced model performance on low-availability datasets. The GRU model, which integrates diverse architectures, was the most effective due to its ability to leverage knowledge from larger, high-quality datasets.

Future work

- Further enhance deep learning models and transfer learning techniques.
- Investigate the use of advanced methods such as transformer models.
- Broaden the range of input features by incorporating additional climate variables, such as ocean heat content (OHC).
- Extend the application to other climate factors.

References

•Cane, M. A., Zebiak, S. E., & Dolan, S. C. (1986). Experimental forecasts of El Niño. *Nature*, 321(6073), 827-832.

•Latif, M., & Barnett, T. P. (1994). Causes of decadal climate variability over the North Pacific and North America. *Science*, 266(5185), 634-637.

•Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2019). Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. *Hydrology and Earth System Sciences*, 23(11), 5109-5129.

•Frame, J., Sun, X., Anderson, W., Hossain, F., & Lettenmaier, D. P. (2021). Post-processing tools for National Water Model outputs using Long Short-Term Memory networks. *Journal of Hydrology*, 596, 125731.

•Barnes, E. A., Hurrell, J. W., & Ebert-Uphoff, I. (2019). Advances in the use of artificial neural networks for the analysis and prediction of atmospheric circulation patterns. *Geophysical Research Letters*, 46(22), 13470-13480.

•Cane, M. A., Zebiak, S. E., & Dolan, S. C. (1986). Experimental forecasts of El Niño. Nature, 321(6073), 827-832.

•Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2019). Rainfall-runoff modelling using Long Short-Term Memory (LSTM) networks. Hydrology and Earth System Sciences, 23(11), 5109-5129.

•Frame, J., Sun, X., Anderson, W., Hossain, F., & Lettenmaier, D. P. (2021). Post-processing tools for National Water Model outputs using Long Short-Term Memory networks. Journal of Hydrology, 596, 125731.

•Barnes, E. A., Hurrell, J. W., & Ebert-Uphoff, I. (2019). Advances in the use of artificial neural networks for the analysis and prediction of atmospheric circulation patterns. Geophysical Research Letters, 46(22), 13470-13480.

•Kumar, A. (2015). A Study of Dimensionality Reduction Techniques and its Analysis on Climate Data.

•Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development, 9(5), 1937-1958. DOI:10.5194/gmd-9-1937-2016

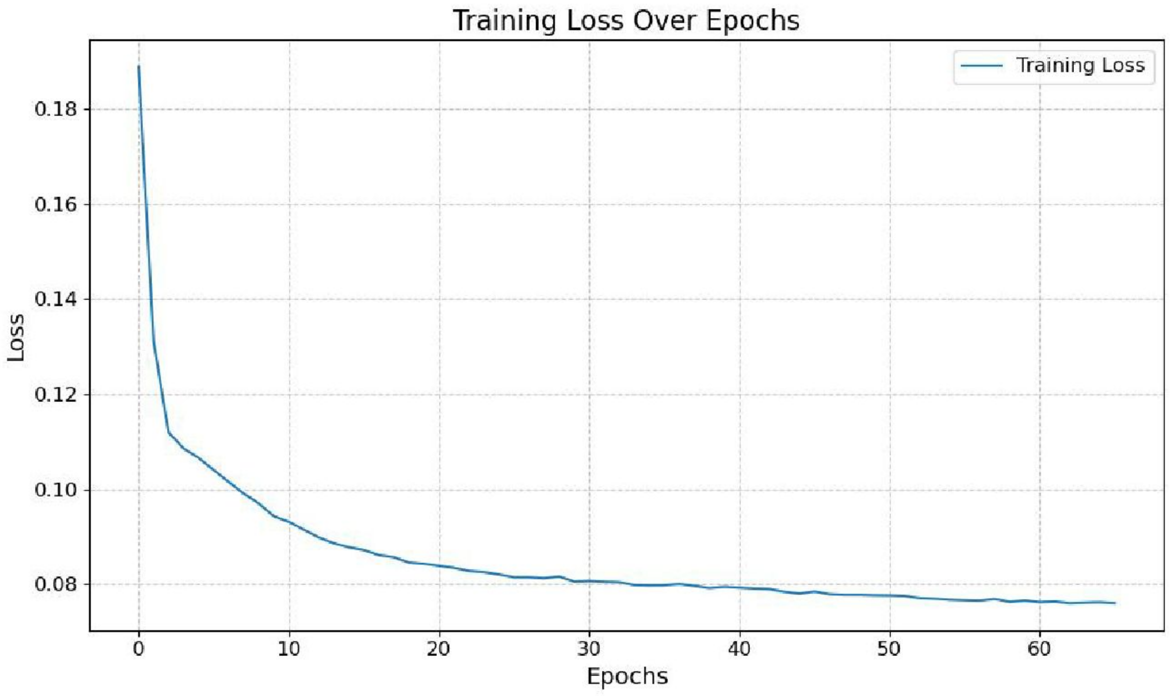
•Rohde, R., et al. (2013). "A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011." Geoinformatics & Geostatistics. DOI: 10.4172/gigs.1000101

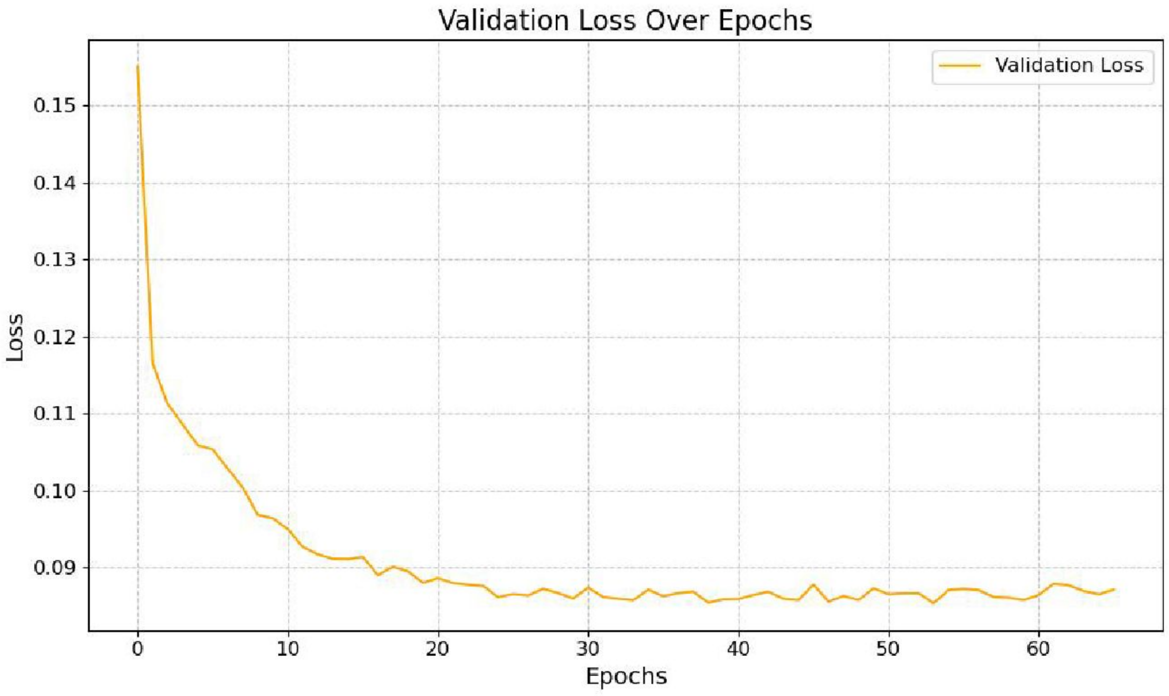
• Danabasoglu, G., et al. (2020). "The Community Earth System Model Version 2 (CESM2)." Journal of Advances in Modeling Earth Systems. DOI: 10.1029/2019MS001916

•Rodgers, K. B., et al. (2021). "The second phase of the CESM2 Large Ensemble Community Project." Earth System Dynamics. DOI: 10.5194/esd-2021-50

•Dueben, P. D., & Bauer, P. (2023). Challenges and design choices for global weather and climate models based on machine learning. Geoscientific Model Development, 16(1), 6433-6445. https://doi.org/10.5194/gmd-16-6433-2023

•Kim, D.-H., Moon, I.-J., Lim, C., & Woo, S.-B. (2024). Improved prediction of extreme ENSO events using an artificial neural network with weighted loss functions. Frontiers in Marine Science, 10. https://doi.org/10.3389/fmars.2023.1309609


Training Loss Over Epochs
Training Loss
Loss
Epochs
0.18
0.16
0.14
0.12
0.10
0.08
0
10
20
30
40
50
60


Validation Loss Over Epochs
Validation Loss
Loss
Epochs
0.15
0.14
0.13
0.12
0.11
0.10
0.09
0
10
20
30
40
50
60