Predictability of Summer Surface Maximum Temperature Extremes on Sub-seasonal Scale over CONUS

M. M. Nageswararao[1] and Vijay Tallapragada[2]

[1]UCAR visiting scientist at NOAA/WCWCP/NCEP/EMC,

[2] NOAA/WCWCP/NCEP/EMC

Email: murali.n.malasala@noaa.gov ; murali@ucar.edu

Acknowledgements:

This research was supported by the NCEP Visiting Scientist Program, managed by UCAR CPAESS. The authors are grateful to the NOAA NCEP EMC Ensemble Group.

Session: Innovative Forecasting Techniques and Tools

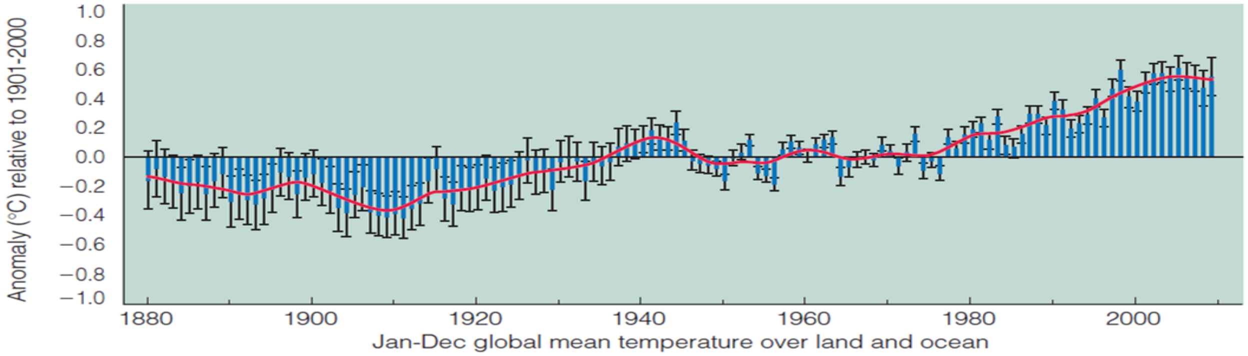
Date and Time : June 7, 2024 , 1:30-3:30 MT

NOAA's S2S Community Workshop: Toward Minimizing Early Model Biases and Errors in S2S Predictions, Boulder, CO, 5–7 June 2024

Background and Motivation:

- Human-induced climate change is not only increasing global temperatures but also intensifying extreme weather events worldwide, notably affecting developed and developing countries.

- In recent decades, global temperatures have risen significantly, with the last three decades being the warmest since 1850.



- Hot days in the USA have increased since the mid-20th century and are expected to rise further, straining health services, causing deaths, stressing water resources, transportation, and energy systems, and increasing wildfires (Seneviratne et al. 2012; Changnon et al. 1996).
- Accurate prediction of T_{max} and it extremes on a sub-seasonal scale over CONUS is crucial. But it is challenging compared to short-term forecasts and seasonal outlooks. This is due to the loss of initial atmospheric condition memory on this time scale, affecting prediction skills.
- In September 2020, NOAA NCEP implemented GEFSv12 to support stakeholders with sub-seasonal forecasts and hydrological applications. GEFSv12 reforecast data for 2000-2019 are initialized at 00 UTC once per day for up to 16 days with 5 ensembles, except on Wednesdays when the integration extends to 35 days with 11 members.
- NWP's direct raw products are rarely used due to high uncertainty and lack of prediction skill, especially on the sub-seasonal scale. Statistical post-processing is essential for improving forecast guidance and usability.

Data Used

Model : GEFSv12 (Zhou et al. 2019; 2021)

The period used : 2000-2019

Horizontal Resolution: 0.25° X 0.25° for Day-1 to 10 and 0.5° X 0.5° for Day-11 to 35. The entire data is interpolated by using bilinear interpolation over CONUS with 0.5° X 0.5°

Members used : 11 members (c00, p01, p02, p03, p04, p05, p06, p07, p08, p09, and p10) based on every Wednesday 00 UTC initial conditions.

Reference data set used : CPC T_{max}

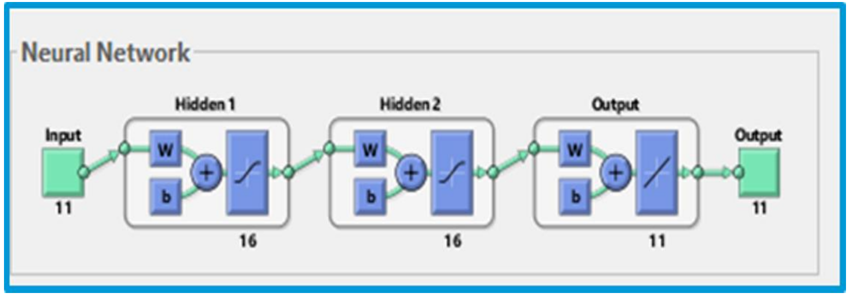
Detrended Quantile Mapping (QQ)

Detrended quantile mapping works by first removing linear trends from both observed and simulated data to focus on variability. Then, it adjusts the quantiles of detrended observed data to match those of detrended simulated data, ensuring consistency in variability. Finally, it makes adjustments to the original simulated data based on the quantile mapping results, correcting biases and discrepancies. This method enhances the realism of model simulations, making them more useful for various climate applications.

.

Artificial Neural Network (ANN)

- ANNs, inspired by the human brain, use multi-layered structures to process complex data and identify important patterns.
- ANN techniques are excellent at grouping data, sorting categories, making predictions, and forecasting time-series patterns using connected nodes.
- ANN algorithms combine inputs at each node with adjustable weights and biases, which are tweaked during training to reduce errors.
- Among ANN architectures, the Feedforward Neural Network (FNN) is one of the simplest, with data flowing in one direction from input to output layers.
- Activation functions like ReLU and Tanh enable the model to learn from input data and perform complex tasks accurately.
- The Tanh function outputs values between -1 and +1, handling negative values better than the sigmoid function, which ranges from 0 to 1. Unlike sigmoid, tanh is zero-centered, making its output symmetric around the origin.
- Loss functions measure the difference between predicted output and actual targets, helping to improve ANN accuracy during training.

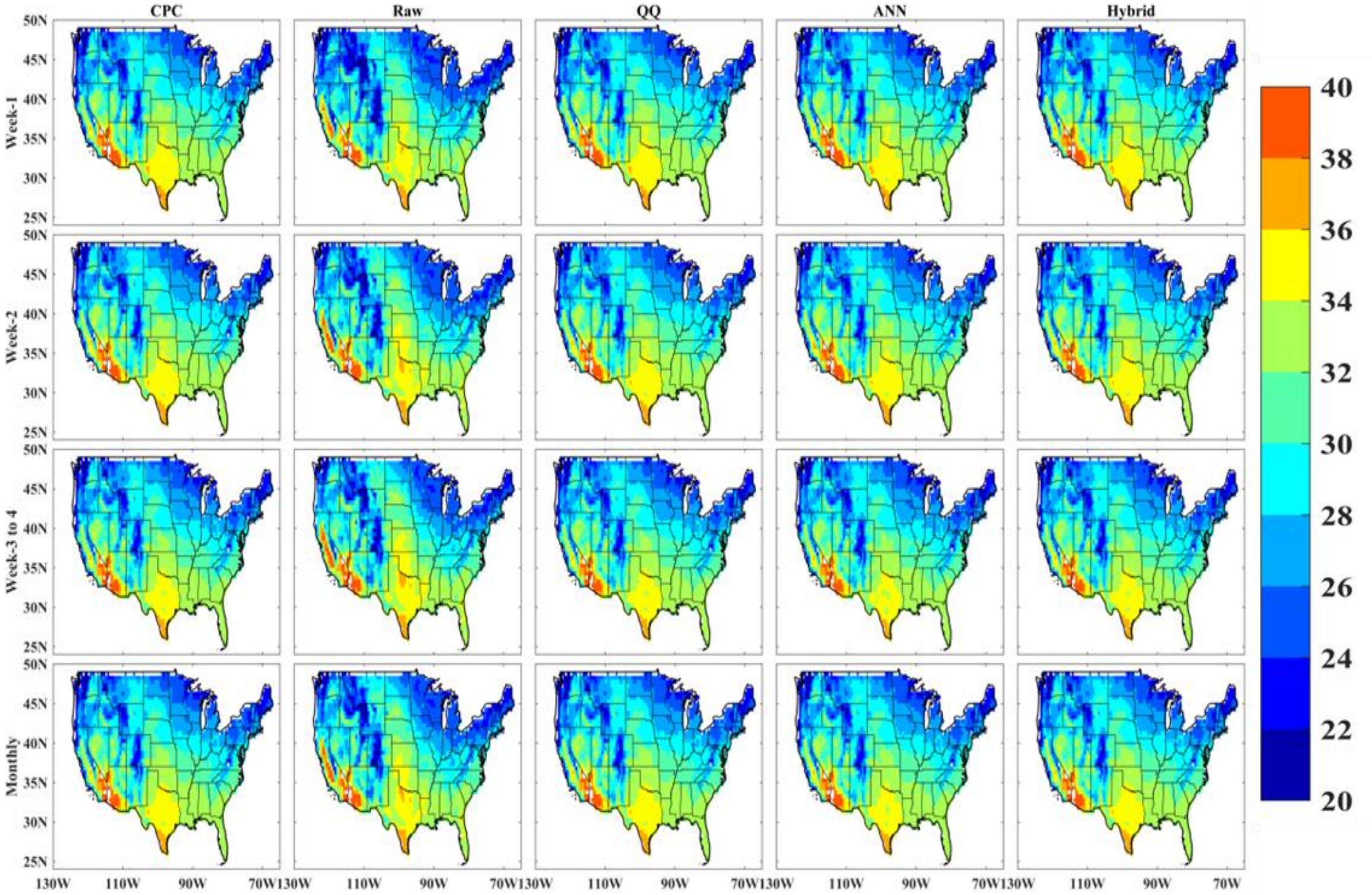


The following are considered when developing a basic Artificial Neural Network (ANN) model:

No. Hidden layers:	2
No. of nodes/neurons in the hidden layer	[16,16]
Neural Network used	Feedforward network
Activation Functions in Neural Networks	Hidden layer 1: ReLU Hidden layer 2: Tanh
Data divided function	70% data for training and 30% data for validation in random way
Learning rate	0.001
Max number of iterations/epochs used	1000
Error tolerance for stopping criterion	1e-14
Training function used	Supervised weight/bias training function with Sequential order weight/bias training (trains)
Neural Network Performance Functions used	Mean squared error performance function

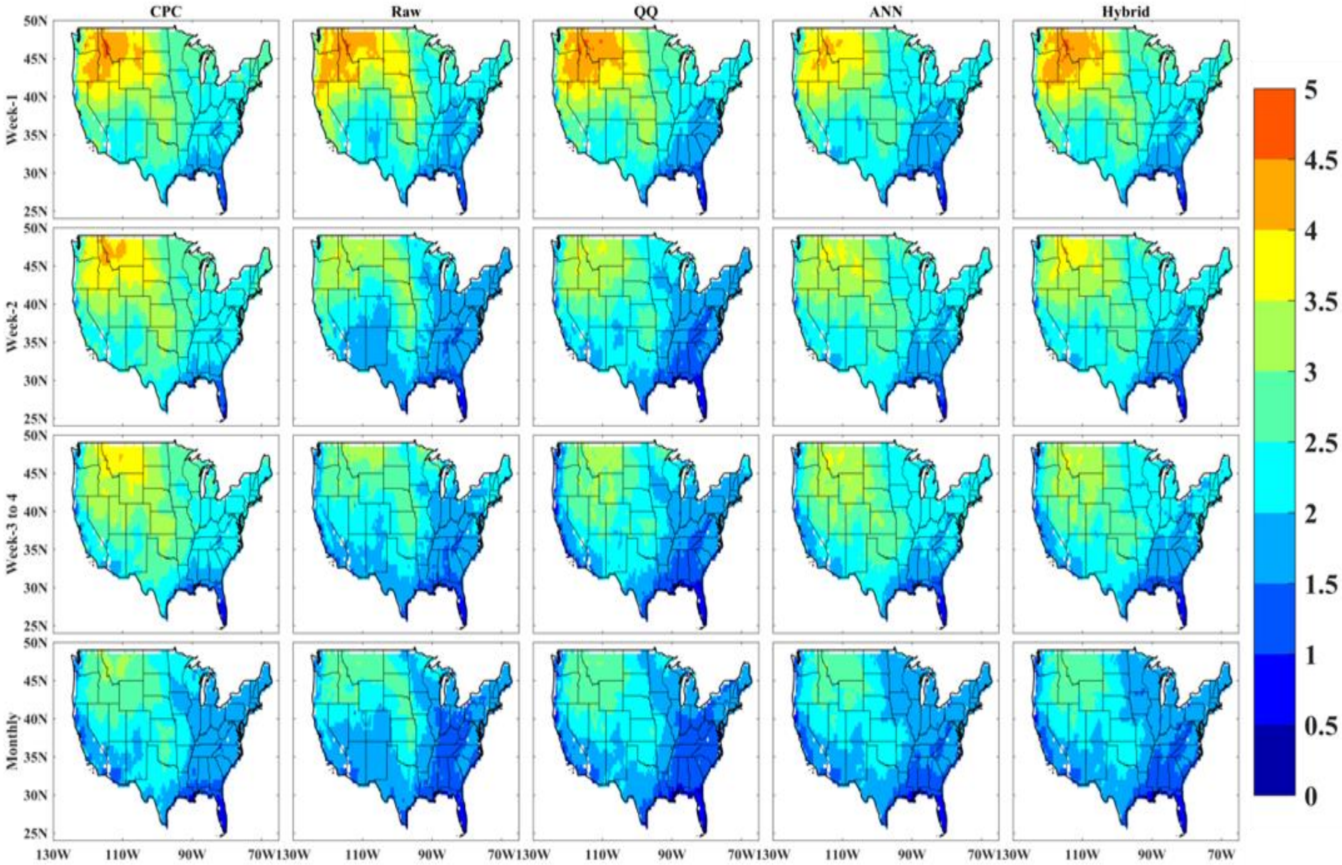
- The QQ method is applied to ANN output for further improvement. This method as hereafter mentioned Hybrid Post-processing (ANN-QQ; Hybrid).

Spatial pattern of JJA T_{max} (°C)



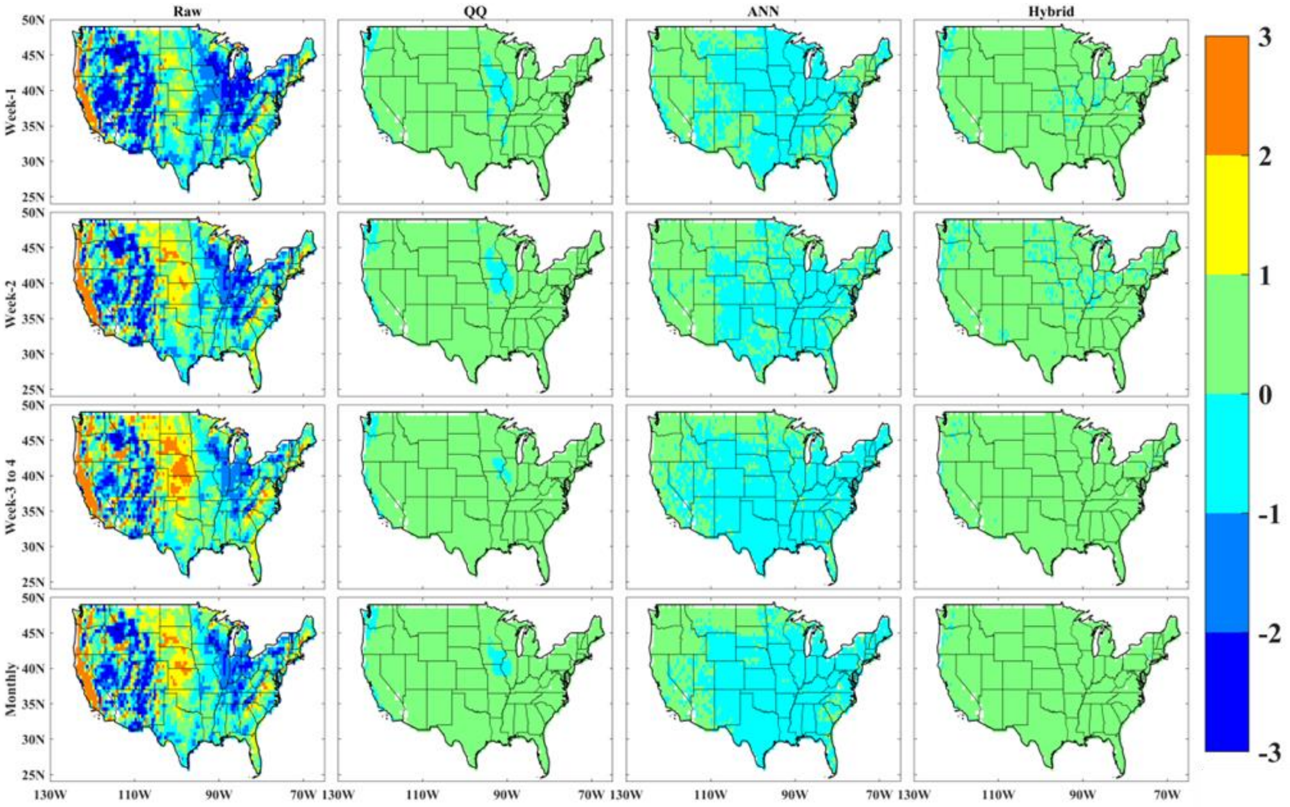
- GEFSv12 captures T_{max} patterns across CONUS for all forecast time scales.
- Cold bias noted for shorter lead times, while warm bias observed for longer lead times.
- Calibration methods significantly reduce biases across all lead time forecasts.

IAV of JJA T_{max} (°C)



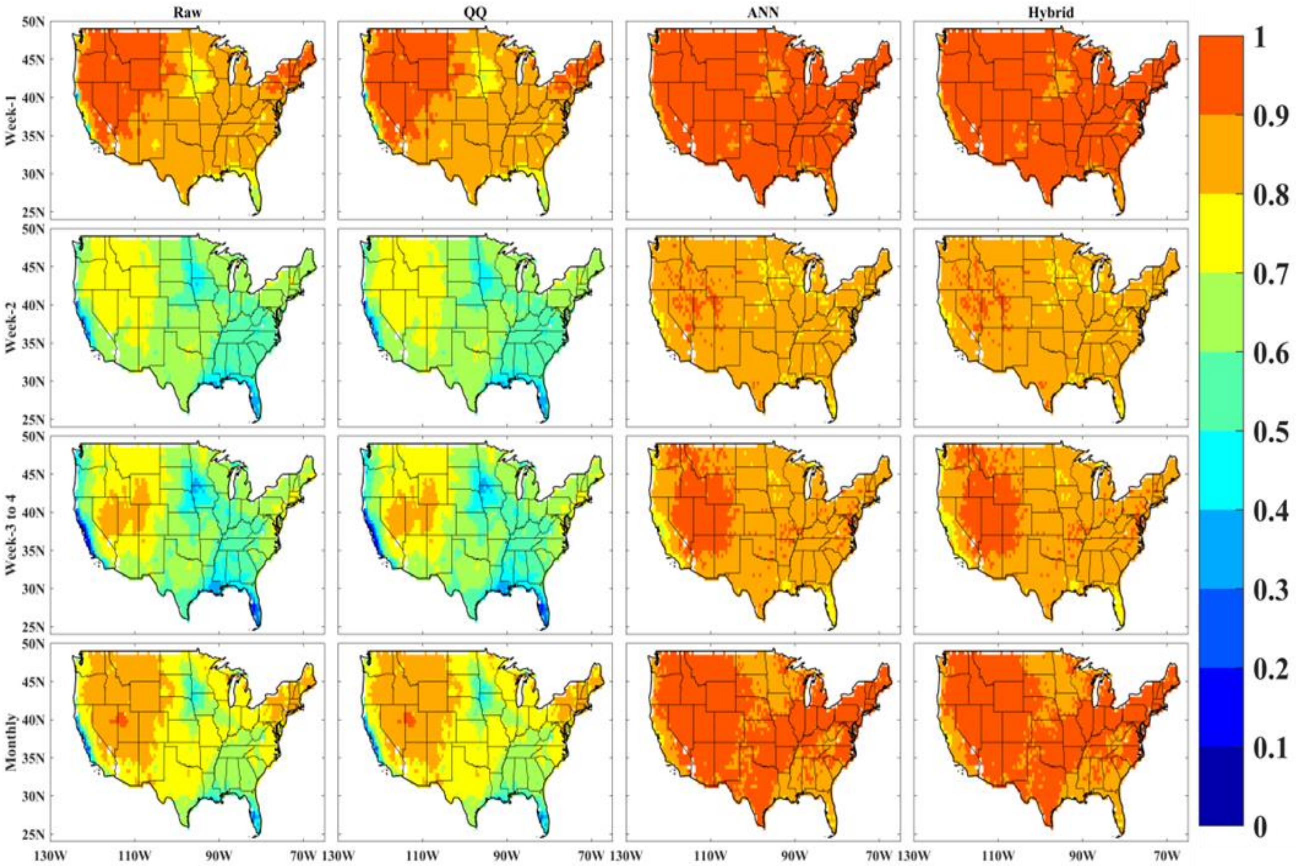
- GEFSv12 model replicates spatial patterns of IAV in T_{max} across CONUS at all forecast scales.
- Underestimation of IAV in T_{max} observed for longer forecast lead times.
- Various calibration methods increase IAV, notably ANN outperforms QQ.
- Hybrid method surpasses both QQ and ANN in enhancing IAV.

Mean Bias for JJA T_{max} (◦C)

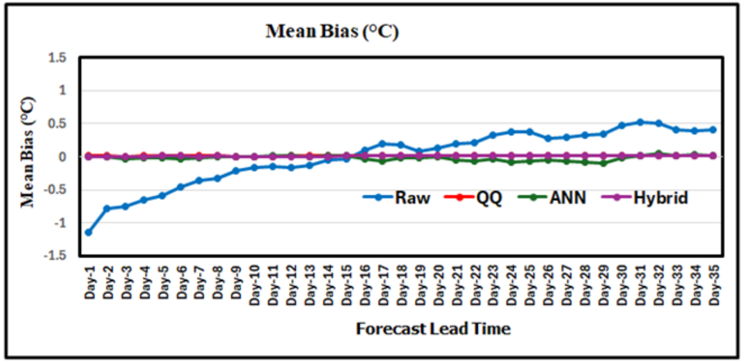


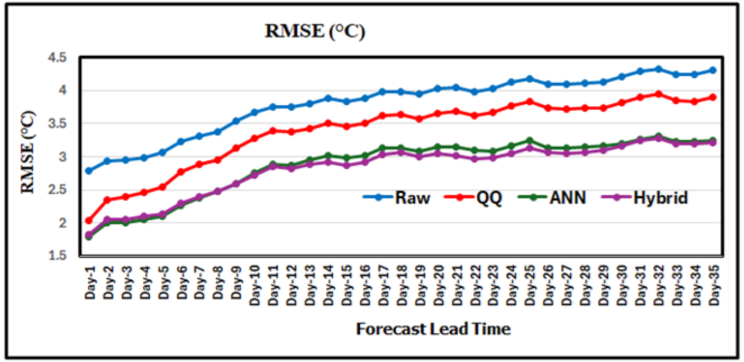
- Raw-GEFSv12 exhibits warm bias along the west coast, East Coast, and Central CONUS for all forecast lead times, with a significant cold bias elsewhere.
- Cold bias is more pronounced and spread out for shorter lead time forecasts.
- Calibration methods effectively mitigate biases, although ANN retains a slight cold bias in eastern part of CONUS; Hybrid method outperforms both ANN and QQ.

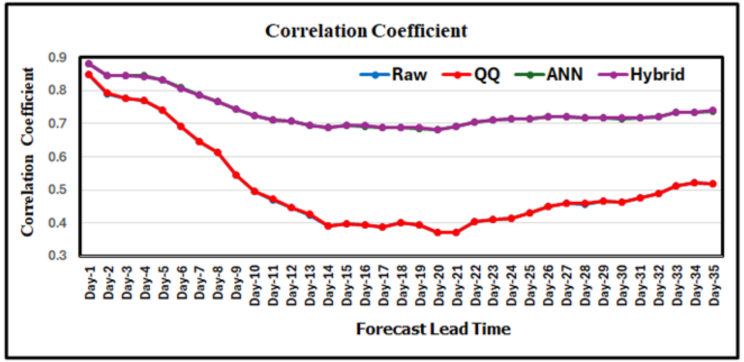
Correlation Coefficient (CC)

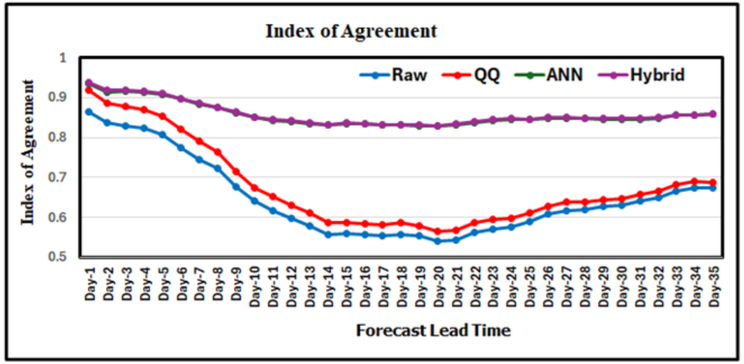


- Shorter lead times yield higher CC in predicting CONUS T_{max} with the Raw model.
- CC decreases as lead time increases, with Week-1 surpassing Week-2 forecasts.
- Longer forecast scales show higher CC, led by monthly forecasts, followed by Week-3 to 4, and Week-2. ANN and Hybrid methods notably boost CC, especially for longer lead times, while QQ CC remains consistent with the Raw product.



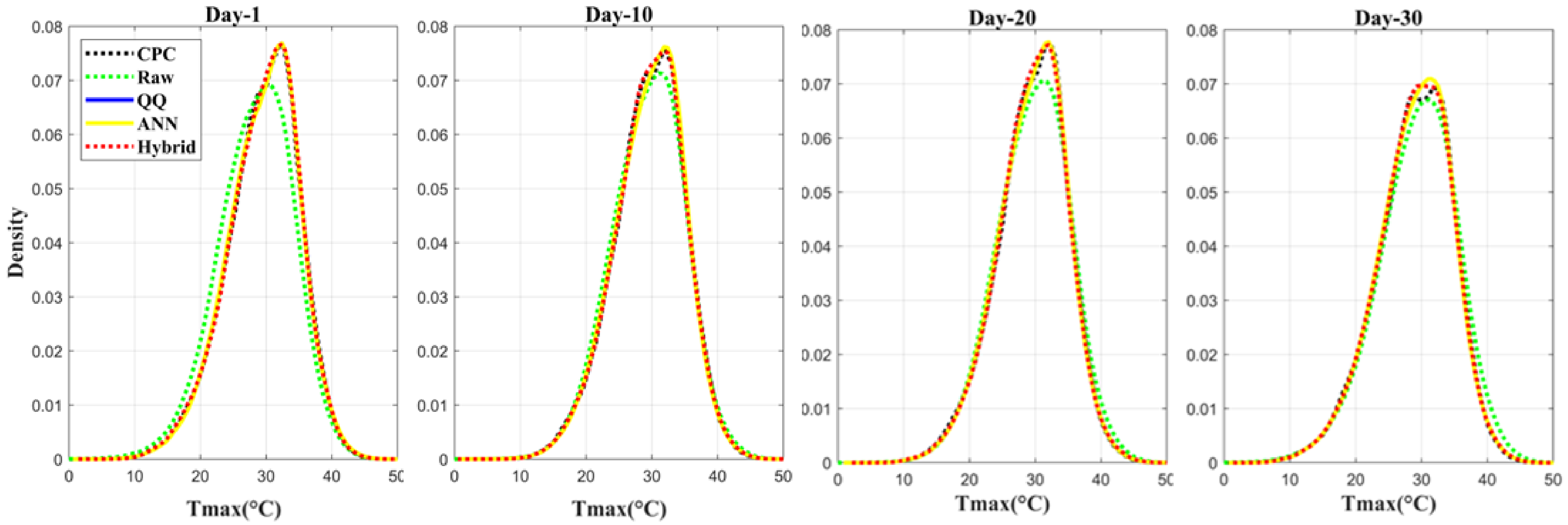






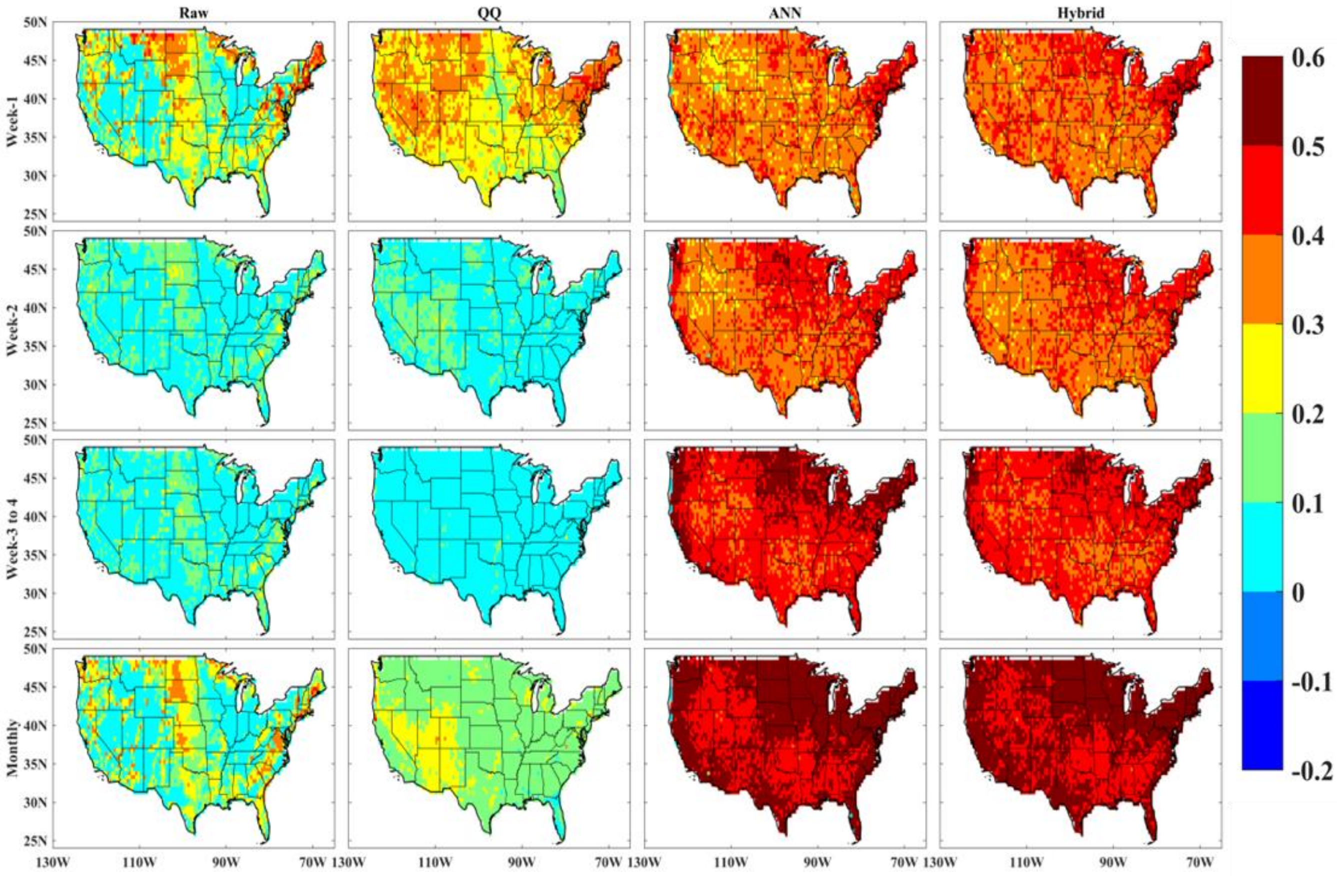
- Raw model exhibits decreasing cold bias after Week-2 and increasing warm bias with lead time. Calibration methods effectively mitigate biases, but ANN shows slight cold bias after Week-2.
- RMSE from Raw and all calibration methods is low for shorter lead times, increasing with lead time. All methods notably reduce RMSE across all lead times, with Hybrid outperforming ANN and QQ.
- CC and IOA are higher for shorter lead times, decreasing with lead time. Hybrid and ANN methods significantly enhance CC and IOA for all forecast lead times, while QQ skill scores remain similar to Raw product.

PDF of JJA daily T_{max} (°C) over CONUS with different lead time forecasts



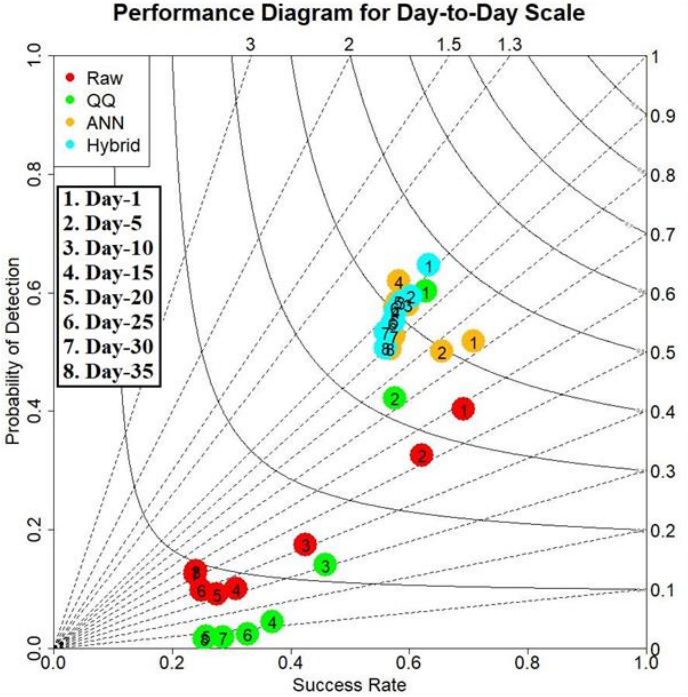
- Raw model's PDF patterns mimic observations but display slight left skew for shorter lead times and right skew for longer lead times in JJA T_{max} over CONUS, causing cold bias for shorter lead times and warm bias for longer lead times.
- All three calibration methods effectively match T_{max} intensity probabilities with observations. Hybrid and ANN distributions perform better than QQ, with the Hybrid method outperforming ANN

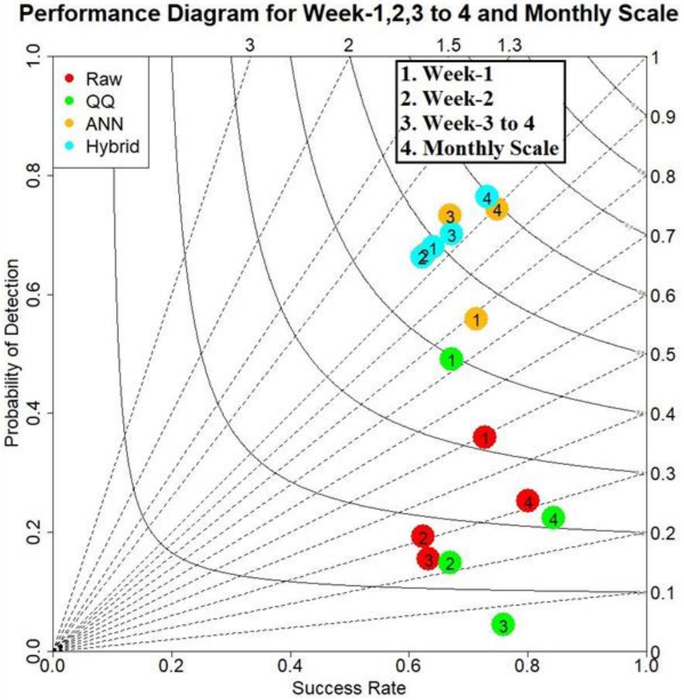
ETS for T_{max} Extreme days over CONUS from Raw, QQ, ANN and Hybrid Methods



- GEFSv12 demonstrates better ETS for shorter lead time forecasts.
- Both ANN and Hybrid methods notably enhance ETS across all forecast time scales compared to QQ. The improvement in ETS particularly more for longer lead time forecasts.
- ETS increases with forecast scale length, with ANN and Hybrid methods showing higher ETS for Monthly forecasts, followed by Week-3 to 4, Week-2, and Week-1.

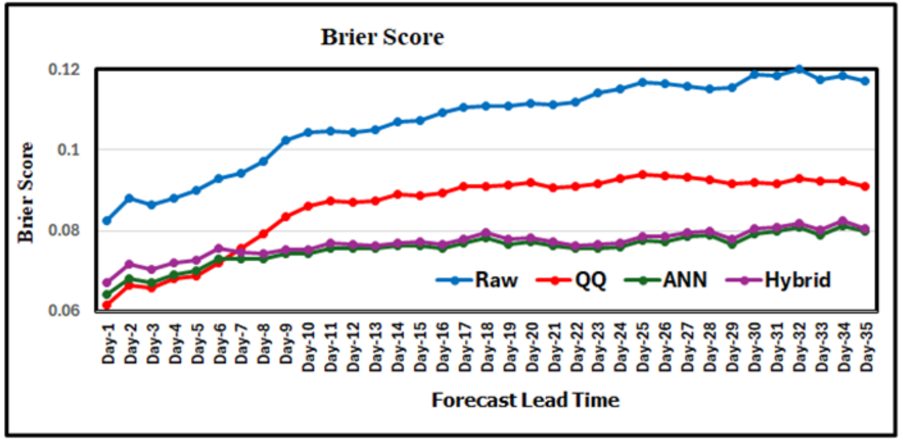
Performance Diagram for T_{max} Extreme days over CONUS On Day-to-Day and Sub-seasonal scales

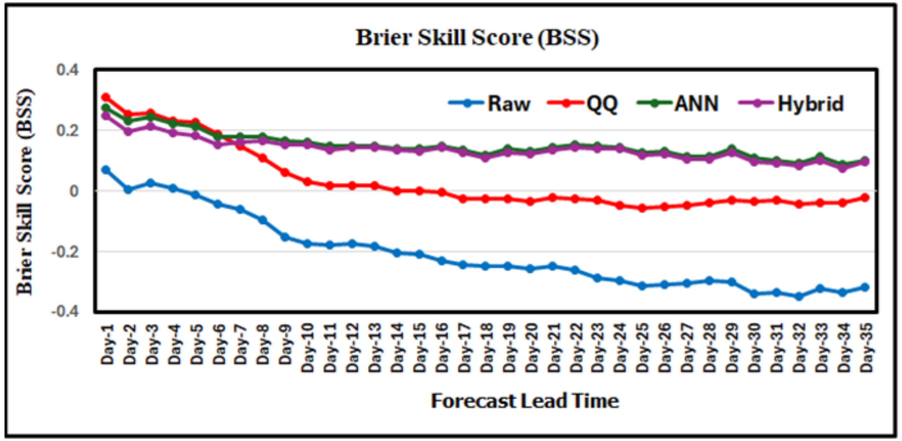




- Raw model consistently underestimates T_{max} extremes across all forecast time scales over CONUS, with lower statistical skill scores.
- ANN and Hybrid methods notably enhance most statistical categorical skill scores compared to the Raw model.
- Hybrid method surpasses both ANN and QQ across day-to-day and sub-seasonal time scales.

BS and BSS of Raw, QQ, ANN, and Hybrid methods against CPC for Ensemble probabilistic forecast of JJA daily T_{max} Extremes over CONUS

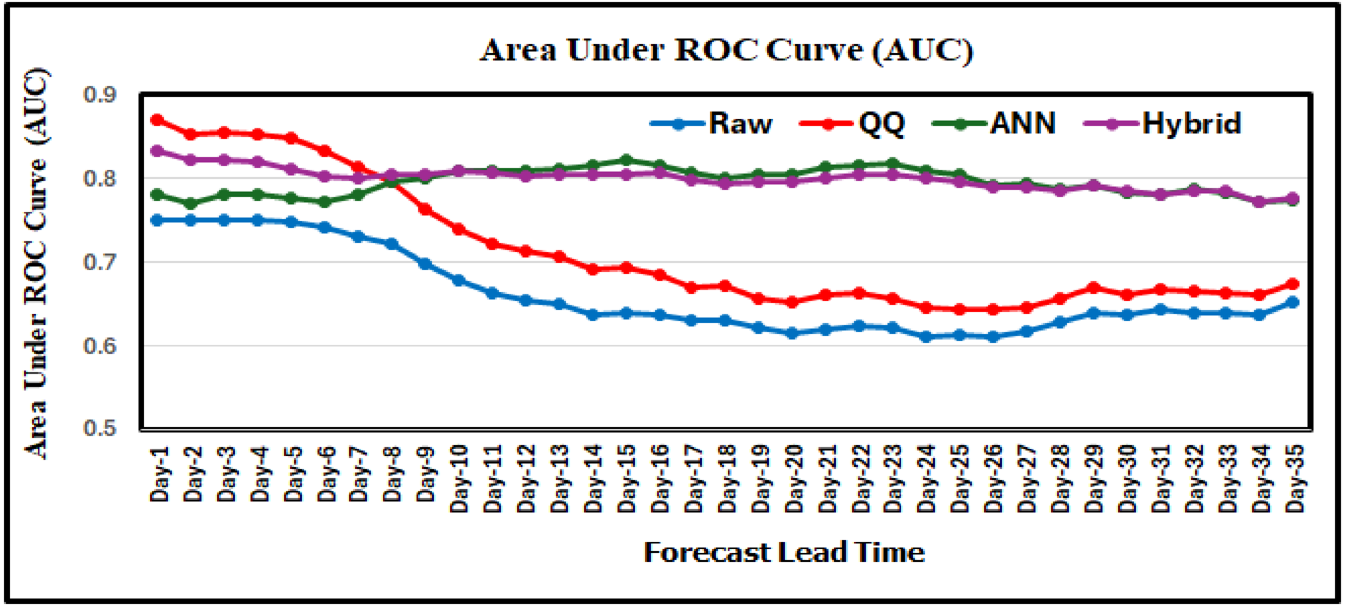




- GEFSv12's ensemble forecasts for daily JJA T_{max} extremes in CONUS exhibit high confidence (BS <0.12). Furthermore, all three calibration methods improve accuracy (BS <0.1), with ANN and Hybrid methods surpassing QQ, achieving consistently high accuracy across all lead times (BS<0.08).

- The Raw forecast shows inferior performance (BSS<0) compared to the reference forecast for JJA daily T_{max} extremes over CONUS. The QQ forecast performs better (BSS>0) for up to 10 days, but deteriorates beyond that, becoming worse than the reference forecast. In contrast, both the Hybrid and ANN methods consistently outperform the reference forecast (BSS>0) for all forecast lead times

Area Under ROC Curve (AUC) of Raw, QQ, ANN, and Hybrid methods against CPC for Ensemble Probabilistic Forecast of JJA daily T_{max} Extremes over CONUS



- The raw data and all three calibration methods demonstrate superiority (AUC>0.5) over the climatological/reference forecast for JJA daily T_{max} Extremes in CONUS.
- All three calibration methods improve the ensemble probabilistic forecast skill for predicting T_{max} extremes over CONUS.
- The ensemble probabilistic skill score of the Hybrid and ANN methods notably exceeds that of the QQ method across all forecast lead times.

Summary & Conclusions

- **GEFSv12 shows cold bias for short lead times and warm bias for long lead times in predicting daily JJA T_{max} over CONUS.**
- **Calibration methods, especially the Hybrid approach, improve skill scores, outperforming raw models and the QQ method.**
- **Shorter lead times have higher CC and IOA in predicting T_{max}, decreasing with longer lead times, but ANN and Hybrid methods boost these metrics.**
- **The raw model consistently underestimates T_{max} extremes, while ANN and Hybrid methods enhance skill scores for day-to-day and sub-seasonal scales.**
- **Ensemble forecasts for daily JJA Tmax extremes show high confidence, with calibration methods improving accuracy, especially beyond 10 days lead time.**
- **ANN and Hybrid methods outperform reference forecasts for all lead times, enhancing forecast skill for T_{max} extremes.**
- **Calibration methods, particularly the Hybrid approach, are crucial for improving forecast performance for T_{max} extremes.**

Thank You