Improving short to medium range GEFS precipitation forecasts in India using post-processing techniques

Sakila Saminathan[1], Subhasis Mitra[1], Di Tian[2], and Hanoi Medina[2]

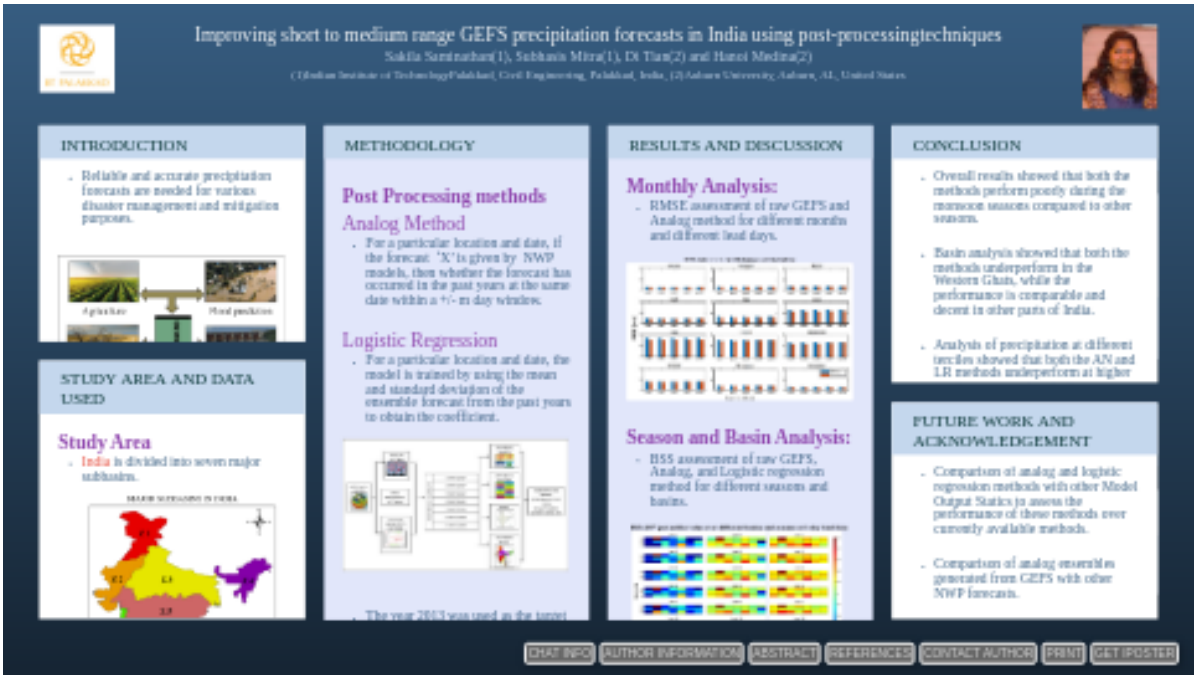
[1]Indian Institute of Technology Palakkad
[2]Auburn University

November 26, 2022

Abstract

This study aims to enhance the accuracy and reliability of the Global Ensemble Forecast System's (GEFS) precipitation forecasts over the Indian subcontinent using two post-processing techniques, namely the Analog method (AN) and Logistic Regression method (LR). The post-processing techniques and GEFS Numerical Weather Prediction Model (NWP) outputs were evaluated against the observed dataset using probabilistic and deterministic evaluation metrics. Results found that both the methods considerably improves short range to medium range (1-15 day) precipitation forecasts over India. Overall results showed that both the methods perform poorly during the monsoon seasons compared to other seasons. Basin analysis showed that both the methods underperform in the Western Ghats, while the performance is comparable and decent in other parts of India. Analysis of precipitation at different terciles showed that both the AN and LR methods underperforms at higher terciles compared to the lower ones. This is because the GEFS model itself was performing poorly in detecting the heavy precipitation events. The comparison of logistic regression and analog methods shows that the LR method outperforms the AN method in almost all the locations and lead times.

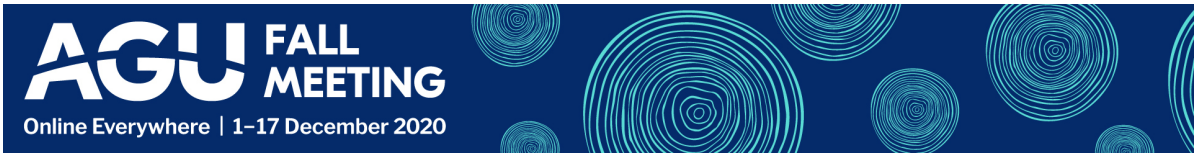
Improving short to medium range GEFS precipitation forecasts in India using post-processingtechniques



Sakila Saminathan(1), Subhasis Mitra(1), Di Tian(2) and Hanoi Medina(2)

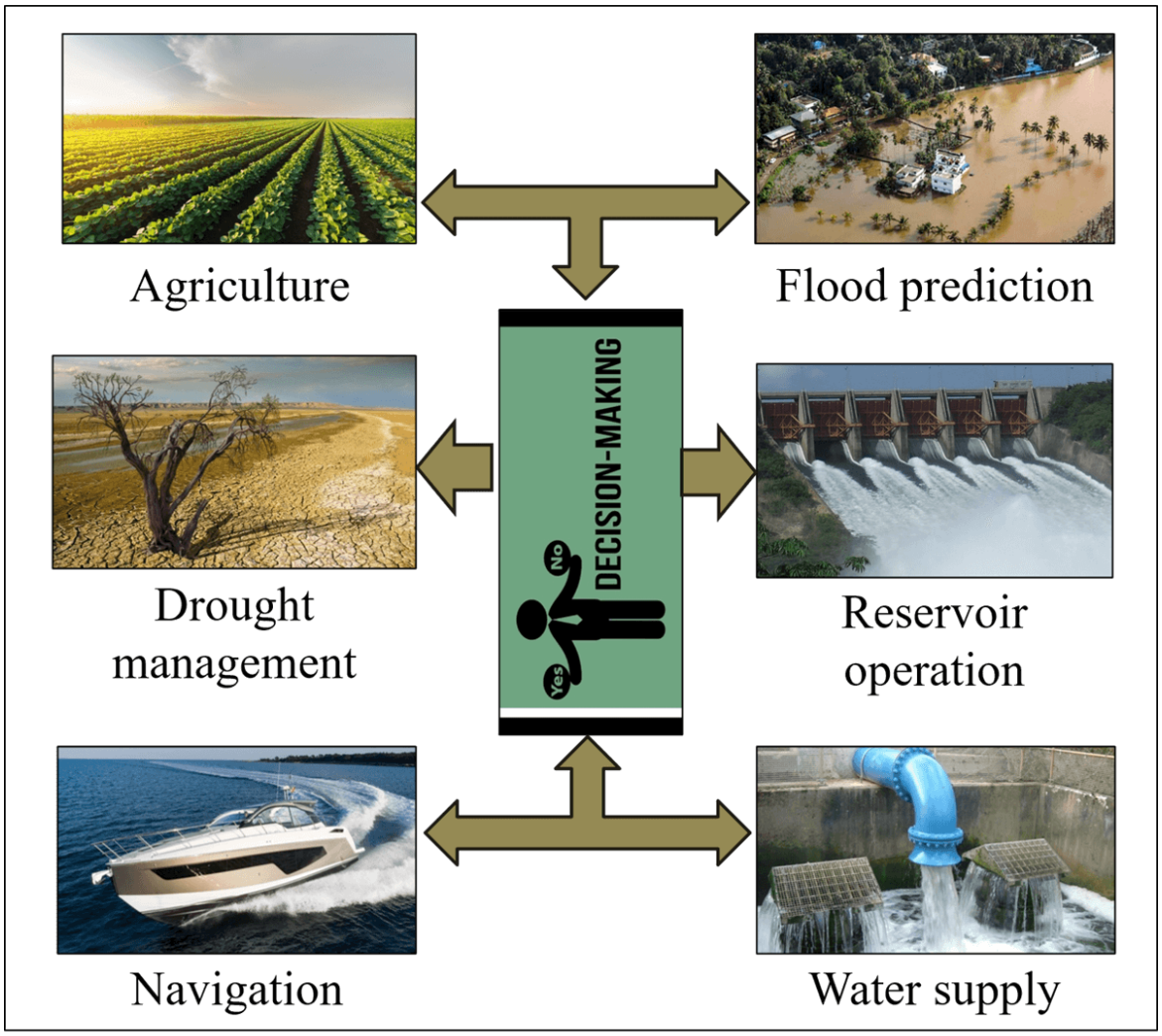
(1)Indian Institute of TechnologyPalakkad, Civil Engineering, Palakkad, India, (2)Auburn University, Auburn, AL, United States

PRESENTED AT:



INTRODUCTION

- Reliable and accurate precipitation forecasts are needed for various disaster management and mitigation purposes.



- Precipitation forecasts from numerical weather prediction models (NWP) exhibit systematic biases and hence needs to be post-processed before being used for further applications.
- In this study, a simple analog method and logistic regression are implemented to post-process short to medium range precipitation forecasts from the Global Ensemble Forecast System (GEFS) forecasts for the entire Indian sub-continent.

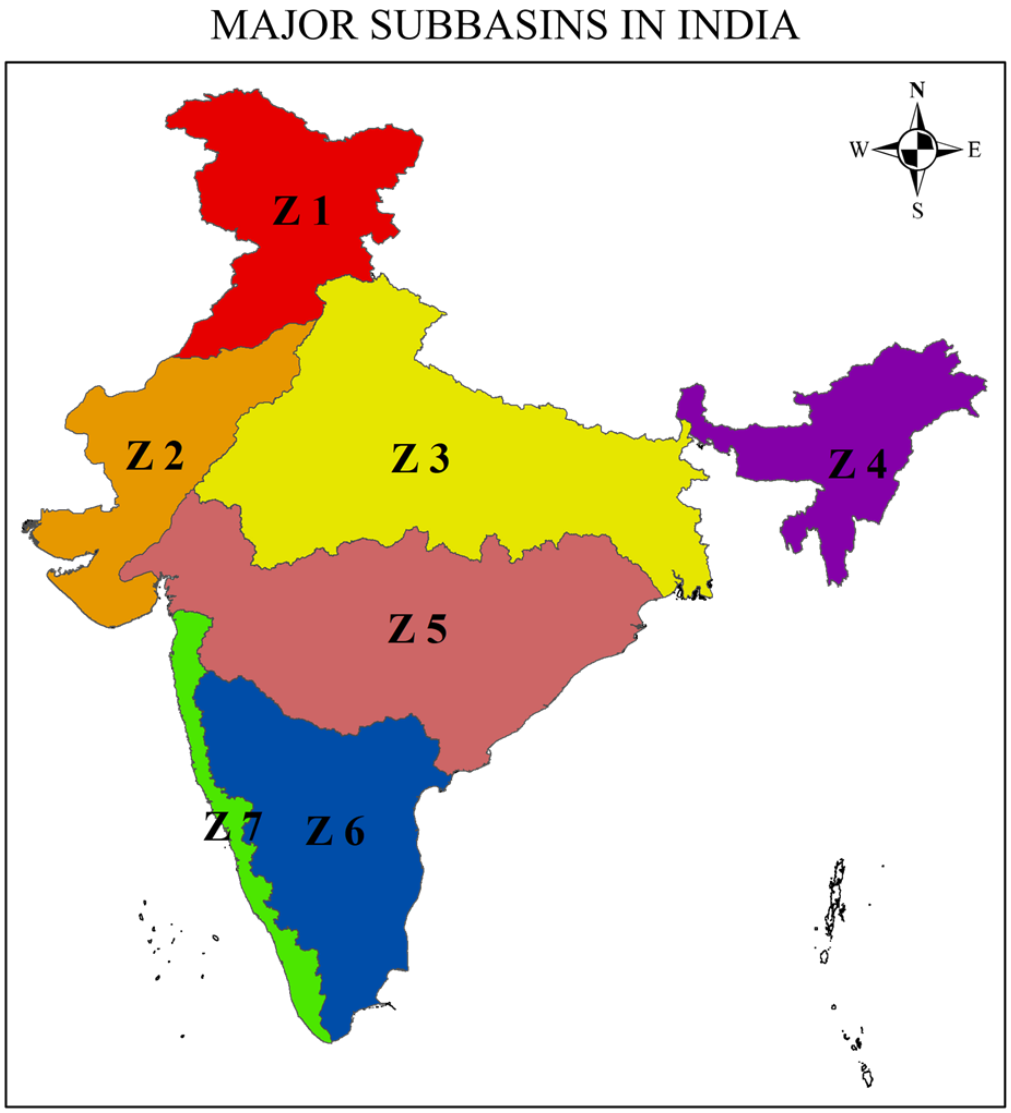
- The analog method (AN) uses the current forecast information and searches for closely related forecasts within a specified search window in the hindcast data to produce an ensemble of precipitation forecasts.

- Logistic regression (LR) is a postprocessing technique that yields a probabilistic forecast. The LR model is trained at a specific location and date using different predictors to obtain the predictand.

STUDY AREA AND DATA USED

Study Area

- India is divided into seven major subbasins.



- Northern basin: Z 1;

  Northwestern basin: Z 2;

  Indo Gangetic plain: Z 3;

  Northeastern basin: Z 4;

  Central basin: Z 5;

  Southern basin: Z 6;

  Western Ghats: Z 7;

- Season:

    1. Classification according to India Meteorological Department
    2. Winter, Summer, Monsoon, Postmonsoon

Data Used:

- Daily Precipitation data.
- Duration: 1985-2013.

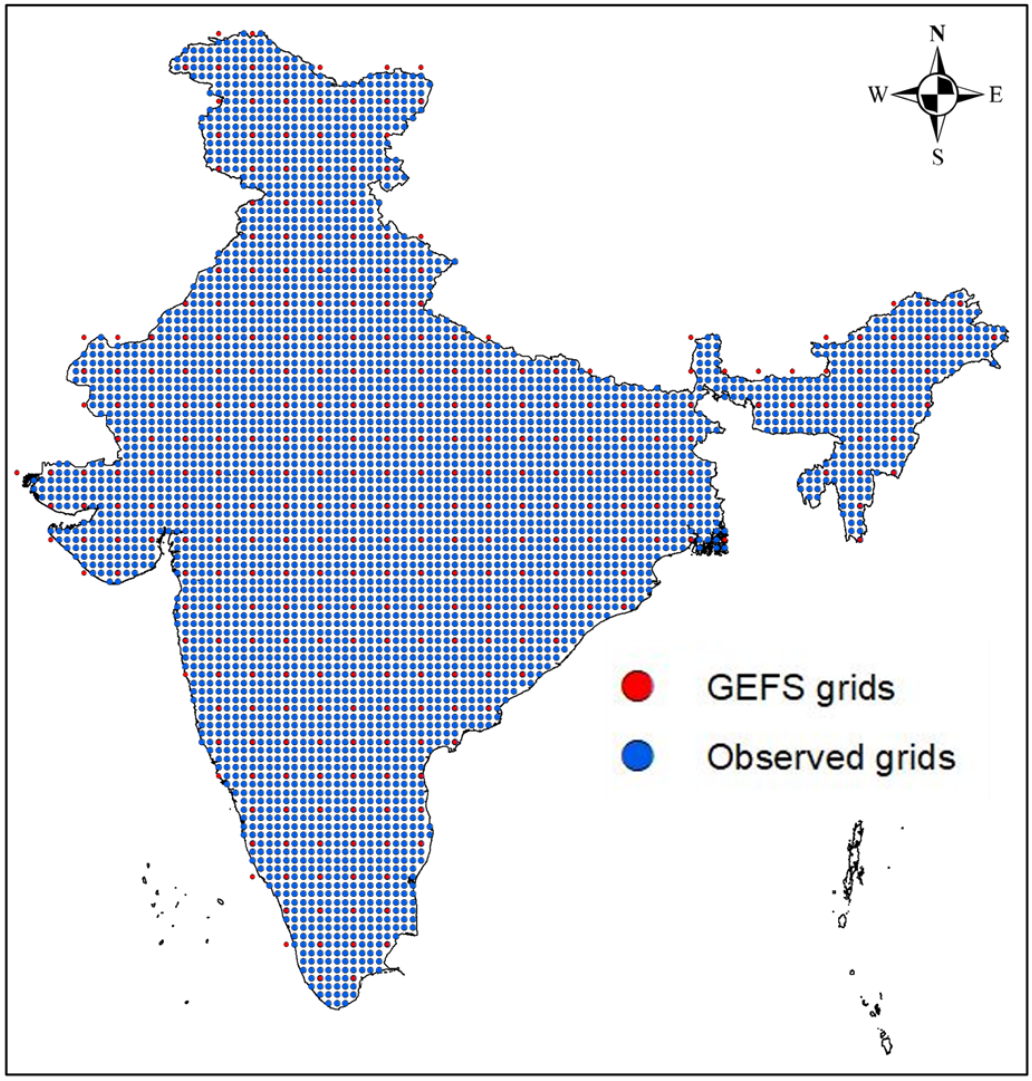
- Observed Data:

    1. India Meteorological Department [IMD]
    2. IMD data for validation.
    3. Spatial resolution: 0.25˚ X 0.25 ˚

- Forecast Data :

    1. Global Ensemble Forecast System [GEFS]
    2. Forecast lead time: 1 – 16 days.
    3. Ensemble members: 11 [1 control + 10 perturbed].
    4. Temporal resolution: 3 hours forecast data for 1st 3days and 6 hours for all 1-15 days.
    5. Spatial resolution: 1 ˚ X 1˚.

GEFS AND IMD OBSERVED GRID POINTS



METHODOLOGY

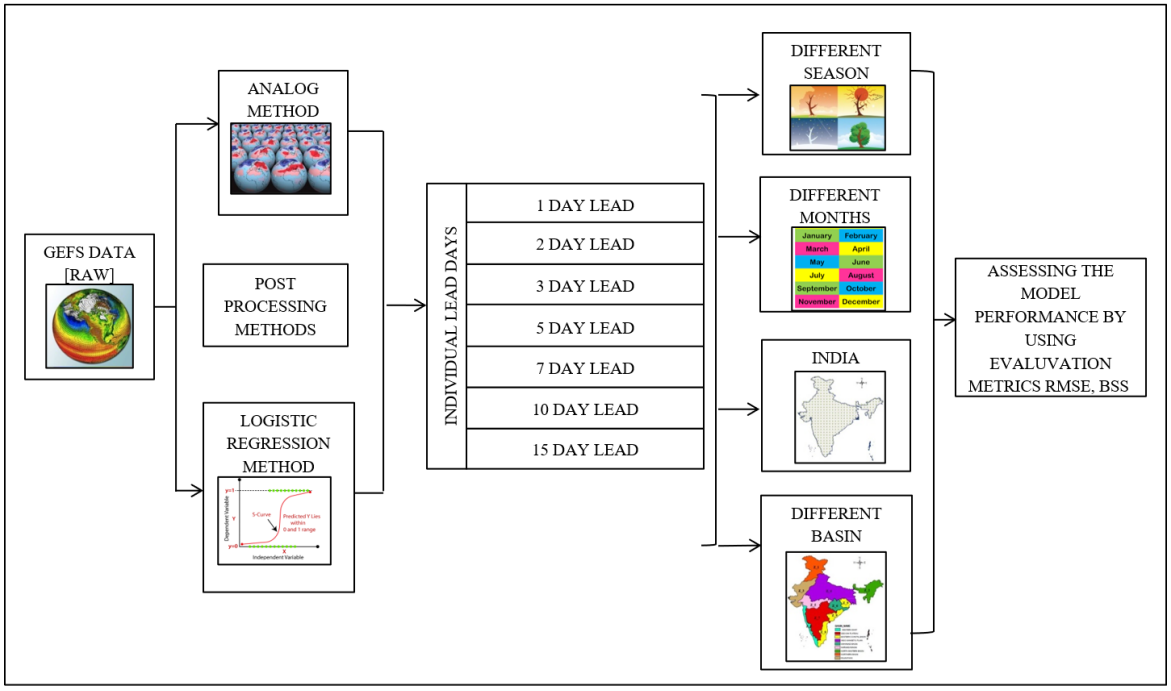
Post Processing methods

Analog Method

- For a particular location and date, if the forecast 'X' is given by NWP models, then whether the forecast has occurred in the past years at the same date within a +/- m day window.

Logistic Regression

- For a particular location and date, the model is trained by using the mean and standard deviation of the ensemble forecast from the past years to obtain the coefficient.



- The year 2013 was used as the target year to estimate the post-processed Precipitation forecast the GEFS forecast using the AN and LR method.

- For each day in the year 2013, logistic regression and analog ensemble forecasts were generated at 1, 2, 3, 5, 7, 10, 15 lead days.

- The analogs were estimated for 4964 grids points in the Indian region for the year 2013.

Evaluation Metrics:

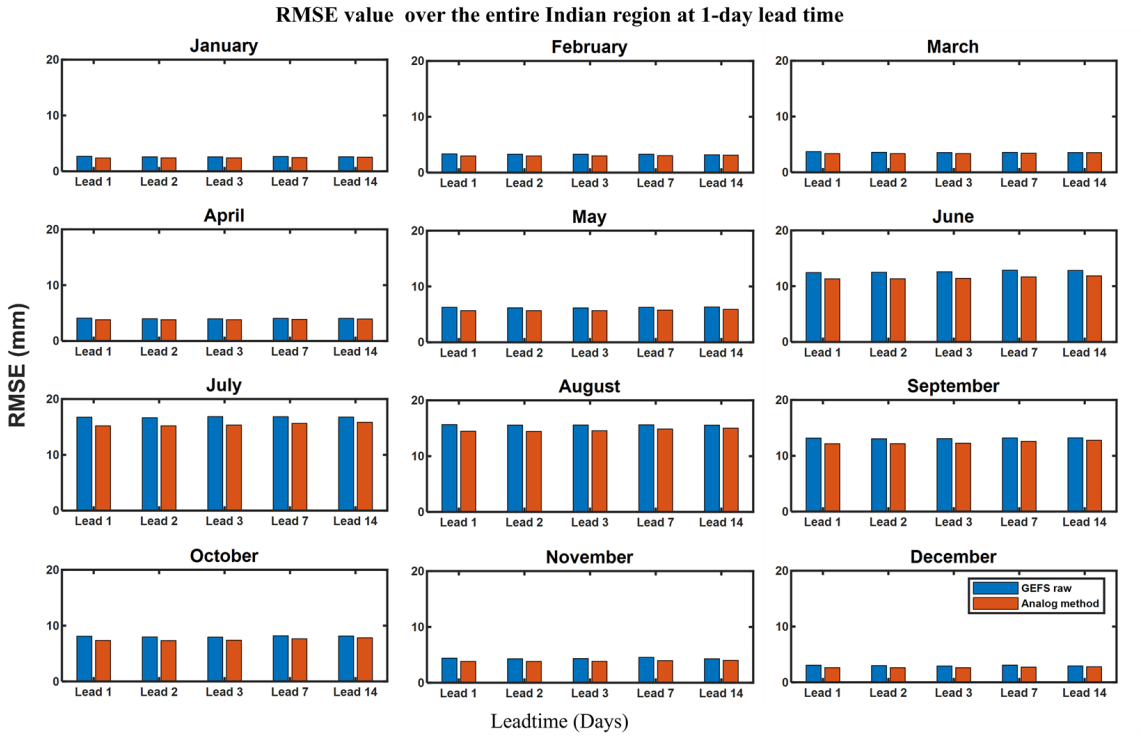
- Brier Skill Score $BSS = 1 - \frac{BS_{forecast}}{BS_{climatology}}$

- Root Mean Square Error $RMSE = \sqrt{\frac{\sum_{i=1}^{n} [F_i - O_i]^2}{n}}$

RESULTS AND DISCUSSION

Monthly Analysis:

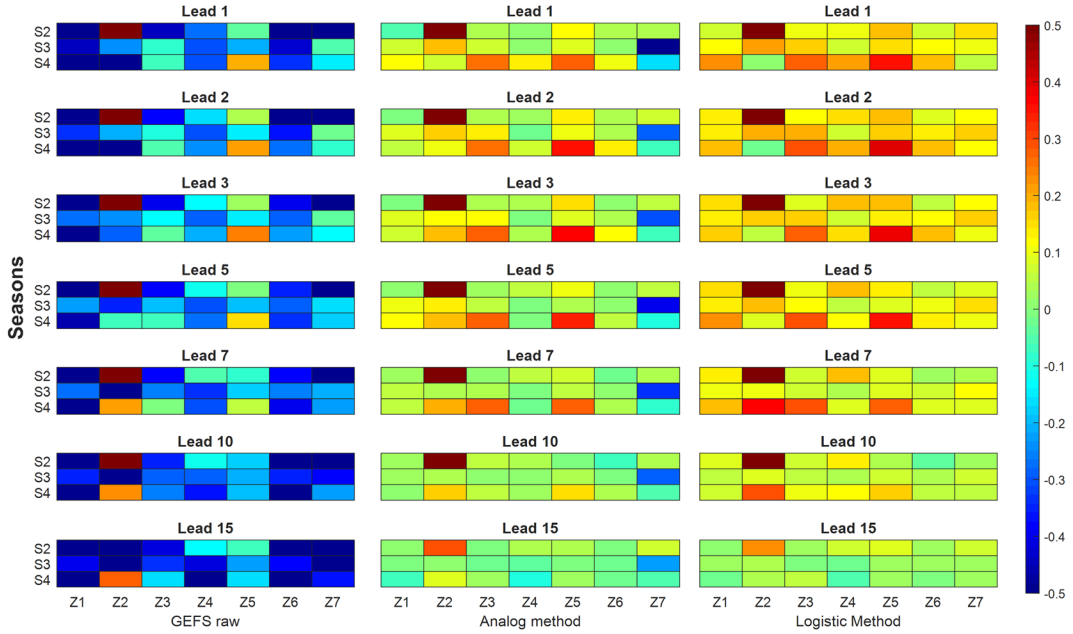
- RMSE assessment of raw GEFS and Analog method for different months and different lead days.



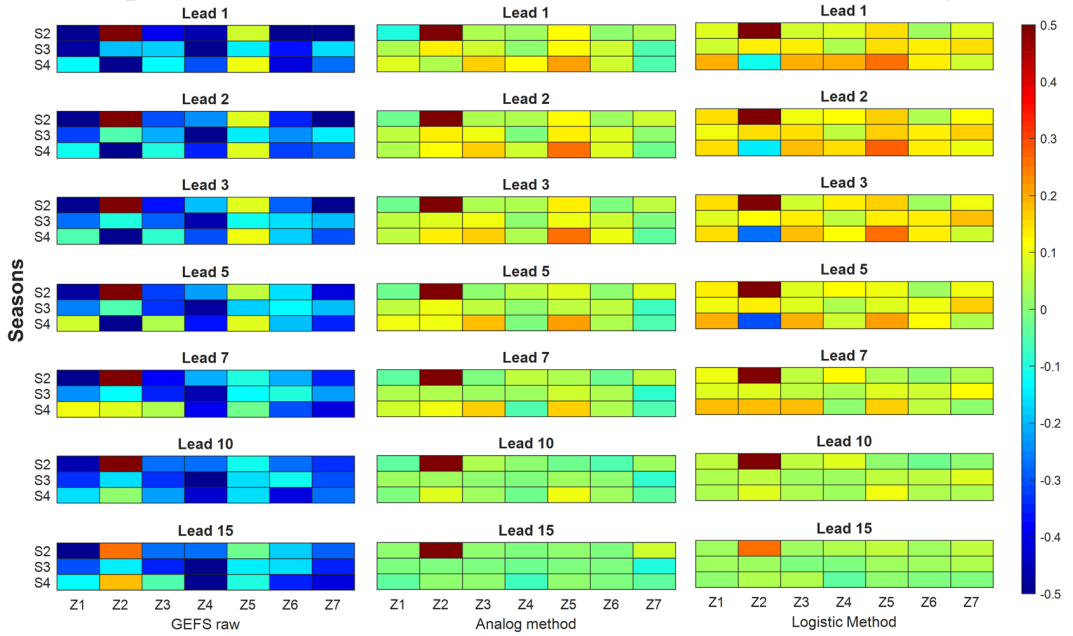
Season and Basin Analysis:

- BSS assessment of raw GEFS, Analog, and Logistic regression method for different seasons and basins.

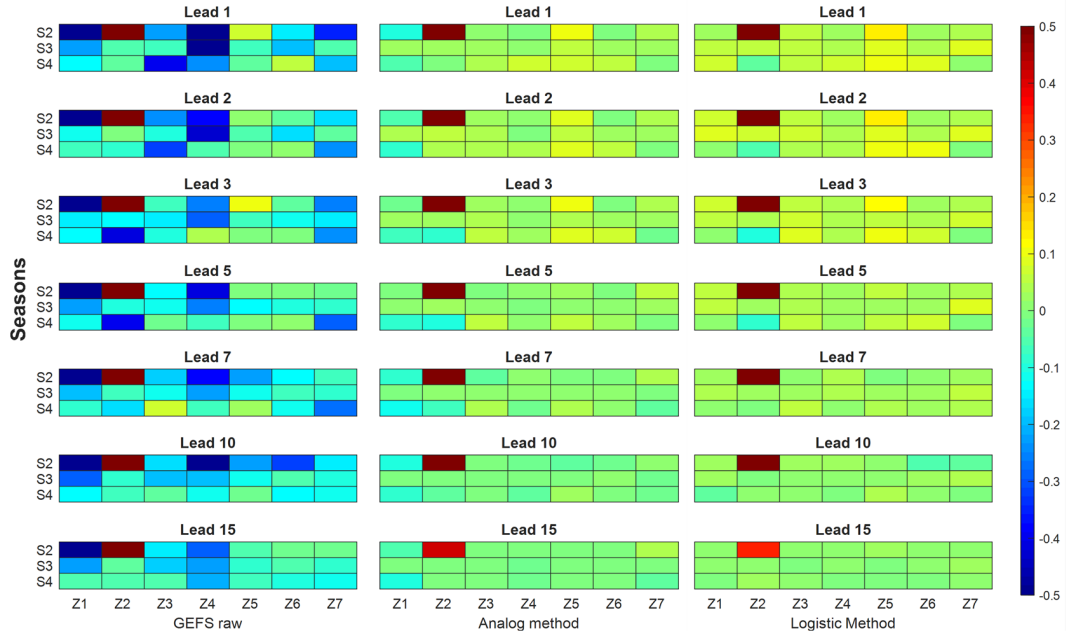
BSS (10th percentile) value over different basins and seasons at 1-day lead time



BSS (50th percentile) value over different basins and seasons at 1-day lead time

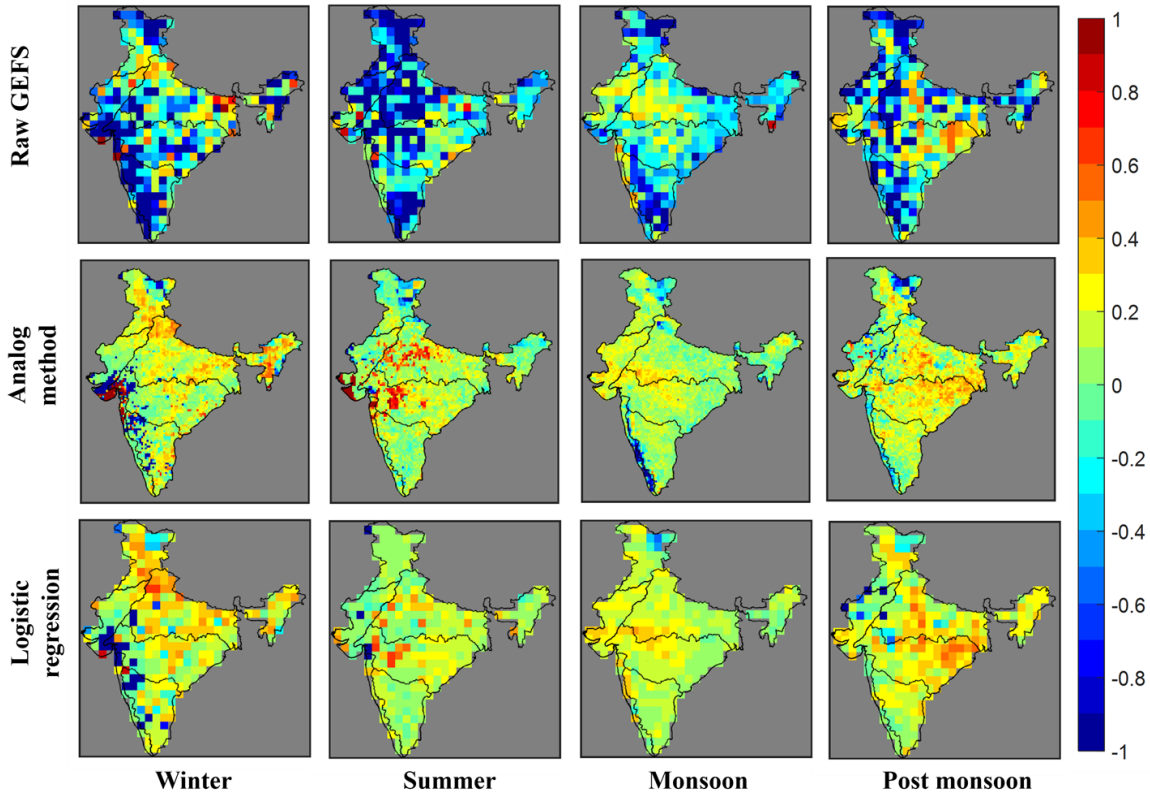


BSS (90th percentile) value over different basins and seasons at 1-day lead time



- BSS (1mm threshold) assessment of raw GEFS, Analog and Logistic regression method over the Indian region at 1-day lead time

BSS (1mm threshold) over the Indian region at 1-day lead time



CONCLUSION

- Overall results showed that both the methods perform poorly during the monsoon seasons compared to other seasons.

- Basin analysis showed that both the methods underperform in the Western Ghats, while the performance is comparable and decent in other parts of India.

- Analysis of precipitation at different terciles showed that both the AN and LR methods underperform at higher terciles compared to the lower ones.

- The comparison of logistic regression and analog methods shows that the LR method outperforms the AN method in almost all the locations and lead times.

FUTURE WORK AND ACKNOWLEDGEMENT

- Comparison of analog and logistic regression methods with other Model Output Statics to assess the performance of these methods over currently available methods.

- Comparison of analog ensembles generated from GEFS with other NWP forecasts.

- Forcing analog ensembles into a Hydrological model for forecasting of Streamflow.

Acknowledgment:

- Indian Meteorological Department for verification data.
- National Oceanic and Atmospheric Administration agency for GEFS NWP data.

AUTHOR INFORMATION

First Presenting Author

Name: Sakila Saminathan

Email: 101814002@smail.iitpkd.ac.in

Affiliation: Department of Civil Engineering, Indian Institute of Technology Palakkad, Kerala, India

Co-Authors

Name: Subhasis Mitra

Email: szm0048@tigermail.auburn.edu

Secondary Email: smitra@iitpkd.ac.in

Affiliation: Department of Civil Engineering, Indian Institute of Technology Palakkad, Kerala, India

Name: Di Tian

Email: tiandi@auburn.edu

Affiliation: Auburn University

Name: Hanoi Medina

Email: hzm0045@tigermail.auburn.edu

Affiliation: Auburn University

ABSTRACT

This study aims to enhance the accuracy and reliability of the Global Ensemble Forecast System's (GEFS) precipitation forecasts over the Indian subcontinent using two post-processing techniques, namely the Analog method (AN) and Logistic Regression method (LR). The postprocessing techniques and GEFS Numerical Weather Prediction Model (NWP) outputs were evaluated against the observed dataset using probabilistic and deterministic evaluation metrics. Results found that both the methods considerably improves short range to medium range (1-15 day) precipitation forecasts over India. Overall results showed that both the methods perform poorly during the monsoon seasons compared to other seasons. Basin analysis showed that both the methods underperform in the Western Ghats, while the performance is comparable and decent in other parts of India. Analysis of precipitation at different terciles showed that both the AN and LR methods underperforms at higher terciles compared to the lower ones. This is because the GEFS model itself was performing poorly in detecting the heavy precipitation events. The comparison of logistic regression and analog methods shows that the LR method outperforms the AN method in almost all the locations and lead times.

REFERENCES

Abhilash, S., Sahai, A. K., Borah, N., Joseph, S., Chattopadhyay, R., Sharmila, S., Rajeevan, M., Mapes, B. E., and Kumar, A. (2015). “Improved Spread–Error Relationship and Probabilistic Prediction from the CFS-Based Grand Ensemble Prediction System.” Journal of Applied Meteorology and Climatology, 54(7), 1569–1578.

Abhilash, S., Sahai, A. K., Pattnaik, S., Goswami, B. N., and Kumar, A. (2014). “Extended range prediction of active-break spells of Indian summer monsoon rainfall using an ensemble prediction system in NCEP Climate Forecast System.” International Journal of Climatology, 34(1), 98–113.

Beria, H., Nanda, T., Singh Bisht, D., and Chatterjee, C. (2017). “Does the GPM mission improve the systematic error component in satellite rainfall estimates over TRMM? An evaluation at a pan-India scale.” Hydrology and Earth System Sciences, 21(12), 6117–6134.

Bhowmik, S. K. R., and Durai, V. R. (2012). “Development of multimodel ensemble based district level medium range rainfall forecast system for Indian region.” Journal of Earth System Science, 121(2), 273–285.

Bisht, D. S., Chatterjee, C., Raghuwanshi, N. S., and Sridhar, V. (2018). “Spatio-temporal trends of rainfall across Indian river basins.” Theoretical and Applied Climatology, 132(1–2), 419–436.

Borah, N., Sahai, A. K., Abhilash, S., Chattopadhyay, R., Joseph, S., Sharmila, S., and Kumar, A. (2015). “An assessment of real-time extended range forecast of 2013 Indian summer monsoon.” International Journal of Climatology, 35(10), 2860–2876.

Dasgupta, S., and De, U. K. (2007). “Binary logistic regression models for short term prediction of premonsoon convective developments over Kolkata (India).” International Journal of Climatology, 27(6), 831–836.

H. M, V. D. D. (1989). “A New Look at Weather Forecasting through Analogues.” Monthly Weather Review, 117, 2230–2247.

Hamill, T. M., Bates, G. T., Whitaker, J. S., Murray, D. R., Fiorino, M., Galarneau, T. J., Zhu, Y., and Lapenta, W. (2013). “NOAA’s Second-Generation Global Medium-Range Ensemble Reforecast Dataset.” Bulletin of the American Meteorological Society, 94(10), 1553–1565.

Hamill, T. M., Scheuerer, M., and Bates, G. T. (2015). “Analog Probabilistic Precipitation Forecasts Using GEFS Reforecasts and Climatology-Calibrated Precipitation Analyses.” Monthly Weather Review, 143(8), 3300–3309.

Hamill, T. M., and Whitaker, J. S. (2006). “Probabilistic Quantitative Precipitation Forecasts Based on Reforecast Analogs: Theory and Application.” Monthly Weather Review, 134(11), 3209–3229.

Hamill, T. M., Whitaker, J. S., and Mullen, S. L. (2006). “Reforecasts: An Important Dataset for Improving Weather Predictions.” Bulletin of the American Meteorological Society, 87(1), 33–46.

Jain, S. K., Mani, P., Jain, S. K., Prakash, P., Singh, V. P., Tullos, D., Kumar, S., Agarwal, S. P., and Dimri, A. P. (2018). “A Brief review of flood forecasting techniques and their applications.” International Journal of River Basin Management, 16(3), 329–344.

Medina, H., and Tian, D. (2020). “Comparison of probabilistic post-processing approaches for improving numerical weather prediction-based daily and weekly reference evapotranspiration forecasts.” Hydrology and Earth System Sciences, 24(2), 1011–1030.

Medina, H., Tian, D., Marin, F. R., and Chirico, G. B. (2019). “Comparing GEFS, ECMWF, and Postprocessing Methods for Ensemble Precipitation Forecasts over Brazil.” Journal of Hydrometeorology, 20(4), 773–790.

Piani, C., Haerter, J. O., and Coppola, E. (2010). “Statistical bias correction for daily precipitation in regional climate models over Europe.” Theoretical and Applied Climatology, 99(1–2), 187–192.

Prasad, K., Dash, S. K., and Mohanty, U. C. (2009). “A logistic regression approach for monthly rainfall forecasts in meteorological subdivisions of India based on DEMETER retrospective forecasts.” International Journal of Climatology.

Rajeevan, M., Pai, D. S., Anil Kumar, R., and Lal, B. (2007). “New statistical models for long-range forecasting of southwest monsoon rainfall over India.” Climate Dynamics, 28(7–8), 813–828.

Shah, R., Sahai, A. K., and Mishra, V. (2017). “Short to sub-seasonal hydrologic forecast to manage water and agricultural resources in India.” Hydrology and Earth System Sciences, 21(2), 707–720.

Tian, D., and Martinez, C. J. (2012). “Comparison of two analog-based downscaling methods for regional reference evapotranspiration forecasts.” Journal of Hydrology, 475, 350–364.

Whitaker, J. S., Wei, X., and Vitart, F. (2006). “Improving Week-2 Forecasts with Multimodel Reforecast Ensembles.” Monthly Weather Review, 134(8), 2279–2284.

Wilks, D. S. (2006). “Comparison of ensemble-MOS methods in the Lorenz ’96 setting.” Meteorological Applications, 13(03), 243.

Wilks, D. S. (2009). “Extending logistic regression to provide full-probability-distribution MOS forecasts.” Meteorological Applications, 16(3), 361–368.

Wilks, D. S., and Hamill, T. M. (2007). “Comparison of Ensemble-MOS Methods Using GFS Reforecasts.” Monthly Weather Review, 135(6), 2379–2390.