MAUSAM, **74**, 2 (April, 2023), 421-428

MAUSAM

DOI : https://doi.org/10.54302/mausam.v74i2.5997
Homepage: https://mausamjournal.imd.gov.in/index.php/MAUSAM

UDC No. 551.553.21 : 551.577 : 551.509.313.4

A machine learning approach for probabilistic multi-model ensemble predictions of Indian summer monsoon rainfall

NACHIKETA ACHARYA*, ** and KYLE JOSEPH CHEN HALL*, **

**Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado Boulder, Boulder, CO, United States*

***National Oceanic and Atmospheric Administration (NOAA) Physical Sciences Laboratory, Boulder, CO, United States*

e mail : dr.nachiketaacharya@gmail.com

सार – भारतीय ग्रीष्मकालीन मॉनसून वर्षा (ISMR) से जुड़ी अनिश्चितता के कारण, एकल निर्धारणात्मक पूर्वानुमान की तुलना में संभाविता ऋतुनिष्ठ पूर्वानुमान उपयोगकर्ता के लिए अधिक उपयोगी हैं जो ISMR की अंतर्निहित अनिश्चितता को व्यक्त कर सकते हैं। हालाँकि इस तरह के संभाविता ऋतुनिष्ठ पूर्वानुमान सामान्य परिसंचरण मॉडल (GCM) आउटपुट से उत्पादित किए जा सकते हैं, और आम तौर पर एकल मॉडल त्रुटि के सभी स्रोतों का प्रतिनिधित्व नहीं करता है। संभाविता बहु मॉडल एंसेंबल (PMME) व्यक्तिगत GCM द्वारा संभाविता पूर्वानुमान के कौशल में सुधार के लिए एक स्वीकृत तरीका है। PMME को दो दृष्टिकोणों में से एक के साथ बनाया जा सकता है: ऋतुनिष्ठ कुल वर्षा की तीन श्रेणियों के संबंध में गैर-पैरामीट्रिक, या पैरामीट्रिक-जलवायु आधार अवधि द्वारा परिभाषित सामान्य से कम, निकट और सामान्य से अधिक। दोनों विधियों की अपनी सीमाएँ हैं। गैर-पैरामीट्रिक PMME एक छोटे एंसेंबल आकार का उपयोग करता है जिसके परिणामस्वरूप अतिविश्वसनीय पूर्वानुमान होते हैं, और पैरामीट्रिक PMME गलत धारणा बनाते हैं कि कुल वर्षा गॉसियन वितरण का अनुसरण करती है। इन समस्याओं से बचने के लिएहम ISMR पूर्वानुमान हेतु PMME के निर्माण के लिए एक नई मशीन लर्निंग (ML) दृष्टिकोण एक्सट्रीम लर्निंग मशीन (ELM) के उपयोग को प्रस्तावित करते हैं। ELM सिंगल-हिडन-लेयर फीड-फॉरवर्ड न्यूरल नेटवर्क का अत्याधुनिक सामान्यीकृत रूप है। हालाँकि, पारंपरिक ELM नेटवर्क केवल एक निर्धारणात्मक परिणाम उत्पन्न करता है, इसलिए हम ELM के संशोधित संस्करण का उपयोग करते हैं जिसे प्रोबेबिलिस्टिक आउटपुट एक्सट्रीम लर्निंग मशीन (PO-ELM) कहा जाता है। PO-ELM संभाविता पूर्वानुमान करने के लिए सिग्मॉइड एडिटिव न्यूरॉन्स और थोड़ा अलग रैखिक प्रोग्रामिंग का उपयोग करता है। इस तरह के PO-ELM आधारित PMME के प्रदर्शन का मूल्यांकन सामान्यीकृत रिसीवर ऑपरेटिंग विशेषता स्कोर और विश्वसनीयता आरेखों के संदर्भ में 1982 से 2018 तक की 37 वर्षों की अवधि के दौरान लीव-थ्री-ईयर-आउट क्रॉस-वैलिडेशन स्कीम के बाद किया जाता है। यह बताया गया है कि ML पर आधारित PMME के लिए हमारी नई रणनीति भारत के बड़े क्षेत्रों में कुशल MME पूर्वानुमान तैयार करने में सक्षम है।

ABSTRACT. Due to the uncertainty associated with Indian summer monsoon rainfall (ISMR), probabilistic seasonal forecasts which can convey the inherent uncertainty of ISMR are more useful to the user community than a single deterministic forecast. While such probabilistic seasonal forecasts can be produced from general circulation model (GCM) output, one single model generally does not represent all sources of error. The probabilistic multi model ensemble (PMME) is a well-accepted way to improve on the skill of probabilistic forecasts by individual GCMs. PMME can be constructed with one of two approaches: non-parametric, or parametric with respect to the occurrence of three categories of seasonal total rainfall-below, near and above normal as defined by the climatological base period. However, both the methods have their limitations. Non-parametric PMME use a smaller ensemble size which results in overconfident forecasts and parametric PMME make the inaccurate assumption that total rainfall follows a Gaussian distribution. To avoid these problems, we propose the use of the Extreme Learning Machine (ELM), a novel machine learning (ML) approach, to construct PMME for ISMR forecasting. ELM is a state-of-the-art generalized form of single-hidden-layer feed-forward neural network. However, since the traditional ELM network only produces a deterministic outcome, we use a modified version of ELM called Probabilistic Output Extreme Learning Machine (PO-ELM). PO-ELM uses sigmoid additive neurons and slightly different linear programming to make probabilistic predictions. The performance of such PO-ELM based PMME is assessed rigorously in terms of Generalized Receiver Operating Characteristic scores and reliability diagrams over a 37 years period spanning from 1982 to 2018 following a leave-three-year-out cross-validation scheme. It is demonstrated that our new strategy for PMME based on ML is capable of producing skillful MME forecasts over large regions of India.

Key words – Indian Monsoon, Probabilistic prediction, Machine learning, Multimodal ensemble.

1. Introduction

Predictions communicating the intrinsic uncertainty of Indian summer monsoon rainfall (ISMR) become more useful to the user community, especially those stakeholders involved in agricultural/hydrological planning and climate risk management, because of ISMR's characteristic intricacy. Probabilistic forecasts which can convey the prediction uncertainty can be considered a better way to disseminate seasonal forecast information, compared to single deterministic forecasts. A plethora of studies exist in the literature which focus on the deterministic approach to the seasonal prediction of ISMR. However, only a few studies have described probabilistic prediction systems (Kulkarni *et al*., 2012; Acharya *et al*., 2013; Acharya *et al*., 2014a; Acharya, 2018). These studies are mostly based on probabilistic multi-model ensemble (PMME) prediction using outputs of general circulation model (GCM). As one single model generally does not represent all sources of error, the PMME is a well-accepted way to improve the skill of probabilistic forecasts by individual GCMs.

PMME forecasts can be made with one of two approaches: non-parametric, or parametric with respect to the occurrence of three equi-probable, mutually exclusive and collectively exhaustive categories *viz*., below-normal (BN), near-normal (NN) and above-normal (AN) of seasonal mean rainfall as defined by the climatological base period. There are some basic differences between non-parametric and parametric methods for making PMME. In the non-parametric approach, weights will be assigned after the estimation of probabilities from individual GCM without assuming the form of forecast distribution (Acharya *et al*., 2013). Whereas, in a parametric method, weights are assigned to the deterministic prediction obtained from GCM and then converted to probabilistic space using Gaussian predictive probability density function (Acharya *et al*., 2014a). However, both the methods have their limitations. For example, Non-parametric PMME uses a smaller ensemble size which results in overconfident forecasts and parametric PMME makes the inaccurate assumption that total rainfall follows a Gaussian distribution.

To avoid the problems with the non-parametric and parametric approaches, in this study, we propose using a novel Machine Learning (ML) method based on Artificial Neural Network (ANN) to construct PMME for seasonal predictions of ISMR for the first time. Based on our previous work constructing deterministic MME using the Extreme Learning Machine (ELM) (Acharya *et al*., 2014b), a generalized form of the single-hidden-layer feed-forward ANN, we explore ELM's potential for making PMME. However, as the traditional ELM network can only produce a deterministic outcome, we implemented a modified version of ELM called Probabilistic Output Extreme Learning Machine (PO-ELM) and used it to construct PMME-based forecasts of ISMR with the state-of-the-art coupled GCMs of the North American Multi-Model Ensemble (NMME) (Kirtman, *et al*., 2014). This work basically reports an extension of the work of Acharya *et al*. (2013); Acharya *et al*. (2014a and 2014b) and the ultimate interest of the present work is to develop a skillful probabilistic seasonal forecast of ISMR which is useful to multiple sectors of society, including agriculture, hydrology and health.

Below, Section 2 provides a short description of the GCMs and reference datasets used. Section 3 outlines the methodology and implementation of the PO-ELM approach for making PMME. Section 4 presents the evaluation of the performance of the proposed method and section 5 concludes with a future scope of this research.

2. Datasets

2.1. *NMME datasets*

This study uses five GCM's May-initialized, lead-1 seasonal mean rainfall hindcasts for the June-July-August-September (JJAS) season spanning over the 1982-2018 time period. These five global ocean-atmosphere coupled models are members of the North American Multi-Model Ensemble (NMME) project (Kirtman *et al*., 2014). The five models include one from NOAA's Centers for Environmental Prediction (NCEP-CFSv2), two from the Canadian Meteorological Center (CanSIPS-IC3 GEM5-NEMO and CanCM4i), one from the Geophysical Fluid Dynamics Laboratory (GFDL-CM2p5-FLOR-B01) and one from the National Aeronautics and Space Administration (NASA-GEOSS2S). These NMME monthly hindcast datasets are available at a common 1° resolution grid at http://iridl.ldeo.columbia.edu/

SOURCES/.Models/.NMME. The GCMs have different numbers of ensemble members representing different initialization methods, which were averaged to generate an ensemble mean to represent the forecast signal. More details about these GCMs can be found in Acharya *et al*. (2021). The capacity of these NMME models to predict ISMR is well documented in several studies (Pillai *et al*., 2018; Singh *et al*., 2019).

2.2. *Observational reference data*

The 1° gridded daily rainfall dataset from India Meteorological Department (IMD) is used as an observational reference (Rajeevan *et al*., 2006). This daily data is further transformed into seasonal mean rainfall for JJAS for the same time period of the GCMs' hindcast (1982 to 2018).

3. Methodology and implementation

3.1. *Overview of ELM*

The ANN, an ML technique based on the human nervous system, is the most popular artificial intelligence-based method for prediction, pattern recognition and pattern classification. The single-hidden-layer feedforward network (SLFN), one of the most popular types of feed forward ANNs, has been extensively studied from both theoretical and practical perspectives for its learning capacity and fault-tolerance. It has been used extensively to forecast convective hazards since the mid-1990s, however, after the pioneering work by Acharya *et al*. (2014b), there is strong interest in exploring the potential of ML based MME to improve seasonal forecasts.

The efficacy of SLFN-based methods is highly dependent on the appropriate tuning of their adjustable hyperparameters, *e.g*., transfer function, learning rate and the number of nodes in each layer. There are also several disadvantages of traditional SLFN-based methods, including long computation time, over-fitting and vanishing gradient. The Extreme Learning Machine (ELM), proposed by Huang *et al*. (2006), appears to solve these problems. ELM can unquestionably learn much faster than conventional learning algorithms like the back- propagation algorithm and the Levenberg-Marquardt algorithm that leverage gradient descent-based methods to optimize the weights in the neural network.

In ELM, the parameters of hidden nodes, like their input weights and biases, are randomly assigned and need not be tuned. Alternately, the output weights are determined analytically using the Moore-Penrose (MP) generalized inverse. This simplified approach makes ELM thousands of times faster than traditional learning algorithms. Moreover, ELM is able to produce very good generalization performance with less human intervention. The detailed mathematical foundations of ELM can be found in Huang *et al.* (2006). Acharya *et al*. (2014b) proposed the use of the ELM for generating the deterministic MME based prediction.

3.2. *Description of PO-ELM*

Since the successful introduction of the basic ELM, a number of further extensions and improvements to the structure of ELM have been proposed by researchers. However, probabilistic prediction using ELM has not yet been explored much since the traditional ELM only produces unbounded continuous deterministic output, which is not well suited for probabilistic forecasting. In a very recent study by Wong *et al.* (2020), the original ELM model was modified to produce probabilistic outputs. The PO-ELM (Wong *et al*., 2020) modifies the original ELM model such that it optimizes a sigmoid objective function and produces bounded continuous output on the interval (0, 1). When trained with a binary target vector, this sigmoid output can be interpreted as a binary probability-the probability that a given input sample will be a member of the class of interest (represented by 1) or not (represented by 0).

Where, however, we have also made a further modification to the PO-ELM model, in order to produce relative probabilities for the BN, NN and AN classes, rather than binary probabilities for each. A traditional PO-ELM, trained on a one-hot encoded multi-class target vector, is akin to three PO-ELM models with the same weights, trained on each of the independent column vectors representing each class. Each of those models would produce binary probabilities for the class they were trained on. This is not useful in tercile probability forecasting, since the three probabilities have no meaning relative to one another - they could all be 0.99, a forecast of 'yes' for all three tercile categories, a difficult to interpret the result. We address this by applying normalization to the PO-ELM network's binary probabilities. More mathematical details of PO-ELM can be found in Wong *et al*. (2020). Fig. 1 shows the schematic structure of PO-ELM.

3.3. *Implementation procedure of PO-ELM for making PMME*

In this study, PO-ELM was trained to make PMME on the hindcast from five GCMs. The whole procedure consisted of the following sequential steps:

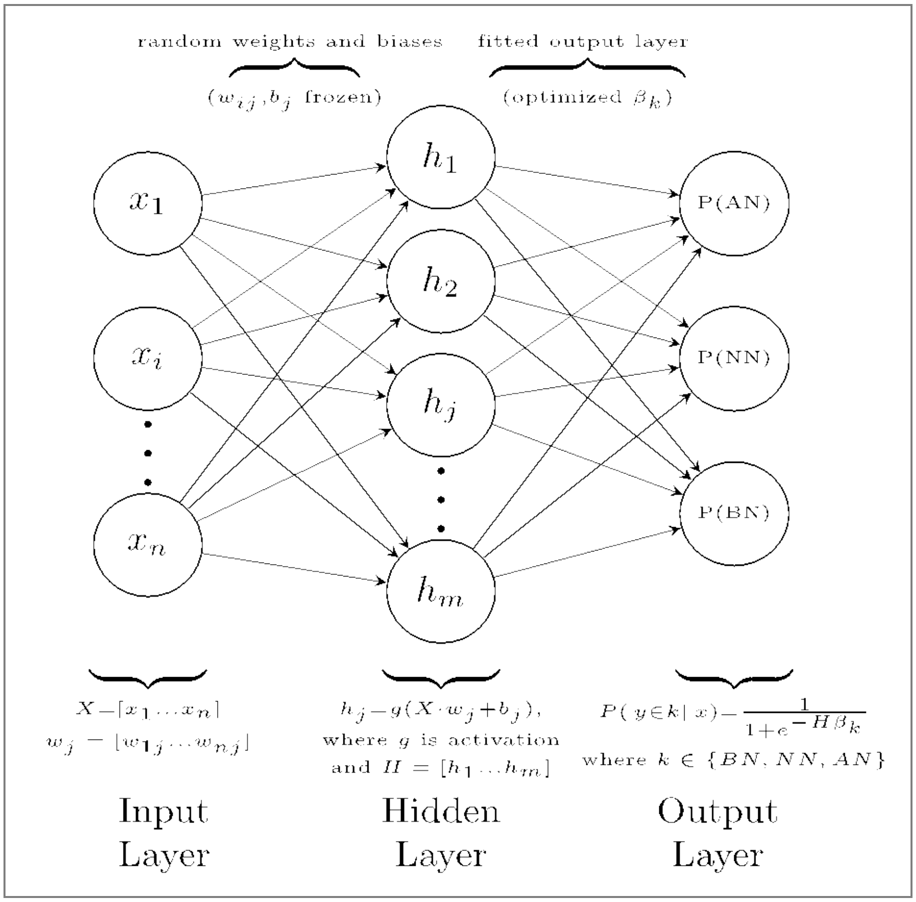

Fig. 1. Schematic representation of the structure of Probabilistic Output Extreme Learning Machine (PO-ELM)

3.3.1. *Scaling the input neurons*

First the rainfall outputs from five GCMs were used as input neurons and were scaled to the range of [-1, 1] *via* the "Min-Max" scaling method (Acharya *et al.*, 2014b).

3.3.2. *One-hot encoding the target data*

The observed rainfall data (IMD) was treated as target data and was "one-hot encoded" according to tercile boundaries defined by the 33rd and 66th percentiles of the observed data. This creates three separate target vectors representing BN, NN and AN respectively. With one-hot encoding, observed rainfall is converted into a binary vector (1 or 0) for each tercile category, with one representing a sample's membership in a tercile category and zero representing a lack of membership in that category.

3.3.3. *Apply PO-ELM algorithm in training and testing mode*

ML models are data-hungry - meaning that a large sample size is needed for ML models to learn relationships and maintain good predictive accuracy. However, due to the short length (1982-2018; only 37 years) of the dataset used here, it cannot be divided independently for training and testing the network. So, we used leave-three-out cross validation where three years among the total dataset was reserved for the ''test'' dataset and residual data (in our case 34 samples) was used as the ''training'' dataset. After training, or fitting the network with PO-ELM, a final weight matrix was obtained which was further applied to the independent inputs in the ''test'' phase. Then the final outcomes were compared with the original observation.

3.3.4. *Hyperparameter tuning and final architecture*

To find an optimal network, we performed informal hyper-parameter tuning using cross-validation, loosely targeting high values and broad geographical spread of verification metrics. We found that the most skillful architecture, presented in section 4, consisted of 5 hidden nodes with linear activation functions. As the weights and bias are randomly chosen in the first part of the network, to avoid the over or under-fitting, we ran the network 30 times and finally averaged the final probabilities.

3.3.5. *Spatial Smoothing*

As this PMME approach was implemented at each grid point individually, the resultant probabilities tended to be noisy spatially. Therefore, the final forecast probabilities were smoothed spatially using local kernel-function smoothing. In particular, the PMME prediction were spatially smoothed using 3 × 3 kernel smoothers with weights that followed a bivariate Gaussian function with standard deviation = 1° and were normalized to sum to 1. Coastal area weights were normalized according to the number of land grid points included in the smoothing.

All above discussed steps required for the generation of a PO-ELM-based PMME forecast system were done using XCast, a high-performance Python data science toolkit for climate forecasting, designed by the authors (Hall and Acharya, 2022).

4. Results and discussion

We demonstrate the skill of PO-ELM-based PMME tercile probabilistic forecasts of Indian Summer Monsoon Rainfall using Generalized Receiver Operating Characteristic scores (GROCS) and reliability diagram as skill measures. Both of the skill metrics were estimated within leave-three-out cross-validation for 1982-2018. The detailed discussion of skill assessment is presented below:

4.1. *Generalized receiver operating characteristic scores*

In probabilistic forecasts, ROC curves describe the degree to which forecasts are able to correctly discriminate between classes, as well as their ability to distinguish one categorical outcome from another (Mason and Graham, 1999). The ROC curve is basically a signal detection curve obtained by plotting a graph of hit rate (on the vertical axis) against corresponding false alarm rate (on the horizontal axis) over a range of different thresholds to assess a probabilistic forecasts system. A hit implies the occurrence of an event of interest, such as above normal precipitation, while a false alarm implies the nonoccurrence of such an event. An ROC curve demonstrating some level of skill generally follows a curve beginning at the bottom left corner of a square and intercepting both the top right corner and a point in the top left quadrant of the figure. A diagonal line indicates no skill, *i.e.*, the hit rate and false-alarm rate are equal. The ROC curve will lie above the 45° line from the origin if the forecast system is skillful (when the hit rate exceeds

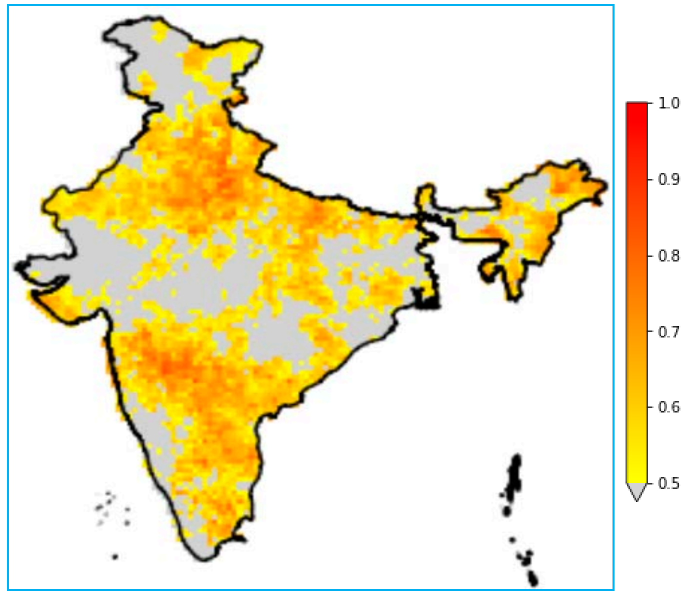

Fig. 2. Generalized Relative Operating Characteristic Score (GROCS) of the PO-ELM based PMME for June-July-August-September during 37 years hindcast period. Areas with GROCS values greater than 0.5 (exhibiting some skill) have been shaded

the false-alarm rate). The area beneath the ROC curve which has become known as the ROC score (ROCS) summarizes the performance of the forecast. ROCS above 0.5 reflect positive discrimination skill while 1.0 representing the maximum possible score. If the ROCS is less than 0.5 (*i.e.*, the same as a no-skill forecast), then the model is less skillful than a random or constant forecast. Generally, ROCS is calculated for each of the tercile categories (BN or NN or AN) individually. However, the generalized ROC score (GROCS) is generalized to encompass all forecast categories collectively, rather than being specific to a single category.

The spatial patterns of GROCS for PO-ELM based PMME have been shown in Fig. 2. It is interesting to notice that the large portions of North, Northeast, Central and South India exhibit more skill (GROCS>0.5) than climatological forecasts (*i.e.*, 0.33 for each tercile category). Especially high GROCS are observable over the Deccan Plateau in South India, the Northern Plains and Western Himalayas in Northern India and along the Eastern Himalayas and North Eastern Range near Purvachal and the Assam Valley. Interestingly, the spatial pattern of high GROCS loosely appears related to the spatial distribution of climatological rainfall during the ISMR season. Therefore, in view of the GROCS analysis, it can be interpreted that for the majority of the grid points over the country, the probabilistic predictions produced by PO-ELM based PMME successfully discriminated between events of occurrences and non-occurrences.

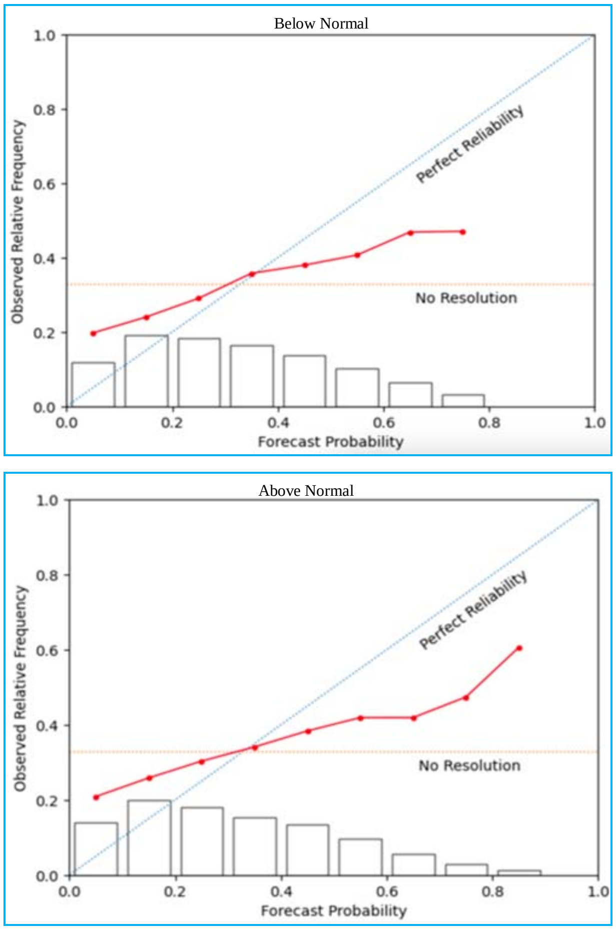

Fig. 3. Reliability diagram for below and above normal categories of the PO-ELM based PMME for June-July-August-September during 37 years hindcast period

4.2. *Reliability diagrams*

One limitation of GROCS is that it only shows the degree of correct probabilistic forecast discrimination. However, for a skillful probability forecast other factors such as the reliability (degree of correspondence between issued forecast probabilities and relative observed frequencies), resolution (ability of the forecast system to assign probabilities different from the climatological probability) and sharpness (tendency of probability forecasts to approach 0 and 1) is equally important. Reliability diagram (Murphy and Winkler, 1977) is widely used to summarized all the above discussed factors to assess the forecast probabilities for each of the forecast categories individually. The primary interest in the diagram where the *x*-axis shows the forecast probability intervals and the *y*-axis shows the corresponding observed relative frequencies of occurrence, is to show the observed relative frequency of occurrence that is associated with the issuance of each of a set of specific probability intervals (bins), shown individually for each of the tercile categories. An inset plot is often shown within the reliability plot, representing the frequency of issuance of each probability interval for each of the forecast categories. This plot tells how strongly and frequently the issued forecast probabilities depart from the climatological probabilities (*i.e*., 0.33 for each tercile category) as a measure of sharpness. A reliability diagram requires a large sample of data to show a wealth of detailed information about the forecasts and their correspondence with the observations for each of a set of issued probability intervals. In the ideal situation a reliability diagram should plotted for each individual grid point, but when considering a smaller sample size (in our case it is only 37 years of hindcast), that becomes impossible, hence it requires a large sample of data by pooling all hindcast years and all grid points. Additionally, since users are unlikely to rely on a forecast for a region exhibiting zero skill and because these diagrams are intended to further enhance a user's understanding of the forecast that they intend to rely on, here we choose to limit the calculation of the reliability to data from the geographical subregions exhibiting GROCS greater than or equal to 0.5. Including grid points whose forecasts display no skill in the calculations would unfairly spoil the reliability for the skillful forecasts.

The reliability diagram is plotted in Fig. 3 for BN and AN forecast categories individually by pooling all forecasts for 37 years and all grid points where GROCS>0.5 over Indian land mass.

Note that only bins with more than 1% of the total number of forecasts in each category are plotted in the diagrams. The vertical boundary between the two represents a line of “no skill” while the diagonal boundary between the two represents climatological probability and the horizontal red line in each diagram represents a line of “no resolution”. The 45-degree line bisecting the angles between the “no skill” and “climatology” boundaries represents a perfectly calibrated forecast, in which case one could expect the forecasted percentage probability exactly represent the portion of the time that forecast was accurate (*i.e*., a 100% probability forecast always comes true 100% of the time, a 90% probability forecast comes true 90% of the time, etc.). Fig. 3 shows that both BN and AN category, the red lines (representing the forecast probability versus observed frequency) are close to the diagonal line (perfect forecast), indicating high reliability and well calibrated forecasts across all probabilities. The resolution has also improved as the angle of red lines with the horizontal (no resolution) is higher. However, the NN category closely mirrors the level of calibration of climatology (figure not shown). Both BN and AN also observe a much more uniformly distributed set of forecast probabilities, whereas most of the NN probabilities are centered around 0.3 and do not exceed 0.7. In conclusion, PO-ELM based PMME provides high reliable forecasts with good resolution and better strategy compared to climatological probabilities for BN and AN category.

5. Concluding remarks and future scopes

A single deterministic rainfall forecast is likely to communicate over-confidence for predicting seasonal Indian summer monsoon rainfall (ISMR) which can undermine the trust between stakeholders and forecasters. Probabilistic seasonal forecasts can provide an effective way of conveying inherent uncertainty within the prediction which are useful to multiple sectors of society, including agriculture, hydrology and health.

Probabilistic seasonal forecasts can be produced from outputs of general circulation models (GCMs), however, one single model generally does not represent all sources of error. The probabilistic multi model ensemble (PMME) is a well-accepted way to improve on the skill of probabilistic forecasts by individual GCMs. While PMME can be constructed by non-parametric or parametric methods, both the methods have their limitations. To overcome the limitations, we have proposed the use of an innovative method based on Machine Learning (ML), *viz*., Probabilistic Output Extreme Learning Machine (PO-ELM), a state-of-the-art generalized form of single-hidden-layer feed-forward neural network, to construct PMME for the seasonal predictions of ISMR using the outputs of the GCMs from the NMME project. The performance of such PO-ELM based PMME was assessed rigorously in terms of GROCS and reliability diagram over a 37 years period spanning from 1982 to 2018 under leave-three-year-out cross-validation. In view of the GROCS analysis, we have found that PO-ELM based PMME successfully discriminated between events of occurrences and non-occurrences for the majority of grid points over the country. Moreover, PO-ELM based PMME provides highly reliable forecasts with good resolution and better sharpness than climatological probabilities, for the BN and AN category. Hence, it is demonstrated that our new strategy for PMME based on ML is capable of producing skillful MME forecasts over large regions of India.

Though this paper concentrated only on the traditional tercile probability events, in the future, we will extend this work to explore further structural modifications to the PO-ELM network to develop a more flexible forecast (probability of exceedance) format that allows users to glean information from those part of forecast distribution what matters most to them such as the probability of extremely dry/wet conditions.

Acknowledgements

The authors gratefully acknowledge India Meteorological Department for providing the gridded rain data and the IRI Data Library for maintaining and hosting NMME datasets.

Disclaimer : The contents and views expressed in this study are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

References

Acharya, N., 2018, "Calibration of Probabilistic Ensemble Forecasts for Indian Summer Monsoon Rainfall: A Non-gaussian Approach", Proceedings of the 8th International Workshop on Climate Informatics : CI 2018, NCAR Technical Note NCAR/TN-550+PROC, p151. http://dx.doi.org/10.5065/D6BZ64XQ.

Acharya, N., Chattopadhayay, S., Mohanty, U. C. and Sahoo, L. N., 2014a, "Prediction of Indian Summer Monsoon Rainfall: a weighted multi-model ensemble to enhance probabilistic forecast skills", *Meteorological Application*, **21**, 3, 724-732.

Acharya, N., Ehsan, M. A., Admasu, A., Teshome, A. and Hall, K. J. C., 2021, "On the Next Generation (NextGen) Seasonal Prediction System to enhance Climate Services over Ethiopia", *Climate Services*, **24**, 100272. https://doi.org/10.1016/j.cliser. 2021.100272.

Acharya, N., Mohanty, U. C. and Sahoo, L. N., 2013, "Probabilistic multi-model ensemble prediction of Indian summer monsoon rainfall using General Circulation Models : A non-parametric approach", *Comptes Rendus Geoscience*, **345**, 3, 724-732.

Acharya, N., Srivastava, N. A., Panigrahi, B. K. and Mohanty, U. C., 2014b, "Development of an artificial neural network based multi-model ensemble to estimate the northeast monsoon rainfall over south peninsular India: an application of extreme learning machine", *Climate Dynamics*, **43**, 5, 1303-1310.

Hall, K. and Acharya, N., 2022, "XCast : A python climate forecasting toolkit", *Frontiers in Climate*, 4. https://doi.org/10.3389/ fclim.2022.953262.

Huang, G. B., Qin-Yu, Z. and Chee-Kheong, S., 2006, "Extreme learning machine: theory and applications", *Neurocomputing* **70**, 1-3, 489-501.

Kirtman, B. P., D. Min, J. M. Infanti, J. L. Kinter, D. A. Paolino, Q. Zhang, H. van den Dool, S. Saha, M. P. Mendez, E. Becker, P. Peng, P. Tripp, J. Huang, D. G. DeWitt, M. K. Tippett, A. G. Barnston, S. Li, A. Rosati, S.D. Schubert, M. Rienecker, M. Suarez, Z. E. Li, J. Marshak, Y. Lim, J. Tribbia, K. Pegion, W. J. Merryfield, B. Denis and E. F. Wood, 2014, "The North American Multimodel Ensemble : Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction", *Bull. Amer. Meteor. Soc*., **95**, 585-601.

Kulkarni, M. A., Acharya, N., Kar, S. C., Mohanty, U. C., Tippett, M. K., Robertson, A. W., Luo J. J. and Yamagata, T., 2012, "Probabilistic prediction of Indian summer monsoon rainfall using global climate models", *Theor. Appl. Climatol*., **107**, 441-450.

Mason, S. J. and Graham, N. E., 1999, "Conditional Probabilities, Relative Operating Characteristics and Relative Operating Levels", *Wea. Forecast*., **14**, 713-725.

Murphy, A. H. and Winkler, R. L., 1977, "Reliability of subjective probability forecasts of precipitation and temperature", *Journal of Royal Statistical. Society*, **26C**, 41-47.

Pillai, P. A., Rao, S. A., Ramu, D. A., Pradhan, M. and George, G., 2018, "Seasonal prediction skill of Indian summer monsoon rainfall in NMME models and monsoon mission CFSv2", *Int. J. Climatol.*, 38:e847-e861.

Rajeevan, M., Bhate, J., Kale, J. and Lal, B., 2006, "High resolution daily gridded rainfall data for the Indian region : Analysis of break and active monsoon spells", *Curr. Sci. India*, **91**, 3, 296-306.

Singh, B., Cash, B. and Kinter, J. L., 2019, "Indian summer monsoon variability forecasts in the North American multimodel ensemble", *Climate Dynamics*, **53**, 7321. https://doi.org/10.1007/s00382-018-4203-6.

Wong, S. Y., Yap, K. S. and Li, X., 2020, "A new probabilistic output constrained optimization extreme learning machine", *IEEE Access*, **8**, 28, 934-28, 946.