E-ISSN: 2618-0618
P-ISSN: 2618-060X

NAAS Rating (2026): 5.20
www.agronomyjournals.com
2026; 9(3): 429-440
Received: 18-01-2026
Accepted: 23-02-2026

Neha Chand
Department of Agrometeorology, College of Agriculture, GBPUA&T, Pantnagar, Uttarakhand, India

Anurag Tripathi
Department of Agrometeorology, College of Agriculture, GBPUA&T, Pantnagar, Uttarakhand, India

Anurag Satpathi
Sher-e-Kashmir University of Agricultural Sciences and Technology, Shalimar, Srinagar, Jammu & Kashmir, India

Shailesh Yadav
Department of Mechanical Engineering, College of Technology, GBPUA&T, Pantnagar, Uttarakhand, India

Ravi Kiran
Department of Agrometeorology, College of Agriculture, GBPUA&T, Pantnagar, Uttarakhand, India

Rajeev Ranjan
ICAR-Central Marine Fisheries Research Institute, Kochi, Kerala, India

Anil Shukla
Department of Agronomy, GBPUA&T, Pantnagar, Uttarakhand, India

A.P. Singh
Department of Soil Science, GBPUA&T, Pantnagar, Uttarakhand, India

Corresponding Author:
Neha Chand
Department of Agrometeorology, College of Agriculture, GBPUA&T, Pantnagar, Uttarakhand, India

District-level rice yield forecasting in Uttarakhand, India: A comparative study of regularization, kernel-based, tree-based and neural network models

Neha Chand, Anurag Tripathi, Anurag Satpathi, Shailesh Yadav, Ravi Kiran, Rajeev Ranjan, Anil Shukla and AP Singh

DOI: https://www.doi.org/10.33545/2618060X.2026.v9.i3f.5279

Abstract
Accurate pre-harvest forecasting of rice yields is critical for food security and climate adaptation in Uttarakhand, a topographically complex Himalayan state. This study presents a rigorous comparative assessment of five machine learning algorithms-Random Forest (RF), Elastic Net (ELNET), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Artificial Neural Network (ANN) for kharif rice yield prediction across 13 districts. Models were calibrated and validated using a 25-year dataset (1998-2023) of district-level yields and monthly agrometeorological variables from the NASA Power website employing a chronological split to simulate real-world forecasting conditions. Performance was evaluated using R^2, RMSE, nRMSE, and mean bias error (MBE). Random Forest emerged as the most robust and generalizable model, consistently delivering superior validation performance across districts ($R^2 = 0.55$-0.89) with low error and minimal bias. XGBoost exhibited near-perfect calibration but was prone to overfitting, particularly in data-sparse high-altitude regions. SVM demonstrated competitive, district-specific utility, while ELNET provided stable, interpretable predictions. ANN performance was inconsistent, limited by the modest training sample size. The findings establish that no single model is universally optimal; rather, a district-specific selection strategy or hybrid ensemble approach is recommended for maximizing predictive accuracy in heterogeneous *agroclimatic* zones.

Keywords: Rice yield forecasting; Machine learning; Random Forest; XGBoost; Uttarakhand

1. Introduction

Global food production needs to double by 2050 in order to meet the demand of the rapidly growing population (Tilman *et al*., 2011; FAO, 2009) [12, 41]. On the other hand, the current yield growth rates for the major cereals grown across the globe are not high enough to meet this target (Ray *et al*., 2013) [34]. Environmental changes, particularly global warming and climate variability, are key concerns that have a negative impact on agriculture (Rosenzweig *et al*., 2014) [36]. This may result in a decline in crop production (Lobell & Gourdji, 2012) [25], making the world more food insecure. As the global population is projected to reach 9 billion people by 2050, governments all over the world need to be well prepared to deal with supply shocks of major cereals.

The Food and Agriculture Organization (FAO) reports that the demand for and consumption of grains have grown significantly relative to production in developing nations such as India. From 1964 to 2030, there will have been a sustained increase in demand for rice, wheat, and other coarse grains (FAO, 2003) [11]. Cereal imports in developing countries increased significantly to meet rising demand, growing from 39 million tons annually in 1970 to 130 million tons annually by 1997-1999. These developing nations are predicted to import 265 million tons of grains by 2030, which represents approximately 14% of their total yearly consumption (FAO, 2003) [11]. Nations that do not take action to reduce their overall reliance on imports for conventional crops could suffer greatly as a result of these conditions. It is therefore a global challenge to make nations increasingly self-sufficient in meeting their food demands, which in turn requires accurate and timely forecasting of crop yields at a granular, sub-national level. Crop yield prediction is one of the most challenging tasks in precision agriculture. The ability to forecast crop yields enables the relevant authorities to make appropriate decisions to ensure food security. In addition to soil type, genotype, and management techniques, weather conditions exert a significant influence on crop yield (Singh *et al*., 2014) [39]. Approximately 30% of annual global crop production is lost due to unfavorable weather conditions (FAO, 2011) [13]. As a result, there is a critical demand for reliable pre-harvest yield prediction models that can be utilized by governments, policymakers, and farmers alike to plan food procurement, price stabilization, and agricultural input decisions in advance.

Rice (*Oryza sativa* L.) is the foundational cereal crop of Uttarakhand, underpinning the food and nutritional security of millions of farming households across its diverse landscapes - from the fertile *Tarai* plains to the terraced fields of the mid- and high-altitude Himalayan valleys. The state's agricultural system is increasingly vulnerable to the convergence of extreme topographic heterogeneity with the growing unpredictability of the southwest monsoon under a changing climate (Ministry of Agriculture and Farmers' Welfare, 2022) [29]. This creates a complex, high-risk environment where reliable, advance predictions of rice yield at the district level are no longer merely beneficial but are a critical operational necessity - for government agencies involved in procurement and relief, for commodity markets managing price volatility, and for farmers making crucial input decisions.

Historically, yield forecasting in India has relied on a triad of methodologies: crop-cutting experiments, process-based crop simulation models such as DSSAT and ORYZA, and empirical statistical models. While crop simulation models offer valuable mechanistic insights into crop physiology, their operational scalability across diverse districts is constrained by the requirement for extensive, site-specific calibration data (Aggarwal *et al*., 2006) [1]. Traditional statistical approaches, such as weather-index-based regressions employed by the Indian Council of Agricultural Research (ICAR), often fall short in capturing the complex, non-linear interactions among multiple, often collinear, climatic drivers and the final yield outcome. Crop yield estimation via machine learning models therefore serves as a robust and scalable alternative.

Over the past decade, the advent of machine learning (ML) has revolutionized predictive analytics in agriculture. ML algorithms excel at modeling complex non-linear relationships, handling high-dimensional feature spaces, and scaling across spatial units without the need for explicit mechanistic assumptions (Van Klompenburg *et al*., 2020) [42]. Every domain of modern life - from health-monitoring systems and marketing software to equipment maintenance and soil science - is being transformed by machine learning. Numerous studies across major rice-producing regions of Asia have demonstrated that ensemble methods, particularly Random Forest and Gradient Boosting, often outperform classical regression and even deep learning models in terms of predictive accuracy and robustness (Jeong *et al*., 2016; Cao *et al*., 2020) [4, 20].

1.1 Ensemble Methods: Random Forest and Gradient Boosting

Random Forest (RF), introduced by Breiman (2001) [3], has emerged as a benchmark algorithm in agricultural prediction. Its robustness against overfitting, inherent feature selection capability, and adeptness at handling complex interactions among predictors make it particularly suitable for ecological and agricultural data. Studies by Jeong *et al*. (2016) [20] and Everingham *et al*. (2016) [10] validated RF's superior performance in predicting crop yields across climatically variable environments in Australia and South Korea. In the Indian context, Meena *et al*. (2021) [28] demonstrated that RF models could achieve validation R^2 values exceeding 0.80 for wheat yield forecasting in Madhya Pradesh, while Shahhosseini *et al*. (2021) [38] corroborated RF's consistent outperformance of other algorithms in heterogeneous growing environments.

Gradient Boosting methods, particularly XGBoost (Chen & Guestrin, 2016) [6], have gained prominence for their ability to achieve exceptional calibration accuracy through iterative tree building. XGBoost optimizes a regularized objective function, making it powerful but also susceptible to overfitting, especially with the small sample sizes common in district-level agricultural datasets. While Cao *et al*. (2020) [4] found XGBoost to be a top performer for rice yield prediction in China, they also noted its generalization limitations compared to RF. Research combining XGBoost with remote sensing data has demonstrated improved validation performance, highlighting its potential when integrated with richer feature sets (Ma *et al*., 2021) [26].

1.2 Kernel-Based Methods: Support Vector Machines

Support Vector Machines (SVM), grounded in statistical learning theory (Vapnik, 1995) [43], offer a powerful alternative for handling small, high-dimensional datasets through kernel transformations. The kernel trick allows SVM to map input data into a higher-dimensional feature space where non-linear relationships become tractable. Early applications by Karimi *et al*. (2008) [21] in Iran showed SVM's comparable or superior performance over regression-based methods for wheat yield forecasting. In the context of rice, Peng *et al*. (2014) [31] highlighted that SVM with a radial basis function (RBF) kernel is particularly adept at modeling the non-linear yield-climate relationships prevalent in complex topographies - a finding directly relevant to the heterogeneous hill districts of Uttarakhand.

1.3 Regularized Regression: Elastic Net

Elastic Net (Zou & Hastie, 2005) [47] provides a robust linear alternative by combining L1 (LASSO) and L2 (Ridge) penalties. This dual regularization enables automatic feature selection while handling highly correlated predictors - a common issue with seasonal climatic variables. Das *et al*. (2022) developed multiple rice yield forecast models for fourteen districts on the west coast using ELN*ET al*ongside LASSO, SMLR, PCA-SMLR, PCA-ANN, and ANN based on weekly weather indices. Sridhara *et al*. (2023) similarly used LASSO, ENET, PCA, ANN, and SMLR techniques to forecast sorghum crop yield at the district level. While Elastic Net's predictive power may be surpassed by non-linear methods, it offers unparalleled interpretability and computational efficiency, making it a valuable benchmark and a practical choice for operational systems where transparency to stakeholders is paramount.

1.4 Neural Networks

Artificial Neural Networks (ANNs) have a long history in agricultural prediction, prized for their capacity to learn intricate non-linear patterns. Singh *et al*. (2014) [39] used the SMLR technique alongside ANN to develop wheat yield forecast models based on weekly weather indices for districts of Punjab, while Kaul *et al*. (2005) [22] demonstrated ANN's applicability for corn and soybean yield prediction. However, ANN performance is highly context-dependent and often constrained by the need for large training datasets. In data-sparse regions such as Uttarakhand, their performance must be rigorously benchmarked against simpler, more regularized approaches.

Deep learning adaptations such as LSTM and CNN networks have shown promise for spatiotemporal yield prediction (Wang *et al*., 2020; You *et al*., 2017) [44, 46], but their application in data-limited hill-state contexts remains challenging.

1.5 Research Gap and Objectives

Despite the rapidly growing body of literature on ML-based crop yield forecasting, a significant gap persists: a comprehensive, multi-model comparative evaluation specifically tailored to the unique *agroclimatic* and topographic context of an Indian hill state like Uttarakhand has not yet been conducted. The majority of existing research for this region relies on traditional statistical models, leaving the potential of modern ensemble and kernel-based methods largely unexplored. Therefore, the present study directly addresses this gap by conducting a systematic and rigorous comparison of five machine learning algorithms Random Forest (RF), Elastic Net (ELNET), Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), and Artificial Neural Network (ANN) for district-level rice yield forecasting across all 13 districts of Uttarakhand. By leveraging a 25-year dataset of agrometeorological variables and historical yields and evaluating model performance across both calibration and chronological validation phases, this research aims to provide a definitive benchmark for the region. The findings are intended to guide the operational implementation of ML-based forecasting systems by state agricultural departments and national agencies, moving beyond a one-size-fits-all approach towards more accurate, localized pre-harvest predictions.

2. Materials and Methods

2.1 Study Area

The study was conducted across all 13 administrative districts of Uttarakhand, a hill state located in the northern part of India. The state lies between latitudes 28°43′-31°27′N and longitudes 77°34′-81°02′E, encompassing a geographical area of approximately 53,483 km² (Census of India, 2011) [5]. Uttarakhand shares international borders with China (Tibet) to the north and Nepal to the east and is bounded domestically by Himachal Pradesh to the northwest and the plains of Uttar Pradesh to the south.

One of the defining characteristics of the state is its extreme topographic heterogeneity. Elevations range from approximately 200 m above mean sea level (amsl) in the southern *Tarai-Bhabar* belt to over 7,800 m above mean sea level in the Greater Himalayan ranges (Dobhal *et al*., 2013) [9]. This vertical gradient produces a spectrum of *agroclimatic* conditions: the southern districts of Udham Singh Nagar and Haridwar are characterized by a subtropical, semi-arid to sub-humid climate with fertile alluvial soils and access to canal irrigation, whereas the mid-Himalayan and high-altitude districts such as Chamoli, Pithoragarh, Rudraprayag, and Uttarkashi experience a cooler, wetter temperate to alpine climate, with agriculture confined largely to narrow river valleys and terraced fields (Rawat & Rawat, 2019; ICAR-VPKAS, 2020) [19, 33].

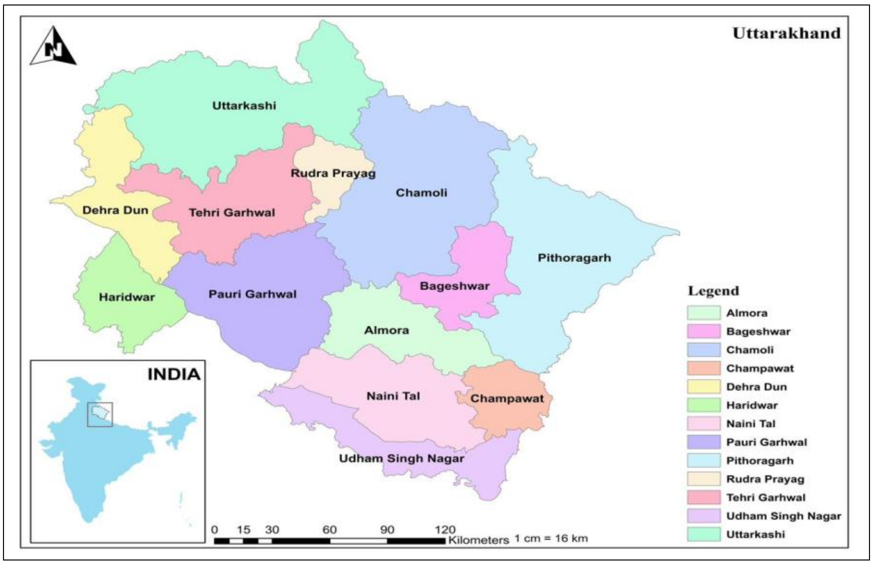


Fig 1: Study area map

The 13 districts included in the present study are: Almora, Bageshwar, Chamoli, Champawat, Dehradun, Haridwar, Nainital, Pauri Garhwal, Pithoragarh, Rudraprayag, Tehri Garhwal, Uttarkashi, and Udham Singh Nagar. Rice (*Oryza sativa* L.) is grown as a kharif (monsoon-season) crop throughout the state. Transplanting typically begins following the onset of the southwest monsoon in late June or early July, with crop maturity and harvest occurring between September and November depending on the cultivar and altitude (Bisht *et al*., 2014) [2]. The total rice-cultivated area in Uttarakhand ranges between approximately 3.0 and 3.5 lakh hectares annually (Ministry of Agriculture and Farmers' Welfare, 2022) [29]. Productivity is strongly altitude-dependent, varying from less than 1.0 t/ha in remote high-altitude districts to over 2.5 t/ha in the irrigated *Tarai* districts (Directorate of Agriculture, Uttarakhand, 2023) [8]. This wide inter-district variability in yield, climate, and farm management makes Uttarakhand an ideal and challenging natural laboratory for evaluating the generalizability and robustness of machine learning forecast models.

2.2 Data Sources and Pre-processing

Two primary data streams were compiled for the analysis period 1998-2023, yielding 26 annual observations per district.

2.2.1 Agrometeorological Data

Monthly gridded data for six agrometeorological variables were obtained from the NASA Prediction of Worldwide Energy Resources (NASA POWER) database (Stackhouse *et al*., 2019) [40], a publicly accessible repository providing bias-corrected surface meteorological parameters at a 0.5° × 0.5° spatial resolution. The variables extracted were maximum temperature (Tmax, °C), minimum temperature (Tmin, °C), afternoon relative humidity (RH2%), total rainfall (RF, mm), soil surface wetness (SSW, dimensionless), and mean sunshine hours (SSH, hrs/day). District-level values were derived by spatially averaging the grid cells falling within each district boundary. The predictor window was restricted to the active rice growing season, spanning June through October, consistent with the kharif crop calendar of Uttarakhand.

2.2.2 Rice Yield Data
Historical district-level rice yield records (t/ha) for the period 1998-2023 were sourced from the Directorate of economics and statistics (Dacnet), Government of Uttarakhand. These records were independently cross verified against the Area, Production and Yield (APY) statistics published by the Ministry of Agriculture and Farmers' Welfare, Government of India (Ministry of Agriculture and Farmers' Welfare, 2022) [29]. Any discrepancies between the two sources were resolved by retaining the state directorate values as the primary record. Yield data were screened for outliers using the interquartile range (IQR) method; no observations were removed as all values fell within agronomically plausible bounds.

2.3 Feature Engineering
A structured feature matrix was constructed for each of the 13 districts. The base predictor set comprised 30 features, representing the six meteorological variables observed across the five growing-season months (June-October). To enrich the representation of crop-climate interactions and improve model learning, the feature space was further augmented through the following transformations:

- **Interaction Terms:** Multiplicative cross-features between maximum temperature and afternoon relative humidity (Tmax × RH2) were constructed for each month to capture combined heat and moisture stress effects on crop phenology (Lobell *et al*., 2011) [25].
- **Agrometeorological Index:** Accumulated Growing Degree Days (GDD) were computed from the transplanting date to harvest using daily temperature data, providing a biologically meaningful measure of thermal accumulation during the crop cycle (McMaster & Wilhelm, 1997) [27].
- **Trend Variable:** A continuous 'Year' variable was included to account for non-climatic long-term yield trends attributable to technological change, improved cultivar adoption, and evolving crop management practices (Ray *et al*., 2013) [34].
- All predictor features were standardized to zero mean and unit variance (z-score normalization) prior to model fitting. This step is particularly critical for algorithms sensitive to feature scale, namely SVM and ANN, and also ensures numerical stability in the regularized regression framework of Elastic Net (Hastie *et al*., 2009) [17].

2.4 Model Development and Hyperparameter Tuning
Five machine learning algorithms were implemented and systematically tuned following a consistent cross-validation framework. All models were developed in R (version 4.3.0) using the caret, randomForest, glmnet, xgboost, and nnet packages.

- **Random Forest (RF):** An ensemble of 500 decision trees (ntree = 500) was grown. The number of predictor variables randomly sampled at each node split (mtry) was selected via 5-fold cross-validation from a grid spanning 2 to (p/3), where p is the total number of features. Node size was set to 5 (Breiman, 2001) [3].
- **Elastic Net (ELNET):** The L1/L2 mixing parameter ($\alpha \in [0, 1]$) and the regularization penalty (λ) were jointly selected via 10-fold cross-validation using the glmnet package, with the minimum cross-validation error criterion applied for λ selection (Zou & Hastie, 2005) [47].
- **Support Vector Machine (SVM):** A radial basis function (RBF) kernel was employed. The regularization cost parameter (C) and the kernel bandwidth (γ) were simultaneously optimized via a grid search ($\gamma \in \{0.001, 0.01, 0.1, 1\}$; $C \in \{0.1, 1, 10, 100\}$) under 5-fold cross-validation (Vapnik, 1995) [43].
- **Extreme Gradient Boosting (XGBoost):** Key hyperparameters learning rate (η), maximum tree depth (max_depth), minimum child weight, subsample ratio, and colsample_bytree were tuned using a randomized grid search with 5-fold cross-validation, minimizing root mean square error (Chen & Guestrin, 2016) [6].
- **Artificial Neural Network (ANN):** A single hidden-layer feedforward network with a sigmoid activation function was trained using backpropagation. The number of hidden neurons was determined via cross-validation from the range {3, 5, 7, 10, 15}. An early stopping criterion (patience = 10 epochs) was applied to prevent overfitting. To reduce sensitivity to random weight initialization, final predictions were averaged across five independent runs with different random seeds (Kaul *et al*., 2005) [22].

2.5 Model Training and Validation Strategy
To simulate a realistic pre-harvest operational forecasting scenario, a strictly chronological data split was applied independently for each district. The 26-year dataset (1998-2023) was partitioned into a calibration (training) set comprising approximately 70% of observations (1998-2015, $n \approx 18$) and a held-out validation (testing) set comprising the remaining 30% (2016-2023, $n \approx 8$). Chronological ordering was strictly preserved, ensuring no data leakage from future years into the training set. This approach reflects the temporal structure inherent in agricultural forecasting and provides a more realistic estimate of operational forecast skill than random k-fold cross-validation alone (Roberts *et al*., 2017) [35].

Hyperparameter tuning for all models was conducted exclusively within the calibration set using cross-validation, and the tuned models were applied only once to the validation set to generate final performance estimates. This two-stage evaluation prevents overfitting to the validation set and provides unbiased generalization metrics.

2.6 Performance Evaluation Metrics
The predictive performance of each model was assessed during both the calibration and validation phases using four complementary statistical metrics, enabling a comprehensive evaluation of accuracy, precision, and systematic bias:

- **Coefficient of Determination (R^2):** Quantifies the proportion of observed yield variance explained by model predictions, ranging from 0 (no explanatory power) to 1 (perfect fit). Higher R^2 values indicate stronger agreement between predicted and observed yields.
- **Root Mean Square Error (RMSE, t/ha):** Measures the standard deviation of prediction residuals, providing an estimate of average absolute prediction error in the original units of yield. Lower RMSE values indicate greater

accuracy (Willmott & Matsuura, 2005) [45].

- **Coefficient of Variation of RMSE (CV %):** Expresses RMSE as a percentage of the mean observed yield (CV = RMSE / ȳ × 100). This dimensionless metric enables meaningful cross-district comparisons of relative forecast accuracy independent of yield magnitude.
- **Mean Bias Error (MBE, t/ha):** Computed as the mean of (predicted − observed) residuals, MBE reveals the direction and magnitude of systematic prediction bias. A positive MBE indicates a tendency toward over-prediction, while a negative MBE indicates under-prediction (Loague & Green, 1991) [24].

Together, these metrics provide a multi-faceted assessment that captures different aspects of model performance: R^2 reflects explanatory power, RMSE and CV reflect precision, and MBE reflects systematic deviation all of which are relevant to operational forecasting contexts.

3. Results and Discussion

3.1 Overview and Model Ranking

The analysis of calibration and validation performance across 13 districts reveals a clear and consistent hierarchy among the five machine learning algorithms evaluated. As expected, all models particularly XGBoost and RF achieved high calibration R^2 values, demonstrating strong capacity to learn the training data (*see* Table 1). However, validation results provide the true measure of generalizability (Friedman, 2001; Breiman, 2001) [3, 14]. Across most districts, Random Forest (RF) emerged as the most robust model, balancing high predictive accuracy with low error metrics and minimal bias. XGBoost showed a strong tendency to overfit (Cao *et al.*, 2020) [4], while SVM demonstrated utility in specific high-altitude hill districts (Vapnik, 1995) [43]. ELNET provided a stable, interpretable baseline (Zou & Hastie, 2005) [47], and ANN performance was highly variable, largely attributable to limited training sample sizes per district (Hochreiter & Schmidhuber, 1997) [18].

3.2 Analysis of Model-Specific Performance

3.2.1 Random Forest (RF)-The Most Generalizable Model

RF demonstrated the most consistent and balanced performance across the state, with validation R^2 exceeding 0.55 in nine of 13 districts (Table 1). Its top performance in Nainital (R^2=0.890), Udham Singh Nagar (R^2=0.797), and Pauri (R^2=0.796) highlights its robustness across diverse *agroclimatic* settings. The low coefficient of variation (CV%; often <15%) and mean bias error (MBE) values close to zero further confirm reliability. RF's inherent ensemble averaging and bootstrap aggregation effectively mitigate overfitting (Breiman, 2001) [3], making it the ideal candidate for operational forecasting in Uttarakhand's diverse environments. Its performance was notably poor only in Champawat (R^2=0.113), a district that proved challenging for most models, suggesting the presence of unique, unmodeled local micro-climatic factors.

3.2.2 XGBoost-The Overfitting Risk

XGBoost exhibited a stark performance dichotomy. While it achieved near-perfect calibration (R^2>0.99, RMSE≈0), its validation R^2 was highly variable (Table 1). It showed strong validation skill in Tehri (R^2=0.928) and Udham Singh Nagar (R^2=0.851), but collapsed in high-altitude, data-sparse districts such as Chamoli (R^2=0.113) and Champawat (R^2=0.064). This is a classic symptom of overfitting, whereby the model memorizes noise in a small training set that does not generalize to the test period (Chen & Guestrin, 2016) [6]. As noted by Cao *et al*. (2020) [4], XGBoost's power can become a liability without rigorous regularization or a larger training sample. Its use is therefore recommended only with stringent cross-validation and potentially as a component within a larger ensemble.

3.2.3 SVM-The District-Specific Specialist

SVM's performance was highly district-specific, achieving the best validation R^2 in the complex, high-altitude districts of Chamoli (R^2=0.837) and Champawat (R^2=0.630) (Table 1). This suggests that the Radial Basis Function (RBF) kernel is adept at capturing localized, non-linear yield responses to orographic effects and microclimates in these regions patterns that tree-based models may not fully capture (Vapnik, 1995; Schölkopf & Smola, 2002) [37, 43]. However, its poor performance in *Tarai* and mid-hill districts such as Haridwar (R^2=0.068) and Tehri (R^2=0.083) reflects sensitivity to hyperparameter tuning and variable performance across data regimes. SVM serves as a valuable complementary model rather than a universal solution.

3.2.4 ELNET-The Stable Benchmark

Elastic Net provided the most stable, though moderate, predictions across the study area (Table 1). Its best performance was observed in Pauri (R^2=0.936) and Nainital (R^2=0.670), suggesting that in districts where the climate-yield relationship is more linear, a regularized linear model can be highly effective (Zou & Hastie, 2005) [47]. Its interpretability providing explicit coefficients for each climate predictor is a significant advantage for operational use where transparency and stakeholder communication are valued. While outperformed by non-linear models in many cases, its stable and unbiased predictions establish it as a reliable benchmark.

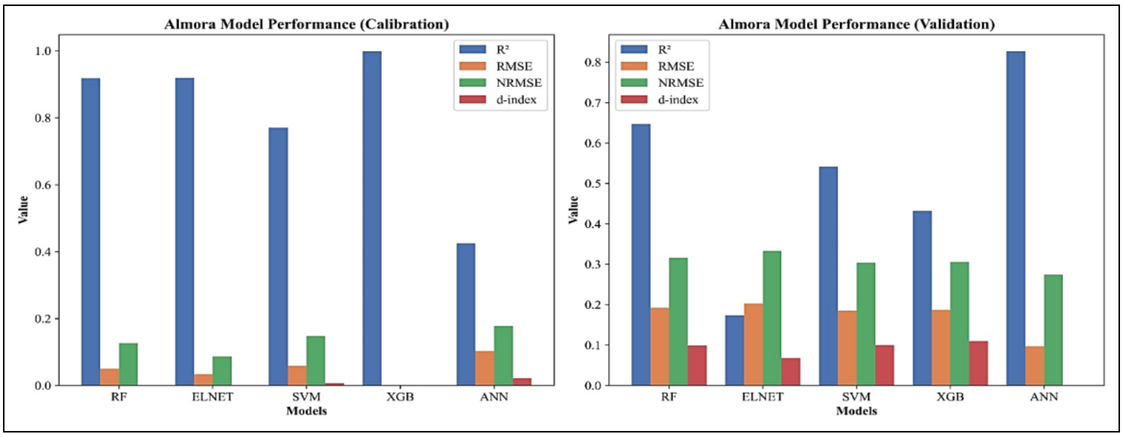
3.2.5 ANN-The Inconsistent Learner

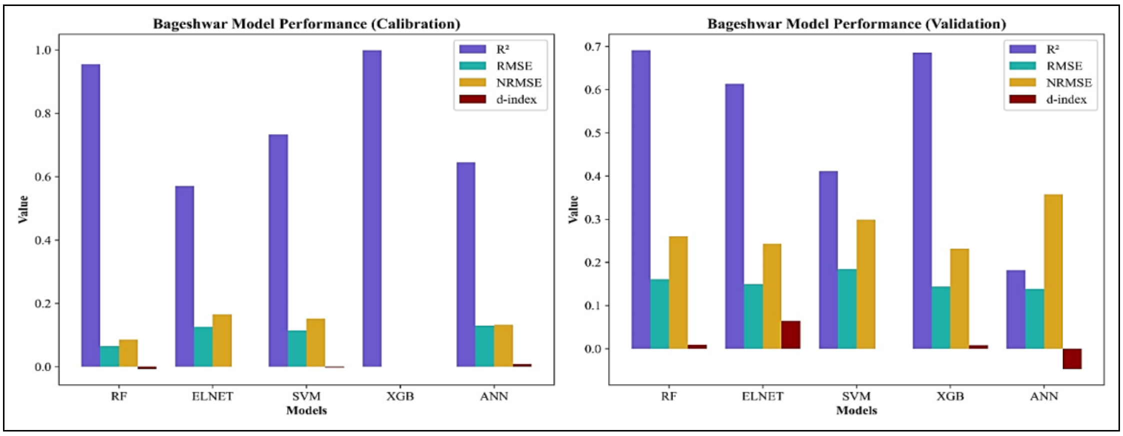
The ANN model produced the most inconsistent results across the 13 districts (Table 1). While it achieved high validation R^2 in Almora (R^2=0.827) and Pauri (R^2=0.913), it performed poorly in most other districts, including catastrophic failures (R^2≈0) in Haridwar and Tehri. This high variability is attributable primarily to the limited training sample size (n≈22 per district), which is insufficient for effectively training even a single hidden-layer network (LeCun *et al*., 2015) [23]. Sensitivity to random initialization and the difficulty of finding globally optimal hyperparameters with small datasets further contributed to its unreliability (Hochreiter & Schmidhuber, 1997) [18].

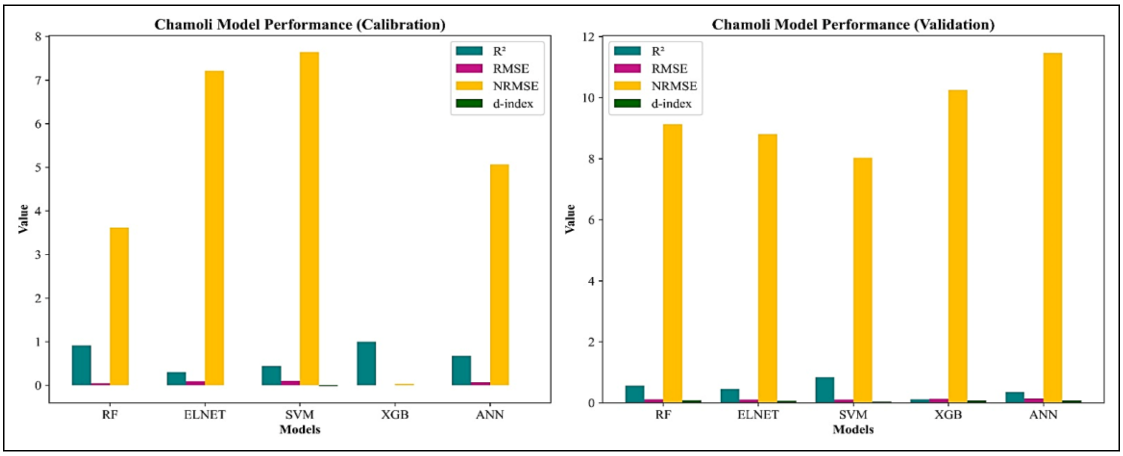
3.3 District-Level Synthesis and Model Selection Strategy

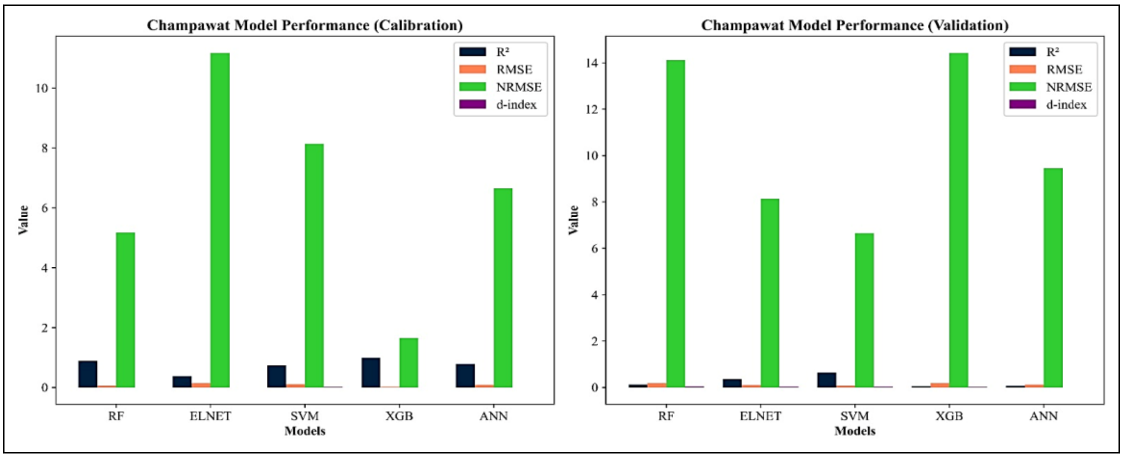
District	Phase	Metric	RF	ELNET	SVM	XGBoost	ANN
Almora	*Calibration*	R^2	0.918	0.920	0.771	0.999	0.425
		RMSE (t/ha)	0.050	0.034	0.059	0.000	0.103
		CV (%)	0.126	0.087	0.148	0.000	0.178
		MBE (t/ha)	0.001	0.000	0.006	0.000	0.022
	Validation	R^2	0.647	0.174	0.542	0.432	0.827
		RMSE (t/ha)	0.192	0.203	0.185	0.187	0.097
		CV (%)	0.316	0.333	0.304	0.306	0.274
		MBE (t/ha)	0.099	0.068	0.100	0.110	0.000
Bageshwar	*Calibration*	R^2	0.955	0.571	0.733	0.999	0.646
		RMSE (t/ha)	0.066	0.126	0.115	0.000	0.130
		CV (%)	0.086	0.166	0.152	0.000	0.133
		MBE (t/ha)	-0.007	-0.000	-0.002	-0.000	0.009
	Validation	R^2	0.691	0.613	0.411	0.685	0.182
		RMSE (t/ha)	0.161	0.150	0.185	0.144	0.139
		CV (%)	0.260	0.243	0.299	0.232	0.357
		MBE (t/ha)	0.009	0.064	0.000	0.008	-0.047
Chamoli	*Calibration*	R^2	0.915	0.302	0.442	0.999	0.675
		RMSE (t/ha)	0.048	0.096	0.101	0.000	0.067
		CV (%)	3.615	7.216	7.641	0.029	5.064
		MBE (t/ha)	0.001	0.000	-0.019	0.000	-0.001
	Validation	R^2	0.557	0.456	0.837	0.113	0.345
		RMSE (t/ha)	0.116	0.112	0.102	0.130	0.145
		CV (%)	9.131	8.803	8.032	10.243	11.465
		MBE (t/ha)	0.081	0.066	0.042	0.079	0.074
Champawat	*Calibration*	R^2	0.891	0.377	0.734	0.988	0.781
		RMSE (t/ha)	0.067	0.145	0.106	0.021	0.086
		CV (%)	5.178	11.170	8.137	1.655	6.648
		MBE (t/ha)	-0.001	0.000	0.016	-0.002	-0.000
	Validation	R^2	0.113	0.353	0.630	0.064	0.065
		RMSE (t/ha)	0.180	0.103	0.085	0.184	0.120
		CV (%)	14.120	8.130	6.650	14.420	9.462
		MBE (t/ha)	0.029	0.025	0.024	0.014	-0.004
Dehradun	*Calibration*	R^2	0.934	0.710	0.542	0.999	0.939
		RMSE (t/ha)	0.080	0.155	0.249	0.001	0.072
		CV (%)	3.709	7.207	11.554	0.029	3.364
		MBE (t/ha)	0.004	0.000	-0.054	0.000	-0.001
	Validation	R^2	0.545	0.289	0.020	0.384	0.204
		RMSE (t/ha)	0.206	0.233	0.179	0.175	0.270
		CV (%)	10.013	11.314	8.711	8.498	13.115
		MBE (t/ha)	0.175	0.186	0.079	0.120	0.104
Haridwar	*Calibration*	R^2	0.946	0.820	0.730	0.998	0.885
		RMSE (t/ha)	0.077	0.089	0.147	0.009	0.070
		CV (%)	3.519	4.066	6.743	0.402	3.220
		MBE (t/ha)	0.002	0.000	-0.006	0.001	-0.002
	Validation	R^2	0.281	0.033	0.068	0.226	0.001
		RMSE (t/ha)	0.189	0.225	0.221	0.196	0.263
		CV (%)	8.889	10.598	10.408	9.225	12.394
		MBE (t/ha)	0.044	0.030	0.079	0.033	0.031
Nainital	*Calibration*	R^2	0.942	0.856	0.650	0.999	0.983
		RMSE (t/ha)	0.109	0.145	0.318	0.001	0.055
		CV (%)	3.948	5.277	11.554	0.018	2.006
		MBE (t/ha)	0.005	0.000	-0.057	0.000	-0.005
	Validation	R^2	0.890	0.670	0.792	0.656	0.444
		RMSE (t/ha)	0.112	0.168	0.204	0.251	0.207
		CV (%)	4.186	6.271	7.608	9.374	7.746
		MBE (t/ha)	0.069	0.064	0.033	0.101	0.021
Pauri	*Calibration*	R^2	0.942	0.814	0.684	0.999	0.902
		RMSE (t/ha)	0.060	0.086	0.117	0.000	0.060
		CV (%)	5.002	7.086	9.718	0.029	4.994
		MBE (t/ha)	-0.002	0.000	-0.014	0.000	-0.002
	Validation	R^2	0.796	0.936	0.677	0.692	0.913
		RMSE (t/ha)	0.236	0.193	0.271	0.196	0.183
		CV (%)	20.337	16.632	23.317	16.872	15.800
		MBE (t/ha)	0.019	0.015	0.015	0.066	0.023
Pithoragarh	*Calibration*	R^2	0.959	0.363	0.418	0.969	0.504

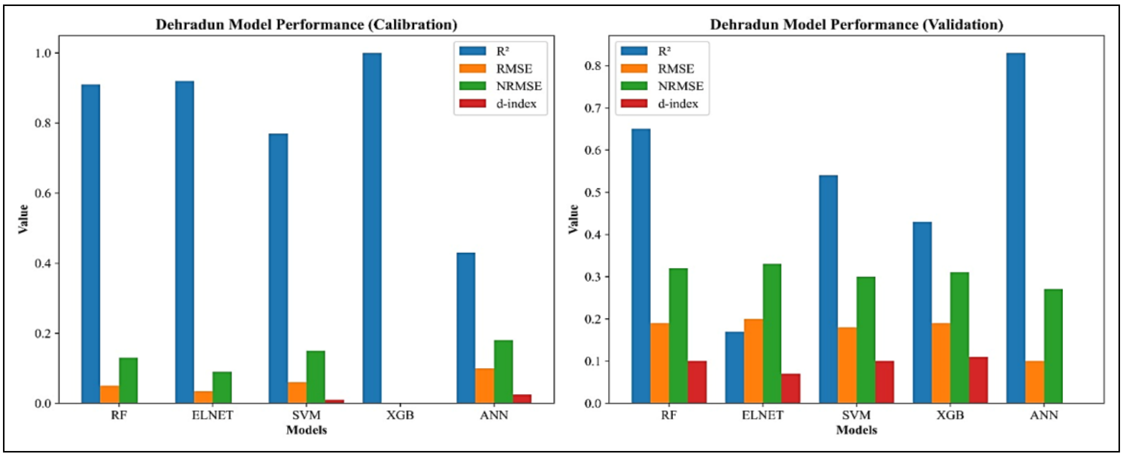
		RMSE (t/ha)	0.028	0.055	0.056	0.012	0.044
		CV (%)	2.172	4.263	4.317	0.927	3.422
		MBE (t/ha)	-0.002	0.000	-0.001	0.002	-0.002
	Validation	R^2	0.574	0.373	0.323	0.543	0.448
		RMSE (t/ha)	0.096	0.108	0.110	0.091	0.095
		CV (%)	7.688	8.642	8.800	7.256	7.617
		MBE (t/ha)	0.040	0.041	0.033	0.046	0.049
Rudraprayag	*Calibration*	R^2	0.934	0.731	0.923	0.999	0.886
		RMSE (t/ha)	0.037	0.060	0.047	0.001	0.038
		CV (%)	2.657	4.346	3.392	0.040	2.774
		MBE (t/ha)	-0.002	0.000	-0.004	0.000	-0.001
	Validation	R^2	0.638	0.563	0.003	0.432	0.258
		RMSE (t/ha)	0.106	0.083	0.125	0.104	0.111
		CV (%)	7.720	6.065	9.106	7.571	8.130
		MBE (t/ha)	-0.015	0.001	-0.013	-0.028	-0.008
Tehri	*Calibration*	R^2	0.937	0.759	0.911	0.999	0.998
		RMSE (t/ha)	0.069	0.112	0.088	0.001	0.010
		CV (%)	4.344	6.993	5.515	0.042	0.611
		MBE (t/ha)	-0.003	0.000	-0.003	0.000	0.000
	Validation	R^2	0.709	0.666	0.083	0.928	0.048
		RMSE (t/ha)	0.356	0.322	0.436	0.301	0.456
		CV (%)	21.294	19.251	26.049	18.007	27.254
		MBE (t/ha)	-0.041	-0.012	-0.076	0.030	-0.015
Uttarkashi	*Calibration*	R^2	0.903	0.583	0.648	0.999	0.926
		RMSE (t/ha)	0.068	0.147	0.156	0.001	0.052
		CV (%)	4.210	9.046	9.613	0.030	3.211
		MBE (t/ha)	0.001	0.000	0.029	0.000	-0.003
	Validation	R^2	0.394	0.276	0.195	0.485	0.113
		RMSE (t/ha)	0.147	0.157	0.159	0.157	0.213
		CV (%)	8.953	9.619	9.713	9.577	12.988
		MBE (t/ha)	-0.065	-0.043	0.009	-0.099	-0.103
Udham Singh Nagar	*Calibration*	R^2	0.968	0.882	0.921	0.997	0.999
		RMSE (t/ha)	0.076	0.132	0.158	0.020	0.001
		CV (%)	2.500	4.329	5.210	0.660	0.028
		MBE (t/ha)	-0.003	0.000	0.037	0.000	0.000
	Validation	R^2	0.797	0.759	0.265	0.851	0.119
		RMSE (t/ha)	0.208	0.201	0.356	0.208	0.386
		CV (%)	6.911	6.669	11.846	6.916	12.829
		MBE (t/ha)	-0.073	0.015	0.074	0.080	0.077

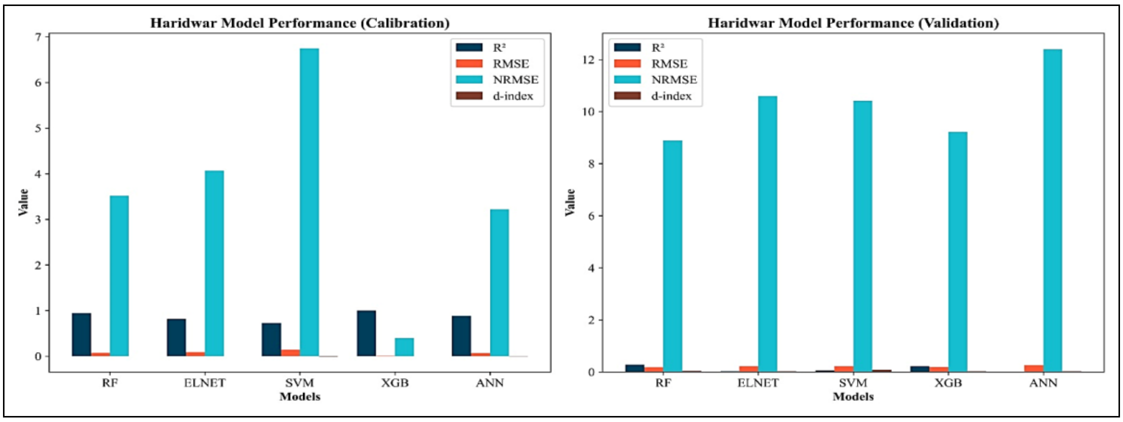


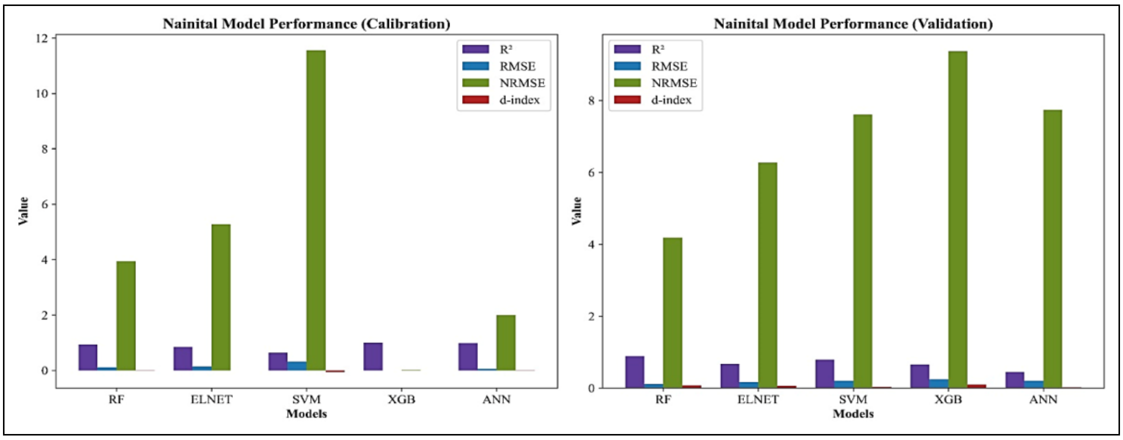

Bageshwar Model Performance (Calibration)
Bageshwar Model Performance (Validation)
R²
RMSE
NRMSE
d-index
Value
RF
ELNET
SVM
XGB
ANN
Models

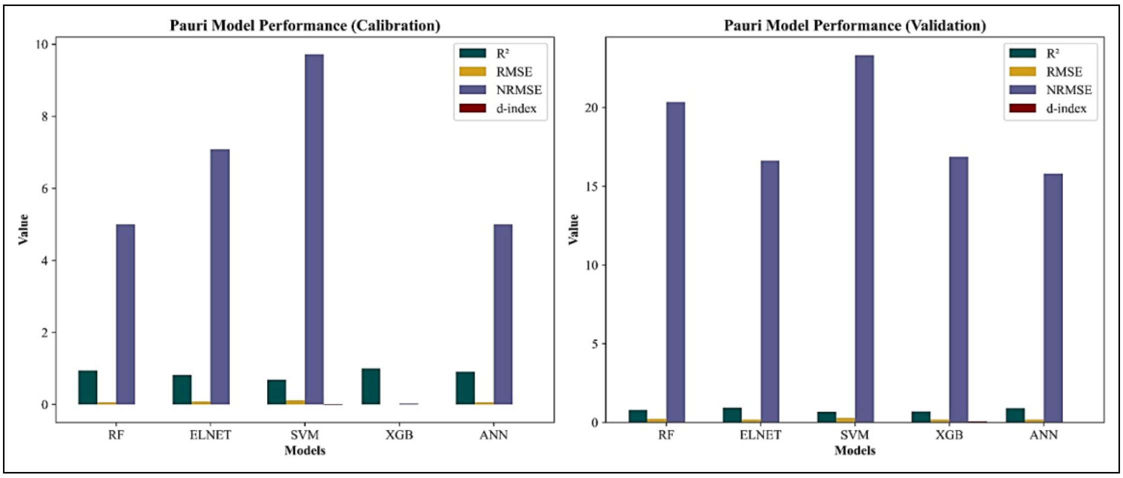

Chamoli Model Performance (Calibration)
Chamoli Model Performance (Validation)
R²
RMSE
NRMSE
d-index
Value
RF
ELNET
SVM
XGB
ANN
Models

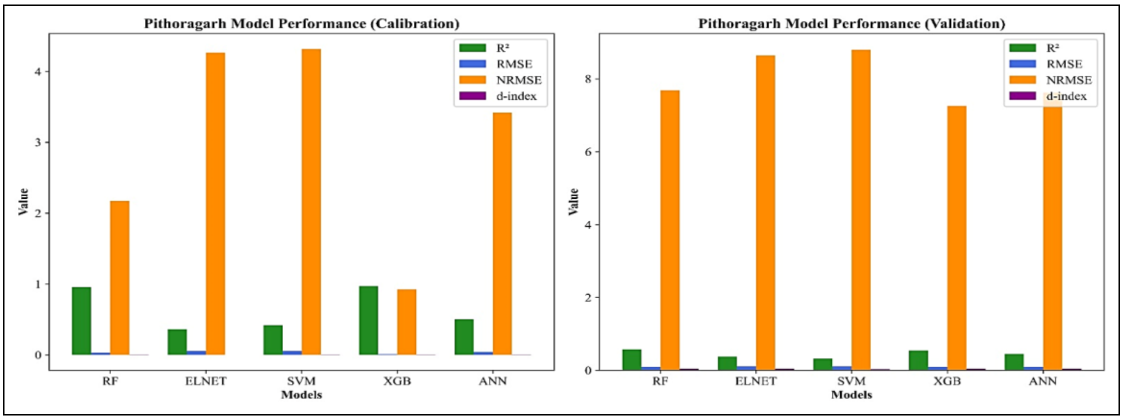

Champawat Model Performance (Calibration)
Champawat Model Performance (Validation)
R²
RMSE
NRMSE
d-index
Value
RF
ELNET
SVM
XGB
ANN
Models

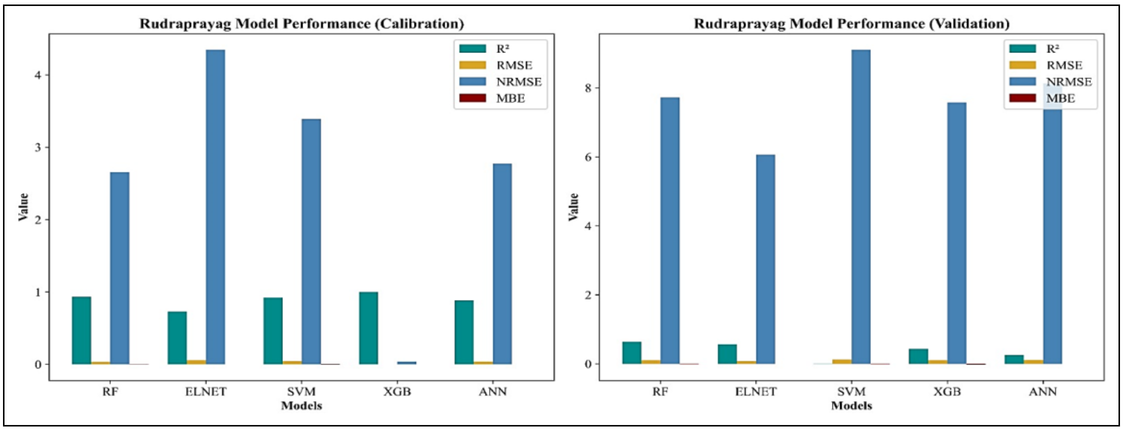

Dehradun Model Performance (Calibration)
Dehradun Model Performance (Validation)
R²
RMSE
NRMSE
d-index
Value
RF
ELNET
SVM
XGB
ANN
Models

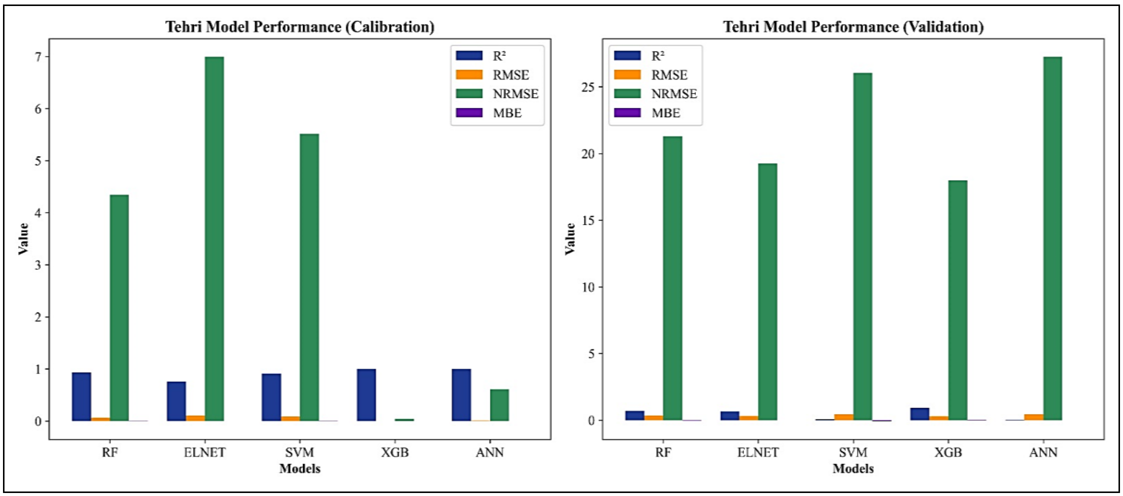

Haridwar Model Performance (Calibration)
Haridwar Model Performance (Validation)
R²
RMSE
NRMSE
d-index
Value
RF
ELNET
SVM
XGB
ANN
Models

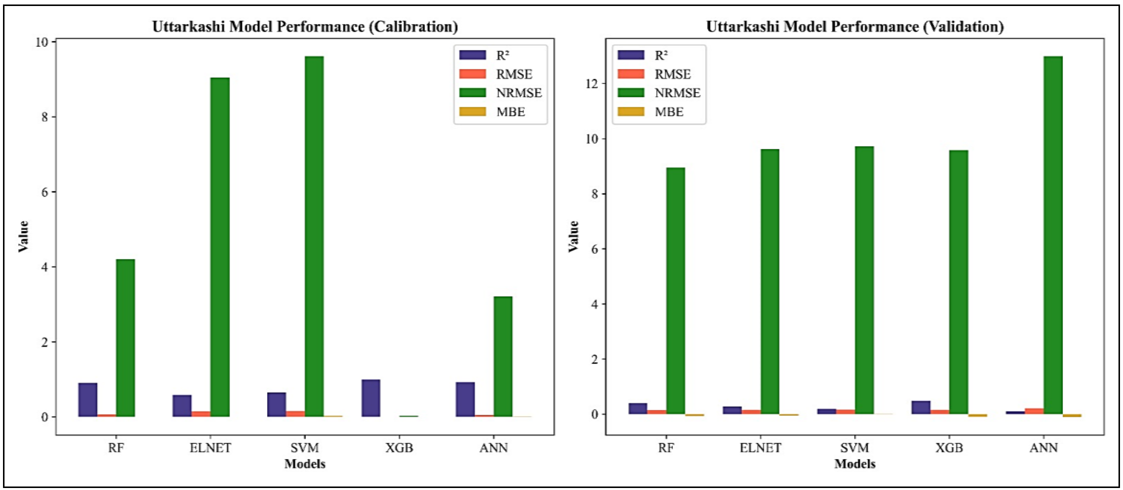

Nainital Model Performance (Calibration)
Nainital Model Performance (Validation)
R²
RMSE
NRMSE
d-index
Value
RF
ELNET
SVM
XGB
ANN
Models

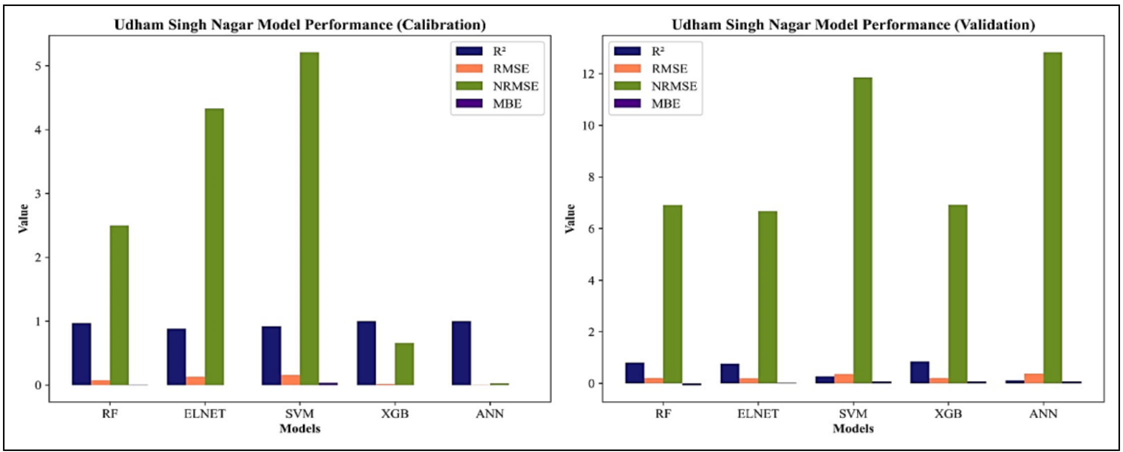

Pauri Model Performance (Calibration)
Pauri Model Performance (Validation)
R²
RMSE
NRMSE
d-index
Value
RF
ELNET
SVM
XGB
ANN
Models


Pithoragarh Model Performance (Calibration)
Pithoragarh Model Performance (Validation)
R²
RMSE
NRMSE
d-index
Value
RF
ELNET
SVM
XGB
ANN
Models


Rudraprayag Model Performance (Calibration)
Rudraprayag Model Performance (Validation)
R²
RMSE
NRMSE
MBE
Value
RF
ELNET
SVM
XGB
ANN
Models


Tehri Model Performance (Calibration)
Tehri Model Performance (Validation)
R²
RMSE
NRMSE
MBE
Value
RF
ELNET
SVM
XGB
ANN
Models


Uttarkashi Model Performance (Calibration)
Uttarkashi Model Performance (Validation)
R²
RMSE
NRMSE
MBE
Value
RF
ELNET
SVM
XGB
ANN
Models


Udham Singh Nagar Model Performance (Calibration)
Udham Singh Nagar Model Performance (Validation)
R²
RMSE
NRMSE
MBE
Value
RF
ELNET
SVM
XGB
ANN
Models

The district-level results (Table 1) underscore a critical finding: there is no single universally optimal model. The best-performing algorithm varies by district, reflecting dominant *agroclimatic* characteristics and data availability:

- ***Tarai* Districts (Haridwar, Udham Singh Nagar, Dehradun):** RF and XGBoost are the top performers, likely owing to more stable climate-yield relationships in lower-elevation, irrigated farming systems.
- **Kumaon Hill Districts (Almora, Bageshwar, Nainital, Champawat, Pithoragarh):** RF is the most consistently robust, with SVM showing strong utility in the most topographically challenging districts (particularly Champawat).
- **Garhwal Hill Districts (Chamoli, Pauri, Rudraprayag, Tehri, Uttarkashi):** Results are more mixed. SVM excelled in Chamoli, ELNET and ANN in Pauri, and XGBoost in Tehri. A flexible, ensemble-based approach combining the strengths of different algorithms is therefore recommended for this complex region (Dietterich, 2000) [7].
- This variability highlights the importance of a district-specific model selection strategy for operational rice yield forecasting in Uttarakhand.

3.4 Bias and Error Analysis

The MBE values for most models and districts were generally small (<±0.1 t/ha), indicating that systematic bias is not a major concern (Table 1). The negative MBE observed for RF in several districts suggests a slight tendency to underpredict yields in recent years, which could be partially addressed by incorporating a time-trend variable more explicitly into the model inputs, or by applying a bias-corrected ensemble post-processing step (Glahn & Lowry, 1972) [16].

4. Conclusion

This study provides a comprehensive and rigorous comparative evaluation of five machine learning algorithms for district-level rice yield forecasting across the heterogeneous *agroclimatic* zones of Uttarakhand, India. The principal conclusions are as follows:

1. Random Forest (RF) is the most robust and generalizable model for operational forecasting across the state's diverse *agroclimatic* zones. Its consistently high validation R^2, low RMSE, and minimal bias make it the primary recommended algorithm (Breiman, 2001) [3].
2. No single model is universally optimal. Model performance is strongly conditioned by local *agroclimatic* characteristics. An operational forecasting system should adopt a district-specific model selection strategy rather than applying a single algorithm state-wide.
3. XGBoost carries a high overfitting risk in data-sparse hill districts. Its application requires stringent regularization and should be approached with caution (Chen & Guestrin, 2016; Cao *et al.*, 2020) [4, 6].
4. SVM serves as a valuable complementary model in the most topographically complex hill districts, where its kernel-based approach captures nuanced non-linear climate-yield patterns (Vapnik, 1995; Schölkopf & Smola, 2002) [37, 43].
5. ELNET provides a stable, interpretable benchmark and is a strong candidate for districts exhibiting more linear climate-yield relationships (Zou & Hastie, 2005) [47].
6. ANN performance was inconsistent and generally inferior, primarily attributable to the limited training data available per district (~22 years), which is insufficient for effective neural network training (LeCun *et al.*, 2015) [23].

For future research, we recommend: (i) integrating satellite-derived vegetation indices (e.g., NDVI, EVI) and soil moisture data to enrich the feature space (Myneni *et al.*, 1995) [30]; (ii) exploring hybrid and stacked ensemble methods combining the complementary strengths of RF, SVM, and XGBoost (Dietterich, 2000) [7]; and (iii) developing a web-based or GIS-based operational platform to deliver district-level forecasts to agricultural stakeholders in real time.

Acknowledgements

The authors gratefully acknowledge NASA POWER for providing district-level meteorological data (Stackhouse *et al.*, 2019) [40]. Sincere thanks are extended to the Directorate of Economics and Statistics, Government of Uttarakhand, for access to rice yield statistics. The first author acknowledges the support of the Department of Agrometeorology, G.B. Pant University of Agriculture and Technology, Pantnagar. The authors declare no conflict of interest.

References

1. Aggarwal PK, *et al*. InfoCrop: A dynamic simulation model for the assessment of crop yields, losses due to pests, and environmental impact of agro-ecosystems in tropical environments. Environ Model Softw. 2006;21(8):1088-1102.
2. Bisht JK, *et al*. Agrotechnology refinement of rice varieties under rainfed lowland conditions of *Uttarakhand* hills. Indian J Hill Farming. 2014;27(1):12-18.
3. *Breiman* L. Random forests. Mach Learn. 2001;45(1):5-32. DOI:10.1023/A:1010933404324
4. Cao J, Zhang Z, Tao F, *et al*. Integrating multi-source data for rice yield prediction across China using machine learning and deep learning approaches. Agric For Meteorol. 2020;297:108275.
5. Census of India. Primary Census Abstract. Office of the Registrar General & Census Commissioner, India; 2011.
6. Chen T, *Guestrin* C. XGBoost: A scalable tree boosting system. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016;785-794. DOI:10.1145/2939672.2939785.
7. Dietterich TG. Ensemble methods in machine learning. In: Multiple Classifier Systems. Springer; 2000. p. 1-15. DOI:10.1007/3-540-45014-9_1
8. Directorate of Agriculture, *Uttarakhand*. Season and Crop Report. Government of *Uttarakhand*; 2023.
9. Dobhal DP, Mehta M, Srivastava D. Influence of terrain morphology and aspect on the retreat of *Dokriani Glacier*, Garhwal Himalaya. J Earth Syst Sci. 2013;122(1):133-144.
10. Everingham Y, Sexton J, *Skocaj* D, Inman-Bamber G. Accurate prediction of sugarcane yield using a random forest algorithm. Agron Sustain Dev. 2016;36(2):27.
11. FAO. World agriculture: Towards 2015/2030. An FAO perspective. Food and Agriculture Organization of the United Nations; 2003. Available from: https://www.fao.org/3/y4252e/y4252e.pdf
12. FAO. How to feed the world in 2050. Food and Agriculture Organization of the United Nations; 2009. Available from: https://www.fao.org/fileadmin/templates/wsfs/docs/expert_paper/How_to_Feed_the_World_in_2050.pdf
13. FAO. Save and grow: A policymaker's guide to the sustainable intensification of smallholder crop production.

Food and Agriculture Organization of the United Nations; 2011. Available from: https://www.fao.org/3/i2215e/i2215e.pdf
14. Friedman JH. Greedy function approximation: A gradient boosting machine. Ann Stat. 2001;29(5):1189-1232. DOI:10.1214/aos/1013203451
15. GADM. Global Administrative Areas database (Version 4.1). University of California, Davis; 2023. Available from: https://gadm.org
16. Glahn HR, Lowry DA. The use of model output statistics (MOS) in objective weather forecasting. J Appl Meteorol. 1972;11(8):1203-1211. DOI:10.1175/1520-0450(1972)011<1203:TUOMOS>2.0.CO;2
17. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. 2nd ed. Springer; 2009.
18. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735-1780. DOI:10.1162/neco.1997.9.8.1735
19. ICAR-VPKAS. Annual Report 2019-20. Vivekananda Parvatiya Krishi Anusandhan Sansthan, Almora; 2020.
20. Jeong JH, Resop JP, Mueller ND, *et al*. Random forests for global and regional crop yield predictions. PLOS ONE. 2016;11(6):e0156571.
21. Karimi M, Shahedi K, Azizi G. Application of support vector machine and neural network for wheat yield prediction in *Iran*. Agric Eng Int: CIGR J. 2008;10:1-15.
22. Kaul M, Hill RL, Walthall C. Artificial neural networks for corn and soybean yield prediction. Agric Syst. 2005;85(1):1-18.
23. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444. DOI:10.1038/nature14539
24. Loague K, Green RE. Statistical and graphical methods for evaluating solute transport models. J Contam Hydrol. 1991;7(1-2):51-73.
25. Lobell DB, Gourdji SM. The influence of climate change on global crop productivity. Plant Physiol. 2012;160(4):1686-1697. DOI:10.1104/pp.112.208298
26. Ma Y, Zhang Z, Kang Y, Ozdogan M. Corn yield prediction and uncertainty analysis based on remotely sensed variables using a Bayesian neural network approach. Remote Sens Environ. 2021;259:112408.
27. McMaster GS, Wilhelm WW. Growing degree-days: One equation, two interpretations. Agric For Meteorol. 1997;87(4):291-300.
28. Meena VS. The use of random forests for wheat yield prediction and its relationship with weather variables. J Agron Crop Sci. 2021;207(2):243-257.
29. Ministry of Agriculture and Farmers' Welfare. Annual report 2021-22. Government of India; 2022. Available from: https://agricoop.nic.in/en/annual-report-2021-22
30. Myneni RB, Hall FG, Sellers PJ, Marshak AL. The interpretation of spectral vegetation indexes. IEEE Trans Geosci Remote Sens. 1995;33(2):481-486. DOI:10.1109/36.377948
31. Peng B, Guan K, Tang J, *et al*. Towards a multiscale crop modelling framework for climate change adaptation assessment. Nat Plants. 2014;6:338-348.
32. QGIS Development Team. QGIS Geographic Information System (Version 3.34). Open Source Geospatial Foundation Project; 2023. Available from: https://qgis.org
33. *Rawat* L, *Rawat* S. Characterization of *agroclimatic* zones of *Uttarakhand* for crop planning. J Agrometeorol. 2019;21(2):220-226.
34. Ray DK, Mueller ND, West PC, Foley JA. Yield trends are insufficient to double global crop production by 2050. PLOS ONE. 2013;8(6):e66428.
35. Roberts DR, *et al*. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography. 2017;40(8):913-929.
36. Rosenzweig C, Elliott J, Deryng D, *et al*. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. Proc Natl Acad Sci USA. 2014;111(9):3268-3273. DOI:10.1073/pnas.1222463110
37. Schölkopf B, Smola AJ. Learning with kernels: Support vector machines, regularization, optimization, and beyond. MIT Press; 2002.
38. Shahhosseini M, Martinez-Feria RA, Hu G, Archontoulis SV. Maize yield and nitrate loss prediction with machine learning algorithms. Environ Res Lett. 2021;14(12):124026.
39. Singh A, Ganapathy Subramanian B, Singh AK, Sarkar S. Machine learning for high-throughput stress phenotyping in plants. Trends Plant Sci. 2014;21(2):110-124. DOI:10.1016/j.tplants.2015.10.015
40. Stackhouse PW Jr, Zhang T, Westberg D, *et al*. POWER release 8 (with GIS applications): NASA Surface Meteorology and Solar Energy Datasets. NASA Technical Reports Server; 2019. Available from: https://power.larc.nasa.gov
41. Tilman D, Balzer C, Hill J, Befort BL. Global food demand and the sustainable intensification of agriculture. Proc Natl Acad Sci USA. 2011;108(50):20260-20264. DOI:10.1073/pnas.1116437108
42. Van Klompenburg T, Kassahun A, Catal C. Crop yield prediction using machine learning: A systematic literature review. Comput Electron Agric. 2020;177:105709.
43. Vapnik VN. The nature of statistical learning theory. Springer; 1995. DOI:10.1007/978-1-4757-2440-0
44. Wang X, Huang J, Feng Q, Yin D. Winter wheat yield prediction at county level and uncertainty analysis in main wheat-producing regions of China with deep learning approaches. Remote Sens. 2020;12(11):1744.
45. Willmott CJ, Matsuura K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim Res. 2005;30(1):79-82.
46. You J, Li X, Low M, Lobell D, Ermon S. Deep Gaussian process for crop yield prediction based on remote sensing data. Proc AAAI Conf Artif Intell. 2017;31(1).
47. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc B. 2005;67(2):301-320.