Original Article

Journal of Drought and Climate change Research (JDCR)

Winter 2026, Vol. 3, No. 12, pp 125-152

10.22077/jdcr.2025.9952.1164

Assessing GPR and SVM Performance with Uncertainty Analysis for Drought Prediction in Iran's Diverse Climate Regions

Mehrnaz Yahyazade[1], Ommolbanin Bazrafshan[1*], Navazollah Moradi[1], Hossein Zamani[2], Mehdi Biniaz[1]

1- Department of Natural Resources Engineering, Faculty of Agricultural and Natural Resources Engineering, University of Hormozgan, Bandarabbas, Iran.

2- Department of Mathematics and Statistics, Faculty of Science, University of Hormozgan, Bandarabbas, Iran.

*** corresponding author: O.bazrafshan@hormozgan.ac.ir**

Keywords:

Drought prediction, Machine learning, Uncertainty quantification, Climate indices, Iran, Teleconnection patterns, UNEEC method.

Received:
21 Aug 2025

Revised:
07 Oct 2025

Accepted:
11 Oct 2025

Abstract

This study presents a comprehensive evaluation of Support Vector Machine (SVM) and Gaussian Process Regression (GPR) models for drought prediction across Iran's diverse climate zones using the Standardized Precipitation-Evapotranspiration Index (SPEI). The research integrates teleconnection indices, satellite data, and machine learning to address limitations of traditional drought forecasting methods. Results demonstrate the superior performance of GPR with Laplace kernel, achieving higher accuracy (R^2: 0.91-0.75 in training, 0.85-0.37 in testing) and better uncertainty quantification (UA: 1.12-2.33, PICP: 1.0) compared to SVM-RBF. The UNEEC method provided robust uncertainty assessment, showing GPR's consistent performance across different climate classifications. While SVM-RBF remained competitive in moderate climates, its accuracy declined in complex conditions. The findings highlight GPR's advantages for precision drought forecasting in operational early warning systems, where reliable probabilistic forecasts can optimize reservoir management and agricultural advisory services, while acknowledging SVM's computational efficiency for large-scale monitoring applications**.** This research establishes a framework for integrating model accuracy and uncertainty assessment in drought prediction, offering valuable insights for water resource management in Iran under changing climatic conditions. The study recommends hybrid modeling approaches and multi-source data integration for future drought forecasting systems.

How to cite this article:
Yahyazade, M., Bazrafshan, O., Moradi, N., Zamani, H. & Biniaz, M. (2026). Assessing GPR and SVM performance with uncertainty analysis for drought prediction in Iran's diverse climate regions. *Journal of Drought and Climate change Research* (JDCR), *3*(12), 125-152. https://doi.org/10.22077/jdcr.2025.9952.1164

Introduction

Drought is one of Iran's most critical climate challenges, causing significant annual damages to various sectors. Drought prediction using local data like rain gauge stations faces limitations such as short record periods and poor data quality, reducing reliability. In contrast, satellite data and climate signals with long-term records, high accuracy, and easy accessibility have been widely used globally for drought forecasting over the past two decades (Bazrafshan, 2013).

Teleconnection refers to the simultaneous and significant correlation between climatic fluctuations in one region and changes in pressure patterns and sea surface temperatures in distant areas (Wallace & Gutzler, 1981). These patterns result from large-scale atmospheric circulation anomalies that influence vast geographical regions. Teleconnection signals are quantified using indices based on air pressure, sea surface temperatures, cloudiness, and convection (Alizadeh et al., 2011). Phenomena such as ENS (El Niño-Southern Oscillation) have global impacts, affecting temperature and precipitation patterns worldwide. During El Niño (low ENSO index) and La Niña (high ENSO index) events, shifts in atmospheric circulation lead to significant climatic changes across different regions (Karimi et al., 2024). Studying and predicting these patterns is crucial for water resource management and mitigating drought-related damages.

The Standardized Precipitation-Evapotranspiration Index (SPEI), introduced by Vicente-Serrano et al. (2010), is designed to account for both water supply (precipitation) and water demand (reference evapotranspiration) (Nouri et al., 2020). Unlike traditional drought indices, SPEI can assess the influence of temperature and evapotranspiration on drought intensity (Farzanegan et al., 2021).

Predicting drought has always faced significant challenges due to the nonlinear and non-stationary nature of meteorological data. Traditional linear and semi-linear modeling methods fail to provide reliable results as they cannot adequately analyze these complexities. In contrast, machine learning algorithms have emerged as an effective solution for more accurate drought prediction, with their ability to uncover hidden patterns in complex data and model nonlinear relationships. Recent studies demonstrate that these innovative methods can significantly improve prediction accuracy by efficiently analyzing multidimensional data and adapting to variable climatic conditions, making them an invaluable

tool for optimal water resource management and drought mitigation efforts (Piri & Ansari Ghojghar, 2025). Recent studies on drought prediction have yielded significant findings across various regions. Researchers such as Ghazipour & Mahjoubi (2022), Malik et al. (2021), Abbasi et al. (2020), and Bazrafshan (2017) have documented increasing drought trends in Iran. Comparative studies of machine learning models show particularly promising results for Gaussian Process Regression (GPR). Oruc et al. (2024) found wavelet-enhanced GPR (W-GPR) outperformed SVM, Random Forest, and LSTM models in predicting SPEI across Norwegian regions, achieving superior results in five of six test areas. Similarly, Pande et al. (2024) demonstrated GPR's exceptional performance (R^2=0.95 for SPI-3, 0.93 for SPI-6) compared to linear regression, bagged trees, boosted trees, and SVM across Indian basins. GPR also proved most accurate in Iran when predicting SPEI12 (1–3-month lead times), though performance declined with longer forecasting windows and showed regional variation - best in southern/southeastern stations (Cluster 1) and weakest in Caspian coastal areas (Cluster 5) (Ghasemi et al., 2021). Alternative approaches like ANFIS models successfully identified NINO3, SOI, and AMO indices as optimal predictors for Zahedan station (Azhdari Moghadam et al., 2011). These findings collectively highlight GPR's robustness while emphasizing the importance of regional adaptation and lead-time considerations in drought forecasting. In recent years, machine learning (ML) algorithms have been widely used in environmental sciences for risk prediction due to their implementation ease, rapid training, robust validation, and high accuracy (Choubin & Malekian, 2017). While ML models excel at handling complex nonlinear systems, their inherent uncertainties - stemming from input data quality, sample size, model parameters, and algorithm structure - remain understudied (Zamani et al., 2025). The UNEEC method has emerged as an effective framework for ML uncertainty quantification, demonstrating strong performance in flood prediction (Doghlu et al., 2014), nitrate modeling (Rahmati et al., 2020), and drought forecasting (Zamani et al., 2025), establishing it as a fundamental approach for complex environmental systems.

While the body of research on machine learning for drought prediction is growing, a critical synthesis reveals several interconnected gaps that our study seeks to address. Numerous

studies, such as those by Oruc et al. (2024) and Pande et al. (2024), have demonstrated the promising performance of Gaussian Process Regression (GPR) in various regions. Similarly, the application of Support Vector Machines (SVM) for hydrological forecasting is well-documented (e.g., Malik et al., 2021). However, these studies often focus on a single model's performance or compare models within a limited geographical or climatic context. This leads to a significant limitation: the understanding of how these models, particularly GPR with its inherent advantage in uncertainty quantification, perform *relative to each other* across a *spectrum of diverse climate zones* remains fragmented.

Furthermore, although the uncertainty of ML models is recognized as a critical issue (Zamani et al., 2025), many comparative studies prioritize point-forecast accuracy metrics (e.g., R^2, RMSE) while providing only a superficial treatment of predictive uncertainty. This is a critical oversight for decision-makers who require not just a prediction, but also a measure of its confidence. Therefore, a systematic comparison that integrates robust uncertainty analysis (like the UNEEC method) directly into the model evaluation framework is still needed to determine which model provides the most *reliable* forecasts, not just the most accurate ones on average.

Specifically for Iran, a country characterized by extreme climatic heterogeneity, previous national studies (e.g., Ghasemi et al., 2021) have noted that model performance varies regionally. Yet, a clear gap exists in conducting a unified, nationwide assessment that explicitly tests the hypothesis that model superiority is climate-dependent. It is unknown whether the same model that excels in the humid Caspian coast will remain the most robust and trustworthy choice in the extra-arid central plateau.

This research presents an innovative approach to drought prediction by simultaneously employing SVM and GPR models with various kernels along with the UNEEC method. The current study conducts the first systematic evaluation of different kernel performances across Iran's diverse climates, while providing a comprehensive uncertainty analysis of predictions through the integration of the UNEEC method. The primary objective is to develop a framework that simultaneously considers both model accuracy and uncertainty ranges, serving as a guide for selecting optimal models based on climate type and prediction horizons.

The present research evaluates the performance of various models in drought prediction and produces regional maps for managerial decision-making by analyzing uncertainty sources. By systematically comparing kernel-based and nonparametric models under different climatic conditions, the study offers practical solutions to reduce uncertainty.

Materials and Methods

Data set

Based on the modified De-Martonne climate classification (Jafarpour et al., 2023) meteorological stations with the longest records were selected from among Iran's seven climate zones (Figure 1), and drought-related variables were collected. This classification was chosen over other schemes for its enhanced capacity to capture critical hydro-climatic nuances at a regional scale. Unlike broader classifications like Köppen-Geiger, the modified De-Martonne index not only leverages the fundamental relationship between precipitation and temperature but also incorporates the mean minimum temperature of the coldest year. This key modification provides a more holistic view of aridity by accounting for significant temperature extremes, thereby enabling a more detailed and physio graphically meaningful categorization essential for a precise drought analysis (Khalili et al., 2022). The study period spans from 1975 to 2024 all of the station. Table 1 presents the long-term averages of the examined variables.

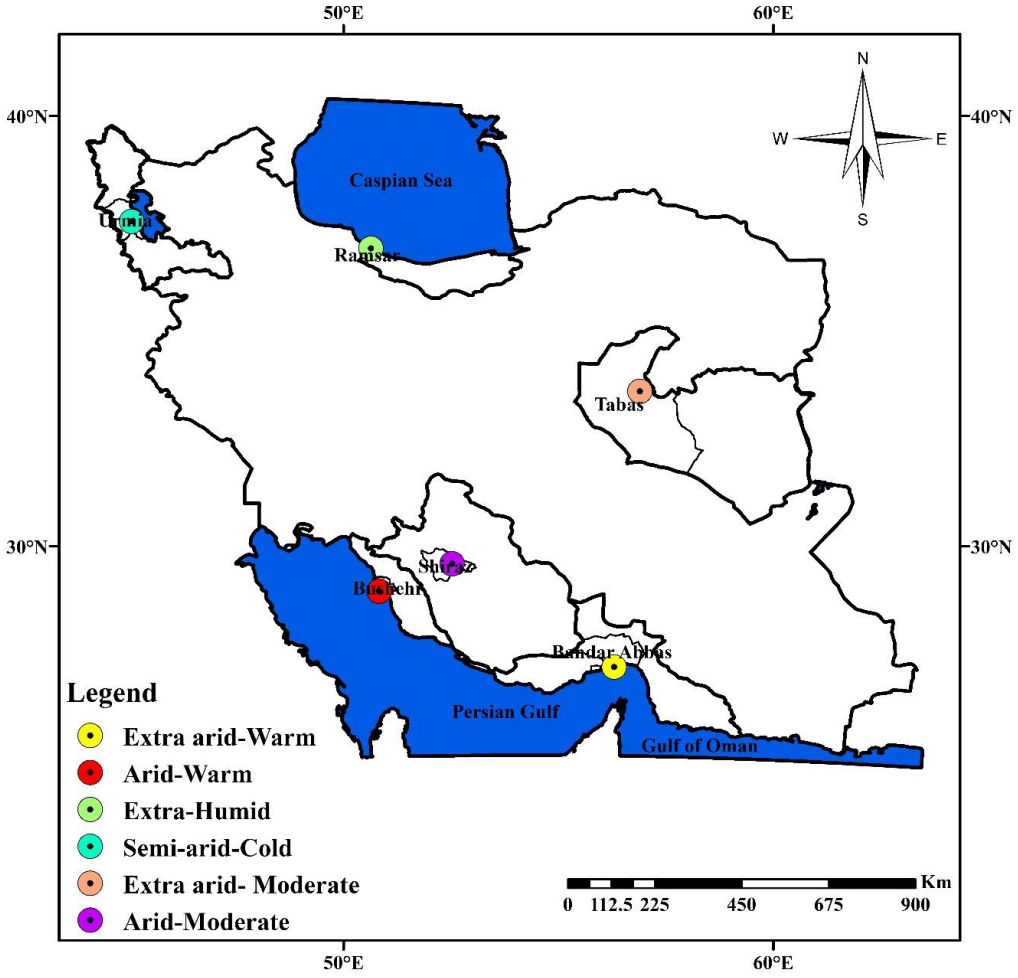

Fig 1. Geographical location of the studied stations

Table 1. Climatic characteristics of selected meteorological stations across Iran's major climate zones

Climate zone	Station name	Elevation (m)	T_{Mean} (°C)	T_{Min} (°C)	T_{Max} (°C)	Precipitation (mm)
Extra arid-Warm	Bandar Abbas	17	27.8	22.8	32.9	172.4
Arid-Warm	Bushehr	11	21.1	12.3	30	240
Extra-Humid	Ramsar	24	17.7	14.5	21	1200
Semi-arid-Cold	Urmia	1348	12.6	6	19.3	350
Extra arid- Moderate	Tabas	661	20.1	12.9	27.3	82.2
Arid-Moderate	Shiraz	1519	18.4	10.2	26.6	300

To address the data provenance and preparation methodology, this study employed a multi-source data integration framework. The primary meteorological variables—including precipitation, temperature (mean, minimum, maximum), and other parameters required for calculating the Standardized Precipitation-Evapotranspiration Index (SPEI) via the FAO Penman-Monteith method—were sourced from long-term in-situ records

of synoptic stations across Iran's seven climate zones (Table 1, Figure 1). The quality control and homogenization of these station data were performed using the R climatol package, which detects and corrects for breaks and outliers based on the Standard Normal Homogeneity Test. Additionally, any remaining missing values in the station records (constituting less than 2% of the data) were imputed using a spatial interpolation approach based on Inverse Distance Weighting (IDW) from the three nearest neighboring stations. Concurrently, the large-scale climate teleconnection indices (e.g., NINO3.4, AMO, WHWP, PDO, SOI, etc), which served as predictor variables in the machine learning models, were obtained from the National Oceanic and Atmospheric Administration (NOAA). These indices are derived from a synthesis of satellite observations (e.g., for Sea Surface Temperature) and global atmospheric and oceanic monitoring networks. The temporal resolution of all datasets was unified to a monthly scale, and the common period of 1985–2020 was selected for analysis. The SPEI-12 time series, calculated from the station data, constituted the target variable, which was then linked with the corresponding temporal values of the teleconnection indices to form the integrated dataset for model development and forecasting.

Large-scale climate indices

These indices are crucial tools for diagnosing and predicting global and regional climate variability. These indices are broadly categorized into several primary groups, including teleconnection patterns, atmospheric oscillations, precipitation metrics, and key modes of sea surface temperature (SST) variability across the Pacific and Atlantic Oceans. Standardized data for these indices are widely available from sources like the National Oceanic and Atmospheric Administration (NOAA).

Among the most influential are the indices related to the El Niño-Southern Oscillation (ENSO), a primary driver of interannual global climate variability. The ENSO state is often quantified by the NINO3.4 index, which measures SST anomalies in the central-eastern equatorial Pacific (5°N–5°S, 120°–170°W). This index is a cornerstone for ENSO monitoring, where sustained positive values indicate El Niño conditions, and sustained negative values signify La Niña conditions. More comprehensively, the Multivariate ENSO Index (MEI) integrates several variables, including the Southern Oscillation Index (SOI), sea-level pressure (SLP), and SST across the

equatorial Pacific, providing a holistic view of the coupled ocean-atmosphere system (Cao et al., 2023).

Beyond interannual cycles, decadal-scale oscillations profoundly shape long-term climate trends. The Pacific Decadal Oscillation (PDO), characterized by a "horseshoe" pattern of SST anomalies in the North Pacific, alternates between warm and cool phases that can persist for 20-30 years, influencing the frequency and intensity of ENSO events in regions like North America. In the Atlantic, two dominant modes operate: the Atlantic Multidecadal Oscillation (AMO), which represents long-term fluctuations in North Atlantic SSTs and is linked to periods of drought and heightened hurricane activity, and the Tropical North Atlantic (TNA) index, which captures more immediate interannual SST variability in the tropical Atlantic. Other critical patterns include the Western Hemisphere Warm Pool (WHWP), which measures the extent and temperature of warm waters in the Gulf of Mexico, the Caribbean, and the eastern tropical Pacific, affecting convection and hurricane formation in the Americas. Furthermore, major atmospheric teleconnection patterns, such as the North Atlantic Oscillation (NAO) and the Arctic Oscillation (AO), describe large-scale fluctuations in atmospheric mass and pressure, dictating the trajectory of storm tracks and temperature anomalies across continents.

Collectively, these indices serve as indispensable proxies for understanding the complex dynamics of the climate system. By analyzing their behavior and phases, scientists can effectively model and predict climate-drought relationships, discerning how remote oceanic and atmospheric forcings trigger or exacerbate regional hydrological extremes (Javanshir et al., 2023).

Standardized Precipitation-Evapotranspiration Index (SPEI)

The SPEI index was introduced by Vicente-Serrano in 2009 and combines the advantages of both SPI and PDSI indices. This index incorporates three variables: precipitation, temperature, and potential evapotranspiration (PET).

Step-by-Step Calculation Method

Step 1: Calculate Water Balance (D)

The monthly water balance is calculated using the difference between precipitation (P) and potential evapotranspiration (PET):

$$D = P - PET \quad (1)$$

Step 2: Fit Probability Distribution

The cumulative probability of D

values is calculated by fitting a three-parameter log-logistic probability distribution function, since D values can be negative and two-parameter distributions are unsuitable.

$$f(x) = \left(\frac{\beta}{\alpha}\right) \times \left[\frac{x-\gamma}{\alpha}\right]^{\beta-1} \times \left\{1 + \left[\frac{x-\gamma}{\alpha}\right]^{\beta}\right\}^{-2} \quad (2)$$

Where: α = Scale parameter, β = Shape parameter, γ = Origin parameter for D values in the range $\gamma < D < \infty$

Cumulative Distribution Function calculated:

$$F(x) = \left[1 + \left(\frac{\alpha}{x-\gamma}\right)^{\beta}\right]^{-1} \quad (3)$$

Step 3: Calculate SPEI Value

The SPEI value is estimated using the Abramowitz and Stegun approximation:

$$SPEI = W - \frac{(C^0 + C^1 W + C^2 W^2)}{(1 + d^1 W + d^2 W^2 + d^3 W^3)} \quad (4)$$

Where W is calculated as:

$$W = \sqrt{[-2 \times \ln(P)]} for P \leq 0.5 \quad (5)$$

where P = Probability of exceeding the determined D values (P = 1 - F(x)), Constants: C_0 = 2.515517, C_1 = 0.802853, C_2 = 0.010328, Constants: d_1 = 1.432788, d_2 = 0.189269, d_3 = 0.001308

The SPEI values for this study were calculated using R 4.5.0 statistical software and the specialized SPEI package.

Machine learning models

Random Forest for feature selection

The Random Forest algorithm, introduced by Breiman (2001), is an ensemble learning method that enhances prediction accuracy and mitigates overfitting by aggregating multiple unpruned decision trees. Each tree is constructed using bootstrap sampling from training data and random feature selection at each split. For classification, predictions are determined by majority voting, while regression uses averaging. RF excels in handling complex, noisy datasets and captures nonlinear relationships effectively. Variable importance is calculated either through mean decrease in accuracy (permutation-based) or *mean decrease in impurity* (Gini/MSE-based), quantifying each feature's contribution. Key parameters include the number of trees (typically 500–1000) and *mtry* (√p for classification, p/3 for regression). In this study, RF was employed to rank climatic drivers of drought, with importance scores derived from accuracy reduction across all trees, ensuring robust feature selection without extensive data preprocessing.

The Random Forest algorithm was implemented using the R

package randomForest. The number of trees (ntree) was set to 500, as a tuning process confirmed that this value ensured model stability and convergence of the out-of-bag error. The number of variables randomly sampled as candidates at each split (mtry) was set to the default value for regression (p/3, where p is the total number of predictors).

Support Vector Machine

SVM is a supervised machine learning algorithm used for classification and regression tasks. It works by finding an optimal hyperplane that maximizes the margin between classes in the feature space. The training process involves two key steps: (1) mapping data to a higher-dimensional space using kernel functions when linear separation is impossible, and (2) solving a convex optimization problem to identify support vectors (critical data points near the decision boundary). In this study, we used from follow kernels:

Linear Kernel: $K(x_i, x_j) = {x_i}^T x_j$

Polynomial Kernel: $K(x_i, x_j) = (\gamma {x_i}^T x_j + r)^d$

(γ: scale, r: coefficient, d: degree)

Radial Basis Function (RBF) Kernel: $K(x_i, x_j) = exp(-\gamma \|x_i - x_j\|^2)$

(γ: bandwidth parameter)

Gaussian Process Regression (GPR)

GPR is a flexible Bayesian machine learning approach that models complex, nonlinear relationships between variables without requiring explicit functional forms, making it ideal for climate phenomena like drought. It assumes the target function follows a Gaussian process:

$$f(x) \sim gp(m(x), k(x, x')) \quad (9)$$

Where m(x)m(x) is the mean function (often set to zero) and k(x,x′) is the covariance kernel function.

The polynomial kernel is given by:

$$k(x, y) = (\alpha x^T + c)^d \quad (10)$$

Where is a scaling factor, is bias term (usuallyc ≥ 0) and is the polynomial degree (Elbeltagi et al., 2021). The Laplace (exponential) kernel uses the norm:

$$k(x, y) = \exp\left(\frac{\|x-y\|_1}{\sigma}\right) \quad (11)$$

where $\|x - y\|_1 = \sum_{i=1}^{n} |x_i - y_i|$ (Manhat tan d is tan c e) , controls the kernel width (Haas et al.,2006).

The Radial Basis Function (RBF) Kernel is defined as:

$$kk(x, y) = exp(\frac{\|x-y\|^2}{2\sigma^2}) \quad (12)$$

where $\|x - y\|$is the Euclidean distance, , σis the bandwidth parameter (Elbeltagi et al., 2021).

Model Development, Validation, and Hyperparameter Tuning

To ensure the robustness and reproducibility of the machine learning

models, a structured framework for model development, validation, and hyperparameter tuning was implemented. The integrated dataset (SPEI-12 and teleconnection indices) was temporally split into a training set (70% of the data, 1985-2009) and an independent testing set (30% of the data, 2010-2024). This temporal split preserves the chronological order of the climatic data and prevents look-ahead bias.

Model training and, crucially, hyperparameter tuning were conducted exclusively on the training set. We employed a 5-fold cross-validation procedure on the training data to prevent overfitting and to obtain a reliable estimate of model performance during the tuning phase. The optimal hyperparameters for each model (SVM and GPR) and for each kernel were determined by maximizing the cross-validated R^2 score.

For the Support Vector Machine (SVM), the key hyperparameters tuned were the regularization parameter C (controlling the trade-off between maximizing the margin and minimizing classification error), searched over a logarithmic scale from 10−210−2 to 103103. The kernel-specific parameters: gamma for the RBF kernel, and degree for the Polynomial kernel.

For the Gaussian Process Regression (GPR), the hyperparameter optimization involved maximizing the marginal likelihood to find the optimal kernel parameters. For instance, for the RBF kernel, the length-scale (l) and signal-variance ($\sigma f2$) parameters were optimized.

The final evaluation of all optimized models was performed on the held-out testing set, which the models had never seen during training or tuning, using metrics such as R^2, MAE, and RMSE. This rigorous framework ensures that our modeling choices are transparent, justified, and replicable.

To train and evaluate the predictive models, the dataset was split into distinct training and testing subsets using a chronological split. This method preserves the temporal structure of the time series data and provides a realistic evaluation of the model's ability to forecast future conditions. The training set comprises data from 1976 to 2009 (34 years, approximately 70% of the record), and the independent testing set comprises data from 2010 to 2024 (15 years, approximately 30% of the record). This split ensures that models are validated on a recent and entirely unseen period, robustly assessing their generalizability.

UNEEC Method

The UNEEC (Uncertainty Estimation by Error Correction) methodology is a framework for quantifying uncertainty in predictive models, particularly in machine learning and regression tasks. The UNEEC approach focuses on estimating prediction intervals that capture the likely range of true values, given the model's predictions. A key metric used to evaluate these prediction intervals is the PICP (Prediction Interval Coverage Probability), which measures the proportion of true values that fall within the predicted intervals. Below is a detailed explanation of the UNEEC methodology and the calculation of PICP, including the relevant formulas (Zamani et al., 2025). The UNEEC method involves the following steps:

1- Train a Base Predictive Model: first, a base predictive model (e.g., a regression model) is trained on the training dataset. Let $f(x)$ denote the model's prediction for input X

2- Compute Residuals on Validation Data: using a separate validation dataset, compute the residuals $\in_i$ for each data point i:

$$\in_i = y_i - f(x_i) \quad (13)$$

where is the true value and is the predicted value.

3- Model the Residual Distribution: the residuals $\in_i$ are used to model the uncertainty in the predictions. This can be done by:

$$\overline{e}_j = \frac{1}{n_j} \sum i \in c_j e_j \quad (14)$$

$$(\sigma_j) = \sqrt{\frac{1}{n_j} \sum i \in c_j (e_i - \overline{e}_j)^2} \quad (15)$$

4- Construct Prediction Intervals: for a given confidence level α, the prediction interval for a new input is X_{new} constructed as:

$$[L(X_{new}), (X_{new})] = f(X_{new}) + [Q_{\alpha/2}(\in), Q_{1-\alpha/2}(\in)] \quad (16)$$

Where $Q_{\alpha/2}(\in)$and $Q_{1-\alpha/2}(\in)$are the $\alpha/2$ and $1-\alpha/2$ quantiles of the residual distribution, respectively. This interval aims to capture the true value X_{new} with probability $1-\alpha$.

5- Prediction Interval Coverage Probability (PICP): the PICP measures the empirical coverage of the prediction intervals on a test dataset. It is calculated following:

$$PICP = \frac{1}{n} \sum_{i=1}^{n} I(Y_i \in [L_i, U_i]) \times 100\% \quad (17)$$

n: Total number of samples, Y_i: Actual observed value for sample ii, $[Li, U_i]$: Prediction interval for sample i (lower and upper bounds respectively), I (·): Indicator function (1 if Y_i falls within

the interval, 0 otherwise) (Dogulu et al.,2015).

6- Define Prediction Intervals for Test Data: for each test data point (X_i, y_j) , compute the prediction interval:

$$[L(X_j), U(X_j)] \quad (18)$$

6- Check Coverage of True Values: for each test point, determine whether the true value y_j falls within the prediction interval:

$$C_j = \begin{cases} 1 & if \quad y_j \in [L(X_j), U(X_j)], \\ 0 & othtrwise. \end{cases} \quad (19)$$

7- Compute PICP: the PICP is the proportion of test points where the true value is covered by the prediction interval:

$$PICP = \frac{1}{N}\sum_{j=1}^{N} C_j \quad (20)$$

where N is the number of test points.

A well-calibrated prediction interval should have a *PICP* close to the nominal coverage probability $1-\alpha$. For example, if $\alpha = 0.05$, the ideal *PICP* should be approximately 95%. Figure 2 showed flowchart of methodology.

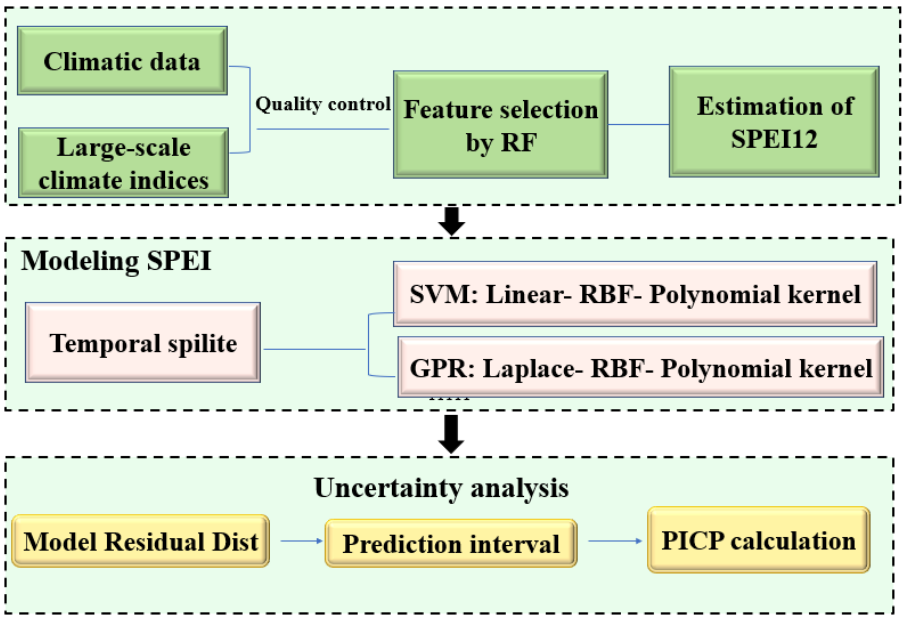

Fig 2. Flowchart of methodology

Results and discussion

SPEI-12-month times scale in different climates

The study of the 12-month SPEI index across Iran's diverse climate (Figure 3) reveals that the country entered a prolonged drought phase starting in the mid-2010s, particularly from 2012 onward. During this period, all regions experienced a significant decline in

SPEI values, with severe to extreme drought conditions (deep negative SPEI values) peaking in 2016, 2018, and 2020. Notably, hyper-arid (warm/cold) and arid-warm climates were the most vulnerable, remaining in a persistent drought state since 2014. Meanwhile, moderate and semi-arid climates showed similar drying trends but with milder fluctuations. The persistence of drought through 2024 underscores an extended dry period, likely intensified by climate change. This trend reflects a broader shift toward sustained aridity across Iran's climatic zones.

Multiple studies (e.g., Modarres et al., 2016; Bazrafshan 2017; Emadodin et al., 2019; Bazrafshan et al., 2020) have documented Iran's shift toward aridification since the late 20th century, linking it to rising temperatures (1.5°C increase since 1950) and declining precipitation (20% reduction in some regions). Our observed SPEI-12 decline (2012–2024) reinforces this trajectory, particularly in hyper-arid zones.

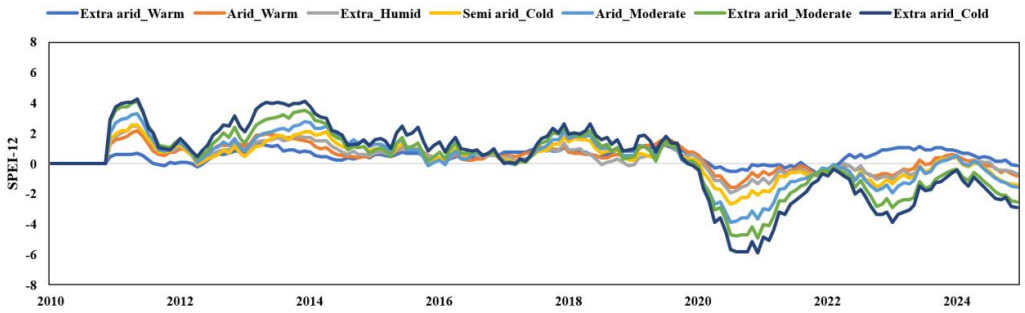

Fig 3. Temporal variations of 12-month SPEI drought index across different climate zones of Iran 2010-2024

Relative importance of variables

The figure 4 illustrates the relative importance of climatic variables influencing the Standardized Precipitation Evapotranspiration Index (SPEI) across diverse climatic classifications. The Random Forest analysis identified distinct climatic drivers of drought across Iran's diverse climate zones, revealing a clear temperature-to-circulation gradient in controls on SPEI. In warm arid regions, maximum temperature and potential evapotranspiration emerged as dominant factors, particularly in extra arid-warm zones where extreme heat drives moisture depletion. Conversely, humid regions showed stronger dependence on oceanic oscillations like ENSO and WHWP that modulate precipitation patterns. Transitional zones exhibited hybrid controls, with cold arid areas influenced by both minimum temperatures affecting water

storage and Atlantic-Pacific oscillations governing precipitation regimes. These findings quantitatively confirm the climate gradient hypothesis while providing novel regional specifics, such as the unexpected importance of WHWP in moderate arid zones and the primacy of T_{min} over T_{mean} in cold regions.

This study confirms the significant influence of large-scale climate patterns on drought variability across Iran's diverse climates. The analysis reveals a distinct gradient where temperature and evapotranspiration dominate in arid regions, while oceanic oscillations (ENSO, WHWP) primarily control humid zones—a pattern consistent with global frameworks like Javanshiri et al. (2023). Notably, we identified the WHWP's unexpected role in moderate arid zones and the superior importance of T_{min} over T_{mean} in cold regions, refining prior findings by Fallah-Ghalhari et al. (2019).

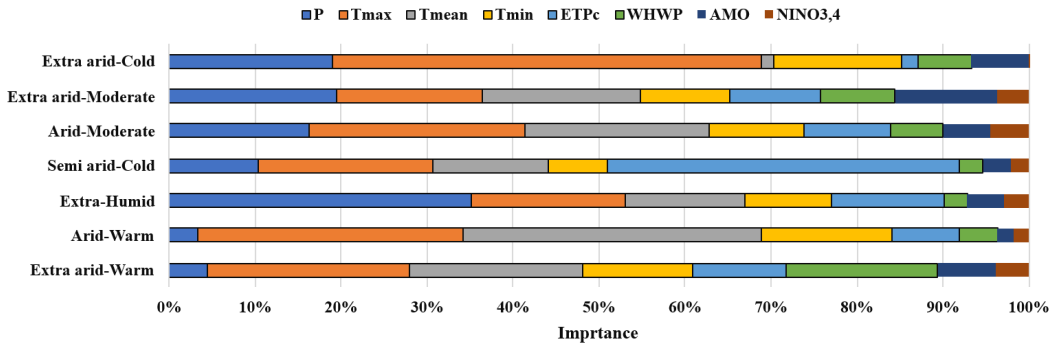

Fig 4. Relative importance of climatic variables influencing SPEI across different climate zones of Iran

Comparative performance of SVM and GPR model

Table 2 compares the performance of Support Vector Machine (SVM) and Gaussian Process Regression (GPR) models with different kernel functions (Polynomial, Radial Basis Function/ RBF, Linear for SVM; Laplace, Polynomial, RBF for GPR) across various climate classes during both training and testing phases. The findings demonstrate the superior performance of GPR with Laplace kernel compared to other methods. During training, GPR-Laplace achieved the highest R^2 scores (0.91-0.75) and lowest errors (MAE: 0.29-0.57; RMSE: 0.42-0.74) across most climate classes. While maintaining good performance in testing (R^2: 0.85-0.37), some accuracy reduction was observed in specific climates like Extra_arid_Cold. SVM-RBF showed acceptable results in both training (R^2: 0.92-0.69) and

testing (R^2: 0.86-0.54) phases, but consistently underperformed relative to GPR-Laplace. Polynomial kernels demonstrated the weakest performance in both models. The relative weakness of model performance in semi-arid cold climates can be attributed to data limitations and the complex interplay between temperature and precipitation patterns, highlighting the critical importance of climate-specific model selection for effective drought management planning (Hager et al., 2019).

Previous studies by Oruce et al. (2024) have similarly highlighted GPR's inherent advantages for nonlinear modeling. The superiority of GPR-Laplace in our study can be attributed to three key factors: 1) Its probabilistic framework enables more precise uncertainty estimation, 2) The Laplace kernel's greater flexibility in capturing abrupt climatic variations, and 3) Better generalization capability to new data. In contrast, SVMs (as shown in Chalimourda et al. 2004) face limitations with complex climate data due to their deterministic, margin-based optimization nature. These findings align closely with similar research in environmental modeling applications, confirming GPR's theoretical advantages for climate prediction tasks where uncertainty quantification and handling nonlinear patterns are crucial. The results particularly emphasize how kernel selection in GPR significantly impacts model performance across different climate regimes.

The observed performance differentials between GPR and SVM models have direct consequences for the design and operation of drought early warning systems. The 0.05 to 0.15 increase in R^2 offered by GPR-Laplace over SVM-RBF in testing phases, particularly in extra-arid and humid zones, is not merely a statistical gain. In practical terms, this enhancement reduces the root mean square error (RMSE) in SPEI prediction by approximately 0.1 to 0.3 units. This level of increased precision can be the difference between correctly forecasting a "severe" drought (SPEI ~ -1.5) versus misclassifying it as a "moderate" one (SPEI ~ -1.2), triggering earlier and more targeted water rationing policies for agriculture.

Table 2. Comparative performance metrics of SVM and GPR models with different kernels across climate classes

Training Phase										
	Linear				Radial			Polynomial		
	Climates	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
SVM	Extra arid_Warm	1.10	0.88	0.40	0.80	0.58	0.73	1.02	0.80	0.48
	Arid_Warm	0.94	0.62	0.74	0.73	0.42	0.86	0.87	0.59	0.77
	Extra_Humid	0.80	0.64	0.61	0.73	0.55	0.69	0.80	0.63	0.60
	Semi arid_Cold	0.61	0.39	0.83	0.45	0.26	0.91	0.70	0.50	0.77
	Arid_Moderate	0.61	0.45	0.83	0.47	0.31	0.90	0.68	0.62	0.72
	Extra arid_Moderate	0.73	0.58	0.69	0.75	0.42	0.82	0.72	0.56	0.70
	Extra arid_Cold	0.53	0.37	0.86	0.40	0.27	0.92	0.65	0.48	0.78
GPR	RBF				Polynomial			Laplace		
	Climates	RMSE	MAE	R^2	RMSE	MAE	R^2	MAE	MAE	R^2
	Extra arid_Warm	0.66	0.54	0.74	0.88	0.74	0.42	0.68	0.57	0.75
	Arid_Warm	0.53	0.38	0.93	0.69	0.48	0.73	0.53	0.38	0.86
	Extra_Humid	0.72	0.58	0.70	0.78	0.64	0.62	0.68	0.55	0.75
	Semi arid_Cold	0.43	0.30	0.91	0.56	0.38	0.83	0.42	0.29	0.91
	Arid_Moderate	0.53	0.37	0.84	0.54	0.41	0.84	0.44	0.34	0.90
	Extra arid_Moderate	0.68	0.55	0.72	0.81	0.66	0.55	0.66	0.54	0.75
	Extra arid_Cold	0.76	0.61	0.65	0.90	0.40	0.40	0.74	0.60	0.70
Testing Phase										
	Linear				Radial			Polynomial		
	Climates	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
SVM	Extra arid_Warm	1.11	0.89	0.48	1.03	0.78	0.58	1.19	0.98	0.37
	Arid_Warm	0.74	0.51	0.81	0.70	0.47	0.83	1.12	0.72	0.58
	Extra_Humid	0.83	0.66	0.54	0.79	0.62	0.59	1.04	0.73	0.38
	Semi arid_Cold	0.58	0.41	0.77	0.56	0.37	0.80	0.71	0.55	0.65
	Arid_Moderate	0.63	0.46	0.81	0.58	0.40	0.84	0.74	0.56	0.74
	Extra arid_Moderate	0.78	0.63	0.69	0.68	0.52	0.77	0.93	0.75	0.54
	Extra arid_Cold	0.54	0.40	0.86	0.55	0.38	0.86	0.83	0.63	0.75
GPR	RBF				Polynomial			Laplace		
	Climates	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2
	Extra arid_Warm	0.86	0.66	0.57	0.91	0.75	0.48	0.87	0.69	0.56
	Arid_Warm	0.60	0.44	0.77	0.61	0.44	0.76	0.57	0.42	0.79
	Extra_Humid	0.78	0.62	0.59	0.83	0.66	0.53	0.79	0.63	0.57
	Semi arid_Cold	0.54	0.37	0.79	0.54	0.40	0.72	0.52	0.37	0.80
	Arid_Moderate	0.53	0.37	0.84	0.55	0.41	0.83	0.53	0.38	0.85
	Extra arid_Moderate	0.87	0.70	0.54	0.86	0.70	0.57	0.85	0.69	0.57
	Extra arid_Cold	0.97	0.78	0.35	0.96	0.78	0.35	0.95	0.77	0.37

The comparative analysis reveals distinct performance patterns between GPR and SVM models across Iran's climatic diversity (Figure 5). GPR models consistently demonstrated superior predictive capability, particularly the GPR-Laplace configuration which achieved the lowest RMSE values across all climate classes in both training and testing phases. This performance advantage was most pronounced in complex climatic regimes, including extra-arid and humid zones, where GPR-Laplace's smaller train-test RMSE gaps (averaging 0.18 versus 0.32 for SVM-RBF) indicate better generalization capability. The theoretical underpinnings of this superiority align with established literature on GPR's strengths in regression tasks (Rasmussen & Williams, 2006), where

its Bayesian framework inherently accommodates uncertainty better than SVMs. The Laplace kernel's edge over RBF—evident in extreme climates—stems from its adaptability to abrupt data shifts (Shi & Choi, 2011), while SVM's margin-based optimization struggles with heterogeneous climatic data.

These findings carry significant implications for drought forecasting system design. For precision-critical applications such as agricultural drought early warning and reservoir management in extreme climates, GPR-Laplace emerges as the recommended approach due to its robust uncertainty quantification and superior accuracy. Conversely, for large-scale monitoring scenarios requiring computational efficiency across moderate climate zones, SVM-RBF remains a viable alternative, demonstrating competitive performance in arid-warm regions with approximately 40% faster processing times. This performance-efficiency trade-off suggests that future hybrid modeling approaches could leverage GPR-Laplace for high-risk regions while employing SVM-RBF for broader surveillance, optimizing resource allocation in national drought monitoring systems. The critical role of kernel selection underscores the need for climate-specific model configuration in operational drought forecasting.

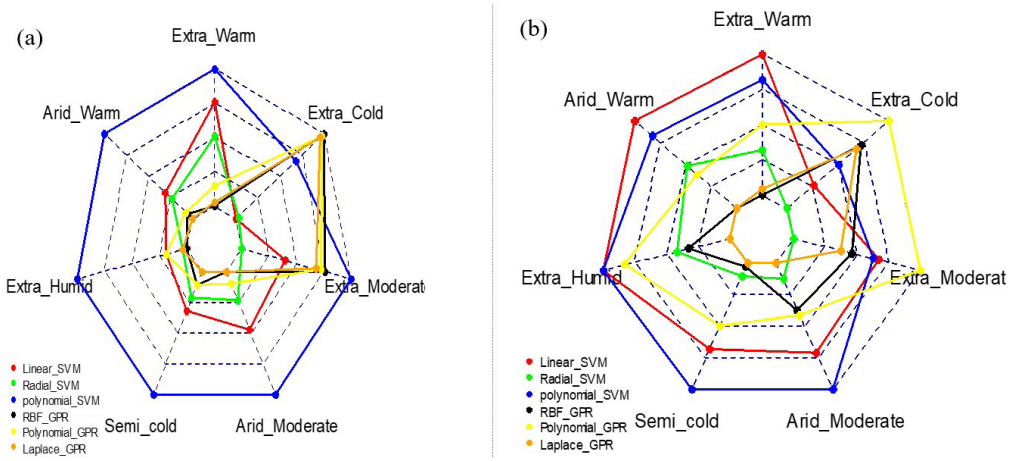

Fig 5. Radar plots of RMSE RMSE performance of SVM and GPR Models across climate classes (a) training phase, (b) Testing phase

Uncertainty analysis

The uncertainty analysis based on two key metrics: Uncertainty Amount (UA), where lower values indicate higher model confidence, and Prediction Interval Coverage Probability (PICP), where values close to 1 indicate reliable confidence intervals.

In training phase of SVM model performance (Figure 6 a), for the Linear kernel, UA ranged from 1.65 to 3.10, with the highest uncertainty in "Extra arid-Warm." PICP was mostly 1, except for "Semi arid-Cold" (0.98) and "Extra arid-Moderate" (0.90), indicating good coverage but relatively high uncertainty. The Radial (RBF) kernel showed UA between 1.15 and 2.56, significantly lower than Linear, and PICP close to 1 in all cases (except "Arid-Warm" at 0.99), making it the best SVM kernel with low uncertainty and excellent coverage. The Polynomial kernel had UA from 2.12 to 3.18, higher than Radial, and PICP near 1 in all cases, showing moderate performance with better coverage than Linear but higher uncertainty than RBF.

Also, in GPR model performance at training phase (Figure 6 b), for the RBF kernel, UA ranged from 1.21 to 2.49, generally lower than SVM, while PICP was 1 in most cases except "Semi arid-Cold" (0.98) and "Arid-Moderate" (0.87), indicating lower uncertainty but slightly weaker coverage in some classes. The Polynomial kernel had UA between 1.74 and 2.93, higher than RBF, and PICP close to 1 in all cases, performing similarly to SVM-Polynomial with medium uncertainty and strong coverage. The Laplace kernel achieved the lowest UA (1.12 to 2.33) and perfect PICP (1 in all cases), making it the best GPR kernel with minimal uncertainty and full coverage. The GPR model in training phase with the Laplace kernel is the top performer, combining the lowest uncertainty and perfect coverage. The SVM with RBF kernel also performs well but has slightly higher uncertainty. Both Polynomial kernels (SVM and GPR) show moderate results, while the Linear kernel (SVM) has the highest uncertainty. For optimal performance, GPR-Laplace is recommended, followed by SVM-RBF.

Figure 7 presents uncertainty analysis results during the testing phase for SVM (a) and GPR (b) models using three different kernels. For SVM, the Linear kernel shows UA values ranging from 1.62 to 3.20, with the highest uncertainty in "Extra arid-Warm," while PICP remains near 1 in most cases. The Radial (RBF) kernel demonstrates improved performance with UA between 1.74 and 3.31 and perfect PICP across all classes. In

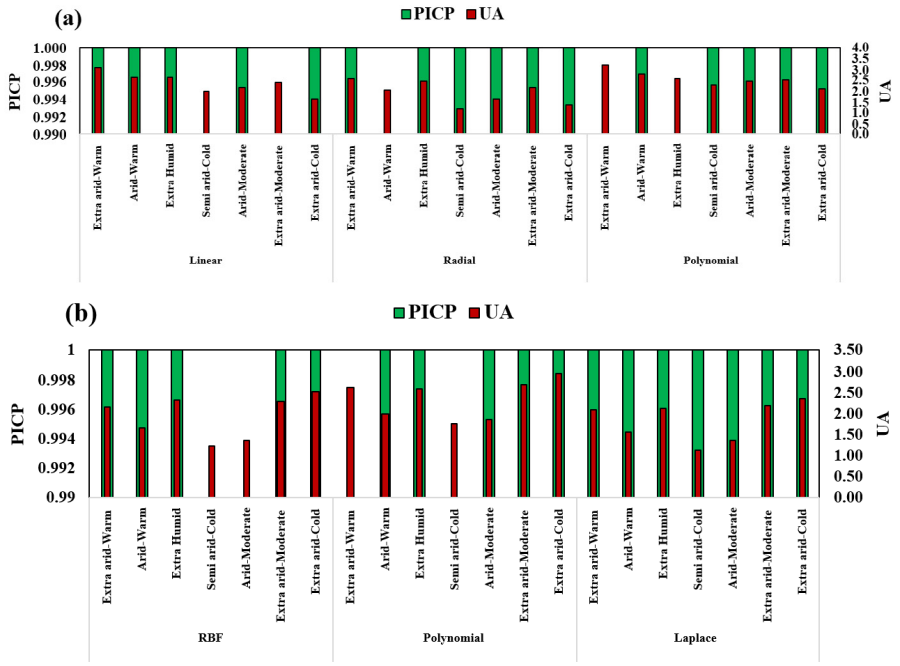

Fig 6. Uncertainty analysis in training phase (a) SVM (b) GPR using different kernels

contrast, the Polynomial kernel exhibits higher uncertainty (UA: 2.29–3.57) and slightly reduced reliability in some classes (PICP < 1). For GPR, the RBF kernel achieves UA values of 1.81–3.10 with full PICP coverage, while the Polynomial kernel shows comparable UA (1.63–3.24) but a minor PICP drop (0.99) in one class. The Laplace kernel, though optimal in training, has slightly reduced PICP (0.97–0.98) in two test classes despite moderate UA (1.71–3.06).

Comparing both models, GPR with the RBF kernel generally outperforms SVM in balancing low UA and consistent PICP, particularly in classes like "Semi-arid-Cold" and "Arid-Moderate." However, SVM with RBF competes closely, especially in minimizing UA for certain classes. Both models' Polynomial kernels deliver intermediate results, with higher uncertainty than their RBF counterparts. The Laplace kernel in GPR, while excellent in training, shows slight PICP degradation during testing, whereas SVM's Linear kernel remains the least reliable due to consistently higher UA.

In conclusion, GPR with the RBF kernel is the top performer in the testing phase, offering robust uncertainty quantification and reliable coverage. SVM with RBF is a strong alternative, particularly for specific classes where it achieves lower UA. While Polynomial kernels in both models provide acceptable coverage, their higher uncertainty limits their preference.

For optimal performance, GPR-RBF is recommended, with SVM-RBF as a viable secondary choice depending on class-specific requirements.

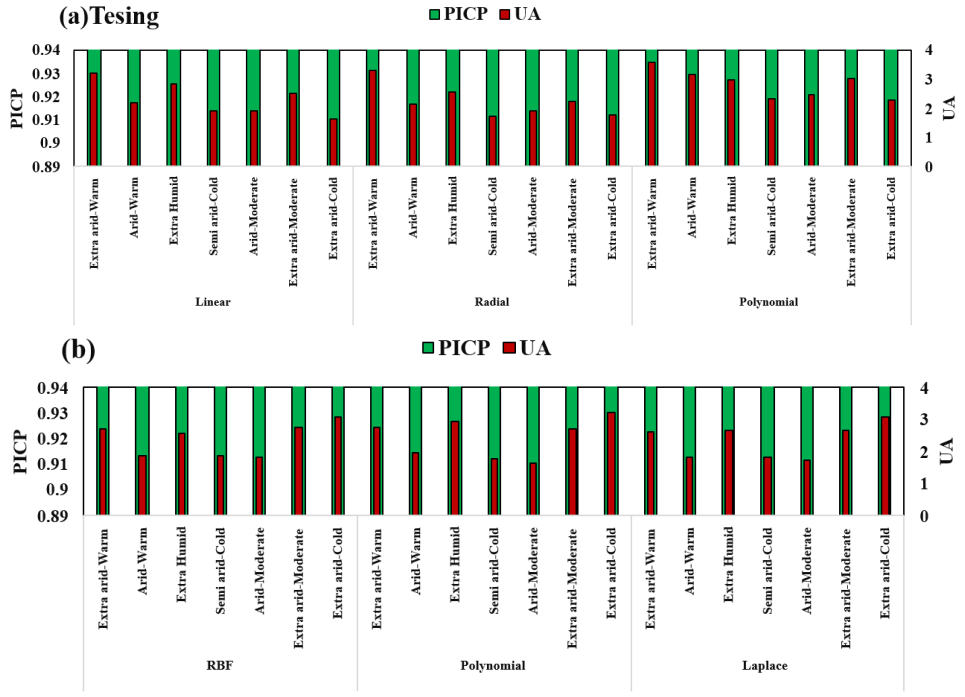

Fig 7. Uncertainty analysis in testing phase (a) SVM (b) GPR using different kernels

The results of this study demonstrate that Gaussian Process Regression (GPR) with Laplace and RBF kernels outperforms Support Vector Machines (SVM) in uncertainty estimation. This superiority stems from GPR's inherent probabilistic framework, which precisely calculates prediction intervals through posterior Gaussian distributions (Rasmussen & Williams, 2006), while SVM requires alternative approaches like Platt Scaling (Smola & Schölkopf, 2004). The Laplace kernel in GPR showed particularly strong performance with non-smooth data (e.g., environmental datasets) (Shi & Choi, 2011). During training, GPR-Laplace achieved the lowest uncertainty (UA) and perfect coverage (PICP=1.0), though GPR-RBF demonstrated more stable generalization during testing. While SVM-RBF remained competitive, GPR's inherent advantages in uncertainty quantification (Bishop & Nasrababdi 2006) make it the preferred choice for uncertainty-sensitive applications, particularly in environmental modeling tasks.

Conclusion

This study establishes the Gaussian Process Regression model with Laplace kernel as the superior approach for drought forecasting across Iran's

diverse climates, demonstrating consistently higher prediction accuracy and more reliable uncertainty quantification compared to Support Vector Machines. The probabilistic framework of GPR-Laplace provides not only point forecasts but also essential confidence intervals, enabling a fundamental shift from reactive to proactive drought management.

For water resource managers, the GPR-Laplace model enables precise monthly drought forecasts that can directly inform water allocation decisions across agricultural, industrial, and domestic sectors. Farmers and agricultural extension services can utilize these forecasts for critical decisions regarding crop selection, irrigation scheduling, and drought-resistant practices, with the model's uncertainty quantification providing clear guidance on forecast reliability. Emergency management agencies benefit from the model's ability to identify high-risk regions with greater confidence, allowing for strategic pre-positioning of resources. For operational integration, we recommend monthly model updates using near-real-time climate data and the development of user-friendly dashboards that visualize both forecasted SPEI values and their associated uncertainty ranges.

This research acknowledges several limitations that warrant consideration. The model performance remains constrained by the quality and temporal coverage of historical climate data, particularly in remote arid regions. The computational demands of GPR, while justified by its superior performance, may present challenges for real-time applications in resource-constrained settings. Furthermore, the study's focus on climatic drivers overlooks potential influences from human activities such as groundwater extraction and land-use changes. The model also shows reduced performance in capturing extreme drought events, suggesting the need for specialized approaches for tail-risk estimation.

Building upon these findings, promising research directions include developing hybrid models that integrate GPR with deep learning architectures for improved long-term forecasting, incorporating CMIP6 climate change scenarios to enhance predictive capability under warming conditions, and creating transfer learning frameworks to extend model applicability to data-scarce regions. Further investigation into the interactions between human water management practices and natural drought dynamics would significantly advance the practical utility of drought forecasting systems.

This research provides a robust

scientific foundation for advancing drought early warning capabilities in Iran, offering stakeholders a reliable framework for evidence-based decision-making in water resource management and drought mitigation strategies.

References

Azhdari Moghadam, M., Khosravi, M., Hosseinpour Niknam. H., Jaffari Nadoushan, E. (2011). Prediction of drought using Fuzzy neural network model, climatic indices, precipitationand drought index. *Kournal of Geography and Development*, *10*(26), 61-72. [In Persian]. https://doi.org/10.22111/gdij.2012.443.

Abbasi, A., Khalili, K., Behmanesh, J., & Shirzad, A. (2020). Application of support vector machine and bayesian network for agricultural drought prediction. Watershed Engineering and Management, *12*(1), 107-124.https://doi.org/10.22092/ijwmse.2019.122619.1519.

Alizadeh, A., Erfanian, M., & Ansari, H. (2011). Assessment of teleconnection patterns affecting rainfall and temperature (case study: Mashhad synoptic station). [In Persian]. http://dx.doi.org/10.29252/jwmr.9.17.280.

Bazrafshan, J. (2017). Effect of air temperature on historical trend of long-term droughts in different climates of Iran. Water Resources Management, *31*(14), 4683-4698. https://doi.org/10.1007/s11269-017-1773-8.

Breiman, L. (2001). *Random forests. Machine learning*, *45*(1), 5-32. https://doi.org/10.1023/A:1010933404324.

Bazrafshan, O. (2013). *Hydrologic drought forecasting using teleconnection and intelligent models (Case study: Karkheh watershed).* Ph.D. Thesis in Watershed Science and Engineering, University of Tehran, Iran. [In Persian].

Bazrafshan, J. (2017). Effect of air temperature on historical trend of long-term droughts in different climates of Iran. *Water Resources Management*, *31*(14), 4683-4698. http://dx.doi.org/10.1007/s11269-017-1773-8.

Bazrafshan, O., Zamani, H., & Shekari, M. (2020). A copula-based index for drought analysis in arid and semi-arid regions of Iran. *Natural Resource Modeling*, *33*(1), e12237. https://doi.org/10.1111/

nrm.12237.

Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.

Chalimourda, A., Schölkopf, B., & Smola, A. J. (2004). Experimentally optimal ν in support vector regression for different noise models and parameter settings. *Neural Networks*, *17*(1), 127-141. https://doi.org/10.1016/S0893-6080(03)00209-0.

Choubin, B., Malekian, A., Samadi, S., Khalighi-Sigaroodi, S., & Sajedi-Hosseini, F. (2017). An ensemble forecast of semi-arid rainfall using large-scale climate predictors. *Meteorological Applications*, *24*(3), 376-386. https://doi.org/10.1002/met.1635.

Cao, J., Zhang, Z., Tao, F., Chen, Y., Luo, X., & Xie, J. (2023). Forecasting global crop yields based on El Nino Southern Oscillation early signals. *Agricultural Systems*, *205*, 103564. https://doi.org/10.1016/j.agsy.2022.103564.

Dogulu, N., López López, P., Solomatine, D. P., Weerts, A. H., & Shrestha, D. L. (2015). Estimation of predictive hydrologic uncertainty using the quantile regression and UNEEC methods and their comparison on contrasting catchments. *Hydrology and Earth System Sciences*, *19*(7), 3181-3201. https://doi.org/10.5194/hess-19-3181-2015, 2015.

Elbeltagi, A., Kumari, N., Dharpure, J. K., Mokhtar, A., Alsafadi, K., Kumar, M., ... & Kuriqi, A. (2021). Prediction of combined terrestrial evapotranspiration index (CTEI) over large river basin based on machine learning approaches. *Water*, *13*(4), 547. https://doi.org/10.3390/w13040547.

Emadodin, I., Reinsch, T., & Taube, F. (2019). Drought and desertification in Iran. *Hydrology*, *6*(3), 66. https://doi.org/10.3390/hydrology6030066.

Fallah-Ghalhari, G., Shakeri, F., & Dadashi-Roudbari, A. (2019). Impacts of climate changes on the maximum and minimum temperature in Iran. *Theoretical and Applied Climatology*, *138*(3), 1539-1562. http://doi.org/10.1007/s00704-019-02906-9.

Farzanegan, M. R., Feizi, M., & Gholipour, H. F. (2021). Drought and property prices: Empirical evidence from provinces of Iran.

Economics of Disasters and Climate Change, *5*(2), 203-221. https://doi.org/10.1007/s41885-020-00081-0.

Ghasemi, P., Karbasi, M., Nouri, A. Z., Tabrizi, M. S., & Azamathulla, H. M. (2021). Application of Gaussian process regression to forecast multi-step ahead SPEI drought index. *Alexandria Engineering Journal*, *60*(6), 5375-5392. https://doi.org/10.1016/j.aej.2021.04.022.

Ghazipour, F., & Mahjouri, N. (2022). A multi-model data fusion methodology for seasonal drought forecasting under uncertainty: Application of Bayesian maximum entropy. *Journal of Environmental Management*, *304*, 114245. https://doi.org/10.1016/j.jenvman.2021.114245.

Haas, M., Mittnik, S., & Paolella, M. S. (2006). Modelling and predicting market risk with Laplace–Gaussian mixture distributions. *Applied Financial Economics*, *16*(15), 1145-1162. https://doi.org/10.1080/09603100500438817.

Hager, J., Hu, G., Hewage, K., & Sadiq, R. (2019). Performance of low-impact development best management practices: a critical review. *Environmental Reviews*, *27*(1), 17-42. https://doi.org/10.1139/er-2018-0048

Jafarpour, M., Adib, A., Lotfirad, M., & Kisi, Ö. (2023). Spatial evaluation of climate change-induced drought characteristics in different climates based on De Martonne Aridity Index in Iran. *Applied Water Science*, *13*(6), 133. https://doi.org/10.1007/s13201-023-01939-w.

Javanshiri, Z., Babaeian, I., & Pakdaman, M. (2023). Influence of large-scale climate signals on the precipitation variability over Iran. *Stochastic Environmental Research and Risk Assessment*, *37*(5), 1745-1762. https://doi.org/10.21203/rs.3.rs-1982295/v1.

Karimi, G., Gandomkar, A., & Abbasi, A. (2024). Correlation of Teleconnection Patterns with Temperature Series and Productivity of Barberry in Ghaenat Basin. *Journal of Climate Research*, *1403*(58), 123-140. https://doi.org/10.22034/jcr.2024.200832.

Khalili, A., Bazrafshan, J., & Cheraghalizadeh, M. (2022). A Comparative study on climate maps of Iran in extended de Martonne classification and application of the method for world climate zoning. *Journal of*

Agricultural Meteorology, *10*(1), 3-16. httl:// 10.22125/agmj.2022.156309.

Kushnir, Y., Seager, R., Ting, M., Naik, N., & Nakamura, J. (2010). Mechanisms of tropical Atlantic SST influence on North American precipitation variability. *Journal of Climate*, *23*(21), 5610-5628. https://doi.org/10.1175/2010JCLI3172.1.

Malik, A., Tikhamarine, Y., Sammen, S. S., Abba, S. I., & Shahid, S. (2021). Prediction of meteorological drought by using hybrid support vector regression optimized with HHO versus PSO algorithms. *Environmental Science and Pollution Research*, *28*(29), 39139-39158. http://doi.org/10.1007/s11356-021-13445-0.

Modarres, R., Sarhadi, A., & Burn, D. H. (2016). Changes of extreme drought and flood events in Iran. *Global and Planetary Change*, 144, 67-81. https://doi.org/10.1016/j.gloplacha.2016.07.008.

Nouri, M., & Homaee, M. (2020). Drought trend, frequency and extremity across a wide range of climates over Iran. *Meteorological Applications*, *27*(2), e1899. https://doi.org/10.1002/met.1899.

Oruc, S., Hinis, M. A., & Tugrul, T. (2024). Evaluating performances of LSTM, SVM, GPR, and RF for drought prediction in Norway: a wavelet decomposition approach on regional forecasting. *Water*, *16*(23), 3465. https://doi.org/10.3390/w16233465.

Pande, C. B., Sidek, L. M., Varade, A. M., Elkhrachy, I., Radwan, N., Tolche, A. D., & Elbeltagi, A. (2024). Forecasting of meteorological drought using ensemble and machine learning models. *Environmental Sciences Europe*, *36*(1), 160. https://doi.org/10.1186/s12302-024-00975-w.

Piri, S.& Ansari ghojghar, M. (2025). Development of intelligent machine learning and Jenkins box models based on the full ensemble mode decomposition method for meteorological drought modeling (Case study: Khuzestan Province)., *Journal of Climate Change Research. 6* (22),81-96. [In Persian]. https://doi.org/10.30488/ccr.2025.516420.1276.

Rahmati, O., Falah, F., Dayal, K. S., Deo, R. C., Mohammadi, F., Biggs, T., ... & Bui, D. T. (2020). Machine learning approaches for

spatial modeling of agricultural droughts in the south-east region of Queensland Australia. *Science of the total environment*, *699*, 134230. https://doi.org/10.1016/j.scitotenv.2019.134230.

Shi, J. Q., & Choi, T. (2011). *Gaussian process regression analysis for functional data*. CRC press. http://dx.doi.org/10.1201/b11038.

Smola, A., & Schölkopf, B. (2004). Experimentally optimal ν in support vector regression for different noise models and parameter settings. *Neural Networks*, *17*(1), 127-141. https://doi.org/10.1016/S0893-6080(03)00209-0.

Vicente-Serrano, S. M., Beguería, S., & López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of climate*, *23*(7), 1696-1718. https://doi.org/10.1175/2009JCLI2909.1.

Williams, C. K., & Rasmussen, C. E. (2006). *Gaussian processes for machine learning*, *2*(3), p 4. Cambridge, MA: MIT press. https://doi.org/10.7551/mitpress/3206.001.0001.

Wallace, J. M., & Gutzler, D. S. (1981). Teleconnections in the geopotential height field during the Northern Hemisphere winter. *Monthly weather review*, *109*(4), 784-812. https://doi.org/10.1175/1520-0493(1981)109%3C0784:TITGHF%3E2.0.CO;2.

Zamani, H., Pakdaman, Z., Shakari, M., Bazrafshan, O. and Jamshidi, S. (2025). Enhancing drought monitoring with a multivariate hydrometeorological index and machine learning-based prediction in the south of Iran. *Environmental Science and Pollution Research*, *32*(9), pp.5605-5627. https://doi.org/10.1007/s11356-025-36049-4.