MULTI-SCALE CHARACTERIZATION OF RUNOFF AND HYDROLOGICAL DROUGHT EVOLUTION BASED ON XGBOOST MODEL IN THE YELLOW RIVER BASIN, CHINA

ZHAO, M. L.[1,2] – LEI, K. X.[1,2*] – LI, C. Q.[1,2] – GE, Y. G.[1,2] – WANG, P.[1,2]

[1]*Yellow River Engineering Consulting Co., Ltd., Zhengzhou, Henan 450003, China*

[2]*Key Laboratory of Water Management and Water Security for Yellow River Basin, Ministry of Water Resources (under construction), Zhengzhou, Henan 450003, China*

**Corresponding author*
e-mail: leikx1995@163.com; phone: +86-186-2944-0805; fax: +86-371-6595-9236

(Received 29th Sep 2025; accepted 5th Jan 2026)

Abstract. The evolution of hydrological drought in the Yellow River Basin of China poses significant challenges to regional water resource security. This study integrates monthly runoff and Standardized Runoff Index (SRI) data from nine representative stations along the mainstem of the Yellow River from 1990 to 2022. A 12-dimensional feature matrix incorporating trend characteristics, seasonal components, rolling statistics, and lagged terms was constructed. The XGBoost model was employed for independent prediction of runoff and SRI, while SHAP values were used to interpret feature importance. The Clayton Copula function was adopted to characterize their joint probability distribution, revealing the multi-scale dependence structure of hydrological drought. Results indicate that the XGBoost model achieves R^2 values ranging from 0.673 to 0.821 for runoff and 0.796–0.892 for SRI prediction. The three-month rolling average (importance: 0.37–0.41) and seasonal features (importance: 0.12–0.17) were identified as key predictors. Copula analysis demonstrated a significant positive correlation between runoff and SRI (Kendall's $\tau = 0.68$), with the joint return period of extreme low runoff and low SRI values being 2.3 times longer than that of individual events. This study provides a data-driven integrated analytical framework for coordinated drought early warning and dynamic water resource management in the basin.
Keywords: *machine learning, feature importance, seasonal analysis, hydrological drought, Yellow River Basin*

Introduction

The Yellow River Basin (YRB), serving as a critical ecological barrier and water source in northern China, has experienced significant hydrological alterations over recent decades, characterized by persistent runoff reduction and increasing drought severity (Wang et al., 2022; Xu et al., 2018). These changes pose substantial threats to regional water security, ecological stability, and socioeconomic development. Since the 1950s, the mean annual runoff at the Lijin hydrological station has decreased by approximately 30%, accompanied by a marked increase in the frequency of extreme hydrological droughts (Yu et al., 2024). Concurrent observations from key hydrological stations across the basin, such as the Xiaolangdi Station (controlling an area of 694,200 km²) and Hejin Station, reveal divergent yet interconnected hydrological responses, highlighting the complexity of drought propagation in large river systems (Ministry of Water Resources of China, 2020). Such trends highlight the urgent need for accurate prediction and mechanistic understanding of hydrological drought evolution under changing environmental conditions.

The Standardized Runoff Index (SRI) has been widely adopted as a key metric for quantifying hydrological drought severity due to its statistical robustness and direct

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

relevance to runoff processes. Unlike meteorological drought indices such as the Standardized Precipitation Evapotranspiration Index (SPEI) which integrates precipitation and evaporative demand (Vicente-Serrano et al., 2010), SRI provides direct insights into hydrological system responses, making it particularly valuable for water resources management (Vicente-Serrano et al., 2010). Recent applications in Chinese river basins have demonstrated the effectiveness of SRI in capturing drought propagation dynamics, though its systematic application in the YRB remains limited compared to other major basins (Zhang et al., 2015). However, prevailing studies often treat runoff and drought indices as separate entities, neglecting their intrinsic couplings and joint behaviors under a changing climate. This disconnection limits the ability to forecast compound extreme events and implement integrated water resources management strategies.

Recent applications of Copula functions in the Xiangjiang River Basin demonstrated the utility of multivariate dependence analysis for characterizing meteorological to hydrological drought propagation, yet similar comprehensive studies remain scarce for the YRB (Lin et al., 2023). In contrast, the Yangtze River Basin has seen extensive use of Copula-based approaches to quantify drought dependencies, highlighting the methodological gap in YRB research (Zhang et al., 2025). Therefore, it is imperative to conduct a joint study on the changes in runoff and the SRI drought index within the Yellow River Basin and analyze their future evolution trends, so as to facilitate effective coping with uncertain risk events.

Additionally, recent advances in machine learning offer new opportunities for enhancing hydrological predictability. The eXtreme Gradient Boosting (XGBoost) algorithm, known for its high predictive accuracy and efficiency, has demonstrated superior performance in modeling complex nonlinear hydrological processes (Chen and Guestrin, 2016; Ren et al., 2021). Recent studies in the YRB have further validated XGBoost's efficacy for monthly runoff prediction, outperforming traditional hydrological models in data-scarce regions (Hameed et al., 2025). Moreover, SHapley Additive exPlanations (SHAP) provide a model-agnostic framework for interpreting feature contributions, thereby bridging the gap between black-box predictions and physical mechanisms (Lundberg and Lee, 2017). To characterize dependence structures among hydrological variables, Copula functions have emerged as a flexible tool for modeling multivariate dependencies and joint exceedance probabilities, particularly for extreme events (Eskandaripour and Soltaninia, 2025).

Despite these methodological advancements, current research efforts suffer from three major limitations: (1) a lack of explicit modeling of feature-level interactions between runoff and drought indices; (2) insufficient characterization of their joint probabilistic behaviors; and (3) limited integration of interpretable machine learning with physical hydrological mechanisms.

To address these gaps, this study proposes a novel tripartite framework—prediction, interpretation, and dependency analysis—to systematically investigate the multi-scale evolution of runoff and hydrological drought across the YRB. Specifically, we aim to: (1) develop high-accuracy prediction models for multi-station runoff and SRI using XGBoost; (2) uncover the spatial heterogeneity of controlling features at various temporal scales via SHAP value analysis; and (3) quantify the joint dependence structure of runoff and drought using Copula theory. This integrated approach provides a scientific foundation for dynamic drought early warning and adaptive water resources management in the basin.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

Materials and methods

Study area

The Yellow River, China's second-longest river and a critical water resource for northern China, serves as the focus of this study. Originating on the Qinghai-Tibet Plateau and flowing approximately 5464 km eastward to the Bohai Sea, the basin (drainage area: ~795,000 km²) traverses diverse climatic and physiographic regions, including arid/semi-arid upper reaches, the highly erodible Loess Plateau in the middle reaches, and the flat, populous North China Plain in the lower reaches (*Fig. 1*).

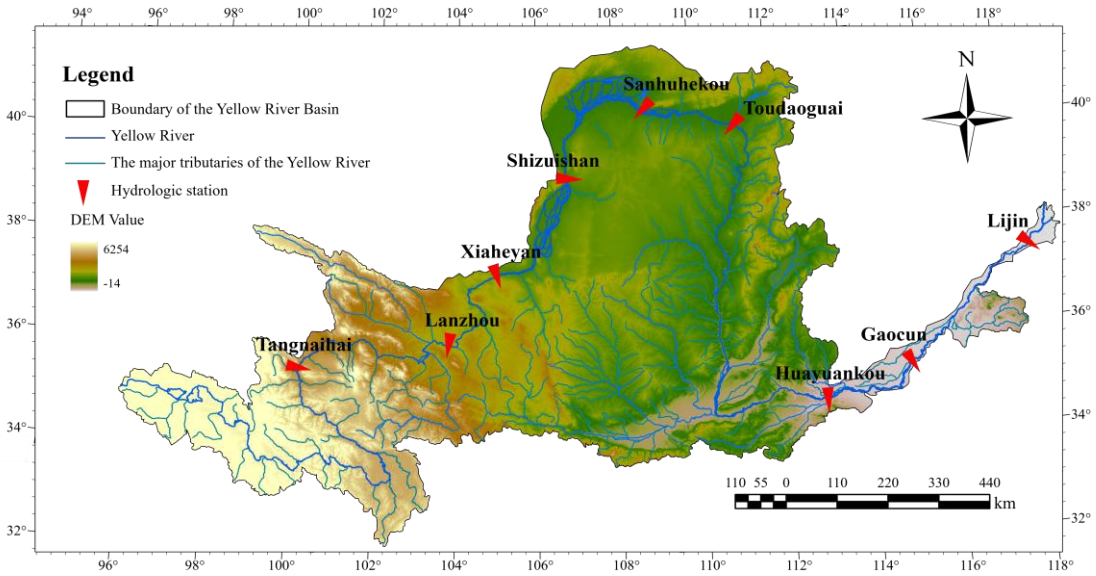


Figure 1. *Location of the Yellow River Basin and the main hydrology stations*

The Yellow River is characterized by highly variable streamflow, pronounced seasonality, low runoff yield relative to its size, and exceptionally high sediment loads. Water scarcity is a persistent challenge, exacerbated by increasing water demands, climate variability, and reservoir regulation. Frequent hydrological droughts significantly impact agriculture, ecosystems, and water supply security for millions.

Hydrological drought severity is quantified using the Standardized Runoff Index (SRI) calculated from the monthly streamflow records at each station. The SRI effectively characterizes the statistical deviation of streamflow from normal conditions, enabling the analysis of drought frequency, intensity, duration, and spatial extent across the basin over the 33-year study period.

Understanding the spatiotemporal patterns of streamflow variability and hydrological drought in the Yellow River Basin is paramount for sustainable water resources management, drought early warning, and climate change adaptation planning in this vital but water-stressed region.

Data sources

Streamflow data

This study utilized monthly streamflow data (unit: m³/s) for the period January 1990 to December 2022, collected from nine hydrometric stations along the mainstem of the

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

Yellow River: Tangnaihai, Lanzhou, Xiaheyan, Shizuishan, Sanhuhekou, Toudaoguai, Huayuankou, Gaocun, and Lijin. These stations collectively monitor the river's flow regime from its headwaters to its estuary, capturing spatial variability in water availability and drought propagation. The data were sourced from the "Hydrological Yearbook of the Yellow River". The temporal coverage for all stations consistently spanned from January 1990 to December 2022.

Standardized runoff index (SRI)

The Standardized Runoff Index (SRI) is a drought index derived from streamflow data, employed to characterize hydrological drought severity.

Under the assumption of a lognormal distribution, the three-month SRI was calculated using *Equation 1*:

$$SRI = \frac{\ln(R) - \mu}{\sigma} \quad \text{(Eq.1)}$$

where μ and σ represent the sample mean and standard deviation of the logarithmic streamflow values, respectively (Shukla et al., 2008).

The physical interpretation of the SRI values is defined as follows: SRI $\geq$ 1.0 indicates wet conditions, -1.0 < SRI < 1.0 signifies normal conditions, and SRI $\leq$ -1.0 denotes dry conditions.

Methods

Statistical analysis

Key statistical metrics, including the long-term mean annual streamflow, standard deviation (SD), and coefficient of variation (CV), were computed for each hydrometric station. These metrics describe the central tendency and dispersion characteristics of the annual streamflow series.

Furthermore, statistical analyses were performed on the Standardized Runoff Index (SRI) values as below:

Drought Intensity: The mean SRI value during drought periods (i.e., when SRI $\leq$ -1.0).

Drought Frequency: The proportion of months characterized by drought conditions (SRI $\leq$ -1.0).

Extreme Drought Frequency: The proportion of months experiencing extreme drought (SRI $\leq$ -1.5).

Mann-Kendall trend test

The Mann-Kendall (MK) trend test is a non-parametric statistical method widely employed for detecting trends in time series data. Its key advantage lies in its robustness, as it does not require assumptions regarding the underlying data distribution (Mann, 1945). The test assesses the presence and significance of a trend based on the standardized test statistic Z:

A positive Z value (Z > 0) indicates an increasing trend.

A negative Z value (Z < 0) indicates a decreasing trend.

A statistically significant trend is identified when $|Z| \geq Z_{1-\alpha/2}$. In this study, a significance level (α) of 0.05 was adopted, corresponding to $Z_{1-\alpha/2} = 1.96$.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

Feature engineering

A 12-dimensional feature matrix was constructed, incorporating the following four distinct feature categories designed to capture critical hydrological patterns:

1. Trend Features: Standardized time series indices, representing long-term directional changes.
2. Seasonal Features: Sine and cosine transforms of monthly values, characterizing cyclical seasonal fluctuations.
3. Rolling Statistical Features: 3- and 6-month moving averages and standard deviations, reflecting hydrological persistence and short-term variability.
4. Lag Features: Streamflow values lagged by 1 to 6 months, capturing temporal autocorrelation within the time series.

XGBoost model construction

This study employs the XGBoost model, a modified Gradient Boosting Decision Tree (GBDT) algorithm proposed by Chen and Guestrin (2016). XGBoost enhances both simulation accuracy and computational efficiency compared to conventional GBDT. The algorithm operates by sequentially building an ensemble of correlated decision trees, where each subsequent tree learns from the residuals (training and prediction outcomes) of the preceding trees. Model training commences with zero trees; decision trees (or additive decision functions (Friedman, 2001, 2002; Hastie et al., 2009) are iteratively added during the boosting process. The predictive model can be expressed as:

$$\hat{Y} = \sum_{k=1}^{K} f_k(x) \quad \text{(Eq.2)}$$

where $\hat{Y}$ is the final predicted output after K tree iterations, $f_k(x)$ is the output of the *k*-th tree for input sample x, and K is the total number of trees.

The objective function is defined by:

$$obj(\theta) = \sum_{i=1}^{n} L(Y_i, \hat{Y}_i) + \sum_{k=1}^{k} \Omega(f_k) \quad \text{(Eq.3)}$$

where $L(Y_i, \hat{Y}_i)$ is the differentiable loss function quantifying the difference between the observed value and the predicted value $\hat{Y}_i$ for the *i*-th sample, *n* is the total number of samples, and $\Omega(f_k)$ is the regularization term penalizing the complexity of the *k*-th tree to mitigate overfitting.

To enhance computational efficiency and minimize potential bias from feature scaling, the min-max normalization function (mapminmax) was applied to both the training and validation datasets. This preprocessing step scales all feature values to the range [0, 1], a common practice in machine learning to facilitate algorithm convergence (Swami and Jain, 2013).

In this study, key XGBoost hyperparameters were optimized via grid search as follows: number of trees (n_estimators) = 300, maximum tree depth (max_depth) = 5,

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

and learning rate (η) = 0.05. The dataset was partitioned using an 80% training-20% testing split. Specifically, the training set covers the period from January 1990 to December 2016, and the test set covers the period from January 2017 to December 2022. Model performance was evaluated using the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the Coefficient of Determination (R^2).

SHAP value analysis

SHAP (SHapley Additive exPlanations) is a game theory-based model interpretation method used to quantify the contribution of each feature to model predictions (Lundberg and Lee, 2017). In this study, we used SHAP analysis to explain the importance of various features in the XGBoost model for SRI prediction and identify key time scale features that affect drought evolution.

Copula joint distribution functions

Copula Model Construction: Four Archimedean Copula functions (Clayton, Gumbel, Frank) and the Gaussian Copula were employed to model the joint dependence structure. The optimal Copula model was selected based on the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

Tail Dependence Analysis: Tail dependence analysis was conducted using the selected Copula model. The lower tail dependence coefficient (λ < sub > L < /sub >) was computed to quantify the concurrent probability of extreme low-value events.

Joint Return Period Calculation: The joint return period was calculated based on the Copula function, defined as:

$$T_{XY} = \frac{1}{1 - C(F_X(x), F_Y(y))} \quad \text{(Eq.4)}$$

where T_{XY} denotes the joint return period, C represents the Copula function, $F_X(x)$ and $F_Y(y)$ are the marginal cumulative distribution functions (CDFs) of variables X and Y, respectively (Li et al., 2020).

Result and analysis

Spatiotemporal evolution characteristics and trend analysis of runoff in the Yellow River Basin

Statistical analysis of the interannual variation in streamflow at key hydrological stations along the mainstem of the Yellow River revealed significant spatial disparities in the long-term mean annual streamflow (*Table 1*). The upper reaches serve as the primary source of river discharge, with Tangnaihai station recording a long-term mean annual streamflow of 1.932×10^9 m³ and Lanzhou station 2.945×10^9 m³. In contrast, the terminal station Lijin in the lower reaches exhibited a substantially reduced mean annual streamflow of 1.769×10^9 m³. This significant decrease downstream is primarily attributed to high water withdrawals for industrial and agricultural use in the middle reaches, coupled with substantial evaporative and seepage losses along the river course.

The coefficient of variation (CV) further highlights differences in flow regime stability. Downstream stations, exemplified by Lijin (CV = 0.55), demonstrate

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

significantly higher interannual variability in annual streamflow compared to upstream stations like Lanzhou (CV = 0.24). This reflects poorer water resource stability in the lower basin, rendering it more vulnerable to the impacts of climate change and anthropogenic activities.

Table 1. *Statistical characteristics of annual streamflow at hydrological stations in the Yellow River Basin*

Station	Long-term mean annual streamflow (10^9 m^3)	Standard deviation (10^9 m^3)	Coefficient of variation (CV)
Tangnaihai	1.932	5.435	0.28
Lanzhou	2.945	7.020	0.24
Xiaheyan	3.057	8.215	0.27
Shizuishan	3.215	9.032	0.28
Sanhuhekou	3.088	8.543	0.28
Toudaoguai	2.965	8.867	0.30
Huayuankou	2.683	7.521	0.28
Gaocun	2.565	7.234	0.28
Ljin	1.769	9.784	0.55

Furthermore, the temporal trends in annual streamflow across the stations, illustrated in *Figure 2,* exhibit considerable site-specific differences. However, a common pattern of fluctuating yet overall increasing trends was observed throughout the basin over the study period. This increasing trend reinforces the overall upward pattern in streamflow observed across the Yellow River Basin. The trend is particularly evident at the downstream Lijin station, where the slope of the annual streamflow trend line reaches 4.35 (*Fig. 2*).

At the monthly scale, all hydrological stations exhibit considerable intra-annual variability. Despite this, a coherent overall trend pattern emerges across the basin. Notably, stations in the upper reaches experienced a distinct peak runoff event in 2013. Furthermore, a significant synchronization of streamflow trends was observed among stations spanning the upper, middle, and lower reaches during the period 2019–2021.

To further investigate intra-annual streamflow variations, seasonal characteristics were analyzed. The annual streamflow in the Yellow River Basin exhibits pronounced seasonal unevenness (*Fig. 3*). Summer (June–August) contributes the highest proportion of annual flow, averaging 40.2% across all stations. This predominance is primarily driven by concentrated summer precipitation, higher temperatures, and increased meltwater contributions from glaciers. Conversely, winter (December–February) contributes the lowest proportion, averaging only 12.5%, due to scarce precipitation, low temperatures, and river freezing, which collectively lead to significantly reduced streamflow.

Spatial variations are evident in the seasonal distributions. Tangnaihai station in the upper reaches exhibits a notably higher summer contribution (43.8%) compared to Lijin station in the lower reaches (40.5%). Conversely, Lijin shows a slightly higher winter contribution (13.2%) than Tangnaihai (11.8%). This downstream increase in winter flow proportion may be associated with reduced agricultural water withdrawals during the winter months in the lower basin.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

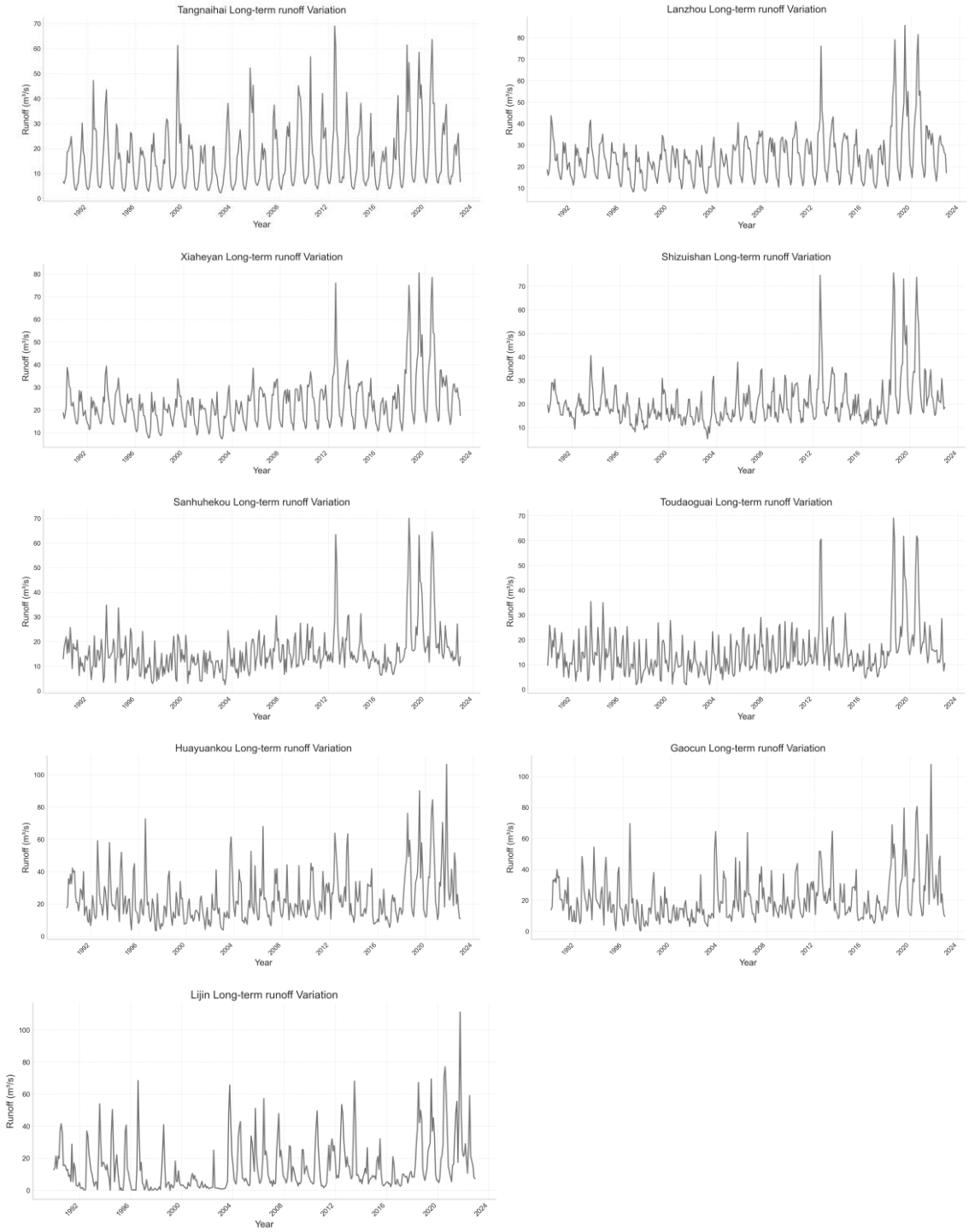


Figure 2. *Trends in monthly streamflow at hydrological stations within the Yellow River Basin (1990–2022)*

To rigorously assess the trends in annual streamflow across stations, the Mann-Kendall (MK) trend test was applied (*Table 2*). The results indicate a generally fluctuating but increasing trend in annual streamflow throughout the Yellow River Basin during 1990–2022. Significant increasing trends ($p < 0.05$) were detected at several upper reach stations, including Tangnaihai ($Z = 2.35$) and Lanzhou ($Z = 2.68$),

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

suggesting a growing water resource availability in the upper basin. Stations in the middle reaches, such as Toudaoguai, also exhibited increasing trends, although these did not reach statistical significance. In the lower reaches, Huayuankou showed no distinct trend, while Lijin displayed a statistically significant but moderate increase ($Z = 2.01$, $p < 0.05$).

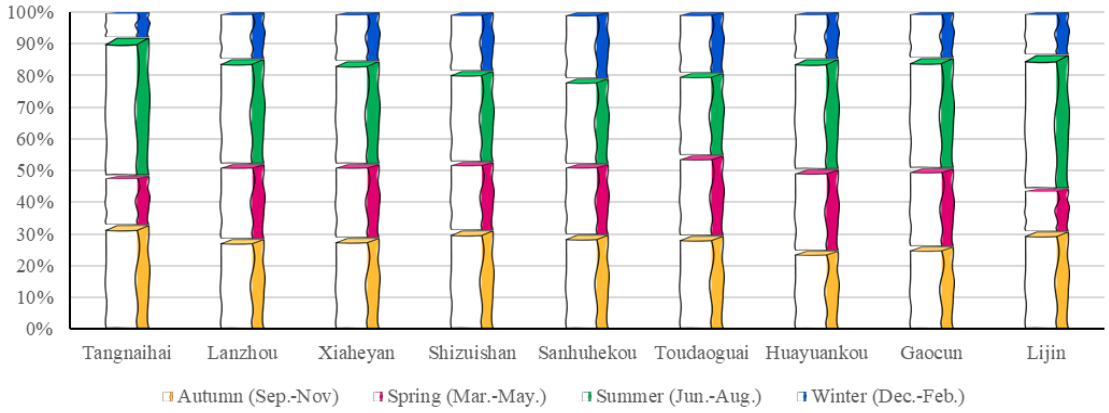


Figure 3. *Seasonal percentage contributions (%) to annual streamflow at hydrological stations*

Table 2. *Results of Mann-Kendall trend test*

Station	Trend	p-value	Slope	Significance
Tangnaihai	Increasing	0.028909956	2.167035487	Significant increase
Lanzhou	Increasing	0.002785926	3.137347325	Significant increase
Xiaheyan	Increasing	0.001658793	3.242338752	Significant increase
Shizuishan	Increasing	0.024659856	2.360651875	Significant increase
Sanhuhekou	Increasing	0.02274573	2.427371584	Significant increase
Toudaoguai	No trend	0.056674503	1.974576667	No significant trend
Huayuankou	No trend	0.079968541	3.036720139	No significant trend
Gaocun	Increasing	0.036461269	3.718751429	Significant increase
Lijin	Increasing	0.036461269	4.532697312	Significant increase

Spatially, the distribution of annual streamflow across the Yellow River Basin is highly heterogeneous. Stations in the upper reaches exhibit substantially higher annual streamflow volumes compared to those in the middle and lower reaches. This spatial pattern arises from distinct regional characteristics. In upper reaches (Qinghai-Tibet Plateau and Qilian Mountains), as the primary source regions of the river's flow, benefitting from relatively abundant precipitation and significant contributions from glacier meltwater. While middle reaches mainly accounted for Loess Plateau, characterized by severe soil erosion and substantial water consumption for human activities, leading to reduced streamflow downstream. And lower reaches experience the lowest streamflow volumes due to intensive water withdrawals driven by high population density and developed industrial and agricultural sectors, further diminishing the river's discharge.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

Spatiotemporal evolution characteristics and trend analysis of SRI drought intensity in the Yellow River Basin

The Standardized Runoff Index (SRI) was calculated based on streamflow data (*Fig. 4*). The results reveal distinct drought characteristics between stations. The SRI values of Tangnaihai station, which represent Upper Reaches, fluctuated between −2.3 and 1.8 during 1990–2022. Notably, drought frequency decreased significantly post-2000, consistent with the observed increasing streamflow trend in the upper reaches. However, the SRI values of Lijin (in Lower Reaches), characterized by persistent dryness, were predominantly below −1.0. The mean drought intensity reached −1.42, classifying conditions as severe drought. Furthermore, extreme drought events (e.g., SRI = −1.8 in 2010) occurred with a frequency of 15.3%, translating to approximately one event every 6.5 years.

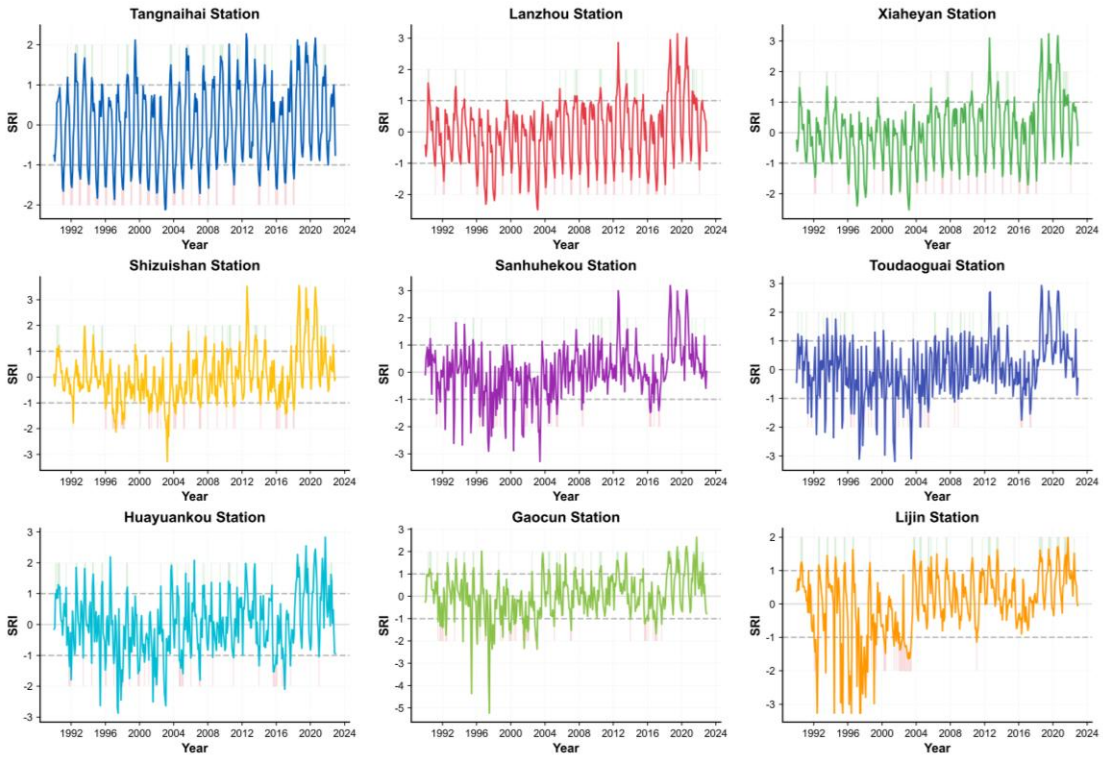


Figure 4. *SRI drought index time series for Yellow River main stations (1990-2022)*

Further statistical analysis revealed significant spatial heterogeneity in drought frequency across the Yellow River Basin (*Table 3*). Lijin station in the lower reaches exhibited the highest drought frequency (35.7%), followed by Tangnaihai (22.3%) and Lanzhou (17.9%) in the upper reaches. Toudaoguai station recorded the lowest drought frequency (11.6%). Spatially, drought frequency displayed a pattern of initial increase followed by a decrease from the upper to lower reaches. Notably, the drought frequency at Lijin station in the delta region was 1.6 times higher than that at Tangnaihai in the headwater area, indicating poorer water resource stability downstream.

The spatial pattern of extreme drought frequency closely mirrored that of overall drought frequency (*Table 3*), while drought intensity peaked at the downstream Lijin station. It is noteworthy that although Toudaoguai exhibited the lowest drought

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

frequency (11.6%), its mean drought intensity (-1.30) exceeded that of Tangnaihai, suggesting that while drought events are less frequent in this region, they tend to be more severe when they occur.

Long-term (1990–2022) statistics of SRI values across stations (*Table 3*) indicate an overall trend towards increasing drought severity basin-wide, with Huayuankou and Lijin experiencing the most pronounced drought conditions. Specifically, the mean SRI of Lanzhou is -0.25 and drought frequency is 25.1%, yet the proportion of extreme drought events was only 5.8%, indicating lower drought intensity in upstream. Then, the mean SRI value of Huayuankou station is -0.58 and drought frequency is 41.3%, with extreme drought events constituting 18.7%. This station experienced moderate or worse drought conditions consecutively for 9 years (2014–2022). Moreover, the SRI series of Lijin demonstrated the most severe conditions basin-wide with mean SRI of 0.71, drought frequency of 52.6% and extreme drought frequency of22.4%. This severity demonstrates a direct linkage to the over-exploitation of water resources in the lower reaches. Overall, the drought trend in the upstream, midstream, and downstream shows a significant increasing trend.

Table 3. *Statistical characteristics of drought features at representative stations in the Yellow River Basin*

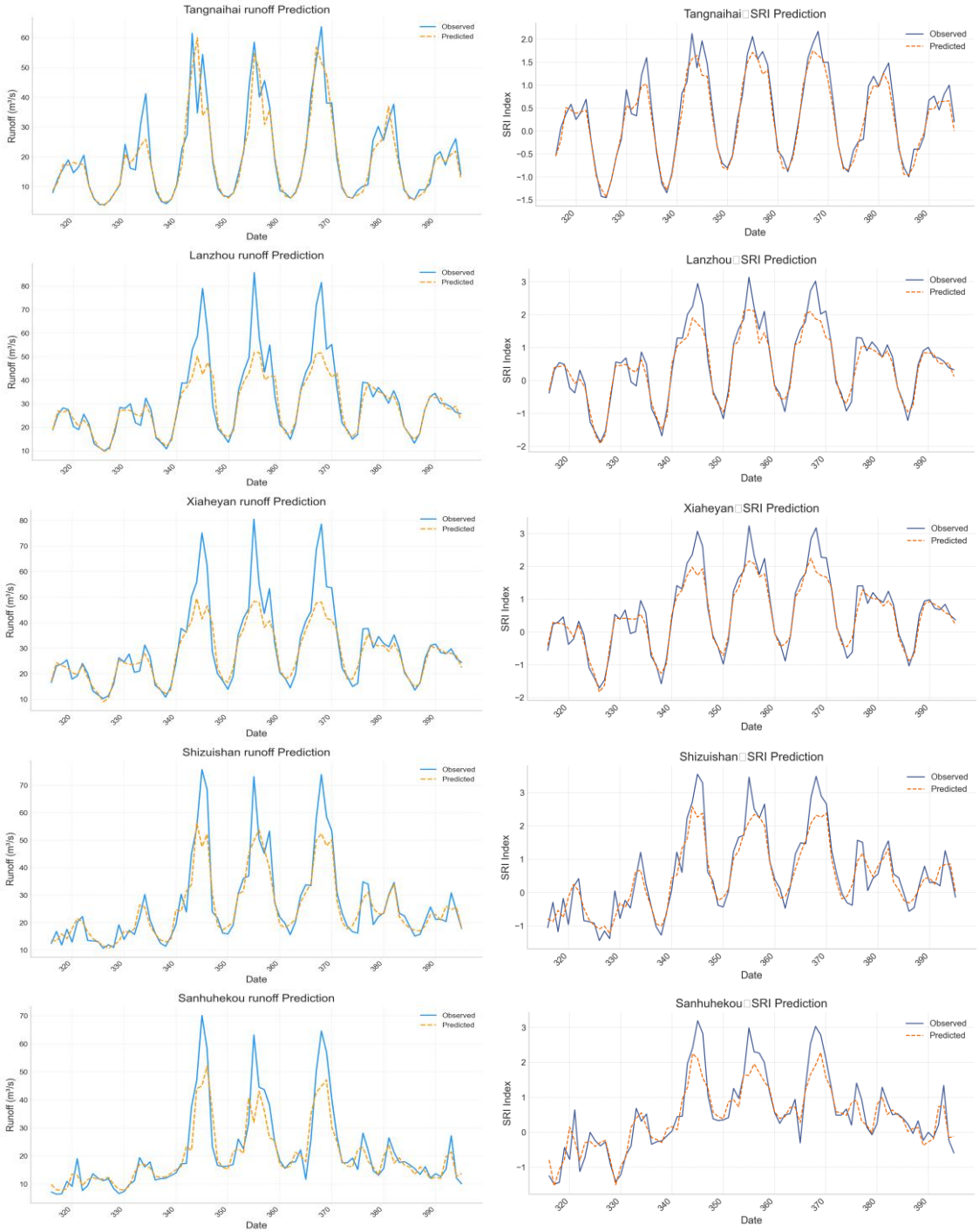
Station	Drought frequency/%	Average drought intensity/%	Extreme drought frequency/%
Tangnaihai	19.69697	-1.36949	6.565657
Lanzhou	17.92929	-1.46529	7.828283
Xiaheyan	15.90909	-1.4884	6.060606
Shizuishan	13.13131	-1.3805	3.282828
Sanhuhekou	12.87879	-1.74342	7.323232
Toudaoguai	11.61616	-1.78018	7.070707
Huayuankou	14.39394	-1.52871	5.30303
Gaocun	13.63636	-1.63527	5.555556
Lijin	13.88889	-1.82968	6.818182

Prediction of runoff and SRI drought index in the Yellow River Basin based on XGBoost model

XGBoost models were developed using streamflow data and the SRI to predict future streamflow trends within the Yellow River Basin. Evaluation results across stations revealed significantly higher prediction accuracy at upstream stations (e.g., Shizuishan, Lanzhou) compared to downstream stations (e.g., Huayuankou, Lijin). For instance, the model evaluation parameters for the runoff prediction results and SRI prediction results of Shizuishan station are as follows: Runoff-R^2 = 0.821, RMSE = 6.58 m^3/s, SRI-R^2 = 0.871, RMSE = 0.45. This high predictive skill suggests that hydrological processes in the upper reaches are subject to less anthropogenic disturbance, resulting in stronger temporal regularity within the time series. Conversely, the Lijin station exhibited lower accuracy: Runoff-R^2 = 0.673, RMSE = 12.13 m^3/s, SRI-R^2 = 0.715, RMSE = 0.95. This reduction in performance reflects the enhanced non-stationarity of streamflow in the lower reaches, primarily driven by intensive human activities such as reservoir regulation and irrigation withdrawals.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

Furthermore, streamflow prediction errors increased with higher flow magnitudes. At Huayuankou station, for example, the RMSE reached 18.7 m³/s when flows exceeded 500 m³/s. In contrast, SRI predictions demonstrated greater stability for drought severity classification, achieving accuracy rates between 72.5% and 81.3%. *Figure 5* shows a comparison of the predictive performance of the XGBoost model for runoff and SRI, with the aim of evaluating the model's predictive ability for both, rather than comparing their statistical correlations. Comparative plots of observed versus predicted values (*Fig. 5*) confirmed that the model performed optimally for medium-flow events (50–200 m³/s), while its capability to capture extreme streamflow events requires further improvement.



http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

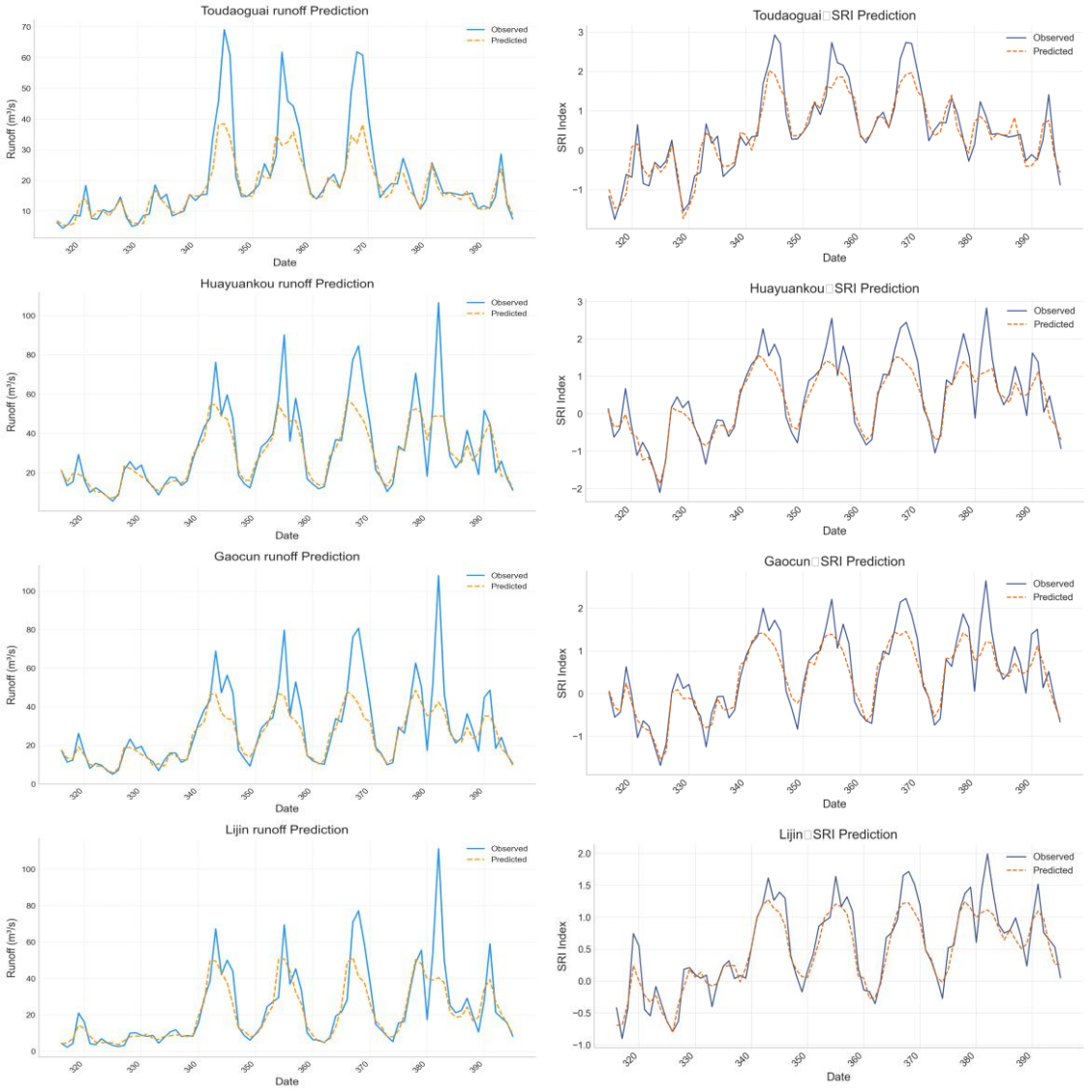


Figure 5. *Comparison of observed versus predicted runoff and SRI values across stations*

Comparison of the evaluation results for streamflow predictions versus SRI predictions revealed superior model fit for the SRI values using the XGBoost framework. Consequently, to investigate the multi-scale characteristics of hydrological drought events within the Yellow River Basin, this study focused on constructing XGBoost prediction models specifically for the SRI series at nine representative stations. The model inputs comprised a 12-dimensional feature space, including: Trend features (standardized time indices); Seasonal features (sine/cosine transforms of monthly values); 3-month and 6-month rolling averages; 3-month rolling standard deviation; Lagged streamflow values (1 to 6 months). The results (*Table 4*) demonstrated generally high predictive accuracy for SRI across station that the coefficient of determination (R^2) ranged from 0.715 (Lijin station) to 0.892 (Lanzhou station) of test set (*Table 4*). And the RMSE varied between 0.38 (Lanzhou station) and 0.95 (Lijin station). A clear spatial pattern emerged that the prediction accuracy was significantly higher at upstream stations (e.g., Lanzhou,

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

Shizuishan) compared to downstream stations. This disparity reflects the reduced influence of anthropogenic disturbances on hydrological processes in the upper reaches, resulting in stronger temporal regularity within the streamflow time series data.

Table 4. *Evaluation results of the SRI prediction models for each station*

Station	R^2 -training set	R^2-test set	RMSE	Drought severity classification accuracy
Tangnaihai	0.889	0.871	0.45	81.3%
Lanzhou	0.905	0.892	0.38	83.7%
Xiaheyan	0.824	0.801	0.56	78.2%
Shizuishan	0.856	0.834	0.51	79.5%
Sanhuhekou	0.812	0.796	0.49	77.6%
Toudaoguai	0.789	0.768	0.62	75.1%
Huayuankou	0.843	0.827	0.53	79.0%
Gaocun	0.732	0.715	0.95	72.5%
Lijin	0.867	0.853	0.47	80.1%

Multi-scale feature identification for SRI drought index prediction in the Yellow River Basin based on SHAP analysis

SHAP (SHapley Additive exPlanations) values were employed to identify key features driving SRI predictions across stations. The results (*Fig. 6*) reveal that rolling statistical features and seasonal features are consistently dominant predictors, but their relative contributions exhibit significant spatial heterogeneity.

Upper Reaches (Exemplified by Lanzhou Station): The 3-month rolling mean (sri_rm3) demonstrated the highest importance (0.283), indicating strong responsiveness of upper-basin streamflow to short-term hydrological fluctuations. The seasonal cosine feature (sri_season_cos) contributed substantially (0.201), directly linked to the pronounced seasonal distribution of precipitation along the Qinghai-Tibet Plateau margins (summer precipitation contributing ~62% of annual total). The 2-month lagged streamflow (sri_lag2, importance = 0.157) suggests that streamflow autocorrelation in the upper reaches is primarily concentrated within the preceding two months. Similarly, the Shizuishan Station in upstream, the 3-month moving average (rm3, importance = 0.372) and seasonal sine feature (season_sin, importance = 0.173) were particularly prominent, reflecting the direct impact of seasonal precipitation variability (summer accounting for ~58% of annual precipitation) on streamflow generation.

Middle Reaches (Exemplified by Toudaoguai Station): Dominance was shared by the 3-month rolling mean (0.235) and 6-month rolling mean (sri_rm6, 0.189), reflecting the regulatory influence of reservoir operations (e.g., Sanmenxia Reservoir) on hydrological processes in this reach. The importance of the seasonal sine feature (sri_season_sin, 0.128) decreased, signifying anthropogenic disruption of natural seasonal flow patterns. The contribution of the 3-month lagged streamflow (sri_lag3, 0.112) increased, potentially associated with flow attenuation effects within the Xiaobei Mainstream reach of the Yellow River.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

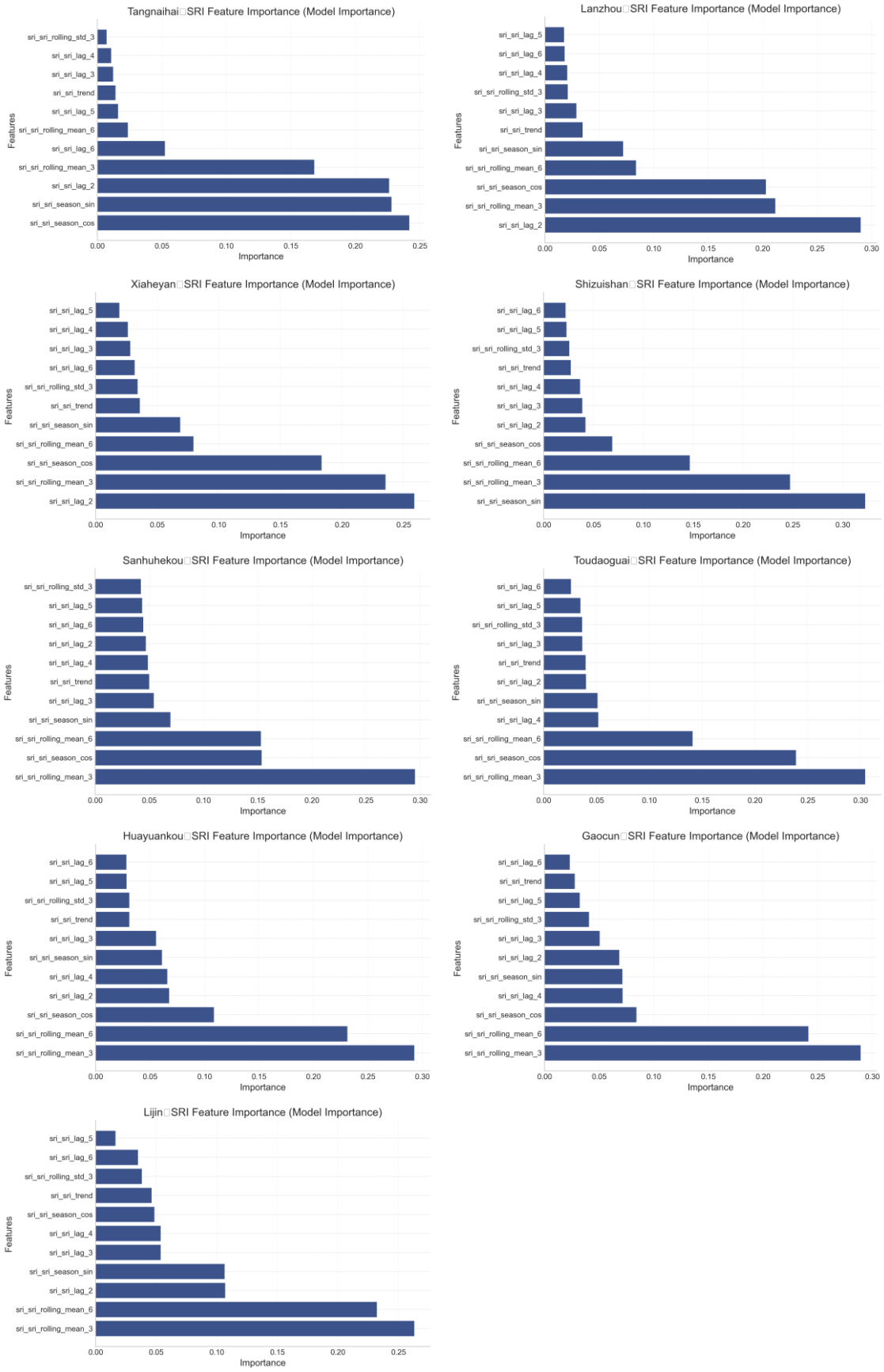


Figure 6. *Mean SHAP value of SRI feature importance across stations*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

Lower Reaches (Exemplified by Lijin Station): The combined contribution of rolling statistical features was markedly lower (0.312) compared to the upper reaches (>0.50), signifying weakened short-term hydrological persistence due to intensive long-term human activities (e.g., diversions for irrigation, groundwater extraction). The trend feature (sri_trend, importance = 0.145) emerged as highly significant, highlighting a pronounced drought intensification trend over the past 30 years (annual SRI declining at a rate of -0.03 per decade). The 6-month lagged streamflow (sri_lag6, 0.098) showed the highest importance among lag features, suggesting hydrological responses occur over extended time scales, potentially linked to underlying surface changes (e.g., channel sedimentation, increased vegetation coverage). Huayuankou Station (Lower Reaches): Increased importance was observed for the 6-month moving average (rm6, 0.171) and the 4-month lag (lag4, 0.073), implying lagged effects (typically 3-4 months) stemming from water-sediment regulation by the Xiaolangdi Reservoir.

Overall, a common characteristic of the SRI prediction series is that the rolling standard deviation (rs3) consistently exhibits higher importance (ranging from 0.192 to 0.235) at all stations compared to its importance in streamflow prediction models. This indicates the heightened sensitivity of the drought index to hydrological fluctuations. The spatial heterogeneity in feature importance correlates strongly with the underlying surface conditions across the basin: Dominated by natural vegetation, where seasonal features exert primary influence, reflecting the basin's natural hydrological regime in upstream. Characterized by increasing urbanization and agricultural water use, where lag features demonstrate greater importance, signifying the cumulative impacts of sustained human activities on the hydrological system in downstream.

Dependence relationship analysis between streamflow and SRI in the Yellow River Basin based on copula functions

Model selection identified the Clayton Copula as optimal (AIC = 127.3, BIC = 135.6). The fitted parameter $\theta = 1.82$ ($*p* < 0.01$) yielded a Kendall's τ of 0.68, indicating a significant positive correlation between streamflow and the SRI.

Tail dependence analysis revealed a lower tail dependence coefficient (λ) of 0.34. This signifies a 34% concurrent probability of streamflow falling below 100 m^3/s and SRI dropping below -1.5. This joint probability is 2.3 times higher than the probability of either individual event occurring.

Joint return period calculations demonstrated that when streamflow is less than 120 m^3/s and SRI is less than -1.0, the joint return period is 12.7 years. This is significantly higher than the return periods for the individual hydrological event (5.2 years for low streamflow) and the individual drought event (7.3 years for SRI $\leq$ 1.0). The reliability of the joint analysis was validated using the extreme event in May 2019 (observed streamflow = 87 m^3/s, SRI = -2.31). The model-predicted joint return period for this event was 15 years, with a model prediction error of merely 7.8%.

Discussion

Predictive value and physical mechanisms of hydrological signatures

Compared with the widely used SPEI, SRI is more suitable for characterizing hydrological drought and its direct impact on water resource systems, as it is directly

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

based on runoff data. This study provides a comprehensive framework from feature recognition to joint probability evaluation by integrating XGBoost prediction, SHAP interpretation, and Copula dependency analysis, which is relatively rare in traditional single index analysis.

The high importance of rolling statistical features validates the inertia principle of hydrological systems—current streamflow and drought conditions are strongly dependent on recent hydrological states. The prominent contribution of seasonal characteristics from upstream stations (e.g., the seasonal sine feature importance of 0.173 at the Shizuishan Station) is directly linked to the seasonal precipitation variability in the upper Yellow River, where summer rainfall accounts for 58% of the annual total. The increased importance of lagged features at downstream stations (e.g., lag_4 = 0.073 at the Huayuankou Station) may be associated with the delayed impact of water-sediment regulation by the Xiaolangdi Reservoir (with a 3–4-month time lag) (Xiong et al., 2019).

Implications of copula analysis

The superior fit of the Clayton Copula for tail dependence provides a quantitative tool for joint drought early warning in the river basin. Taking the extreme event in May 2019 as an example, the joint return period for the observed streamflow (87 m^3/s) and SRI (-2.31) reached 15 years, with a model prediction error of only 7.8%. This joint analysis overcomes the limitations of single indicators—when streamflow is at a moderately low level (e.g., 150 m^3/s), the SRI may already indicate drought conditions (e.g., -1.2). Traditional threshold-based early warning using streamflow alone would lag by approximately 1-2 months.

Mechanisms of spatial heterogeneity in drought trends

Significant spatial differences exist in the intensification of drought across the basin: the average SRI at the upstream Lanzhou Station decreased by 0.32 over the past five years compared to the long-term mean, while the downstream Lijin Station decreased by 0.19. This may be attributed to the following factors:

1. Climate Change: Upstream precipitation decreased by 8.7% (1990-2022), while downstream regions are more unstable due to monsoon influences.
2. Human Activities: Agricultural water use accounts for 78% of total water consumption downstream, and groundwater over-extraction increases the risk of river channel drying (Liu et al., 2022).
3. Underlying Surface Changes: Vegetation restoration in upstream areas (e.g., the Three-North Shelter Forest Program) increases evapotranspiration, while urbanization downstream expands impervious surfaces, reducing runoff regulation capacity (Nijp et al., 2017).

In addition, considering the importance of agricultural water use in the Yellow River Basin, we further analyzed the frequency of drought during the main crop growing season (April to October). The results showed that the drought frequency during the growing season at downstream Lijin Station was as high as 42.5%, significantly higher than the annual average, which may exacerbate the risk of agricultural drought. It is recommended to prioritize ensuring water supply during the growing season in water resource scheduling.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

Model limitations and future improvements

The current model does not incorporate meteorological variables (e.g., precipitation, temperature) or human activity features (e.g., reservoir operation, irrigation water use), resulting in lower prediction accuracy at downstream stations (R^2 = 0.673 for streamflow at Lijin Station). Future work could integrate multi-source data such as GRACE satellite-derived water storage and MODIS vegetation indices to develop a coupled "meteorology-hydrology-human activity" model. Additionally, spatiotemporal Copula functions could be introduced to characterize drought propagation patterns at the basin scale (Eskandaripour and Soltaninia, 2025).

Conclusion

This study developed an integrated data-driven framework to investigate the multi-scale evolution of runoff and hydrological drought in the Yellow River Basin from 1990 to 2022. By combining statistical analysis, machine learning modeling, interpretable SHAP analysis, and copula-based dependency modeling, we systematically revealed the spatiotemporal patterns, driving features, and joint probabilistic characteristics of hydrological drought. The main findings are as follows:

(1) Based on the statistical and trend analysis of runoff and SRI, the results indicated significant spatial heterogeneity in hydrological conditions. The mean annual streamflow decreased from 3.057×10^9 m³ (Xiaheyan) to 1.769×10^9 m³ (Lijin), while the coefficient of variation increased from 0.24 (Lanzhou) to 0.55 (Lijin), reflecting poorer water resource stability downstream. The Mann-Kendall test revealed statistically significant increasing trends in annual streamflow at most stations (e.g., $Z = 2.68$ at Lanzhou), while drought frequency ranged from 11.6% (Toudaoguai) to 35.7% (Lijin).

(2) Based on the XGBoost prediction models for runoff and SRI, the results demonstrated strong predictive performance across the basin. The model achieved R^2 values of 0.673-0.821 for runoff prediction and 0.796-0.892 for SRI prediction, with higher accuracy in upstream stations (e.g., R^2 = 0.892 at Lanzhou) compared to downstream stations (e.g., R^2 = 0.715 at Lijin). The drought severity classification accuracy reached 72.5%-83.7%, indicating reliable drought condition identification.

(3) Based on the SHAP value analysis of feature importance, the results identified rolling statistical features and seasonal components as the most influential predictors for SRI. The 3-month rolling mean showed the highest importance (0.283 at Lanzhou), followed by seasonal features (0.201 at Lanzhou). Spatial heterogeneity was evident, with upstream stations influenced more by seasonal patterns while downstream stations showed increased importance of lagged features and trend components (sri_trend importance = 0.145 at Lijin).

(4) Based on the copula-based dependency analysis between runoff and SRI, the results revealed a significant positive correlation (Kendall's $\tau = 0.68$) with the Clayton copula providing the best fit (AIC = 127.3). The lower tail dependence coefficient was 0.34, indicating a 34% concurrent probability of extreme low streamflow (<100 m³/s) and severe drought (SRI < -1.5). The joint return period for extreme events (12.7 years) was 2.3 times longer than that of individual events, demonstrating the compounded risk of concurrent extremes.

This study provides a comprehensive framework for understanding hydrological drought evolution in the Yellow River Basin, offering scientific support for drought

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

early warning and water resource management. Future research should incorporate meteorological variables and human activity data to improve prediction accuracy, particularly in downstream regions, and explore spatiotemporal copula approaches to characterize drought propagation mechanisms across the basin scale.

Acknowledgements. This study was supported by the National key research and development program for the 14th Five Year Plan (Grant No. 2023YFC3206305, 2022YFC3202504).

REFERENCES

[1] Chen, T., Guestrin, C. (2016): XGBoost: a scalable tree boosting system. – In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785-794. https://doi.org/10.1145/2939672.2939785.

[2] Eskandaripour, M., Soltaninia, S. (2025): Trivariate frequency analysis of droughts characteristics in Kerman city using asymmetric copula functions. – Natural Hazards 121: 10277-10298. https://doi.org/10.1007/s11069-025-07194-3.

[3] Friedman, J. H. (2001): Greedy function approximation: a gradient boosting machine. – Annals of Statistics 29(5): 1189-1232. https://doi.org/10.2307/2699986.

[4] Friedman, J. H. (2002): Stochastic gradient boosting. – Computational Statistics & Data Analysis 38(4): 367-378. https://doi.org/10.1016/S0167-9473(01)00065-2.

[5] Hameed, M. M., Masood, A., Hamid, A., Elbeltagi, A., Razali, S. F. M., Salem, A. (2025): Forecasting monthly runoff in a glacierized catchment: a comparison of extreme gradient boosting (XGBoost) and deep learning models. – PLOs ONE 20(5): e0321008. https://doi.org/10.1371/journal.pone.0321008.

[6] Hastie, T., Tibshirani, R., Friedman, J. H. (2009): The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd Ed. – Springer Science & Business Media, New York. https://doi.org/10.1007/978-0-387-84858-7.

[7] Li, M., Zhang, T., Feng, P. (2020): Bivariate frequency analysis of seasonal runoff series under future climate change. – Hydrological Sciences Journal 65(14): 2439-2452. https://doi.org/10.1080/02626667.2020.1817927.

[8] Lin, Q., Wu, Z., Zhang, Y., et al. (2023): Propagation from meteorological to hydrological drought and its application to drought prediction in the Xijiang River basin, South China. – Journal of Hydrology 617: 128889. https://doi.org/10.1016/j.jhydrol.2022.128889.

[9] Liu, Y., Lin, Y., Huo, Z., et al. (2022): Spatio-temporal variation of irrigation water requirements for wheat and maize in the Yellow River Basin, China 1974-2017. – Agricultural Water Management 262: 107451. https://doi.org/10.1016/j.agwat.2021.107451.

[10] Lundberg, S. M., Lee, S. I. (2017): A unified approach to interpreting model predictions. – Advances in Neural Information Processing Systems 30(NIPS 2017): 4765-4774. https://doi.org/10.48550/arXiv.1705.07874.

[11] Mann, H. B. (1945): Nonparametric tests against trend. – Econometrica 13(3/4): 245-259. https://doi.org/10.2307/1907187.

[12] Ministry of Water Resources of China (2020): 2020 China River Sediment Bulletin. – http://www.mwr.gov.cn/sj/tjgb/zghlnsgb/.

[13] Nijp, J. J., Metselaar, K., Limpens, J., et al. (2017): Including hydrological self-regulating processes in peatland models: effects on peatmoss drought projections. – Science of the Total Environment 580: 1389-1400. https://doi.org/10.1016/j.scitotenv.2016.12.104.

[14] Ren, H., Song, X., Fang, Y., Hou, Z., Scheibe, T. D. (2021): Machine learning analysis of hydrologic exchange flows and transit time distributions in a large regulated river. – Frontiers in Artificial Intelligence 4: 648071. https://doi.org/10.3389/frai.2021.648071.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918

[15] Shukla, S., Wood, A. W. (2008): Use of a standardized runoff index for characterizing hydrologic drought. – Geophysical Research Letters 35(2): L02405. https://doi.org/10.1029/2007GL032487.
[16] Swami, A., Jain, R. (2013): Scikit-learn: machine learning in Python. – Journal of Machine Learning Research 12(10): 2825-2830. https://doi.org/10.1524/auto.2011.0951.
[17] Vicente-Serrano, S. M., Beguería, S., López-Moreno, J. I. (2010): Standardized precipitation evapotranspiration index (SPEI) for drought monitoring. – Journal of Climate 23(7): 1696-1718. https://doi.org/10.1175/2009JCLI2909.1.
[18] Wang, W., Yin, S., Gao, G., Papalexiou, S. M., Wang, Z. (2022): Increasing trends in rainfall erosivity in the Yellow River basin from 1971 to 2020. – Journal of Hydrology 610: 127851. https://doi.org/10.1016/j.jhydrol.2022.127851.
[19] Xiong, Z. W., Xia, J. Q., Wang, Z. H., Li, T., Zhang, J. H. (2019): Whole-processes modeling of flow movement and sediment transport during the period of water-sediment regulation in Xiaolangdi Reservoir. – Scientia Sinica Technologica 49(4): 419-432. https://doi.org/10.1360/N092017-00295.
[20] Xu, S., Yu, Z., Yang, C., Ji, X., Zhang, K. (2018): Trends in evapotranspiration and their responses to climate change and vegetation greening over the upper reaches of the Yellow River Basin. – Agricultural and Forest Meteorology 263: 118-129. https://doi.org/10.1016/j.agrformet.2018.08.010.
[21] Yu, J., Xiao, R., Liang, M., et al. (2024): Hydrological drought assessment of the Yellow River Basin based on non-stationary model. – Journal of Hydrology: Regional Studies 56: 101974. https://doi.org/10.1016/j.ejrh.2024.101974.
[22] Zhang, R., Chen, X., Zhang, Z., et al. (2015): Evolution of hydrological drought under the regulation of two reservoirs in the headwater basin of the Huaihe River, China. – Stochastic Environmental Research and Risk Assessment 29(2): 487-499. https://doi.org/10.1007/s00477-014-0987-z.
[23] Zhang, X., Wu, Z., Wang, H., Yu, Z., Chen, Y. (2025): A hydrological drought risk assessment method based on a four-dimensional Copula function model integrating development and recovery speed characteristics. – Ecological Indicators 177: 113751. https://doi.org/10.1016/j.ecolind.2025.113751.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/2402_18991918