

Earth and Environmental Sciences

Special Topic: Climate Change Impacts and Adaptation

Development of a random-forest-copula-factorial analysis (RFCFA) method for predicting propagation between meteorological and hydrological droughtHao Wang¹, Yongping Li^{1,2*}, Guohe Huang^{1,2}, Quan Zhang¹, Yuan Ma¹ & Yangfeng Li¹¹State Key Joint Laboratory of Environmental Simulation and Pollution Control, School of Environment, Beijing Normal University, Beijing 100875, China;²Institute for Energy, Environment and Sustainable Communities, University of Regina, Regina Sask S4S 0A2, Canada*Corresponding author (email: yongping.li@iseis.org (Yongping Li))

Received 31 March 2023; Revised 20 November 2023; Accepted 29 November 2023; Published online 5 January 2024

Abstract: In the context of global climate warming, the propagation of meteorological drought (MD) may aggravate the devastating impact of hydrological drought (HD) on water security and sustainable development. There are challenges in accurately predicting the propagation of drought and effectively quantifying the effects of uncertainty, especially in data-deficient regions. In this study, a novel method called RFCFA is developed through integrating random forest (RF), copula, and factorial analysis (FA) into a general framework as well as applied to the Aral Sea Basin (a typical arid and data-scarce basin in Central Asia) under considering the impact of climate change. Several findings can be summarized: (1) the projected future drought propagation probability of ASB is 39.2%, which is about 8% higher than historical level; (2) drought propagation is mainly affected by mean climate condition, catchment characteristics (i.e., elevation, LUCC, and slope), and human activities (i.e., irrigation and reservoir operation); (3) the lower propagation probability in spring is expected under SSP1-2.6 due to increased snow meltwater, and the drought propagation probability in autumn is the highest (reaching 45.4%) under the influence of reservoir operation; (4) the combined effects of meteorological conditions and agricultural irrigation can lead to a higher probability of future propagation in the upper river basin in summer. Findings are valuable for predicting drought propagation risk, revealing main factors and inherent uncertainties, as well as providing support for drought management and disaster prevention.

Keywords: drought propagation, random forest, copula, factorial analysis, climate change, Aral Sea Basin**INTRODUCTION****Motivation**

Drought, characterized by prolonged periods of abnormal water deficit, such as reduced precipitation and runoff, has significant and far-reaching impacts on societies, ecosystems, and economies worldwide. Over the past century, major drought events have resulted in millions of deaths and huge economic losses. From 1970 to 2019, approximately 650,000 deaths were attributed to drought, while global economic losses reached around \$124 billion between 1998 and 2017 [1]. The risk of drought is expected to increase due to

climate change, particularly in arid and semi-arid regions like Central Asia, which already face water scarcity [2,3]. Extreme disasters (e.g., drought) triggered by climate change can exacerbate water stress and hinder the sustainable development of these regions [4–6]. However, the impacts of global warming on different types of droughts and their propagation are not fully understood. Drought poses increasing challenges in the context of global warming, and it is critical to develop more robust methods to reveal drought propagation mechanisms and mitigate the adverse effects of drought hazards.

Literature review

The process of drought development, known as drought propagation, involves the progression of water deficit [7]. Meteorological drought (MD), characterized by a precipitation deficit during a particular period, propagates through the various components of the hydrological cycle, leading to soil moisture depletion and ultimately resulting in hydrological drought (HD) [8–10]. Drought assessment research relies on various drought indices, such as the standardized precipitation index (SPI) and the standardized streamflow index (SRI) [11–14]. Drought indices derived from the global climate models (GCMs) dataset can serve as valuable tools for assessing the impact of global warming on different types of droughts and their propagation [15–19]. For example, Oguntunde *et al.* [16] investigated drought propagation and found a time lag of 2–3 months for MD drought. Wu *et al.* [17] developed an ensemble hydrological prediction framework using multiple GCMs to study global drought under specific warming scenarios, revealing the potential for increased severity in drought propagation [20].

In recent years, the introduction of copula in drought research has advanced the field of probabilistic drought modeling [21–32]. For instance, Yang *et al.* [21] developed a Bayesian copula method using multiple GCMs to generate reliable ensemble projections of drought risk. Yue *et al.* [22] integrated copula and Bayesian network techniques to investigate the spatiotemporal dynamics and meteorological triggering conditions of hydrological drought. Farrokhi *et al.* [29] proposed a methodology that combined copula and machine learning algorithms to model the multivariate dependence of drought characteristics, providing comprehensive insights into future droughts. Jiang *et al.* [27] developed a copula-based machine learning method for calculating the propagation of drought from meteorological to ecological droughts, offering a detailed understanding of the propagation process.

Despite extensive research, the assessment of drought propagation between MD and HD is still an emerging field. In regions with limited data availability and sparse observation networks, physically based hydrological models (e.g., SWAT) face challenges in utilizing mechanistic simulation of hydrological processes and water resources management [33]. In such cases, data-driven hydrological models (e.g., random forest, RF) offer a suitable alternative due to their flexibility, adaptability, and fast execution speed [34–36]. The results of drought propagation between MD and HD are subject to various uncertainties arising from multiple sources, such as GCM, future scenario, and MD condition. Hydrological models and copula are unable to quantify the individual and interactive effects of these uncertainty sources [37]. Factorial analysis (FA), an efficient statistical tool, can be utilized to identify the dominant uncertainty source that has the most significant impact on the system response [38,39]. For instance, Duan *et al.* [40] developed a factorial-analysis-based method to investigate the population exposure to drought over the Pearl River Basin under climate change, revealing the GCM and RCP are the major uncertainty sources of drought exposure in

different periods.

Research gap

Although a number of research were reported on assessing drought and drought propagation, there are still difficulties in accurately predicting the propagation of drought and effectively quantifying the effects of uncertainty, especially in data-deficient regions. In watersheds with limited observed datasets, it is challenging to obtain streamflow data from physically based hydrological models for studying drought propagation between MD and HD. Both RF and physically based hydrological models (e.g., SWAT) lack the capability to quantify the individual and interactive effects of factors on drought propagation results. FA can effectively reveal the interactive effects of multiple factors through calculating the variations in simulation responses under different combinations of factors' levels and has been widely used in experiments and hydrological modeling; however, no previous study was conducted to reflect the effects of multiple factors and their interactions on drought propagation through integrating RF, copula, and FA into a general framework.

OBJECTIVE AND CONTRIBUTION

The objective of this study is to develop a random-forest-copula-factorial analysis (RFCFA) method for predicting the drought propagation between MD and HD and quantifying their uncertainties in data-scarce watersheds (Figure 1). The main novelty and contribution of this study are: (1) This is the first attempt to develop such an integrated RFCFA method through incorporating RF, copula, and FA within a general framework. (2) RFCFA has an advantage in predicting drought propagation (between MD and HD) under climate change and analyzing the effects of multiple factors and their interactions on propagation probabilities. (3) RFCFA is firstly applied to the Aral Sea Basin (a typical arid and data-scarce basin in Central Asia) for validating its feasibility and capability, and the obtained results are valuable for predicting drought propagation risk, revealing main factors and inherent uncertainties, as well as providing support for drought management and disaster prevention.

RESULTS AND DISCUSSION

Performance of RF and SWAT

Figure 2 shows the preference of observed and simulated streamflow from SWAT and RF for ASB. For upper ADRB (SDRB), the mean values of NSE for calibration are 0.82 (0.78) and 0.92 (0.87) for SWAT and RF, respectively. Similarly, for validation, the mean values are 0.80 (0.69) and 0.76 (0.67). It demonstrates that the RF and SWAT are feasible for the streamflow simulation in upper ADRB.

The effectiveness of RF models for five sub-basins of ASB (upper/lower ADRB, and upper/middle/lower SDRB) is also presented in Figure S2 and Table S2. For validation, RF performs better for ADRB compared with SDRB, with average NSE values of 0.75 ± 0.03 , 0.75 ± 0.01 , 0.68 ± 0.02 , 0.68 ± 0.02 , and 0.68 ± 0.01 for

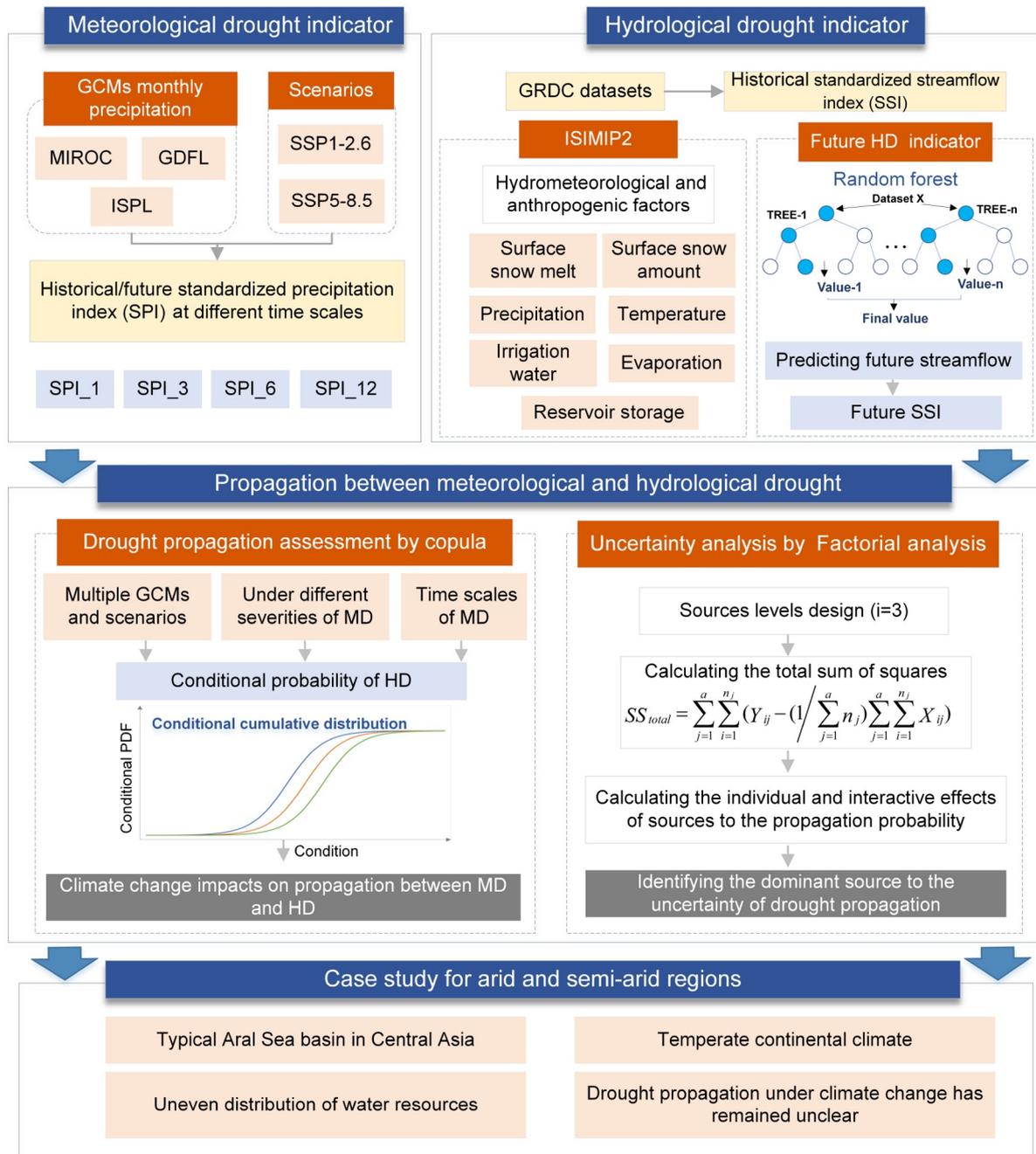


Figure 1 The framework of RFCFA method.

each sub-basin, respectively. The satisfactory performance of RF suggests that it effectively captures the relationship between streamflow and its predictors, including snow-related, meteorological, and anthropogenic factors. Additionally, the parameters SMFMX and SFTMP, which are related to snow melt, are found to be sensitive for the SWAT model (Figures S5 and S6). The simulation results from SWAT and RF can be used as a reference for adjusting SWAT parameters or selecting modeling factors for RF.

The degradation in performance of SWAT for validation is considerably lower compared with RF. Given

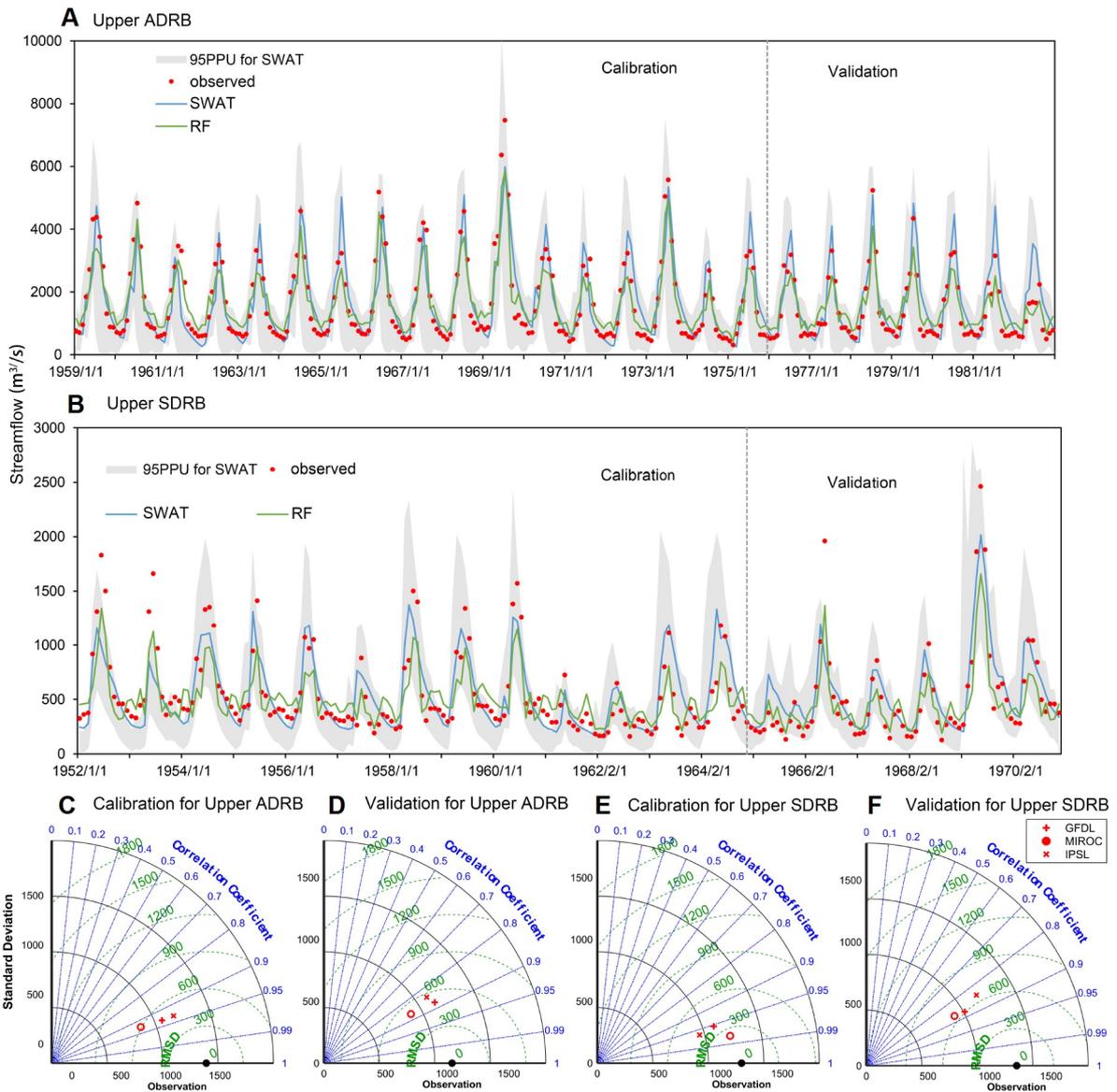


Figure 2 Comparison of simulated streamflow of RF and SWAT for upper Amu Darya River basin (ADRB) and the Syr Darya River basin (SDRB).

that all types of data sets are adequate, it is expected that the performance of SWAT is superior to RF. Consequently, streamflow prediction based on SWAT is more reliable. This limitation hampers the reservoir module of SWAT in accurately estimating reservoir releases and ultimate streamflow. For the plain areas of middle and lower ASB, the absence of detailed datasets of reservoir operation and pre-defined stream may lead to SWAT losing its ability to simulate streamflow and assess water resources. Data-driven models, such as RF, possess the advantage of flexible modeling, which helps mitigate the problem of data sparsity encountered in streamflow simulation. In addition, for the middle and lower ASB, the reservoir operation data provided by ISIMIP (although it does not meet the input requirements of the SWAT reservoir module) can be incorporated into the RF modeling. Assuming that the impact of the reservoir on streamflow has

remained relatively unchanged, RF can be considered a viable alternative method for hydrological simulation and can provide valuable data support for analyzing HD.

HD and MD under climate change

Figure S7 presents that, in the spring, the correlation between snowmelt and runoff is stronger in the upper ADRB ($R=0.69$) compared with the upper SDRB ($R=0.55$). This indicates that snowmelt has a more direct and consistent impact on runoff in the upper ADRB region. In other seasons, the correlation between the snowmelt index and the runoff index is stronger overall. The R value for the upper ADRB is 0.72, indicating a closer relationship, while the R value for the upper SDRB is 0.34, suggesting a weaker correlation. By examining the probability density function diagram, we can observe that there is some overlap between the SSI for runoff and SSI for snowmelt distributions in both the upper ADRB and upper SDRB, but there are also noticeable differences. This suggests that although snowmelt affects runoff in both regions, the specific pattern of these effects may vary. Figure 3 and Table S5 present the historical and future trends of indicators for MD and HD. In the historical period, the decreases of short-term (i.e., time scale n of one or three months) SPI_n and SSI series were mostly not significant for each sub-basin. However, the historical relatively long-term (i.e., time scale n of six or one year) SPI_n series showed a positive trend for lower ADRB and the entire SDRB, with average values of 0.0043/month, 0.0034/month, 0.0036/month, and 0.0032/month, respectively. This indicates a gradual increase in wet meteorological conditions for these regions over time. Compared with the historical results, the future SPI_n series would be projected to have mostly accelerated growth trends under SSP5-8.5, with average values of 0.0038/month, 0.0006/month, 0.0026/month, 0.0023/month, and 0.0024/month for the five sub-basins (upper ADRB, lower ADRB, upper SDRB, middle SDRB, and lower SDRB). Similar results can be observed for the SSI series, suggesting that both HD and MD will be relieved compared with historical values. Figure 4 illustrates the higher ratios of months when MD and HD occur ($SSI-1$ or $SSP-n < -0.5$) during the historical period. For example, the historical ratio for HD was 34.2%, and the future values are projected to be 31.4% and 30.9% under SSP1-2.6 and SSP5-8.5 for ASB, respectively. This implies that the occurrence of meteorological and HD will be less frequent in the future. The results align with previous research and can be attributed to increased precipitation due to climate change [17]. However, adapting to the drought hazards remains a primary ecological and social challenge for ASB.

Drought propagation and impact factors

Figure 5 presents the varying annual propagation between MD and HD. In the historical period, the propagation probabilities for five sub-basins were 32.5%, 32.3%, 31.6%, 32.5%, and 32.3%, respectively. As the MD time scale grew, the propagation probabilities decreased by 1.5% and 1.8% for ADRB and SDRB, respectively. This indicates that short-term (one month) MD had slightly higher effects on HD risk than long-term (12 months) ones in the historical period. As the MD severities developed, the propagation between MD and HD increased by 1.8% and 1.5% for ADRB and SDRB, respectively. Overall, the difference in HD conditional probabilities among sub-basins was not distinct, suggesting relatively homogeneous drought propagation between MD and HD in the historical period for the Aral Sea Basin. The future mean propagation probabilities would be 38.1% and 40.4% under SSP1-2.6 for ADRB and SDRB, respectively.

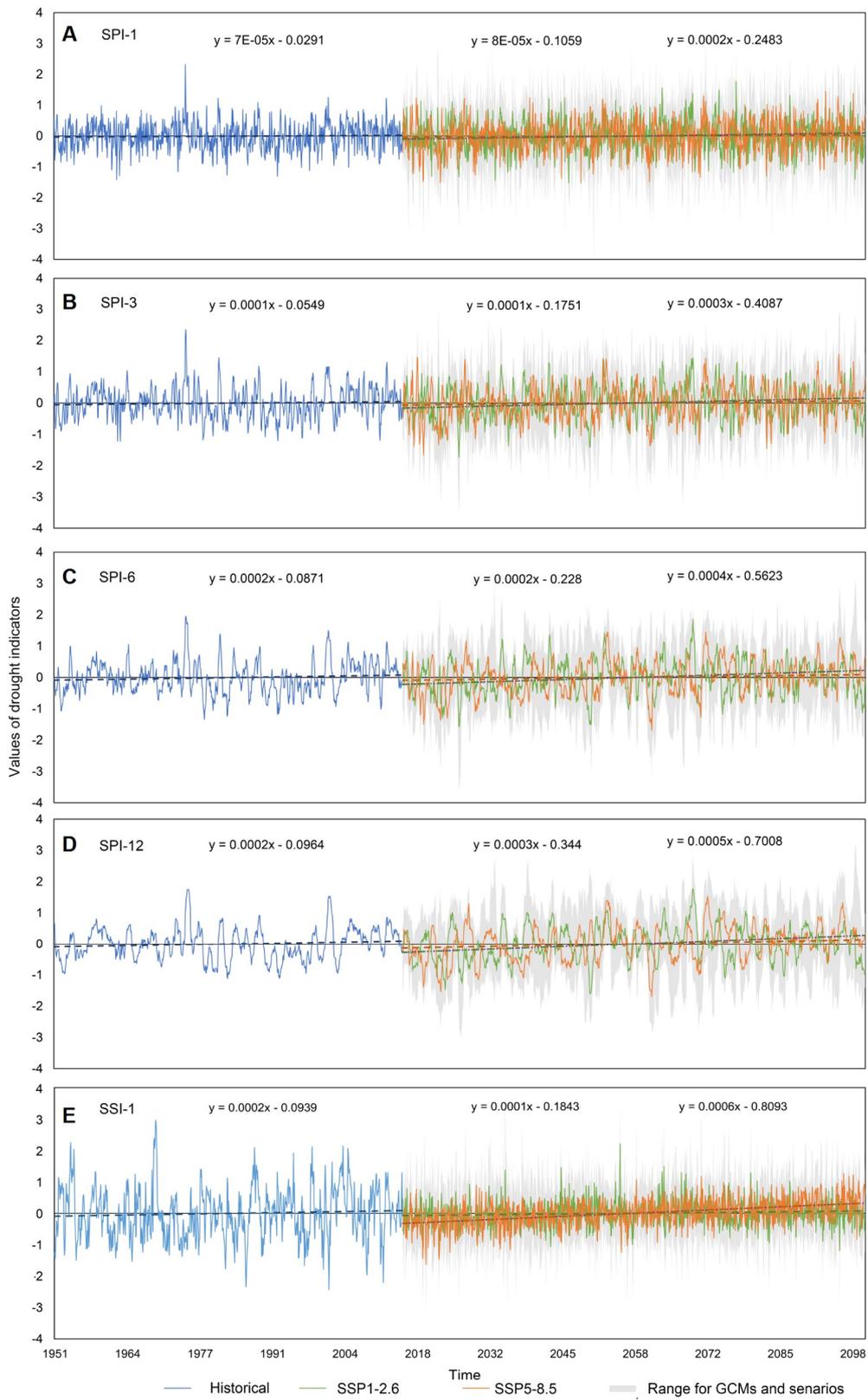


Figure 3 Temporal variation in meteorological and hydrological droughts at monthly scales for Aral Sea Basin during 1950–2099.

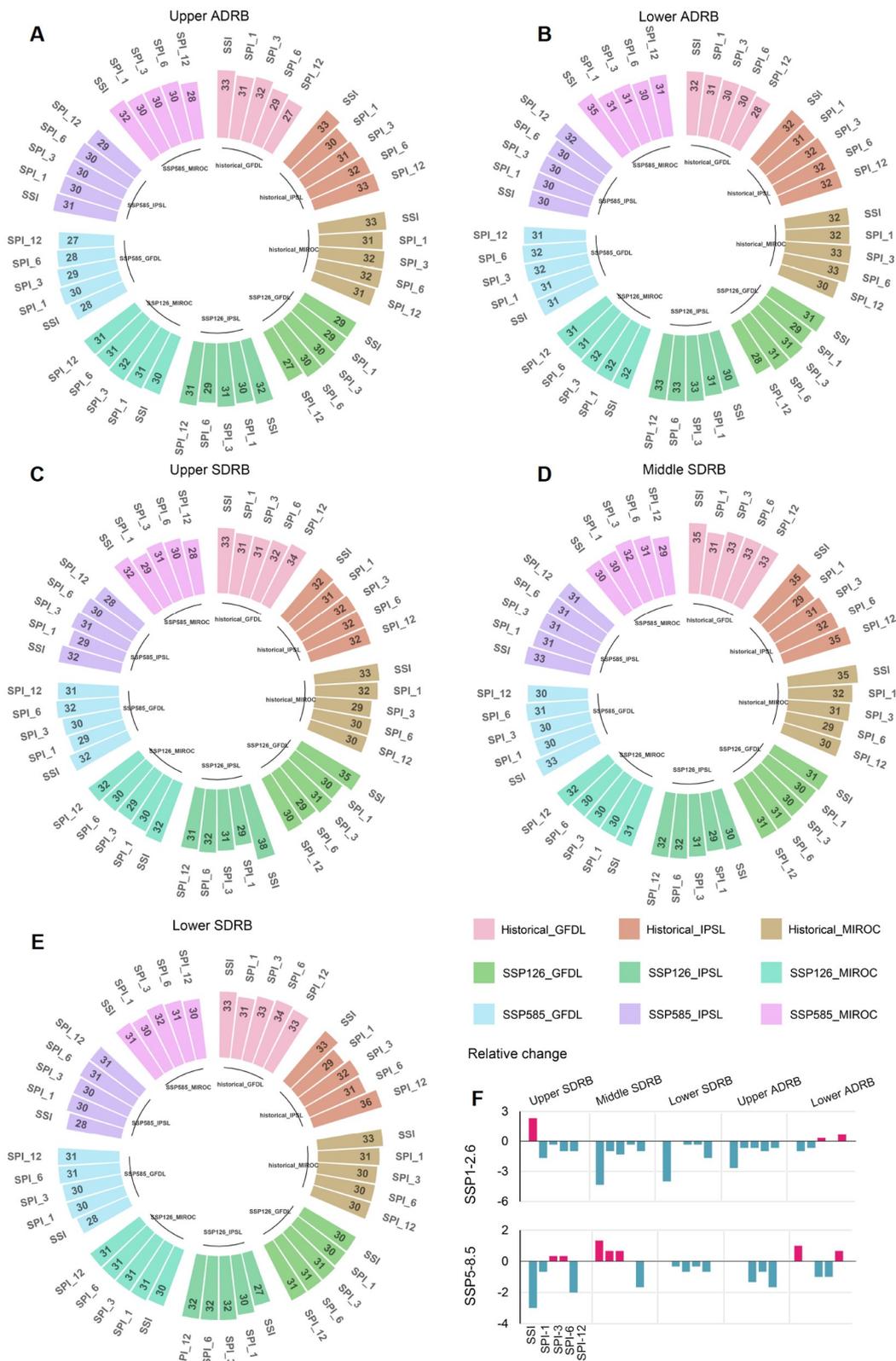


Figure 4 The ratios of the months when droughts occur during 1950–2099 for Amu Darya River basin (A, B) and the Syr Darya River basin (C–E), and their relative change between future and historical periods (F).

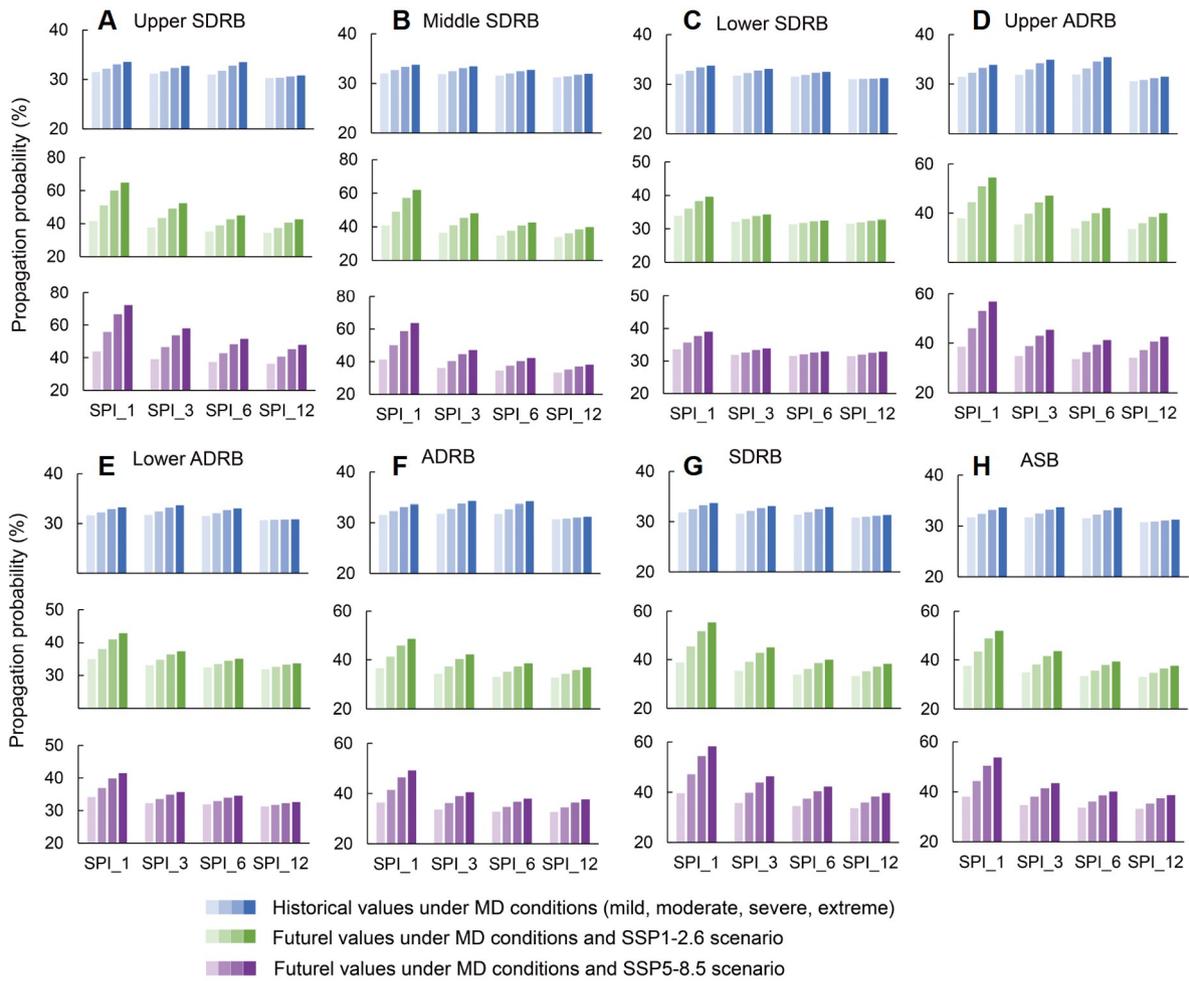


Figure 5 Comparison of annual propagation probability for Aral Sea Basin (H) and its sub-basins (A–G).

Meanwhile, the future propagation between MD and HD would be higher than historical ones, with an average increase of 5.6% and 9.7% under SSP5-8.5 for ADRB and SDRB, respectively. It is noteworthy that this result is consistent with the previous study and revealed the intensification of drought propagation between MD and HD under climate change [17]. The HD conditional probabilities would decrease with the MD time scale growing, especially for the upper and middle sub-basins. For example, the propagation between MD and HD would decrease by 10.1% and 17.2% under SSP5-8.5 for upper ADRB and upper SDRB, respectively. This result revealed that the risks from HD and short-term scale MD had strong linkages for upper and middle basins. The higher probability of HD would occur synchronously with the aggravation of MD severities, with increases of 7.5% and 10.1% under the SSP5-8.5 for ADRB and SDRB, respectively. These findings not only enhance our understanding of drought propagation mechanisms, but also have practical implications for improving and supplementing drought predictions. The variations in propagation probabilities across different spatiotemporal scales (e.g., 3 and 6 months) can assist decision makers in assessing future hydrological drought risks based on the severity of meteorological droughts in the past. This information can support hydrological drought prediction and drought disaster prevention.

Meteorological conditions are typical natural factors affecting drought propagation. Figure 6 shows that, in spring, the higher propagation between MD and HD would be obtained under SSP1-2.6 rather than SSP5-8.5, with the mean values of 33.4%, 31.7%, 43.6%, 39.8%, and 32.7 for five sub-basins. This could be attributed to the fact that the higher temperature projected under SSP5-8.5 would result in an 11% increase in glacier/snow meltwater, which would help alleviate the water shortage experienced in spring for ASB. Overall, the average climate conditions can significantly influence the hydrological cycle by affecting water transport and terrestrial water storage dynamics. Consequently, these factors can impact the propagation of droughts. Previous studies have shown that the propagation from MD to HD occurs slower in dry and continental climates [41]. Similar patterns have been observed in regional studies. In some arid or semiarid regions, relatively low and irregular rainfall leads to extended dry periods and limited water storage. As a result, HD tends to be influenced by long-term precipitation patterns and it should get more attention [7].

Human activities (e.g., reservoir operation) are also important factors affecting drought propagation. Figure 6 also presents that the highest seasonal future drought propagation probabilities would be obtained in autumn, with ADRB and SDRB having percentages of 44.5% and 46.2%, respectively. The seasonal future propagation probabilities are expected to be higher than historical ones, particularly in autumn. Under SSP5-8.5, these probabilities would increase by 3.2%, 9.7%, 14.0%, and 4.5% in the four seasons for ASB. When comparing with the historical results, the greatest decrease in seasonal future propagation between MD and HD is observed in autumn as the MD time scale increases, with values of 12.9% and 14.1% for ADRB and SDRB, respectively. Similarly, the greatest increase in seasonal future propagation between MD and HD is seen in autumn as MD severities worsen, with values of 13.7% and 14.8% for ADRB and SDRB, respectively. These findings indicate that the impacts of future short-term regional meteorological conditions on HD will be significantly amplified in autumn. It has been found that reservoir operation can reduce the frequency and magnitude of drought, indicating positive effects in managing drought [42]. Previous studies have shown

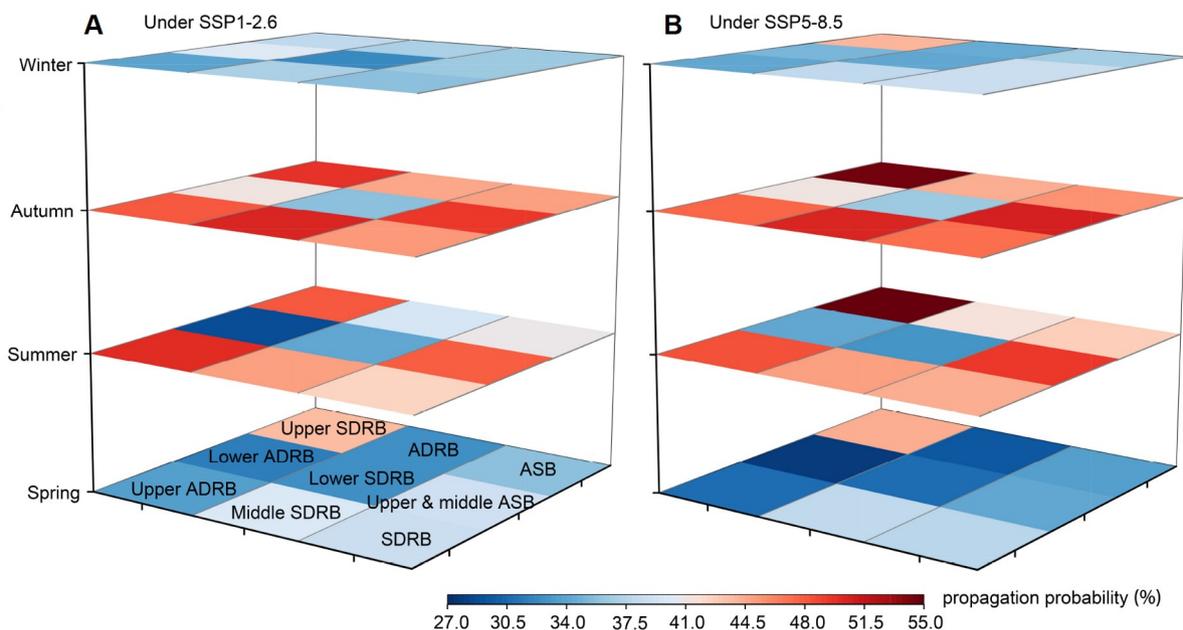


Figure 6 Comparison of seasonal propagation probability under SSP1-2.6 (A) and SSP5-8.5 (B).

that if reservoirs are operated to increase water availability during dry periods in a specific basin or location, the duration and severity of HD can be decreased [43]. During dry periods, if the stored water resources in reservoirs are used by upstream areas for industrial, agricultural, and domestic demands, it can exacerbate water shortage downstream prolong the duration, and intensify the severity of HD [44]. Therefore, it is crucial to pay more attention to meteorological/hydrological monitoring of major rivers during autumn. Additionally, the determination of the reservoir storage period should take these results into consideration.

Catchment characteristics, including slope, elevation, Land Use and Land Cover Change (LUCC), and land use, play an important role in the propagation of HD. The runoff coefficient is a key factor that characterizes these characteristics. Figure 7 demonstrates that the average annual runoff coefficient in the upper ASB is significantly higher (0.49) compared with the middle and lower ASB (0.14) under SSP5-8.5. This difference can be attributed to the mountainous areas in the upper ASB, which receive abundant precipitation and have

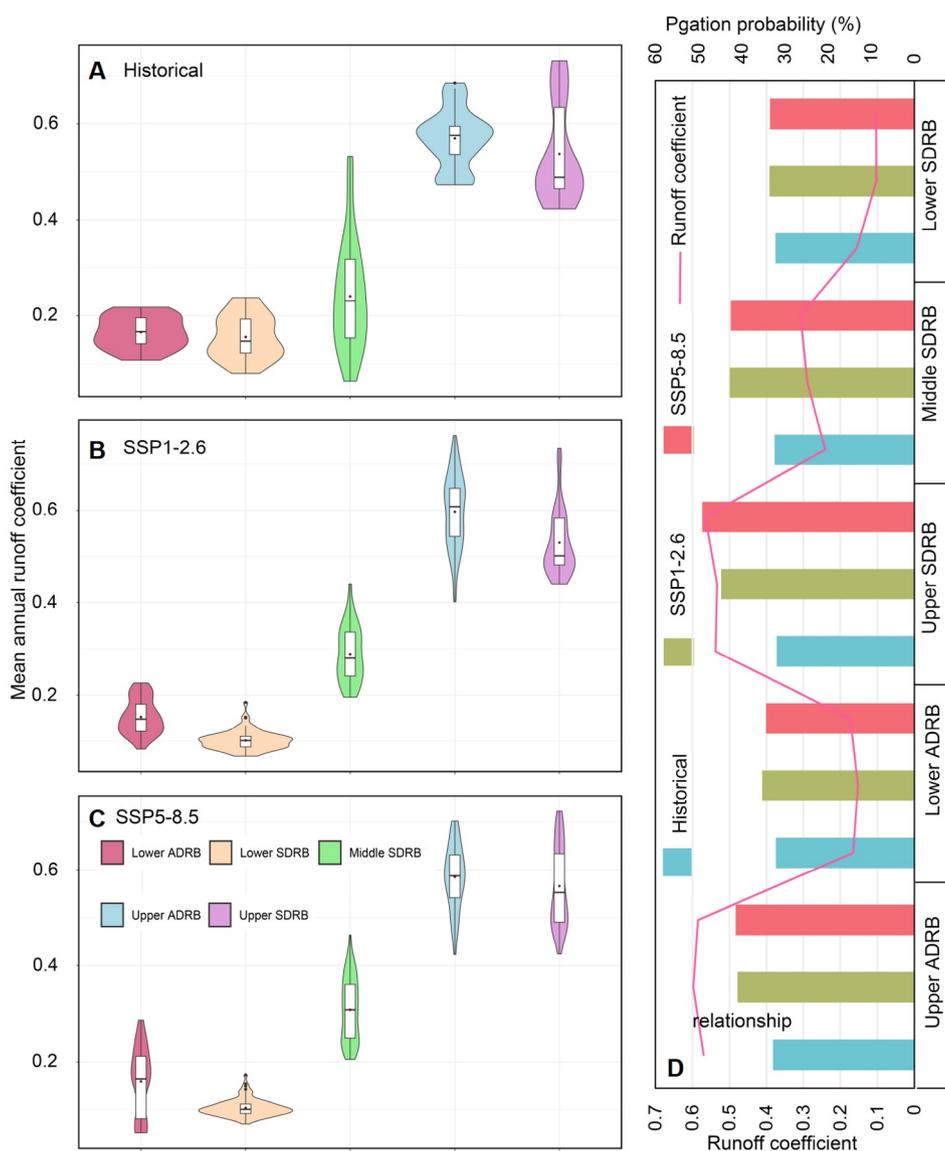


Figure 7 Mean runoff coefficient for different periods (A–C) and its relationship (D) with propagation probability.

steep terrain, facilitating the generation and confluence of runoff. On the other hand, the middle and lower ASB have flat terrain and limited precipitation, resulting in a scarcity of runoff and significant dissipation. Additionally, Figure 7 depicts the relationship between the mean annual runoff coefficient and propagation probability. It is observed that both variables exhibit similar variation patterns across different geographical locations. This pattern is influenced not only by elevation and terrain characteristics, as mentioned earlier, but also by the impact of LUCC on drought propagation. In the mountainous areas of upper ASB, there are extensive forests, while the lower sub-basins primarily consist of grassland and desert. The type of land cover plays a role in the propagation of drought by affecting water dissipation. For example, increased vegetation can lead to enhanced rainfall interception and plant leaf transpiration, which can reduce river flow. Changes in vegetation also impact the partitioning of precipitation or water deficit into surface water, which in turn affects water resource generation. The effect of vegetation coverage on meteorological and HD depends on the type of vegetation, with forests consuming more water compared with grass or shrubs. Previous studies have highlighted the influence of vegetation changes on the relationship between MD and HD [45]. Therefore, it is important to improve drought warning that conducting meteorological and hydrological observations in the alpine forest coverage area within the runoff-producing region of the river basin.

Irrigation is also a significant factor that influences drought propagation. Within the five sub-basins, noticeable geographic variations in future propagation between MD and HD are observed across the four seasons, particularly in summer and autumn (Figure 6). For instance, during summer, the propagation between MD and HD for upper ADRB would be 1.6 times higher (50.2%) compared with the lower ADRB (32.7%), while the values for upper SDRB (51.5%) would be 1.5 times higher than the lower SDRB (34.6%). This can be attributed to the high evapotranspiration rates in summer, which result in significant water resource consumption. Additionally, there is a greater demand for irrigation water in the middle of the growing season. In other words, the combination of meteorological conditions and agricultural irrigation contributes to the relatively higher future propagation between MD and HD in the upper sub-basins during summer. Irrigation from groundwater withdrawals can affect groundwater levels and soil moisture content through irrigation return flows, thereby influencing drought propagation [46]. Previous studies have explored the aggregated hydrological drought resulting from irrigation at regional and global scales [47]. Incorporating human activities into hydrological model simulations is an important approach to quantifying the impacts of irrigation or reservoir operation. However, previous studies on irrigation impacts on drought propagation have been limited in their quantitative evaluation of the impacts due to the low accessibility of irrigation data. They are often qualitative or combined with other human activities. Thus, policymakers should promptly devise effective measures to combat drought in response to these findings.

Uncertainty analysis

Figure 8A–H presents the FA results in future periods (2015s to 2099s). For upper ADRB, the primary uncertainty source comes from GCM, with contribution rates of 54%. For upper and middle SDRB, the main source is time scales of MD (TSM) with contribution rates of 36% and 54%, respectively. For lower ADRB and SDRB, the interactions dominate the uncertainty of drought propagation between MD and HD, with contribution percentages of 66% and 58%. In contrast, the future scenario (FS) is the minor uncertainty source for all basins, with contribution percentages of less than 5%. The results indicate that the short-term

MD effects of the year make a greater impact on annual HD risk than the long-term effects of FSs. Meanwhile, the significant effect of interactions on uncertainty demonstrates the anthropic, climatic, and hydrological factors are associated closely with the drought risk propagation for the lower Aral Sea Basin, and all these factors should be considered in drought risk assessment research.

Figure 8A–H provides the seasonal uncertainty analysis results. The main uncertainty sources for most sub-basins of ADRB and SDRB vary in different seasons, implying that the impacts of factors on uncertainty are complex in a year. For the middle SDRB, TSM is the primary source in four seasons with an average contribution rate of 42%. This result illustrates the short-term precedent that MD has a major influence on HD risk. It is noteworthy that GCM is always the prominent uncertainty source for upper ADRB, with the contribution of 56%, 30%, 26%, 38%, and 76% to the annual and seasonal results, respectively. Generally, GCM is the important uncertainty source for ASB at annual and seasonal scales, with contribution ratios of 19%, 21%, 36%, 16%, and 36%, respectively. This result demonstrates that climate change has significant impacts on the uncertainty of drought propagation between MD and HD, and the impacts would be highest in the summer. GCM was found as the dominant uncertainty source in the previous studies, and it highlights the importance of ensemble projection based on multiple GCMs datasets for the research on the projection of drought propagation under climate change [48–50]. Severitie of MD (SM) is a notable uncertainty source for the Aral Sea Basin, especially in upper SDRB with an average contribution rate of 26%. This finding indicates that regional meteorological conditions would exert a markable influence on HD risk for upper SDRB. Being similar to the annual results, the interaction is a noteworthy uncertainty source, with average contribution rates of 22%, 33%, 25%, 24%, and 44% for five sub-basins, respectively. Figure 8I–K presents the complex response of drought propagation between MD and HD to the interaction between FS and SM, FS, and GCM, as well as FS and TSM. For instance, propagation between MD and HD would get their peak value of 45% when FS is SSP5-8.5 and SM is at the extreme level. This indicates the interactions between factors have effect impacts on propagation probability.

CONCLUSIONS

In this study, a random-forest-copula-factorial analysis (RFCFA) method has been developed to predict the drought propagation characteristics and quantify uncertainty in data-scarce watersheds. RFCFA incorporates RF, copula function, and FA within a general framework. RFCFA is applied to ASB for various seasons to unveil the mechanism of drought propagation and the impact of multiple uncertainty sources.

The main findings are as follows: (1) RF is robust for streamflow simulation and a suitable alternative method for SWAT in data-scarce watersheds. (2) The future HD and MD would be relieved under SSP1-2.6 and SSP5-8.5. (3) The future mean annual propagation probability would have an increase of about 8% above historical levels. (4) Drought propagation is mainly affected by mean climate condition, catchment characteristics (i.e., elevation, LUCC, slope), and human activities (i.e., irrigation and reservoir operation). (5) GCM is the important uncertainty source for ASB at annual and seasonal scales.

Overall, this study illustrates that RFCFA is an appropriate method for analyzing drought propagation in data-scarce watersheds under climate change. The findings of this study can enhance understanding of drought propagation mechanism and provide support for future drought warning and mitigation strategies.

However, several limitations exist in this study including: (1) The impact of data uncertainty resulting from the limited number of samples in the observed GRDC data is not investigated. (2) The lack of publicly available data sets on human activities (e.g., reservoir operation, irrigation, and LUCC) hinders the quantification of their role in drought propagation. (3) The physical mechanisms involved in the propagation of drought remain unclear.

METHODOLOGY

Development of RFCFA

The framework of RFCFA is presented in Figure 1. Two drought indices (SPI and SSI) are used to identify MD and HD. RF is adopted to simulate and predict the streamflow. Copula is employed to model the dependence structure between MD and HD. FA is employed to quantify the uncertainty of the drought propagation arising from multiple sources.

RF is an ensemble technique that utilizes multiple decision trees trained through bootstrap aggregating [51]. RF offers the advantage of generating reasonable predictions without requiring hyper-parameter tuning and mitigating overfitting issues commonly observed in decision trees [52–54]. To ensure the universality of RF models, the historical datasets, including observed streamflow datasets of global runoff data center (GRDC), GCMs datasets from Coupled Model Intercomparison Project Phase 6 (CMIP6) and, Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) are divided into two subsets by random sampling: The subset containing 70% of the data is used to calibrate the model, and the subset containing the remaining 30% data is used for validation. The predictors include two snow-related factors (i.e., surface snow melt, surface snow amount), three meteorological factors (i.e., near-surface air temperature, precipitation, and evaporation), and two anthropogenic (i.e., reservoir storage, potential irrigation water withdrawal) (Figure S1). The hydrological data utilizes GRDC observed datasets, and the datasets of predictors utilize the time series of raster averages within the shapefile data range for each sub-basin.

RF is trained for five sub-basins separately. The trained models are then used to predict future streamflow (from 2015 to 2099) under SSP1-2.6 and SSP5-8.5. Finally, the SSI series are calculated from historical and simulated streamflow results. To assess the effectiveness of RF for hydrological simulation in ASB and compare its performance with physically based hydrological models (such as SWAT) in data-scarce watersheds, this study conducted hydrological simulations for upper ASB using both SWAT and RF models. More detailed information can be found in Figures S2 and S3, and Tables S3 and S4.

The indicator of meteorological drought that the SPI has the advantage of being applied on different timescales (n), which allows for the consideration of past impacts of drought in calculation. Building on the concept of SPI, the SSI is introduced as an indicator of hydrological drought [55]. The gamma probability density function is a suitable model for describing the distribution of water cycle variables, including streamflow and precipitation [56]:

$$g(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, x > 0 \quad (1)$$

where $\Gamma(\alpha)$ is the gamma function; x represents the accumulation of a variable. The SSI (timescale n is 1 month) and SPI_n (timescale $n = 1, 3, 6, 12$ months) are introduced to assess the drought propagation between

MD and HD. The selection of time scales for meteorological and hydrological drought Indexes refers to previous studies [57–59]. This strategy is useful for studying drought propagation at different lag times based on meteorological drought indices at different time scales. The drought indicators that each SSI and SPI_n series are classified into four classes: mild drought ($-0.5 < SPI_n \leq -1$), moderate drought ($-1.5 < SPI_n \leq -1$), severe drought ($-2 < SPI_n \leq -1.5$), and extreme drought ($SPI_n \leq -2$) [60]. Similarly, an SSI-1 value less than -0.5 is considered a signal for HD.

The dependence structure of each SSI and its SPI_n monthly time series is modeled by the copula functions. Let SSI and SPI_n be two discrete random variables with marginal distributions $F_{SPI_n}(spi)$ and $F_{SSI}(ssi)$, respectively, on the interval $[0,1]$, their joint probability distribution $F_{SPI_n,SSI}(spi, ssi)$ satisfies the following copula function C in Eq. (2), as follows [61]:

$$P(SPI_n \leq spi, SSI \leq ssi) = F_{SPI_n,SSI}(spi, ssi) = C[F_{SPI_n}(spi), F_{SSI}(ssi)], \quad (2)$$

where $F_{SPI_n}(spi)$ and $F_{SSI}(ssi)$ are cumulative distribution functions (CDFs) for bivariate variables SPI-1 and SSI-1 (Figure S8).

Six copula functions from two copula families, including the elliptical copula and the Archimedean copula, are selected as the candidate model for calculating the joint distribution of SPI_n and SSI (Table S2). The bivariate copula $C(u, v)$ function is expressed as

$$C(u, v) = \varphi^{-1}(\varphi(u), \varphi(v)), \quad (3)$$

where $C(u, v)$ is the copula function, φ denotes the convex function, and u and v represent the SPI_n and SSI. According to the joint distribution formulated by the best-fit copula functions, the conditional probabilities of HD occurrences under different severities (mild, moderate, severe, and extreme) and time scales of MD can be formulated respectively as follows:

$$\begin{aligned} P(SSI < -0.5 \mid -1 < SPI_n \leq -0.5) &= \frac{P(-1 < SPI_n \leq -0.5, SSI < -0.5)}{P(-1 < SPI_n \leq -0.5)} \\ &= \frac{F_{SPI_n,SSI}(-0.5, -0.5) - F_{SPI_n,SSI}(-1, -0.5)}{F_{SPI_n}(-0.5) - F_{SPI_n}(-1)}, \end{aligned} \quad (4)$$

$$\begin{aligned} P(SSI < -0.5 \mid -1.5 < SPI_n \leq -1) &= \frac{P(-1.5 < SPI_n \leq -1, SSI < -0.5)}{P(-1.5 < SPI_n \leq -1)} \\ &= \frac{F_{SPI_n,SSI}(-1, -0.5) - F_{SPI_n,SSI}(-1.5, -0.5)}{F_{SPI_n}(-1) - F_{SPI_n}(-1.5)}, \end{aligned} \quad (5)$$

$$\begin{aligned} P(SSI < -0.5 \mid -2 < SPI_n \leq -1.5) &= \frac{P(-2 < SPI_n \leq -1.5, SSI < -0.5)}{P(-2 < SPI_n \leq -1.5)} \\ &= \frac{F_{SPI_n,SSI}(-1.5, -0.5) - F_{SPI_n,SSI}(-2, -0.5)}{F_{SPI_n}(-1.5) - F_{SPI_n}(-2)}, \end{aligned} \quad (6)$$

$$\begin{aligned} P(SSI < -0.5 \mid SPI_n \leq -2) &= \frac{P(SPI_n \leq -2, SSI < -0.5)}{P(SPI_n \leq -2)} \\ &= \frac{F_{SPI_n,SSI}(-2, -0.5)}{F_{SPI_n}(-2)}. \end{aligned} \quad (7)$$

The conditional probabilities of HD under the given MD drought severities and time scales are defined as

the drought propagation between MD and HD. To estimate the uncertainty of drought propagation between MD and HD, modeling chains are analyzed by factorial analysis, which consists of two FSs, three GCMs, four SMs, and four TSMs. Based on many experimental samples, the change in the effect of drought propagation $\Delta y_{i,j,k,l}$ is assumed to follow the following model [62]:

$$\Delta y_{i,j,k,l} = \mu + F_i + G_j + S_k + T_l + I_{i,j,k,l} \quad (8)$$

where μ represents the mean change of the propagation probabilities; F_i , G_j , S_k , and T_l represent the effects on the drought propagation between MD and HD of the i th FS, the j th GCM, the k th SM, and l th TSM, respectively; $I_{i,j,k,l}$ represents the sum of the effects due to the interactions between different sources. The total variance (overall uncertainty, VT) can be decomposed into contributions from different sources as follows:

$$VT = VF + VG + VS + VT_{FG} + VI_{FS} + VI_{FT} + VI_{GS} + VI_{GT} + VI_{ST} + VI_{FGS} + VI_{FGT} + VI_{FST} + VI_{GST} + VI_{FGST}, \quad (9)$$

where VF, VG, VS, and VT represent the variance contributed by the effects of FSs, GCMs, SMs, and TSMs, respectively; and VI_{FG} , VI_{FS} , VI_{FT} , VI_{GS} , VI_{GT} , VI_{ST} , and VI_{FGST} represent the variance from interaction effects between FS-GCM, FS-SM, FS-TSM, GCM-TSM, SM-TSM, FS-GCM-SM, FS-GCM-TSM, FS-SM-TSM, GCM-SM-TSM, and FS-GCM-SM-TSM, respectively. By dividing the variance from different sources by the total variance, the fractional contributions of different sources to the overall uncertainty can be obtained.

Case study and data

The RFCFA method is used to predict drought propagation in the Aral Sea Basin (ASB), which spans an area (35°9′–47°48′N, 55°35′–77°42′E) of about 1.8 million km² in Kazakhstan, Uzbekistan, Turkmenistan, Kyrgyzstan, and Tajikistan. The ASB has a typical temperate continental climate and a fragile ecosystem influenced by the Siberian High. The main water sources for rivers in ASB are precipitation and glacier/snow meltwater from mountainous regions. Additionally, the annual evaporation in this region is as high as 1700 mm [63]. The ASB is divided into the Amu Darya River basin (ADRB) and the Syr Darya River basin (SDRB), covering areas of 860,000 and 550,000 km², respectively (Figure 1). The annual precipitation in the eastern and southeastern high mountainous regions ranges from 1000 to 3000 mm, while the rainfall on the southern and southwestern slopes of these mountains varies between 600 and 800 mm. In contrast, the lowlands and valleys receive an annual rainfall of 80–200 mm [64]. The annual snowmelt water volume is estimated to be approximately 53.3 km³ for the SDRB and 60.1 km³ for the ADRB. Furthermore, glaciers contribute to about 2% and 8% of the total annual water volume in the SDRB and ADRB, respectively [65]. The proportion of snow/glacier meltwater is approximately 64% and 62% for SDRB and ADRB, respectively [65]. The water supply of the Aral Sea and local agricultural irrigation rely heavily on the runoff from two major sources. Drought disasters have been occurring frequently in ASB [66]. Due to limited observational and predictive data, there has been little research on the propagation between MD and HD for ASB. Therefore, it is crucial to investigate the drought propagation between MD and HD, reveal the main factors affecting drought propagation, as well as analyze the climate change impact. Considering the catchment area, the ASB is classified into five sub-basins, including the upper and lower ADRB, as well as the upper, middle, and lower SDRB.

In this study, we focused on meteorological and hydrological drought by predicting and analyzing SPI and SSI. SPI is calculated from historical and future precipitation datasets from GCMs datasets (Table S1). The historical SSI is calculated from the Global Runoff Data Center (<https://www.bafg.de>), and the future SSI is calculated from the streamflow projected by random forest. To train RF models, seven monthly meteorological, snow-related, and anthropogenic factors are chosen as the predictors to predict streamflow (Table S1). The reason for choosing these factors is mainly based on the runoff-producing mechanism. The meteorological and snow-related factors are derived from GCMs datasets of the Coupled Model Intercomparison Project (CMIP6, <https://esgf-node.llnl.gov/search/cmip6/>) datasets archive [67]. Based on the Global Land Data Assimilation System (GLDAS, <https://disc.gsfc.nasa.gov/>) Noah Land Surface Model L4 V2.0 reanalysis dataset, The GCMs datasets for three variables (temperature, precipitation, and evaporation) are corrected by quantile mapping with the assumption that the relationship will be maintained for the future [68,69]. Meanwhile, anthropogenic factors are obtained from the two datasets (i.e., WaterGAP2 and H08) from Inter-Sectoral Impact Model Intercomparison Project (ISIMIP, <https://www.isimip.org/>) datasets archive [70]. The datasets of CMIP6 and ISIMIP are commonly employed in previous studies, and they are proven to be reliable [71–73]. Three GCMs datasets (i.e., GFDL-CM4, MIROC6, IPSL-CM6A-LR) are chosen in this study, and the reason for selecting them is that the driver datasets of WaterGAP2 and H08 are GFDL-ESM2M, IPSL-CM5A-LR, MIROC5, respectively. Besides, the scenarios for datasets of CMIP6 are SSP1-2.6 and SSP5-8.5, and those for datasets of ISIMIP are RCP1.6 and RCP8.5.

Data availability

The original data are available from corresponding authors upon reasonable request.

Acknowledgements

The authors are grateful to the editors and the anonymous reviewers for their insightful comments and suggestions.

Funding

This work was supported by the Innovative Research Group of the National Natural Science Foundation of China (52221003), the Strategic Priority Research Program of Chinese Academy of Sciences (XDA20060302), and the National Natural Science Foundation of China (52279003 and 52279002).

Author contributions

H.W., Y.L. and G.H. designed the research. H.W. and Q.Z processed the data and calculated the drought indexes. H.W. carried out relevant estimations and analyses. H.W. wrote the manuscript. Y.L., G.H. and Y.L. revised and improved the manuscript.

Conflict of interest

The authors declare no conflict of interest.

Supplementary information

The supporting information is available online at <https://doi.org/10.1360/nso/20230022>. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

- 1 Santini M, Noce S, Antonelli M, *et al.* Complex drought patterns robustly explain global yield loss for major crops. *Sci Rep* 2022; **12**: 5792.
- 2 Su B, Huang J, Fischer T, *et al.* Drought losses in China might double between the 1.5°C and 2.0°C warming. *Proc Natl Acad Sci USA* 2018; **115**: 10600–10605.
- 3 Huang WW, Chen XJ, Fan YR. Management of contaminated drinking water source in rural communities. *J Environ Inform* 2021; **39**: 136–151.
- 4 Schubert SD, Stewart RE, Wang H, *et al.* Global meteorological drought: A synthesis of current understanding with a focus on SST drivers of precipitation deficits. *J Clim* 2016; **29**: 3989–4019.
- 5 Barker LJ, Hannaford J, Chiveron A, *et al.* From meteorological to hydrological drought using standardised indicators. *Hydrol Earth Syst Sci* 2016; **20**: 2483–2505.
- 6 Wan W, Zhao J, Li HY, *et al.* Hydrological drought in the anthropocene: Impacts of local water extraction and reservoir regulation in the U.S. *JGR Atmos* 2017; **122**: 11313–11328.
- 7 Van Loon AF, Stahl K, Di Baldassarre G, *et al.* Drought in a human-modified world: reframing drought definitions, understanding, and analysis approaches. *Hydrol Earth Syst Sci* 2016; **20**: 3631–3650.
- 8 Wu J, Chen X, Love CA, *et al.* Determination of water required to recover from hydrological drought: Perspective from drought propagation and non-standardized indices. *J Hydrol* 2020; **590**: 125227.
- 9 Hellwig J, de Graaf IEM, Weiler M, *et al.* Large-scale assessment of delayed groundwater responses to drought. *Water Resources Res* 2020; **56**: e2019WR025441.
- 10 Van Loon AF. Hydrological drought explained. *WIREs Water* 2015; **2**: 359–392.
- 11 Rajbanshi J, Das S. The variability and teleconnections of meteorological drought in the Indian summer monsoon season: Implications for staple crop production. *J Hydrol* 2021; **603**: 126845.
- 12 Shiru MS, Shahid S, Dewan A, *et al.* Projection of meteorological droughts in Nigeria during growing seasons under climate change scenarios. *Sci Rep* 2020; **10**: 10107.
- 13 Meque A, Abiodun BJ. Simulating the link between ENSO and summer drought in Southern Africa using regional climate models. *Clim Dyn* 2015; **44**: 1881–1900.
- 14 Mondal SK, Huang J, Wang Y, *et al.* Doubling of the population exposed to drought over South Asia: CMIP6 multi-model-based analysis. *Sci Total Environ* 2021; **771**: 145186.
- 15 Ahmed K, Shahid S, Chung ES, *et al.* Climate change uncertainties in seasonal drought severity-area-frequency curves: Case of arid region of Pakistan. *J Hydrol* 2019; **570**: 473–485.
- 16 Oguntunde PG, Abiodun BJ, Lischeid G. Impacts of climate change on hydro-meteorological drought over the Volta Basin, West Africa. *Glob Planet Change* 2017; **155**: 121–132.
- 17 Wu G, Chen J, Shi X, *et al.* Impacts of global climate warming on meteorological and hydrological droughts and their propagations. *Earths Future* 2022; **10**: e2021EF002542.
- 18 Sun CX, Huang GH, Fan Y, *et al.* Drought occurring with hot extremes: Changes under future climate change on loess plateau, China. *Earths Future* 2019; **7**: 587–604.
- 19 Duan K, Mei Y. Comparison of meteorological, hydrological and agricultural drought responses to climate change and uncertainty assessment. *Water Resour Manage* 2014; **28**: 5039–5054.
- 20 Jia Q, Jia H, Li Y, *et al.* Applicability of CMIP5 and CMIP6 models in China: Reproducibility of historical simulation and uncertainty of future projection. *J Clim* 2023; **36**: 5809–5824.
- 21 Yang X, Li YP, Huang GH, *et al.* Development of a multi-GCMs Bayesian copula method for assessing multivariate drought risk under climate change: A case study of the Aral Sea basin. *Catena* 2022; **212**: 106048.
- 22 Yue S, Sheng X, Yang F. Spatiotemporal evolution and meteorological triggering conditions of hydrological drought in the Hun River basin, NE China. *Nat Hazards Earth Syst Sci* 2022; **22**: 995–1014.
- 23 Najib MK, Nurdianti S, Sopaheluwakan A. Multivariate fire risk models using copula regression in Kalimantan, Indonesia. *Nat Hazards* 2022; **113**: 1263–1283.

- 24 Hasan IF, Abdullah R. Agricultural drought characteristics analysis using copula. *Water Resour Manage* 2022; **36**: 5915–5930.
- 25 Van de Vyver H, Van den Bergh J. The Gaussian copula model for the joint deficit index for droughts. *J Hydrol* 2018; **561**: 987–999.
- 26 Sadegh M, Ragno E, AghaKouchak A. Multivariate copula analysis toolbox (MvCAT): Describing dependence and underlying uncertainty using a Bayesian framework. *Water Resources Res* 2017; **53**: 5166–5183.
- 27 Jiang T, Su X, Zhang G, *et al.* Estimating propagation probability from meteorological to ecological droughts using a hybrid machine learning copula method. *Hydrol Earth Syst Sci* 2023; **27**: 559–576.
- 28 Jahannemaei N, Khosravinia P, Sanikhani H, *et al.* Toward analyzing meteorological droughts in western Iran: A multivariate approach based on vine copulas. *Nat Hazards* 2023; **116**: 1903–1929.
- 29 Farrokhi A, Farzin S, Mousavi SF. Meteorological drought analysis in response to climate change conditions, based on combined four-dimensional vine copulas and data mining (VC-DM). *J Hydrol* 2021; **603**: 127135.
- 30 Erhardt TM, Czado C. Standardized drought indices: A novel univariate and multivariate approach. *J R Stat Soc Ser C- Appl Stat* 2018; **67**: 643–664.
- 31 Wang T, Tu X, Singh VP, *et al.* A composite index coupling five key elements of water cycle for drought analysis in Pearl River basin, China. *J Environ Manage* 2022; **320**: 115813.
- 32 Ansell L, Dalla Valle L. Social media integration of flood data: A vine copula-based approach. *J Environ Inform* 2022; **39**: 97–110.
- 33 Yaseen ZM, Awadh SM, Sharafati A, *et al.* Complementary data-intelligence model for river flow simulation. *J Hydrol* 2018; **567**: 180–190.
- 34 Jia QM, Li YP, Li YF, *et al.* Analyzing variation of inflow from the Syr Darya to the Aral Sea: A Bayesian-neural-network-based factorial analysis method. *J Hydrol* 2020; **587**: 124976.
- 35 Gao PP, Li YP, Huang GH, *et al.* An integrated Bayesian least-squares-support-vector-machine factorial-analysis (B-LSVM-FA) method for inferring inflow from the Amu Darya to the Aral Sea under ensemble prediction. *J Hydrol* 2021; **594**: 125909.
- 36 Wang H, Li YP, Liu YR, *et al.* Analyzing streamflow variation in the data-sparse mountainous regions: An integrated CCA-RF-FA framework. *J Hydrol* 2021; **596**: 126056.
- 37 Pan Y, Zeng XK, Gao XY, *et al.* Assessing human health risk to dnaps exposure in bayesian uncertainty analysis. *J Environ Inform* 2021; **39**: 67–80.
- 38 Yin J, Gentine P, Slater L, *et al.* Future socio-ecosystem productivity threatened by compound drought-heatwave events. *Nat Sustain* 2023; **6**: 259–272.
- 39 Su YY, Li YP, Liu YR, *et al.* Development of an integrated PCA-SCA-ANOVA framework for assessing multi-factor effects on water flow: A case study of the Aral Sea. *Catena* 2021; **197**: 104954.
- 40 Duan R, Huang G, Zhou X, *et al.* Ensemble drought exposure projection for multifactorial interactive effects of climate change and population dynamics: Application to the pearl river basin. *Earths Future* 2021; **9**: e2021EF002215.
- 41 Gevaert AI, Veldkamp TIE, Ward PJ. The effect of climate type on timescales of drought propagation in an ensemble of global hydrological models. *Hydrol Earth Syst Sci* 2018; **22**: 4649–4665.
- 42 Zhang R, Chen X, Zhang Z, *et al.* Evolution of hydrological drought under the regulation of two reservoirs in the headwater basin of the Huaihe River, China. *Stoch Environ Res Risk Assess* 2015; **29**: 487–499.
- 43 Wu D, Zhao X, Liang S, *et al.* Time-lag effects of global vegetation responses to climate change. *Glob Change Biol* 2015; **21**: 3520–3531.
- 44 Al-Faraj FAM, Scholz M. Assessment of temporal hydrologic anomalies coupled with drought impact for a transboundary river flow regime: The Diyala watershed case study. *J Hydrol* 2014; **517**: 64–73.
- 45 Peña-Angulo D, Vicente-Serrano SM, Domínguez-Castro F, *et al.* Unravelling the role of vegetation on the different trends between climatic and hydrologic drought in headwater catchments of Spain. *Anthropocene* 2021; **36**: 100309.
- 46 Wang J, Wang W, Cheng H, *et al.* Propagation from meteorological to hydrological drought and its influencing factors in

- the huaihe river basin. *Water* 2021; **13**: 1985.
- 47 Van Loon AF, Rangelcroft S, Coxon G, *et al.* Using paired catchments to quantify the human influence on hydrological droughts. *Hydrol Earth Syst Sci* 2019; **23**: 1725–1739.
- 48 Sun L, Sun J, Li Y, *et al.* Development of an ensemble Bayesian inference-based copula approach for bivariate risk evaluation of extreme precipitation under climate change. *Intl J Climatol* 2022; **42**: 8755–8776.
- 49 Xu K, Wu C, Zhang C, *et al.* Uncertainty assessment of drought characteristics projections in humid subtropical basins in China based on multiple CMIP5 models and different index definitions. *J Hydrol* 2021; **600**: 126502.
- 50 Wu C, Yeh PJF, Ju J, *et al.* Assessing the spatiotemporal uncertainties in future meteorological droughts from CMIP5 models, emission scenarios, and bias corrections. *J Clim* 2021; **34**: 1903–1922.
- 51 Breiman L. Random forests. *Machine Learn* 2001; **45**: 5–32.
- 52 Zhou B, Li Y, Huang G, *et al.* Assessing distributed solar power generation potential under multi-gcms: A factorial-analysis-based random forest method. *ACS Sustain Chem Eng* 2022; **10**: 12588–12601.
- 53 Schoppa L, Disse M, Bachmair S. Evaluating the performance of random forest for large-scale flood discharge simulation. *J Hydrol* 2020; **590**: 125531.
- 54 Shiri N, Shiri J, Kazemi MH, *et al.* Estimation of CO₂ flux components over northern hemisphere forest ecosystems by using random forest method through temporal and spatial data scanning procedures. *Environ Sci Pollut Res* 2022; **29**: 16123–16137.
- 55 Vicente-Serrano SM, López-Moreno JI, Beguería S, *et al.* Accurate computation of a streamflow drought index. *J Hydrol Eng* 2012; **17**: 318–332.
- 56 Kao SC, Govindaraju RS. A copula-based joint deficit index for droughts. *J Hydrol* 2010; **380**: 121–134.
- 57 Zhang Q, Li YP, Huang GH, *et al.* Copula function with Variational Bayesian Monte Carlo for unveiling uncertainty impacts on meteorological and agricultural drought propagation. *J Hydrol* 2023; **622**: 129669.
- 58 Xu Y, Zhang X, Hao Z, *et al.* Characterization of agricultural drought propagation over China based on bivariate probabilistic quantification. *J Hydrol* 2021; **598**: 126194.
- 59 Zhu Y, Liu Y, Wang W, *et al.* A global perspective on the probability of propagation of drought: From meteorological to soil moisture. *J Hydrol* 2021; **603**: 126907.
- 60 Ionita M, Scholz P, Chelcea S. Assessment of droughts in romania using the standardized precipitation index. *Nat Hazards* 2016; **81**: 1483–1498.
- 61 Nelsen RB. Copulas and quasi-copulas: An introduction to their properties and applications. In: *24th Linz Seminar on Fuzzy Set Theory*. Linz, 2003, 391–413.
- 62 Chegwiddden OS, Nijssen B, Rupp DE, *et al.* How do modeling decisions affect the spread among hydrologic climate change projections? Exploring a large ensemble of simulations across a diversity of hydroclimates. *Earths Future* 2019; **7**: 623–637.
- 63 Chen Y, Li Z, Fang G, *et al.* Large hydrological processes changes in the transboundary rivers of central asia. *JGR Atmos* 2018; **123**: 5059–5069.
- 64 Yang T, Li Q, Chen X, *et al.* Spatiotemporal variability of the precipitation concentration and diversity in Central Asia. *Atmos Res* 2020; **241**: 104954.
- 65 Armstrong RL, Rittger K, Brodzik MJ, *et al.* Runoff from glacier ice and seasonal snow in High Asia: Separating melt water sources in river flow. *Reg Environ Change* 2019; **19**: 1249–1261.
- 66 Yang X, Li YP, Liu YR, *et al.* A MCMC-based maximum entropy copula method for bivariate drought risk analysis of the Amu Darya River Basin. *J Hydrol* 2020; **590**: 125502.
- 67 O’Neill BC, Tebaldi C, van Vuuren DP, *et al.* The scenario model intercomparison project (scenariomip) for CMIP6. *Geosci Model Dev* 2016; **9**: 3461–3482.
- 68 Thrasher B, Maurer EP, McKellar C, *et al.* Technical note: Bias correcting climate model simulated daily temperature extremes with quantile mapping. *Hydrol Earth Syst Sci* 2012; **16**: 3309–3314.
- 69 Rodell M, Houser PR, Jambor U, *et al.* The global land data assimilation system. *Bull Amer Meteor Soc* 2004; **85**: 381–

394.

- 70 Warszawski L, Frieler K, Huber V, *et al.* The inter-sectoral impact model intercomparison project (ISI-MIP): Project framework. *Proc Natl Acad Sci USA* 2014; **111**: 3228–3232.
- 71 Ukkola AM, De Kauwe MG, Roderick ML, *et al.* Robust future changes in meteorological drought in CMIP6 projections despite uncertainty in precipitation. *Geophys Res Lett* 2020; **47**: e87820.
- 72 Ju J, Wu C, Li J, *et al.* Global evaluation of model agreement and uncertainty in terrestrial water storage simulations from ISIMIP 2b framework. *J Hydrol* 2023; **617**: 129137.
- 73 Li Y, Chen A, Mao G, *et al.* Multi-model analysis of historical runoff changes in the Lancang-Mekong River Basin: Characteristics and uncertainties. *J Hydrol* 2023; **619**: 129297.