Reviewer 1：

Lakes in arid and semi-arid regions are good indicators for the regional environmental changes. The manuscript develops a series of technologies to construct a continuous monthly record of Bahannao Lake based on remote sensing data, further reveals the temporal shifts and nonlinear controls of hydro-climatic drivers on lake dynamics with the help of multi-factor analysis using the XGBoost model. A large number of data analysis have been carried out; the methods are sound and the results are reliable. However, there are shortcomings in the presentation and interpretation of the results, which prohibit the publication of this manuscript in the current version.
Overall, the manuscript offers significant value and is suitable for publication in HESS. The topic is of interest and fits the journal scope, but I have several suggestions and comments before publication in HESS.

We would like to sincerely thank the reviewer for taking the time to review the manuscript and for their positive assessment.

Major comments:

1. In the introduction part, it is necessary to address why you carry out the present work and the innovation of your work. But in the current form, it seems to be unclear in this part.

Response:

Thank you for this insightful comment. We agree that, in the original manuscript, the motivation and innovation of this study were not sufficiently clear in the Introduction. Following the reviewer’s suggestion, we have substantially revised the Introduction to explicitly clarify why this study is necessary and what its key innovations are. The main revisions are summarized as follows (**L82-L87**、**L182-L192**、**L207-L220**).

L82-L87:

However, despite increasing attention to global lake changes, small and medium-sized closed-basin lakes in arid and semi-arid regions remain poorly characterized in long-term observations. These lakes are highly sensitive to climate variability but are often underrepresented in existing global or regional datasets, highlighting an urgent need for improved long-term monitoring.

L182-L192:

These studies demonstrate the feasibility of large-scale lake monitoring, but also highlight persistent limitations related to temporal continuity, cloud dependence, and the applicability of existing products to small lakes. As a result, many existing lake-area studies rely on annual or seasonal snapshots derived from a limited number of cloud-free images, which may obscure important intra-annual variability, abrupt changes, and short-term climate responses, particularly for small lakes with strong seasonal dynamics.

To address this limitation, we construct a continuous 40-year monthly lake-area time series for Bahannao Lake by integrating multi-source Landsat imagery and applying a tailored image-processing workflow. This higher-temporal-resolution dataset enables a more detailed assessment of seasonal and interannual lake dynamics.

L207-L220:

Despite substantial progress in global lake monitoring, significant gaps remain for lakes in arid and semi-arid regions. Long-term and continuous lake area records are often interrupted by cloud contamination, seasonal ice cover, and striping artifacts, while the role of hydro-climatic drivers—particularly their nonlinear interactions—remains insufficiently understood.

To address these challenges, this study develops an optimized lake area extraction framework that integrates seasonal index selection, adaptive thresholding, connectivity analysis, and mutual information–based gap filling to construct a continuous monthly ake-area record for Bahannao Lake from 1984 to 2024. By coupling this reconstructed time series with multi-factor analysis using the XGBoost model, we quantify the relative importance and nonlinear effects of key hydro-climatic drivers on lake dynamics. This framework not only improves the reliability of long-term lake monitoring under complex conditions, but also provides new insights into seasonal and interannual climate controls on small lakes in arid and semi-arid regions.

2. The authors need to further specify the main findings of the paper and make comparison with related research. The results need to be compared with the former research.

Response:

Thank you for this constructive comment. We fully agree with the reviewer that the main findings and their comparison with previous studies needed to be clarified and strengthened.

Following this suggestion, we have substantially rewritten the Discussion section to more clearly articulate the key results of this study and to systematically compare them with existing research on lake dynamics in arid and semi-arid regions. Specifically, the revised Discussion now:

1.Explicitly summarizes the main findings, including the pronounced seasonal differences, the stage-dependent evolution of lake dynamics, and the shift in dominant hydro-climatic controls revealed by the long-term monthly lake-area record;

2.Compares our results with previous studies, highlighting both consistencies (e.g., the dominant role of precipitation and evaporation in arid-region lakes) and extensions beyond earlier work based mainly on annual-scale analyses;

3.Discusses the added value of the combined use of linear correlation analysis and XGBoost, particularly in identifying nonlinear and transitional controls that are not fully captured by traditional methods.

These revisions have been incorporated into the rewritten Discussion section (**L960–L1045**).

L960-L1045:

This study constructed a continuous monthly lake-area time series for Bahannao Lake spanning 1984–2024 using an optimized lake-area extraction framework that integrates seasonal water-index selection, adaptive thresholding, maximum connectivity analysis, and mutual information–based gap filling. Compared with widely used long-term products such as the JRC Global Surface Water dataset, which are often constrained by cloud contamination, seasonal ice cover, and temporal discontinuities, the proposed framework substantially improves temporal continuity and robustness under complex environmental conditions. This improvement is particularly important for small lakes in arid and semi-arid regions, where data gaps and seasonal disturbances are pervasive in existing datasets.

At the methodological level, this study introduces targeted improvements at several critical steps relative to previous approaches. First, the seasonal application of NDWI and MNDSI for non-freezing and freezing periods, respectively, enhances the stability of water-body identification under varying surface conditions, outperforming *traditional single-index methods (McFeeters, 1996; Yao et al., 2015). Second, the combination of Otsu thresholding with DEM-based terrain constraints effectively reduces misclassification caused by topographic shadows and complex terrain, which is a common challenge for inland lakes in arid environments. Third, the mutual information–based image-filling strategy reconstructs cloud- and stripe-contaminated pixels by matching historically most similar cloud-free images, thereby extending the usability of long-term Landsat archives. Compared with approaches relying solely on interpolation (Zhao and Gao, 2018), this strategy substantially improves the completeness and reliability of multi-decadal lake-area records. Collectively, these methodological enhancements systematically address key challenges repeatedly identified in previous studies, including cloud contamination, seasonal variability, topographic interference, and spectral complexity of inland waters (Mouw et al., 2015; Palmer et al., 2015; Shen et al., 2017; Cao et al., 2019), and establish a transferable framework suitable for lake monitoring in arid and data-scarce regions.*

From a hydro-climatic perspective, the reconstructed long-term record provides important insights into the mechanisms controlling lake dynamics in arid environments. Consistent with previous studies, precipitation and evaporation emerge as the primary factors regulating lake-area variability, particularly during the warm season when both water inputs and evaporative losses are enhanced (Tao et al., 2015; Li et al., 2017). The correlation analysis indicates that lake area is significantly positively correlated with precipitation and relative humidity in summer, whereas atmospheric moisture conditions exert a more pronounced influence during spring and winter. These findings reinforce the view that lake dynamics in arid regions are governed by the seasonal balance between water supply and evaporative demand.

However, compared with many existing studies that rely primarily on annual-scale analyses, the monthly lake-area time series developed here reveals pronounced seasonal heterogeneity and transitional behavior. In spring and autumn, linear correlations between lake area and individual climatic variables are generally weak, whereas XGBoost feature-importance analysis consistently identifies relative humidity and potential evapotranspiration as influential factors. This discrepancy suggests that lake responses during transitional seasons may be governed by nonlinear processes or threshold effects that cannot be fully captured by linear statistical methods alone. The combined use of correlation analysis and XGBoost therefore provides complementary perspectives on lake–climate relationships across different temporal scales.

At the decadal scale, both correlation analysis and XGBoost results indicate a clear evolution in dominant climatic controls on lake-area variability. During 2000–2014, precipitation and relative humidity exhibit increased importance and significant positive associations with lake area, indicating a moisture-dominated control regime. In contrast, during 2015–2024, the importance of air temperature and potential evapotranspiration increases markedly, while the contribution of precipitation weakens. This shift reflects a transition toward evaporation-dominated control under sustained warming conditions and highlights a dynamic reorganization of hydro-climatic drivers. Such temporal evolution extends existing understanding by explicitly demonstrating how dominant controls on arid-region lakes can shift under intensified climate variability.

These results have broader implications for studies of lakes in arid and semi-arid regions. The fragile water balance and limited buffering capacity of dryland lakes render them highly sensitive to even modest changes in precipitation, atmospheric moisture, and evaporative demand. The observed transition from precipitation-dominated to evaporation-dominated control suggests increasing vulnerability of arid-region lakes under ongoing climate warming. Even in the absence of a pronounced decline in precipitation, enhanced evaporation and atmospheric drying may offset or exceed water inputs, thereby accelerating lake shrinkage. This finding underscores the necessity of considering multiple hydro-climatic factors simultaneously when assessing future lake trajectories in arid environments.

From a water-resources management perspective, the results indicate that lake conservation and management strategies in arid regions should not focus solely on precipitation trends but must also account for changes in evaporative demand, drought intensity, and atmospheric moisture conditions. The lake-area extraction framework and the insights into evolving climatic controls presented here provide a robust technical foundation for long-term lake monitoring, risk assessment, and adaptive water-management strategies in data-sparse dryland regions.

Several limitations of this study should be acknowledged. First, while remote sensing reliably captures surface-area dynamics, subsurface processes such as groundwater inflow and outflow were not explicitly quantified and may influence lake water balance. Second, the 30 m spatial resolution of Landsat data limits detection of fine-scale shoreline changes, and future studies could benefit from integrating higher-resolution sensors such as Sentinel-2. Third, although XGBoost effectively captures nonlinear *relationships, its data-driven nature limits physical interpretability relative to process-based hydrological models. Future research could integrate remote sensing, machine learning, ecohydrological modeling, and socioeconomic data to further advance understanding of lake dynamics in arid regions.*

3. Relationship between climate change elements and the lake area need to be further investigated, e.g., using correlation analysis or wavelet coherence analysis.

Response:

Thank you for this valuable suggestion. We agree that a more explicit investigation of the relationships between climatic variables and lake-area changes is necessary.

Following the reviewer's recommendation, we have added a dedicated correlation analysis to systematically quantify the relationships between lake area and key hydro-climatic variables (including precipitation, air temperature, relative humidity, and potential evapotranspiration). This analysis provides a direct statistical assessment of the strength, direction, and significance of climate–lake linkages across seasonal and interdecadal scales.

The new correlation analysis and its results have been incorporated into Section 3.3.2 (**L845–L876**). In addition, the correlation results are interpreted jointly with the XGBoost-based nonlinear analysis to provide complementary insights into both linear and nonlinear climate controls on lake-area variability.

L845-L876:

The seasonal correlation analysis reveals pronounced differences in lake–climate relationships across seasons (Figure 17(a)). In spring, lake area exhibits a significant positive correlation with relative humidity (RH) ($r = 0.403$, $p < 0.01$) and a significant negative correlation with temperature (T) ($r = -0.352$, $p < 0.05$), indicating that spring lake-area variability is sensitive to atmospheric moisture conditions and warming processes. In contrast, correlations with precipitation (P) and potential evapotranspiration (PET) are weak and not statistically significant.

During summer, the lake–climate relationships are strongest. Lake area shows significant negative correlations with temperature ($r = -0.549$, $p < 0.01$) and PET ($r = -0.315$, $p < 0.05$), and significant positive correlations with precipitation ($r = 0.437$,

$p < 0.01$) and RH ($r = 0.468$, $p < 0.01$). These results indicate that summer lake-area variability is jointly controlled by moisture supply and enhanced evaporative demand. In autumn, lake area is significantly negatively correlated only with temperature ($r = -0.315$, $p < 0.05$), whereas correlations with precipitation, RH, and PET are not significant, suggesting that autumn lake-area variations may reflect cumulative effects of antecedent hydro-climatic conditions. In winter, lake area shows a significant positive correlation with RH ($r = 0.315$, $p < 0.05$), while correlations with other climatic variables remain weak, reflecting reduced hydrological activity during the cold season.

At the interdecadal scale, lake–climate correlations exhibit clear stage-dependent characteristics (Figure 17(b)). During the period 1984–1999, lake area shows no significant correlation with temperature, precipitation, PET, or RH, indicating a relatively weak response to individual climatic factors.

During 2000–2014, lake area becomes significantly positively correlated with precipitation ($p < 0.05$) and RH ($p < 0.01$), suggesting an enhanced sensitivity of lake-area variability to moisture conditions during this period. In the most recent period (2015–2024), lake area maintains a significant positive correlation only with RH ($p < 0.05$), while correlations with other climatic variables weaken, implying a dominant role of atmospheric moisture conditions in regulating recent lake-area changes.

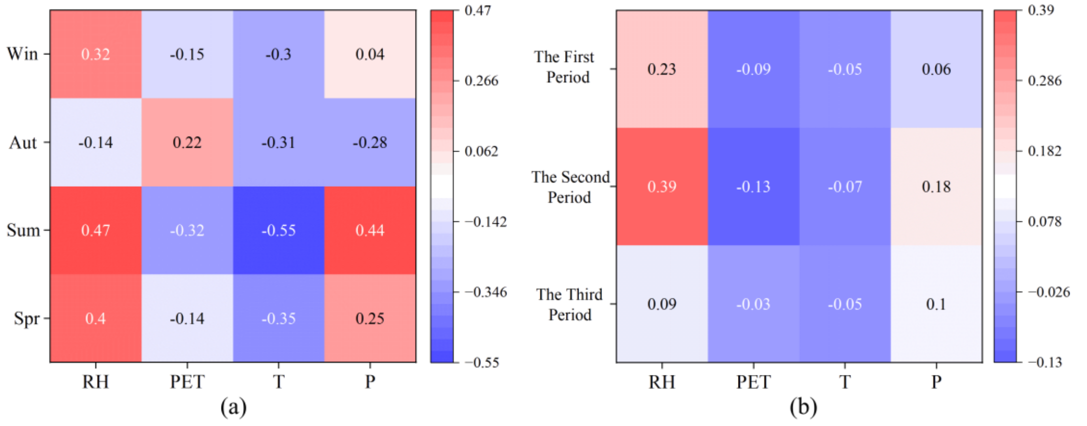


Figure 17 Seasonal and interdecadal differences in correlations between lake area and climatic drivers. (a) Seasonal correlations between lake area and RH, PET, T, and P.(b) Correlations between lake area and climatic variables across three sub-periods (1984–1999, 2000–

2014, and 2015–2024).

4. What are the advantages of the XGOOST method, why is it suitable for the analysis in this study?

Response:

Thank you for this important question. We agree that the advantages and suitability of the XGBoost method should be clearly explained.

Following this comment, we have substantially clarified the role, advantages, and interpretation strategy of the XGBoost model in both the Methods and Results sections. The main revisions are summarized below:

1.Clarification of the role and advantages of XGBoost (**L328-L339**):

L328-L339:

In this study, the XGBoost model is employed primarily as an interpretative tool. The objective is to quantify the relative importance of different hydro-climatic factors and to explore potential nonlinear relationships between lake-area variability and climatic drivers. Given the limited sample size, strong interannual variability, and high nonlinearity characteristic of arid-region lake systems, model performance metrics (e.g., R^2) are used as auxiliary indicators, while greater emphasis is placed on feature-importance rankings for mechanism interpretation.

Compared with linear correlation analysis, the XGBoost results highlight the importance of nonlinear and season-dependent controls, particularly during transitional seasons when linear correlations are weak. This demonstrates the added value of XGBoost in revealing climatic influences that cannot be fully captured by linear statistical methods alone.

2.Explanation of feature-importance interpretation (**L383-L389**):

L383-L389:

The feature importance derived from the XGBoost model reflects the relative contribution of each climatic variable in reducing prediction error across all decision trees. It should be noted that this importance ranking does not imply direct causality, but rather indicates the sensitivity of lake-area variability to different climatic factors *under nonlinear interactions. Therefore, feature importance is interpreted in conjunction with linear correlation analysis to provide a more robust understanding of hydro-climatic controls.*

3.Justification of model evaluation strategy and suitability (**L878-L890**):

L878-L890:

To further quantify the relative importance of climatic variables and explore potential nonlinear effects beyond linear correlations, an XGBoost model was applied using precipitation, temperature, relative humidity, and potential evapotranspiration as predictors.

Model evaluation indicates that training performance generally exceeds testing performance, and testing R² values are relatively low or even negative in some cases. This behavior reflects the limited sample size, strong interannual variability, and inherent nonlinearity of lake-area dynamics in arid regions, rather than model inadequacy. Therefore, in this study, XGBoost is primarily used as an interpretative tool to assess the relative importance of climatic drivers rather than as a predictive model.

XGBoost-derived feature importance exhibits clear seasonal contrasts that broadly agree with the correlation analysis while providing additional insights into nonlinear controls.

Minor comments:

1. The equations should be listed in numbers. And a map of the study area should be given.

Response:

Thank you for the helpful comment. In the revised manuscript, we have numbered all equations to improve clarity and consistency. In addition, we have added a map of the study area and expanded the description of the study region, including its geographical location, climatic conditions, and hydrological setting (**L223-L245\L266-L267**), to provide clearer contextual information for the analysis.

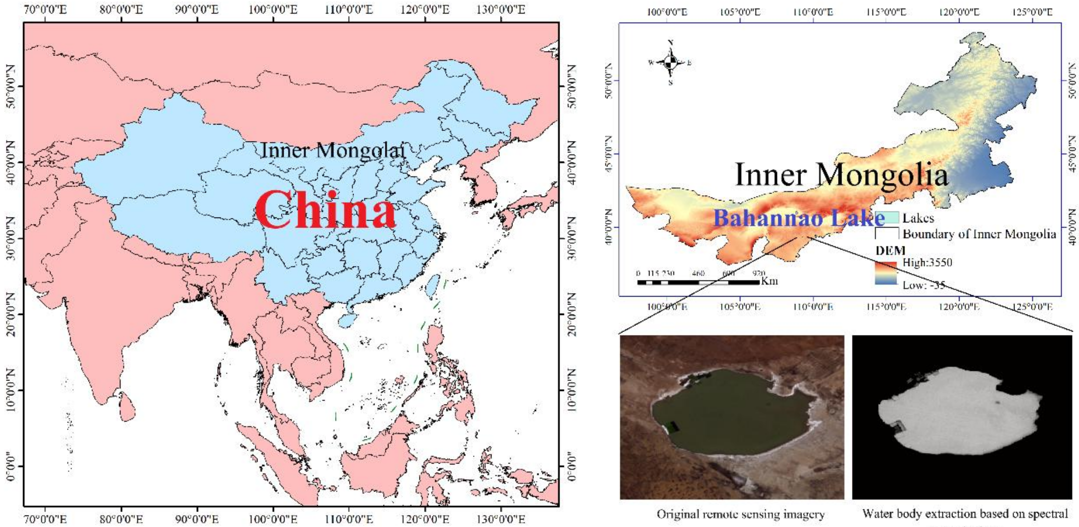
L223-L245:

Closed-basin lakes of various sizes are widely distributed across the Ordos Plateau, formed since the late Quaternary through combined aeolian and fluvial erosion processes. Bahannao Lake is the terminal basin of a chain of seven bead-like erosional lake depressions that developed along an ancient river valley. Bahannao Lake(109°16'E, 39°19'N) is located in the central Ordos Plateau at an elevation of 1278 m, with a lake-basin area of 26.50 km². The basin is underlain by a continuous and intact Lower Cretaceous sandstone formation, which provides a closed geomorphic setting primarily recharged by atmospheric precipitation. The sandstone contains abundant sodium- and calcium-rich carbonates, serving as the major source of dissolved salts in Bahannao Lake. Administratively, the study area belongs to Wushen Banner of the Ordos region in Inner Mongolia (Figure 1).

The zonal vegetation is dominated by arid to semi-arid desert steppe. The region is controlled for most of the year by the northwesterly monsoon, resulting in a cold and dry climate, while the southeasterly monsoon occasionally influences the area and plays a decisive role in seasonal precipitation. The mean annual temperature ranges from 6 to 9 °C, and the mean annual precipitation is only 200–300 mm, concentrated mainly from June to September with short-duration high-intensity rainfall events. In contrast, the annual potential evaporation reaches 2500–3000 mm, approximately ten times the precipitation amount, and the regional aridity index ranges from 3.5 to 4.0.

Because of the extremely fragile water balance and rapid hydrological response to climatic anomalies, Bahannao Lake and other nearby lakes are widely recognized as important natural indicators of climate variability, drought intensification, and land–*atmosphere interactions in the arid and semi-arid regions of northern China.*

L266-L267:



Original remote sensing imagery

Water body extraction based on spectral water indices

Figure 1. Overview map of the study area

2. The first key words should be “remote sensing”, instead of “R mote sensing”.

Response:

Thank you for pointing this out. We have corrected the typographical error in the keywords, and “remote sensing” is now written correctly in the revised manuscript.

L41:

Keywords: remote sensing, lake area extraction, XGBoost, arid region, hydro-climate

3. The expression of “Bahanao Lake” should be consistent all through the manuscript.

Response:

Thank you for this comment. We have carefully checked the entire manuscript and ensured that the name “Bahanao Lake” is used consistently throughout the revised manuscript.

4. Section 3.1 has too many subtitles and need to be merged into 4-5 subtitles. Subtitle (3) should be (2). Subtitle (4) Monthly Image Download need to be deleted.

Response:

Thank you for this helpful comment. We agree that the original Section 3.1 contained too many subtitles, which reduced the clarity of the methodological description. Following the reviewer’s suggestion, we have substantially reorganized and streamlined Section 3.1 by merging related steps, reducing the total number of subtitles, and removing redundant items.

Specifically, the subtitle “Monthly Image Download” has been deleted, and the remaining steps have been reorganized into four coherent subsections under **Section 3.1 Remote sensing interpretation and monthly lake image synthesis (L391)**:

3.1.1 Selection of water indices and image preprocessing (**L392**)

3.1.2 Threshold-based water segmentation and noise removal (**L471**)

3.1.3 Cloudy and striped image reconstruction (**L555**)

3.1.4 Monthly synthesis and time-series construction (**L601**)

This revised structure presents the methodological workflow in a clearer and more compact manner, improves logical continuity.

5. In Figure 9, 14, 15, 16, 17, 18, 19, 20, 21 fitting lines and P values need to be given.

Response:

Thank you for this valuable comment. Following the reviewer’s suggestion, we have added fitting (regression) lines and the corresponding p-values to all figures involving trend analysis in the revised manuscript, in order to improve the statistical interpretation and clarity of the results.

During the revision process, several figures were merged to reduce redundancy and improve presentation clarity. As a result, the figure numbering has changed compared with the original manuscript. Specifically, the original Figures 7, 14–15, 13 and 19, 17–18, and 20–21 have been reorganized and merged into the current Figures 9, 12, 13, 14, and 15, respectively.

All merged and renumbered figures now explicitly include fitted trend lines and the corresponding p-values, ensuring that the statistical significance of observed trends is clearly presented in accordance with the reviewer’s request.

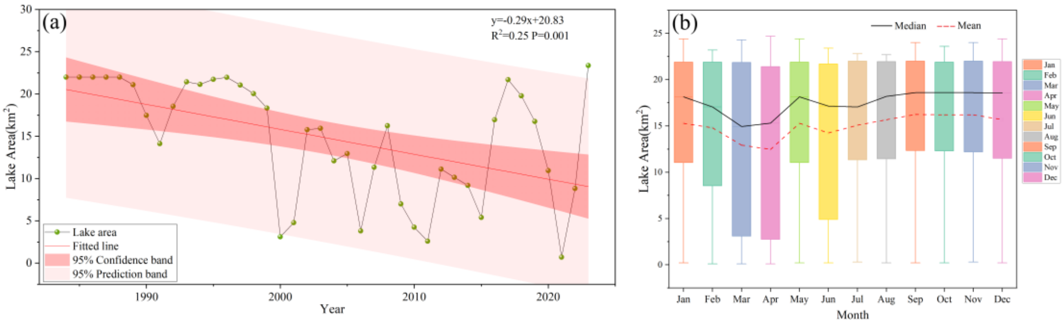


Figure 7 Interannual and intra-annual variation of Bahannao Lake area. (a) Interannual variations of lake area; (b) multi-year mean monthly variations of lake area.

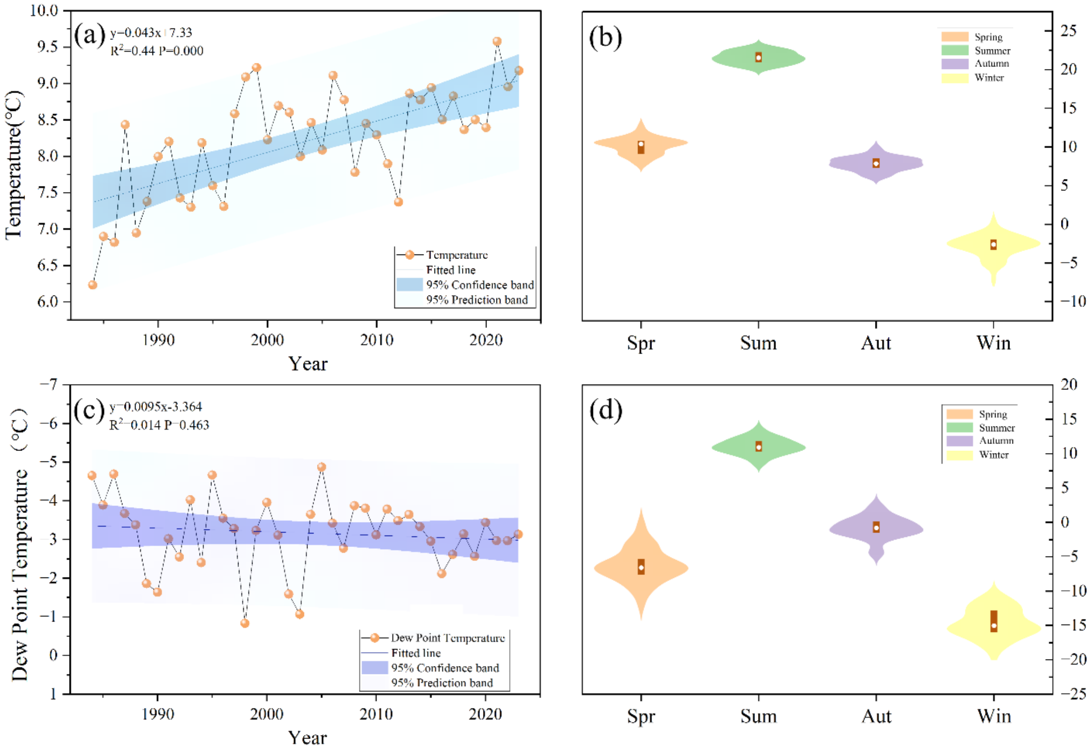


Figure 12 Temporal and seasonal variations in air temperature and 2 m dew point temperature over the study area during the period of 1984-2024 ((a) Interannual variations in air temperature; (b) Multi-year mean seasonal cycle of air temperature;(c) Interannual variations in 2 m dew point temperature;(d) Multi-year mean seasonal cycle of 2 m dew point temperature)

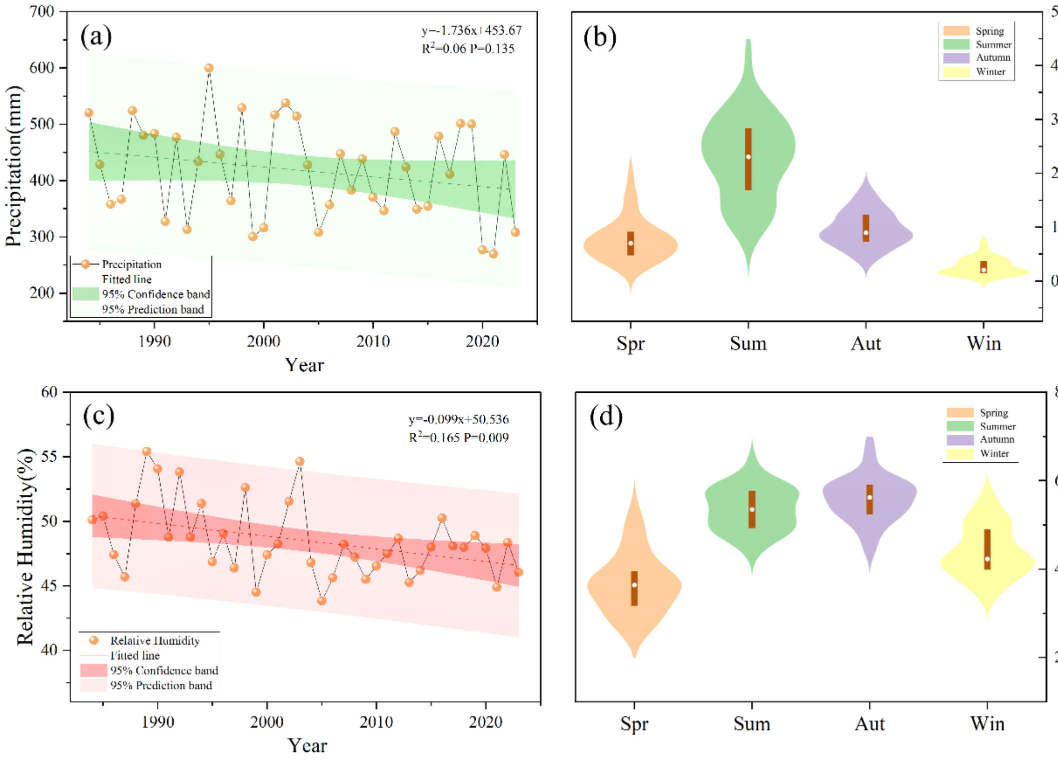


Figure 13 Temporal and seasonal variations in precipitation and relative humidity over the study area during the period of 1984-2024 ((a) Interannual variations in precipitation; (b) Multi-year mean seasonal cycle of precipitation;(c) Interannual variations in relative humidity;(d) Multi-year mean seasonal cycle of relative humidity)

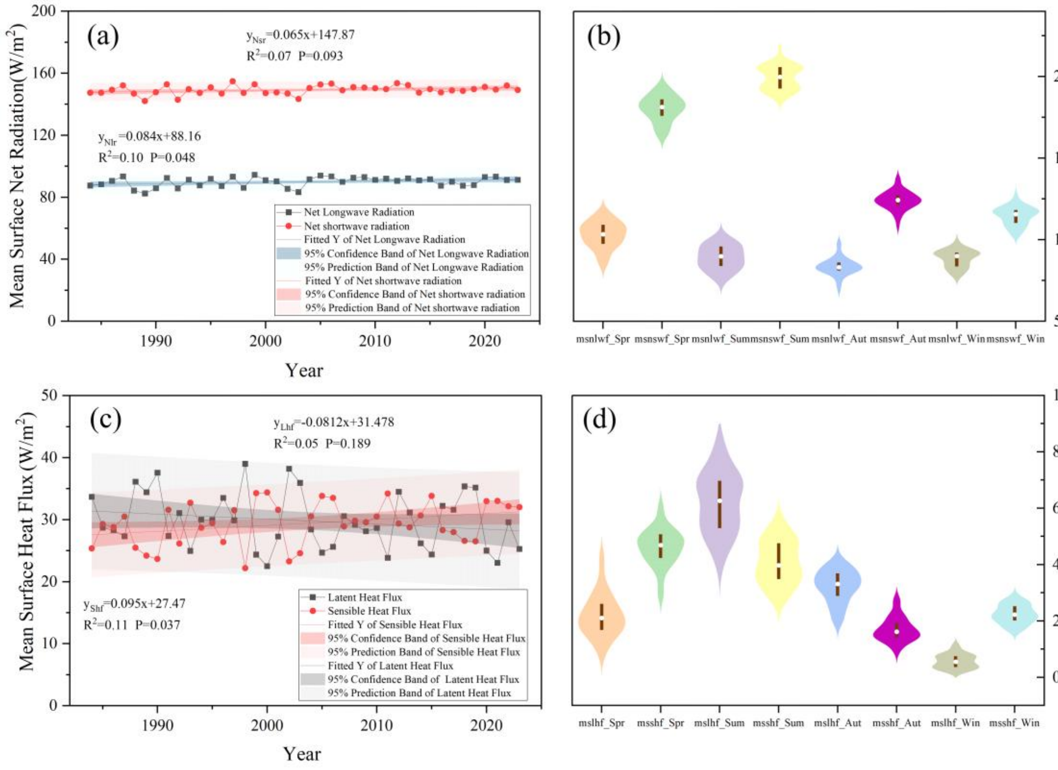


Figure 14 Regional variations of surface net radiation and surface heat flux during 1984–2024. (a) Interannual variations of mean surface net radiation; (b) multi-year mean seasonal variations of surface net radiation; (c) interannual variations of mean surface heat flux; (d) multi-year mean seasonal variations of surface heat flux.

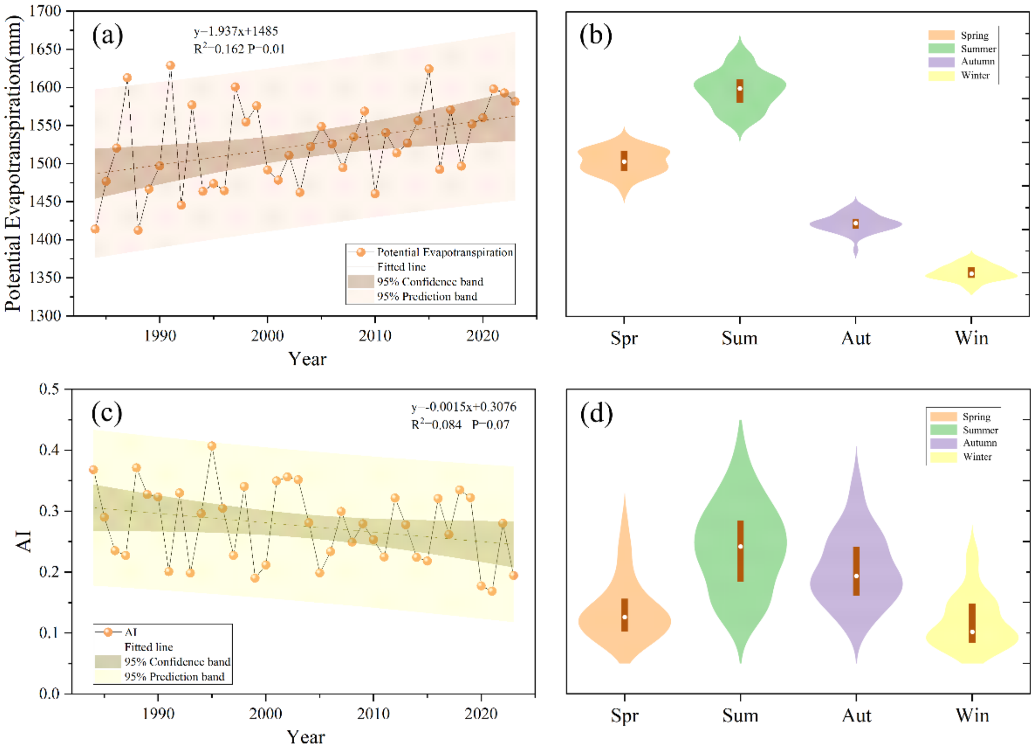


Figure 15 Regional variations in evaporation and drought conditions during 1984–2024. (a) Interannual variations of potential evapotranspiration;(b) multi-year mean seasonal variations of potential evapotranspiration; (c) interannual variations of the AI; (d) multi-year mean seasonal variations of the AI.

6. Some of the figures need to be merged. It is suggested to merge figures 23 and 24. Fig 11 and Fig 13 should be integrated into one figure; figures 14-16 should be merged into one figure.

Response:

Thank you for this constructive suggestion. Following the reviewer's recommendation, we have carefully reorganized and merged several figures in the revised manuscript to reduce redundancy and improve the clarity and coherence of figure presentation.

Specifically, multiple figures with overlapping content or complementary information were integrated as follows:

1) Original Figures 2 and 3 have been merged into the current Figure 3 to jointly illustrate water-body identification during non-freezing and freezing periods. (**L427**)
2) Original Figures 7 and 8 have been merged into the current Figure 6 to present cloud and stripe reconstruction results in a unified manner. (**L594**)
3) Original Figures 9 and 13 have been merged into the current Figure 7, and original Figures 10 and 11 into the current Figure 8, in order to streamline the presentation of lake-area validation and intra-annual variability. (**L623\L642**)
4) Original Figures 14 and 15 have been merged into the current Figure 12, while original Figures 13 and 19 are now combined as Figure 13, integrating related hydro-climatic variables into consolidated panels. (**L732\L747**)
5) Original Figures 17 and 18 have been merged into the current Figure 14, and original Figures 20 and 21 into the current Figure 15, to present energy flux components and their variations more concisely. (**L775\L831**)
6) Finally, original Figures 23 and 24 have been merged into the current Figure 18, as suggested by the reviewer. (**L955**)

These revisions substantially reduce the total number of figures while preserving all key information. The merged figures are now organized with clear subpanels and revised captions, improving readability and strengthening the logical flow of the Results section.

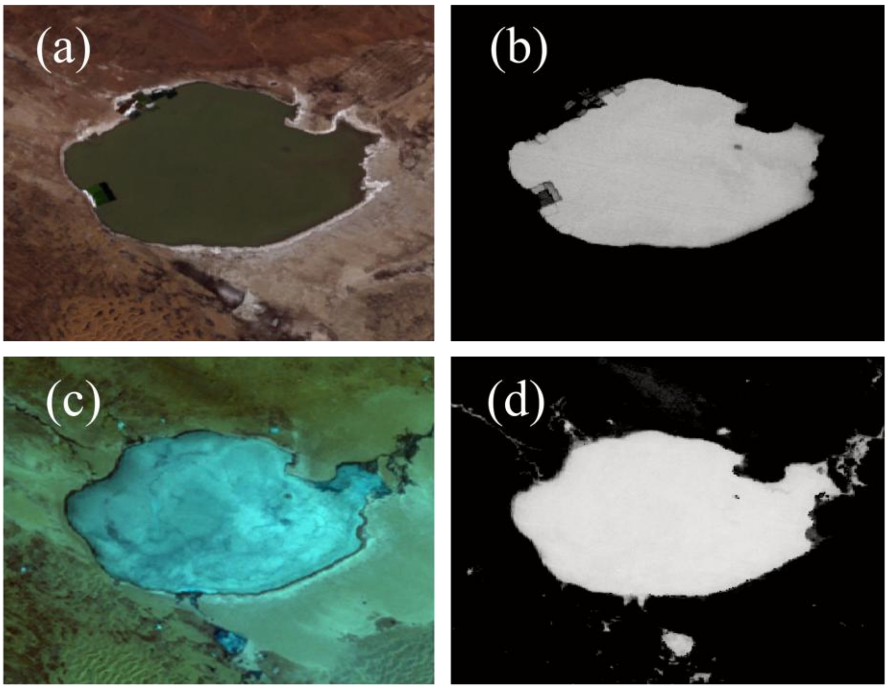


Figure 3 Lake extraction from Landsat imagery during non-freezing and freezing periods. (a) Original Landsat image during the non-freezing period;(b) Lake area identified using NDWI;(c) Original Landsat image during the freezing period;(d) Lake area identified using MNDSI. Source: Landsat imagery courtesy of the U.S. Geological Survey (USGS), processed and interpreted by the authors.

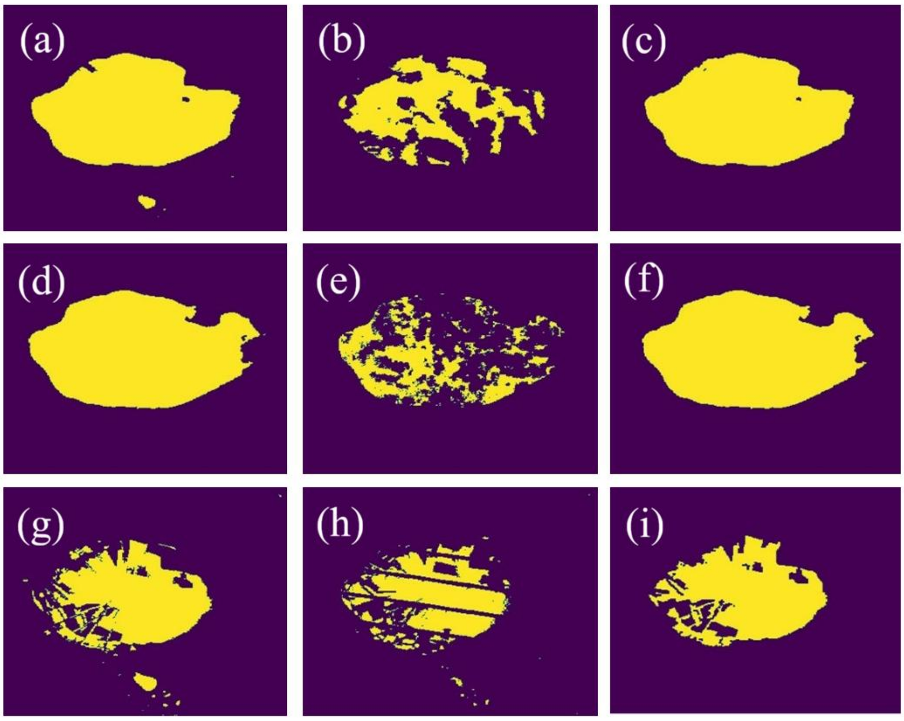


Figure 6: Filling processing of cloudy and striped-interference lake images using similar cloud-free references. (a)\(d) Cloud-free reference images identified as most similar to the cloudy images; (b)\(e) Original cloudy images;(c)\(f) Cloud-filled results after processing; (g) Cloud-free image most similar to the striped-interference image;(h) Original striped-interference image;(i) Result after stripe-filling processing. Source: Landsat imagery courtesy of the U.S. Geological Survey (USGS), processed and interpreted by the authors.

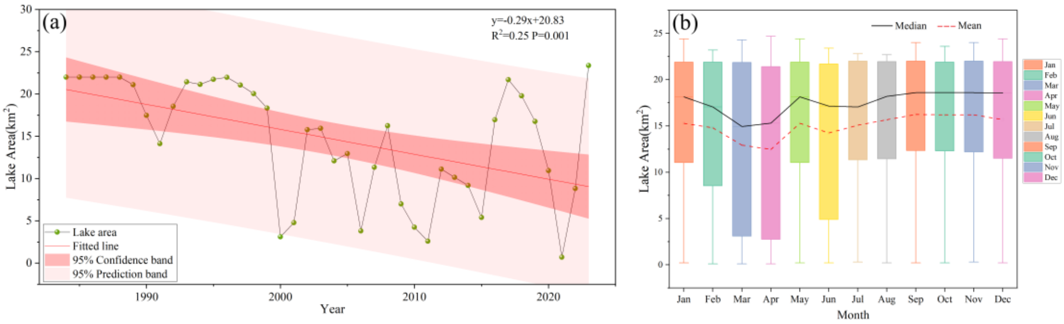


Figure 7 Interannual and intra-annual variation of Bahannao Lake area. (a) Interannual variations of lake area; (b) multi-year mean monthly variations of lake area.

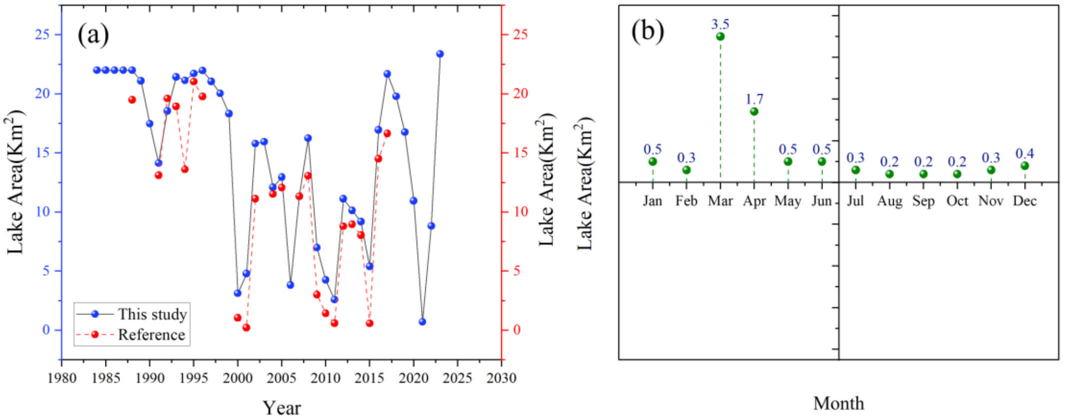


Figure 8 Validation of lake area estimates and intra-annual variability in a typical year (2021). (a) Comparison between lake area derived in this study and reference datasets;(b) monthly variations of lake area in 2021, selected as a representative year to illustrate intra-annual dynamics.

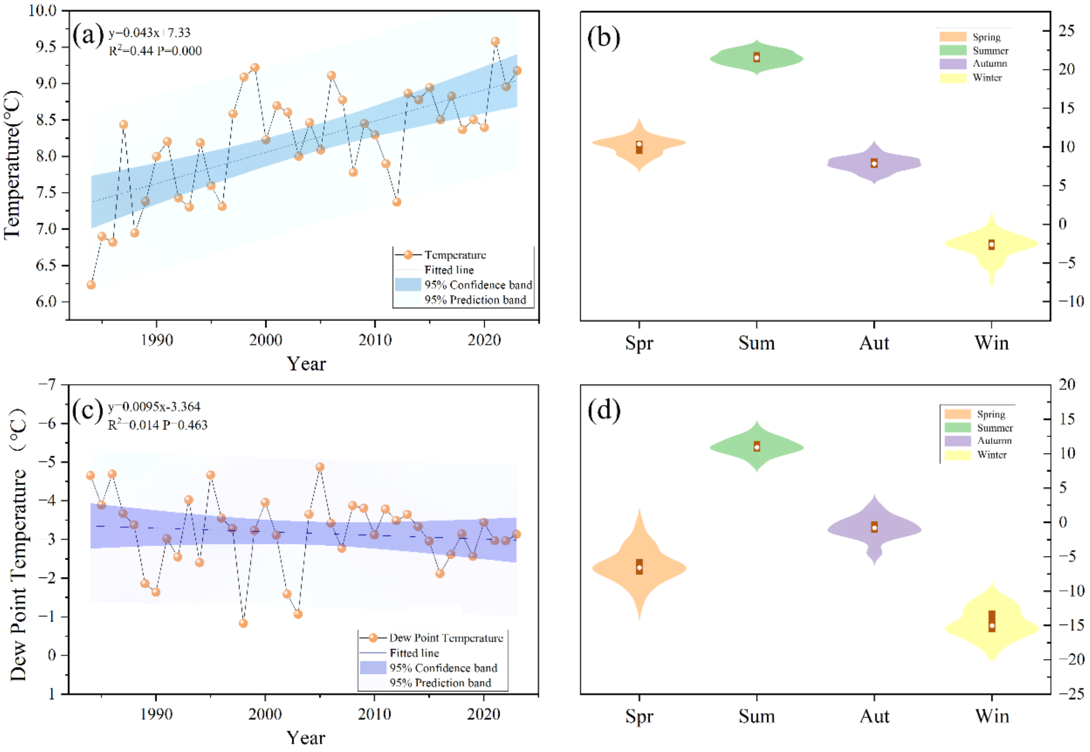


Figure 12 Temporal and seasonal variations in air temperature and 2 m dew point temperature over the study area during the period of 1984-2024 ((a) Interannual variations in air temperature; (b) Multi-year mean seasonal cycle of air temperature;(c) Interannual variations in 2 m dew point temperature;(d) Multi-year mean seasonal cycle of 2 m dew point temperature)

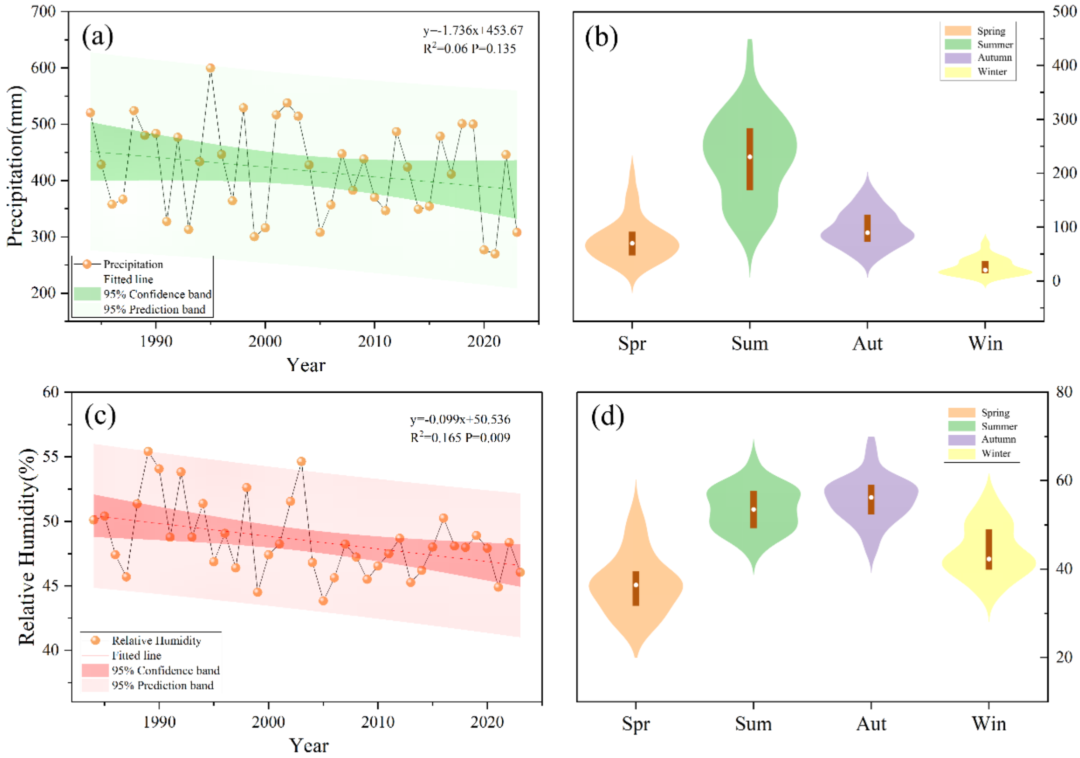


Figure 13 Temporal and seasonal variations in precipitation and relative humidity over the study area during the period of 1984-2024 ((a) Interannual variations in precipitation; (b) Multi-year mean seasonal cycle of precipitation;(c) Interannual variations in relative humidity;(d) Multi-year mean seasonal cycle of relative humidity)

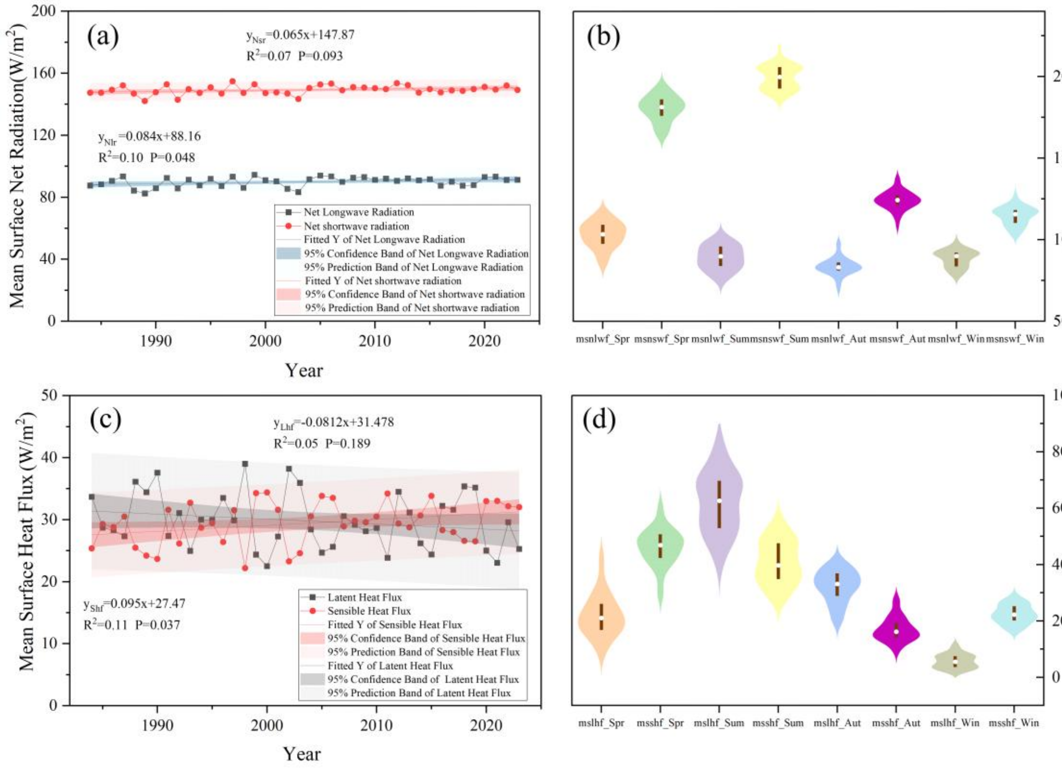


Figure 14 Regional variations of surface net radiation and surface heat flux during 1984–2024. (a) Interannual variations of mean surface net radiation; (b) multi-year mean seasonal variations of surface net radiation; (c) interannual variations of mean surface heat flux; (d) multi-year mean seasonal variations of surface heat flux.

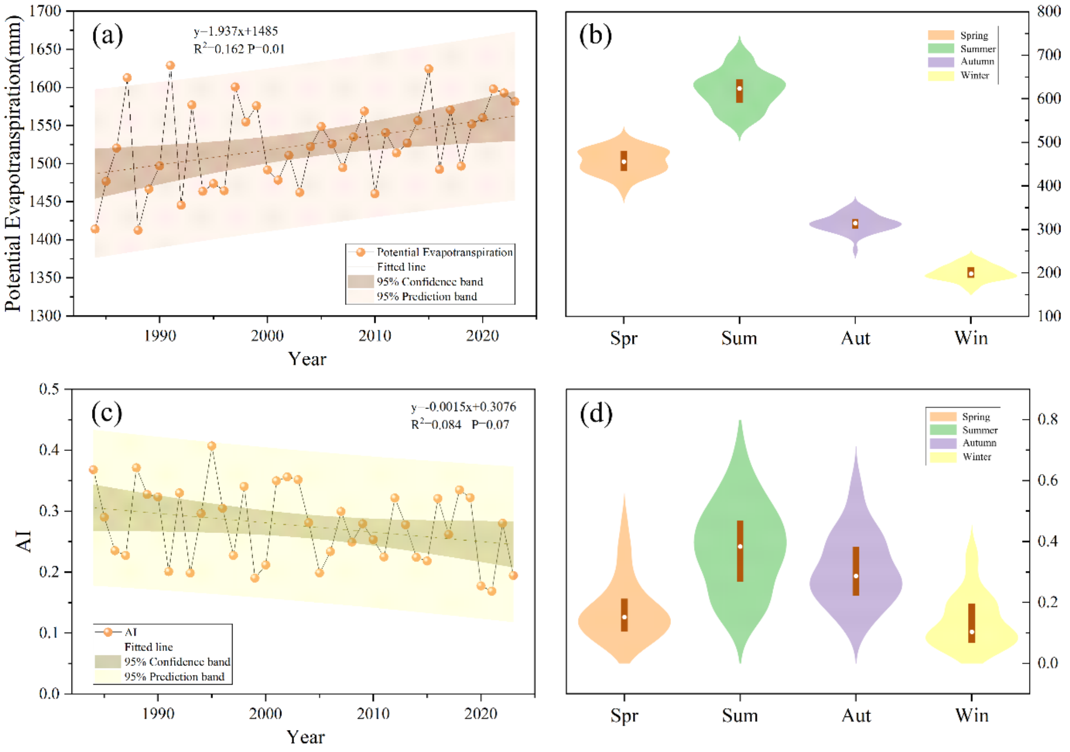


Figure 15 Regional variations in evaporation and drought conditions during 1984–2024. (a) Interannual variations of potential evapotranspiration;(b) multi-year mean seasonal variations of potential evapotranspiration; (c) interannual variations of the AI; (d) multi-year mean seasonal variations of the AI.

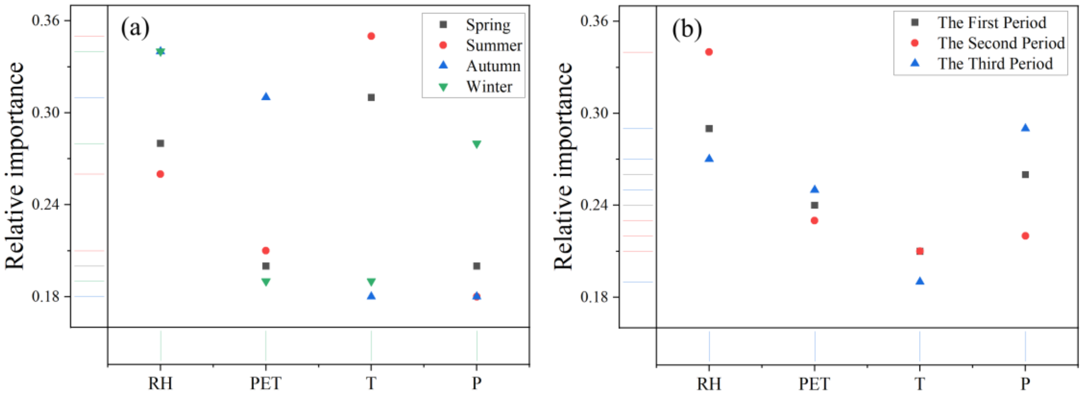


Figure 18 Weight of influencing factors by season

7. Section 3.3 can be divided into 3.3.1 Changes of hydro-climate series and 3.3.2 Impacts of hydro-climate elements on lake area. The second part may be from line 646 to line 758.

Response:

Thank you for this helpful suggestion. Following the reviewer's recommendation, we have reorganized Section 3.3 to clearly separate the description of hydro-climatic changes from the analysis of their impacts on lake-area variations.

Specifically, Section 3.3 has been restructured into two subsections:

Section 3.3.1 Changes of hydro-climate series (L708), which focuses on the temporal evolution of key hydro-climatic variables. This subsection is further organized into three thematic components:

(1) Temperature and moisture conditions (**L709**), including air temperature, 2 m dew point temperature, precipitation, and relative humidity;

(2) Surface radiation and heat flux components (**L752**), including net longwave radiation, net shortwave radiation, latent heat flux, and sensible heat flux;

(3) Evaporative demand and aridity conditions (**L784**), including potential evapotranspiration and drought index.

Section **3.3.2 Impacts of hydro-climate elements on lake area** (**L836**), which corresponds to the content originally located around Lines 646–758 in the previous version, and now explicitly examines the lake-area response to climatic forcing. This subsection includes:

(1) Linear relationships between lake area and climatic variables (**L837**);

(2) Nonlinear hydro-climatic controls revealed by the XGBoost (**L877**).

8. Line 156, "Recently, Pekel et al. (Pekel et al., 2014) utilized---", should be "Recently, Pekel et al. (2014) utilized---"

Response:

Thank you for pointing out this citation format issue. We have corrected the expression from "Pekel et al. (Pekel et al., 2014)" to "Pekel et al. (2014)" in the revised manuscript. The corrected text now appears at **L157-L158**.

L157-L158:

Recently, Pekel et al. (2014) utilized a large training dataset, combined with expert systems and visual analysis

9. It is suggested to give more description of Factor analysis in the method section of the manuscript. And XGBOOST model need to be mentioned in the result section.

Response:

Thank you for this constructive suggestion. Following the reviewer's recommendation, we have revised the manuscript to more clearly describe the factor analysis framework in the Methods section and to explicitly present the XGBoost model outputs in the Results section. The main revisions are summarized below.

(1) Expanded description of factor analysis in the Methods section

In Section 2.2.3 (XGBoost Model), we now provide a clearer explanation of the selection and physical meaning of the input factors. Specifically, precipitation (P), air temperature (T), relative humidity (RH), and potential evapotranspiration (PET) are explicitly defined as the primary components of the lake water balance in arid and semi-arid regions, directly regulating lake-area variability through water input and evaporative loss (**L346-L352**).Energy-related variables (e.g., radiation and heat fluxes) are clarified as background indicators of atmospheric conditions rather than direct driving forces of lake-area change, and this conceptual distinction is now explicitly stated in the Methods section.

L346-L352:

The input factors $x_i = \{x_1, x_2, \ldots, x_n\}$ *include precipitation (P), air temperature (T), relative humidity (RH), and potential evapotranspiration (PET), which represent the primary components of the lake water balance in arid and semi-arid regions. These variables directly or indirectly regulate lake-area changes through their influence on water input and evaporative loss. Energy-related variables (e.g., radiation and heat fluxes) are included as background indicators of atmospheric conditions and are not interpreted as direct driving forces of lake-area change.*

We further clarify that feature importance derived from the XGBoost model represents the relative contribution of each factor to reducing prediction error across all decision trees, rather than implying direct causality. Feature importance is therefore interpreted in combination with linear correlation analysis to provide a more robust understanding of hydro-climatic controls (**L383-L389**).

L383-L389:

The feature importance derived from the XGBoost model reflects the relative contribution of each climatic variable in reducing prediction error across all decision trees. It should be noted that this importance ranking does not imply direct causality, but rather indicates the sensitivity of lake-area variability to different climatic factors under nonlinear interactions. Therefore, feature importance is interpreted in conjunction with linear correlation analysis to provide a more robust understanding of hydro-climatic controls.

(2) Explicit presentation of XGBoost results in the Results section

In Section 3.3.2 (Impacts of hydro-climate elements on lake area), we now explicitly report the performance and outputs of the XGBoost model. Model evaluation results indicate that training performance generally exceeds testing performance, and that testing R^2 values are relatively low or even negative in some cases. We clarify that this behavior reflects limited sample size, strong interannual variability, and inherent nonlinearity of lake-area dynamics in arid regions, rather than model inadequacy (**L882-L890**). Accordingly, XGBoost is used primarily as an interpretative tool to assess the relative importance of climatic drivers rather than as a predictive model.

L882-L890:

Model evaluation indicates that training performance generally exceeds testing performance, and testing R^2 values are relatively low or even negative in some cases. This behavior reflects the limited sample size, strong interannual variability, and inherent nonlinearity of lake-area dynamics in arid regions, rather than model inadequacy. Therefore, in this study, XGBoost is primarily used as an interpretative tool to assess the relative importance of climatic drivers rather than as a predictive model.

XGBoost-derived feature importance exhibits clear seasonal contrasts that broadly agree with the correlation analysis while providing additional insights into nonlinear controls.

10. The title of figure 12 is better to be "--- during the period of 1984-2024". And the four seasons need to be written as Spring, Summer, Autumn and Winter.

Response:

Thank you for this helpful suggestion. Following the reviewer's comment, we have revised the figure accordingly.

Due to figure reorganization in the revised manuscript, the original Figure 12 has been renumbered as Figure 9. The figure title has been updated to "Figure 9. Seasonal variation of lake area during the period of 1984–2024".

In addition, the seasonal labels within the figure have been standardized to Spring, Summer, Autumn, and Winter to ensure consistency and clarity.

L646-L647:

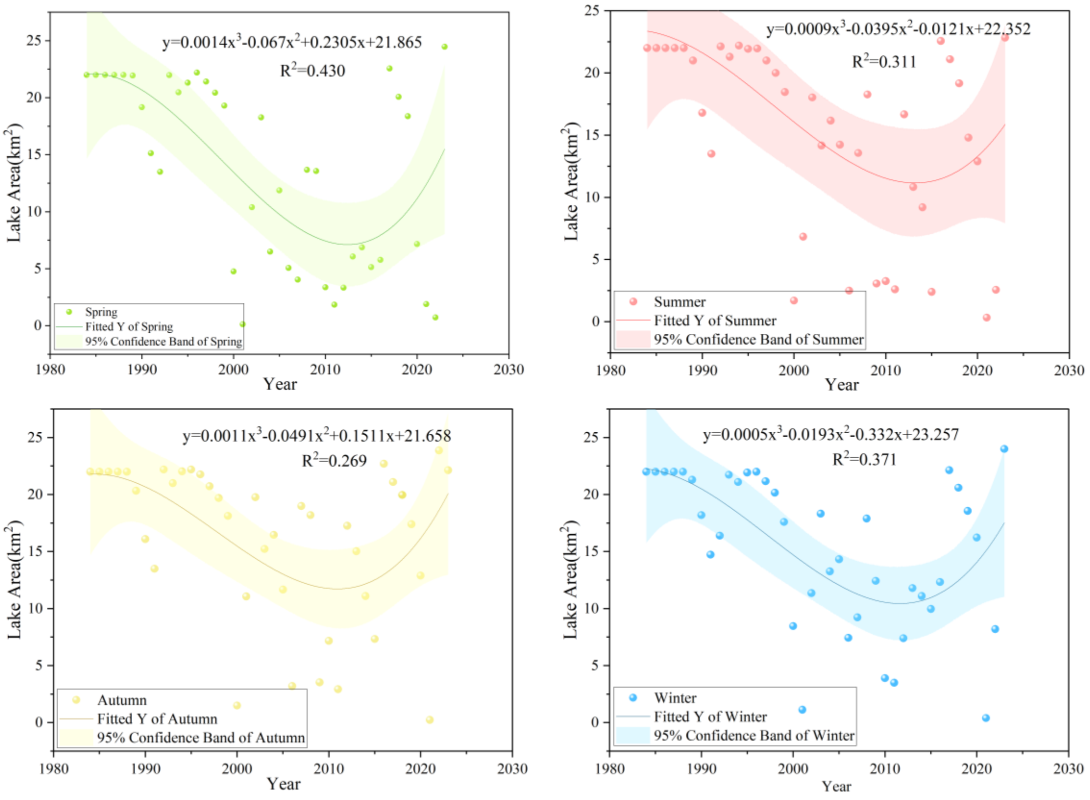


Figure 9 Seasonal variation of lake area during the period of 1984-2024

11. In section 3.3 Impact of Climate Chang, (1) Temperature Variation, "1) Temperature" and "2) 2m Dew Point Temperature" need to be merged. So do "(3) Radiation and Energy Exchang" and "(4) Humidity and Evapotranspiration".

Response:

Thank you for this constructive suggestion. We have revised the structure of Section 3.3 to reduce redundancy and improve conceptual coherence, following the reviewer's recommendation.

Specifically, Section 3.3 has been reorganized as “3.3 Impact of Climate Change”, with two main subsections. The hydro-climatic variables are now grouped according to their physical relevance rather than treated as separate, fragmented components:

3.3.1 Changes of hydro-climate series (**L708**), which includes:

(1) Temperature and moisture conditions (**L709**), combining
    1) Temperature and 2 m dew point temperature (**L710**) and
    2) Precipitation and relative humidity (**L737**);
(2) Surface radiation and heat flux components (**L752**), combining
    1) Net longwave radiation and net shortwave radiation at the surface (**L753**) and
    2) Mean surface latent heat flux and sensible heat flux (**L767**);
(3) Evaporative demand and aridity conditions (**L784**), including
    1) Potential evapotranspiration (**L785**) and
    2) Drought (**L794**).

12. Line 655, “---as shown in the figure 23---”, “the” should be deleted.

Response:

Thank you for pointing this out. We have removed the definite article “the” and updated the figure reference to reflect the revised figure numbering after figure merging. The text now reads “as shown in Figure 18(b)” in the revised manuscript.

L926-L927:

At the decadal scale, XGBoost results reveal a clear temporal shift in the dominant climatic controls on lake-area variability, as shown in figure 18(b).

Reviewer 2：

Lakes are barometers of climate change. Changes in the area of lakes have significant indicative significance for regional climate. This paper adopts a new approach, using remote sensing data to study the changes of lakes in arid areas and their relationship with climatic factors, which has certain value. However, the current state is not suitable for publication and still requires major revisions.

We would like to sincerely thank the reviewer for taking the time to review the manuscript and for their positive assessment.

1. Where exactly is Bahannao Lake? What kind of climate environment does it belong to? What kind of representativeness does it have? It is strongly recommended that the author explain clearly.

Response:

Thank you for this important comment. We fully agree that the location, climatic background, and representativeness of Bahannao Lake need to be clearly explained. To address this concern, we have added a detailed description of the study area in L**223-L245**, explicitly outlining the geographic setting, climatic conditions, and the representativeness of Bahannao Lake as a small closed-basin lake that serves as a sensitive indicator of hydro-climatic change in dryland regions. Furthermore, an overview map of the study area has been added as Figure 1 (**L266-L267**), providing a clear visual reference.

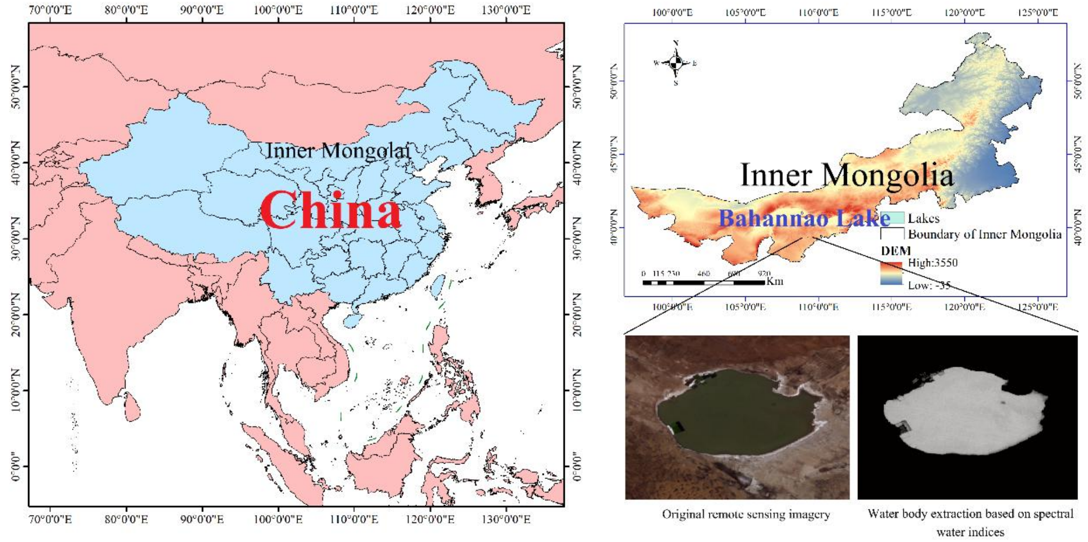
L223-L245:

Closed-basin lakes of various sizes are widely distributed across the Ordos Plateau, formed since the late Quaternary through combined aeolian and fluvial erosion processes. Bahannao Lake is the terminal basin of a chain of seven bead-like erosional lake depressions that developed along an ancient river valley. Bahannao Lake(109°16'E, 39°19'N) is located in the central Ordos Plateau at an elevation of

1278 m, with a lake-basin area of 26.50 km². The basin is underlain by a continuous and intact Lower Cretaceous sandstone formation, which provides a closed geomorphic setting primarily recharged by atmospheric precipitation. The sandstone contains abundant sodium- and calcium-rich carbonates, serving as the major source of dissolved salts in Bahannao Lake. Administratively, the study area belongs to Wushen Banner of the Ordos region in Inner Mongolia (Figure 1).

The zonal vegetation is dominated by arid to semi-arid desert steppe. The region is controlled for most of the year by the northwesterly monsoon, resulting in a cold and dry climate, while the southeasterly monsoon occasionally influences the area and plays a decisive role in seasonal precipitation. The mean annual temperature ranges from 6 to 9 °C, and the mean annual precipitation is only 200–300 mm, concentrated mainly from June to September with short-duration high-intensity rainfall events. In contrast, the annual potential evaporation reaches 2500–3000 mm, approximately ten times the precipitation amount, and the regional aridity index ranges from 3.5 to 4.0.

Because of the extremely fragile water balance and rapid hydrological response to climatic anomalies, Bahannao Lake and other nearby lakes are widely recognized as important natural indicators of climate variability, drought intensification, and land–atmosphere interactions in the arid and semi-arid regions of northern China.



Original remote sensing imagery

Water body extraction based on spectral water indices

Figure 1. Overview map of the study area

2. There are too many figures in the entire manuscript. It is recommended to compress them. Each figure contains too little information, especially Figures 2 to 8. It is suggested to integrate them.

Response:

Thank you for this helpful comment. We agree that the original manuscript contained too many figures with fragmented information. Following the reviewer's suggestion, we have substantially compressed and reorganized the figures by merging multiple related figures, particularly those originally numbered 2–8, as well as several later figures.

Following the reviewer's suggestion, we have comprehensively reorganized and compressed the figures throughout the manuscript to improve clarity and information density. Specifically:

Original Figures 2 and 3 have been merged into the current Figure 3 (**L427**);

Original Figures 7 and 8 have been merged into the current Figure 6 (**L594**);

Original Figures 9 and 13 have been integrated into the current Figure 7 (**L623**);

Original Figures 10 and 11 have been merged into the current Figure 8 (**L642**);

Original Figures 14 and 15 have been merged into the current Figure 12 (**L732**);

Original Figures 13 and 19 have been merged into the current Figure 13 (**L747**);

Original Figures 17 and 18 have been merged into the current Figure 14 (**L775**);

Original Figures 20 and 21 have been merged into the current Figure 15 (**L831**);

Original Figures 23 and 24 have been merged into the current Figure 18 (**L955**).

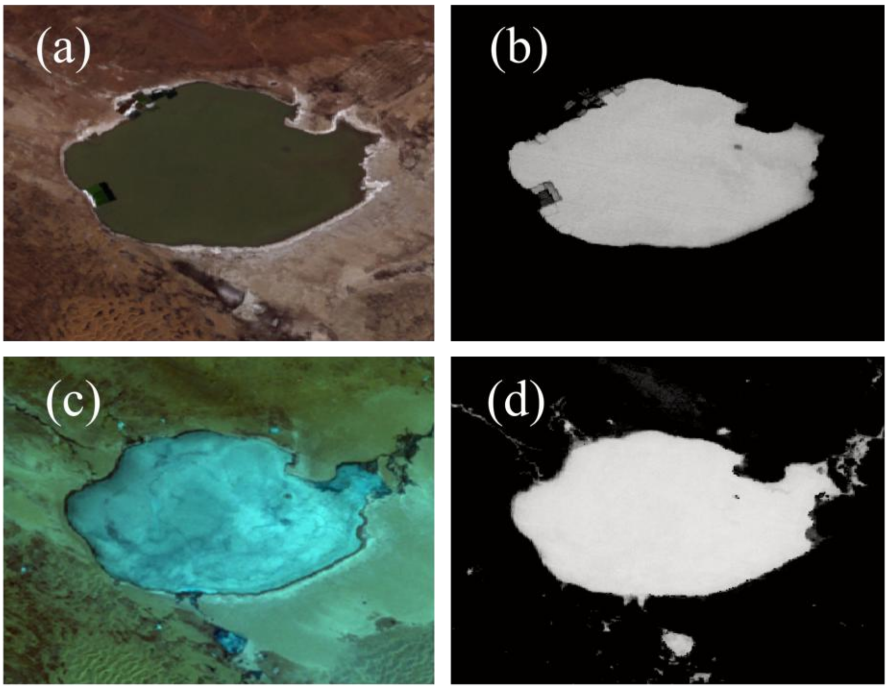


Figure 3 Lake extraction from Landsat imagery during non-freezing and freezing periods.(a) Original Landsat image during the non-freezing period;(b) Lake area identified using NDWI;(c) Original Landsat image during the freezing period;(d) Lake area identified using MNDSI. Source: Landsat imagery courtesy of the U.S. Geological Survey (USGS), processed and interpreted by the authors.

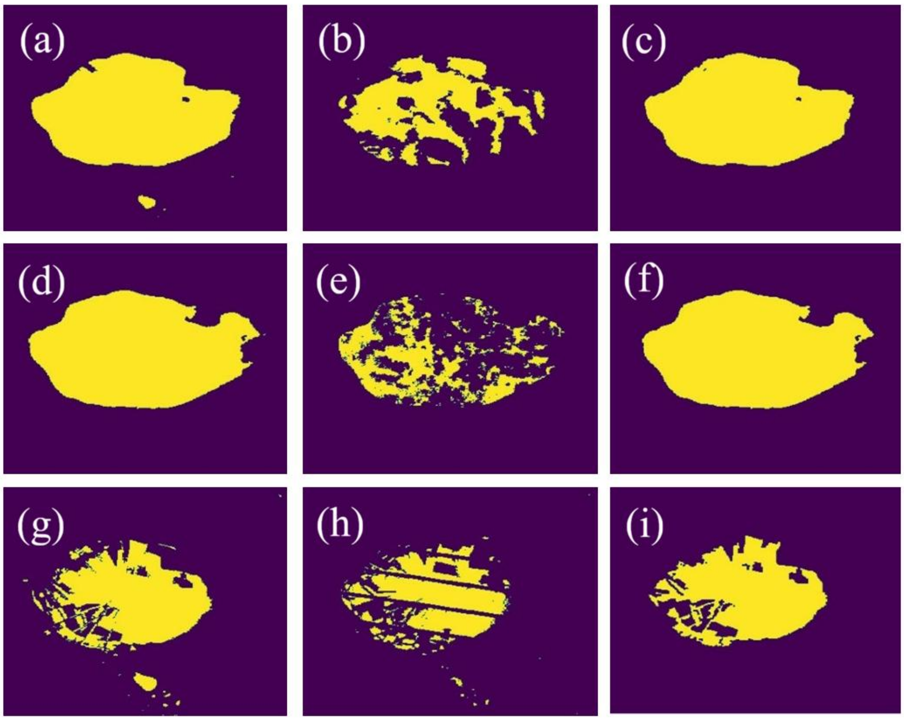


Figure 6: Filling processing of cloudy and striped-interference lake images using similar cloud-free references. (a)\(d) Cloud-free reference images identified as most similar to the cloudy images;

(b)\(e) Original cloudy images;(c)\(f) Cloud-filled results after processing; (g) Cloud-free image most similar to the striped-interference image;(h) Original striped-interference image;(i) Result after stripe-filling processing. Source: Landsat imagery courtesy of the U.S. Geological Survey (USGS), processed and interpreted by the authors.

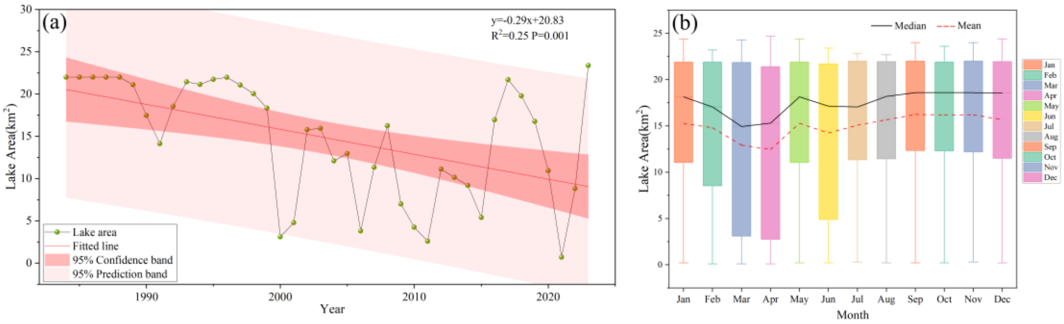


Figure 7 Interannual and intra-annual variation of Bahannao Lake area. (a) Interannual variations of lake area; (b) multi-year mean monthly variations of lake area.

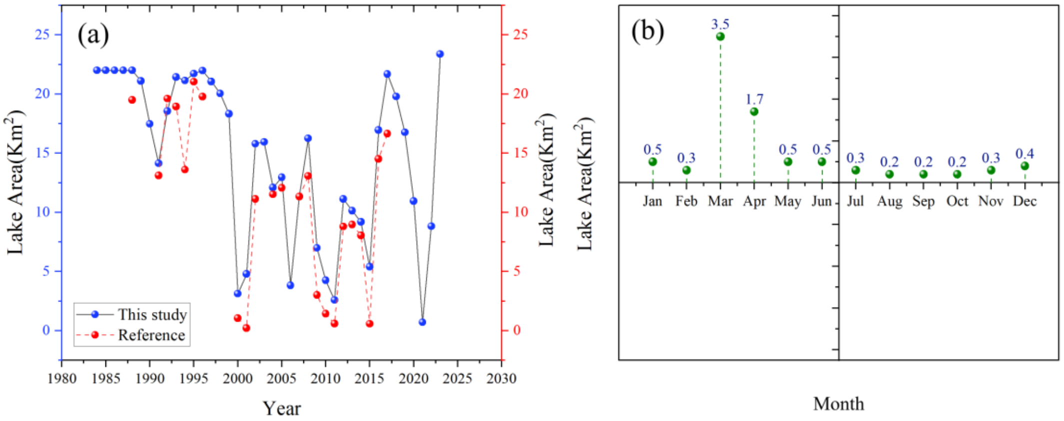


Figure 8 Validation of lake area estimates and intra-annual variability in a typical year (2021). (a) Comparison between lake area derived in this study and reference datasets;(b) monthly variations of lake area in 2021, selected as a representative year to illustrate intra-annual dynamics.

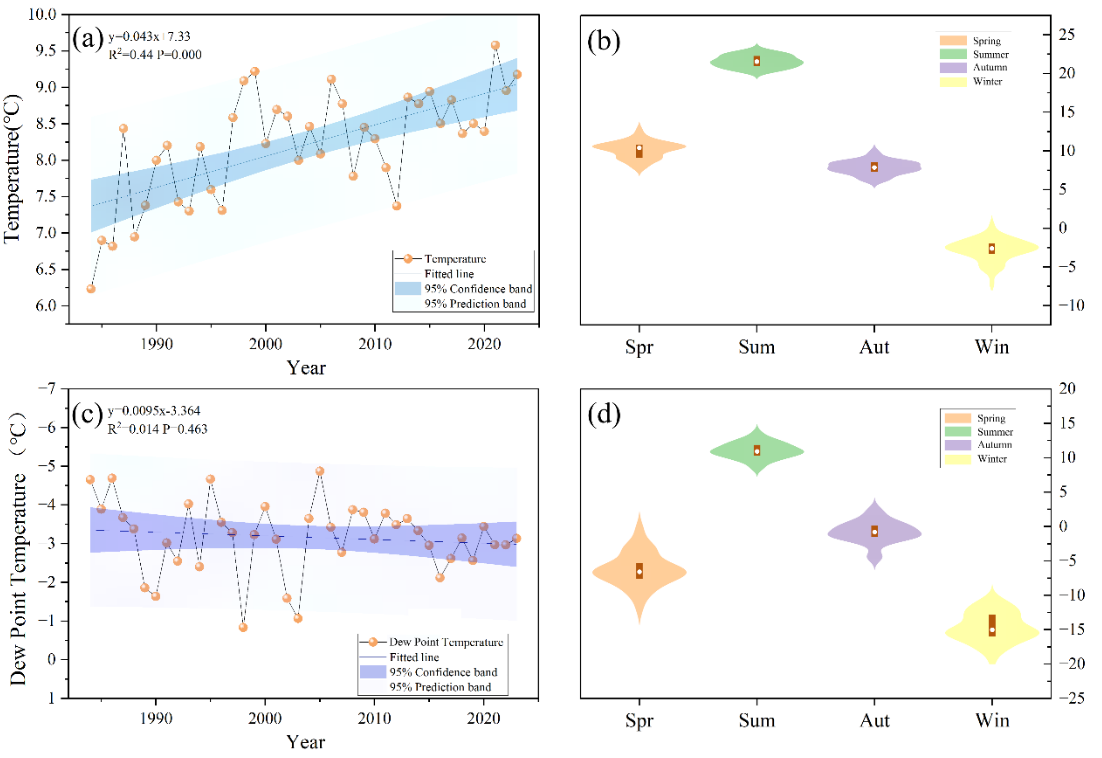


Figure 12 Temporal and seasonal variations in air temperature and 2 m dew point temperature over the study area during the period of 1984-2024 ((a) Interannual variations in air temperature; (b) Multi-year mean seasonal cycle of air temperature;(c) Interannual variations in 2 m dew point temperature;(d) Multi-year mean seasonal cycle of 2 m dew point temperature)

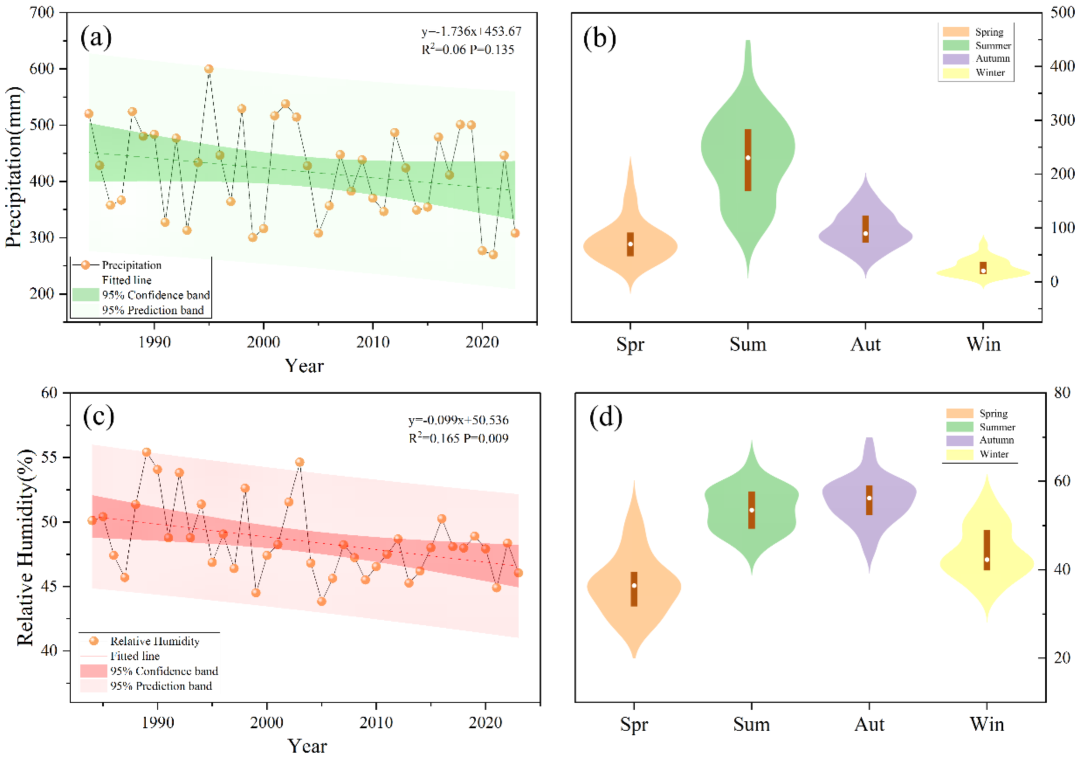


Figure 13 Temporal and seasonal variations in precipitation and relative humidity over the study area during the period of 1984-2024 ((a) Interannual variations in precipitation; (b) Multi-year

mean seasonal cycle of precipitation;(c) Interannual variations in relative humidity;(d) Multi-year mean seasonal cycle of relative humidity)

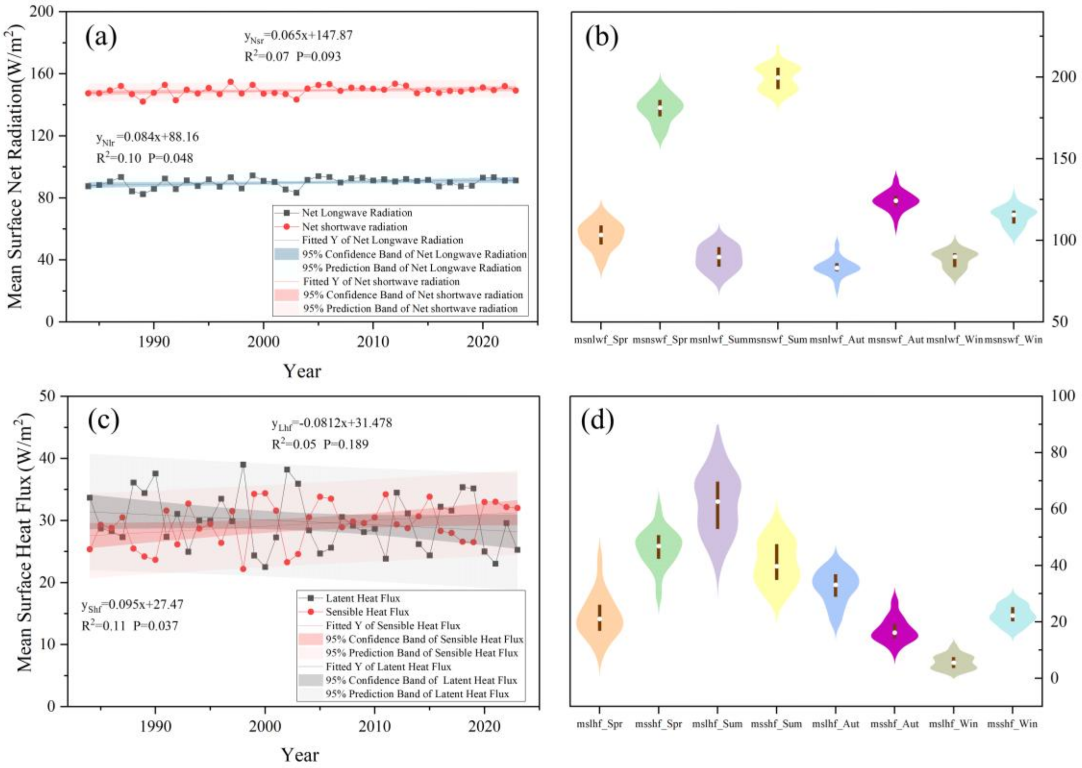


Figure 14 Regional variations of surface net radiation and surface heat flux during 1984–2024. (a) Interannual variations of mean surface net radiation; (b) multi-year mean seasonal variations of surface net radiation; (c) interannual variations of mean surface heat flux; (d) multi-year mean seasonal variations of surface heat flux.

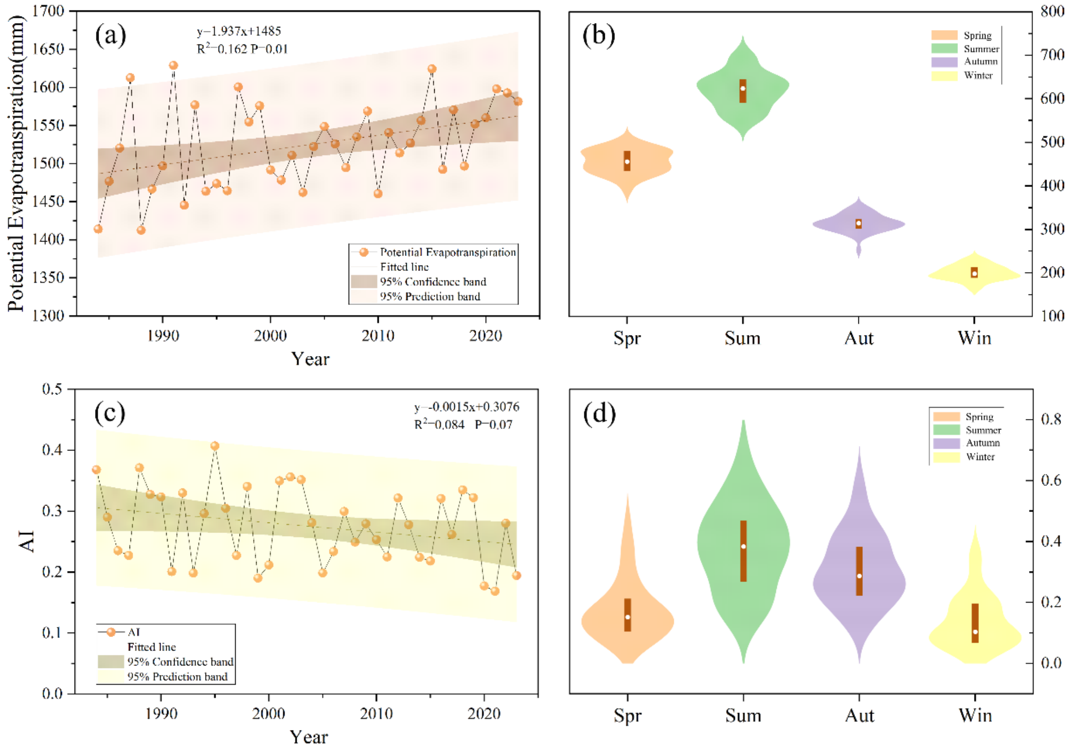


Figure 15 Regional variations in evaporation and drought conditions during 1984–2024. (a) Interannual variations of potential evapotranspiration;(b) multi-year mean seasonal variations of potential evapotranspiration; (c) interannual variations of the AI; (d) multi-year mean seasonal variations of the AI.

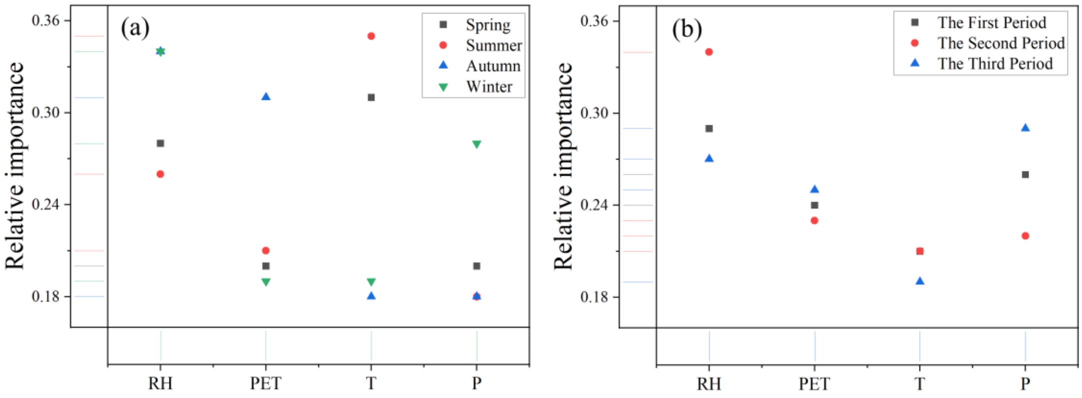


Figure 18 Weight of influencing factors by season

3. In the process of analyzing the driving factors of lake changes, energy flux is mainly affected by meteorological factors and land use, and it is a passive changing quantity. It is not appropriate to use it to explain the changes in lake area. From the perspective of water balance, precipitation, evaporation, runoff and groundwater

are the direct influencing factors. In terms of mechanism analysis, the influence of other factors is all exerted through these factors.

Response:

Thank you for this important comment. Following the reviewer's suggestion, we did not remove the energy-flux variables, but systematically redefined their role in the analysis. In the revised manuscript, radiation and heat-flux variables are no longer treated as direct driving factors of lake-area change, but are instead interpreted as passive indicators reflecting background atmospheric and evaporative conditions.

To ensure that the mechanism analysis is firmly grounded in a water-balance framework, we reselected precipitation, air temperature, relative humidity, and potential evapotranspiration as the primary explanatory variables and conducted XGBoost-based feature-importance analysis using these water-balance–related factors. In addition, linear correlation analysis between lake area and these climatic variables was added to provide complementary evidence for the interpretation of driving mechanisms.

These revisions clarify the distinction between direct hydrological controls and indirect atmospheric indicators, and ensure that lake-area changes are interpreted primarily through water input and evaporative loss processes. The corresponding revisions have been incorporated in the Methods and Results sections (**L845-L958**).

L845-L958:

The seasonal correlation analysis reveals pronounced differences in lake–climate relationships across seasons (Figure 17(a)). In spring, lake area exhibits a significant positive correlation with relative humidity (RH) ($r = 0.403$, $p < 0.01$) and a significant negative correlation with temperature (T) ($r = -0.352$, $p < 0.05$), indicating that spring lake-area variability is sensitive to atmospheric moisture conditions and warming processes. In contrast, correlations with precipitation (P) and potential evapotranspiration (PET) are weak and not statistically significant.

During summer, the lake–climate relationships are strongest. Lake area shows significant negative correlations with temperature ($r = -0.549$, $p < 0.01$) and PET ($r = -0.315$, $p < 0.05$), and significant positive correlations with precipitation ($r = 0.437$,

$p < 0.01$) and RH ($r = 0.468$, $p < 0.01$). These results indicate that summer lake-area variability is jointly controlled by moisture supply and enhanced evaporative demand. In autumn, lake area is significantly negatively correlated only with temperature ($r = -0.315$, $p < 0.05$), whereas correlations with precipitation, RH, and PET are not significant, suggesting that autumn lake-area variations may reflect cumulative effects of antecedent hydro-climatic conditions. In winter, lake area shows a significant positive correlation with RH ($r = 0.315$, $p < 0.05$), while correlations with other climatic variables remain weak, reflecting reduced hydrological activity during the cold season.

At the interdecadal scale, lake–climate correlations exhibit clear stage-dependent characteristics (Figure 17(b)). During the period 1984–1999, lake area shows no significant correlation with temperature, precipitation, PET, or RH, indicating a relatively weak response to individual climatic factors.

During 2000–2014, lake area becomes significantly positively correlated with precipitation ($p < 0.05$) and RH ($p < 0.01$), suggesting an enhanced sensitivity of lake-area variability to moisture conditions during this period. In the most recent period (2015–2024), lake area maintains a significant positive correlation only with RH ($p < 0.05$), while correlations with other climatic variables weaken, implying a dominant role of atmospheric moisture conditions in regulating recent lake-area changes.

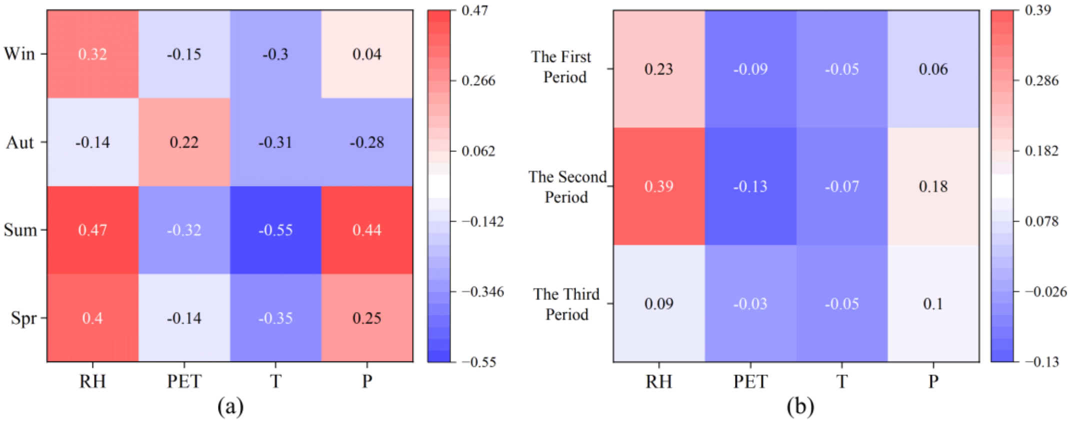


Figure 17 . Seasonal and interdecadal differences in correlations between lake area and climatic drivers. (a) Seasonal correlations between lake area and RH, PET, T, and P. (b) Correlations between lake area and climatic variables across three sub-periods (1984–1999, 2000–2014, and 2015–2024).

(2) Nonlinear hydro-climatic controls revealed by XGBoost

To further quantify the relative importance of climatic variables and explore potential nonlinear effects beyond linear correlations, an XGBoost model was applied using precipitation, temperature, relative humidity, and potential evapotranspiration as predictors.

Model evaluation indicates that training performance generally exceeds testing performance, and testing R^2 values are relatively low or even negative in some cases. This behavior reflects the limited sample size, strong interannual variability, and inherent nonlinearity of lake-area dynamics in arid regions, rather than model inadequacy. Therefore, in this study, XGBoost is primarily used as an interpretative tool to assess the relative importance of climatic drivers rather than as a predictive model.

XGBoost-derived feature importance exhibits clear seasonal contrasts that broadly agree with the correlation analysis while providing additional insights into nonlinear controls (Figure 18(a)). In spring, XGBoost feature importance indicates that air temperature is the most influential predictor (T, 0.31), followed by relative humidity (RH, 0.28), whereas precipitation (P, 0.20) and potential evapotranspiration (PET, 0.20) play secondary roles. This finding is consistent with the correlation analysis, which shows a significant positive correlation between lake area and RH ($r = 0.403$, $p < 0.01$) and a significant negative correlation with air temperature ($r = -0.352$, $p < 0.05$). Together, these results highlight the sensitivity of springtime lake dynamics to atmospheric moisture conditions and evaporative demand. Long-term trend analysis further indicates a significant increase in air temperature at a rate of 0.043 °C yr^{-1} ($p < 0.001$) and a significant decline in RH ($-0.099\ yr^{-1}$, $p = 0.009$), reinforcing the role of enhanced evaporation and atmospheric drying in shaping spring lake-area changes. In summer, air temperature (T, 0.35) and relative humidity (RH, 0.26) dominate the feature-importance rankings, with precipitation (P, 0.18) and potential evapotranspiration (PET, 0.21) also contributing substantially. This aligns well with the correlation results, which indicate that summer lake area is positively correlated with precipitation ($r = 0.437$, $p < 0.01$) and RH ($r = 0.468$, $p < 0.01$), and negatively correlated with temperature ($r = -0.549$, $p < 0.01$) and PET ($r = -0.315$, $p < 0.05$). Trend analysis shows that although precipitation exhibits a decreasing tendency (-1.736 mm yr^{-1}, $p = 0.135$, not significant), PET increases significantly at a rate of 1.937 mm yr^{-1} ($p = 0.01$). This suggests that summer lake-area variability is *increasingly constrained by enhanced evaporative demand, with the balance between water input and evaporation losses playing a dominant role.*

In autumn, linear correlations between lake area and most climatic variables are weak and statistically insignificant. However, XGBoost results still indicate relatively high importance for relative humidity (RH, 0.34) and potential evapotranspiration (PET, 0.31), suggesting that autumn lake dynamics may be governed by nonlinear processes or threshold effects that are not adequately captured by linear methods alone. Considering the significant upward trend in PET and the declining tendency of the aridity index (AI; $-0.015\ yr^{-1}$*,* $p = 0.07$*), autumn lake systems appear to be transitioning toward evaporation-dominated control.*

In winter, overall feature importance values are relatively low due to reduced hydrological activity, yet relative humidity (RH, 0.34) remains the most influential variable in the XGBoost model. This is consistent with the correlation analysis showing a significant positive relationship between winter lake area and RH ($r = 0.315$*,* $p < 0.05$*), indicating that background atmospheric moisture conditions still serve as an important indicator of lake variability during the frozen period.*

At the decadal scale, XGBoost results reveal a clear temporal shift in the dominant climatic controls on lake-area variability, as shown in figure 18(b). During 1984–1999, the importance of individual climatic variables is generally low and dispersed, consistent with the weak correlations observed during this period. This suggests that the lake system exhibited relatively low sensitivity to climatic fluctuations in the early stage.

During 2000–2014, precipitation (P, 0.22), potential evapotranspiration (PET, 0.23) and relative humidity (RH, 0.34) show markedly higher importance in the XGBoost model, in agreement with correlation results indicating significant positive relationships between lake area and precipitation ($r = 0.179$*,* $p < 0.05$*) and RH (*$r = 0.388$*,* $p < 0.01$*). This period is therefore characterized by a precipitation- and moisture-dominated control regime.*

In the most recent period (2015–2024), the importance of temperature and PET increases noticeably, while the contribution of precipitation weakens. Combined with the observed warming trend and enhanced evaporative demand, these results indicate a transition toward an evaporation-dominated climatic control on lake-area dynamics in recent years.

By integrating long-term trend analysis, linear correlation analysis, and XGBoost-based nonlinear feature importance, this study demonstrates that lake-area variability in arid regions is not governed by a single climatic factor, but rather by the interplay between water supply and evaporative demand across different seasons and time scales. Linear correlation analysis effectively captures the summer lake–climate relationship dominated by water balance, whereas the nonlinear XGBoost approach provides complementary insights into more complex control mechanisms during transitional seasons such as spring and autumn. Overall, the results indicate that with continued regional warming, increasing PET, and intensifying aridity, evaporative processes are playing an increasingly important role in controlling lake-area variability, offering important implications for understanding the response of arid-region lakes to future climate change.

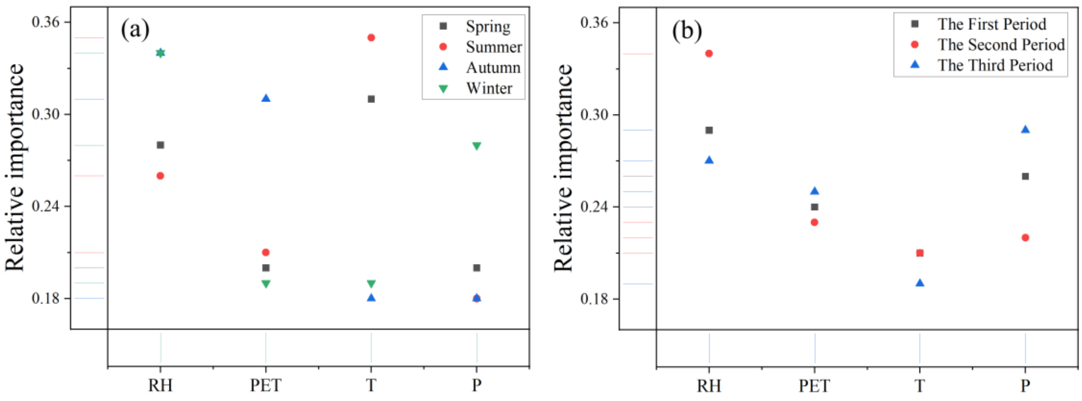


Figure 18 Weight of influencing factors by season

4. How is the drought index calculated? What does the size of a numerical value represent? It needs to be explained clearly in the method section.

Response:

Thank you for this helpful comment. In the revised manuscript, we have clarified both the calculation method and the physical meaning of the drought index.

Specifically, we have added a detailed description of the Aridity Index (AI) in the Methods section, including its definition, calculation procedure, and data sources (**L301-L326**). The AI is calculated as the ratio between precipitation and potential evapotranspiration, representing the balance between water supply and atmospheric evaporative demand. Lower AI values indicate drier conditions, while higher values correspond to more humid conditions.

In addition, we have expanded the analysis of regional drought conditions in the Results sections (**L794-L830**). Based on the long-term trend of the AI series, we describe the evolution of aridity in the study area and discuss its implications for lake-area shrinkage and hydro-climatic controls.

L301-L326:

2.2.2 Aridity index (AI)

The aridity index (AI) was used to quantify regional drought conditions. AI is defined as the ratio of precipitation (P) to potential evapotranspiration (PET), expressed as:

$$AI = \frac{P}{PET}$$

where P represents precipitation and PET denotes potential evapotranspiration. AI reflects the balance between atmospheric water supply and evaporative demand.

In this study, PET was calculated using the FAO Penman–Monteith method, which is widely recognized as a physically based and robust approach for estimating atmospheric evaporative demand. PET was computed as:

$$PET = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34 u_2)}$$

Where R_n *is the net radiation at the surface (MJ m-2 day-1), G is the soil heat flux (MJ m-2 day-1), T is the mean air temperature at 2 m height (℃),* u_2 *is the wind speed at 2 m height (m/s),* e_s *is the saturation vapor pressure (kPa),* e_a *is the actual vapor pressure (kPa), Δ is the slope of the saturation vapor pressure–temperature curve (kPa ℃-1), and* γ *is the psychrometric constant (kPa ℃-1).*

All meteorological variables required for PET estimation were obtained from the ERA5 reanalysis dataset and spatially averaged over the study area to ensure consistency with basin-scale analysis.

Based on AI values, climatic conditions were classified following the United Nations Environment Programme (UNEP) scheme: AI < 0.05 indicates hyper-arid conditions,

$0.05 \leq AI < 0.20$ represents arid conditions, $0.20 \leq AI < 0.50$ corresponds to semi-arid conditions, $0.50 \leq AI < 0.65$ indicates dry sub-humid conditions, and $AI \geq 0.65$ represents humid conditions. This classification allows a quantitative interpretation of regional dryness and facilitates comparison with previous studies in arid and semi-arid regions.

L794-L830:

2) Drought

The aridity index (AI) exhibits pronounced interannual variability over the study period, with values generally fluctuating between approximately 0.18 and 0.40 (Figure 15(c)), indicating persistently dry climatic conditions in the study area. Although a decreasing trend is observed ($-0.0015\ yr^{-1}$), the trend is not statistically significant at the 0.05 level ($p = 0.07$), suggesting that long-term aridity intensification is moderate rather than abrupt.

Despite the weak linear trend, the consistently low AI values (< 0.5) confirm that Bahannao Lake is located within a semi-arid to arid climatic regime, where water availability is inherently limited and highly sensitive to changes in hydro-climatic forcing. The wide prediction band further reflects strong year-to-year variability in regional moisture conditions, likely driven by fluctuations in precipitation and evaporative demand.

Importantly, the combination of a marginally decreasing AI trend and a significant increase in potential evapotranspiration implies a gradual shift toward enhanced atmospheric water demand, even in the absence of a statistically significant drying trend in AI alone. This suggests that lake-area dynamics are more strongly controlled by evaporative processes than by precipitation-driven moisture supply, particularly in recent decades.

In addition to interannual variability, the aridity index (AI) exhibits pronounced seasonal contrasts (Figure 15(d)). Summer shows the highest AI values, with a relatively wide distribution and higher median, indicating comparatively wetter conditions driven by concentrated precipitation during the warm season. Autumn presents intermediate AI values, reflecting a transition from moisture input to increasing evaporative demand.

In contrast, spring and winter are characterized by distinctly lower AI values. Spring exhibits low median AI and limited dispersion, indicating persistent moisture deficit *during the lake recharge period. This seasonal dryness coincides with rising temperatures and increasing evaporative demand, which constrains lake expansion despite episodic precipitation events. Winter shows the lowest AI values overall, reflecting extremely dry atmospheric conditions dominated by minimal precipitation and suppressed moisture availability.*

The seasonal pattern of AI highlights that Bahannao Lake is subject to strong intra-annual asymmetry in hydro-climatic conditions, with relatively favorable moisture supply confined to summer, while prolonged dry conditions prevail during spring and winter. Such seasonal dryness amplifies the sensitivity of lake area to evaporative processes.

5. What variables do the abbreviations in Figures 18 and 23 represent? It needs to be explained in the figure captions.

Response:

Thank you for this helpful comment. We agree that the abbreviations used in the figures should be clearly explained in the figure captions to avoid ambiguity.

In the revised manuscript, we have added explicit explanations of all abbreviations in the corresponding figure captions.

Figure 14 (formerly Figure 18): The following explanation has been added to the figure caption (**L780-L783**):

Figure 18 (formerly Figure 23): After revising the input variables, the XGBoost analysis now includes four hydro-climatic factors (RH, PET, T, and P). Their meanings have been explicitly clarified in the figure caption (**L957-L958**):

L780-L783:

Abbreviations: msnlwf denotes mean surface net longwave radiation; msnswf denotes mean surface net shortwave radiation; mslhf denotes mean surface latent heat flux; msshf denotes mean surface sensible heat flux. The suffixes Spr, Sum, Aut, and Win represent spring, summer, autumn, and winter, respectively.

L957-L958:

Abbreviations: RH denotes relative humidity (%); PET denotes potential evapotranspiration (mm); T denotes air temperature (℃);P denotes precipitation (mm).

6. Remote sensing monitoring of lake area changes has become relatively mature. What are the innovative aspects of the methods and data used in this paper? A clearer explanation needs to be given through comparison.

Response:

Thank you very much for this important comment. We agree that remote sensing–based monitoring of lake area has become a relatively mature research field, and therefore the novelty of the present study needs to be clarified more explicitly through comparison with existing studies.

Following the reviewer's suggestion, we have revised the manuscript to clearly articulate the innovation of this study from the perspectives of data, methodology, and analytical framework, and to explicitly position our work relative to previous research. The main revisions are summarized below.

(1) Clarification of data-related innovation in the Introduction (**L182–L192; L207–L220**):

L182-L192:

These studies demonstrate the feasibility of large-scale lake monitoring, but also highlight persistent limitations related to temporal continuity, cloud dependence, and the applicability of existing products to small lakes. As a result, many existing lake-area studies rely on annual or seasonal snapshots derived from a limited number of cloud-free images, which may obscure important intra-annual variability, abrupt changes, and short-term climate responses, particularly for small lakes with strong seasonal dynamics.

To address this limitation, we construct a continuous 40-year monthly lake-area time series for Bahannao Lake by integrating multi-source Landsat imagery and applying a tailored image-processing workflow. This higher-temporal-resolution dataset enables a more detailed assessment of seasonal and interannual lake dynamics.

L207-L220:

Despite substantial progress in global lake monitoring, significant gaps remain for lakes in arid and semi-arid regions. Long-term and continuous lake area records are often interrupted by cloud contamination, seasonal ice cover, and striping artifacts, while the role of hydro-climatic drivers—particularly their nonlinear interactions—remains insufficiently understood.

To address these challenges, this study develops an optimized lake area extraction framework that integrates seasonal index selection, adaptive thresholding, connectivity analysis, and mutual information–based gap filling to construct a continuous monthly ake-area record for Bahannao Lake from 1984 to 2024. By coupling this reconstructed time series with multi-factor analysis using the XGBoost model, we quantify the relative importance and nonlinear effects of key hydro-climatic drivers on lake dynamics. This framework not only improves the reliability of long-term lake monitoring under complex conditions, but also provides new insights into seasonal and interannual climate controls on small lakes in arid and semi-arid regions.

(2) Explicit methodological positioning in the Methods section (**L270–L274**):

L270-L274:

Although water-index-based lake extraction from Landsat imagery is well established, long-term monthly monitoring of small lakes in arid regions poses specific challenges, including frequent cloud contamination, striping artifacts in ETM+ data, and strong intra-annual variability. To address these issues, we developed an optimized processing workflow tailored to long-term monthly lake monitoring.

(3) Comparative discussion of methodological advantages in the Discussion section (**L960-L987**):

L960-L987:

This study constructed a continuous monthly lake-area time series for Bahannao Lake spanning 1984–2024 using an optimized lake-area extraction framework that integrates seasonal water-index selection, adaptive thresholding, maximum connectivity analysis, and mutual information–based gap filling. Compared with widely used long-term products such as the JRC Global Surface Water dataset, which are often constrained by cloud contamination, seasonal ice cover, and temporal discontinuities, the proposed framework substantially improves temporal continuity and robustness under complex environmental conditions. This improvement is particularly important for small lakes in arid and semi-arid regions, where data gaps and seasonal disturbances are pervasive in existing datasets.

At the methodological level, this study introduces targeted improvements at several critical steps relative to previous approaches. First, the seasonal application of NDWI and MNDSI for non-freezing and freezing periods, respectively, enhances the stability of water-body identification under varying surface conditions, outperforming traditional single-index methods (McFeeters, 1996; Yao et al., 2015). Second, the combination of Otsu thresholding with DEM-based terrain constraints effectively reduces misclassification caused by topographic shadows and complex terrain, which is a common challenge for inland lakes in arid environments. Third, the mutual information–based image-filling strategy reconstructs cloud- and stripe-contaminated pixels by matching historically most similar cloud-free images, thereby extending the usability of long-term Landsat archives. Compared with approaches relying solely on interpolation (Zhao and Gao, 2018), this strategy substantially improves the completeness and reliability of multi-decadal lake-area records. Collectively, these methodological enhancements systematically address key challenges repeatedly identified in previous studies, including cloud contamination, seasonal variability, topographic interference, and spectral complexity of inland waters (Mouw et al., 2015; Palmer et al., 2015; Shen et al., 2017; Cao et al., 2019), and establish a transferable framework suitable for lake monitoring in arid and data-scarce regions.

7. For such study, merely focusing on a single lake, especially an unknown small one, is not of much significance and also reduces the universality of the method and the representativeness of the conclusion. It is suggested to increase the number of lakes in this study or include similar lakes in this area.

Response:

Thank you for this valuable comment. We agree that focusing on a single small lake may raise concerns regarding representativeness and methodological universality. Following the reviewer's suggestion, we have revised the manuscript in two complementary ways.

First, in the Introduction, we have added a clearer explanation of why Bahannao Lake was selected as the study object and clarified its representativeness and typicality among small closed-basin lakes in arid and semi-arid regions (**L193-L206**). We emphasize that Bahannao Lake is a data-scarce but climatically sensitive desert lake, making it a suitable test case for evaluating long-term remote-sensing monitoring methods under challenging conditions.

Second, to further verify the robustness and regional applicability of the proposed method, we have added a new section entitled "3.2.2 Method validation using representative lakes in arid regions" (**L655-L706**). In this section, the same remote-sensing workflow is applied to two additional representative lakes—Hongjiannao Lake and Wuliangsuhai Lake—and the derived lake-area time series are quantitatively compared with published results from previous studies. The good agreement between our results and independent reference datasets demonstrates that the proposed framework is transferable and applicable beyond a single lake case.

These revisions strengthen the representativeness of the study and support the broader applicability of both the method and the conclusions.

L193-L206:

Bahannao Lake is a small closed-basin lake located in a semi-arid desert region of northern China. Owing to its remote location and the long-term absence of systematic in situ observations, continuous records of lake area are lacking. Nevertheless, as a *water body embedded in a fragile desert ecosystem, variations in lake area are highly sensitive to hydro-climatic changes and play an important role in regional eco-hydrological stability.*

In recent decades, intensified warming and drying have caused pronounced lake shrinkage, characterized by strong interannual variability and multiple abrupt changes. However, compared with larger or well-monitored lakes, the dynamic behavior and driving mechanisms of Bahannao Lake remain poorly understood due to the lack of long-term, high-temporal-resolution observations. As a typical but underrepresented small lake in arid regions, Bahannao Lake provides an ideal case for testing robust remote-sensing monitoring methods and investigating hydro-climatic controls on dryland lake dynamics.

L655-L706:

3.2.2 Method validation using representative lakes in arid regions

To further evaluate the robustness and regional applicability of the proposed lake-area extraction method, we applied the same remote-sensing workflow to two representative lakes in arid and semi-arid northern China: Hongjiannao Lake and Wuliangsuhai Lake. These lakes differ markedly in size, hydrological conditions, and degree of human influence, and have been widely investigated in previous remote-sensing studies, providing independent reference datasets for method validation.

Using the identical image-processing procedures and water-body extraction criteria as those employed for Bahannao Lake, we constructed annual lake-area time series for both Hongjiannao Lake and Wuliangsuhai Lake (Figure 10). The derived time series capture the major interannual fluctuations and long-term trends of lake-area variability for both lakes.

To quantitatively assess consistency with existing studies, the lake-area estimates obtained in this study were compared with previously published lake-area datasets (Figure 11). For both lakes, the temporal evolution and long-term trends derived in this study show good agreement with reference datasets reported in the literature.

For Hongjiannao Lake, quantitative comparison indicates that the relative differences *between lake-area estimates derived in this study and published datasets generally remain within a reasonable range. Specifically, the maximum and minimum relative differences are 14.65% and 9.12% when compared with Ji et al. (2023), 18.70% and 9.57% with Xie et al. (2021), 11.82% and 8.29% with Ma et al. (2020), 11.30% and 7.94% with Wang et al. (2018), and 10.57% and 3.15% with Liu et al. (2016).*

For Wuliangsuhai Lake, the relative differences are generally smaller, with maximum differences of 8.50% (minimum 1.74%) compared with Guan et al. (2022), 8.02% (minimum 1.79%) compared with Li et al. (2023), and 18.12% (minimum 1.09%) compared with Tan et al. (2021). These results indicate a high level of consistency between the lake-area estimates derived in this study and those reported in previous literature.

Although minor discrepancies in absolute lake-area values are observed, these differences can be attributed to variations in image selection, water-index thresholds, temporal coverage, and post-processing strategies among different studies. An additional source of discrepancy arises from differences in temporal aggregation strategies. In this study, annual lake area is calculated as the mean of monthly lake-area estimates derived from all available images within a year, which reduces the influence of short-term fluctuations and image-specific noise. In contrast, many previous studies report lake area based on a single image or a limited number of images selected for each year. Such differences in temporal representation can lead to systematic deviations in absolute lake-area values, particularly for lakes exhibiting strong intra-annual variability.

Overall, the consistency between our results and independent reference datasets supports the robustness and transferability of the proposed lake-area extraction method across different lake types in arid and semi-arid regions. This validation provides confidence that the method is suitable for long-term lake-area monitoring and comparative analysis in data-sparse dryland environments.

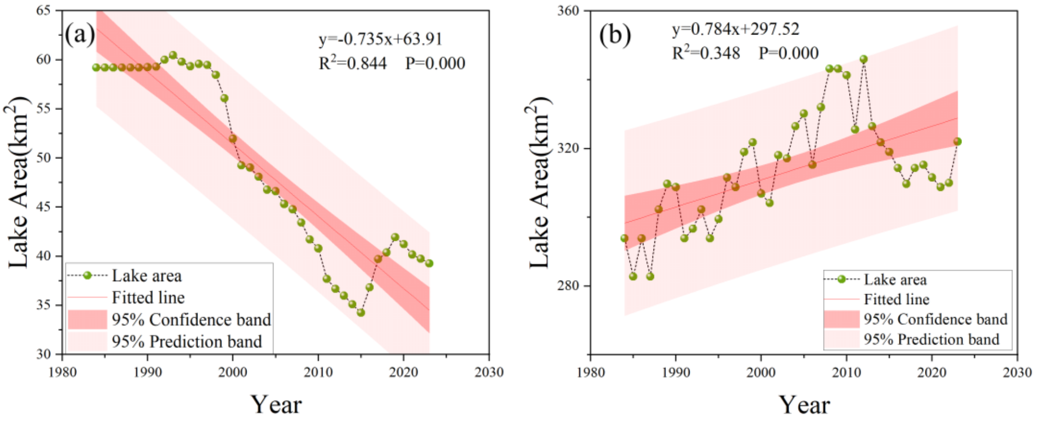


Figure 10 Interannual variations in lake area for Hongjiannao Lake (a) and Wuliangsuhai Lake(b).

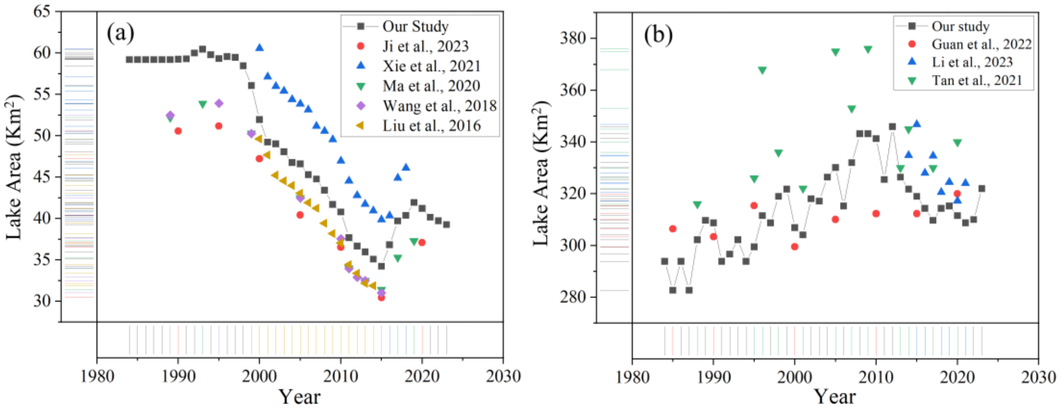


Figure 11. Comparison of lake area estimates derived in this study with published reference datasets.

(a) Comparison of Hongjiannao Lake area with lake-area estimates reported in previous studies;

(b) Comparison of Wuliangsuhai Lake area with lake-area estimates reported in previous studies.