Reply to the Handling Editor

Ref: Manuscript ID egusphere-2025-4778

Title: Disentangling the Key Drivers of Water Balance in Central Asia's Lake Balkhash: A Relative Contribution Assessment

Dear Handling Editor,

We sincerely thank you for handling the review process of our manuscript and for giving us the opportunity to further revise it. We also thank the two reviewers for their careful reassessment and constructive comments.

In this revision, we have mainly addressed the remaining concerns regarding the calibration, training, and validation strategy of the hybrid hydrological model. We clarified in Section 2.3.1 that the SEGSWAT+ calibration and ML residual-correction training were conducted using the same chronological 70/30 partition, and that no random shuffling was applied at any stage. We also revised Figure 6 and its caption to clearly distinguish the calibration/training and validation periods, and added detailed station- and period-specific calibration/validation information in Supplementary Tables S1 and S2. In addition, we addressed the remaining minor comments by moderating the statement about overfitting and removing the phrase “As noted by the reviewers” from the manuscript. We also improved the scientific support for several statements by adding relevant references on endorheic lake water balance and glacier melt/peak-water dynamics. Finally, we made minor language edits to improve clarity and readability, removed the LogNSE description to maintain consistency with the metrics shown in Figure 6, and standardized the decimal precision in Table B3.

We believe that these revisions have improved the clarity, rigor, and transparency of the manuscript. We hope that the revised version is now suitable for publication in HESS.

Sincerely,

Jinglu Wu (on behalf of all co-authors)

Reply to Reviewers' comments (Reviewer#1)

Ref: Manuscript ID egusphere-2025-4778

Title: Disentangling the Key Drivers of Water Balance in Central Asia's Lake Balkhash: A Relative Contribution Assessment

Dear Reviewer,

We sincerely thank you for your positive evaluation of our revised manuscript and for your helpful remaining comments. We have addressed both points in the revised manuscript, as detailed below.

Specific comments

(1) L273–276: The statement that overfitting is mitigated by training the ML model on residuals is not fully supported. While modelling residuals instead of the full signal may reduce the variance and apparent complexity of the target, it does not inherently prevent overfitting. Residuals can still contain noise and unresolved variability, which may be captured by flexible ML models, particularly in the absence of explicit regularization or rigorous validation. This point should be rephrased to avoid overinterpretation.

Response

Thank you for this precise comment. We agree that residual-based learning alone does not prevent overfitting, and our previous wording was too strong. We have revised the statement in Section 2.3.1 to avoid overinterpretation and to clarify how overfitting risk was reduced.

The revised text is:

"Training the ML component on residuals rather than total streamflow reduces the magnitude and apparent complexity of the target variable and helps restrict the ML correction to the residual component. However, residual-based learning alone does not eliminate the risk of overfitting. To reduce this risk, we used chronological validation, hyperparameter tuning, and independent performance evaluation based on KGE, NSE, PBIAS, and R^2."

(2) L812: It is recommended to remove the phrase "As noted by the reviewers," as this is not appropriate for the final manuscript.

Response

Thank you for pointing this out. We agree that this phrase is inappropriate for the final manuscript. We have removed "As noted by the reviewers" and revised the relevant sentence so that the statement is supported by the manuscript content and cited literature rather than by reference to the review process.

Reply to Reviewers' comments (Reviewer#2)

Ref: Manuscript ID egusphere-2025-4778

Title: Disentangling the Key Drivers of Water Balance in Central Asia's Lake Balkhash: A Relative Contribution Assessment

Dear Reviewer,

We sincerely thank you for your careful reassessment of our manuscript and for raising important concerns regarding the calibration, training, and validation strategy of the hybrid modeling framework. We fully agree that, for a hybrid model in which the ML component learns residuals from a process-based model, the training and validation design must preserve the temporal structure of the hydrological records and avoid information leakage.

Your comment made us realize that our previous description was not sufficiently clear and could be misinterpreted as using a random 70/30 split. We apologize for this ambiguity. In the revised manuscript, we have clarified that the model was not trained or validated using random partitioning. Instead, both the SEGSWAT+ calibration and the ML residual-correction training were based on the same chronological 70/30 partition, and no random shuffling was applied at any stage.

Major comments:

1. Model Calibration & Training Strategy

I have serious concerns regarding the model calibration and training strategy, which fundamentally undermine the credibility of the proposed hybrid modeling framework.

Based on the authors' response, the SEGSWAT+ model appears to be calibrated using historical data presumably 1931–1969; however, the manuscript does not explicitly state the calibration period, aside from an gray box in Figure 6. The calibration and validation periods should be clearly specified, including in the figure caption. The machine learning ML component is then trained using a 70/30 split, which appears to be randomly partitioned.

This design is methodologically inconsistent and inappropriate for a hybrid framework. The ML model is intended to learn residuals or corrections from the process-based model; therefore, its training data must be strictly aligned with the calibration period of SEGSWAT+. Instead, the current approach mixes data across time via random splitting, which violates the temporal structure of the problem.

More critically, the use of random splits introduces potential information leakage: streamflow variation pattern after 1970 was implicitly learned by ML. As a result, the reported model performance is likely inflated and does not reflect true temporal extrapolation capability. This issue is not minor, it directly affects the validity of the conclusions drawn from the modeling results, especially the future projection.

A rigorous and appropriate approach would require:

(1) calibrating SEGSWAT+ over a clearly defined historical period e.g., 1931–1969,

(2) training the ML model exclusively on that same period, and

(3) evaluating both the original and ML-corrected models on a temporally independent validation period e.g., post-1970.

Without such a temporally consistent framework, the current results cannot be considered a reliable assessment of model performance or generalization. I strongly recommend a complete revision of the calibration–training–validation strategy.

Response

We sincerely thank the reviewer for raising this critical concern. We fully agree that, for a hybrid framework in which the ML component learns residuals from a process-based model, the training period of the ML residual-correction model must be temporally consistent with the calibration period of the process-based model. We also fully agree that random splitting of hydrological time series would be inappropriate because it could introduce information leakage and inflate validation performance.

We apologize that our previous description did not make the calibration, training, and validation strategy sufficiently clear. The reviewer's interpretation that the ML training used a randomly partitioned 70/30 split is understandable based on the earlier wording, but this was not the procedure used in our analysis.

To address this issue, we have substantially revised Section 2.3.1 and Figure 6. The revised manuscript now explicitly states that:

(1) SEGSWAT+ calibration and ML residual-correction training were conducted using the same chronological 70/30 partition.

(2) The first 70% of available observations within a given period were used for both SEGSWAT+ calibration and ML residual-correction training.

(3) The remaining subsequent 30% were used for temporally independent validation.

(4) No random shuffling was applied at any stage.

The revised text in Section 2.3.1 is as follows:

"For each historical period and each station with available observations, the SEGSWAT+ calibration and the ML residual-correction training were conducted using the same chronological 70/30 partition. Specifically, the first 70% of the available observed records within a given period were used for SEGSWAT+ calibration and ML residual training, and the remaining subsequent 30% were used for temporally independent validation. No random shuffling was applied at any stage. For the naturalized inflow reconstruction, both the SEGSWAT+ parameter set and the ML residual-correction model calibrated/trained during P1 were applied to the full 1931–2024 meteorological forcing series."

We believe this directly addresses the reviewer's concern about possible random partitioning and temporal leakage. The ML residual-correction component was not trained by randomly mixing pre- and post-1970 data. Instead, the calibration/training and validation subsets followed chronological ordering.

In addition, we revised the Figure 6 caption to explicitly specify the calibration/training and validation periods.

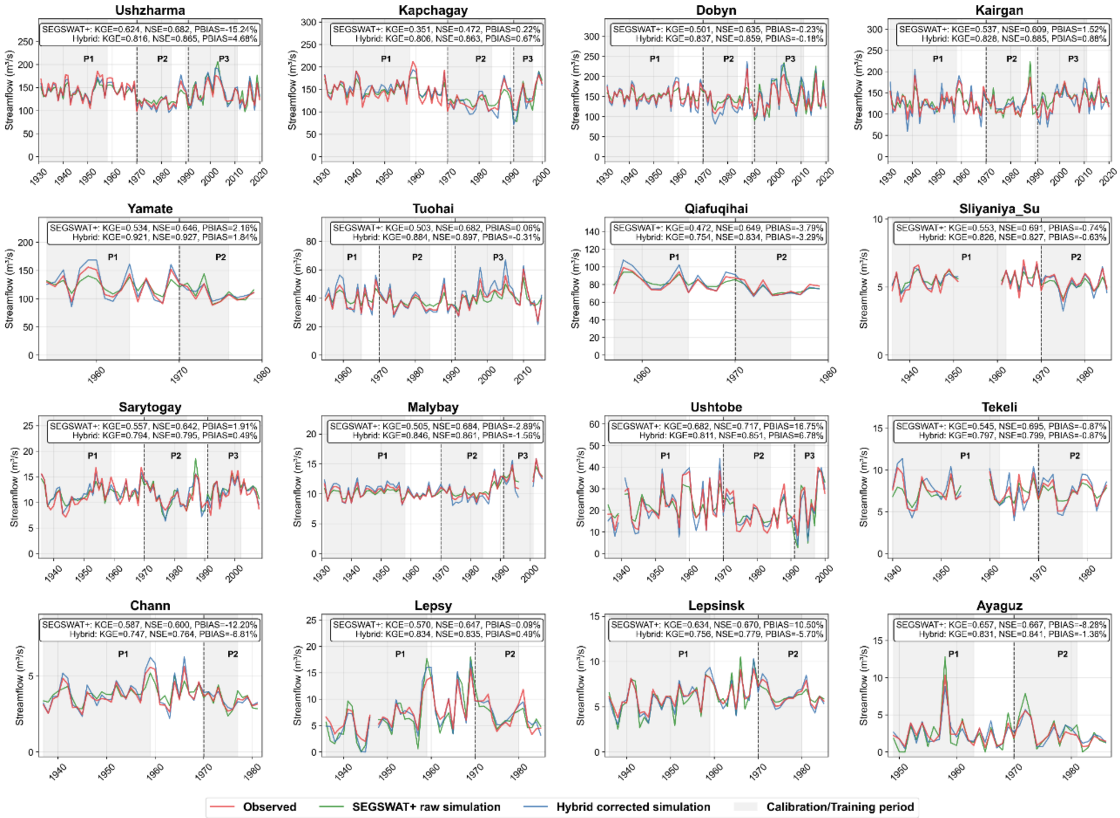

Figure 6. Comparison between observed and simulated streamflow for all hydrological stations with available observations. The figure shows observed streamflow, raw SEGSWAT+ simulation, and hybrid corrected simulation. The shaded gray areas indicate the chronological calibration/training periods used consistently for both SEGSWAT+ calibration and ML residual-correction training, whereas the unshaded subsequent segments indicate temporally independent validation periods. No random shuffling was applied at any stage. Vertical dashed lines mark the boundaries of the three historical periods: P1, 1931–1969; P2, 1970–1990; and P3, 1991–2024. The performance metrics shown in each panel are the mean values across the validation periods of P1, P2, and P3 for the corresponding station.

We have also added a clearer connection between Figure 6, Appendix Table B3, and Supplementary Tables S1 and S2. The revised text now states:

"To quantitatively demonstrate the improvement achieved by this hybrid approach, representative performance summaries for the five main river systems are provided in Appendix Table B3, while detailed station- and period-specific calibration/validation information and performance metrics are provided in Supplementary Tables S1 and S2. For station-period combinations without sufficient P3 observations, station-specific P3 validation metrics were not reported. In these cases, simulations used the most recent parameter sets constrained by available discharge observations."

We also recognize the reviewer's point that, for naturalized inflow reconstruction, using a model trained under the reference period and applying it to later periods involves temporal extrapolation. To make this distinction clear, we now explicitly state that the P1-calibrated/trained model was used for naturalized inflow reconstruction, whereas the actual streamflow reconstruction used period-specific calibration/training where observations were available.

We agree that hydrological modeling over such a long historical period remains subject to uncertainty, particularly because climate, glacier conditions, and human water use changed substantially after 1970. For this reason, we retained and clarified the discussion of cryospheric dynamics and methodological limitations in the Discussion section. This section acknowledges that using fixed P1 parameters to reconstruct naturalized inflow may not fully capture evolving glacier geometry and meltwater response in recent decades.

To address the reviewer's concern, we revised not only the methodological description but also the figure caption, supporting tables, and discussion of limitations, so that the temporal structure of the hybrid modeling framework is transparent and reproducible.

Table B3. Representative mean performance comparison of SEGSWAT+ raw simulations and hybrid corrected simulations across calibration/training and validation phases.

River System	Station	Phase	Metric	SEGSWAT+ (Raw)	Hybrid Model (Corrected)
Ili	Ushzharma	Calibration	KGE	0.671	0.864
			NSE	0.796	0.903
			PBIAS (%)	-15.74	5.35
		Validation	KGE	0.624	0.816
			NSE	0.682	0.865
			PBIAS (%)	-15.24	4.68
Karatal	Ushtobe	Calibration	KGE	0.704	0.853
			NSE	0.787	0.894
			PBIAS (%)	4.66	3.10
		Validation	KGE	0.682	0.811
			NSE	0.717	0.851
			PBIAS (%)	16.75	6.78
Aksu	Chann	Calibration	KGE	0.684	0.757
			NSE	0.696	0.856
			PBIAS (%)	-12.70	-7.31
		Validation	KGE	0.587	0.747
			NSE	0.60	0.764
			PBIAS (%)	-12.20	-6.81
Lepsy	Lepsinsk	Calibration	KGE	0.658	0.806
			NSE	0.712	0.845
			PBIAS (%)	10.87	-6.27
		Validation	KGE	0.634	0.756
			NSE	0.670	0.779
			PBIAS (%)	10.50	-5.70
Ayaguz	Ayaguz	Calibration	KGE	0.667	0.848
			NSE	0.677	0.901
			PBIAS (%)	-18.98	0.36
		Validation	KGE	0.657	0.831
			NSE	0.667	0.841
			PBIAS (%)	-8.28	-1.36

Table S1. Chronological calibration/training and validation periods for each hydrological station and historical period.

River System	Station	Period	Calibration/training years	Validation years	Temporal Resolution
Ili	Ushzharma	P1	1931-1958	1959-1970	Monthly, Yearly
Ili	Ushzharma	P2	1971-1984	1985-1989	Monthly
Ili	Ushzharma	P3	1991-2011	2012-2020	Yearly
Ili	Kapchagay	P1	1931-1958	1959-1970	Monthly, Yearly
Ili	Kapchagay	P2	1971-1984	1985-1989	Monthly
Ili	Kapchagay	P3	1991-1997	1998-2000	Yearly
Ili	Dobyn	P1	1931-1958	1959-1970	Yearly
Ili	Dobyn	P2	1971-1984	1985-1990	Yearly
Ili	Dobyn	P3	1991-2011	2012-2020	Yearly
Ili	Kairgan	P1	1931-1958	1959-1970	Yearly
Ili	Kairgan	P2	1971-1984	1985-1990	Yearly
Ili	Kairgan	P3	1991-2011	2012-2020	Yearly
Ili	Yamate	P1	1953-1964	1965-1970	Monthly
Ili	Yamate	P2	1971-1976	1977-1979	Monthly
Ili	Tuohai	P1	1955-1965	1966-1970	Monthly
Ili	Tuohai	P2	1971-1984	1985-1990	Monthly, Yearly
Ili	Tuohai	P3	1991-2007	2008-2015	Yearly
Ili	Qiafuqihai	P1	1958-1965	1966-1970	Monthly
Ili	Qiafuqihai	P2	1971-1976	1977-1979	Monthly
Ili	Sliyaniya Su	P1	1936-1962	1963-1970	Monthly
Ili	Sliyaniya Su	P2	1971-1980	1981-1986	Monthly
Ili	Sarytogay	P1	1935-1959	1960-1970	Monthly
Ili	Sarytogay	P2	1971-1984	1985-1989	Monthly
Ili	Sarytogay	P3	1991-2002	2003-2008	Monthly
Ili	Malybay	P1	1936-1958	1959-1970	Monthly
Ili	Malybay	P2	1971-1984	1985-1989	Monthly
Ili	Malybay	P3	1991-2001	2002-2004	Monthly
Karatal	Ushtobe	P1	1936-1959	1960-1970	Monthly
Karatal	Ushtobe	P2	1971-1984	1985-1989	Monthly
Karatal	Ushtobe	P3	1991-1997	1998-2000	Yearly
Karatal	Tekeli	P1	1940-1962	1963-1970	Monthly
Karatal	Tekeli	P2	1971-1979	1980-1986	Monthly
Aksu	Chann	P1	1937-1959	1960-1970	Monthly
Aksu	Chann	P2	1971-1977	1978-1983	Monthly
Lepsy	Lepsy	P1	1936-1959	1960-1970	Monthly
Lepsy	Lepsy	P2	1971-1980	1981-1989	Monthly
Lepsy	Lepsinsk	P1	1936-1959	1960-1970	Monthly
Lepsy	Lepsinsk	P2	1971-1980	1981-1985	Monthly
Ayaguz	Ayaguz	P1	1949-1963	1964-1970	Yearly
Ayaguz	Ayaguz	P2	1971-1981	1982-1986	Yearly

Table S2. Period-specific training and validation performance metrics for raw SEGSWAT+ simulations and hybrid corrected simulations.

River System	Station	Period	Phase	KGE (SEGSWAT+)	KGE (Hybrid)	NSE (SEGSWAT+)	NSE (Hybrid)	PBIAS (%) (SEGSWAT+)	PBIAS (%) (Hybrid)
Ili	Ushzharma	P1	Training	0.671	0.913	0.854	0.929	-16.8	2.54
Ili	Ushzharma	P1	Validation	0.606	0.867	0.689	0.923	-16.3	2.04
Ili	Ushzharma	P2	Training	0.624	0.831	0.703	0.923	-18.34	7.01
Ili	Ushzharma	P2	Validation	0.614	0.826	0.693	0.913	-17.84	6.51
Ili	Ushzharma	P3	Training	0.717	0.848	0.83	0.858	-12.08	6.5
Ili	Ushzharma	P3	Validation	0.653	0.756	0.663	0.759	-11.58	5.48
Ili	Kapchagay	P1	Training	0.596	0.873	0.735	0.874	4.81	4.31
Ili	Kapchagay	P1	Validation	0.479	0.863	0.703	0.864	4.17	3.67
Ili	Kapchagay	P2	Training	0.525	0.814	0.654	0.815	4.83	1.79
Ili	Kapchagay	P2	Validation	0.515	0.804	0.644	0.805	-4.33	-1.72
Ili	Kapchagay	P3	Training	0.245	0.761	0.255	0.923	-6.97	-6.54
Ili	Kapchagay	P3	Validation	0.058	0.751	0.068	0.921	0.81	0.06
Ili	Dobyn	P1	Training	0.606	0.901	0.79	0.902	1.23	0.73
Ili	Dobyn	P1	Validation	0.48	0.898	0.705	0.899	0.35	0.26
Ili	Dobyn	P2	Training	0.496	0.824	0.591	0.825	8.06	7.56
Ili	Dobyn	P2	Validation	0.444	0.814	0.581	0.815	-4.78	-4.28
Ili	Dobyn	P3	Training	0.616	0.808	0.69	0.874	6.54	6.04
Ili	Dobyn	P3	Validation	0.579	0.798	0.618	0.864	3.74	3.48
Ili	Kairgan	P1	Training	0.626	0.846	0.689	0.859	2.8	2.3
Ili	Kairgan	P1	Validation	0.58	0.836	0.679	0.849	-0.84	-0.54
Ili	Kairgan	P2	Training	0.546	0.904	0.549	0.905	5.99	5.49
Ili	Kairgan	P2	Validation	0.43	0.894	0.546	0.895	1.69	1.47
Ili	Kairgan	P3	Training	0.607	0.763	0.698	0.92	4.29	1.79
Ili	Kairgan	P3	Validation	0.6	0.753	0.602	0.91	3.71	1.72
Ili	Yamate	P1	Training	0.547	0.927	0.708	0.929	-4.35	-3.85
Ili	Yamate	P1	Validation	0.516	0.917	0.698	0.927	1.27	1.13
Ili	Yamate	P2	Training	0.625	0.928	0.633	0.929	5.66	3.13
Ili	Yamate	P2	Validation	0.552	0.925	0.594	0.926	3.05	2.55
Ili	Tuohai	P1	Training	0.512	0.891	0.697	0.892	6.99	6.49
Ili	Tuohai	P1	Validation	0.451	0.881	0.687	0.882	-3.36	-2.86
Ili	Tuohai	P2	Training	0.536	0.9	0.739	0.901	1.99	0.67
Ili	Tuohai	P2	Validation	0.49	0.882	0.689	0.883	1.25	0.39
Ili	Tuohai	P3	Training	0.579	0.899	0.68	0.926	-9.15	-8.65
Ili	Tuohai	P3	Validation	0.569	0.889	0.67	0.925	2.3	1.53
Ili	Qiafuqihai	P1	Training	0.562	0.801	0.738	0.92	4.85	4.35
Ili	Qiafuqihai	P1	Validation	0.522	0.754	0.701	0.91	-4.19	-3.69
Ili	Qiafuqihai	P2	Training	0.62	0.849	0.805	0.85	-3.47	-2.97
Ili	Qiafuqihai	P2	Validation	0.421	0.755	0.597	0.758	-3.38	-2.88
Ili	Sliyaniya Su	P1	Training	0.598	0.874	0.767	0.875	3.15	1.92
Ili	Sliyaniya Su	P1	Validation	0.522	0.864	0.701	0.865	2.49	1.43
Ili	Sliyaniya Su	P2	Training	0.681	0.804	0.746	0.805	-4.19	-3.07
Ili	Sliyaniya Su	P2	Validation	0.585	0.788	0.681	0.789	-3.98	-2.68

Ili	Sarytogay	P1	Training	0.603	0.806	0.754	0.831	4.39	3.89
Ili	Sarytogay	P1	Validation	0.575	0.796	0.744	0.797	-1.53	-1.02
Ili	Sarytogay	P2	Training	0.686	0.791	0.769	0.801	-1.4	-1.29
Ili	Sarytogay	P2	Validation	0.529	0.756	0.552	0.759	1.13	0.88
Ili	Sarytogay	P3	Training	0.584	0.884	0.719	0.885	6.63	1.84
Ili	Sarytogay	P3	Validation	0.566	0.829	0.631	0.83	6.14	1.61
Ili	Malybay	P1	Training	0.595	0.837	0.777	0.838	-3.62	-3.12
Ili	Malybay	P1	Validation	0.574	0.781	0.752	0.782	-3.01	-2.51
Ili	Malybay	P2	Training	0.47	0.886	0.674	0.887	4.34	3.84
Ili	Malybay	P2	Validation	0.46	0.876	0.664	0.877	-2.86	-1.12
Ili	Malybay	P3	Training	0.511	0.892	0.645	0.925	3.99	3.86
Ili	Malybay	P3	Validation	0.48	0.882	0.635	0.924	-2.81	-1.06
Karatal	Ushtobe	P1	Training	0.796	0.912	0.873	0.921	16.29	7.99
Karatal	Ushtobe	P1	Validation	0.757	0.873	0.767	0.885	15.28	7.7
Karatal	Ushtobe	P2	Training	0.696	0.817	0.853	0.922	18.06	9.9
Karatal	Ushtobe	P2	Validation	0.679	0.786	0.759	0.883	15.1	4.54
Karatal	Ushtobe	P3	Training	0.621	0.83	0.636	0.84	-20.38	-8.59
Karatal	Ushtobe	P3	Validation	0.611	0.774	0.626	0.784	19.88	8.09
Karatal	Tekeli	P1	Training	0.602	0.761	0.652	0.769	-5.15	-4.65
Karatal	Tekeli	P1	Validation	0.57	0.757	0.64	0.759	1.53	1.03
Karatal	Tekeli	P2	Training	0.549	0.884	0.761	0.885	3.53	3.03
Karatal	Tekeli	P2	Validation	0.52	0.838	0.75	0.839	-3.28	-2.78
Aksu	Chann	P1	Training	0.706	0.754	0.721	0.784	-11.55	-7.31
Aksu	Chann	P1	Validation	0.696	0.744	0.711	0.774	-11.05	-6.81
Aksu	Chann	P2	Training	0.662	0.76	0.672	0.927	-13.84	-7.31
Aksu	Chann	P2	Validation	0.479	0.75	0.489	0.753	-13.34	-6.81
Lepsy	Lepsy	P1	Training	0.692	0.921	0.697	0.922	-10.93	-7.58
Lepsy	Lepsy	P1	Validation	0.682	0.911	0.687	0.912	3.1	1.06
Lepsy	Lepsy	P2	Training	0.592	0.826	0.728	0.827	-4.6	-4.17
Lepsy	Lepsy	P2	Validation	0.458	0.756	0.608	0.758	-2.92	-0.08
Lepsy	Lepsinsk	P1	Training	0.664	0.819	0.674	0.829	11.29	-6.36
Lepsy	Lepsinsk	P1	Validation	0.626	0.751	0.636	0.757	10.79	-5.72
Lepsy	Lepsinsk	P2	Training	0.652	0.794	0.75	0.861	10.45	-6.18
Lepsy	Lepsinsk	P2	Validation	0.642	0.761	0.705	0.801	10.22	-5.68
Ayaguz	Ayaguz	P1	Training	0.774	0.87	0.784	0.88	-25.04	3.97
Ayaguz	Ayaguz	P1	Validation	0.764	0.86	0.774	0.87	-4.12	-3.47
Ayaguz	Ayaguz	P2	Training	0.56	0.826	0.57	0.921	-12.93	-3.25
Ayaguz	Ayaguz	P2	Validation	0.55	0.801	0.56	0.811	-12.43	0.75

2. Scientific Rigor and Referencing

I also find the scientific writing and use of references to be below the standard expected for publication. Many statements in the manuscript are presented without appropriate supporting evidence, which weakens the overall rigor of the study.

Examples include:

• Line 42: "As an endorheic lake, all its inflow is ultimately lost to evaporation."
This statement is inaccurate or, at best, incomplete. Other components such as groundwater seepage and human withdrawals can play significant roles in the water balance. This claim requires either qualification or proper referencing.

• Line 43: "While increasing glacier melt can temporarily raise inflow…"
Although revised in response to earlier comments, this statement still lacks any citation. Given that this is a well-established concept, the absence of references is not acceptable.

• Line 498: "As noted by the reviewers…"
This phrasing is inappropriate for a scientific manuscript. Scientific arguments should be supported by peer-reviewed literature, not by referencing the review process itself.

These issues suggest a broader lack of rigor in how claims are formulated and supported throughout the manuscript. A thorough revision is needed to ensure that all scientific statements are properly justified with relevant literature or data.

Response

We thank the reviewer for pointing out these issues. We agree that scientific statements should be supported by appropriate references or by the results of the study, and that the manuscript should not refer to the review process as support for an argument. We have revised the manuscript accordingly.

First, we revised the statement about Lake Balkhash as an endorheic lake. The original sentence was overly simplified because it implied that evaporation is the only relevant loss term. The revised text now reads:

"As an endorheic lake, Lake Balkhash has no surface outflow, and its long-term water balance is primarily regulated by river inflow, direct precipitation, lake evaporation, and, to a lesser extent, groundwater exchange and human withdrawals (Deng et al., 2011; Wang et al., 2022)."

This revision provides a more complete description of the lake water balance and adds supporting references.

Second, we added references for the glacier melt and "peak water" statement. The revised text is:

"Together with recent evidence of changing hydroclimatic conditions in the Tianshan Mountains, increasing glacier melt can temporarily enhance river inflow during the transient 'peak water' phase, but continued glacier mass loss ultimately reduces solid-water storage and may threaten long-term water availability (Huss and Hock, 2018; Jin et al., 2024; Sorg et al., 2012)."

This revision supports the statement with relevant literature and clarifies that the glacier-melt contribution is transient rather than indefinitely sustainable.

Third, we removed the phrase "As noted by the reviewers" from the manuscript. We agree that this wording is not appropriate in the final manuscript. The relevant sentence has been rewritten so that the argument is supported by literature and manuscript analysis rather than by reference to the review process.

In addition to these specific changes, we reviewed the manuscript more broadly to improve scientific rigor and wording. We checked that key claims are either supported by references or directly linked to our results, and we moderated statements that were too strong or insufficiently qualified.

We appreciate the reviewer's comment, which helped us improve the accuracy and professionalism of the manuscript.

References

Deng, M., Wang, Z., and Wang, J.: Analysis of Balkhash Lake ecological water level evolvement and its regulation strategy, Journal of Hydraulic Engineering, 42, 403–413, 2011.

Huss, M. and Hock, R.: Global-scale hydrological response to future glacier mass loss, Nature Climate Change, 8, 135–140, https://doi.org/10.1038/s41558-017-0049-x, 2018.

Jin, C., Wang, B., Cheng, T. F., Dai, L., and Wang, T.: How much we know about precipitation climatology over Tianshan Mountains—the Central Asian water tower, npj Climate and Atmospheric Science, 7, 1–10, https://doi.org/10.1038/s41612-024-00572-x, 2024.

Sorg, A., Bolch, T., Stoffel, M., Solomina, O., and Beniston, M.: Climate change impacts on glaciers and runoff in Tien Shan Central Asia, Nature Climate Change, 2, 725–731, https://doi.org/10.1038/nclimate1592, 2012.

Wang, Z., Huang, Y., Liu, T., Zhong, R., Zan, C., and Wang, X.: Analyzing the water balance of Lake Balkhash and its influencing factors, Arid Zone Research, 39, 400–409, 2022.