Response to Reviewer 3’s Comments:

This manuscript presents a hybrid modelling framework that couples WRF-Hydro with an LSTM-Attention residual correction module to attribute runoff decline and hydrological drought intensification in the Xilin River Basin from 1980 to 2020. The topic is timely and relevant, particularly given increasing interest in combining physically based models with machine learning. I also appreciate the author's effort to use deep learning strictly as a residual corrector rather than a full replacement of physical processes, as well as the attempt to examine intra-annual variability.

That said, I have several concerns regarding the conceptual basis and robustness of the attribution results. These issues do not necessarily invalidate the overall modelling framework, but they do significantly weaken the confidence in the central conclusions. I outline these below.

Major comments

1. Conceptual separation between climate change and human activities

The manuscript treats climate change and human activities as two independent drivers, but in practice these are tightly coupled. For example, irrigation and land use changes can modify evapotranspiration and even regional climate, while anthropogenic emissions are themselves a key driver of climate change. Given this, I find it difficult to interpret the reported attribution (e.g., the dominance of human contribution) as a clean physical separation. At present, it reads more like a methodological partition than a true process-based attribution.

Response:

We thank the reviewer for this thoughtful conceptual comment. We fully agree climate change and human activities are tightly coupled, for example, anthropogenic emissions are the fundamental driver of global climate change, and irrigation and land-use change can also influence climate through land–atmosphere interactions (Thiery et al., 2017). We have clarified and addressed this issue as follows.

First, a clarification of conceptual boundaries is needed. In this study, the "climate change" used as external forcing to the model implicitly incorporates the global climate change signal driven by anthropogenic greenhouse gas emissions, whereas "human activities" refers specifically to direct human interventions within the basin, including water withdrawal, grazing, and land-use change. In other words, the influence of emissions on climate change is already incorporated into the climate term via external forcing, while the influence of local interventions on regional climate is relatively limited at the spatial scale of this basin.

The attribution framework used in this study belongs to the operational attribution paradigm widely applied in hydrology, and is consistent in nature with the Budyko framework (Wang and Hejazi, 2011), the climate elasticity method (Yang and Yang, 2011), and scenario-comparison approaches based on hydrological models (Xu et al., 2023). Under this framework, the "climate change contribution" represents the runoff response driven by climate change when the land surface remains in its baseline-period state, while the "human activity contribution" is obtained as a residual term that reflects the combined effects of land-surface changes and human activities on runoff. This decomposition provides a quantification of contributions under scenario comparison, rather than a process-level causal separation. This is a shared premise of operational attribution methods (Duethmann et al., 2020; Xu et al., 2023) and is not a limitation unique to this study.

For the basin in this study, the framework is well suited. The Xilin River Basin is a semi-arid grassland inland-river basin in which agricultural land accounts for only about 2.9% (Fig. 1 of the original manuscript), and irrigated agriculture occupies an even smaller share. The "irrigation–evapotranspiration–regional climate" feedback therefore plays a relatively limited role at the spatial scale of this study. Human influences on runoff in the basin operate primarily through land-surface hydrological processes and modifications of the land surface, rather than through local climate feedbacks. This supports the application of the quantitative attribution framework described in Section 3.5 to this basin. To avoid the methodological decomposition being misread as a physical causal separation, we have added an explicit statement at the beginning of Section 3.5 clarifying that the attribution framework is a methodological operational decomposition rather than a process-level causal separation, and noting that the residual term inevitably absorbs multiple sources of uncertainty.

The revision to Section 3.5 regarding the nature of the attribution framework reads as follows: "It should be noted that this framework represents a methodological operational decomposition rather than a process-based causal separation. The climate change contribution reflects the runoff response simulated by the model under climate inputs, given the assumption that land-surface parameters remain fixed at their baseline-period values. The human activity contribution is obtained as a residual term and mainly reflects the combined influence of human interventions, including land-surface changes and water resources development, on runoff. This decomposition provides a quantification of contributions under scenario comparison, rather than a process-level causal separation between climate and human drivers."

We also fully recognize that fully characterizing the coupling between climate change and human activities at the Earth-system level remains a challenge. We have therefore added a paragraph in the Discussion to acknowledge the nature of the framework and point toward directions for future work.

The revision to the Discussion regarding this limitation reads as follows: "It should also be noted that the attribution framework adopted in this study represents a methodological operational decomposition rather than a process-level causal separation, because climate change and human activities are intrinsically coupled in the Earth system. Future work can explicitly represent such feedback processes by coupling atmospheric, land-surface, and hydrological models, thereby further enhancing the physical interpretability of the attribution results."

2. Attribution framework likely absorbs model error into the "human" component

The approach defines human impact as the residual between observed runoff and model-simulated runoff under climate forcing. This implicitly assumes that the model fully captures climate-driven processes, which is a strong assumption. Any model bias, missing processes, or structural limitations will be included in this residual. Since the manuscript notes that WRF-Hydro has limitations in low-flow conditions, this is particularly concerning for drought analysis. It raises the possibility that part of the reported human contribution is compensating for model deficiencies.

Response:

We thank the reviewer for this suggestion. Attribution methods based on simulation–observation comparison inherently absorb model errors into the residual, which can affect the estimation of the human activity contribution. This is a well-recognized limitation common to such approaches (Yuan et al., 2022). We fully agree with the reviewer’s concern and have addressed it through targeted measures in both the framework design and the diagnostic analysis.

This study adopts a hybrid framework combining WRF-Hydro with an LSTM-Attention residual correction module (WH-DL), and one of its main purposes is precisely to mitigate the issue of "model errors affecting the quantitative analysis." Specifically, during the baseline period, LSTM-Attention learns the structured bias of WRF-Hydro relative to observed runoff. This residual encompasses model bias, structural limitations, and other sources of uncertainty, and the correction substantially improves the accuracy. As a result, the difference between the WH-DL output and observed runoff—which represents the human activity signal—contains far less residual model error than the difference based on WRF-Hydro alone. The effectiveness of this design is supported by the performance comparison in Section 4.2 of the manuscript: during the validation period, WRF-Hydro yields R^2, NSE, and KGE values of 0.51, 0.49, and 0.62, respectively, whereas WH-DL improves these to 0.73, 0.71, and 0.69. It should be noted that the validation period still falls within the baseline period, when human activity intensity is relatively limited. The residual learned by LSTM-Attention is therefore dominated by the systematic model error of WRF-Hydro, although it also inevitably incorporates a small portion of the baseline-period human activity signal. When this correction pattern is extrapolated to the change period, this absorbed signal may make the estimated human contribution somewhat conservative, but it does not alter the qualitative conclusion that human activities are the dominant driver. If the LSTM-Attention correction were not applied, the model error would be entirely attributed to human activities, leading to a more substantial overestimation. The WH-DL framework is designed to mitigate this problem rather than to fully eliminate it.

Regarding the concern that "model limitations under low-flow conditions could affect the drought analysis," we believe this is in fact a strength of the framework's design. Fig. 5 of the original manuscript shows that the accuracy improvement of WH-DL over WRF-Hydro is concentrated in the low-flow and snowmelt-peak periods, which is consistent with the limited capability of WRF-Hydro to represent low-flow processes in semi-arid basins. In other words, the DL residual correction targets precisely the low-flow segments on which the drought analysis most relies, rather than aggravating the problem. This structural improvement provides a relatively reliable basis for the subsequent SRI-12-based hydrological drought attribution analysis.

We also recognize that no residual-based attribution framework can, in principle, completely eliminate the potential influence of model errors (Duethmann et al., 2020). The aim of WH-DL is to minimize, rather than fully eliminate, this influence. On this basis, we constrain the attribution results using multiple independent lines of evidence: (i) the substantial accuracy improvement of WH-DL over WRF-Hydro; (ii) the human contribution remains within 57.45%–67.99% under a ±2-year perturbation of the change-point year (see our response to Comment 3); (iii) the correction terms exhibit a physically interpretable seasonal structure that aligns with the temporal pattern of human activities (see our response to Comment 3); and (iv) human activity indicators show a significant correlation with hydrological drought (see our response to Comment 5). These mutually reinforcing diagnostic results give the central conclusions a reasonable degree of robustness.

3. Baseline assumption of negligible human influence

The attribution relies heavily on the assumption that human impacts are negligible during the baseline period (1980–2000). However, it is unlikely that grazing, land use change, and water use were absent during this time. The authors do acknowledge this limitation, but the implications are not explored. Given how central this assumption is, I would strongly encourage the authors to at least test how sensitive their results are to this choice (e.g., shifting the baseline period or allowing for gradual human influence).

Response:

We thank the reviewer for this key suggestion. We fully agree with the concern regarding the strictness of the "no human influence" assumption for the baseline period. Human activities such as grazing, land-use change, and water resource use did occur in the basin during 1980–2000, but their intensity was clearly weaker than in the period after 2001 (see Fig. S4 of the original manuscript). Following this suggestion, we adopted a change-point perturbation test to systematically assess the robustness of the central conclusions.

First, we revised the relevant statements to more accurately reflect this fact. We acknowledge that the strict assumption of "no human influence during the baseline period" does not fully hold. A more accurate statement is that human influence during the baseline period (1980–2000) was relatively limited, whereas the change period (2001–2020) reflects the combined effects of climate change and human activities. The revised manuscript has adjusted the relevant statements and avoids absolute terms such as "negligible."

The revisions to the statements regarding the baseline-period assumption are as follows:

In Section 3.5, "During the baseline period, runoff changes were minimally influenced by human activities and can be neglected." has been revised to "During the baseline period, runoff variability was predominantly driven by climate variability, with limited anthropogenic influence."

In Section 4.1.1 of the original manuscript, "Accordingly, 1980–2000 was designated as the baseline period. During this stage, runoff changes were primarily driven by climatic factors, with negligible human activity impacts." has been revised to "Accordingly, 1980–2000 was designated as the baseline period, during which runoff variability was predominantly driven by climate variability, with limited anthropogenic influence."

Following the reviewer's suggestion, we re-defined the baseline and change periods using 1999, 2000, 2002, and 2003 as alternative change-points centered around 2001, and quantified the uncertainty of the attribution results (Figure S1). Within the ±2-year perturbation, human activities consistently remain the dominant driver (>50%), and the original estimate of 61.04% (based on 2001) falls within this perturbation range. The specific uncertainty range and the robustness argument have been incorporated into the Discussion of the revised manuscript.

Expansion of the limitation statement regarding the baseline period in the revised manuscript:

We have expanded the statement on the limitation of the baseline-period assumption in the Discussion of the original manuscript (Line 515) by adding the sensitivity-test results above. "Consequently, it strengthens the robustness and applicability of attribution analysis. Nevertheless, the attribution framework in this study assumes that runoff variations before the change-point year were not influenced by human activities. This assumption introduces uncertainty into the quantitative analysis and may bias the calculation of contribution rates. This represents a limitation of the study." has been revised to "Consequently, it strengthens the robustness and applicability of attribution analysis. Nevertheless, the attribution framework assumes that runoff variations during the baseline period were predominantly driven by climate variability, with limited anthropogenic influence. To test the robustness of the attribution to this baseline assumption, we repeated the attribution with change-point years of 1999, 2000, 2002, and 2003 (Figure S1). Across the ±2-year perturbation, the human contribution remained between 57.45% and 67.99% and the climate contribution between 32.01% and 42.55%, with human activities remaining the dominant driver (>50%) in all scenarios. Across all tested partitions, human activities remain the dominant driver, indicating that the central attribution finding is not contingent on the specific choice of change-point year."

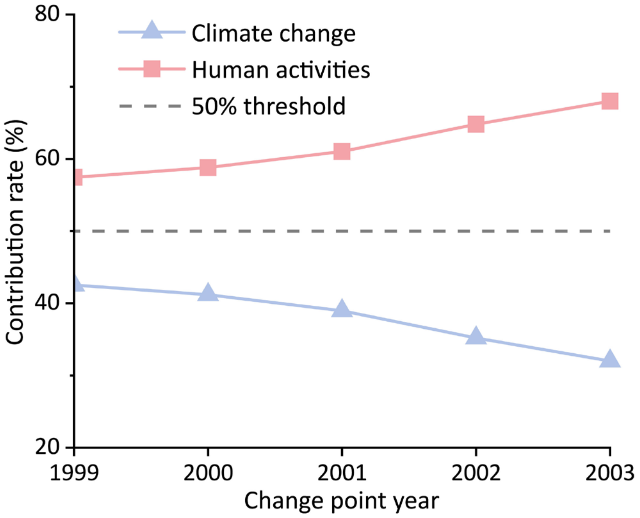


Figure S1. Sensitivity of attribution results to change-point year selection.

Further analysis shows that the human contribution increases as the change-point year moves later, while the climate contribution decreases correspondingly. This trend is consistent with the evolving intensity of human activities in the basin. A later change-point excludes earlier years of moderate human activity from the change period and concentrates the change period on the more recent stage of stronger human activity. This increases the contrast in human activity intensity between the change period and the baseline period, and thereby raises the estimated human contribution. This systematic response of the perturbation also indicates that the sensitivity analysis is physically interpretable rather than a numerical coincidence.

We considered the reviewer's second alternative—assuming a linearly increasing human influence during the baseline period—but did not adopt it after careful evaluation. This approach would require additional parameters that are difficult to constrain independently (such as the starting year, the rate, and the form of the increase), and would instead introduce new sources of uncertainty. In contrast, the change-point perturbation approach requires fewer assumptions, has a clear perturbation structure, and produces results that can be directly compared. We therefore adopted the first alternative suggested by the reviewer.

We thank the reviewer once again for this constructive suggestion. The change-point sensitivity analysis has substantially strengthened the quantitative credibility of the attribution conclusions.

4. Interpretation of the deep learning correction

I understand the intention to preserve physical interpretability by using LSTM as a residual corrector. However, the residual itself is not purely physical. It likely contains a mix of model error, unresolved processes, and possibly human-induced signals. As a result, the corrected runoff series is not strictly a physically consistent output, which complicates its use for attribution. I think this point needs to be discussed more carefully.

Response:

We thank the reviewer for this suggestion on the methodology. The reviewer points out that the residual contains a mixture of model errors, unresolved processes, and potential human signals, and that the corrected runoff is therefore not strictly physically consistent. This is indeed a core issue that any deep-learning-augmented hydrological modeling study must address.

We agree that, in a strict sense, the output of deep-learning correction cannot be equated with the output of a purely physical model, and this is an inherent property of residual-correction methods (Konapala et al., 2020; Reichstein et al., 2019). However, to systematically test whether the WH-DL-corrected runoff (Q_sim) meets basic physical constraints, we have added a series of physical-consistency diagnostics. The simulated runoff is non-negative across all 492 months from 1980 to 2020, with no violations of basic physical constraints. To assess water-balance consistency, given that groundwater observations in the basin are difficult to obtain and that a full ET-based closure involves multiple components that are hard to constrain independently, we adopted the runoff coefficient (Q/P) as a simplified constraint metric. During the baseline period, Q_sim/P (1.58%) is close to Q_obs/P (1.61%), indicating that LSTM-Attention reasonably captures the actual hydrological response within the training domain. During the change period, Q_sim/P (1.44%) is higher than Q_obs/P (0.96%), which is consistent with a counterfactual scenario extrapolated under baseline conditions and not reduced by human activities, and is in line with the attribution logic of this study. The overall runoff coefficient of the basin is consistent with values of about 0.02 reported for the Xilin River Basin in previous studies (Tang et al., 2014), and falls within the typical range for semi-arid basins.

A more diagnostically meaningful aspect is the seasonal structure of the corrections. During the change period, the monthly deviations of Q_sim from Q_obs exhibit a structural pattern that aligns with the seasonal cycle of human activities in the basin. The deviations are mainly concentrated during the snowmelt peak (April), the irrigation season (July–August), and early autumn (September–October), which are precisely the stages most strongly affected by human activities in the basin (anthropogenic water diversion during the spring snowmelt period, agricultural irrigation in summer, the lagged effects of agricultural water consumption, and the sustained influence of grazing). These periods are consistent with the human-activity-dominated stages identified in Section 4.3 of the original manuscript. The deviation in March is close to zero, which is consistent with the conclusion in the original manuscript that the runoff increase during the early snowmelt period is dominated by climate change with a weaker contribution from human activities. If the correction were mainly driven by the seasonal bias of WRF-Hydro itself, the deviation pattern would be expected to align with the seasonal cycle of climate or of model error. The observed deviation, however, aligns specifically with the timing of human activities (spring water diversion, summer irrigation, and the sustained influence of grazing), which is key evidence that the signal captured by the correction is physically interpretable.

The residual indeed contains a mixture of model errors, unresolved processes, and human signals. The core design of the framework in this study is to establish a structural distinction among these signals. Specifically, model errors are identified and corrected to a certain extent through the LSTM-Attention training during the baseline period (see our response to Comment 2). Part of the systematic bias arising from unresolved processes is also compensated by the LSTM-Attention residual correction, thereby improving the simulation accuracy. The human activity signal, in turn, is explicitly identified during the change period through the comparison between Q_sim and Q_obs. This design avoids treating the three types of signals as equivalent, but it cannot perfectly separate them. Any human signal residing within the baseline period would be partially absorbed by LSTM, and this unavoidable limitation has been quantitatively examined through the change-point sensitivity analysis (see our response to Comment 3). The attribution results of this study therefore exhibit a reasonable degree of robustness.

We thank the reviewer once again for this constructive comment. These diagnostic analyses further strengthen the rationale for using WH-DL correction in attribution applications.

5. Lack of explicit representation of human processes

Human impacts are inferred indirectly rather than explicitly represented in the model. Processes such as irrigation return flow, groundwater abstraction, or soil moisture–atmosphere feedback are not included. Without explicitly modelling these, it is difficult to interpret what the human contribution in this study physically represents.

Response:

Thank you for this suggestion. This study does not explicitly represent specific human processes such as irrigation return flow, groundwater pumping, or soil-moisture–atmosphere feedbacks within the hydrological model. Instead, the human influence is inferred indirectly from the difference between observed runoff and the climate-driven simulated runoff. Within this framework, the "human contribution" is not an isolated quantification of any specific human process (such as irrigation water use or groundwater pumping), but rather an integrated measure of all basin runoff responses that cannot be explained by climate forcing. This approach is consistent with operational attribution methods in hydrology (such as the Budyko framework, climate elasticity methods, and scenario-comparison approaches): when basin observations cannot support the independent parameterization of individual human processes, the integrated quantification of contributions under scenario comparison is a more robust strategy (Dey and Mishra, 2017).

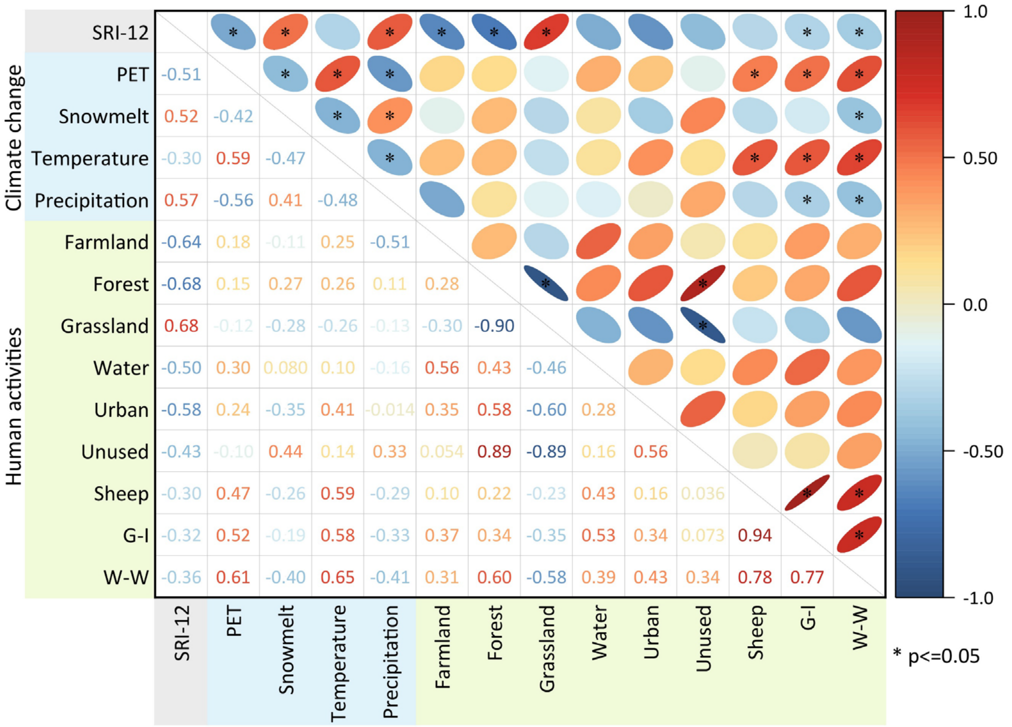
Explicit modeling of the above human processes represents an alternative research pathway with higher process-level interpretability. However, its feasibility relies heavily on the availability of multi-source observations, including metered irrigation data, long-term groundwater-level records, and soil-moisture observations. As a semi-arid grassland inland-river basin, the Xilin River Basin has very limited data of this kind. There is only one national meteorological station in the basin, and systematic groundwater-level and soil-moisture observations are lacking. Irrigation in the basin is dominated by dispersed water withdrawals, for which long-term metered records are difficult to obtain. Under such conditions, incorporating these processes into the physical model would instead introduce numerous parameters that are difficult to constrain independently. This is the reason we adopted the relatively robust approach of residual-based attribution combined with deep-learning correction.

Although these processes are not explicitly represented in the model, we used two independent lines of evidence to indirectly constrain the physical meaning of the "human contribution" measure. First, we added a correlation analysis between multi-source human activity data and hydrological drought to the original manuscript, in order to strengthen the interpretability of the physical meaning of the "human contribution." The results show that hydrological drought is significantly negatively correlated with water withdrawal, grazing intensity, cropland area, and forest area, and significantly positively correlated with grassland area. The physical mechanisms revealed by these correlations—water withdrawal directly reducing runoff, grazing affecting runoff generation by altering the land surface, and forest expansion intensifying water consumption through enhanced evapotranspiration—are consistent with the attribution finding that human activities are the dominant driver. In addition, the seasonal structure of the corrections described in our response to Comment 4 provides further process-level physical correspondence. The human signal identified by the residual-based attribution closely aligns in temporal structure with the activity patterns of known human processes, providing additional support for the physical meaning of the "human contribution" measure.

Newly added content in the Discussion. Revised text in the manuscript:

To enhance the interpretability of the attribution results, we further analyzed the relationships between hydrological drought (SRI-12) and representative climatic factors as well as human activity indicators (Figure S2). The results show that SRI-12 is significantly negatively correlated with PET, cropland, forestland, water withdrawal, and grazing intensity, while it is significantly positively correlated with snowmelt, precipitation, and grassland. These results suggest that hydrological drought is jointly regulated by water supply and atmospheric evaporative demand. Precipitation and snowmelt are key sources of runoff recharge in arid and semi-arid regions and are therefore closely linked to hydrological processes (Berghuijs et al., 2014; Chai et al., 2025; Li et al., 2024). In contrast, enhanced evaporative demand tends to accelerate soil water depletion and reduce basin runoff generation, thereby intensifying hydrological drought.

In addition, grassland and forestland exhibit a significant negative correlation (−0.90). Since 2000, a series of ecological restoration policies, such as afforestation and the Grain-for-Green Program, have been implemented by the government (Li et al., 2023). While these measures have improved the ecological environment, they may also alter basin water balance by enhancing vegetation transpiration (Liu et al., 2026). In particular, under a warming climate and conditions of insufficient water supply, dense vegetation cover may further aggravate hydrological drought by increasing evapotranspiration and consuming deeper soil water storage (Schwärzel et al., 2020; Zhou et al., 2024).



	SRI-12	PET	Snowmelt	Temperature	Precipitation	Farmland	Forest	Grassland	Water	Urban	Unused	Sheep	G-I	W-W
SRI-12		*	*		*	*	*	*					*	*
PET	-0.51		*	*	*							*	*	*
Snowmelt	0.52	-0.42		*	*									*
Temperature	-0.30	0.59	-0.47		*							*	*	*
Precipitation	0.57	-0.56	0.41	-0.48									*	*
Farmland	-0.64	0.18	-0.11	0.25	-0.51									
Forest	-0.68	0.15	0.27	0.26	0.11	0.28		*			*			
Grassland	0.68	-0.12	-0.28	-0.26	-0.13	-0.30	-0.90				*			
Water	-0.50	0.30	0.080	0.10	-0.16	0.56	0.43	-0.46						
Urban	-0.58	0.24	-0.35	0.41	-0.014	0.35	0.58	-0.60	0.28					
Unused	-0.43	-0.10	0.44	0.14	0.33	0.054	0.89	-0.89	0.16	0.56				
Sheep	-0.30	0.47	-0.26	0.59	-0.29	0.10	0.22	-0.23	0.43	0.16	0.036		*	*
G-I	-0.32	0.52	-0.19	0.58	-0.33	0.37	0.34	-0.35	0.53	0.34	0.073	0.94		*
W-W	-0.36	0.61	-0.40	0.65	-0.41	0.31	0.60	-0.58	0.39	0.43	0.34	0.78	0.77	

Figure S2. Correlations of hydrological drought with major climatic factors and human activity indicators.
Note: G-I, grazing intensity; W-W, water withdrawal.

We also fully recognize that indirect constraints cannot substitute for explicit process-level modeling, which represents a higher-level research goal.

The following statement on this limitation has been added to the Discussion of the revised manuscript: "This study quantifies the influence of human activities indirectly, by combining residual-based attribution with correlation analysis using multi-source human activity data. However, processes such as irrigation return flow, groundwater pumping, and soil-moisture–atmosphere feedbacks are not explicitly represented in the present study. In basins where data conditions allow, future work can explicitly represent such processes by coupling atmospheric, land-surface, and hydrological models, thereby further enhancing the process-level interpretability of the attribution results."

We thank the reviewer once again for this constructive suggestion. These clarifications further specify the physical meaning and methodological boundaries of the "human contribution" measure used in this study.

6. Under-constrained attribution problem

More broadly, the study attempts to partition runoff changes into two components (climate vs human) based on a single observed variable. Without additional constraints, this is an underdetermined problem, and some degree of trade-off between components is unavoidable. This makes the quantitative attribution (e.g., 61% vs 39%) less robust than it may appear.

Response:

We thank the reviewer for this valuable comment. We fully agree that quantifying the contributions of climate and human activities from a single observed variable is indeed an underdetermined problem. A certain degree of compensation between the two components is inevitable, and this property cannot be completely eliminated mathematically. This is an inherent issue shared by all simulation–observation-based attribution methods. However, regarding the phrase "without additional constraints," we would like to add some clarification. The attribution in this study is not a simple partitioning of a single observed variable, but an integrated assessment constrained jointly by multiple physical processes and multi-source data. In other words, these constraints provide additional known information for what would otherwise be an underdetermined problem, thereby strengthening the robustness of the attribution conclusion.

From the perspective of physical-process constraints, WRF-Hydro is a distributed physical model whose process equations for water balance and soil-moisture dynamics impose strong constraints on the simulated runoff. The "climate contribution" is therefore not a freely adjustable quantity, but a physical variable strictly governed by the model's controlling equations. From the perspective of input-data constraints, the ERA5 dataset provides multi-variable meteorological forcing, including precipitation, temperature, radiation, and humidity, and has been evaluated against CN05.1 station observations (see the supplementary content at the end of this response for the accuracy assessment). This multi-variable input provides physical constraints on the "climate contribution" that go far beyond what a single observed runoff variable could offer. In terms of parameter calibration, the use of R^2, NSE, and KGE jointly as evaluation criteria reduces the influence of parameter equifinality on the attribution results. In terms of the physical consistency of the simulation results, the correction terms exhibit features that align with the seasonal cycle of human activities (see our response to Comment 4), providing further support for the physical plausibility of the attribution. In terms of other data, indicators of human activities such as water withdrawal, grazing, and land use show clear relationships with hydrological drought (see our response to Comment 5), which further supports the credibility of the "human contribution" conclusion from a perspective beyond runoff observations.

On the other hand, the standard for robustness is not the complete elimination of underdeterminacy, but the stability of the attribution conclusion under perturbations of key assumptions. This study has systematically tested this through the change-point sensitivity analysis described in our response to Comment 3: within the ±2-year perturbation of the change-point year, the human contribution remains between 57.45% and 67.99%, and human activities remain the dominant driver (>50%) in all perturbation scenarios. This range provides a quantifiable bound on the robustness of the attribution conclusion.

It should be noted that, although the quantitative contribution rates carry an uncertainty range, the central qualitative conclusion of this study—that human activities are the dominant driver of the runoff decline in the basin after 2001—holds across all perturbation scenarios. The robustness of this conclusion stems from the cross-validation of the multiple independent constraints described above, rather than from any single line of evidence.

The validation results and corresponding explanation are as follows:

Due to the sparse observational network within the basin, with only one meteorological station (Xilinhot) available, station observations alone are insufficient to fully represent the basin-wide climatic variability. Therefore, this study used the CN05.1 dataset as the reference data to evaluate the reliability of the forcing data. CN05.1 is a high-resolution gridded observational dataset developed by interpolating daily observations from more than 2,400 national meteorological stations across China (Wu and Gao, 2013). Considering that precipitation and temperature are the dominant meteorological drivers of runoff generation and land-surface water balance in WRF-Hydro, the validation of the forcing data was focused on these two variables. Figure S3 presents the comparison between the daily meteorological forcing data and the reference observations over the Xilin River Basin. The results show that WRF-simulated temperature agrees well with the reference data, with a CC of 0.99, a bias of 0.69 °C, and an RMSE of 2.2 °C. For precipitation, WRF also reproduces the daily variability reasonably well, with a CC of 0.65, a bias of 0.12 mm d^{-1}, and an RMSE of 2.3 mm d^{-1}. Overall, these validation results indicate that the WRF downscaled forcing data used in this study have acceptable accuracy in the study area and can provide reliable support for the subsequent WRF-Hydro simulations and attribution analysis.

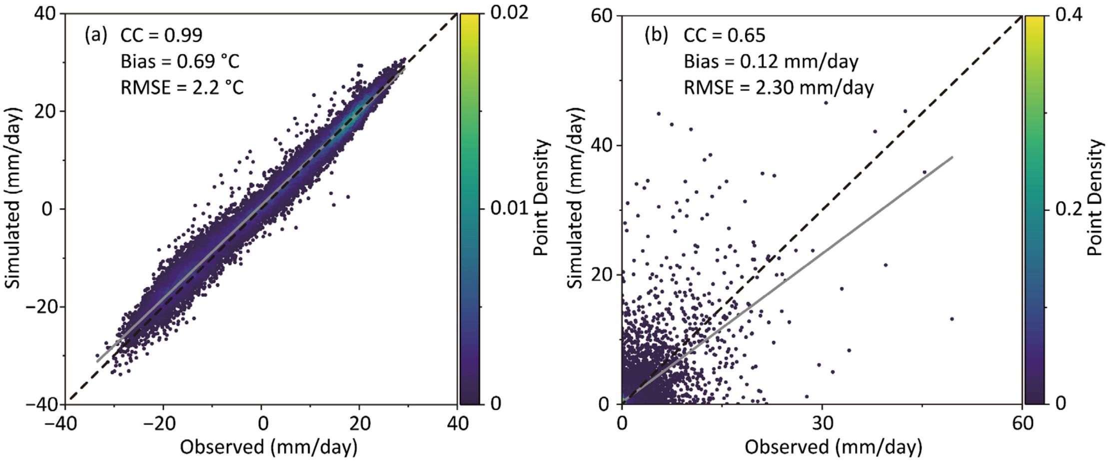


Figure S3. Comparison of daily WRF downscaled temperature (a) and precipitation (b) against reference observations.

Note: CC, correlation coefficient; Bias, mean bias; RMSE, root mean square error.

We thank the reviewer once again for this constructive suggestion. This series of clarifications has strengthened the robustness of the attribution conclusions.

Minor comments

1. Terminology clarity

It would help to more clearly distinguish between anthropogenic climate change and local human interventions, as the current terminology can be confusing.

Response:

We thank the reviewer for this suggestion, which helps improve the clarity of the terminology. The ambiguity raised by the reviewer is conceptually related to our response to Major Comment 1. In this study, "climate change" and "human activities" differ in both their scale and their pathways of influence. The former refers to the overall change of the climate system acting on the basin, driven by both natural climate variability and anthropogenic greenhouse gas forcing. The latter refers specifically to direct human interventions within the basin (such as water withdrawal and land-use change). To avoid confusion, we have systematically clarified the relevant terminology in the revised manuscript.

Specifically, we have added a clarification of the terminology in the description of the attribution framework in Section 3.5, making the meaning of each term explicit. The original sentence "Climate change and human activities jointly drive runoff variability." has been revised to "Climate change (representing the regional manifestation of large-scale climate processes, driven by both natural variability and anthropogenic forcing transmitted through the atmosphere) and local human activities (representing direct in-basin interventions on land surface and water resources, such as grazing, land use change, and water withdrawal) jointly drive runoff variability."

Other parts of the revised manuscript have been checked for consistency with this distinction. The shorter form "human activities" has been retained in contexts where no ambiguity arises, in order to avoid unnecessary verbosity.

2. Uncertainty quantification

The attribution results are presented as point estimates without uncertainty bounds. Given the assumptions involved, some form of uncertainty or sensitivity analysis would greatly strengthen the manuscript.

Response:

We thank the reviewer for this suggestion, which is related to Major Comment 3. We agree that presenting the attribution result as a single point estimate can give readers an impression of overconfidence. In the revised manuscript, we have adjusted the attribution result from a point estimate to a robust expression with an uncertainty range.

Specifically, we re-defined the baseline and change periods using 1999, 2000, 2002, and 2003 as alternative change-points centered around 2001, and quantified the uncertainty of the attribution results (please see our response to Major Comment 3 for details). Within the ±2-year perturbation of the change-point year, the human contribution remains between 57.45% and 67.99%, and the climate contribution between 32.01% and 42.55%. Human activities remain the dominant driver (>50%) in all perturbation scenarios, indicating that the central attribution conclusion is robust to the choice of change-point. The original estimate of 61.04% (based on 2001) falls within this perturbation range; the full uncertainty range is provided in our response to Major Comment 3.

In addition to the sensitivity analysis above, the robustness of the attribution result is also cross-validated through multiple independent lines of evidence. These include the mitigation of model-error contamination by the improved accuracy of WH-DL (see our response to Major Comment 2), the physical-consistency diagnosis of the corrected runoff (see our response to Major Comment 4), and the correlation analysis between indicators of human activities (water withdrawal, grazing, and land use) and hydrological drought (see our response to Major Comment 5). These independent lines of evidence reinforce one another, lending strong support to the central qualitative conclusion that human activities are the dominant driver of the runoff decline in the basin.

3. Sensitivity to methodological choices

The manuscript would benefit from exploring the sensitivity of the results to choices such as the baseline period, change-point detection, or model configuration.

Response:

We thank the reviewer for this suggestion regarding methodological sensitivity. We have constrained the three categories of methodological choices—baseline period, change-point detection, and model configuration—through corresponding sensitivity tests or robustness arguments, as described below.

For the sensitivity to the baseline period, we re-defined the baseline and change periods using 1999, 2000, 2002, and 2003 as alternative change-points centered around 2001, and systematically quantified the uncertainty of the attribution results (please see our response to Major Comment 3 for details). Within the ±2-year perturbation of the change-point year, the human contribution remains between 57.45% and 67.99%, and the climate contribution between 32.01% and 42.55%, with human activities remaining the dominant driver (>50%) in all perturbation scenarios.

For the sensitivity to the change-point detection method, the original manuscript already cross-validated the result by combining the Mann-Kendall test with the Pettitt test (Figure 3 of the original manuscript). Both methods independently identified the change-point as 2001. This dual-method validation indicates that the change-point detected in this study does not depend on a single statistical method and is robust to the choice of detection algorithm.

For the robustness of the model configuration, this study applies multi-level methodological justification and parameter testing. At the level of WRF-Hydro parameter calibration, following previous studies in arid and semi-arid basins (Guo et al., 2024; Yu et al., 2023) and considering the sensitivity characteristics of WRF-Hydro parameters, we selected seven key parameters with strong influence on runoff simulation (covering soil hydraulic properties, surface runoff generation, overland flow routing, and channel routing), and applied the one-at-a-time (OAT) method for systematic parameter-by-parameter scanning. This procedure is essentially a parameter sensitivity test, and the final parameter combination was selected based on the joint evaluation criteria of R^2, NSE, and KGE. The physical meaning, default value, candidate range, and final calibrated value of each parameter are summarized in Table S1.

Revised description of parameter calibration in the manuscript:

Referring to previous studies conducted in arid and semi-arid basins (Guo et al., 2024; Yu et al., 2023), and considering the sensitivity characteristics of WRF-Hydro parameters, this study selected several parameters with strong influences on runoff simulation for calibration (Table S1). The calibration was performed using a one-at-a-time (OAT) approach, in which each target parameter was adjusted individually while keeping all other parameters unchanged, and the optimal value was determined based on the runoff simulation performance.

We sincerely thank the reviewer once again for the constructive suggestions at both the conceptual and methodological levels. These suggestions have helped us improve the manuscript in multiple aspects.

Table S1 Model parameters considered in the calibration.

Name	Description	Default	Candidate values	Optimal value
bexp	Pore size distribution index	×1	0.1, 0.4, 0.7, 1, 3, 5, 7, 10	0.7
dksat	Saturated hydraulic conductivity	×1	0.1, 0.5, 0.7, 1, 2, 3, 5, 10	1
smcmax	Saturation soil moisture content	×1	0.1, 0.5, 0.7, 1, 1.5, 2, 3, 5	0.5
REFKDT	Surface runoff parameter	3	1, 2, 3, 3.5, 4, 4.5, 5, 10	4.5
slope	Openness of Bottom drainage boundary	0.1	0.01, 0.03, 0.05, 0.07, 0.1, 0.2, 0.3, 0.5	0.05
OVROUGHRTAC	Overland Manning roughness multiplier	1	0.01, 0.1, 0.3, 0.5, 0.7,1, 3, 5	0.5
mann	Channel Manning roughness	×1	0.1, 0.2, 0.3, 0.5, 0.7, 1, 2, 5	0.5

Note: For parameters with a default value of ×1, the candidate and optimal values also represent multipliers applied to the spatially distributed defaults.

References

Dey, P. and Mishra, A.: Separating the impacts of climate change and human activities on streamflow: a review of methodologies and critical assumptions, Journal of Hydrology, 548, 278–290, https://doi.org/10.1016/j.jhydrol.2017.03.014, 2017.

Duethmann, D., Blöschl, G., and Parajka, J.: Why does a conceptual hydrological model fail to correctly predict discharge changes in response to climate change?, Hydrology and Earth System Sciences, 24, 3493–3511, https://doi.org/10.5194/hess-24-3493-2020, 2020.

Konapala, G., Kao, S.-C., Painter, S. L., and Lu, D.: Machine learning assisted hybrid models can improve streamflow simulation in diverse catchments across the conterminous US, Environmental Research Letters, 15, 104022, https://doi.org/10.1088/1748-9326/aba927, 2020.

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., and Prabhat: Deep learning and process understanding for data-driven Earth system science, Nature, 566, 195–204, https://doi.org/10.1038/s41586-019-0912-1, 2019.

Thiery, W., Davin, E. L., Lawrence, D. M., Hirsch, A. L., Hauser, M., and Seneviratne, S. I.: Present-day irrigation mitigates heat extremes, Journal of Geophysical Research: Atmospheres, 122, 1403–1422, https://doi.org/10.1002/2016JD025740, 2017.

Wang, D. and Hejazi, M.: Quantifying the relative contribution of the climate and direct human impacts on mean annual streamflow in the contiguous United States, Water Resources Research, 47, W00J12, https://doi.org/10.1029/2010WR010283, 2011.

Xu, R., Qiu, D., Wu, C., Mu, X., Zhao, G., Sun, W., and Gao, P.: Quantifying climate and anthropogenic impacts on runoff using the SWAT model, a Budyko-based approach and empirical methods, Hydrological Sciences Journal, 68, 1358–1371, https://doi.org/10.1080/02626667.2023.2218551, 2023.

Yang, H. and Yang, D.: Derivation of climate elasticity of runoff to assess the effects of climate change on annual runoff, Water Resources Research, 47, W07526, https://doi.org/10.1029/2010WR009287, 2011.

Yuan, F., Zhao, C., Jiang, Y., Ren, L., Shan, H., Zhang, L., Zhu, Y., Chen, T., Jiang, S., Yang, X., and Shen, H.: Attribution of climate change and human activities to streamflow variations with a posterior distribution of hydrological simulations, Hydrology and Earth System Sciences, 26, 5315–5339, https://doi.org/10.5194/hess-26-5315-2022, 2022.