Responses to Referee #2 comments:

Manuscript number: **egusphere-2025-4213**

Title: **Is earlier always better? A comparative assessment of rainfall replenishment timing for multiyear drought mitigation**

We sincerely thank you for the valuable comments and suggestions, which have greatly improved our manuscript. We have addressed all the comments and made the necessary revisions. Our responses to each comment are in blue text, while the referees' comments are in black.

This manuscript investigates the role of precipitation timing on reducing the severity of multiyear droughts. The topic is generally well suited for this journal, and the objective clearly identified.

Response: Thank you sincerely for the careful assessment of our work. We are grateful that the study's objective and relevance were recognized. We also truly appreciate the time taken to evaluate the manuscript. Your comments are highly valuable and have guided us in clarifying the conceptual framework, methodology, and implication of our work.

I found, however, a few major drawbacks that limit the scope of the proposed research.

Response: Thank you for pointing out several major issues that currently limit the clarity and scope of the study. We greatly appreciate these constructive comments. In response, we have carefully reconsidered the conceptual framework, the methodological approach, and the interpretation of the results. Detailed, point-by-point responses to each comment are provided below.

To facilitate understanding and provide a convenient reference, we first present Table S1, which lists the principal symbols and acronyms used throughout the manuscript.

Table S1: A list of principal symbols and acronyms

Symbols and Acronyms	**Description**
MYD	MultiYear Drought
PDSI	Palmer Drought Severity Index, a monthly drought index where each value represents the PDSI-defined "drought severity" for that month (the terminology is distinguished from the event-level drought severity, DS, used in this study)
DS	Drought Severity of a drought event, the cumulative deficiency of a drought index below the critical threshold throughout the event
DD	Drought Duration of a drought event, the period during which a drought index is continuously below the critical threshold during the event
DI	Drought Intensity of a drought event, the maximum deviation of the drought index below the critical threshold during the event
DDP	Drought Development Period of a drought event
PIP	DI Point of a drought event, the PDSI value corresponding to the DI

DMQM	Drought Mitigation Quantitative Model
CAFEC	Climatically Appropriate For Existing Conditions
$pr_{baseline}$	the actual precipitation in the targeted month under the baseline (no hypothetical additional precipitation is applied) scenario
$\overline{pr}$	the long-term mean monthly precipitation
pr_{rate}	a dimensionless proportion (0–1) that controls the relative magnitude of the hypothetical precipitation addition, the amount of hypothetical precipitation addition is given by $pr_{rate} \times \overline{pr}$
$pr_{forcing}$	the precipitation in the targeted month after applying hypothetical additional precipitation, $\mathrm{pr_{forcing}} = \mathrm{pr_{baseline}} + \mathrm{pr_{rate}} \times \overline{\mathrm{pr}}$, $\mathrm{pr_{rate}} = 0, 5\%, 10\%, \ldots, 100\%$
$DS_{baseline}$	the actual DS under the baseline (no hypothetical additional precipitation is applied) scenario
$DS_{forcing}$	the DS after the hypothetical additional precipitation is applied
DS_{rate}	the relative change in DS, $DS_{rate}=\frac{DS_{baseline}-DS_{forcing}}{DS_{baseline}}$
k	MYD attenuation efficiency, k characterizes how rapidly additional precipitation is converted into reductions in DS
$k_{theoretical}$	the theoretical k value, $k_{theoretical}$ is obtained by fitting the DMQM analytical solution $\mathrm{DS_{rate}}=1-\exp(-\mathrm{k}\cdot\mathrm{pr_{rate}})$ to the DS_{rate}-pr_{rate} curve derived from the numerical experiments
k_{actual}	the actual k value, represent the empirical sensitivity of DS to small precipitation perturbations, $k_{actual}=\frac{DS_{rate}(pr_{rate}=0.05)}{0.05}$, $pr_{rate}=0.05$ is the step size in the numerical experiments
ε	relative error, $\varepsilon=\left(\frac{k_{theoretical}-k_{actual}}{k_{actual}}\right)\times 100\%$
k_1	$k_{theoretical}$ value at the first month of a MYD event
k_2	$k_{theoretical}$ value at the second month of a MYD event
k_3	$k_{theoretical}$ value at the third month of a MYD event
k_{max}	maximum $k_{theoretical}$ within a MYD event, obtained as the largest monthly k among all months of this event
Δk_{21}	the changes in $k_{theoretical}$ between the second and the first month of a MYD event
Δk_{31}	the changes in $k_{theoretical}$ between the third and the first month of a MYD event
$\mathrm{t_{forcing}}$	the month of hypothetical additional precipitation is applied
$\mathrm{t_1}$	the first month of a MYD event
$\mathrm{t_2}$	the second month of a MYD event
$\mathrm{t_3}$	the third month of a MYD event
$\mathrm{t_{PIP}}$	the month at PIP of a MYD event
$\mathrm{t_{optimal}}$	the optimal timing for drought alleviation
$\mathrm{t_a}$	$\mathrm{t_{optimal}}$ in Conceptual Model 1
$\mathrm{t_b}$	$\mathrm{t_{optimal}}$ in Conceptual Model 2
$\mathrm{t_c}$	$\mathrm{t_{optimal}}$ in Optimal Solution Model

Note: To avoid ambiguity, the monthly values of the Palmer Drought Severity Index (PDSI) are referred to simply as "PDSI" in this study, whereas DS is used exclusively to denote event-level drought severity.

1. The term mitigation is, in my opinion, improperly used or, at least, not well described. The term mitigation is usually associated with anthropogenic action to reduce the impacts of a natural hazard. Here, terms such as "amelioration" or "alleviation" may be more appropriate. In general, the authors discuss about precipitation as it is usually done for irrigation in agricultural districts. While there are some similarities, precipitation over large areas cannot be discussed in the same terms. I suggest a deep revision of the text to avoid any confusion on the strategy discussed here.

Response: We sincerely thank you for this important and constructive comment. We fully agree that the term "mitigation" can create conceptual ambiguity, as it is commonly associated with anthropogenic actions to reduce the impacts of natural hazards. In our study, this term specifically refers to the reduction in drought severity (DS) of a multiyear drought (MYD) event under hypothetical precipitation increases applied within numerical experiments, rather than to any direct human intervention. To avoid confusion, we now consistently use drought attenuation to describe the system's response and drought alleviation to describe the event-level reduction in DS.

To clarify, we explicitly define this process as index-based drought alleviation and describe the system's internal response using drought attenuation efficiency (k). The experiments involve idealized precipitation forcing, implemented by perturbing precipitation in the PDSI model for a single month while keeping other months unchanged. This design isolates how the system responds to controlled water-input perturbations at different drought stages and does not imply operational control over rainfall.

Expressions that could suggest direct management, such as "precipitation-based drought mitigation" or "early intervention", are replaced with conceptually precise terminology, including timing of hypothetical precipitation increases, idealized precipitation forcing, and model-derived drought attenuation efficiency. These revision ensure that the language consistently reflects the synthetic nature of the numerical experiments and avoids giving the impression that precipitation timing can be artificially controlled.

Although the findings are derived from idealized experiments, they provide useful insights into how the system responds to additional water input. These insights can conceptually inform discussions about real-world drought alleviation, since many water management practices, such as reservoir releases, ecological flow restoration, irrigation, or experimental cloud-seeding trials, aim to reduce water deficits caused by prolonged precipitation shortages. Such actions influence soil moisture, hydrological storage, and overall water availability. Because the PDSI framework treats precipitation as the sole water-input variable, the additional precipitation applied in our numerical experiments functions as an equivalent representation of increased available water, regardless of the specific real-world mechanism through which this water is supplied. This allows the numerical experiments to isolate the system's intrinsic sensitivity to controlled water-input perturbations without tying the analysis to any particular management measure. Our results therefore highlight the potential timing and effectiveness of increased water availability, even though the model input is expressed as hypothetical precipitation. Furthermore, the hypothetical perturbations are constructed based on historical MYDs, ensuring that the imposed conditions are grounded in real drought evolution rather than artificially invented scenarios. By consistently emphasizing attenuation for system response and alleviation for event-level improvement, we ensure conceptual clarity and avoid ambiguity in interpreting the model outcomes.

2. The focus on multiyear drought is never clearly explained nor justified. Why the analysis focuses specifically on multiyear drought and not on drought in general? What is the reasoning behind such choice?

Response: We sincerely thank you for this comment. The decision to focus specially on multiyear drought (MYD) is motivated by both scientific and practical considerations.

MYDs represents the most consequential category of drought events because their impacts accumulate across seasons and years. Shorter droughts often allow for partial natural recovery through subsequent rainfall or seasonal hydrological recharge, whereas MYDs progressively deplete soil moisture, groundwater storage, and ecological resilience. In these long-duration events, the timing of additional water input becomes far more critical for drought alleviation, making MYDs an ideal research subjects for examining how hypothetical precipitation increases influence the drought severity (DS) of MYD events. Events of short duration tend to have limit impacts, are not extreme or typical, and often end naturally before any intervention could meaningful influence their progress.

The cumulative and multi-stage nature of MYDs brings uncertainty regarding the most effective timing of additional water input. Because these events occur across several hydrological cycles, the system's response to water-input perturbations is not uniform across months. This provides a meaningful opportunity to investigate drought attenuation efficiency (k) at different months of a prolonged drought, and to identify which stages are most sensitive to additional water input.

MYDs are also the events where index-based drought alleviation is most relevant. The PDSI framework accumulates deficits over time, and alleviating a deep, long-lasting deficit requires understanding how the system converts additional water input into reductions in DS. Our numerical experiments are therefore particularly well suited to MYDs, where the relative importance of water-input timing is magnified.

Beyond the modeling perspective, these persistent, expansive, and difficult-to-recover MYDs have shown an escalating trend in severity, duration, and frequency under climate change (Dai, 2013; Stevenson et al., 2022; Chen et al., 2025). This further highlights the importance of understanding their dynamics and potential alleviation strategies.

These considerations collectively justify why MYDs serve as the focus of this study. In addition, several of the Specific comments (L121, L148, and L149) raise related questions about the definition and treatment of MYDs in our framework, and detailed responses to these comments are provided in their respective locations.

References

Chen, L., Brun, P., Buri, P., Fatichi, S., Gessler, A., McCarthy, M.J., Pellicciotti, F., Stocker, B. and Karger, D.N.: Global increase in the occurrence and impact of multiyear droughts, Science, 387, 278-284, doi: 10.1126/science.ado4245, 2025.

Dai, A.: Increasing drought under global warming in observations and models, Nat. Clim. Change, 3, 52-58, doi: https://doi.org/10.1038/nclimate1633, 2013.

Stevenson, S., Coats, S., Touma, D., Cole, J., Lehner, F., Fasullo, J., and Otto-Bliesner, B.: Twenty-first century hydroclimate: A continually changing baseline, with more frequent extremes, Proc. Natl. Acad. Sci., 119, e2108124119, doi: https://doi.org/10.1073/pnas.2108124119, 2022.

3. The issue of non-linear response is not well discussed in the probabilistic context of drought indicators. The fact that drought severity changes non-linearly with the usual indicators is a well know factor, well reflected by the commonly used classifications of drought indices. In this context, it is not possible to simply compare a unitary change in severity between events occurring at different severity levels, as (example) a change in SPI from 0 to -1 and from -1 to -2 have completely different probabilistic implications. I suggest the authors to carefully revisit their considerations on the "mitigation" efficiency by explicitly taking into account for these disparities.

Response: Thank you very much for this thoughtful and important comment. We fully acknowledge the point that standardized drought indices exhibit nonlinear and uneven probabilistic spacing, and that a one-unit change (e.g., from 0 to -1 and from -1 to -2 in SPI) carries fundamentally different statistical implications, and our framework is built precisely on this understanding. We appreciate this reminder, as it highlights a key conceptual issue that we realized was not clearly explained in our original manuscript. We did not clearly distinguish between the instantaneous drought severity given by PDSI and the event-level drought severity (DS) that is actually used in our study.

To clarify our approach, the "drought severity" used in our analysis is not the instantaneous severity value provided directly by an index such as PDSI, SPI, or SPEI. Instead, DS is an event-based cumulative indicator, computed using run theory as the integral of PDSI series over the drought duration (DD). DS therefore represents the accumulated water deficit over the entire event rather than a single index value. Importantly, the unequal spacing and probabilistic nonlinearity of PDSI are inherently preserved in the calculation of DS. Thus, our method does not rely on comparing a uniform one-unit change of the index across different severity levels. Instead, DS represents the cumulative effect of all PDSI values that naturally vary in magnitude, spacing, and probabilistic meaning.

Moreover, in light of your suggestion, we have carefully revisited our assessment of attenuation efficiency (k). Because DS integrates PDSI over the entire MYD event, the calculation inherently accounts for the unequal probabilistic spacing and nonlinear response of the underlying index. These nonlinear characteristics are central to understanding why the timing of hypothetical water input does not always follow a monotonic "earlier is better" pattern. The selection of an optimal timing in our results is precisely driven by the nonlinear feature of PDSI and DS, rather than any assumption of linear response. This comment helped us recognize that our explanation of this conceptual link needed improvement.

We sincerely appreciate your insightful observations. They have prompted us to clarify key definitions and improve the conceptual rigor of the manuscript (Table S1). We have incorporated clarifications into the responses to the Specific comments (L68-70, L71, L164, L166, L305 and Fig. 8), and to the Specific comment surrounding Eq. 3, explaining the difference between instantaneous PDSI severity and event-level DS, and reinforcing that our framework is fully compliant with the probabilistic structure of standardized drought indices.

4. The terminology used in the paper is sometime rather confusing. Different terms are used to define the same quantities in different part of the paper (e.g., t_optimal, t_a, t_b, t_c, etc.).

Response: We sincerely thank you for pointing out potential confusion in the terminology. We fully acknowledge that multiple labels for related quantities can be confusing if not clearly explained. So here we clarify the notation, explain the conceptual role of each symbol, and show how the figures and conceptual models connect.

In our study, $t_{optimal}$ represents the general optimal timing of hypothetical additional precipitation at the event scale. The other symbols (t_a, t_b, and t_c) are introduced only to denote the optimal timing implied by different conceptual perspectives, as summarized in Fig. 4:

1. t_a corresponds to the optimal timing in Conceptual Model 1 (Figure 4b), which emphasizes the advantage of earlier water input within a drought event. Because the accumulated deficit is still low during the earlier month of a drought event, hypothetical precipitation added at this month produces a larger proportional reduction in drought severity (DS) of a multiyear drought (MYD) event. Under this perspective, t_a tends to approach the first month of a MYD event (t_1).

2. t_b corresponds to Conceptual Model 2 (Figure 4c), which emphasizes the statistical sensitivity of standardized indices. For example, when PDSI values are far below the drought threshold (e.g., PDSI = -4), a small shift in cumulative probability produces a large numerical change in the index compared with values closer to the threshold (e.g., PDSI = -1) under the same perturbation (Fig. 1). Since PDSI values approach their maximum deviation below the threshold when moving toward the drought intensity point (PIP), this perspective indicates a potential drought alleviation advantage for hypothetical precipitation added timing near the month at PIP of a MYD event (t_{PIP}).

3. t_c is the optimal solution form the Optimal Solution Model (Figure 4d) and represents the balanced timing that reconciles the opposing tendencies of Conceptual Model 1 and 2. t_c denotes the overall optimal timing that our study aims to identify.

To reduce the readers' cognitive burden, we provide a consolidated reference table of principal symbols and acronyms. This table lists $t_{optimal}$, t_a, t_b, t_c, t_{PIP}, and other frequently used symbols, and gives a short description of each symbol (Table S1).

Additionally, Fig. 1 introduces the problem by contrasting event-level characteristics and standardized -index statistical features. Figures 1a and 1b illustrate event-level progression, while Figures 1c and 1d illustrate the probabilistic properties of standardized indices. Fig. 4 follows logically from Fig. 1 by converting the problem posed in Fig. 1 into analytical perspectives (Figures 4b and 4c) and proposed solutions (Fig. 4d). Conceptual Models 1 and 2 in Fig. 4 emphasize the two different mechanisms that drive alleviation advantage, and the Optimal Solution Model integrates them to determine t_c.

Several specific comments (L13, L85, L122, L125-128, L166, L194 and Fig. 1) are directly related to this terminology issue and will be addressed in detail in their corresponding replies.

5. Overall, while a recognize some potential in this research, I suggest a careful revision of the text before considering it for publication. I added some further specific comments below, in the hope that they will be useful to improve the quality of the manuscript.

Response: We sincerely thank you for the overall assessment and for recognizing the potential of our research. We fully acknowledge the importance of clarity and conceptual consistency. In response to the specific comments, we have provided detailed clarifications and explanations in this response document, including revised phrasing, figure descriptions, and terminology definitions. We believe these clarifications improve the clarity and conceptual rigor of the study.

Specific comments

6. L12. Remove "especially".

Response: Thank you for this comment. We have removed the word "especially" from line 12 as recommended.

7. L13. The concept of "optimal rainfall replenishment timing (T_optimal)" needs to be better established here.

Response: Thank you for pointing out this issue. Here we clarify that $t_{optimal}$ refers to the month during a multiyear drought event when a given hypothetical water input achieves the greatest alleviation effect, that is the month corresponding to the maximum value of k under the DMQM framework. This concept is further elaborated through the notations t_a, t_b, and t_c, which represent the optimal timing under Conceptual Model 1 (Fig. 4b), Conceptual Model 2 (Fig. 4c), and the Optimal Solution Model (Fig. 4d), respectively. These help to distinguish different mechanisms influencing drought alleviation (Section 3.4). Among these, t_c represents the final $t_{optimal}$ used throughout the study.

8. L15. Calibration period. I suggest rewording.

Response: Thank you for the suggestion. In our experiment, the baseline period must remain unchanged to avoid affecting the hydrological parameters derived from historical conditions. Although the dataset covers 1961-2020, the calibration period (baseline) is set to 1961-1990, and the numerical experiments adding hypothetical water input are applied only to 1991-2020. This ensures that adding hypothetical precipitation to specific months affects only the experiment period and does not retroactively alter the baseline conditions or the associated hydrological parameters. Subsequent comments related to this issue, including those corresponding to L17, L132, L142, and L185, are addressed accordingly in their respective responses.

9. L16. "351 MYD". These are obviously not droughts, but cells under drought. Many of these cells are under the same MYD contemporarily. Please reword, as this statement seems to suggest that 351 MYD were observed in a period of only 60 years.

Response: Thank you for pointing out this potential misunderstanding. We clarify that our analysis is conducted at the grid scale. For each of the 199 grids during 1991-2020, multiyear drought (MYD) occurrences are identified independently, and 351 grid-level MYD instances were identified in total. Some of these grid-level instances may correspond to the same regional MYD event, but our study focuses on the grid-level drought periods as the statistical units for attenuation efficiency analysis, rather than aggregating drought events at the regional scale.

We also clarify why the grid-based treatment is necessary and appropriate for our framework. Grid identification allows us to directly link drought characteristics to the underlying station observations within each grid. When multiple stations fall inside one grid, their drought attenuation efficiency (*k*) are averaged, which helps reduce biases caused by the uneven spatial distribution of stations across China and ensures that gridded results remains observation-based. Working at the grid scale also provides a consistent spatial structure that aligns with the way most hydroclimatic datasets are distributed worldwide. Many datasets used for drought assessment, especially global and regional data products, are available exclusively in gridded format. Mapping station-derived *k* onto grids therefore makes our approach more easily reproducible and transferable to other regions or for future global-scale applications. Additionally, because run theory, as applied in this study, quantifies drought characteristics based on a single time series record, its formulation is inherently suited to station-level or single-grid calculations, making the grid an appropriate methodological unit for analyzing *k*. At the same time, we also recognize that beyond the station and grid scales, drought processes exhibit coherent dynamics at the regional scale. We are therefore exploring clustering-based approaches to identify and analyze regional drought events. This direction represents an important extension of the present work and is being developed as part of our subsequent research. Given that this study focuses on the patterns and statistics at the grid scale, the regional scale clustering analysis will be presented in a separate follow-up study rather than included here.

A clarification has been added to avoid possible confusion regarding the interpretation of the MYD count. Subsequent comments related to this issue, including those concerning L151-152, L325, and Fig. 8, are addressed accordingly in their respective responses.

10. L17. Is the study period 1991-2020 or 1961-2020? The different periods are confusing.

Response: Thank you for pointing this out. The two periods serve different purposes in our framework. The full dataset covers 1961-2020, but the baseline hydrological parameters are derived from 1961-1990, while the numerical experiments adding hypothetical precipitation are conducted only for 1991-2020. This design ensures that the experiments does not retroactively alter the baseline conditions or the associated hydrological parameters, clarifying the distinction between the calibration and study periods.

11. L17-18. A key coefficient (k)… This does not clarify what "k" is.

Response: Thank you for pointing this out. In our framework, *k* quantifies the system's drought attenuation efficiency, namely how rapidly hypothetical additional precipitation is converted into reductions in drought severity (DS) of multiyear drought (MYD) events. Higher *k* values correspond to faster alleviation, while lower *k* values reflect slower hydrological recovery or more persistent drought conditions.

Given the central role of k in our study, we provide a more complete explanation below.

To clarify our approach, the exponential formulation in DMQM is not arbitrarily postulated. Our modeling framework was developed by performing numerical experiments driven by observations from meteorological stations. Before specifying any functional structure, we first examined how drought severity (DS) of a multiyear drought event responds to incremental precipitation increases within controlled numerical experiment. Across events and stations, the numerical experiments consistently demonstrated a clear diminishing return pattern (Fig. 5). When hypothetical additional precipitation was incrementally increased at the same month of a drought event, the reduction in DS showed progressively weaker marginal benefits. The initial portion of added water input produced the strongest alleviation, while subsequent increments at that same month yielded progressively weaker effects. This empirical feature motivated us to seek a formulation capable of representing a bounded recovery process. To formalize this behavior, we drew an analogy with systems exhibiting saturation-type adjustment, similar to first-order kinetic expressions (e.g. $C_t = C_0 \cdot exp(-k \cdot t)$, where C_t denotes the remaining intensity at time t, C_0 is the initial state at t=0, and k is the first-order rate constant controlling the decay speed. The instantaneous decay rate $\frac{dC}{dt}$ is proportional to the current magnitude C_t.). Following this reasoning, we adopted the formulation: $DS_{rate} = 1 - exp(-k \cdot pr_{rate})$. This formulation captures how incremental water input progressively reduces the residual DS (Fig. 1). In this framework, $exp(-k \cdot pr_{rate})$ represents the proportion of remaining DS after a given additional precipitation increment. When the additional water supply is sufficient, the relative reduction in DS approaches a practical upper limit, a pattern similar to physical constraints such as finite soil water storage capacity or vegetation water use limits.

To demonstrate that the exponential structure represents real system behavior, we compared formulations inspired by zero-order, first-order, and second-order kinetic processes using historical multiyear drought events (MYDs) during 1991-2020. Across most stations and events, the first-order structure more accurately reproduced the curvature observed in the numerical experiments at different month of MYDs. We further extended this comparison to an n-order formulation and found that the estimated n converged toward 0.98 across MYDs, which is essentially unity. This outcome strengthens the conclusion that a first-order structure is the most appropriate representation of the observed pattern. The results are summarized in Table S2. We also compared the exponential form against alternative empirical relationships. Logistic curves did not match the observed pattern. Power-law functions were unable to satisfy the required boundary conditions. We note that the same exponential adjustment structure has been adopted in other fields. For instance, Childs et al. (2025) applied a similar functional pattern in modeling climate-driven disease dynamics. The recurrence of this structure across domains further supports its suitability for representing systems with constrained recovery capacity.

Table S2: Comparison of R^2 values for different functional models

Formulation	t_1	t_2	t_3
Zero-order: $DS_{rate} = k \cdot pr_{rate}$	0.992	0.984	0.989
First-order: $DS_{rate} = 1 - exp(-k \cdot pr_{rate})$	0.985	0.996	0.993
Second-order: $DS_{rate} = 1 - 1/(k \cdot pr_{rate} + 1)$	0.982	0.976	0.961
N-order: $DS_{rate} = 1 - [1 - k \cdot (1-n) \cdot pr_{rate}]^{1/1-n}$	0.989	0.996	0.998

Note: t_1, t_2, and t_3 represent the first to the third month of the MultiYear Drought event.

Subsequent comments related to this issue, including those concerning L20, L194, L209, Fig.3 and Fig. 5, are addressed in their respective replies.

References

Childs, M.L., Lyberger, K., Harris, M.J., Burke, M., and Mordecai, E.A.: Climate warming is expanding dengue burden in the Americas and Asia, Proc. Natl. Acad. Sci., 122, e2512350122, doi: https://doi.org/10.1073/pnas.2512350122, 2025.

12. L18. T_optimal should be a subscript.

Response: Thank you for the suggestion. We will carefully check the notation and ensure that $t_{optimal}$ is consistently presented with a subscript throughout the manuscript.

13. L20. This sentence is not clear if you do not clarify what k is.

Response: Thank you for this comment. As clarified in our response to Comment L17-18, *k* quantifies the system's drought attenuation efficiency, namely how rapidly hypothetical additional precipitation is converted into reductions in drought severity (DS) of multiyear drought (MYD) events. Higher *k* values correspond to faster alleviation, while lower *k* values reflect slower hydrological recovery or more persistent drought conditions. k_{max} refers to the maximum monthly *k* within a MYD event and thus identifies the optimal timing for drought alleviation under the same perturbation.

14. L24. T_optimal should be subscript.

Response: Thank you for the suggestion. We will carefully check the notation and ensure that $t_{optimal}$ is consistently presented with a subscript throughout the manuscript.

15. L24-25. This sentence needs to be clarified, as it is not clear to me how these findings can be used for "intervention".

Response: Thank you for this comment. As clarified in Comment 1 in this document, although the findings are derived from idealized experiments, they provide insights into how the system intrinsically responds to additional water input. The hypothetical additional precipitation applied represents increased water availability, allowing us to identify the timing ($t_{optimal}$) when drought severity (DS) of a multiyear drought (MYD) event is most effectively reduced. By consistently distinguishing MYD attenuation efficiency (system response) and drought alleviation (event-level improvement), the results conceptually inform discussions on real-world MYD alleviation strategies, without implying direct control over rainfall.

16. L30. I was not able to check the reported number of "3000 deaths" in the US, beside the Science paper cited in the text. I suggest checking the source of this information.

Response: Thank you for this comment. We have rechecked this figure and confirmed that the reported number of approximately 3,000 deaths in the United States was obtained from the National Integrated Drought Information System, an official U.S. government platform that coordinates drought monitoring, forecasting, planning, and information at national, state, and local scales. The data are available at: https://www.drought.gov/news/high-cost-drought. The source provides the basis for the reported number.

17. L55. "three key dimensions". Please reword, as the term "dimension" is misused here.

Response: Thank you for pointing this out. The "term" is replaced with "characteristics" to more accurately reflect that drought duration (DD), drought intensity (DI), and drought severity (DS) describe distinct features of a drought event rather than mathematical or spatial dimensions.

18. L55. "the lowest". This is true only for indicator where droughts are negative. Please clarify (largest in absolute value).

Response: Thank you for this comment. In our study, we focus only on drought events. During these events, the standardized drought indices take negative values, and greater severity corresponds to larger departures of the index below the critical threshold. Therefore, drought intensity (DI) is defined as the maximum deviation of the drought index below the critical threshold during a drought event. This revised wording provides a clearer and more general interpretation of DI.

19. L57. These are not indicators but rather drought features.

Response: Thank you for this comment. We agree that these are features describing drought event characteristics rather than drought indicators.

20. L60. "the four stages". These stages have not been introduced yet.

Response: Thank you for this comment. We agree that the four stages had not yet been explicitly stated at this point in the manuscript. To improve clarity, we now briefly introduce the definitions here when they are first mentioned, specifying that the four stages refer to the onset, development, recovery, and termination of drought events.

21. L68-70. This sentence is misleading. A large numerical shift in anomalies (e.g., from 4 to 2) does not necessarily means a large change in drought conditions, as conditions still remains extreme. This is due to the non-linearity in the drought definition according to the indices, but not to a non-linear response of drought to precipitation. I think that this needs to be better clarified, as it has huge implications on your analyses.

Response: Thank you for this valuable comment. The original sentence may have caused a misleading interpretation. Our intention was not to suggest that a numerical change in a standardized drought index (e.g., from 4 to 2) necessarily represents a proportional improvement in actual drought events. Our original sentence was intended to describe the statistical behavior of standardized indices.

Standardized drought indices follow a normal distribution, so values in the tail region are numerically more sensitive to small shifts in cumulative probability. Consequently, when the standardized index is far below its drought threshold, even a small hydrological improvement can lead to a relatively large numerical change in the index. This sensitivity arises from the statistical construction of the standardized index. This is the specific meaning we aimed to convey in the original sentence.

This characteristic was already considered in our framework, which is why we focus on drought severity (DS) derived through run theory to capture the accumulated PDSI values over a drought event rather than a single anomaly value. Here we clarify this point and emphasize that under the same perturbation, PDSI values far below the drought threshold exhibit relatively large numerical changes compared with values closer to the threshold, reflecting the probabilistic structure of standardized indices. At the same time, DS provides a meaningful evaluation of event-level drought response. These two considerations together motivate the design of our numerical experiments.

22. L71. This example is incorrect, as a 1-unit change does not have the same physical and probabilistic meaning at, e.g., -1 or -4. Maybe I misunderstand your statement, and we are saying the same think, but changes in anomalies cannot be treated linearly but they need to account for the non-linearity.

Response: Thank you for this comment. We agree that a 1-unit change in a standardized drought index does not have the same physical and probabilistic meaning across the range of values, and that the non-linearity of the index must be considered. This is exactly the conceptual basis of our study.

In the manuscript, the example using PDSI values of -4 and -1 (Fig. 1c and 1d) was intended to illustrate the statistical properties of standardized indices. Specifically, for PDSI values that are far below the drought threshold (PDSI = -4), a small shift in cumulative probability results in a larger numerical change in index compared with values closer to the threshold (PDSI = -1) under the same perturbation. This demonstrates the non-linear probabilistic spacing of the index, and it is this standardized feature that underlies our approach to analyzing drought attenuation efficiency.

We recognize that the arrangement of Fig. 1 may have led to some confusion. Fig. 1a and 1b describe event-level characteristics. During the early months of a drought event, the PDSI deviation is relatively small and the accumulated deficit is low (Fig. 1a). As the event progresses toward the drought intensity (DI) point, defined as the maximum deviation of the index below the critical threshold during the event, the PDSI deviation increases and the accumulated deficit becomes larger (Fig. 1b). Fig. 1c and 1d focus solely on the statistical behavior of the standardized index and the relative sensitivity of PDSI values under the same perturbation, independent of event-level cumulative deficits. The separation of these panels is intentional because it distinguishes between the standardized-index response (Fig. 1c and 1d) and the event-level progression (Fig. 1a and 1b).

The potential misinterpretation that Fig. 1c and 1d reflect cumulative event-level drought severity is understandable given the layout. Our key point is that early-stage PDSI deviations are smaller and correspond to less accumulated deficit, while later-stage, more extreme deviations are more sensitive to identical perturbations but occur when substantial deficit has already been accumulated. This distinction explains why, in Fig. 4, Conceptual Model 1 and 2 are separated to illustrate the opposite alleviation advantages. The standardized index is more responsive at extreme deviations, whereas the event-level deficit has already progressed significantly.

In summary, the numerical example and Fig. 1c-d are intended to illustrate the statistical, non-linear behavior of the standardized index, which is distinct from the event-level cumulative drought severity discussed in Fig. 1a-b.

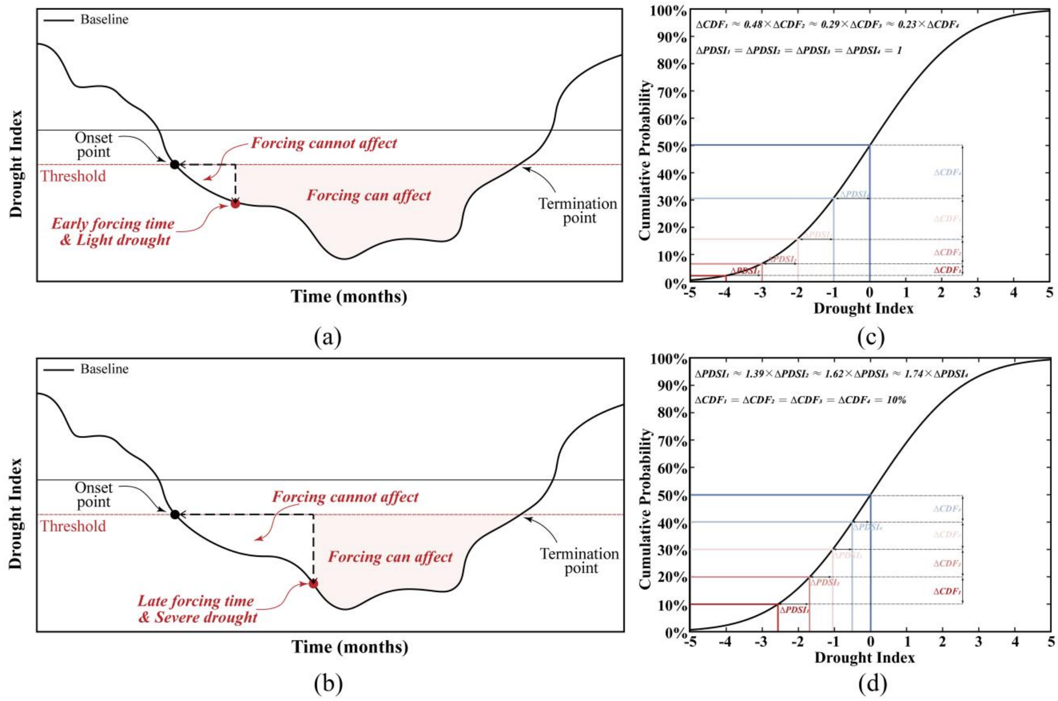

Figure 1: Conceptual and statistical illustrations related to drought mitigation and the normal distribution characteristics of commonly used drought indices (with PDSI as an example). Here, "forcing" refers to a hypothetical incremental water input applied at a given month within a drought event for conceptual illustration. (a) During the early stage of a drought event, PDSI deviation is relatively small and the accumulated deficit is low. (b) As the drought progresses toward the drought intensity (DI) point, which is defined as the maximum deviation of the index below the critical threshold during the event, the PDSI deviation increases and the accumulated deficit becomes larger. (c) Change in cumulative probability induced by a one-unit shift in PDSI under varying drought intensity. (d) PDSI shift required to achieve an equivalent change in cumulative probability under different drought conditions.

23. Fig. 1. The term forcing is used without a proper definition.

Response: Thank you for this comment. In our study, "forcing" specifically refers to a hypothetical incremental water input applied at a given month within a drought event for conceptual illustration. We will ensure that this definition is clearly stated in our manuscript and explicitly indicated in the caption of Fig. 1 so that readers can easily understand its meaning.

24. L85. This concept is here introduced without a proper definition.

Response: Thank you for pointing out this issue. Here we clarify that $t_{optimal}$ refers to the month during a multiyear drought event when a given hypothetical water input achieves the greatest alleviation effect, that is the month corresponding to the maximum value of k under the DMQM framework. This concept is further elaborated through the notations t_a, t_b, and t_c, which represent the optimal timing under Conceptual Model 1 (Fig. 4b), Conceptual Model 2 (Fig. 4c), and the Optimal Solution Model (Fig. 4d), respectively. These help to distinguish different mechanisms influencing drought alleviation (Section 3.4). Among these, t_c represents the final $t_{optimal}$ used throughout the study.

25. L90. Fig. 3 is cited before fig. 2.

Response: Thank you for the comment. We have adjusted the figure citations so that all figures are cited in numerical order.

26. L104. Remove and.

Response: Thank you for this comment. We have removed “and” from the sentence as suggested.

27. L105. Some additional details on the selection procedure are needed.

Response: Thank you for this comment. Daily records from 1960-2020 were examined for each station, and stations with more than 5% missing values were excluded. For the remaining stations, short gaps (≤ 5 days) were filled by linear interpolation, while longer gaps were replaced using the multi-year mean for the corresponding day. This procedure ensures continuity of the time series while minimizing bias in subsequent analyses.

28. Fig. 2. Add region names instead of only numbers.

Response: Thank you for the suggestion. The region names have been added to Fig.2 as requested.

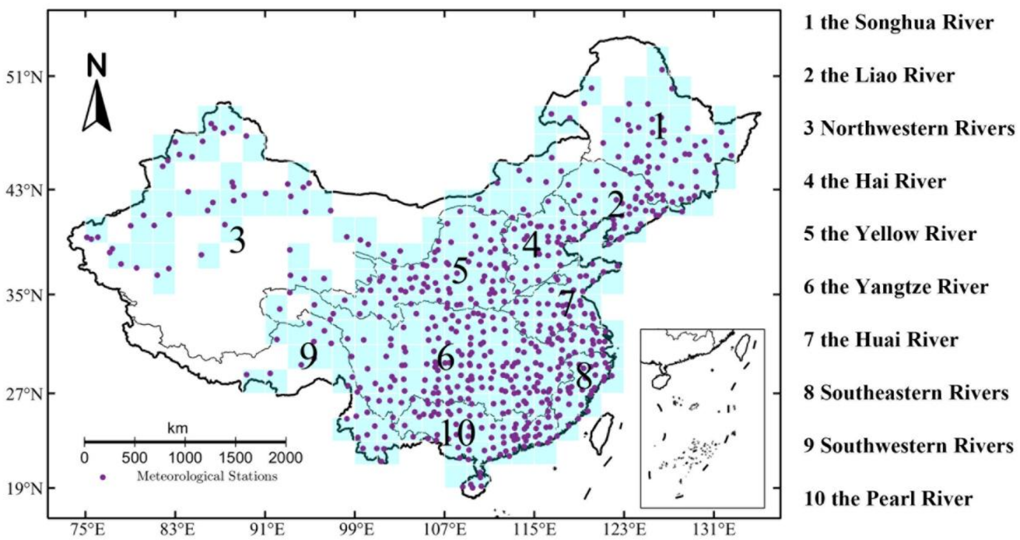

Figure 2: The study area covers 10 large river basins (regions). No. 1, the Songhua River; No. 2, the Liao River; No. 3, Northwestern Rivers; No. 4, the Hai River; No. 5, the Yellow River; No. 6, the Yangtze River; No. 7, the Huai River; No. 8, Southeastern Rivers; No. 9, Southwestern Rivers and No. 10, Pearl River. A total of 199 ($2° \times 2°$) grid boxes (color: cyan) were identified as containing at least one station with continuous instrumental meteorological records. In total, 620 meteorological stations (purple dots) were used.

29. L121. "representative". How?

Response: Thank you for this comment. The representativeness of MYDs in this study is ensured through both duration and intensity criteria. Previous studies have applied different duration thresholds to define MYDs, ranging from at least 13 months (Brunner and Tallaksen, 2019; Mourik et al., 2025) to at least two years (Chen et al., 2025). Considering the objectives of our study and the need to retain sufficient samples for statistical analysis, we adopted the widely used at least 13 months criterion and supplemented it with an intensity requirement.

Specifically, MYDs were identified at the $2° \times 2°$ grid level as events where PDSI remained below -1 for at least 13 consecutive months. To further ensure representativeness and consistency with established MYD characteristics, only events with $DD > 12$ and $DI < -3$ were retained. This combination of duration and intensity criteria guarantees that the selected MYDs reflect significant, prolonged drought events rather than short or marginal deficits, thereby capturing events that are physically meaningful.

References

Brunner, M.I. and Tallaksen, L.M.: Proneness of European catchments to multiyear streamflow droughts, Water Resour. Res., 55, 8881-8894, doi: https://doi.org/10.1029/2019WR025903, 2019.

van Mourik, J., Ruijsch, D., van der Wiel, K., Hazeleger, W., and Wanders, N.: Regional drivers and characteristics of multi-year droughts, Weather Clim. Extremes, 48, 100748, doi: https://doi.org/10.1016/j.wace.2025.100748, 2025.

Chen, L., Brun, P., Buri, P., Fatichi, S., Gessler, A., McCarthy, M.J., Pellicciotti, F., Stocker, B. and Karger, D.N.: Global increase in the occurrence and impact of multiyear droughts, Science, 387, 278-284, doi: 10.1126/science.ado4245, 2025.

30. L122. PIP is not well defined. Is it the value of the PDSI or the time?

Response: Thank you for the comment. We clarify that PIP refers to the drought intensity (DI) point of a drought event, whcih refers to the PDSI value corresponding to the DI. DI is defined as the maximum deviation of the PDSI below the critical threshold during the event. The month at which this PIP occurs is denoted as t_{PIP}.

31. L125-128. This paragraph is confusing, as it relies on several concepts not introduced yet.

Response: Thank you for this helpful comment. We agree that placing this overview paragraph at the beginning of the Methods section could be confusing, because it referenced several concepts (such as *k*,

k_{max}, and $t_{optimal}$) before they had been formally introduced. This ordering may make it difficult for readers to follow the logic on a first reading.

To improve clarity, we now present this overview after all relevant concepts have been introduced, so that it functions as a concise summary of the methodological framework rather than an early preview of undefined terms. In addition, a consolidated table listing principal symbols, acronyms and their description will be provided so that readers can easily consult definitions at any point (Table S1).

32. Fig. 3. Many terms are not explained here (pr_mean, is it the long term of the mean of each month?), k_theoretical and k_actual.

Response: Thank you for this helpful comment. Figure 3 was originally placed early in the Methods section, which could be confusing because it referenced several concepts (such as $\overline{pr}$, $k_{theoretical}$, and k_{actual}) before they were formally introduced. To improve clarity, we now present Figure 3 after all relevant concepts have been introduced, so that it functions as a concise summary of the methodological framework rather than a preview of undefined terms. In addition, a consolidated table listing principal symbols, acronyms and their description will be provided so that readers can easily consult definitions at any point (Table S1).

$\overline{pr}$ is the long-term mean monthly precipitation. The amount of hypothetical precipitation added in the numerical experiments is calculated as $pr_{rate} \times \overline{pr}$, ensuring comparability across grids and months with different climatological precipitation levels.

$k_{theoretical}$ is obtained by fitting the DMQM analytical solution $DS_{rate} = 1 - exp(-k \cdot pr_{rate})$ to the full series of data-based experimental outcomes across all perturbation steps. It represents the theoretical efficiency implied by the DMQM framework, capturing the nonlinear, bounded response of DS to hypothetical additional precipitation. $k_{theoretical}$ also integrating the combined effects of both Conceptual Model 1 and 2 (as clarified in our response to Comment 4 at the beginning of this document).

k_{actual} is directly calculated from the numerical experiments. It represents the empirical (data-based) sensitivity of DS to small precipitation perturbations, approximated from the initial slop of the DS_{rate} and pr_{rate} relationship. Specifically, because the model expansion for small pr_{rate} gives $DS_{rate} \approx k \cdot pr_{rate}$ for $pr_{rate} \ll 1$, we approximate $k_{actual} = \frac{DS_{rate}(pr_{rate}=0.05)}{0.05}$, where 0.05 is the smallest perturbation step used in the experiments. This provides a direct, local estimate of the initial slope.

Our main analysis focuses on $k_{theoretical}$ because it characterizes the system-level response of DS to hypothetical additional precipitation and suitable for spatial and temporal comparison. k_{actual} serves as an independent reference to evaluate the internal consistency and numerical credibility of the DMQM formulation, but it is not used as the primary analysis indicator ($\varepsilon = \left(\frac{k_{theoretical} - k_{actual}}{k_{actual}}\right) \times 100\%$).

33. L132. PDSI requires calilbration. How was this performed? Add details.

Response: Thank you for this comment. The PDSI used in this study was calculated using the Climatically Appropriate For Existing Conditions (CAFEC) formulation (Jacobi et al., 2013), which includes a full calibration process to ensure regional applicability. The calibration was performed over the baseline period 1961-1990. For each grid, monthly climatological values of the water balance components (including precipitation, potential evapotranspiration, recharge, runoff and losses) were first derived over the calibration period. From these, the CAFEC precipitation and the monthly calibration coefficients were computed following the algorithm. Departures between observed precipitation and CAFEC were then multiplied by the calibrated weighting factor to obtain the Z-index and PDSI time series.

References

Jacobi, J., Perrone, D., Duncan, L.L., and Hornberger, G.: A tool for calculating the Palmer drought indices, Water Resour. Res., 49, 6086-6089, doi: https://doi.org/10.1002/wrcr.20342, 2013.

34. L135. Was the FAO-56 applied directly on monthly data? Clarify how, as daily data are previously discussed.

Response: Thank you for this comment. The FAO-56 method was not applied directly to monthly data. Instead, daily potential evapotranspiration (PET) was first calculated using the FAO-56 formulation based on daily meteorological observations. The resulting daily PET values were then aggregated to obtain monthly PET, which was subsequently used as input for the PDSI calculation.

35. L138. How is G modelled?

Response: Thank you for the comment. In this study, because potential evapotranspiration (PET) was calculated at the daily timescale, and the magnitude of the day soil heat flux (G) beneath the grass reference surface is relatively small, it may be ignored and thus $G \approx 0$ (Allen et al., 1998).

References

Allen, R. G., Pereira, L. S., Raes, D., and Smith, M.: FAO Irrigation and Drainage Paper No. 56 – Crop Evapotranspiration, FAO, Rome, Italy, 1998.

36. L142. Why is the period 1961-1990 used for reference? If the analysis focuses on 1991-2020, why wasn't this period (recommended by WMO) used instead.

Response: Thank you for this thoughtful comment. Consistent with our responses to Comments L15 and L17, in our study, 1991-2020 serves as the simulation period for applying hypothetical additional precipitation in the numerical experiments. Using the same period as the baseline would alter the reference climatology and the derived hydrological parameters, potentially confounding the experimental results. To maintain an independent and stable reference, 1961-1990 is adopted as the calibration period, while all numerical experiments are applied only to 1991-2020. This approach ensures that the effects of hypothetical additional precipitation are evaluated consistently against a fixed climatological reference.

37. L148. What is a "typical" MYD? Please clarify.

Response: Thank you for the comment. In this study, a "typical" MYD is defined as an event that captures the key characteristics of persistent and severe droughts, ensuring both temporal continuity and intensity. Following the approach described in our response to L121, at each 2° × 2° grid, drought events are first identified where PDSI remains below -1 for at least 13 consecutive months. To ensure representativeness and focus on significant events, only those with $DD > 12$ and $DI < -3$ are retained.

These criteria guarantee that the selected MYDs are temporally continuous, sufficiently severe, and physically meaningful. By combining both duration and intensity thresholds, the definition of "typical" MYDs aligns with previous studies (Brunner and Tallaksen, 2019; Mourik et al., 2025; Chen et al., 2025) and ensures that the events included in the analysis reflect the most consequential and characteristic prolonged droughts in the study region.

References

Brunner, M.I. and Tallaksen, L.M.: Proneness of European catchments to multiyear streamflow droughts, Water Resour. Res., 55, 8881-8894, doi: https://doi.org/10.1029/2019WR025903, 2019.

van Mourik, J., Ruijsch, D., van der Wiel, K., Hazeleger, W., and Wanders, N.: Regional drivers and characteristics of multi-year droughts, Weather Clim. Extremes, 48, 100748, doi: https://doi.org/10.1016/j.wace.2025.100748, 2025.

Chen, L., Brun, P., Buri, P., Fatichi, S., Gessler, A., McCarthy, M.J., Pellicciotti, F., Stocker, B. and Karger, D.N.: Global increase in the occurrence and impact of multiyear droughts, Science, 387, 278-284, doi: 10.1126/science.ado4245, 2025.

38. L149. Did you perform any pooling? A single month of interruption was enough to stop a MYD? Clarify

Response: Thank you for this comment. In this study, identified MYDs are strictly continuous, meaning that the PDSI remains below -1 for the entire duration without interruption. A single month with PDSI above this threshold would terminate the event. No pooling across non-continuous periods was performed. This strict continuity criterion ensures that all selected events represent uninterrupted multiyear droughts with consistent intensity (DI = -3), allowing us to accurately assess the timing of hypothetical precipitation increases on drought alleviation. By focusing on continuous events, we avoid conflating separate drought stages (including onset, development, recovery, and termination) and ensure that the analysis captures the full temporal structure of prolonged droughts, which is essential for evaluating system responses and attenuation efficiency.

39. L151-152. These are not events, but drought periods for each given grid cells. Speaking about "at least 11" multiyear drought events in each basin over 30 years is misleading.

Response: Thank you for the comment. As clarified in our response to Comment L16, our analysis is conducted at the grid scale. For each of the 199 grids during 1991-2020, multiyear drought (MYD) occurrences are identified independently. Across all grids, a total of 351 grid-level MYD instances were identified across all grids. Some of these may belong to the same regional MYD event, but our study focuses on the grid-level response rather than regional aggregation. This grid-scale approach is appropriate because it captures spatial heterogeneity in drought severity and timing, which is essential for addressing our research questions. A clarification has been added to avoid possible confusion regarding the interpretation of the MYD count.

When summarizing basin-level characteristics, the aggregation is performed across grid-level MYD samples within each basin, which provides a statistically meaningful representation of basin-scale drought attenuation efficiency (k). In this framework, the statement that each basin contained at least 11 MYD samples refers to the number of valid grid-level MYD periods rather than independent basin-wide drought events. This clarification aligns with the explanation provided in our responses to the comments regarding L16 and Fig. 8.

40. L164. "greater negative values". Even if this is true, the relationship is not linear, and the same DS change may mean very different things depending on the conditions. As an example, if you use the corresponding probability rather than the standardize anomalies, the results will be very different. The "mitigation" effect should be independent from the way the metric is presented.

Response: Thank you for this comment. We clarify that in our study, drought severity (DS) represents the sum of monthly PDSI values over the duration of a drought event (DD), capturing both the severity and persistent of drought rather than being a single PDSI value. Therefore, all analyses of drought alleviation efficiency are conducted at the event level, which represents the integrated impact of hypothetical additional precipitation across the drought period.

In our framework, the standardized nature of drought indices is fully considered. As discussed in our response to L71, PDSI values that deviate further below the drought threshold exhibit relatively larger numerical changes under the same perturbation compared with values closer to the threshold. This characteristic affects how a single month of hypothetical additional precipitation contributes to reductions in DS over the event, as illustrated by Conceptual Model 2. In other words, the non-linearity and tail sensitivity of the standardized index influence the conversion efficiency of added precipitation into DS reduction, linking the single-month PDSI response to the event-level DS response.

Accordingly, the original sentence contained a misleading description. The intended meaning is that PDSI values further below the drought threshold exhibit larger numerical changes under the same perturbation, as explained above.

41. L166. I do not follow this logic. Maybe it is still a consequence of the misunderstanding on the meaning of a unitary change in DS. Please reword and clarify, ad this is a key assumption of your methodology.

Response: Thank you for this comment. Before clarifying the logic of Conceptual Model 2, we note that the drought severity (DS) used in our framework is an event-level cumulative quantity, not the instantaneous value of PDSI. DS is calculated as the integral of the PDSI series over the drought duration using run theory. Therefore, the unequal spacing and probabilistic non-linearity of PDSI are inherently preserved in DS. Then the logic under Conceptual Model 2 is rooted in the statistical sensitivity of standardized drought indices such as PDSI. PDSI values that fall far below the drought threshold respond more strongly to the same hypothetical precipitation perturbation. As illustrated in Fig. 1, for example, a shift in cumulative probability produces a much larger numerical change when the PDSI is -4 compared with -1 under identical perturbation.

Within a multiyear drought (MYD) event, PDSI values progressively deviate further below the threshold when approaching the drought intensity point (PIP), which corresponds to the maximum deviation of the PDSI below the critical threshold during the event. Additional hypothetical water input applied at the month at PIP (t_{PIP}) is expected to produce the greatest reduction in the event-level drought severity (DS) of MYD event under Conceptual Model 2. Consequently, the optimal timing predicted by Conceptual Model 2 (t_b) tends to coincide with t_{PIP}.

42. Fig. 4a is the same as in Fig. 3. Avoid repeating the same plot multiple times.

Response: Thank you for this comment. We would like to clarify that in the original structure of the manuscript, Fig. 3 and Fig. 4a served different purposes. Fig. 3 presents a workflow-style overview of the methodological framework, with its subpanels designed to enhance the readability of the framework. In contrast, Fig. 4a illustrated drought event under baseline pattern before applying hypothetical additional precipitation. Because Fig. 4a supports the conceptual derivation of the DMQM, it is not a duplicate of Fig. 3 but plays a distinct explanatory role.

However, after considering this comment together with other suggestions in this document, we agree that the manuscript can be made clearer and more concise by adjusting the figure structure. Therefore, we will move Fig. 3 to the end of the Methods section and simplify it by removing elements that overlap with Fig. 4. This adjustment ensures that each figure has a clearly defined and non-redundant function.

43. L185. Why the absence of overlap between the baseline period and the experiment period is needed?

Response: Thank you for the comment. Consistent with our responses to Comments L15, L17, L132, and L142, the numerical experiments add hypothetical additional precipitation only after the start of the simulation period. If the baseline period overlapped with the experiment period, the added hypothetical precipitation perturbations would also influence the calculation of the reference climate and related hydrological parameters. Having no overlap ensures that the baseline remains fixed and unaffected, so the response to hypothetical additional precipitation can be evaluated relative to a stable and independent reference.

44. L194. Clarify the acronym. Also, why a theoretical model is needed? As seen later, this introduced some unneeded approximations. Cannot be the same analyses made on the empirically computed values Delta_DS and Delta_pr?

Response: Thank you for this comment. DMQM stands for drought mitigation quantitative model. As clarified in our response to Comment L17-18, in our framework, k quantifies the system's drought attenuation efficiency, namely how rapidly hypothetical additional precipitation is converted into reductions in drought severity (DS) of multiyear drought (MYD) events. We constructed DMQM to provide a consistent, system-level measure of DS_{rate}-pr_{rate} sensitivity, integrating the effects of Conceptual Models 1 and 2 (as clarified in our response to Comment 4 at the beginning of this document) and capturing the nonlinear pattern. While empirical ΔDS and Δpr values can be used directly, they are more variable and less comparable across grids and months.

45. Eq. 3. As stated before, a change in DS can have different meaning depending on the severity itself. How did you account for that?

Response: Thank you for this comment. We clarify that that Eq. 3 represents the differential form of the drought alleviation process at the event level, using drought severity (DS) calculated as the cumulative sum of PDSI values over the duration of a multiyear drought. It is not derived from Conceptual Model 2 alone. In our numerical experiments, DS exhibited a consistent nonlinear diminishing-return response to hypothetical added precipitation (Fig. 5). This empirically observed pattern resembles first-order kinetics, where the instantaneous rate of DS reduction is proportional to the remaining DS. Eq. 3 captures this behavior and links the instantaneous rate of DS reduction to the remaining DS. Our testing confirms that this formulation provides the best fit to the observed response pattern (Table S2). The formulation therefore integrates the mechanisms represents in both conceptual models (Fig. 4b and 4c). Within this unified system representation, the parameter k characterizes the efficiency with which hypothetical additional precipitation reduces DS, and comparing k across different intervention timings provides the basis for identifying the optimal timing of drought alleviation.

46. L209. The terms "actual" and "theoretical" are poorly explained. Is the former based on the empirical values and the latter on the model? If you have K_actual, why you should base your analysis on the theoretical?

Response: Thank you for this important comment. As clarified in our response to Comment L17-18, L194, and Fig. 3, k quantifies the system's drought attenuation efficiency, namely how rapidly hypothetical additional precipitation is converted into reductions in drought severity (DS) of multiyear drought (MYD) events. Formally, the differential form is $\frac{\partial DS_{rate}}{\partial pr_{rate}}=k\cdot(1-DS_{rate})$ (eq. 3), k also represents the initial response rate when $DS_{rate}=0$.

In this study, k_{actual} is directly calculated from the numerical experiments. It represents the empirical (data-based) sensitivity of DS to small precipitation perturbations, approximated from the initial slop of the DS_{rate} and pr_{rate} relationship. Specifically, because the model expansion for small pr_{rate} gives $DS_{rate} \approx k \cdot pr_{rate}$ for $pr_{rate} \ll 1$, we approximate $k_{actual} = \frac{DS_{rate}(pr_{rate}=0.05)}{0.05}$, where 0.05 is the smallest perturbation step used in the experiments. This provides a direct, local estimate of the initial slope. In contrast, $k_{theoretical}$ is obtained by fitting the DMQM analytical solution $DS_{rate} = 1 - exp(-k \cdot pr_{rate})$ to the full series of data-based experimental outcomes across all perturbation steps. It represents the theoretical efficiency implied by the DMQM framework, capturing the nonlinear, bounded response of DS to hypothetical additional precipitation. $k_{theoretical}$ also integrating the combined effects of both Conceptual Model 1 and 2 (as clarified in our response to Comment 4 at the beginning of this document).

Our main analysis focuses on $k_{theoretical}$ because it characterizes the system-level response of DS to hypothetical additional precipitation and suitable for spatial and temporal comparison. k_{actual} serves as an independent reference to evaluate the internal consistency and numerical credibility of the DMQM formulation, but it is not used as the primary analysis indicator.

47. Fig. 5. It is not clear to me how these plots were derived. Are these from eq. 6. Using which k? This section of the results and the previous methodology need to be restructured.

Response: Thank you for this comment. In Fig. 5, the DS_{rate} and pr_{rate} response curves are derived from the numerical experiments aggregated over all grids in the study region. The horizontal axis represents pr_{rate} (a dimensionless proportion that controls the relative magnitude of the hypothetical precipitation addition, the amount of hypothetical precipitation addition is given by $pr_{rate} \times \overline{pr}$, $\overline{pr}$ is the long-term mean monthly precipitation), and the vertical axis represents DS_{rate} (the corresponding relative change in DS). For each month of a multiyear drought event during the drought development period (Month 1 to 8), we used the experimental outcomes to fit the DMQM analytical form: $DS_{rate} = 1 - exp(-k \cdot pr_{rate})$, which yields a single $k_{theoretical}$ for the entire study region for that month. This parameter represents the system-level drought attenuation efficiency and captures the nonlinear dependence implied by the differential formulation of DMQM, $\frac{\partial DS_{rate}}{\partial pr_{rate}} = k \cdot (1 - DS_{rate})$. The resulting curves therefore represents the nonlinear, diminishing-return behavior observed in the experiments: strong attenuation at low DS_{rate} followed by progressively weaker gains as $1 - DS_{rate}$ weakens ($DS_{rate} = \frac{DS_{baseline} - DS_{forcing}}{DS_{baseline}}$, $DS_{baseline}$ is the actual DS under the baseline scenario that no hypothetical additional precipitation is applied, and $DS_{forcing}$ is the DS after the hypothetical additional precipitation is applied).

For comparison, a linear reference line is plotted for Month 1. Its slope is computed from the marginal response at very small hypothetical precipitation perturbation ($pr_{rate} = 0.05$, the step size in the numerical experiments): $k_{actual} = \frac{DS_{rate}(pr_{rate}=0.05)}{0.05}$. This linear reference represents the hypothetical scenario where attenuation efficiency remains constant with additional water. The clear divergence between this linear reference and the fitted nonlinear DMQM curves represents the intrinsic non-linearity of the drought attenuation process, which cannot be captured by a constant-slope formualtion.

48. Fig. 6. Errors in the theoretical seem larger in the North-east, which is the same region where with several Kmax at month >= 2. Is there an effect of the error in the theoretical model in these results?

Response: Thank you for the comment. We would first like to clarify that in Fig. 6a, the relative errors of $k_{theoretical}$ are not larger in the North-east. Instead, they are generally smaller than in many parts of the study area. This indicates that the DMQM model performs well in this region. Therefore, the results in these North-east region are more reliable.

To further avoid misunderstanding, we provide additional clarification regarding the spatial distribution of relative error. In northern China more broadly, especially in the northwestern arid region, relatively larger errors do appear. This pattern can be explained by the physical environment of these areas. In arid regions, drought variability is strongly influenced by factors other than precipitation, most notably the high evaporative demand. Because precipitation amounts are very low and potential evapotranspiration (PET) is large, changes in precipitation alone cannot fully capture the main drivers of drought development. Our numerical experiment perturbs only hypothetical precipitation while keeping PET fixed. Consequently, in arid regions, the dominance of non-precipitation factors can be transmitted into the estimation of drought attenuation efficiency (k). It is important to note that evaporative demand is not directly controllable in practical drought management, whereas water availability is. This is precisely the motivation for focusing on hypothetical water-input perturbations. The slightly higher relative errors in arid regions are therefore physically reasonable and consistent with expectations. In the humid southern region, the larger relative errors also occur in some grids. The underlying cause of these localized large relative errors is different from that in arid regions. As clarified in our response regarding Fig. 3 and L209, k_{actual} represents the initial slope. In humid regions, baseline precipitation is high. Increasing hypothetical precipitation by the minimum experiment step of 0.05 therefore corresponds to a relatively large increment of hypothetical water input. This leads to a greater deviation from the initial slope and produces relative larger differences between $k_{theoretical}$ and k_{actual}. Even in these arid and humid grids, however, the relative errors remain modest (mostly around 6%), suggesting that the DMQM still performs reasonably well across the study area. In semi-arid regions (North-eastern region of our study area), precipitation is neither extremely low nor excessively high, and it remains the dominant factor governing drought development (Feng et al., 2025). The 0.05 increment represents a relatively small amount of hypothetical water input compared with the local baseline precipitation, and the agreement between theoretical and actual results is therefore the strongest in these regions.

Subsequent comments related to this issue, including those concerning Fig. 7, and L355-356, are addressed accordingly in their respective responses.

References

Feng, Y., Sun, F. and Deng, X.: Attributing the divergent changes of drought from humid to dry regions across China, J. Hydrol., 660, 133363, doi: https://doi.org/10.1016/j.jhydrol.2025.133363, 2025

49. Fig. 7. Same as before. Most of the large differences seem to be in the north-east.

Response: Thank you for the comment. As clarified in our response to Comment Fig. 6, the North-east region does not exhibit the large differences. Instead, the relative errors between $k_{theoretical}$ and k_{actual} in this region are generally smaller than in many other parts of the study area. This indicates that the DMQM performs well in the North-east, and the corresponding results are therefore reliable. The detailed explanation provided in our response to Fig. 6 applies here as well.

50. Fig. 8. Why reporting the map of the duration. Is this the median duration of the droughts (as in Table 1)? How are the probability plots made? Based on how many data? Is there any smoothing applied?

Response: Thank you for the constructive comment. Fig. 8 presents the spatial distribution of the median drought duration (DD) for each grid cell, consistent with Table 1. In our framework, drought severity (DS) is the dependent variable, computed using run theory as the integral of PDSI series over DD. DS therefore represents the accumulated water deficit over the entire event rather than a single PDSI value. Since DS is defined as the cumulative PDSI over DD, and k is derived from DS, DD represents a key factor influencing k in addition to the standardized characteristics of PDSI itself. Therefore, we present the DD map to help interpret differences in k across basins.

To remain consistent with our clarification in response to Comment L16, we emphasize that our analysis is conducted at the grid scale. The probability plots are constructed using k estimates from all 351 grid-level multiyear drought (MYD) samples across China. Here, a "MYD sample" refers to a grid-level MYD period during 1991-2020. For each of the 199 grids, MYD occurrences are identified independently. Across all grids, a total of 351 grid-level MYD samples were identified. Some of these samples may belong to the same regional MYD event, but our analysis focuses on the grid-level response as the statistical unit rather than aggregating events regionally. Each river basin contains at least 11 valid grid-level MYDs, ensuring basic statistical robustness. For each grid, k_1, k_2, and k_3 represent how rapidly hypothetical additional precipitation is converted into reductions in DS at Month 1, 2, and 3. The empirical distributions (i.e., probability distributions derived directly from the observed Δk) of $\Delta k_{21} = k_2 - k_1$ and $\Delta k_{31} = k_3 - k_1$ are then evaluated for each basin and China as a whole. The probability curves are generated using kernel density estimation (ksdensity in MATLAB), which provides a smooth empirical approximation of the underlying probability distribution. No further smoothing, filtering, or curve fitting is applied beyond this standard KDE procedure.

51. Table 1. Add the number of pixels for each basin, and also the DS.

Response: Thank you for the comment. We have added the number of grids for each basin to Table 1.

In our framework, DS is the dependent variable, computed using run theory as the integral of PDSI series over the drought duration (DD). DS therefore represents the accumulated water deficit over the entire event rather than a single PDSI value.

Because Table 1 is intended to summarize basin characteristics relevant to interpreting the variations in Δk (the changes in drought attenuation efficiency k between different months of a MYD event), and DS is the response variable, it is not conceptually appropriate to include DS in this table.

Table 1: Summary of the median values of DD and Δk, including $\Delta k_{21}(=k_2-k_1)$, $\Delta k_{31}(=k_3-k_1)$, and $\Delta k_{32}(=k_3-k_2)$, across the ten river basins and all of China.

River basin	Number of grids	DD (Months)	$P_{50}(\Delta k_{21})$	$P_{50}(\Delta k_{31})$	$P_{50}(\Delta k_{32})$
Songhua River	26	22	-0.091	-0.235	***-0.142***
Liao River	10	28	**0.022**	**0.003**	**-0.024**
Northwestern Rivers	44	17	-0.071	-0.152	-0.087
Hai River	8	29	-0.083	-0.110	-0.053
Yellow River	20	21	-0.069	-0.111	-0.090
Yangtze River	39	19	-0.145	***-0.250***	-0.127
Huai River	11	23	-0.081	-0.141	-0.042
Southeastern Rivers	8	14	-0.024	-0.196	-0.077
Southwestern Rivers	15	21	-0.148	-0.148	-0.083
Pearl River	18	19	***-0.162***	-0.200	-0.063
All of China	199	21	-0.080	-0.166	-0.094

Note: The bold value marks the maximum P_{50} among the ten river basins in China, and the bold-italic value marks the minimum P_{50}.

52. L305. Why not DS? Is it possible that this result is just the effect of DS as mentioned before (not all the delta have the same effect, as they depend on DS itself).

Response: Thank you for the comment. We note that there may be some confusion between DS (the drought severity characteristic of each MYD) and PDSI. In our framework, DS is the dependent variable, computed using run theory as the integral of PDSI series over the drought duration (DD). DS therefore represents the accumulated water deficit over the entire event rather than a single index value.

The analysis in this section focus on Δk (the changes in drought attenuation efficiency k between different months of a MYD event) variation with respect to the drought duration (DD), which is an independent event characteristic potentially influencing sensitivity. Using DS itself as a grouping factor would be inappropriate because DS is the quantity whose response to hypothetical precipitation perturbations we are evaluating. Allowing DS to serve simultaneously as both response and grouping variable would confound the interpretation. By contrast, binning by DD allows us to isolate the effect of DD on Δk, consistent with the event-based framework of our study.

53. L325. Some basins are just one pixel, or close to. How these values can be compared among basins?

Response: Thank you for this comment. We clarify that each river basin in our research contains at least eight girds (Table 1 in this document), with each grid having at least one identified grid-level multiyear drought (MYD) sample. The parameter k_{max}/k_1 is computed independently for each grid-level MYD sample. The basin-level k_{max}/k_1 is then obtained as the median of the grid-level k_{max}/k_1 values within the basin. This approach captures the spatial variability within the basin while providing a representative indicator for the overall basin response. This explanation is consistent with our clarifications in response to Comments L16, L151-152, and Fig. 8.

54. L339. "precipitation-based drought mitigation". What does this mean?

Response: Thank you for the comment. In this place, "precipitation-based drought mitigation" refers to studies focusing on the contribution of precipitation to drought recovery, which generally dominates over other factors.

55. L355-356. How much of this result is due to the error in the theoretical values?

Response: Thank you for the comment. As clarified in our response to Comment Fig. 6, although Figure 6a prensents the relative error at the grid level, a more integrated view from the river basin perspective (Figures 6c-6e) indicates that $k_{theoretical}$ closely match k_{actual} across the ten major river basins. The red 1:1 line and black fitted line illustrate the strong overall agreement, and the orange crosses indicate that individual basins are also well captured. This suggests that the results are more robust when considered at the basin scale, and the observed regional patterns in attenuation efficiency are not driven by error in the theoretical values.

56. L366. This is not true everywhere in the world.

Response: Thank you for the comment. Our original statement implied a global uniformity that is not accurate. The sentence has been revised to clarify that regional differences exist. Our study focuses on China, and no conclusion regarding global uniformity is intended.