Anonymous referee #1

Li et al. present an interesting study that employs deep learning models to predict groundwater levels during freezing and thawing periods, as well as to classify the underlying dynamic drivers. The paper is mostly well-written. Still, considerable issues require significant revision to make the paper clearer. The most important ones are related to the current structure; the results and discussion are in the same chapter, which is recommended for modification. The authors are encouraged to include a separate discussion section to discuss the main groundwater level types and the most significant implications from these different types. Second, some issues should be more precisely defined in the method. Finally, the authors should consider to present the conclusion in a more structured and clear way.

Response: We sincerely appreciate your recognition of our research work and your valuable suggestions. In response to the major concerns you raised, we have made corresponding revisions and improvements. Regarding the issue of merging the "Results" and "Discussion" sections, we have retained the overall logical coherence of the chapters while supplementing the discussion content. We have also clearly distinguished the discussion topics in the form of subsections to enhance the structure and readability of the manuscript. Specifically, we added "4.1 Implications of Groundwater Level Dynamics Classification for Water Resources Management" and "4.2 A New Perspective on Identifying Groundwater Level Dynamic Mechanisms" in the discussion section. For the inaccuracies in the description of the methodology, we have reorganized and precisely restated the relevant content. Finally, we have revised the "Conclusions" section to make it more concise and well-structured.

General comments:

1. Abstract:

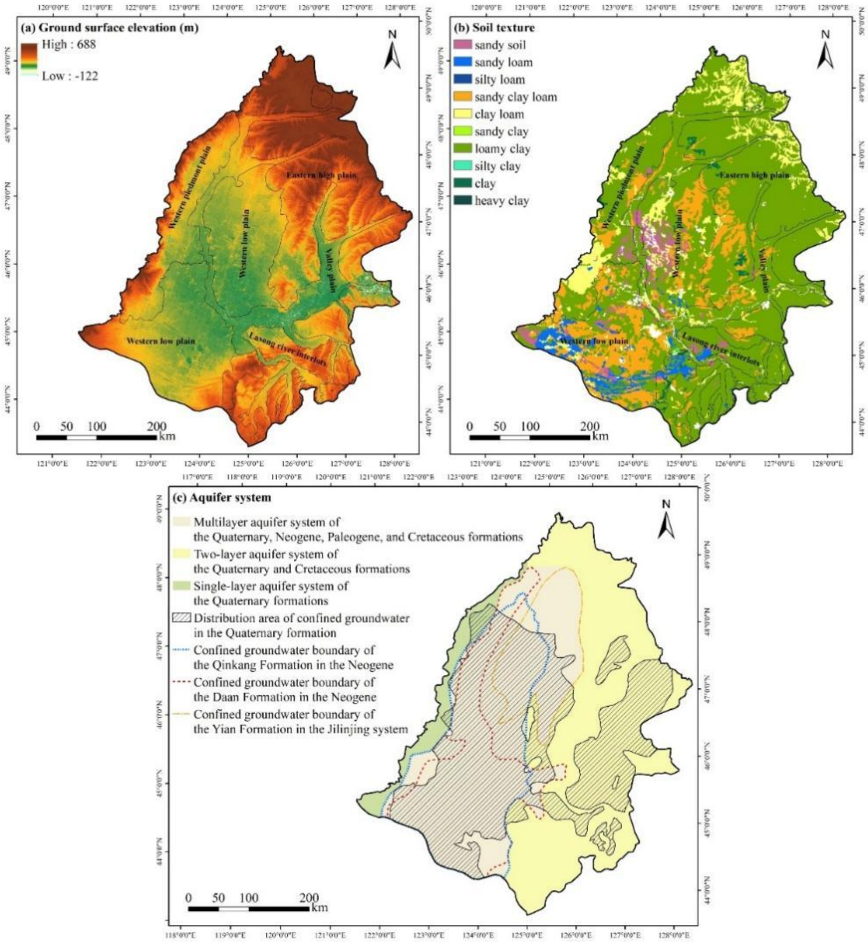
It would be worth rephrasing to make the message clear and better reflect the key findings and the value of this study.

Response: We sincerely thank the reviewer for the valuable comments on the abstract. We have revised the abstract accordingly. The revised content is located from line 14 to line 42.

2. Method:

a) Figure 2 What do the solid circles in Fig. 2(a) represent? Additional description on these labels should be added to the figure caption. Since some similar information is presented in panels (a), (b), and (c), consider merging some of them.

Response: We thank the reviewer for the comment. Regarding the solid circles in Figure 2(a), we have confirmed that this marker was mistakenly added during the drawing process and has been removed in the revised manuscript. In addition, in response to the suggestion concerning the redundancy of certain information across the subplots in Figure 2(a), (b), and (c), we have merged and adjusted the figure contents accordingly. The revised figure is shown below:



Spatial distribution of the ground surface elevation (a), topography (b) and aquifer system (c) in the Songnen Plain, China.

b) Lines 147-149 How do you determine the exact timing of the beginning and end of the freezing period for each well? A precise definition of the freezing period should be provided, similar to the one you gave for the 'Beginning of winter' in Lines 194-198.

Response: Thank you very much for your comments. In the original manuscript, lines 147–149 only provided an overview of the general conditions of soil freezing and thawing in the study area, without giving a precise definition of the start and end times of the freezing period for each monitoring well. According to your suggestion, we have added the start and end times of the freezing period for each monitoring well in the revised manuscript, lines 238–241.

c) Lines 167-169 Please detail the method to estimate the groundwater extraction volume. Given that the groundwater extraction volume is a key component of the proposed mechanism, its estimation accuracy may have an impact on the results. Also, the well depth and screened interval of the monitoring wells might also influence the response rate of the observed groundwater levels, but this aspect does not appear to be addressed in the paper.

Response: Thank you very much for your comments. We have provided additional explanations of the method for estimating groundwater extraction in the revised manuscript, lines 196–214. In this study, groundwater extraction was approximated using the crop water deficit method. This approach is based on crop planting area, crop water requirements, and precipitation data, which allows quantification of groundwater demand during the irrigation period and is therefore theoretically justified. To assess the reliability of the method, we compared the annual total crop water deficit with the annual groundwater supply in local areas. The results show a strong correlation between the two, and the interannual variation of total crop water deficit is consistent with that of groundwater supply, supporting the effectiveness of this method in reflecting groundwater extraction.

We fully acknowledge that the well depth and screened interval of the monitoring wells may affect the groundwater level response. However, due to limited data availability, we were unable to obtain relevant structural parameters for all wells and thus could not conduct a detailed analysis in this study. In future work, we will further explore the impact of these factors when sufficient data become available.

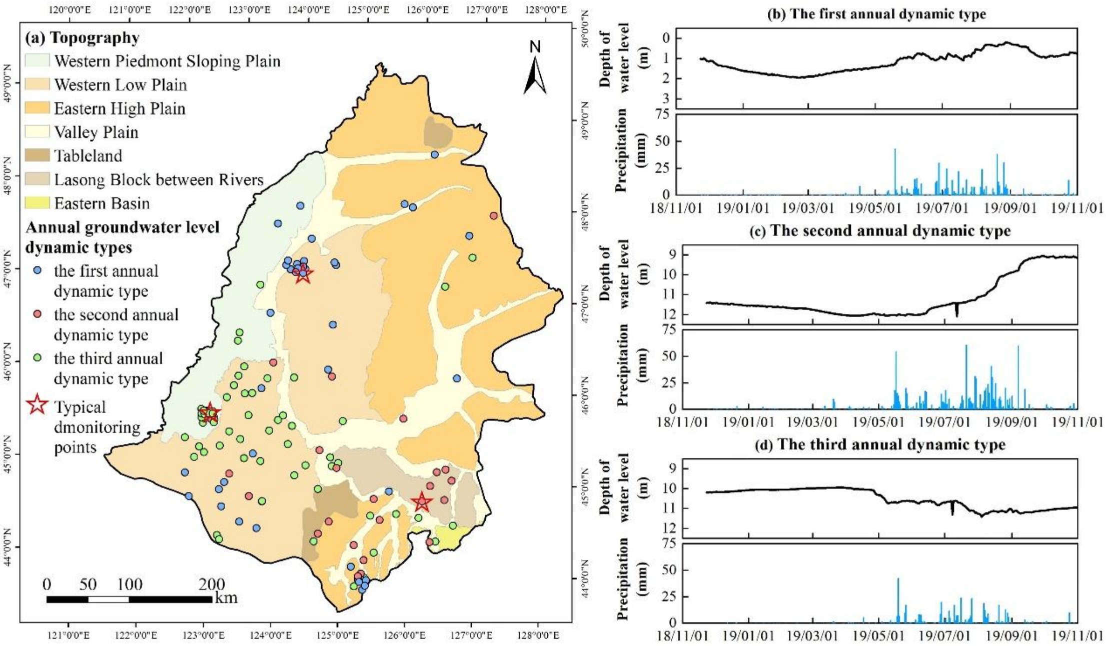
3. Result and discussion:

a) Lines 313-314 This statement is unclear or lacks significance. Could you provide a quantitative indicator to support it?

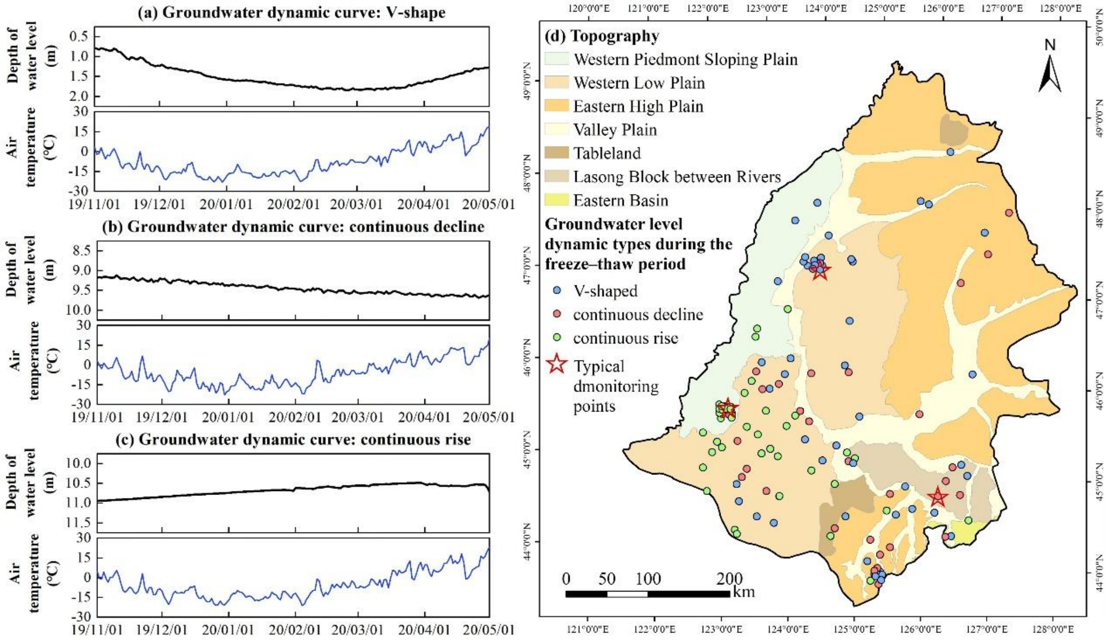
Response: We thank the reviewer for the comment. To address the lack of clarity in lines 313–314 of the original manuscript, we computed the Pearson correlation coefficients at lags of 0–7 days between the simulated and observed water levels for the four representative monitoring wells shown in Figure 4, and added the corresponding quantitative indicators in the revised manuscript, lines 358–363.

b) Lines 359 The authors are strongly recommended to label the three monitoring wells representing the three types of groundwater level dynamics (panels b, c, and d) in Figure 5a. The well numbers mentioned here are not very informative since the locations of the wells are not indicated. There are similar cases later on as well.

Response: We thank the reviewer for the suggestion. We have marked the locations of the three representative monitoring wells in Figure 5a. In addition, the locations of the representative wells have also been added to Figure 6. The revised figures are as follows:



(a) Spatial distribution of different annual groundwater level dynamic types in the Songnen Plain, China; (b–d) Dynamic curves of different annual groundwater types and their corresponding precipitation variations. (b) The first annual dynamic type is represented by an unconfined aquifer monitoring well, numbered 230204210070, located in the western low plain; (c) The second annual dynamic type is represented by an unconfined aquifer monitoring well, numbered 220182210411, located in the Lasong Block between rivers; (d) The third annual dynamic type is represented by an unconfined aquifer monitoring well, numbered 220802210145, located in the western piedmont sloping plain.



(a–c) Dynamic curves of different groundwater types during the freeze–thaw period and corresponding changes in air temperature; (d) Spatial distribution of different groundwater level dynamic types during the freeze–thaw period in the Songnen Plain, China. The dynamic curves of the groundwater level exhibiting patterns of (a) V-shaped, (b) continuous decline, and (c) continuous rise correspond to the unconfined aquifer monitoring wells numbered 230204210070, 220182210411, and 220802210145, respectively.

c) Lines 388-395 I am not sure I fully understand the authors' meaning here. They state that continuous groundwater level decline mostly occurs in areas with deep groundwater level, but actually, the groundwater depth is greater in areas where the groundwater level shows a continuous rise. Moreover, I think some of the mechanism for the "continuous rising" type should be discussed further, that could enhance the implication of this study.

Response: We sincerely thank the reviewer for the constructive suggestion. The original statement that "sustained declines in groundwater levels mostly occur in areas with greater groundwater depths" could indeed be misleading, as some areas with sustained rising trends actually have even deeper groundwater levels. We have revised and rephrased this section The revised content is as follows, located on lines 435-438 of the revised manuscript:

"Monitoring points with the continuous decline in the groundwater level were mainly distributed in areas, such as the eastern high plain and the Lasong Block between rivers, where the groundwater level depth ranged from 4.52 to 11.51 m at the start of the freezing period (Fig. 6d)."

In Section 3.3.2, based on the results of the EG method, we explicitly pointed out that the "continuously rising" groundwater level type is not affected by the freeze–thaw process, and its cause is mainly attributed to the intra-annual groundwater level recovery process. However, its corresponding intra-annual groundwater level dynamic type had not yet been identified. In Section 3.3, we clarified the correspondence between intra-annual and freeze–thaw period groundwater level dynamic types, indicating that the "continuously rising" type corresponds to the extraction type in intra-annual dynamics. Therefore, in the revised manuscript (lines 645-653), we have provided a more detailed explanation of the formation mechanism of the "continuously rising" groundwater level type.

d) Line 427 It is confusing to see the sentence "Precipitation directly recharged the groundwater" here.

Response: We thank the reviewer for pointing out the issue with the statement "Precipitation directly recharged the groundwater." We acknowledge that this expression was logically ambiguous and lacked terminological rigor. We have revised the sentence by linking it more clearly to the preceding one. The revised content is as follows, located on lines 471-474 of the revised manuscript:

" When a pronounced precipitation peak occurred (Figure 9b), the EG score increased significantly (exceeding 0.15), corresponding to a rise in groundwater level (Figure 9e), indicating that precipitation infiltration made a substantial contribution to the groundwater level increase."

e) Some subheadings are a bit too long and very similar, e.g., Sections 3.2, 3.2.1, and 3.2.2, as well as 3.3, 3.3.1, and 3.3.2, I suggest the authors refine them.

Response: We thank the reviewer for the suggestion. We have simplified and refined the subheadings of Sections 3.2 and 3.3 by removing redundant words and highlighting the core content. The revised subheadings are as follows:

3.2. Dynamic Characteristics of Regional Groundwater Level and their Distribution Laws

3.2.1. Annual Dynamics Variations and Spatial Distribution

3.2.2. Freeze–Thaw Period Dynamics Variations and Spatial Distribution

3.3. Main Controlling Factors and Identification of Causes for Various Groundwater Level Dynamic Types

3.3.1. Annual Dynamics: Influencing Factors and Dynamics Mechanisms

3.3.2. Freeze–Thaw Dynamics: Influencing Factors and Dynamics Mechanisms

f) The authors are encouraged to strengthen the discussion by connecting this research to relevant studies and highlighting its potential implications.

Response: Thank you for the valuable comments from the reviewers. Based on the suggestions, we have added a new "Discussion" section in the revised manuscript, focusing on Section 4.1 "Implications of Groundwater Level Dynamics Classification for Water Resources Management" and Section 4.2 "A New Perspective on Identifying Groundwater Level Dynamic Mechanisms." Section 4.1 further validates the accuracy of our study results by citing and comparing relevant existing research and discusses the application value of the study in detail. Section 4.2 mainly introduces the Expected Gradient (EG) method used in this study, explaining its differences and advantages compared to traditional research on groundwater level dynamic mechanisms. The revised content is located from lines 697 to 755.

4. Conclusion:

The conclusion section is considerably longer than necessary and could be more concise.

Response: We sincerely thank the reviewer for the valuable suggestions regarding the conclusion. In response, we have revised the conclusion accordingly. The revised content is located from line 788 to line 820.

Minor comments:

Line 49 There are formatting issues with some references, which also appear throughout the rest of the paper.

Response: We sincerely thank the reviewer for the review. We have conducted a comprehensive check of all references cited in the manuscript and have standardized their formatting in accordance with the journal's guidelines to ensure accuracy and consistency.

Line 137 delete "topography of the"

Response: Thank you for the suggestion. We have removed "topography of the".

I'm not sure if it's due to image resolution, but some of the colors in the figures are difficult to distinguish. For example, in Fig. 2a, the colors of the solid circles are too similar to those used in the base map.

Response: We thank the reviewer for the comment. Regarding the solid circles in Figure 2(a), we have confirmed that this marker was mistakenly added during the drawing process and has been removed in the revised manuscript . In addition to Fig. 2(a), we also noticed a similar issue with insufficient color contrast in Fig. 11(a) of the original manuscript. To improve the readability and visual clarity of the figure, we have adjusted the color of the solid circles in Fig. 11(a) to enhance their contrast against the background map and minimize potential misinterpretation.

Anonymous referee #2

This manuscript applies a machine learning (ML) approach to predict time-varying groundwater levels in seasonally freezing regions of China. The topic is timely and of high importance for groundwater resource management and environmental protection. However, the study overlooks several critical factors that could significantly influence the results and interpretations. By incorporating additional hydrogeological and environmental variables, the model's accuracy could be greatly improved, leading to a more comprehensive understanding of groundwater dynamics.

Response: Thank you for your valuable comments on our study. We have carefully reviewed your suggestions and made corresponding revisions, and we hope these modifications meet your expectations. We agree that key factors such as hydrogeological conditions and environmental variables may significantly influence the model outputs and their interpretation. However, the core focus of this study is on building LSTM models for each monitoring site individually, aiming to simulate the temporal variation of groundwater level at the point scale. Within this framework, spatially fixed attributes such as aquifer properties and topography remain relatively stable over the short term and are unlikely to exert dynamic influence on the time series at a single site. Additionally, factors such as vertical leakage and surface water interactions are difficult to quantify due to limited data availability. In future work, if data conditions permit, we will consider incorporating these variables to enhance the physical interpretability and predictive accuracy of the model.

Specific comments:

Line 25: Please define NSE upon first mention to ensure clarity for readers unfamiliar with the metric.

Response: Thank you for pointing out this issue. We have added the full term “Nash-Sutcliffe Efficiency” when “NSE” first appears in the abstract.

Line 39: Provide more detailed justification of why monitoring groundwater levels is crucial, not only for managing water resources but also for protecting ecological systems. Additionally, consider using the ML-predicted results to present a case study with quantitative analysis to better illustrate the implications.

Response: Thank you for the valuable comments from the reviewer. Following the suggestions, we have comprehensively revised the relevant parts of the manuscript to further elaborate on the importance of groundwater level depth, especially emphasizing its role in water resource management and ecosystem protection. Additionally, we supplemented the citations with the study by Liu et al. (2022), which used machine learning to predict groundwater level depth in the lower Tarim River, providing a quantitative case validation of the practical significance of groundwater level prediction. The revised content is located from line 46 to line 61.

Lines 62–67: The key disadvantage of physical models, compared to ML models, lies in their time-consuming setup, calibration, and validation processes. However, physical models have the advantage of offering more mechanistic insight into underlying hydrological processes, which ML models often lack.

Response: We thank the reviewer for highlighting the insufficient discussion on the comparison between physical models and machine learning models in the current manuscript. In response to your suggestion, we have revised and supplemented the relevant content accordingly. In the updated version, we have clearly stated the advantages of physical models in revealing the physical mechanisms of hydrological processes, while also acknowledging their limitations in regions with complex geological conditions due to high modeling complexity and substantial data requirements. The revised content is located from line 74 to line 84.

Line 118: The model would benefit from incorporating a wider range of influencing factors, such as aquifer properties, topography, hydraulic conditions (e.g., lateral flow, vertical leakage, groundwater storage, surface water interactions), and anthropogenic variables like population density. Spatial heterogeneity in evapotranspiration and precipitation should also be considered to improve model realism.

Response: We sincerely thank the reviewer for the professional and insightful comments. We fully agree that a variety of natural and anthropogenic factors—such as aquifer properties, topography, groundwater dynamics, population density, and evapotranspiration—can exert significant influence on regional groundwater level changes. However, considering the design rationale and actual data availability in this study, we have carefully reflected on and responded to this point from the following two perspectives:

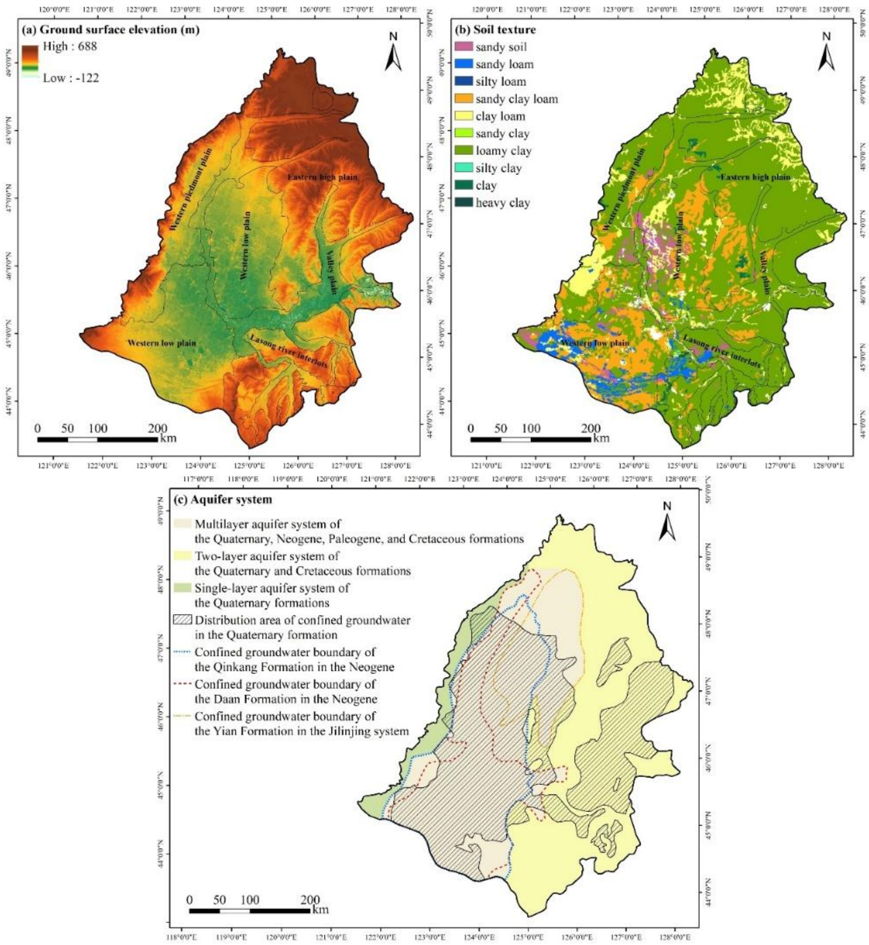
First, the core framework of our study is to independently construct an LSTM model for each monitoring well to simulate the temporal variation of groundwater level at that specific location. The model uses historical meteorological variables and anthropogenic dynamic factors (including air temperature, precipitation, snow depth, and groundwater extraction) as inputs, aiming to capture the nonlinear response relationship between these temporally dynamic factors and groundwater level changes. Under this modeling strategy, spatially static attributes such as aquifer properties and topography remain constant over short periods at a given site and thus cannot provide dynamic explanatory power for the temporal evolution of groundwater levels at that point. Additionally, the spatial heterogeneity of factors such as evapotranspiration and precipitation primarily influences regional-scale patterns or spatial distributions. Since our study focuses on site-specific time series modeling and identification of dominant influencing factors, it is relatively less dependent on spatially heterogeneous variables. We have clarified this limitation in Section 3.5 “Model Limitations” of the revised manuscript.

Second, regarding the absence of variables related to groundwater dynamics (e.g., lateral flow, vertical leakage, and surface–groundwater interactions), we fully acknowledge their critical roles in groundwater system evolution. Although in theory, groundwater flow fields could be constructed through spatial interpolation of observed water levels, in our study the groundwater level is the target output variable of the model. Thus, prior to obtaining the model predictions, it cannot serve as an input driver.

Moreover, in practice, there is a lack of independent observational data (such as hydraulic gradients or recharge–discharge rates) that directly reflect groundwater dynamics, making it currently unfeasible to incorporate these factors into the model. In future work, if data availability improves, we intend to include such variables as key supplementary inputs to enhance the model's physical interpretability.

Figure 2: Consider including a geological map that shows the distribution of geological formations or aquifer types. This would help contextualize the results spatially.

Response: We thank the reviewer for the suggestion. In response, we have added a new subfigure to Figure 2 showing the distribution of the aquifer system. The corresponding description of the aquifer system distribution has been added in the revised manuscript (Lines 168–172). The revised figure is as follows:



Spatial distribution of the ground surface elevation (a), topography (b) and aquifer system (c) in the Songnen Plain, China.

Figure 4: The observed and simulated groundwater levels do not align well; the simulated series appears overly variable. Please explain the possible causes of this discrepancy, such as overfitting, lack of key input variables, or limitations in the model's temporal resolution.

Response: We thank the reviewer for the valuable comments. Although the original manuscript included an explanation for the poor model performance at certain monitoring wells, the reasoning lacked clarity and failed to accurately convey the sources of model error. In response, we have revised and reorganized the relevant paragraph to enhance its logical structure and coherence. The revised content is located from line 341 to line 363.

Lines 373–376 and 557–558: These sections are overly descriptive. Instead of simply stating observations, clarify what the results reveal about the status or trends of water resources. Quantitative insights or implications for water management should be emphasized.

Response: We sincerely thank the reviewer for the valuable comments. Based on the suggestions, we conducted an in-depth analysis of the implications of groundwater level dynamic classification for regional water resources management and supplemented the relevant content in Section 4.1 of the Discussion (lines 697–730). For different groundwater dynamic types, we explicitly proposed differentiated water resources management strategies. In addition, by comparing and linking our findings with previous related studies, we strengthened the scientific basis of the conclusions and demonstrated the consistency between the dynamic types identified in this study and existing empirical research results.

References:

Liu, Q., Gui, D., Zhang, L., et al.: Simulation of regional groundwater levels in arid regions using interpretable machine learning models, Sci. Total Environ., 831, 154902, https://doi.org/10.1016/j.scitotenv.2022.154902, 2022.