Reference Number: egusphere-2025-2734

RESPONSES TO REVIEWER ONE'S COMMENTS

We would like to express our sincere appreciation for your professional and insightful remarks on our paper. The comments are valuable and helpful for us to improve the quality of the manuscript. All the concerns raised have been carefully treated and an itemized reply to the reviewer's comments is presented in the revision files.

COMMENT 1:

The main objective of this study is to address the impact of spatial scale on the spatial heterogeneity of water use. However, this study treats water use as a whole, without separating the spatial heterogeneity of water use by different sectors (e.g., irrigation, industrial and domestic). Actually, water use by different sectors shows large spatial variation which I think is important in water use modelling.

COMMENT 2:

In the model framework, water use grid maps at different spatial scales are first prepared as inputs to the CA model. To generate the water use grid maps, the water use data at administrative survey scale is processed and downscaled to grid-based formats. Here, the iterative input selection algorithm is used to select the most relevant variables for water use, while the CNN model captures the relationships between input variables and water use. This step is important for the model performance, So I am wondering what are the most relevant variables for sectoral water use? This result should appear in this paper. Commonly, irrigation water use is mostly related to irrigated cropland area, and industrial/domestic water use are relevant to GDP/population density, as well as night light intensity. Whether the result of this study aligns with previous results?

COMMENT 3:

L145: "the appropriate spatial scale for water use simulation is identified using an end-to-end deep learning-based spatial scale adaptive selection model". What is the definition of "appropriate spatial scale"? In my view, the appropriate spatial scale should be a clear spatial resolution (e.g., 5km, 10km), and may vary across water use sectors or limited water use dataset. It is suggested to clarify the result of the appropriate spatial scale in gridded water use simulation.

RESPONSE to COMMENTS 1-3:

We appreciate the reviewer's comments regarding the treatment of different water use sectors, the selection of relevant variables, and the explanation of the appropriate spatial scale. In the revised manuscript, we have substantially expanded Section 2.1 to clarify these points. In preparing the water use grid maps, we first generated separate maps for irrigation, domestic, and industrial water use, based on prefecture-level statistical survey data and sector-specific predictor variables. These sectoral grids were then aggregated into total water use maps for modeling. This choice is based on two main considerations: (1) during the study period (1998–2013), China's total water use at the national scale entered a relatively stable stage in its temporal trend, with an average annual growth rate of only about 0.87% (from 505.53 billion m^3 in 1998 to 575.44 billion m^3 in 2013), mainly due to policy interventions, technological improvements, and changes in industrial structure. Since the primary objective of this study is to investigate the effects of spatial scale on the spatial heterogeneity of water use, it is more appropriate to adopt a temporally stable water use indicator as the simulation target, minimizing the confounding effects of sector-specific temporal fluctuations.

We also now explicitly list in Section 2.1 the most relevant input variables for each sector identified through the iterative input selection algorithm. These variables serve as inputs to the subsequent Convolutional Neural Network (CNN) model for downscaling prefecture-level data to grid-based water use maps.

The "appropriate spatial scale" in this study is a concept proposed in our previous work, designed to balance simulation accuracy with the spatial information density of rasterized water use data. In earlier studies, fixed spatial scales were applied across different prefectures, which failed to reflect the variability in land area, natural endowments, and water use structures among cities. This often led to discrepancies in information density across the simulated rasters. Excessively high resolutions could cause over- or underestimation due to data limitations, whereas overly coarse resolutions could obscure critical spatial variations. To address this, we previously developed a deep learning-based spatiotemporal scale adaptive selection (SSAS) model, in which the spatial scale selection module identifies the optimal spatial resolution of input variables by maximizing information density, quantified using Conditional Entropy, while balancing simulation accuracy through Kullback-Leibler Divergence Loss and Relative Error. This approach enables each prefecture to have its own optimal spatial scale rather than adopting a uniform resolution.

In the revised manuscript, we have substantially expanded Section 2.1 ("Water Use Grid Maps Generating") to address these points. The revised section now reads:

Revised Section 2.1:

The spatial scale of water use simulation is determined by the spatial scale of the input data, so water use grid maps at different spatial scales were prepared as input to the simulation model. To obtain the water use grid maps, several steps should be done to convert the water use data at administrative survey scale into spatially explicit grids of varying resolutions.

The grid maps of irrigation, domestic, and industrial water use are generated from the prefecture-level statistical survey data and water use sector-specific predictor variables. For each sector, the most relevant input variables are identified through an iterative input variables selection algorithm (Zhang et al., 2023; Zhang et al. 2025). Specifically, irrigation water use was modeled by the potential evapotranspiration, normalized difference vegetation index (NDVI), rainfall and soil moisture; domestic water use was modeled by population, rainfall, temperature and night-light; industrial water use was modeled by GDP, night-light, population and rainfall. And then these sectoral gird maps were aggregated to form total water use grid maps for modeling. This aggregation is done for two reasons: the first one is that the temporal trend of the total water use has become stable during the study period (1998–2013) and future. The average annual growth rate is only about 0.87% (from 505.53 billion m3 in 1998 to 575.44 billion m3 in 2013) due to the policy interventions, technological improvements, and industrial structure changes. Since the primary objective of our study is to examine the influence of spatial scale on the spatial heterogeneity of water use, a temporally stable indicator helps minimize the confounding effects of sector-specific temporal fluctuations; the second reason is that the total water use can figure out the scale effects across regions instead of the sector-level temporal variability while the sectoral differences are implicitly in the inputs before the aggregation.

Earlier studies often applied a fixed spatial resolution in different regions, which could not account for differences in land area, natural endowments, and water use structures, and leaded to the discrepancies in information density and potential over- or underestimation of water use. To address this issue, an appropriate spatial scale can be determined by the deep learning-based spatiotemporal scale adaptive selection model (Liu et al., 2021; Zhang et al., 2025). And the model can balance the accuracy of the simulation based on the spatial information density of gridded water use data, and its results vary across prefectures. The spatial scale selection module in the selection model figures out the appropriate spatial scale by maximizing information density while balancing simulation accuracy in terms of the Conditional Entropy, Kullback–Leibler Divergence Loss and Relative Error performance metrics. This selection module enables each prefecture to adopt its own appropriate spatial scale rather than a fix resolution. Finally, total water use grid maps are generated at three spatial resolutions: the small scale (e.g., 1 km), the appropriate spatial scale as determined by the selection module, and the prefecture scale as the statistical survey water use data.

***COMMENT* 4:**

Water use simulation from the probability rule CA model (Section 4.1.1) are not validated. This part uses the Akaike Information Criterion (AIC) to determine the most suitable probability distributions for water use grids across various prefectures. However, the optimal probability distributions also rely on the input data (e.g., the long-term gridded water use data). As water use in China shows significant spatial and temporal variation between different periods, it is doubtable that the probability rule CA model can used for water use prediction.

RESPONSE:

We appreciate the reviewer's attention to the use of the probability rule in CA-based water use simulation and the concern that significant spatial and temporal variation in water use across different periods may challenge its applicability. In our

CA-based simulation, the probability rule is designed to represent the stochastic state transitions of water use at the grid level, and this choice is both theoretically and practically justified for the following reasons.

First, the evolution of water use is influenced by both deterministic drivers and inherent variability. A purely deterministic CA update rule risks over-smoothing or ignoring this randomness, whereas the probability rule allows structured temporal dependence and stochastic variation to be represented in a unified framework. This is particularly important in China, where substantial variations between periods exist—rather than assuming temporal stability, the probability rule directly incorporates these variations by deriving transition probabilities from observed historical changes in each prefecture. Second, the method is locally adaptive. The state transition matrix is calibrated independently for each prefecture, so local variation patterns are preserved. Equal-frequency categorization ensures the k intervals reflect each cell's own historical variability, and the most suitable probability distribution for each interval is selected using the Akaike Information Criterion (AIC) from candidates including normal, lognormal, exponential, gamma, and uniform. This ensures that both the magnitude and volatility of water use in different regions and periods are reflected in the fitted distributions. Third, the calibrated parameter k controls the granularity of the state representation and is tuned using Root Mean Squared Error (*RMSE*) and Relative Error (*RE*) in an independent validation period, balancing resolution and generalizability. This calibration, conducted for each prefecture, ensures the model captures temporal changes without overfitting.

As for validation, we compared the simulated gridded water use against observed water use for prefectures and years, using *RE* and *RMSE* as metrics. The results showed good agreement in both distributional shape and spatial patterns, though localized deviations exist in some regions.

In the revised manuscript, we have clarified these point both in Section 2.2 (overview of the CA framework) and in Section 2.2.1 (details of the probability rule), ensuring that the theoretical basis and empirical support are both addressed. And the validation procedure and results are included in Section 4.1.1.

Revised First Paragraph in Section 2.2:

The CA model, grounded in complexity theory, is widely used in land use and urban growth modeling. It provides a robust platform for simulating spatial phenomena governed by local interactions and transition rules (Sapino et al. 2023, Tariq et al. 2023). Each cell in a CA model represents a discrete spatial unit that updates its state over time based on predefined rules and the states of its neighboring cells. It's decentralized, bottom-up modeling structure enables the simulation of complex global behaviors emerging from simple local dynamics (Al-Shaar et al. 2022, Wang et al. 2020). Both the probability and the linear update rules are designed and tested to capture the dual nature of water use dynamics. The probability rule has been widely applied in significant spatial and temporal variation areas in land use simulation and other fields. It will be designed here for the water use at different scales. Rather than assuming temporal stability, the probability rule explicitly incorporates the variations through calibrating the state transition matrix and probability distributions for each prefecture independently by the own historical water use record. This rule enables the simulation to capture both the structured temporal dependence and the inherent randomness in water use, ensuring adaptability to local conditions. The linear update rule assumes that changes in water use are more deterministic and can be approximated as a linear combination of the cell's own state and those of its neighbors. This rule is more appropriate for long-term, high spatial autocorrelation, and persistent patterns. After implementing and comparing the water use simulation results of the two rules in the CA framework, their results can assess the relative effectiveness of stochastic versus deterministic update mechanisms across different spatial scales. These two rules not only strengthen the robustness of the modeling framework but also provide insights into the dominant processes shaping water use dynamics in different regions.

Revised Section 2.2.1:

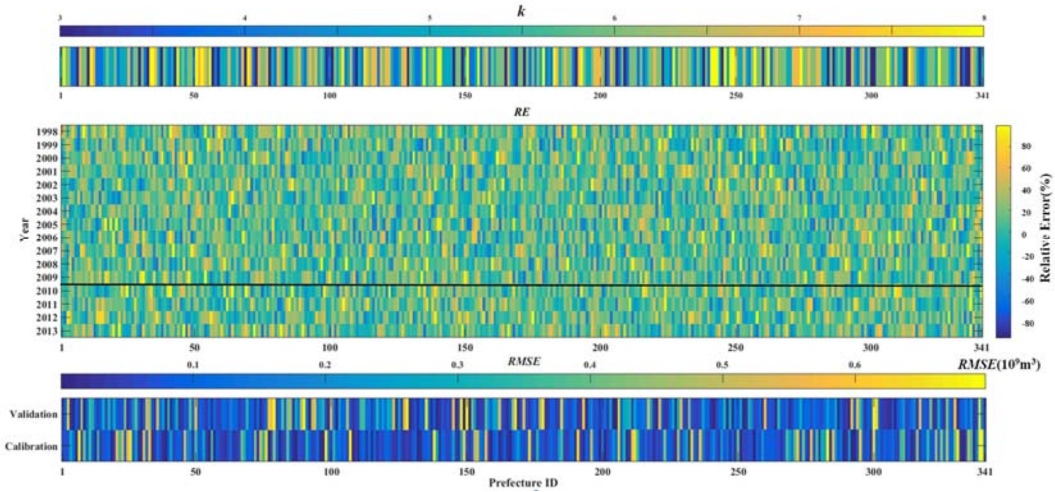
The probability rule in the CA model is designed to represent the stochastic state transitions of water use over time. It abstracts the temporal dynamics of water use at the grid level into a probabilistic transition framework that can be applied consistently across different spatial scales and regions, while remaining adaptable to significant spatial and temporal variations. This adaptability is achieved by calibrating the update rule separately for each prefecture by its own historical water use record for appropriately capturing both long-term trends and localized fluctuations. In this approach, the state of each grid cell (i.e., representing the amount of water use) is divided into *k* distinct intervals using equal-frequency categorization based on the cell's historical water use record. This categorization ensures that the intervals reflect the variations in water use over time. For each interval, the most suitable statistical distribution is selected using the Akaike Information Criterion (AIC). The selection process enables the model to represent the probabilistic characteristics of water use within each intensity class. And the distribution is chosen from a set of candidate distributions, including normal, lognormal, exponential, gamma, and uniform.

Once the optimal distribution is found for each interval, a state transition matrix is constructed based on observed transitions of grid cells between intervals from one year to the next. The transition matrix captures the likelihood of a grid cell moving from its current water use state to another in the subsequent time step. The model first generates the next state probabilistically through the transition matrix, and then the water use samples are generated from the corresponding probability distribution. These two steps incorporate both the structured temporal dependence and the inherent randomness in future water use patterns.

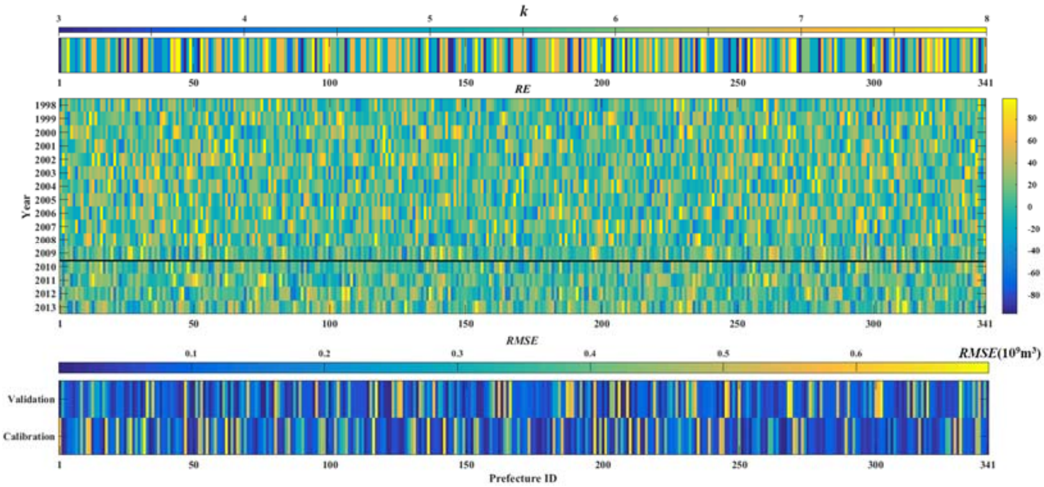
In the probability rule, the calibrated parameter is the number of state intervals and is denoted as *k*. The value of k directly affects the granularity of the state categorization and the accuracy of the state transition matrix. A larger k increases the resolution of the state representation and captures the finer variations in water use, but a larger k can also lead to overfitting. And a smaller k oversimplifies the demand pattern. To calibrate the parameter k, the historical and observed water use data is divided into a calibration and a validation sets. And the performances of the model with different k values is then evaluated by the Root Mean Squared Error (*RMSE*) and Relative Error (*RE*) metrics. The optimal k can be calibrated by the minimums of the *RMSE* and *RE* in the validation period, ensuring a balance between model accuracy and generalizability.

Added Validation Procedure and Results in Section 4.1.1:

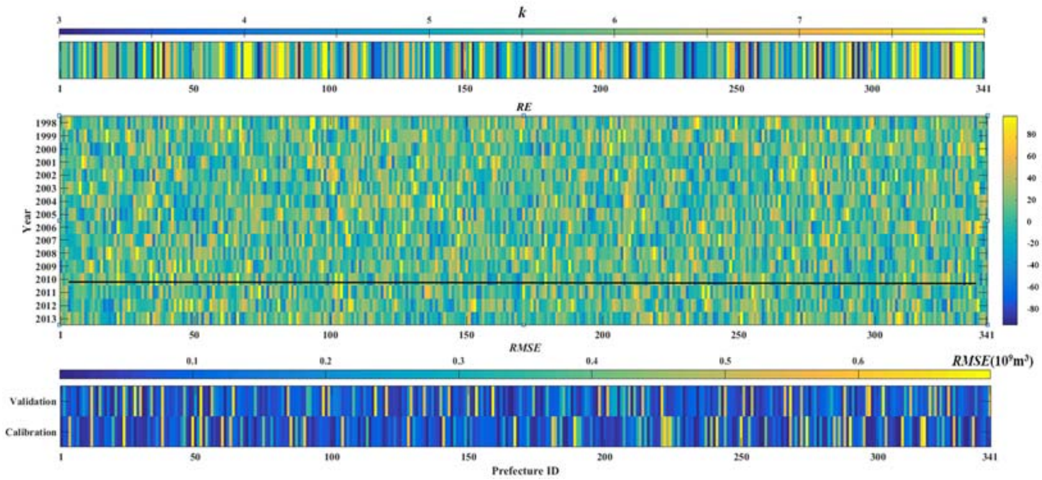
In the CA model with the probability rule, the number of state intervals (k) is the only parameter to be calibrated. The dataset from 1998–2009 is used for calibration and 2010–2013 for validation, with *RMSE* and *RE* as performance metrics. The optimal value of k at three spatial scales (1 km, appropriate scale, prefecture scale) for each prefecture is determined by minimizing *RMSE* and *RE* in the validation period. The calibrated k values for each prefecture, along with the corresponding *RMSE* and *RE* during the calibration (1998–2009) and validation (2010–2013) periods at the three spatial scales, are presented in Figure 3.



(a)



(b)



(c)

Figure .3 Optimal parameters of the probability rule CA model and the model performances at: (a) 1km scale; (b) appropriate spatial scale; (c) prefecture scale

According to the results shown in Figure 3, the calibrated parameter k exhibits clear spatial heterogeneity across prefectures and varies with spatial scales. At the 1 km scale (Figure 3(a)), most prefectures show k values concentrated around 5–6, corresponding to relatively low *RMSE* and *RE* values. This suggests that a moderate number of state intervals can effectively capture local water use variability while avoiding overfitting. In these areas, the probability distributions and transition probabilities appear to reflect stable temporal patterns, resulting in more accurate simulations. At the appropriate spatial scale (Figure 3(b)), the distribution of k becomes more diversified among prefectures. Some regions require larger k values (>=7) to preserve finer distinctions in water use states, while others perform better with smaller k values (<=4) that smooth out excessive variability. Different from the results at the 1 km scale, the overall *RMSE* and *RE* values are slightly higher, indicating that while the appropriate scale balances detail and generalization, it may not fully capture abrupt local changes in some prefectures. At the prefecture scale (Figure 3(c)), k values are generally smaller (mostly 3–4), reflecting the reduced spatial detail at coarser resolution. Thus, their accuracies of the simulation decrease, with higher RMSE and RE values. When the input variability of small scale is strongly aggregated at large scales, fewer state intervals oversimplify the temporal transitions, leading to greater deviations from observed water use patterns.

After determining the optimal k values at each scale, the next step is to characterize the statistical nature of water use within each state interval. The Akaike Information Criterion (AIC) is taken as performance metric to select the most suitable probability distribution for each interval in every prefecture. The AIC can balance the model fitness and the complexity through penalizing excessive parameters, it can reduce the risk of overfitting. The selected distribution types not only fit the historical data well but also is used to generate the future scenarios. The results of the optimal probability distributions for water use grids at the three different spatial scales (i.e., 1 km scale, appropriate spatial scale, and prefecture scale) are shown in Figure 4. These distributions, combined with the calibrated k values, form the basis of the probability rule CA model's ability to reproduce the spatial and temporal heterogeneity of water use.

COMMENT 5:

This study calibrates the parameters in the linear rule CA model for the 1998–2009 while the dataset from 2010–2013 is for its validation. However, the calibration and validation processes are not clear. Which datasets are used for model evaluation, the prefecture water use data or the gridded water use maps?

RESPONSE:

We appreciate the reviewer's request for clarification. The calibration and validation were based on prefecture-level statistical water use data collected from water resources bulletins and related surveys, which were used as reference (truth) values. The CA model simulations were run at the gridded scale, and the simulated water use was subsequently aggregated to prefecture boundaries to produce prefecture-level totals. These aggregated totals were then compared with the observed prefecture-level statistics to compute *RMSE* and *RE*. The calibration period (1998–2009) was used to optimize the model parameters by minimizing these metrics, while the validation period (2010–2013) applied the calibrated parameters without modification for independent evaluation. This clarification will be explicitly added to Section 4.1.2 in the revised manuscript.

Revised Section 4.1.2:

There are three parameters to be calibrated in the linear rule CA model: the self-influence coefficient α, the neighboring influence coefficient β, and the spatial decay exponent p. The calibration and validation are taken the statistics water use at the prefecture-level as the reference (i.e., observed water use data). Specifically, for a given parameter set, the gridded water use is firstly simulated. The simulated grids are then aggregated into the total water use at each prefecture scale along their boundaries. These total water uses are assessed by the observed water use data from water resources bulletins and related statistical surveys. The calibration period covers 1998–2009 and the parameter values are determined by minimizing *RMSE* and *RE* between the simulated and observed total water use at the prefecture scale. The validation period covers 2010–2013 and the performance of the model is also assessed by *RMSE* and *RE*. The optimal parameters at three spatial scales during the calibration and validation periods, are illustrated in Figure 5.

COMMENT 6:

The main objective of the model framework is to generate water use data at multiple spatial scale. There are many gridded water use products at both global or country scale for China (e.g., Hou et al., 2024, ESSD; Huang et al., 2018, HESS; Zhang et al., 2025, Scientific Data),a s well as the high-resolution hydrological model simulations. It is necessary to compare the water use simulation with previous products, which helps to evaluate the reliability of the model framework of this study.

RESPONSE:

We appreciate the reviewer's suggestion regarding the comparison with existing gridded water use datasets. We fully agree that such a comparison is important to assess the reliability and added value of our model framework. We downloaded the water use raster datasets from the three previous studies mentioned by the reviewer for examination. Huang et al. (2017) provides global-scale water use data at a 0.5° resolution; Hou et al. (2024) focuses only on industrial water use at a 0.1° resolution; and Zhang et al. (2025) provides water use data at a 0.1° resolution. In contrast, our water use simulation is at a finer spatial resolution (1km × 1km, appropriate spatial scale), enabling more detailed representation of spatial heterogeneity in water use.Due to these substantial differences in spatial resolution, direct comparison of spatial distribution patterns is not feasible. So we conducted comparisons in terms of statistical performance metrics. We have added comparison with previous studies in section 5.3 in the Discussion section to present these performance-based comparisons and elaborate on the implications of the differences in spatial resolution.

Added Comparison with Previous Studies in Section 5.3:

There had been some water use simulation results at previous studies. Fox example, Huang et al. (2018) produced a global-scale monthly water withdrawal dataset at 0.5° resolution, distinguishing six sectors (e.g., irrigation, domestic, electricity generation, livestock, mining, manufacturing) over the period 1971–2010; Hou et al. (2024) developed China's industrial water withdrawal dataset (CIWW), providing gridded monthly data from 1965 to 2020 at 0.1° and 0.25° resolutions; Zhang et al. (2025) presented a high-resolution sectoral water use dataset (HSWUD) for mainland China, covering irrigation, manufacturing, thermal power cooling, and domestic use at 0.1° × 0.1° resolution, with strong consistency to prefecture-level statistics ($R^2 \approx 0.88$). As the results shown in the previous sections, our dataset is generated at different spatial resolutions (e.g., 1 km×1 km, appropriate spatial scale), enabling detailed representation of spatial heterogeneity within prefectures. Due to the substantial differences in spatial resolutions between these datasets, it is not easy to compare the differences of the spatial distribution patterns. But the relative values of performance metrics such as *RMSE* and *RE* can figure out the better one among them. The values of *RMSE* within 0.1 (i.e., normalized by mean water use) and a *RE* within −20 % to +30 % are found across all prefectures according to the results of our simulation. And all these results are consistent.

Thus, the results show that, relative to the three reference datasets, our model's prefecture-level water use estimates achieve a *RMSE* within 0.1 (normalized by mean water use) and a *RE* within −20 % to +30 % across all prefectures. These results within the range generally are considered acceptable for large-scale water use modeling, indicating that our estimates are consistent with these previous studies while offering finer spatial details.

COMMENT 7:

Figure 4 & 6: water use is high in many irrigated areas. However, water use in the North China Plain which is marked with intensive irrigation and population, shows moderate level of water use, lower than that of the northeastern China. This result is contrary with previous estimates.

RESPONSE:

We appreciate the reviewer's observation and understand the concern regarding the relatively moderate water use estimates for the North China Plain (NCP) compared to northeastern China in our results. Several factors may explain this phenomenon:

(1) Statistical survey data trends – Our model is calibrated and validated against prefecture-level statistical survey data (e.g., water resources bulletins). In recent years, a decreasing trend of agricultural water use has been reported in many NCP prefectures, partly due to improvements in irrigation efficiency, the implementation of water-saving policies, and adjustments in cropping structures. These changes are reflected in our gridded estimates.

(2) Sectoral aggregation effects – Total water use in our study is the aggregation of irrigation, industrial, and domestic sectors. While water use in the NCP remains dominated by irrigation, many northeastern prefectures show a growing contribution from industrial water use, particularly heavy industries and thermal power generation, which elevate their overall totals relative to the NCP.

(3) Climatic and water availability factors – Despite a shorter growing season, northeastern China often supports water-intensive crops and benefits from relatively abundant local water resources, resulting in higher water use per unit area.

In addition, some previous studies that reported higher water use in the NCP were sector-specific (mainly irrigation) or based on different temporal baselines, which partly explains the discrepancy with our aggregated, multi-sector results. We have added a paragraph in Section 4.1.2 to clarify these points, so that readers can understand the reasons for the observed differences.

Added Paragraph in Section 4.1.2:

Moderate water use levels have been found in the North China Plain (NCP), compared with northeastern China, even though the NCP features intensive irrigation and high population density. Several factors contribute to this pattern. First, the calibration and validation rely on prefecture-level statistical survey data (e.g., water resources bulletins). In recent years, a decreasing trend of agricultural water use has been reported in many NCP prefectures, partly due to improvements in irrigation efficiency, the implementation of water-saving policies, and adjustments in cropping structures. Second, total water use in our study is an aggregate of irrigation, domestic, and industrial sectors. While total water use in the NCP remains dominated by irrigation, many northeastern prefectures show a growing contribution from industrial water use, particularly heavy industries and thermal power generation, leading to higher overall totals. Third, climatic and water availability differences also play a role. Despite a shorter growing season, northeastern China often supports water-intensive crops and benefits from relatively abundant local water resources, which result in higher water use per unit area. In addition, some previous studies that reported higher water use in the NCP were sector-specific (mainly irrigation) or based on different temporal baselines, which partly explains the discrepancy with our aggregated results.

COMMENT 8:

L39: Key words: This study is all about water demand/water use, and “water resources management” and “water scarcity assessment” are not suitable for the keywords.

RESPONSE:

We appreciate the reviewer’s suggestion regarding the keywords. We have revised the keywords in the manuscript, replacing “water resources management” and “water scarcity assessment” with more appropriate terms directly related to water demand and water use.

Updated keywords: water use; spatial scale; cellular automata; multi-scale simulation

COMMENT 9:

L145： I don’t find the reference for Liu et al., 2022.

RESPONSE:

We thank the reviewer for pointing this out. We have corrected the citation year for *Liu et al.* to ensure consistency between the in-text citation and the reference list.

We have also carefully checked all other references to avoid similar inconsistencies.