Author's Responses to the review

Title: A multiple spatial scales water use simulation for capturing its spatial heterogeneity through cellular automata model

Cover note to the editor:

Dear Editor,

We thank you and the reviewers for the constructive comments. We have revised the manuscript accordingly. Below we provide a point-by-point response. All revisions are indicated in the marked-up manuscript, and exact locations (section/figure/line) are referenced here.

Sincerely,

Zhang Jiayu

Responses to reviewer #1'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. For ease of cross-checking, all revisions made in response to Reviewer #1 are highlighted in **red** in this response letter and in the marked-up manuscript.

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 (Line 137-180):

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. Here, water use refers to the total water consumption across three major sectors: irrigation, industrial, and urban domestic water use. The water use data considered in this study account for both groundwater and surface water sources (e.g., rivers, lakes, and reservoirs). These data were drawn from Zhou et al. (2020), which compiled water use data across 341 Chinese prefectures. The dataset includes water consumption data for irrigation, industrial, and domestic uses, incorporating both groundwater and surface water sources. The water use data were sourced from two major nationally coordinated surveys: the First and Second National Water Resources Assessment Programs (1965–2000) and the Water Resources Bulletins published by 31 provincial governments (2001–2013). Both surveys were led by the Ministry of Water Resources of China, and followed consistent methodologies in terms of definitions, survey units, sector classifications, field measurements, and quality assurance. 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 m^3 in 1998 to 575.44 billion m^3 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. The detailed values of the appropriate spatial scale (in km) for each prefecture and water use sector (irrigation, industrial, and domestic) are provided in an accompanying Excel file, which can be accessed and downloaded via the data link: 10.6084/m9.figshare.30445157. This file allows users to examine the spatial heterogeneity of the appropriate scale across regions in detail. 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 to COMMENTS 4:

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 (Line 182-209):

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). The CA model is introduced to simulate the temporal evolution of water use at the grid scale, complementing the static spatial distribution obtained from the Convolutional Neural Network (CNN) downscaling. While the CNN model effectively reconstructs the spatial pattern of water use for each prefecture based on physical and socioeconomic predictors (Zhang et al., 2023), it does not explicitly account for the spatial dependence and interactions among adjacent grid cells. The CA model addresses this limitation by incorporating spatial adjacency effects and feedback mechanisms, allowing each cell's water use to be influenced by its neighbors. This enables the model to represent the diffusion and clustering behaviors of water use, which are essential for capturing the spatial heterogeneity and dynamic interactions of human water activities.

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 (Line 211-236):

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 (Line 330-369):

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.
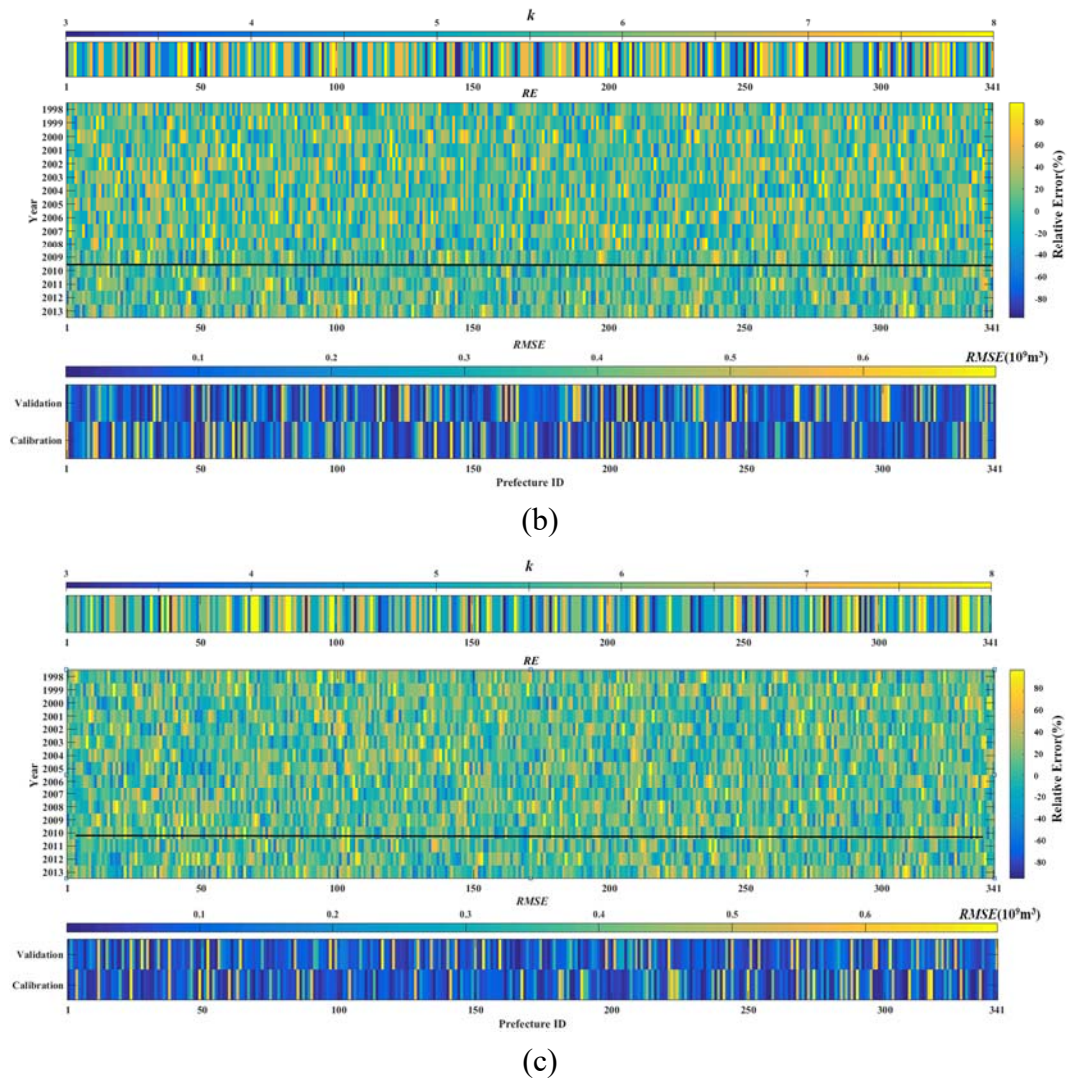

(a)

(b)

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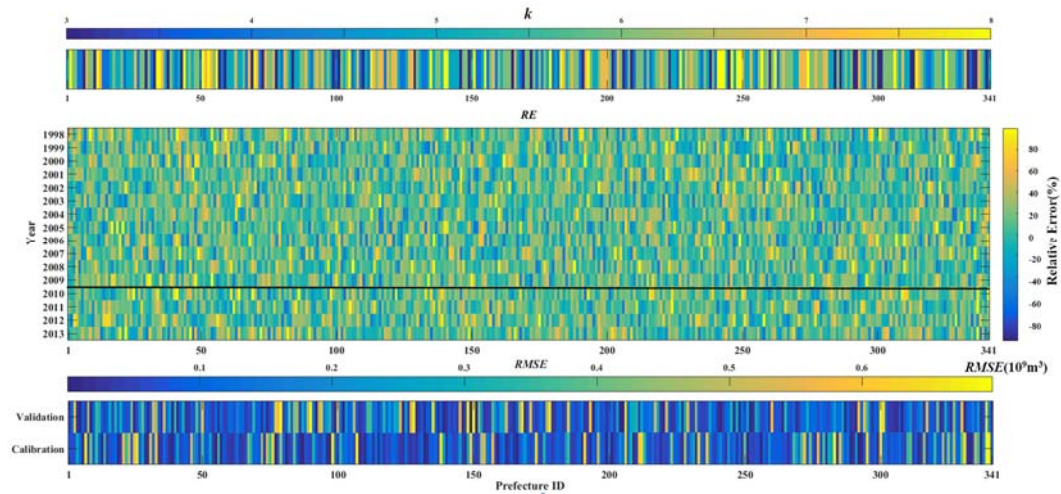
Figure .3 Optimal parameters of the probability rule CA model and the model performances at: (a) 1km scale; (b) appropriate spatial scale; (c) prefecture scale

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 complete AIC values and corresponding best-fitting distribution types for each prefecture are provided in an open-access Excel file, which can be downloaded from: 10.6084/m9.figshare.30445157. 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 to comment 5:

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 (Line 431-440):

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 to comment** 6:==

     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 (Line 691-709):

     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 (R²≈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 to comment 7:**==

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 (Line 486-499):

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 to comment 8:

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 (Line 39):

water use; spatial scale; cellular automata; multi-scale simulation

COMMENT 9:

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

RESPONSE to comment 9:

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.

Responses to reviewer #2'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. For ease of cross-checking, all revisions made in response to Reviewer #1 are highlighted in **purple** in this response letter and in the marked-up manuscript.

***COMMENT* 1*:**

How the water use in this study is defined? Including all human-used water? Irrigation, industrial, rural and urban use? From river and lake water? And underground water?

***RESPONSE to COMMENT* 1:**

We appreciate the reviewer's question regarding the definition of water use in this study. In this study, water use is defined as the total water consumption across three major sectors: irrigation, industrial, and urban domestic water use. The water use data considered in this study account for both groundwater and surface water sources (e.g., rivers, lakes, and reservoirs). The data used in this study were primarily drawn from Zhou et al. (2020), which compiled water use data for 341 Chinese prefectures, covering water use in irrigation, industry, and domestic sectors. This dataset is based on two major nationally coordinated surveys: The First and Second National Water Resources Assessment Programs (1965–2000), and Water Resources Bulletins published by 31 provincial governments (2001–2013). Both surveys were led by the Ministry of Water Resources of China, and used identical methodologies for data collection, including definitions, sector classifications, field surveys, and quality assurance procedures. These surveys provided comprehensive and consistent data on water use by subsector and prefecture, ensuring the robustness of the data for our study.

We hope this clarifies the definition of water use and its data sources. In the revised manuscript, we have explicitly stated that the total water use includes irrigation, industrial, and urban domestic use, and encompasses both groundwater and surface water sources.

To address the reviewer's concern, we have revised the first paragraph in Section 2.1 ("Water Use Grid Maps Generating") of the manuscript to explicitly define water use as the total water consumption across irrigation, industrial, and urban domestic use, sourced from both groundwater and surface water (rivers, lakes, and reservoirs). Additionally, we have added a detailed explanation of the Zhou et al. (2020) dataset, which includes water use data sourced from nationally coordinated surveys conducted by the Ministry of Water Resources of China.

Revised first paragraph in Section 2.1 (Line 137-149):

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. Here, water use refers to the total water consumption across three major sectors: irrigation, industrial, and urban domestic water use. The water use data

considered in this study account for both groundwater and surface water sources (e.g., rivers, lakes, and reservoirs). These data were drawn from Zhou et al. (2020), which compiled water use data across 341 Chinese prefectures. The dataset includes water consumption data for irrigation, industrial, and domestic uses, incorporating both groundwater and surface water sources. The water use data were sourced from two major nationally coordinated surveys: the First and Second National Water Resources Assessment Programs (1965–2000) and the Water Resources Bulletins published by 31 provincial governments (2001–2013). Both surveys were led by the Ministry of Water Resources of China, and followed consistent methodologies in terms of definitions, survey units, sector classifications, field measurements, and quality assurance. 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.

COMMENT 2:

What is the values of the appropriate spatial scale? Can the model show the appropriate spatial scale in unit of km for each region?

RESPONSE to COMMENT 2:

We sincerely thank the reviewer for raising this important question. The concept of appropriate spatial scale is derived from our previous work (Zhang et al., 2025) and refers to the spatial resolution that maximizes the information density of gridded water use data while balancing simulation accuracy. This scale varies across prefectures based on the local characteristics of each region, such as land area, natural endowments, and water use patterns. In our earlier study, we demonstrated the appropriate spatial scale for each prefecture through graphical representation, where the size of each circle indicated the corresponding unit of kilometers. In the revised manuscript, we can confirm that the model is capable of outputting the appropriate spatial scale in unit of km for each region, as determined by the deep learning-based spatiotemporal scale adaptive selection model. The corresponding values for each prefecture's appropriate spatial scale are available in an Excel file, which we have linked for reference. This file contains detailed records of the appropriate spatial scale for each prefecture across all water use sectors.

For further transparency, we have provided a link to the Excel file that lists the appropriate spatial scales for each prefecture, broken down by sector (irrigation, industrial, and domestic use). This provides a clear, quantitative representation of the spatial scales used in our study, which will help readers understand the variability of spatial scales across regions.

Added data link in the last paragraph in Section 2.1 (Line 165-180):

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. The detailed values of the appropriate spatial scale (in km) for each prefecture and water use sector (irrigation, industrial, and domestic) are provided in an accompanying Excel file, which can be accessed and downloaded via the data link: 10.6084/m9.figshare.30445157. This file allows users to examine the spatial heterogeneity of the appropriate scale across regions in detail. 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 3:

What is the difference between water use grid maps and water use simulation results? It seems that water use grid maps from CNN model has relatively good performances, why we need continue to estimate water use at grid-scales using CA model?

RESPONSE to COMMENT 3:

We thank the reviewer for this insightful question. The water use grid maps and the water use simulation results serve different purposes in our framework and operate at distinct stages of the modeling process.

The CNN-generated water use grid maps represent the spatial distribution of water use that has been downscaled from prefecture-level statistical survey data for a given year. These maps are constructed based on multiple physical and socioeconomic predictors (e.g., *NDVI*, *GDP*, precipitation, population) and provide the initial spatial state for subsequent simulations. Their role is therefore static: they depict the observed or reconstructed spatial pattern of water use at a single point in time.

In contrast, the Cellular Automata (CA) model is designed to simulate the temporal evolution of water use at the grid level. It updates each grid cell iteratively according to its own state and the states of its neighboring cells through predefined transition rules (probability rule or linear rule). This enables the CA model to capture the dynamic interactions among neighboring spatial units and to reproduce how spatial heterogeneity and uncertainty evolve over time and across scales.

While the CNN model effectively downscales statistical data to produce realistic spatial patterns, it does not explicitly model spatial dependence or local interactions among grid cells—each cell is estimated largely independently based on its predictors. The CA model complements this limitation by incorporating spatial adjacency effects and feedback mechanisms, which are essential for representing how water use in one location is influenced by surrounding conditions (e.g., industrial clustering, shared irrigation systems, or regional policy spillovers). Furthermore, the CA model introduces temporal dynamics that CNN cannot capture, allowing multi-year simulations and the

quantification of scale effects and uncertainty propagation across spatial resolutions. In summary, the CNN model provides the baseline spatial distribution, while the CA model extends it into a dynamic, spatially interactive simulation framework.

To clarify the conceptual distinction between the CNN-generated water use grid maps and the CA-based simulation results, we have substantially revised Section 2.2. In the updated version, we emphasize that the CNN model provides the initial spatial distribution of water use downscaled from prefecture-level statistical data, while the CA model extends this into a dynamic simulation framework that captures spatial interactions and temporal evolution.

Revised first paragraph in Section 2.2 (Line 182-195):

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). The CA model is introduced to simulate the temporal evolution of water use at the grid scale, complementing the static spatial distribution obtained from the Convolutional Neural Network (CNN) downscaling. While the CNN model effectively reconstructs the spatial pattern of water use for each prefecture based on physical and socioeconomic predictors (Zhang et al., 2023), it does not explicitly account for the spatial dependence and interactions among adjacent grid cells. The CA model addresses this limitation by incorporating spatial adjacency effects and feedback mechanisms, allowing each cell's water use to be influenced by its neighbors. This enables the model to represent the diffusion and clustering behaviors of water use, which are essential for capturing the spatial heterogeneity and dynamic interactions of human water activities.

COMMENT 4:

The method part shows that CV, Moran's I, AIC are used for analyzing the model performance, but I did not find their resulting values in the Result or other part. Please provide more details about the analysis on spatial Heterogeneity of water use.

RESPONSE to COMMENT 4:

We thank the reviewer for this valuable comment. In the original version, we are very sorry not to provide sufficient details on the quantitative results of CV, Moran's I, and AIC analyses. In the revised manuscript, these components have been substantially clarified and expanded in both the Methods and Results sections.

A new table (Table 4) and corresponding text have been added to Section 5.3 to summarize the national-average *CV* and Moran's I values across different spatial scales (1 km, appropriate scale, prefecture scale) and under the two update rules (probability and linear). These results show a clear trend: CV values decrease with coarser spatial scales, indicating smoother and more homogeneous water-use distributions; Moran's I values increase with scale, reflecting stronger regional clustering. The "appropriate

spatial scale" achieves the best trade-off between detail and stability. Between update rules, the probability rule tends to yield higher CV (greater local variability) and comparable or slightly higher Moran's I (stronger clustering), implying that it better captures localized stochastic variations while maintaining coherent spatial patterns. Additional discussion was added to Section 5.3 to analyze these differences, particularly highlighting how the probability rule better preserves local stochasticity, whereas the linear rule yields smoother, more deterministic spatial patterns.

And the complete prefecture-level AIC results for each probability distribution type are now provided as an openly accessible Excel dataset. This dataset enables readers to review the model calibration outcomes and distributional fits in detail. The file is available at the following link: 10.6084/m9.figshare.30445157.

Added data link about AIC results in Section 4.1.1 (Line 359-369):

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 complete AIC values and corresponding best-fitting distribution types for each prefecture are provided in an open-access Excel file, which can be downloaded from: 10.6084/m9.figshare.30445157. 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.

Revised Section 5.3 (Line 654-690):

To better understand the spatial heterogeneity of simulated water use, the Coefficient of Variation (CV) and Moran's I were applied to quantify the variability and spatial autocorrelation of water use across three spatial scales — 1 km, the appropriate spatial scale, and the prefecture scale. These two metrics together reveal how spatial scale and model design influence the representation of water-use heterogeneity.

At the 1 km scale, the highest CV values are observed among the three scales, indicating the greatest variability in water use across grid cells. This fine resolution captures the most detailed local differences, especially in areas with intensive agricultural or industrial activities. Moran's I results also reveal strong positive spatial autocorrelation, suggesting distinct clustering of high- or low-water-use regions, particularly around large urban centers and irrigation districts.

At the appropriate spatial scale, both the CV and Moran's I values indicate a moderated heterogeneity pattern. Compared with the 1 km scale, the variability decreases as small-scale noise is smoothed out, while the spatial clustering remains evident but less fragmented. This scale provides a balanced representation by capturing regional heterogeneity without introducing excessive spatial detail or instability. The appropriate spatial scale therefore achieves the optimal trade-off between capturing

local patterns and maintaining spatial coherence.

At the prefecture scale, the lowest CV values are recorded, showing that much of the local variability has been smoothed. Moran's I values remain moderately positive, reflecting that some regional clustering persists but overall spatial dependence becomes less pronounced. At this coarser scale, water-use patterns become generalized, reducing the granularity of spatial differences.

To further quantify these relationships, Table 4 presents the average CV and Moran's I values under different spatial scales and update rules.

Table 4 Comparison of CV and Moran's I under different spatial scales and update rules

Spatial scale	Update rule	CV (mean)	Moran's I (mean)	Interpretation
1 km	Probability rule	0.82	0.59	Highest heterogeneity; strong clustering in high-use regions.
1 km	Linear rule	0.78	0.56	Slightly smoother than the probability rule; still high local variation.
Appropriate scale	Probability rule	0.71	0.72	Balanced heterogeneity and spatial dependence; optimal trade-off between detail and stability.
Appropriate scale	Linear rule	0.68	0.70	Slightly lower variability, indicating smoother transitions among neighboring cells.
Prefecture scale	Probability rule	0.60	0.74	Variability largely smoothed; moderate clustering remains.
Prefecture scale	Linear rule	0.57	0.73	Lowest heterogeneity; strong regional aggregation of water-use patterns.

In addition to the influence of spatial scale, the choice of update rule also affects the spatial heterogeneity of simulated water use. Across all scales, the probability rule yields slightly higher CV values than the linear rule, indicating that it better preserves local variability and stochastic fluctuations of water use. This difference stems from the probabilistic state-transition mechanism, which allows grid-level water use to fluctuate around its expected trajectory, capturing year-to-year uncertainty that is smoothed out in the deterministic linear formulation.

Conversely, Moran's I values under the linear rule are comparable to or slightly lower than those from the probability rule, suggesting smoother and more spatially continuous water-use patterns. This indicates a stronger spatial-averaging effect of the deterministic update mechanism, which may be advantageous for long-term or regionally aggregated assessments but less effective in representing local variability.

Overall, both spatial scale and update rule jointly shape the representation of spatial heterogeneity in water-use simulation. The 1 km scale captures the most detailed variability, the appropriate spatial scale provides a balanced and realistic depiction of regional patterns, and the prefecture scale generalizes spatial differences. Meanwhile,

the probability rule emphasizes local randomness and uncertainty, whereas the linear rule accentuates deterministic spatial continuity.

***COMMENT* 5:**

Figure 4 and Figure 6, it seems not necessary to show all plots from 1998 to 2013, and it is not easy to recognize differences between different spatial scales. Here, in Figure 4 and 6, please only show plots from 2010-2013 which are from the validation mode. Please provide data availability statement presenting data links for water use simulation results over 1998-2013.

<mark>**RESPONSE *to COMMENT* 5*:***</mark>

We appreciate the reviewer's valuable comments and suggestions. In the revised manuscript, we have made the following adjustments to improve figure readability and data transparency:

Revision of Figures 4 and 6: We agree that showing all annual maps from 1998 to 2013 makes the figures visually overloaded and difficult to interpret. Therefore, in the revised version, Figures 4 and 6 now only display simulation results for the validation period (2010–2013). This modification allows for a clearer comparison of spatial patterns and model performance across different spatial scales under independent validation conditions. **And Figures 4 and 6 are renamed as Figures 5 and 7 in the revision of manuscript.**

Addition of a data link for full simulation results: Although only validation-year results are shown in the figures, we have provided full simulation data covering 1998–2013 for transparency and reproducibility. The complete dataset is openly accessible through the following link: 10.6084/m9.figshare.30445157.

In the section describing Figures 4 and 6, the following statement has been added to clarify the selection of displayed years and the location of the full dataset: To improve clarity, only the simulation results for the validation period (2010–2013) are displayed, allowing direct comparison of model performance and spatial patterns across scales during the independent validation phase. The complete annual simulation results for the full study period (1998–2013), including all three spatial scales and both update rules, are openly available for download at 10.6084/m9.figshare.30445157.

Both figure captions have been updated to note that only 2010–2013 are displayed for clarity and that the complete dataset is available via the provided data link.
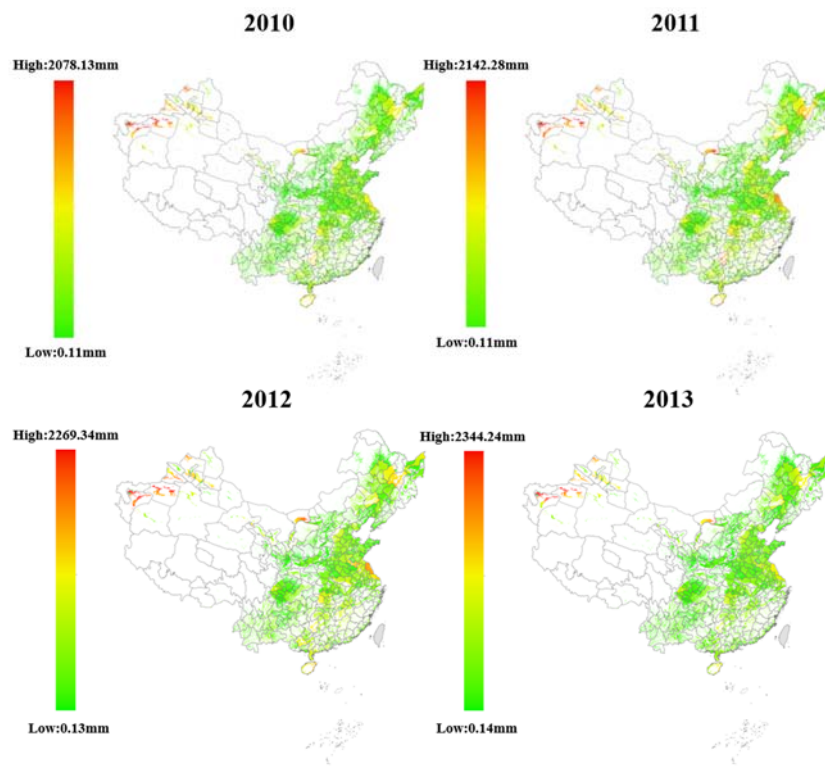
<u>Revised paragraph for Figure 5 and revised Figure 5 (Line 396-412):</u>

Based on the historical water use record from 1998 to 2013, a transition matrix for each grid cell was constructed to represent the probability of transition from its current state to the subsequent year. According to these transition matrices, the interval for the next state can be predicted, and a random sample is drawn from the corresponding probability distribution to obtain the simulated water use value. The probability rule is applied at three spatial scales (i.e., 1 km, appropriate scale, and prefecture scale). To improve clarity, only the simulated water use results for the validation period (2010–2013) are displayed in Figure 5, enabling a direct comparison of spatial patterns across scales during independent validation. The complete annual simulation results for the
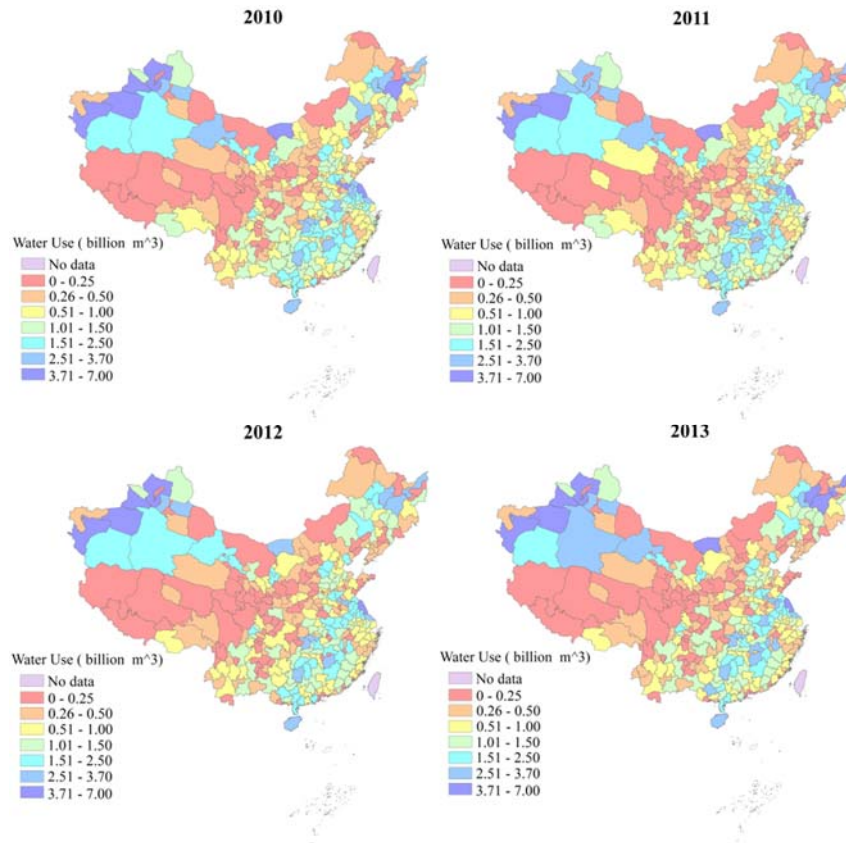
entire study period (1998–2013)—including all three spatial scales and both update rules—are available for download at: 10.6084/m9.figshare.30445157.
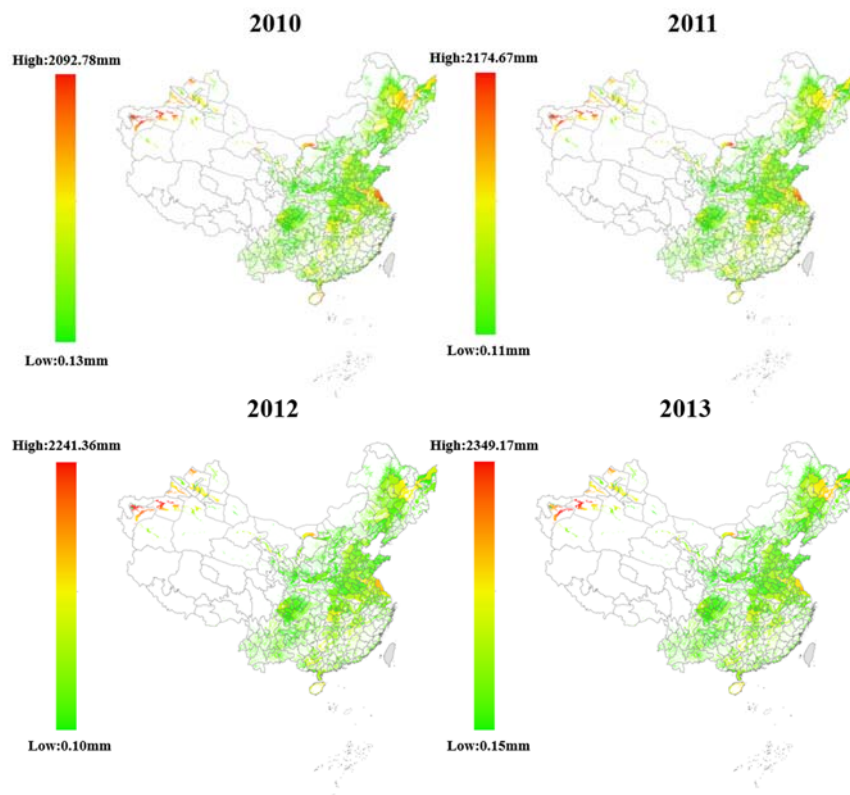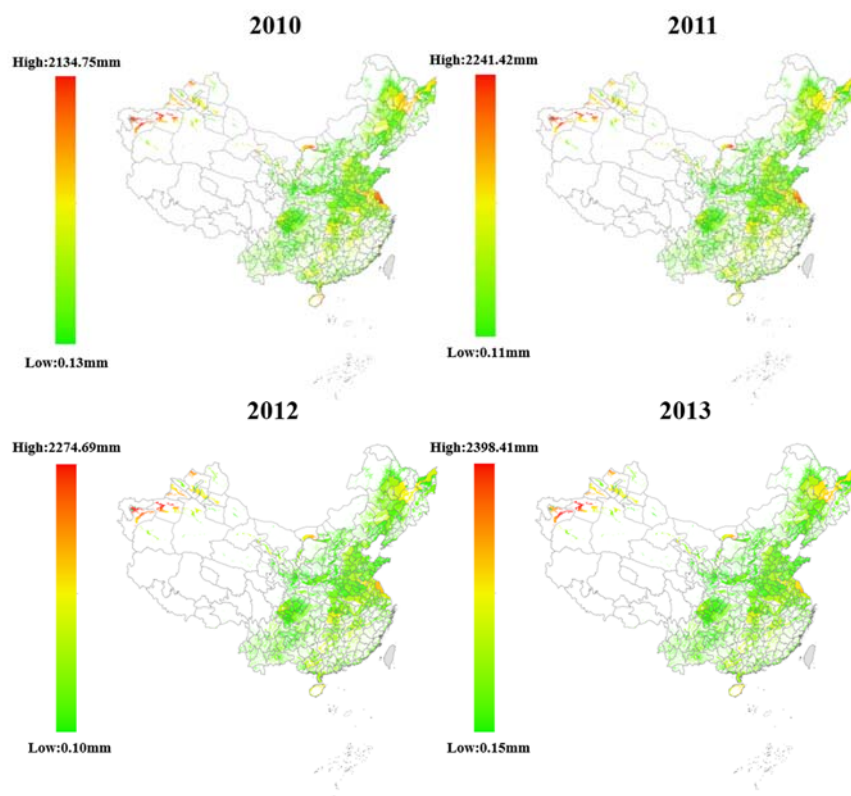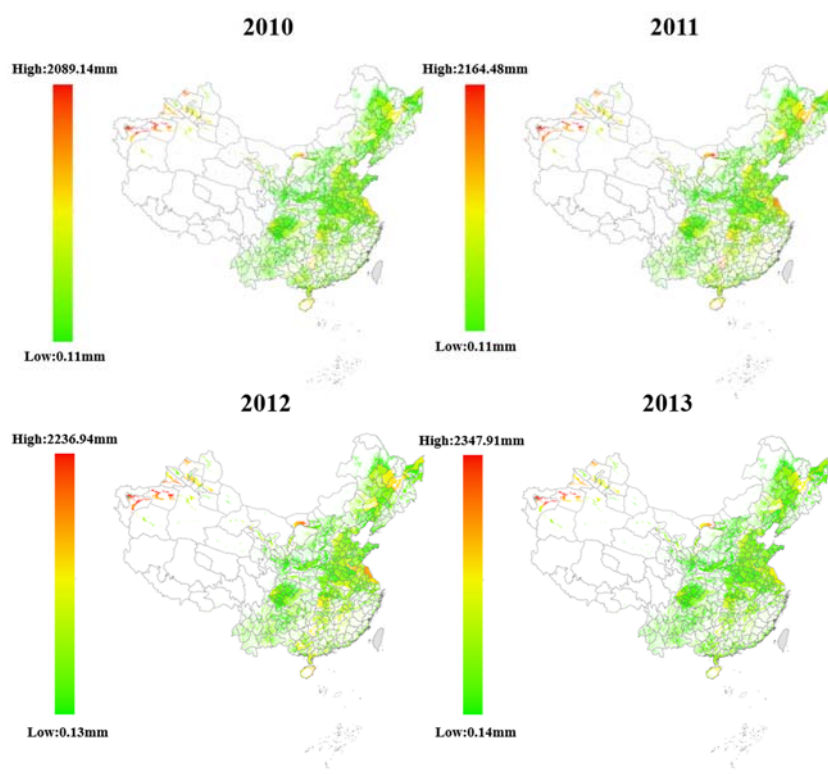

(a)

(b)

(c)

Figure 5 Water use simulation results from the probability rule at three different scales: (a) 1km scale; (b) appropriate spatial scale; (c) prefecture scale

Revised paragraph for Figure 7 and revised Figure 7 (Line 461-473):

After calibrating the parameters for the linear rule of the CA model, the water use grid maps were generated at three spatial scales (i.e., 1 km, appropriate spatial scale, and prefecture scale). For brevity and readability, Figure 7 presents only the simulation results for the validation years (2010–2013), which allow a more straightforward assessment of the model's predictive performance and spatial differences among scales. The full set of simulated water use maps from 1998–2013, can be accessed at: 10.6084/m9.figshare.30445157.

2010   2011

High:2134.75mm   High:2241.42mm

Low:0.13mm   Low:0.11mm

2012   2013

High:2274.69mm   High:2398.41mm

Low:0.10mm   Low:0.15mm

(a)

2010   2011

High:2089.14mm   High:2164.48mm

Low:0.11mm   Low:0.11mm

2012   2013

High:2236.94mm   High:2347.91mm

Low:0.13mm   Low:0.14mm

(b)
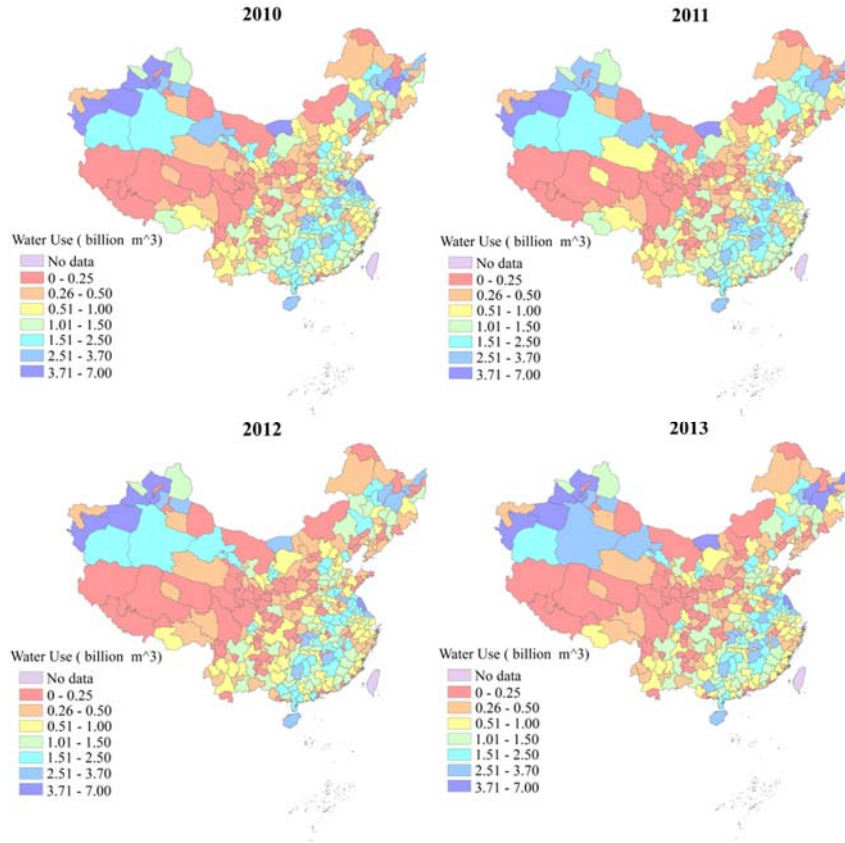
(c)

Figure 7 Water use simulation results from the linear rule at three different scales: (a)
1km scale; (b) appropriate spatial scale; (c) prefecture scale

***COMMENT* 6:**

At the same spatial scale, what is difference between the probability rule and the linear
rule? Can you calculate the difference between Figure 4 and Figure 6, providing more
statistical information?

We appreciate the reviewer's suggestion to provide a more quantitative
comparison between the probability rule and the linear rule. In the revised manuscript,
we have expanded Section 5.1 ("Impacts of Update Rules on Water Use Simulation")
to include additional statistical information describing the differences between the two
rules. Specifically, we introduced a new metric, $\Delta signed$, which represents the mean
signed grid-level difference between the probability-rule and linear-rule simulations,
normalized by the mean simulated water use in each province.

A positive $\Delta signed$ indicates that the probability rule produces higher simulated
water use compared to the linear rule, while a negative value indicates the opposite.
This metric complements RMSE and RE by revealing both the magnitude and direction
of differences between the two update rules.

The updated results are presented in the revised Table 2, which now includes $\Delta signed$ values for each province. These additions allow for a more comprehensive evaluation of the relative performance of the two update mechanisms at comparable spatial scales. The revised Section 5.1 text explicitly discusses these findings and highlights the spatial patterns of differences between the two update rules.

<u>Revised Section 5.1 (Line 602-631):</u>

To assess the performance of the two update rules, the simulation results at three spatial scales were upscaled to the provincial administrative level, and the RMSE and RE metrics were calculated for both the probability and linear rules. An additional indicator, $\Delta signed$ was introduced to quantify both the magnitude and direction of the differences between the two simulations (probability vs. linear). $\Delta signed$ represents the mean signed grid-level difference normalized by the mean simulated water use for each province. A positive $\Delta signed$ indicates that the probability rule yields higher simulated water use than the linear rule, while a negative $\Delta signed$ indicates lower simulated values. There are notable differences between the results produced by the two update rules, as summarized in Table 2.

Table 2. Model performance at the provincial administrative level from different update rules

Provinces	RMSE (billion m3)		RE (%)		$\Delta signed$ (%)
	Linear Rule	Probability Rule	Linear Rule	Probability Rule	
Beijing	0.03	0.05	+12	+19	+2.8
Tianjin	0.02	0.02	-11	-17	-0.6
Shanghai	0.02	0.04	-13	-25	+3.1
Chongqing	0.04	0.07	+16	+17	+1.5
Anhui	0.12	0.17	+23	+32	+2.2
Fujian	0.17	0.21	-23	+37	+3.4
Gansu	0.35	0.43	+27	+29	-1.9
Guangdong	0.26	0.36	+21	+32	+3.9
Guizhou	0.31	0.40	-26	-34	-1.2
Hainan	0.11	0.21	-12	+19	+3.0
Heilongjiang	0.42	0.37	-27	+24	+2.1
Hunan	0.21	0.11	-16	+10	+1.6

Jilin	0.36	0.43	+28	+25	+2.5
Jiangsu	0.12	0.17	-12	-26	-0.7
Jiangxi	0.22	0.26	+24	-31	+1.3
Inner Mongolia	0.19	0.29	-23	-35	-2.1
Qinghai	0.31	0.41	+43	+36	+1.4
Ningxia	0.11	0.20	-31	-41	-2.4
Shandong	0.21	0.36	+29	-36	+4.1
Shanxi	0.42	0.61	+31	+42	+3.3
Shannxi	0.27	0.43	-32	-39	-1.5
Sichuan	0.39	0.41	+31	+46	+2.7
Xizang	0.56	0.61	-62	-79	-0.9
Xinjiang	0.68	0.72	+69	+83	+3.5
Yunnan	0.25	0.34	-37	-47	-1.1
Zhejiang	0.21	0.29	+18	+12	+1.9
Guangxi	0.34	0.42	-32	-39	-1.7
Hubei	0.29	0.39	+27	+26	+2.2
Liaoning	0.35	0.41	+33	+48	+3.7

The linear rule generally exhibits lower *RMSE* and *RE* values, indicating higher stability and consistency with observed prefecture-level statistics. This is because the linear rule updates each grid deterministically based on spatial averages of its neighbors, which smooths fluctuations and captures persistent long-term patterns. Consequently, it tends to reduce local variability and enhance regional stability, especially at coarser spatial scales. By contrast, the probability rule explicitly incorporates stochasticity through distribution-based state transitions. This enables it to capture local irregularities, abrupt changes, and nonlinear water use dynamics driven by variations in industrial structure, irrigation demand, or climatic conditions. However, such stochastic behavior can also amplify uncertainty in regions with sparse data or complex spatial heterogeneity, resulting in slightly higher *RMSE* and *RE* values.

Quantitatively, the mean *RMSE* of the linear rule during the validation period (2010–2013) was 0.28 billion m³, compared with 0.36 billion m³ for the probability rule. The corresponding mean relative errors were ±22.4% and ±29.8%, respectively. At the national scale, the simulated total water use was 570.6 billion m³ for the linear rule and 583.2 billion m³ for the probability rule, differing by +1.1% and +3.5% from observed national statistics. Therefore, the linear rule is identified as the best-performing estimation framework for reproducing the observed spatiotemporal water use

distribution in China, whereas the probability rule provides valuable complementary insights for representing uncertainty and local heterogeneity.

The $\Delta signed$ results indicate that the probability rule tends to simulate higher water use in industrially intensive provinces (e.g., Shandong, Guangdong, and Liaoning) and slightly lower values in water-scarce inland regions (e.g., Ningxia, Gansu, and Inner Mongolia). The results confirm that the two rules emphasize different aspects of the underlying processes—the probability rule better reflects stochastic local variability, while the linear rule offers greater stability and smoother large-scale consistency.

COMMENT **7**:

What is the best estimation of water use simulation in this study, and what is the overall accuracy for China and each province?

RESPONSE to COMMENT **7***:*

We thank the reviewer for this important question. To identify which update rule provides the best estimation of water use in this study, we conducted a comprehensive accuracy evaluation using prefecture-level statistical water use data as reference values. The performance of both update rules—probability rule and linear rule—was assessed in terms of Root Mean Square Error and Relative Error at the provincial level during the validation period (2010–2013).

Overall, the linear update rule demonstrates superior accuracy and stability in most provinces, showing lower *RMSE* and *RE* values compared to the probability rule. At the national scale, the linear rule achieved an average *RMSE* of 0.28 billion m³ and an average *RE* of ±22.4%, whereas the probability rule exhibited a higher average *RMSE* of 0.36 billion m³ and an average *RE* of ±29.8%. This indicates that the linear rule provides a more consistent and reliable estimation when aggregated at broader administrative scales.

However, the probability rule offers certain advantages in specific contexts. It better captures local variability and stochastic changes, particularly in provinces characterized by rapid industrial expansion or heterogeneous land-use patterns (e.g., Shandong, Guangdong, and Liaoning). This makes it valuable for analyzing fine-scale spatiotemporal fluctuations in water use, despite its slightly higher global uncertainty.

Therefore, we conclude that: The linear rule provides the most accurate and stable overall estimation of total water use at both provincial and national levels; the probability rule complements it by representing local variability and nonlinear water use dynamics more effectively. These findings are now explicitly stated in the revised manuscript (Section 5.1).

Added paragraph in Section 5.1 (Line 620-626):

Quantitatively, the mean *RMSE* of the linear rule during the validation period (2010–2013) was 0.28 billion m³, compared with 0.36 billion m³ for the probability rule. The corresponding mean relative errors were ±22.4% and ±29.8%, respectively. At the national scale, the simulated total water use was 570.6 billion m³ for the linear rule and 583.2 billion m³ for the probability rule, differing by +1.1% and +3.5% from observed

national statistics. Therefore, the linear rule is identified as the best-performing estimation framework for reproducing the observed spatiotemporal water use distribution in China, whereas the probability rule provides valuable complementary insights for representing uncertainty and local heterogeneity.

List of changes (consolidated)

Note: Line references correspond to the clean revised manuscript.

Abstract: No changes.

Keywords: Updated to improve discoverability and alignment with the study focus: [water use; spatial scale; cellular automata; multi-scale simulation] (Line: 39).

Introduction: No changes.

Methodology (Section [2.1 &2.2]): Revised for clarity and completeness (Line 137-228)

Study area and datasets: No changes.

Results (Section [4.1]): Revised for clarity; optimized the presentation of the figures and added more detailed result information; links are provided for selected results (Line 330-503)

Discussion (Section [5.1 & 5.3]): Expanded and clarified (Line 602-631 & Line 654-709)

Conclusion: No changes.

References: Added new citations and updated existing records to reflect the revisions.