water

MDPI

Article

Impact of Future Climate Change on Groundwater Storage in China's Large Granary: A Study Based on LSTM and CMIP6 Models

Haiqing Wang [1,2], Peng Qi [1,*], Moran Xu [1,2], Yao Wu [1] and Guangxin Zhang [1]

[1] Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, No. 4888, Shengbei Street, Changchun 130102, China; wanghaiqing@iga.ac.cn (H.W.); xumoran@iga.ac.cn (M.X.); wuyao@iga.ac.cn (Y.W.); zhgx@iga.ac.cn (G.Z.)

[2] Institute of Water Conservancy and Electric Power, Heilongjiang University, Harbin 150080, China

* Correspondence: qipeng@iga.ac.cn

Abstract: Northeast China, as a primary grain-producing region, has long drawn attention for its intensive groundwater extraction for irrigation. However, previous studies on the future spatiotemporal changes of groundwater storage (GWS) are lacking. Utilizing the Global Land Data Assimilation System Version 2.2 (GLDAS-2.2), which simulates groundwater storage (as Equivalent Water Height) using the Catchment Land Surface Model (CLSM-F2.5) and calibrates it with terrestrial water storage data from the GRACE satellite, we analyzed the spatiotemporal variations of GWS in northeast China and employed a Long Short-Term Memory (LSTM) neural network model to quantify the responses of GWS to future climate change. Maintaining current socio–economic factors and combining climate factors from four scenarios (SSP126, SSP245, SSP370, and SSP585) under the CMIP6 model, we predicted GWS from 2022 to 2100. The results indicate that historically, groundwater storage exhibits a decreasing trend in the south and an increasing trend in the north, with a 44° N latitude boundary. Under the four scenarios, the predicted GWS increments in northeast China are 0.08 ± 0.09 mm/yr in SSP126, 0.11 ± 0.08 mm/yr in SSP245, 0.12 ± 0.09 mm/yr in SSP370, and 0.20 ± 0.07 mm/yr in SSP585. Although overall groundwater storage has slightly increased and the model projections indicate a continued increase, the southern part of the region may not return to past levels and faces water stress risks. This study provides an important reference for the development of sustainable groundwater management strategies.

Keywords: groundwater storage; LSTM method; CMIP6; climate change

check for updates

Academic Editors: Chin H Wu and Fernando António Leal Pacheco

Received: 1 October 2024
Revised: 6 January 2025
Accepted: 14 January 2025
Published: 23 January 2025

Citation: Wang, H.; Qi, P.; Xu, M.; Wu, Y.; Zhang, G. Impact of Future Climate Change on Groundwater Storage in China's Large Granary: A Study Based on LSTM and CMIP6 Models. *Water* **2025**, *17*, 315. https://doi.org/10.3390/w17030315

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Groundwater is a significant source of freshwater for humans and ecosystems [1]. It accounts for one-third of the world's total freshwater consumption, and 42%, 36%, and 27% of irrigation, domestic, and manufacturing water use, respectively [2]. The increases in climate change and economic development present significant threats. Model-based projections by [3] suggest that by 2050, water stress will increase in areas where the ratio of water withdrawals to long-term annual average water resources is greater than 0.4. Water stress in areas with decreasing water resources is projected to increase by a factor of two to four. Changes in groundwater storage (GWS) directly affect regional water resources, food production, and human production and life [4]. There is a degree of uncertainty in how GWS will respond to climate change [5], so it is necessary to conduct in-depth research.

Groundwater exploration is primarily accomplished through the measurement of groundwater levels (piezometers, dip meters, and wells), the assessment of groundwater quality (isotopes, carbon-14, and chloride mass balance), geochemical sampling, and seepage measurements [6]. When conducting large-scale analyses, GWS data often lack comprehensiveness and universality due to geographical and economic limitations. NASA's Gravity Recovery and Climate Experiment (GRACE) satellite and its applicability to climate change issues have provided us the opportunity to study possible changing trends in groundwater systems [7]. It serves as an important reference for the spatiotemporal evolution of groundwater resources and water resource management [8]. Groundwater research based on the GRACE satellite has been widely used to characterize aquifer depletion in vast regions, large drainage areas, and transboundary aquifers. It can retrieve groundwater-related parameters, such as specific yield, and develop new indices to identify aquifer depletion and GW stress [9,10]. The Catchment Land Surface Model (CLSM) within GLDAS-2.2 integrates observations of terrestrial water storage variations from the GRACE satellites (GRACE-TWS) to enhance global groundwater storage (GWS) modeling [11]. The GWS values, represented in millimeters (mm), reflect the variations in groundwater storage as an equivalent depth of water accumulated or depleted over a given area. This metric allows for the assessment of spatiotemporal changes in groundwater resources, providing crucial insights into the sustainability of water supplies, especially in regions dependent on groundwater for irrigation and other uses. Therefore, by assimilating TWS-GRACE data, GLDAS-2.2 can effectively represent GWS dynamics [12].

Although GRACE-based groundwater research has received widespread attention, there are still relatively few studies on large-scale predictive analysis [13]. Existing predictive analyses often use hydrological models, such as MODFLOW, Land Surface Models, and integrated hydrological models (IHMs). These models require large amounts of remote sensing data and hydrological parameters [14,15], causing a scarcity of predictive research. Deep learning (DL) methods have recently seen success and are becoming more widely applied. Specifically, artificial neural networks (ANN) have been successfully used in the research of various surface water [16] and groundwater-level problems [17–20]. Among deep learning methods, Long Short-Term Memory (LSTM) networks stand out for their ability to capture long-term dependencies and nonlinear relationships in time-series data, making them especially suitable for groundwater prediction. Unlike traditional models, LSTM can efficiently process complex temporal patterns without requiring extensive hydrological parameters, reducing data dependency while maintaining high prediction accuracy [21].

Northeastern China, as one of China's main grain-producing areas, plays a crucial role in guaranteeing national food security and the supply of agricultural products [22]. However, with the continuous development of the economy and society, the utilization and protection of groundwater resources have become challenging. The issue of groundwater overexploitation is particularly prominent, especially in the main irrigation areas in the eastern part of northeast China [23]. Limited by natural and social factors, the demand and emphasis on groundwater research are relatively low. Most existing research focuses on spatiotemporal variation [24], the evaluation and protection of groundwater resources [25], and groundwater response to water resource allocation and other driving factors [26,27]. There is a lack of systematic prediction for the trend of groundwater reserves in northeast China [28].

Despite the growing concern about groundwater storage (GWS) changes in northeast China, there remains a lack of comprehensive studies addressing both historical trends and future projections under diverse climate scenarios. This study aims to fill these gaps by integrating GWS data from GLDAS-2.2 with an improved Long Short-Term Memory (LSTM) model to forecast GWS changes from 2022 to 2100. The analysis incorporates four

CMIP6 scenarios—SSP126, SSP245, SSP370, and SSP585—to capture the potential variability and uncertainties in future climate impacts. By systematically analyzing historical trends, identifying influencing factors, and simulating future scenarios, this study provides a robust framework for understanding GWS dynamics in the region, thereby contributing to the growing body of research on groundwater responses to climate change.

2. Materials and Methods

2.1. Study Area

The focus area of this study is the black soil region in northeast China, which covers the three provinces of northeast China and the four eastern leagues of the Inner Mongolia Autonomous Region (including the cities of Chifeng, Tongliao, Xing'an, and Hulunbeier). The geographic coordinates range from latitude 41°20′ to 53°40′ N and longitude 116°25′ to 135°05′ E (Figure 1). The region is one of the most important grain-producing areas in China, known for its vast plains and black soil cultivated land. It is relatively flat with a variety of landform types, including plains, mountains, and hills. Typical black-soil arable land is found in most parts of the region, especially in the Liaoning, Heilongjiang, and Jilin provinces. This fertile black soil is known as the "Giant Panda of Cultivated Land", and it provides unique conditions for the growth of crops.

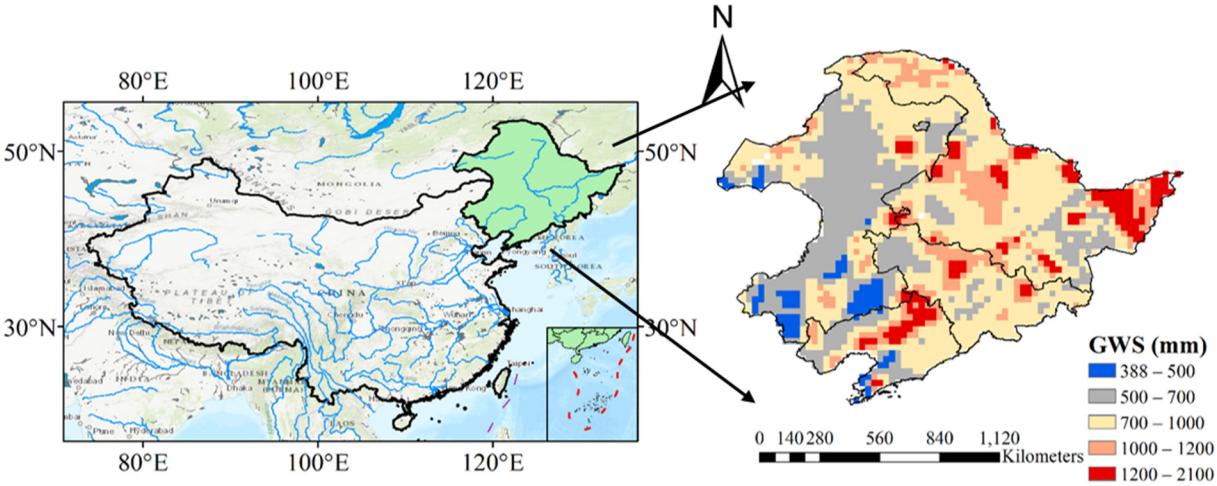


Figure 1. Spatial distribution of mean GWS (as Equivalent Water Height, representing the spatial pattern of groundwater storage) from February 2003 to December 2022 based on GLDAS-2.2 [29].

Agriculture in the northeast is of great importance to the entirety of China. With a cultivated area of about 41.42 million hectares, 1973 million hectares of which are cultivated on black soil, the northeast is a base for the production and processing of agricultural yields in China. The groundwater supply in the three northeastern provinces accounts for 40–45% of the total water supply. The groundwater supplies in the four eastern provinces of Inner Mongolia account for about 65% of the total water supply. In the Heilongjiang province, farmland irrigation expends more than 85% of the total water supply. In Jilin province, farmland irrigation accounts for more than 65% of the total water supply (data from the Water Resources Bulletin).

In northeast China, groundwater storage (GWS) varies significantly across different regions. The plain areas, including the Songnen, Sanjiang, and Liaohe Plains, account for 28.45% of the region's total area. These plains, located at river confluence zones, are rich in groundwater resources despite a long-term declining trend in storage since the Mesozoic era. In these areas, GWS often reflects substantial reserves, indicative of their role as critical groundwater recharge zones. In contrast, the hilly regions, predominantly

composed of Jurassic igneous rocks, clastic rocks, and carbonate rocks, are characterized by well-developed joints and fractures that facilitate groundwater recharge. However, in some areas, such as in the southern part of Liaoning Province, the limited water retention capacity of carbonate rock aquifers leads to lower groundwater availability. The Inner Mongolia region, with its weak hydrological cycle and minimal river flows, presents a stark contrast. Groundwater resources here are scarce and predominantly confined to shallow or localized aquifers, reflecting the challenges of water resource management in this area.

2.2. Data Sources

(1) Data selection

In this study, the GWS from GLDAS-2.2 was selected as the target variable. This dataset has been calibrated using TWS data from GRACE. The validation results show that the NSE between GLDAS and GRACE (CSR GRACE/GRACE-FO RL06.3 Mascon Solutions) for the average GWSA in northeast China is 0.758. These results indicate a good agreement between the GLDAS and GRACE data, suggesting that the GLDAS dataset is reliable for representing groundwater storage changes in the region.

There are many factors affecting GWS that have been explored in detail by previous scholars [30]. In this paper, six significant influencing factors were selected as input data to explore their specific interactions with GWS.

Climatic factors include precipitation [31], potential evapotranspiration [32], and average temperature [33]. These three climatic factors have been verified to possess strong correlations with GWS. Additionally, socio–economic factors such as GDP per capita [27], cultivated area [34], and groundwater supply volume (GWSV) [35] are considered. GDP per capita is usually related to water resource utilization efficiency. Higher GDP may promote improved water resource management. Cultivated land area directly affects water demand. Increased cultivated land may increase groundwater extraction. Groundwater supply volume (GWSV) directly reflects the availability of groundwater resources in the region. The consideration of these factors is intended to improve the accuracy of the model and ultimately enhance the reliability of the prediction results in northeast China.

Traditional hydrological methods, such as the water balance equation, typically use variables like precipitation, evapotranspiration, and groundwater extraction as inputs. The water balance approach focuses on quantifying the change in groundwater storage based on the difference between incoming and outgoing fluxes. In contrast, the LSTM model used in this study incorporates not only climatic factors but also socio–economic factors, such as GDP and groundwater supply, which are often not included in traditional methods. The inclusion of socio–economic factors allows for a more comprehensive understanding of groundwater dynamics, especially considering human-induced factors such as groundwater extraction and water management practices.

The LSTM (Long Short-Term Memory) model, unlike traditional methods such as the water balance equation, is a data-driven approach that is capable of capturing complex, non-linear relationships in time-series data. In this study, the LSTM model processes multiple input variables, including climatic and socio–economic factors, to predict groundwater storage (GWS). The model takes into account temporal dependencies in the data, which allows it to learn patterns and trends from historical observations. Specifically, the LSTM model is trained on past values of input variables to forecast future groundwater storage changes.

While the water balance equation typically calculates groundwater storage change as the difference between fluxes over time (i.e., the difference between inflows like precipitation and outflows like evapotranspiration and groundwater extraction), the LSTM model does not explicitly compute these fluxes. Instead, it relies on the relationship between input

variables and groundwater storage, learning directly from the data. The LSTM model's ability to learn from complex interactions between input variables and historical patterns makes it a powerful tool for predicting groundwater storage dynamics, offering flexibility beyond the assumptions of traditional methods.

The consideration of these factors in the LSTM model aims to improve the accuracy of the predictions, ultimately enhancing the reliability of the groundwater storage change predictions.

(2) Data Availability

The groundwater storage (GWS) dataset used in this study is GLDAS Catchment Land Surface Model L4 daily 0.25° × 0.25° GRACE-DA1 V2.2 (https://doi.org/10.5067/SZYMDFZ7TJ4H) [29] (Figure 2). The precipitation, potential evapotranspiration, and temperature datasets for climate influences were obtained from the Tibetan Plateau Science Data Center. These datasets, compiled and organized by Shouzhang Peng, cover monthly data from 1901 to 2022 [36–38]. Gross domestic production (GDP) datasets were made available by provincial statistical yearbooks (Heilongjiang Province: http://tjj.hlj.gov.cn; Jilin Province: http://tjj.jl.gov.cn; Liaoning Province: https://www.ln.gov.cn; Inner Mongolia Autonomous Region: http://tj.nmg.gov.cn, accessed on 10 October 2023), while groundwater supply data was obtained from provincial water resource bulletins and provincial water authorities (Heilongjiang Province: http://slt.hlj.gov.cn; Jilin Province: http://slt.jl.gov.cn; Liaoning Province: https://slt.ln.gov.cn/; Inner Mongolia Autonomous Region: http://slt.nmg.gov.cn, accessed on 10 October 2023). Cultivated land area data were gathered from "IrriMap_CN: Improved annual irrigation maps across China in 2000–2019 based on satellite imagery and machine-learning method" by Zhang et al. (2022) [39], available on Figshare (https://figshare.com/, accessed on 10 October 2023). Future climate data (FCD) select data from four scenarios (SSP126, SSP245, SSP370, and SSP585) under the seven institutions in the Coupled Model Intercomparison Project Phase 6 (CMIP6) model. CMIP6 is a global climate modeling comparison project aimed at facilitating comparison and cooperation among global climate models [40]. The data can be downloaded from https://esgf-data.dkrz.de/search/cmip6-dkrz/ (accessed on 10 October 2023).

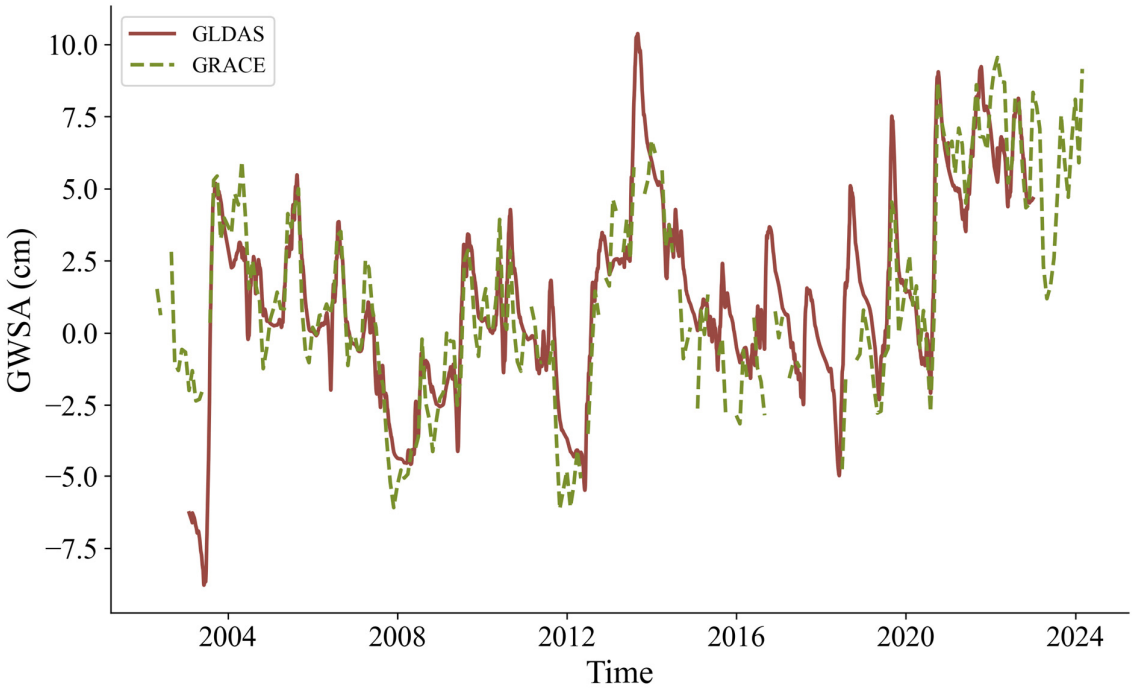


Figure 2. Comparison of GLDAS and CSR GRACE GWSA Data for northeast China, with a NSE of 0.758.

2.3. Methods

(1) Interpolation

The dataset used in this article has different resolutions to obtain datasets at the same points. Specifically, we use the bilinear interpolation method. Based on the groundwater storage dataset with a resolution of 0.25° × 0.25°, the longitude and latitude of each point in the dataset were interpolated to other datasets to obtain the data values corresponding to the same longitude and latitude.

$$\text{Value} = \frac{1}{(x_2 - x_1) \times (y_2 - y_1)} \times [A \times (x_2 - x) \times (y_2 - y) + B \times (x - x_1) \times (y_2 - y) + C \times (x_2 - x) \times (y - y_1) + D \times (x - x_1) \times (y - y_1)] \quad (1)$$

where A, B, C, and D represent the values of four data points, respectively. (x, y) are the coordinates of the target point. (x_1, y_1), (x_1, y_2), (x_2, y_1), and (x_2, y_2) are the coordinates of the four data points. (x_1, y_1) and (x_2, y_2) define a rectangular area containing the four data points.

This formula uses the weight of the target point between four data points and estimates the value of the target point by taking a weighted average of the data point values.

When making model predictions, GDP data and GWSV data were based on the annual scale of prefecture-level cities. The data extraction method corresponding to each grid point found the data of the prefecture-level city corresponding to the point and repeated the corresponding data the corresponding number of times.

(2) CMIP6 data processing

Seven global climate models (GCMs) (ACCESS-ESM1-5, BCC-CSM2-MR, CanESM5, CMCC-ESM2, CNRM-CM6-1, CNRM-ESM2-1, and INM-CM4-8) of CMIP6 were selected for future climate data. Precipitation, temperature, wind speed, and relative humidity were selected for the climate data types. Potential evapotranspiration was calculated according to the Penman–Monteith formula [41]:

$$PET = \frac{0.408 \times \Delta \times (R_s - R_{s_min} - G) + \gamma \times \frac{900}{T + 273} \times U \times (e_s - e_a)}{\Delta + \gamma \times (1 + 0.34 \times U)} \quad (2)$$

PET is potential evapotranspiration. Δ represents the slope of the saturated water vapor pressure curve (kPa/°C), $\Delta = (4098 \times e_s)/(T + 237.3)^2$. *G*, the soil heat flux density (MJ m^{-2} day^{-1}), is negligible. e_s is the saturated water vapor pressure (kPa) and *T* is the air temperature (degrees Celsius). R_s is the downward shortwave radiation (solar radiation). R_{s_min} is the minimum value of shortwave radiation in potential evapotranspiration (usually set to 0.02 × *Rs*). γ is the psychrometric constant (kPa/°C), $\gamma = (0.00163 \times (1 + 0.022 \times T))/0.408$. *U* is the wind speed (m/s). e_s is saturated water vapor pressure (kPa), $e_s = 0.6108 \times e\hat{}((17.27 \times T)/(T + 237.3))$.

e_a is the actual water vapor pressure (kPa), $e_a = (relative\ humidity/100) \times e_s$, where the relative humidity is in percent.

To improve the accuracy of the data, this study adopted the Delta method [42]. The historical data were selected for the time period of 1950–2014 to compare and correct the current data (1950–2014), which was then used to obtain the climate data for the future period of 2015–2050.

$$P_{fut}(m) = P_{obs}(m) \times \frac{\overline{P}_{CMIP6,fut}(m)}{\overline{P}_{CMIP6,his}(m)} \quad (3)$$

$$T_{fut}(m) = T_{obs}(m) + [\overline{T}_{CMIP6,fut}(m) - \overline{T}_{CMIP6,his}(m)] \quad (4)$$

where m stands for month, $P_{obs}(m)$ and $T_{obs}(m)$ are the current status data and different month data, $\overline{P}_{CMIP6,fut}(m)$ and $\overline{T}_{CMIP6,fut}(m)$ are the mean data for different future months

in CMIP6 mode, $\overline{P}_{CMIP6,his}(m)$ and $\overline{T}_{CMIP6,his}(m)$ are the historical monthly average data in CMIP6 mode, and $P_{fut}(m)$ and $T_{fut}(m)$ represent the bias-corrected data. The temperature correction was made using Equation (4), and the precipitation and potential evapotranspiration data correction was made using Equation (3). The corrected historical data were analyzed to obtain the mean value for the whole area. Then, they were compared with the current data to obtain the correlation analysis plot (Figure 3). The correlation coefficient between the corrected precipitation data and the current data stands at 0.71. The fitting equation, $y = 0.94x + 3.21$, indicates a robust fit with a relatively high correlation. This suggests a substantial improvement in precipitation data accuracy, and the corrected potential evapotranspiration and temperature data exhibit exceptional correlation with the current data, surpassing 0.97 in correlation coefficients and slopes. The appropriate equations, coupled with coefficients exceeding 0.97, signify a highly accurate fit. The discreteness of precipitation data is mainly evident during heavy rainfall events, which are relatively rare. However, the differences are smaller during periods of light precipitation. Due to the discretization of model space and time, temperature and potential evapotranspiration are discrete. These results underscore the effectiveness of the calibration process, providing enhanced reliability in potential evapotranspiration and temperature assessments.

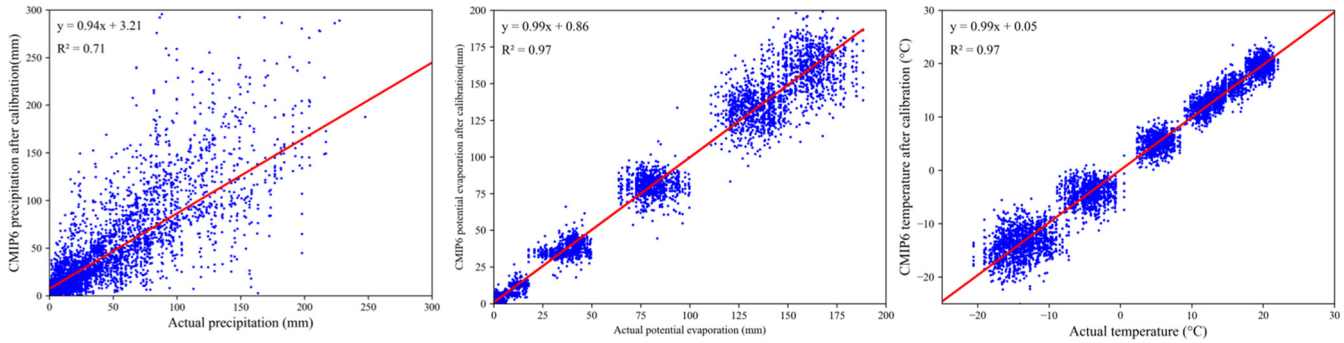


Figure 3. Plot of correlation analysis between current and corrected data for climate influences from 1950 to 2014, from left to right, for precipitation, potential evapotranspiration, and temperature.

(3) Pearson Correlation Coefficient

The Pearson correlation coefficient was calculated to determine the linear relationship between GWS and other influencing factors. The formula used for the calculation is as follows:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \quad (5)$$

r represents the Pearson correlation coefficient.
n is the number of observations.
x and y are the observed values of the two variables being analyzed.
$\sum xy$ is the sum of the product of x and y.
$\sum x$ and $\sum y$ are the sums of the respective variables.
$\sum x^2$ and $\sum y^2$ are the sums of the squares of the respective variables.

(4) Standard Deviation

Standard deviation was calculated to evaluate the dispersion of GWS data. The formulas used for calculation are as follows:

$$s = \sqrt{\frac{1}{n-1}\sum_{n=1}^{n}(x_i - \overline{x})^2} \quad (6)$$

where s is the standard deviation, n is the number of samples, x_i denotes each observed value, n is the number of samples, and $\overline{x}$ is the sample means.

(5) LSTM model construction

Long Short-Term Memory (LSTM) networks are recurrent neural networks that are frequently used for modeling sequential data, such as time series or natural language, due to their advantages with long-term dependencies, strong stackability and scalability, and capability to mitigate gradient vanishing [21]. The LSTM model is a recurrent neural network with three gate units in a memory block (Figure 4). These units are input gates, output gates, and forgetting gates. Under the RNN structure, the valves from each layer are added to the LSTM. Whether the memory state output result of the layer reaches the threshold or not, these valves are opened or closed, and the results are combined to calculate the current layer. Each valve uses the activation function to compute the memory state as input [43].

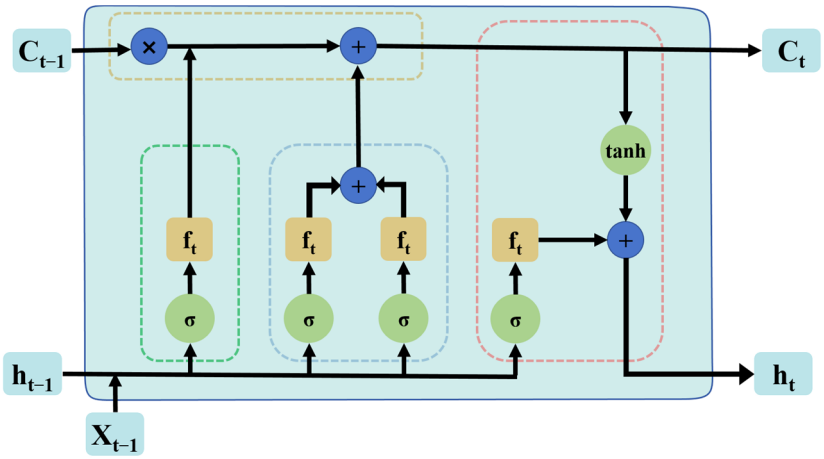


Figure 4. LSTM network structure.

Before we created a LSTM model, the data needed to be processed. This included data normalization, noise processing, merging all feature data by time elements (predictive feature data (2003.02–2019.12) + future feature data (2020.01–2100.12)), filling in the missing values, and normalizing and constructing the time-series inputs and target sequence. The three socio–economic factors used in this study are on an annual scale, and they have been divided into 12 parts to convert them into monthly data. The cultivated land area is calculated by aggregating the total area of each grid cell. For missing values, since the proportion of missing data is relatively small in the time scale, we directly used an artificial neural network to impute them.

In this study, the TensorFlow library was used to create the LSTM model by first defining a sequential layer to add neural network layers. Then, LSTM layers were added to define the LSTM hidden layer, the number of neurons in the input layer, and the shape of the inputs. Then, a fully connected layer was added to determine the number of output layers.

After all the layers had been added to the model, the model needed to be compiled to specify the loss function, optimizer, and performance metrics. Here, the mean square error (MSE) was used as the loss function for the model, and the Adam optimization algorithm was used to train the LSTM model.

The model was then trained using the current state data (2003.02–2019.12), followed by the collated future data. The obtained data were back-normalized to obtain future groundwater storage data (2021.05–2100.12).

Nash–Sutcliffe Efficiency (NSE) and root–mean–square error (RMSE) were chosen to evaluate the merit of the model fitting.

$$NES = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \quad (7)$$

$$RMSE = \sqrt[2]{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (8)$$

where n is the sample size, y_i is the actual observed value, $\hat{y}_i$ is the corresponding predicted value, and $\overline{y}$ is the mean of the observations.

The value of NSE ranges from negative infinity to 1. The closer the NSE is to 1, the better the model fits the observations. A NSE of 0 indicates that the model's performance is comparable to that of the simple average, while a negative value indicates that the model's performance is poor. The smaller the RMSE, the better the model fits the observations. The unit of RMSE is the same as the unit of the observations, so it can be intuitively understood as the average deviation between the predicted and actual values.

3. Results

3.1. Characteristics and Drivers of the Spatiotemporal Evolution of GWS

3.1.1. Analysis of the Spatiotemporal Evolution of GWS

GWS data from GLDAS were analyzed from 2004 to 2022. The annual average GWS data obtained for each raster point were plotted at a distance to produce a spatial distribution of annual GWS changes in the northeast from 2004 to 2022 (Figure 5). The GWS of northeast China is divided, with 2013 as the cut-off time point. The northern region has experienced deficits for 13 consecutive years, with the largest deficit being more than 60 mm in 2008, influenced by the 2007 and 2008 droughts. The worst deficits were mainly located in the central part of the northeast China, with less fluctuation in the southern part. In 2013, the first peak in the GWS was caused by precipitation. In the following years, the GWS gradually increased in the northern region, especially in the central part of the northeast. The southern region was affected by drought and continued to experience a deficit, which reached its maximum in 2015. This deficit state gradually spread to the northern region until abundant precipitation in 2021 alleviated it in the northeast. The effects of drought in the northern region continued until 2021, when the GWS was replenished by abundant precipitation in the northeast, and a second peak occurred.

The GWS data from February 2003 to December 2022 in northeast China were linearly fitted to obtain trend slopes and corresponding p-values for each raster point (Figure 6). The trend of the GWS in northeast China appears to follow the approximate boundary of 44° N latitude. It exhibits a decreasing trend south of 44° N latitude and an increasing trend north of it. The trend was considered insignificant for p-values greater than 0.05. Overall, the percentage of the GWS in northeast China showed a decreasing trend of 29.96% and an increasing trend of 70.04%. In the middle of the northern region, a significant increasing trend of more than 6 mm/year was observed. Conversely, the southernmost part of the southern region showed a decreasing trend, mostly below 2 mm/year, specifically in the southwestern region where the decreasing trend was more than 3 mm/year.

Concerning monthly changes, the GWS in northeast China exhibited a discernible seasonal pattern. Annually, the GWS was at its lowest in April and May, gradually increasing until reaching its peak in August and September (Figure 7). This cyclical change was primarily influenced by the significant impact of the temperate monsoon climate and anthropogenic activities. There are distinctions between the Sanjiang Plain (eastern part of Heilongjiang Province) and other regions in this cycle, which are directly related to the

Sanjiang Plain's use of GWS for irrigating land and improving crop growth cycles. Monthly seasonal changes in GWS were notably significant in the southeastern and northern regions. These regions boasted an average monthly peak difference in GWS of 80 mm, which is more than double the average monthly peak difference in other regions. This was mainly attributed to anthropogenic activities and the quantity of groundwater resources.

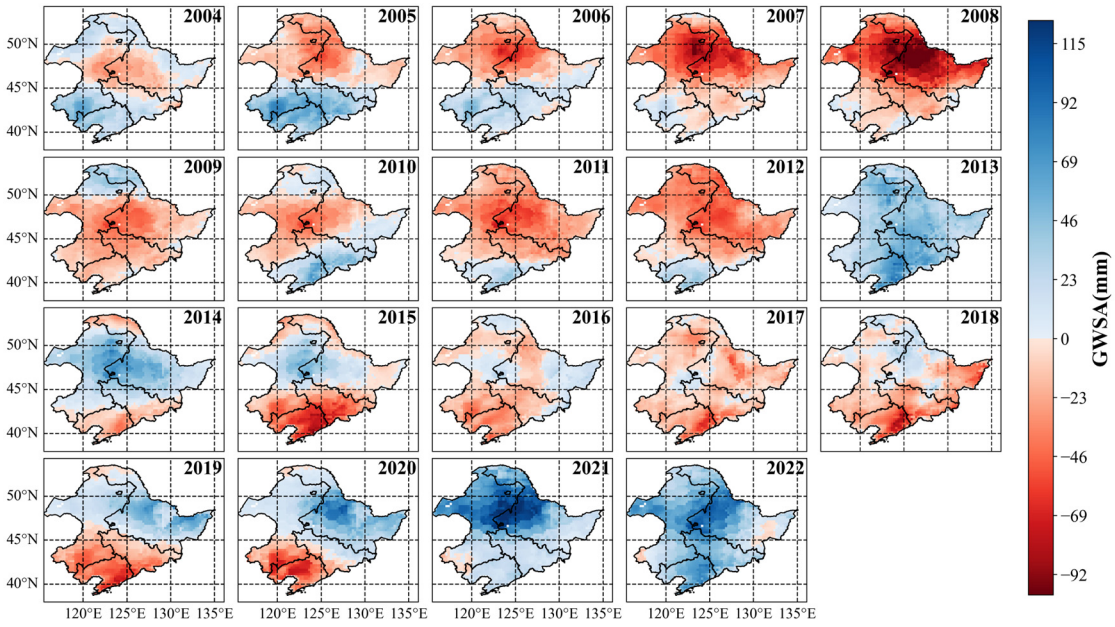


Figure 5. Spatial distributions of GWS interannual spacing values in the northeast China, 2004–2022.

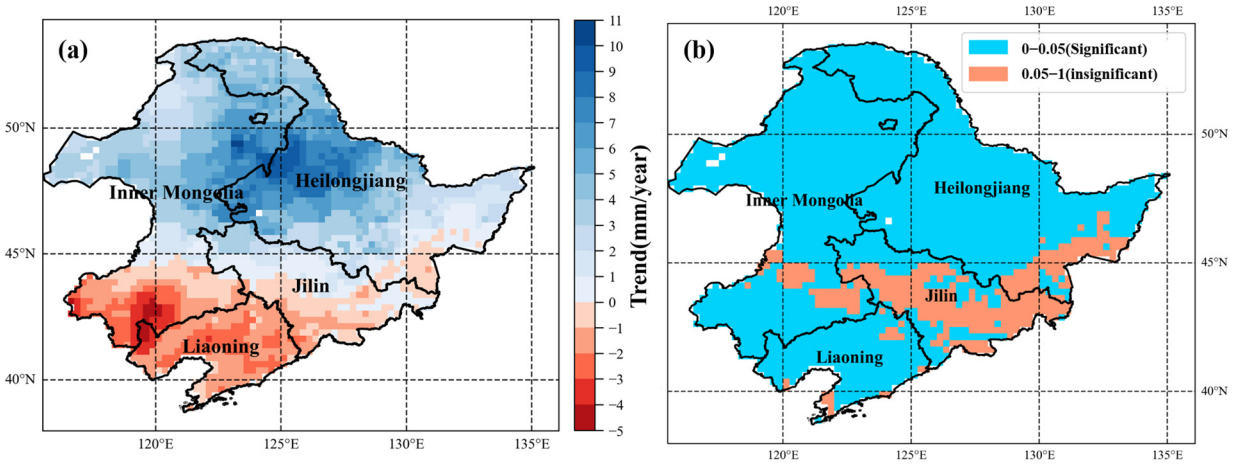


Figure 6. Spatial trend of the GWS in northeast China, 2003.02–2022.12. (**a**) The spatial trend of the GWS in northeast China, (**b**) spatial distribution of *p*-value corresponding to the trend.

Regarding monthly seasonal trends, the changes aligned with the overall trend. Using 45° N as the dividing line, the southern region exhibited a decreasing trend while the northern region showed an increasing trend (Figure 8). Overall, northeast China showed a slight upward trend, with growth rates from May to August being relatively slow (below 2 mm/year). The lowest growth rate was 1.22 mm/year in August, while more pronounced increases occurred in January–April and September–December, peaking in October at 3 mm/year. The area exhibited a decreasing trend in GWS, reduced from September to December, with the lowest value in October. This is predominantly related to the lagging effect of anthropogenic activities and precipitation on GWS.

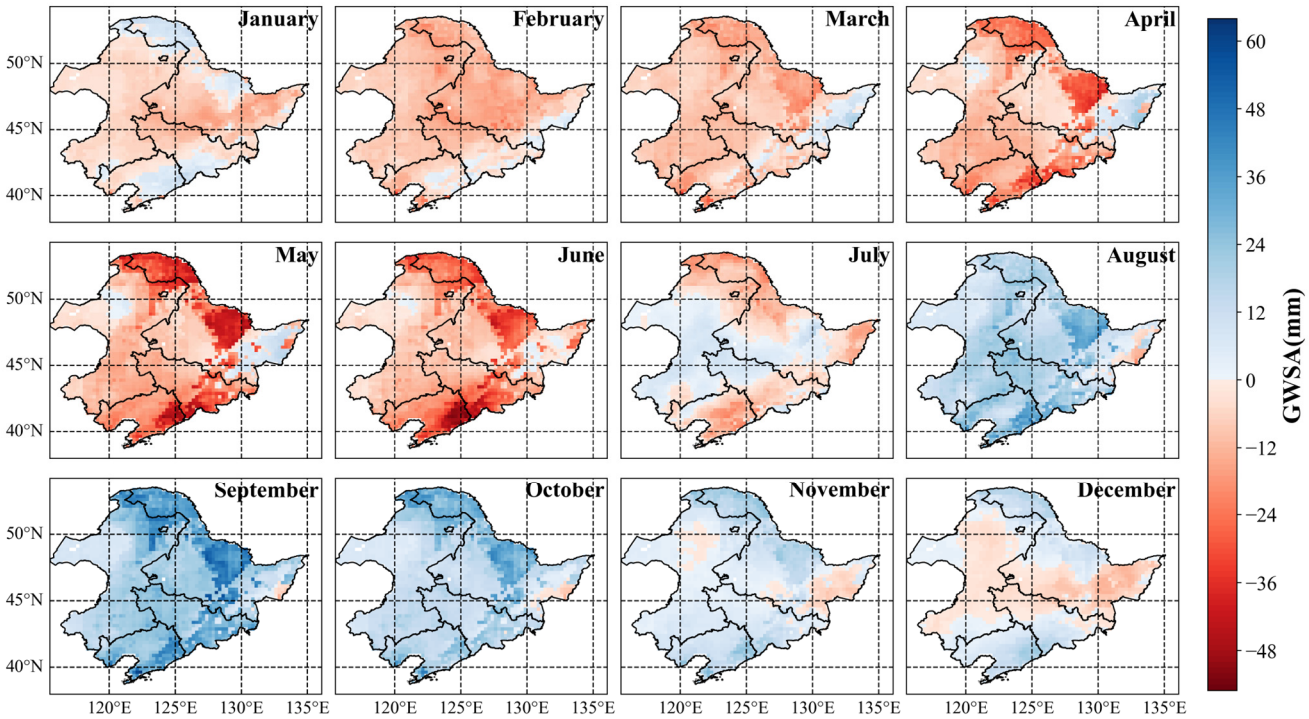


Figure 7. Spatial distribution of monthly-scale groundwater storage anomaly (GWSA) in northeast China.

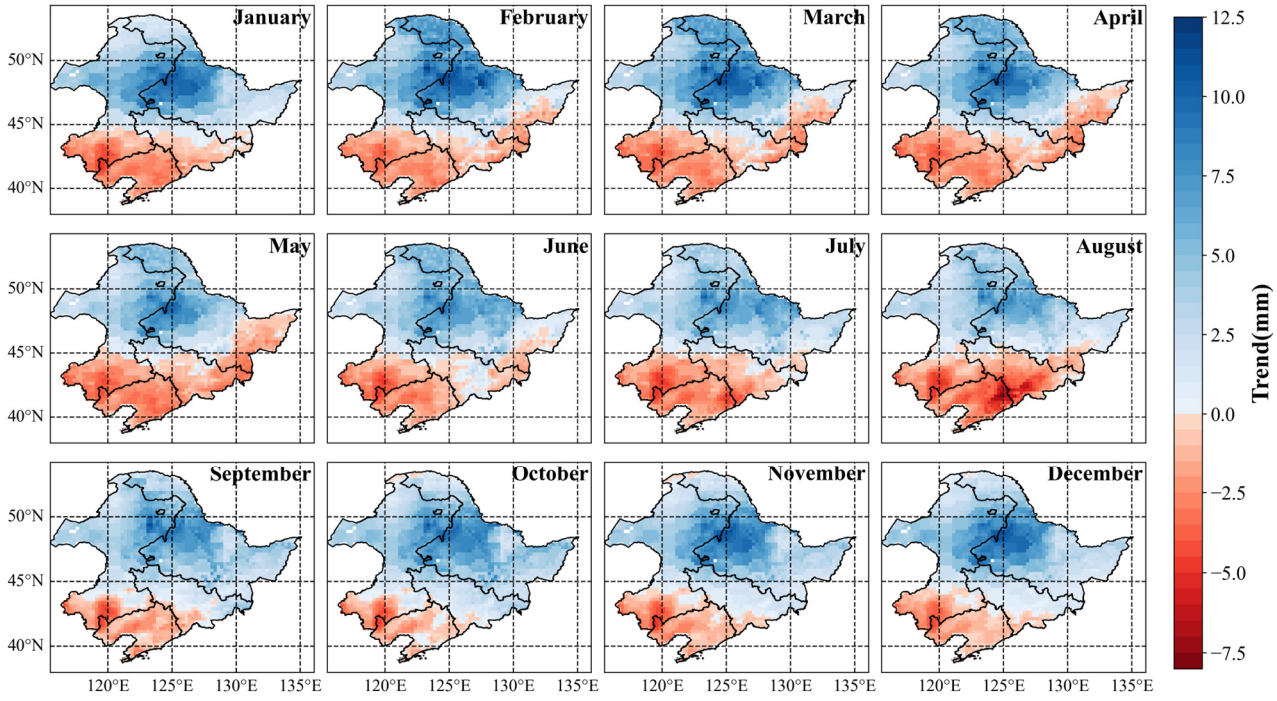


Figure 8. Spatial distributions of monthly trends in GWS in northeast China.

3.1.2. Driver Analysis

The correlations between the six driving factors and GWS were analyzed using Pearson's correlation coefficients, as shown in Figure 9. Overall, climatic factors had a strong influence on the GWS (collection > 0.5) [44], with precipitation being dominant. The influence of socio–economic factors on the GWS was more difficult to analyze, and the influence of anthropogenic factors on the GWS varied in different regions. There were only a few regions where anthropogenic factors had a greater influence on the GWS, and

those regions were primarily in the southwest of northeast China and the central part of Heilongjiang Province.

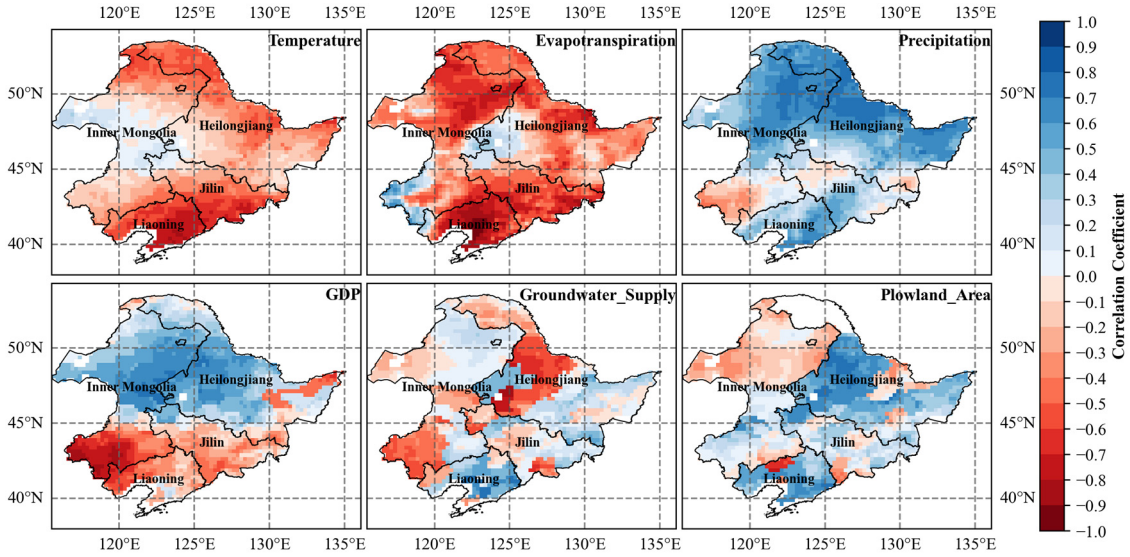


Figure 9. Spatial distribution of correlation coefficients between GWS and corresponding influencing factors in northeast China (inter-annual data), with no arable land in the Daxinganling region.

Potential evapotranspiration and temperature were negatively correlated with groundwater storage (GWS), with correlation coefficients consistently lower than −0.5 in most areas. Precipitation was usually positively correlated with GWS, with correlation coefficients exceeding 0.6 in most areas. The rise and fall of groundwater reserves seen in Figure 5 were closely correlated with GWS, and GWS reached peaks in the northeast during the years with abundant water (e.g., 2013, 2021, 2022). In northeast China, the corresponding GWS reached its peak in the years of abundant water (e.g., 2013, 2021, 2022), and the corresponding GWS reached its low point in the years of depletion (e.g., 2007, 2011, 2017) (data from the provincial water resources bulletin). The influence of socio–economic factors was higher in the southwestern part of northeast China and the central part of Heilongjiang Province. The hydrological recharge in the southwestern region was less affected by climatic factors. The region, characterized by fewer surface water resources, predominantly relies on groundwater, making it more susceptible to anthropogenic activities impacting groundwater reserves. Heilongjiang Province's central region, known for its heavy industrial activity and high groundwater consumption, presented unique challenges. The convergence of factors, including relatively small groundwater reserves (~600 mm) and insufficient recharge, exacerbates the impact of anthropogenic activities on groundwater reserves.

Generally speaking, the impact of human activities on groundwater storage is greater in areas with low groundwater storage and inadequate recharge. This is evident in the monthly seasonal variations, particularly in areas with large and irregular fluctuations, which are more susceptible to the influence of human activities. Areas with decreasing trends may be overexploited and areas with increasing trends may benefit from better management and control. This suggests the need for the integrated management of groundwater resources.

3.2. Characterization of Groundwater Storage Response to Future Climate Change

3.2.1. Model Fitting Performance

In this study, the current dataset from February 2003 to September 2016 was used as the training set, while the data from October 2016 to December 2019 served as the test set.

The influencing factors were precipitation, temperature, potential evapotranspiration, GDP per capita, cultivated land area, and groundwater water supply, and the target was the GWS obtained from the GRACE satellite. The future dataset spans from January 2020 to December 2100. The temperature, precipitation, and potential evapotranspiration were obtained from CMIP6. The GDP per capita, cultivated land area, and groundwater supply were predicted by the latest value of the dataset. As an example, the data value of the influencing factors in December 2019 would be used as the representative data for the future period to predict the future GWS. The time span of the future data obtained would be from May 2021 to December 2100, and the final analysis would utilize data from 2022 to 2100.

To evaluate the merits of the LSTM model's results, this study used the first 80% of the status quo data (2003.02–2019.12) as the training set and the second 20% as the test set. We compared the GWS dataset obtained by predicting the influencing factors using the LSTM model on the test set to the GWS data with the status quo during the same time period. To evaluate the model's performance, we used NSE and RMSE as indicators (Equations (7) and (8)). We obtained the spatial distribution of LSTM model fit scores, as shown in Figure 10. The proportion of points with a NSE greater than 0.3 in the model fit evaluation index was 93.8% of the overall, and the proportion of points greater than 0.5 was 60.2% of the overall. This indicates that the model had a suitable result in most cases.

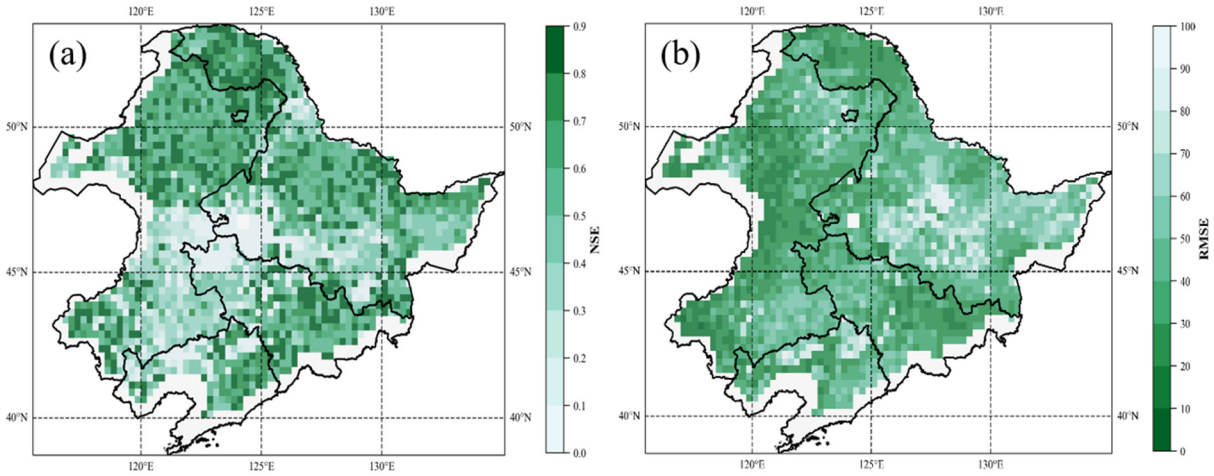


Figure 10. Spatial distribution of LSTM model fit scores. The data from 2003.02 to 2016.09 were used as the training set, and the data from 2016.10 to 2019.12 were used as the test set. (**a**,**b**) were the NSE and RMSE spatial distribution plots of the model test set and the status quo data, respectively.

3.2.2. Spatial Analysis of Future GWS

To conduct a more accurate analysis, the future GWS data obtained for the four scenarios of the seven institutional models are averaged in this paper and divided into the near term (2022–2060) and the far term (2061–2100) to be explored separately. The mean values were derived to obtain trend changes specific to the future period to obtain shared trend changes, and also to eliminate GWS changes at the extremes as a result.

To calculate the future GWS increase and decrease changes more intuitively, we subtracted the mean value of the predicted future GWS data (2021.05–2100.12) from the current situation GWS data (2003.02–2019.12) and obtained the spatial distribution map of the future scenario GWS increases and decreases (Figure 11). From 2020 to 2100, northeast China's precipitation, potential evapotranspiration, and temperature all rose significantly, and precipitation had the greatest influence on GWS among the influencing factors of GWS, leading to a significant overall increase in groundwater in the northeast. Although the southern region, especially the southeastern region, showed an increasing trend, it still

did not reach the historical GWS average, with a difference of more than 90 mm. The GWS increased in the northern region and in the central region of Heilongjiang Province, demonstrating a clear upward trend with an increase of more than 60 mm. This trend was basically the same as that of the current status quo. Generally speaking, the GWS in different time periods under the four scenarios showed an increasing trend. The GWS under different scenarios rose faster in the southern region and slower in the northern region between 2022 and 2060, even decreasing in some scenarios. Between 2060 and 2100, it rose significantly faster in the northern region compared with the southern region.

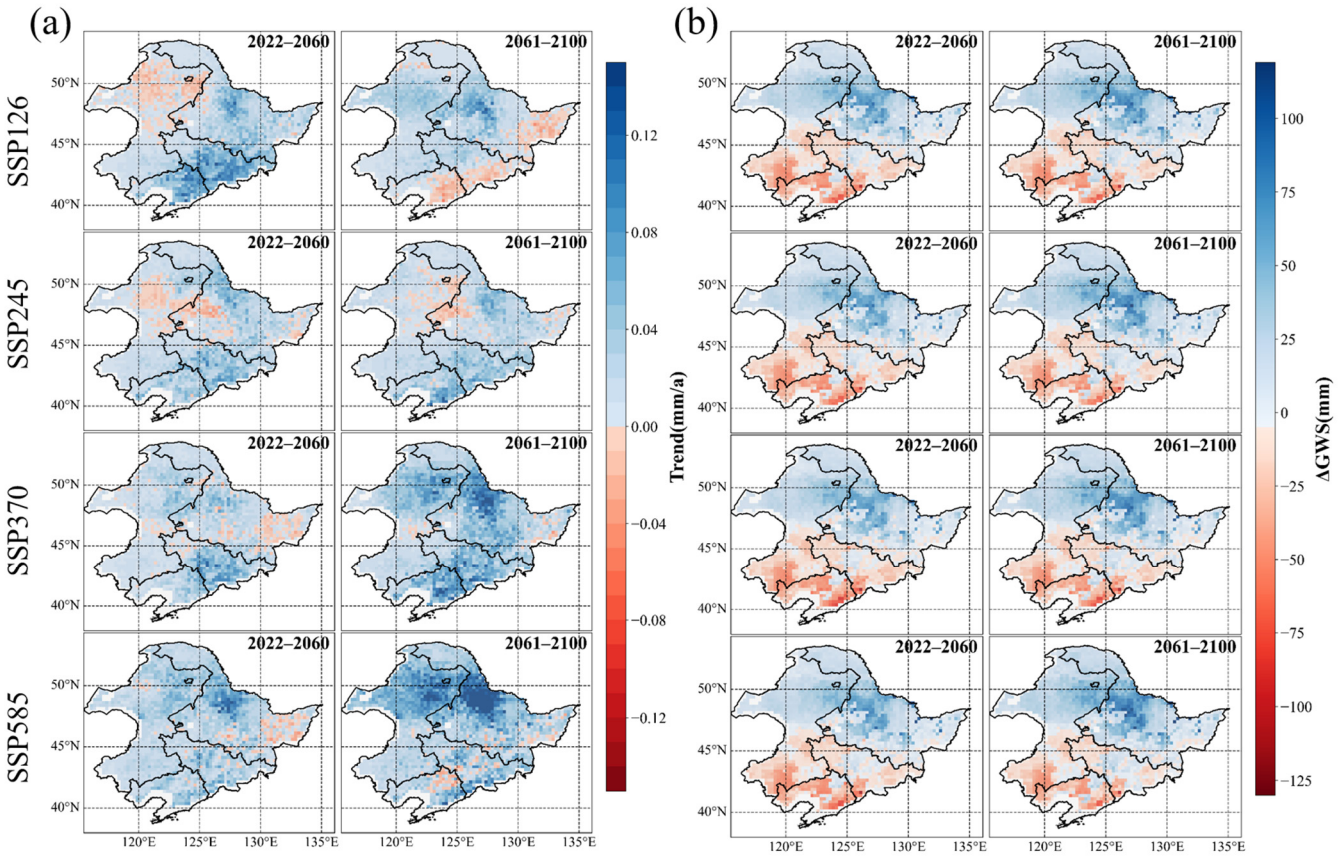


Figure 11. Spatial distribution map of the future GWS. (**a**) Spatial trend of GWS in different contexts over time and its plotting. (**b**) Spatial variation of the mean GWS in different contexts over time and its plotting against the mean GWS in the historical period.

To further analyze the future trends in groundwater storage (GWS) in northeast China, the trends for each grid point under different institutional models were determined, and their averages were calculated (Figure 11a). Spatially, notable differences were observed among various institutions. Under the SSP126 scenario, there was a decline in the northern region, an increase in the southern region, a long-term upward trend in the northern region, and a decline in the southern region. In the SSP245 scenario, the overall trends remained consistent across different periods. A few areas in the central region exhibited a declining trend and the remainder showed an upward trend. The GWS under the SSP370 scenario generally experienced a slight short-term increase and a significant long-term upward trend. In the SSP585 scenario, there was a short-term increase in the southeastern and northwestern regions, a decrease in the central region, and a significant long-term increase. The overall trend aligned with the changing patterns of climate factors. In general, across different periods and scenarios, GWS exhibited an upward trend. Between 2022 and 2060, there was a faster increase in groundwater storage in the southern region and a slower increase or even a decreasing trend in the northern region. From 2060 to 2100,

the northern region experienced a faster increase in GWS, with a significantly rising trend when compared to the southern region.

3.2.3. Temporal Analysis of Future GWS

In the more recent period, groundwater reserves under the four scenarios generally showed an upward trend (Figure 12). This trend showed that groundwater reserves under the SSP126 scenario were the highest and exhibited a continuous upward trend, reaching a peak in 2052. Groundwater reserves under the SSP585 scenario in the longer period showed a significant increase. SSP126 showed an almost stagnant growth trend and was the lowest among the four scenarios. The trends of SSP245 and SSP370 basically remained unchanged, resulting in the groundwater reserves of SSP585 > SSP245 > SSP126 > SSP370 over the longer period (Table 1). The overall groundwater reserves are SSP126 > SSP585 > SSP245 > SSP370.

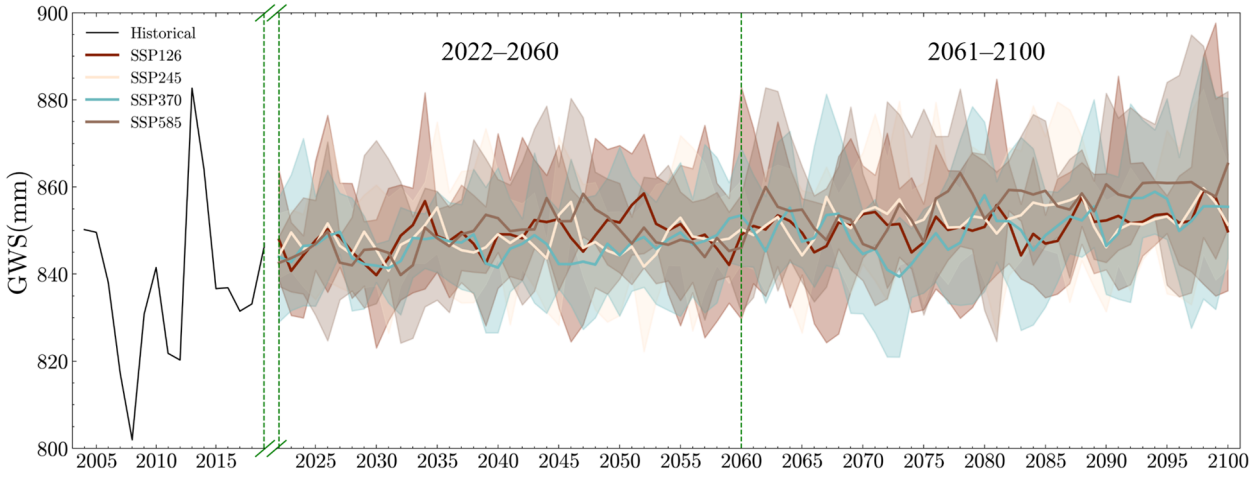


Figure 12. Temporal changes in GWS in the northeast, with shading indicating the range of fluctuations in GWS under different scenarios. 2004–2019 is a plot of interannual changes in GWS for the historical period, and 2002–2100 is a plot of simulated interannual changes in GWS for the future.

Table 1. Mean values of GWS in different scenarios at different times and the difference between their mean values and the historical mean values of GWS.

Scenarios	GWS (mm) (2022–2060)	ΔGWS (mm)	GWS (mm) (2061–2100)	ΔGWS (mm)
SSP126	848.57	9.85	851.34	12.62
SSP245	847.62	8.89	852.82	14.09
SSP370	846.19	7.46	850.55	11.82
SSP585	848.22	9.49	856.26	17.53

Monthly and quarterly analysis of future scenarios' mean GWS in the northeast revealed that, in general, GWS fluctuates more in fall and winter and less in spring and summer (Figure 13). The trend fluctuates the least from April to July and the most in August and September. During the recent period, GWS trended upward in all months, with an overall decrease in the summer and an overall increase in the other seasons. The most pronounced upward values in GWS were seen in the winter. The GWS values were the largest in the SSP126 scenario, showing SSP126 > SSP245 > SSP370 > SSP585. The magnitude of the trend change in different scenarios was SSP245 > SSP126 > SSP370 > SSP585. Under the SSP126 scenario, GWS showed a decreasing trend in August and September and an increasing trend in the other months. The overall trend mean was small, close to zero. Under the SSP245 scenario, the growth rate of GWS in the northeast was the largest,

achieving a peak of 0.302 mm/a in October, and reaching its minimum value in July. The GWS under the SSP370 scenario exhibited a decreasing trend in June and July. The GWS under the SSP585 scenario showed a decreasing trend in November and December and an upward trend in the rest of the months.

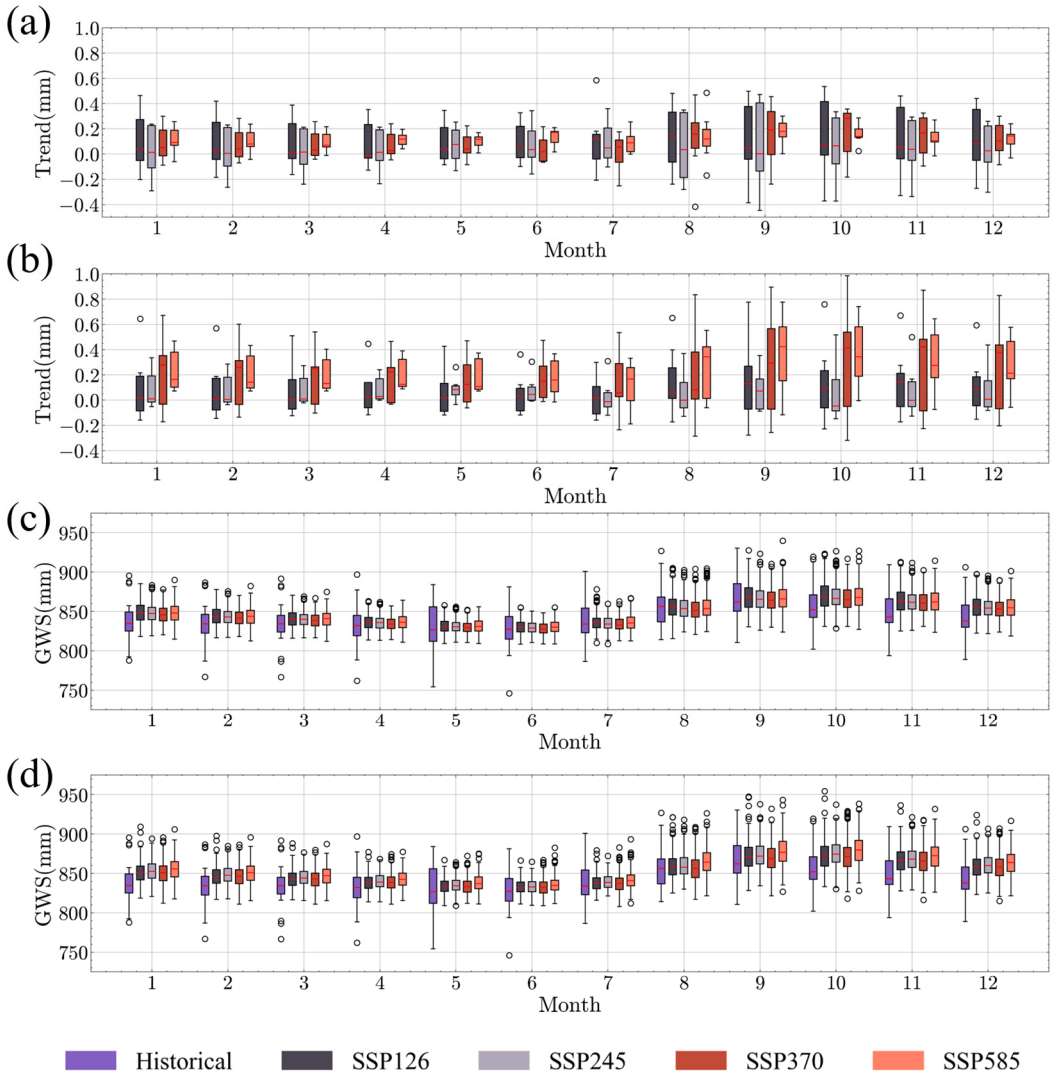


Figure 13. Box plots of monthly trends in mean values of future scenarios of groundwater storage in the northeast and changes in GWS. (**a**,**b**) Changes in inter-monthly trends for different periods. (**c**,**d**) Comparisons of inter-monthly mean values of GWS for different periods with historical GWS. (**a**,**c**) The period in 2022–2060; (**b**,**d**) the period in 2061–2100.

In the future-focused period, GWS under all four scenarios was greater than the current value, and GWS under the SSP585 scenario exceeded that under SSP126. This ranked it first in GWS under the four scenarios. Under the SSP126 scenario, the most recent scenario trend was greater than that in the future-focused period in May, June, and July. Conversely, the most recent scenario trend was smaller than the future-focused period in the other months. Under the SSP245 scenario, the GWS under the future period showed an increasing trend in July, October, November, and December and a decreasing trend in the other months. The trend was similar in the SSP370 scenario but there was a decreasing trend in the other months. The SSP585 scenario showed an increasing trend in all months and was much larger than the other scenarios.

4. Discussion

4.1. Model Uncertainty Analysis

To evaluate the performance of the LSTM simulation, we took the NSE of the GLDAS test set data and the LSTM simulation data as the evaluation index. The results show (Section 3.2.2) that 93.8% of the test set points had a NSE greater than 0.3, with a mean NSE value of 0.582, demonstrating higher suitability. Due to the differences in the study area, study time, and the type of impact factor data selected, the minimum standard of NSE (0.3) selected in this paper, as well as the final NSE mean score result indices obtained, are not exactly the same [45,46].

In studies that utilize deep learning (DL) for simulation, especially large-scale analyses of single or multiple regions, the choice of influencing factors (input data), dataset size, parameter selection, and tuning of the model architecture can have a significant impact on the performance of the model [47,48]. In large-scale studies (with more than 500 data points), complexities arise due to extended model runtime and the challenge of handling numerous grid points. The choices of parameters and dataset superiority are crucial for model performance. Feng et al. (2020) [49] utilized a modified LSTM-based model called data integration (DI) to predict streamflow in 671 catchments over the Continental United States (CONUS). While the LSTM achieved a NSE of 0.73, the improved DI model reached an outstanding NSE of 0.86, emphasizing the significance of high-quality datasets in achieving accurate model predictions. Xiong et al. (2022) [46] simulated dissolved inorganic nitrogen (DIN) in seven basins across Asia, Europe, and North America, achieving NSE values ranging from 0.32 to 0.84. By increasing the dataset length, the NSE for the basin with the lowest value (NSE = 0.32) improved to 0.68. However, the varying impacts of influencing factors on the target variable highlight the need for careful consideration in long-term predictions.

To enhance model accuracy, we performed hyperparameter tuning on 10 grid points to obtain optimal parameters, choosing input layer 16, hidden layer 16, iteration number 130, and batch size 24. Due to the inherent randomness in LSTM models and the limitations in parameter selection, we introduced modifications and iteratively trained models. The final model selection involved a Conv-LSTM [50,51] architecture with parameters set as filters = 32 and kernel size = 3. This reduced the proportion of poorly fitted points from 40% to 5.2%. For regions with consistently poor fits, a region-wide model was trained, resulting in a substantial improvement in NSE from −10.5 to 0.26. In the evaluation metrics of the final model, the mean value of NSE was 0.582—significantly higher than the pre-optimized mean of 0.326 (Figure 14).

4.2. Uncertainty Analysis of Future Climate Factors

The results from historical periods have proven that climate factors play a decisive role in groundwater storage (Section 3.1). Thus, the accuracy of future climate data is crucial in simulating GWS [52]. Although the climate data simulated in the CMIP6 model improved compared with the CMIP5 climate model, their interannual variability did not [42]. Therefore, a downscaling method should be used to downscale the climate data. Two common downscaling methods are dynamic downscaling and static downscaling. The future results obtained by different downscaling methods will obviously differ, even significantly so [53,54]. This study used the Delta method to bias correct future data(Section 2.3). This method has been proven to have good performance in China (see Section 2.3 for details).

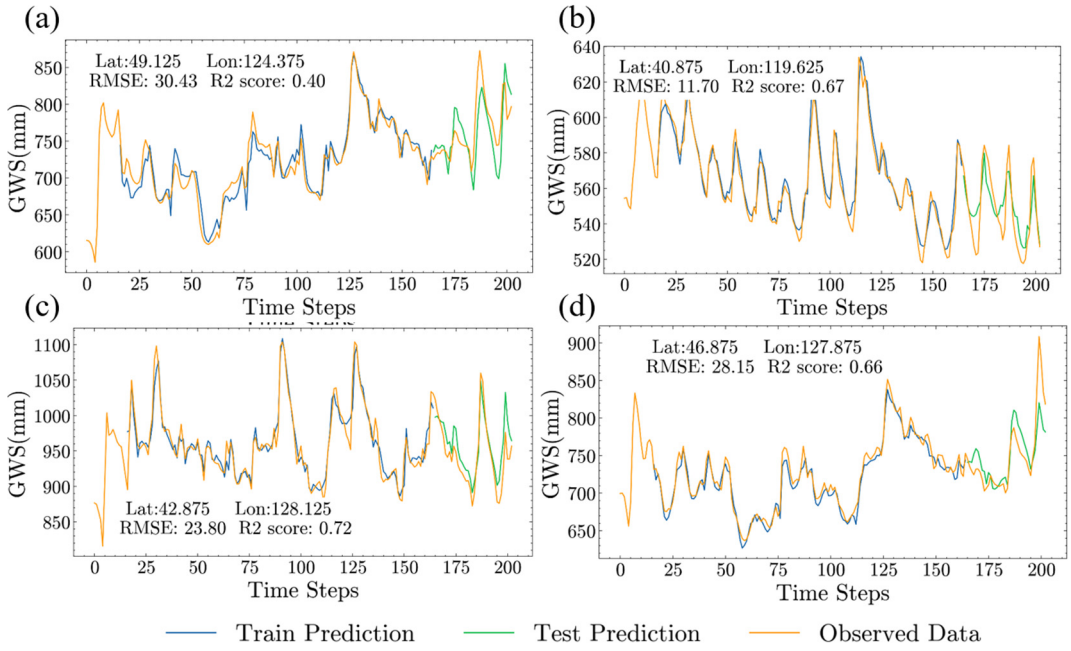


Figure 14. Fitted plots for the four sampled data points. In the figure, the yellow line represents the historical data, the blue line represents the prediction result of the training set, and the green line represents the prediction result of the testing set. (**a**,**b**) LSTM model; (**c**,**d**) Conv-LSTM model.

Figure 15 shows the future changes in annual average climate factors in northeast China that have been corrected by the Delta downscaling method. From the figure, one can see that the values of the modeled future climate factors differed significantly between the different GCMs. However, they exhibited the same upward trend on the time scale. Precipitation and potential evapotranspiration generally showed a fluctuating upward trend, while temperature showed a significant upward trend. This is consistent with the findings derived by other scholars [55]. The mean annual precipitation values for the three scenarios (low, medium, and high) were obtained under the six GCMs developed by Zhang et al. (2010) [56]. They ranged from 564 to 581 mm for the years 2030 to 2059, and from 593 to 637 mm for the years 2070 to 2099. These results are consistent with those from our study.

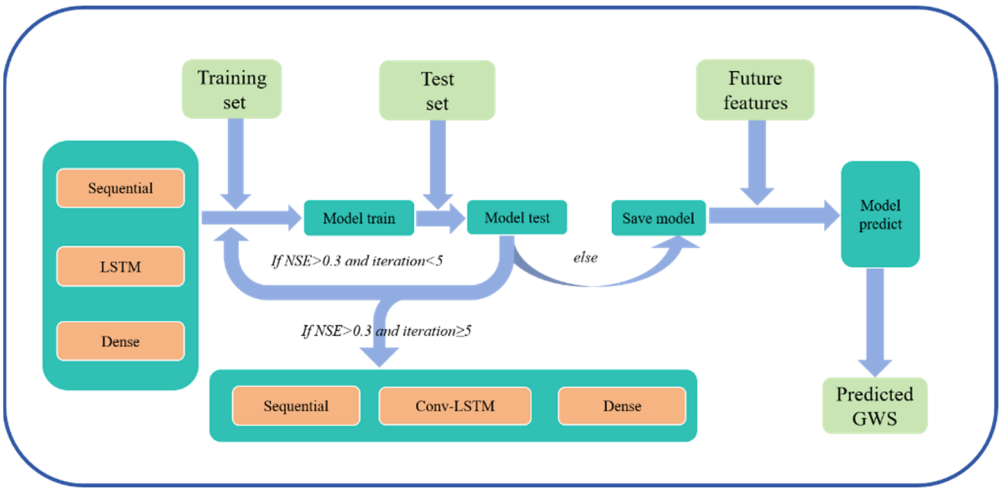


Figure 15. Schematic diagram of LSTM model.

Of the three selected climate factors, temperature experienced a simpler trend change. The differences between the scenarios were large, demonstrating a trend of SSP585 > SSP370 > SSP245 > SSP126 (Figure 16). The changes in the other two climate factors were more

complex, with intertwined differences between scenarios. From 2022 to 2060, potential evapotranspiration in the SSP585 scenario dominated, while precipitation continued to rise, surpassing the other two scenarios after 2035, including the SSP126 scenario. Precipitation and potential evapotranspiration in the SSP370 scenario consistently remained at the lowest levels among the four scenarios. The changes in the trend suggest that GWS in the SSP370 scenario takes the lead among the four scenarios, and the results obtained support this inference. The climatic differences among the four scenarios widened over time. From 2061 to 2100, the SSP585 scenario exhibited the most substantial decrease in groundwater reserves among the four scenarios, significantly larger than that of the other three scenarios [57]. The SSP245 scenario exhibited a significant increase in declining water availability compared to the 2020–2060 period, which could lead to a continuous rise in groundwater stress (GWS) as an absolute factor. Similarly, the SSP126 scenario also showed a notable increase in declining water availability over the same period. However, while the SSP370 scenario displayed an upward trend in declining water, this trend was weaker than in the other two scenarios.

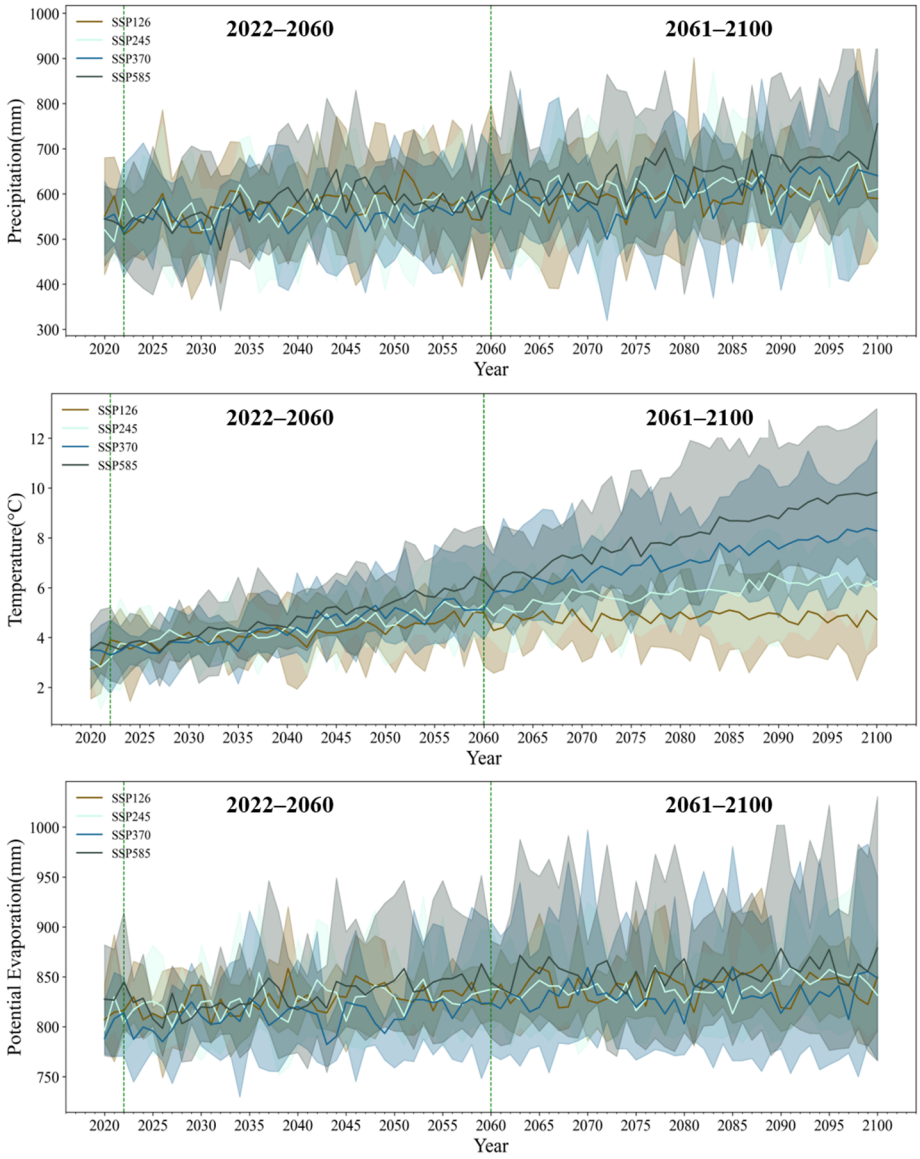


Figure 16. Graph of temporal trends in future climate factors.

4.3. Uncertainty Analysis of Future GWS

Most of the previous studies did not have direct GWS availability or provided GWS data on too short of a time scale. Instead, they used total water storage (TWS) data from the GRACE and GRACE-FO satellites and employed a balance equation that subtracted TWS from soil moisture storage (SMS), snow water equivalent storage (SWES), surface reservoir storage (RESS), and contributions from other categories in the water balance equation [45]. Through this, they obtained the GWS data. However, in this study, the GWS data were obtained directly using the GLDAS-2.2 GRACE-DA product dataset, and the trend conclusions were consistent with the results of previous studies [24,58]. The temporal and spatial GWS data for the entire northeast of China were investigated based on the groundwater storage (GWS) data obtained by the GRACE satellite observations for the period of February 2003 to December 2022. The temporal and spatial variations of GWS in the whole of northeast China are reported in this study, showing that the growth rate in northeast China (including the three eastern provinces and the four eastern leagues of Inner Mongolia) is 2.3 mm/a. These results corroborate previous studies [24].

In this study, although the LSTM model demonstrated good predictive capability, it still has some inherent limitations, particularly in terms of its sensitivity to training data and potential challenges when predicting long-term trends. The LSTM model relies on historical data for training, meaning its predictions can be influenced by the quality and range of the training data. This can be especially problematic for long-term forecasts, where the model may struggle to accurately capture the impact of extreme climate events or unforeseen changes in groundwater storage. Furthermore, when predicting over long periods, the model may face challenges such as overfitting or being unable to account for the influence of external environmental changes. To address these issues and enhance the reliability of the long-term projections, this study utilized data from seven different global climate models (GCMs). By using the mean values from these models, the uncertainty associated with relying on a single climate model is reduced. This approach helps capture a broader range of potential groundwater storage variations under different climate scenarios, thereby improving the stability and accuracy of the predictions.

For the future time groundwater storage model, the future GWS tended to increase in all four scenarios from 2022 to 2100, with 0.08 $\pm$ 0.09 mm/a for SSP126, 0.11 $\pm$ 0.08 mm/a for SSP245, 0.12 $\pm$ 0.09 mm/a for SSP370, and 0.20 $\pm$ 0.07 mm/a for SSP585. Previous studies, like that of Döll et al. (2009), have explored the impacts of global climate change on long-term annual average diffuse groundwater recharge (GWR) from 2041 to 2070 using climate models ECHAM4 and HadCM3. These studies showed that GWR will increase in northeast China in 2041–2070, compared with 1961–1990. GWR in northeast China increased by 30–70%, compared with 1961–1990, showing an upward trend consistent with the conclusions of this paper.

Although the future climate data obtained from different GCMs were consistent in terms of trend, they showed a large variability in values, which can lead to a certain variability in the future GWS obtained through model simulations [59]. Obtaining the average value will have some influence on the extreme value of GWS, weakening the extreme value of GWS variation. Table 2 shows the specific changes in GWS for different scenarios under different GCMs for the period of 2022–2100. Under the future scenarios, the GWS derived by the CanESM5 agency was the maximum value of each GCM for all four scenarios, and the CNRM-CM6-1 was the minimum value of each GCM. The difference between the GWS obtained by both models was more than 14 m. Of the four scenarios, the SSP585 scenario had a more stable trend. It showed a large upward trend and was dominant over the other scenarios and the ΔGWS of the more historical period. The SSP126 and SSP370 scenarios showed large differences in trends across agency models, primarily

due to the slowing of the trend of precipitation in the 2022–2100 time period. The large differences in the trend changes resulted in larger differences in the final total trend changes across agencies. Therefore, the most drastic fluctuations were mainly in the SSP370 scenario. Figure 15 demonstrates the most drastic fluctuation changes in precipitation and potential evapotranspiration under different GCMs in the SSP370 scenario. This indicates the mean of the trend values as being 0.12 mm/a, the difference between the maximum and the minimum as being 0.33 mm/a, and the difference between the peak and the mean as being 0.19 mm/a. To obtain a more accurate value, this study used the seven GWS mean values of seven GCMs as the final results for analysis. This weakened the trend of groundwater storage under extreme climate scenarios, but by obtaining the mean values of groundwater storage under different scenarios with different institutional models, a more stable value of groundwater storage, and thus, a more accurate conclusion, was obtained [44].

Table 2. Trend and GWS change values of GWS means for individual organizations in northeast China.

GCMs	SSP126		SSP245		SSP370		SSP585	
	Trend (mm/a)	ΔGWS (mm)	Trend (mm/a)	ΔGWS (mm)	Trend (mm/a)	ΔGWS (mm)	Trend (mm/a)	ΔGWS (mm)
ACCESS-ESM1-5	0.22	15.68	0.14	17.53	0.05	8.29	0.08	10.82
BCC-CSM2-MR	0.12	6.95	−0.01	9.26	0.11	7.42	0.25	12.63
CanESM5	0.09	19.40	0.25	17.66	0.31	19.59	0.31	24.89
CMCC-ESM2	0.11	13.47	0.17	10.64	0.12	7.74	0.24	12.84
CNRM-CM6-1	0.09	4.12	0.09	3.72	0.16	3.30	0.22	9.34
CNRM-ESM2-1	−0.09	11.32	0.05	12.97	−0.02	12.66	0.14	13.29
INM-CM4-8	0.02	7.81	0.10	8.89	0.12	8.68	0.17	11.15
Average	0.08 ± 0.09	11.25 ± 4.96	0.11 ± 0.08	11.52 ± 4.62	0.12 ± 0.09	9.67 ± 4.77	0.20 ± 0.07	13.57 ± 4.80

5. Conclusions

This study comprehensively explores the spatiotemporal evolution of groundwater storage (GWS) in northeast China. We utilized an enhanced Long Short-Term Memory (LSTM) model and incorporated climate factors from seven global climate models (GCMs) within the CMIP6 framework. The research predicts GWS in the region under four scenarios (SSP126, SSP245, SSP370, SSP585) from 2022 to 2100. The key conclusions drawn from the study are as follows:

(1) Due to the increase in precipitation, groundwater storage (GWS) in northeast China exhibited an overall increasing trend of 2.3 mm/yr during the historical period (2003–2019). This trend is roughly divided by the latitude of 44° N, with GWS showing an increasing trend north of this latitude and a decreasing trend south of it. Additionally, the rate of increase in GWS diminished during the summer months but increased during the fall and winter seasons in northeast China

(2) Climatic factors were the main influences on GWS, dominating its rise and fall; the influence of anthropogenic factors on GWS was greater in the regions with lower groundwater reserves and insufficient recharge of water resources, which were mainly concentrated in the southwestern region of northeast China, as well as the central region.

(3) Benefiting from the increase in precipitation in the coming period, GWS in northeast China is on an upward trend, mainly in the fall and winter. However, GWS in the southern region has not yet reached the level of the historical average for the same period. As such, groundwater resources may face greater pressure and require more protective measures. The trends of GWS changes under the four scenarios in the future period generally show the trend of SSP585 > SSP370 > SSP245 > SSP126. Among them, the SSP126 and SSP585 scenarios fluctuate greatly.

(4) This study provides an in-depth analysis of the changes in groundwater resources in northeast China, revealing potential future trends and spatial distribution of groundwater reserves. This offers crucial insights into the sustainable management of groundwater resources in the face of climate change and socio–economic development. However, there are some limitations to this research. Firstly, although the LSTM model incorporates climate factors from seven GCMs, the accuracy of the model's predictions depends on the reliability of these climate models and the assumptions of different scenarios. Secondly, while the impact of human activities on groundwater storage is considered, there may be some biases or inadequacies in the evaluation process. Future research can address these issues by incorporating more detailed socio–economic data and more precise models of human activity.

Author Contributions: H.W., M.X. and P.Q. conceived the idea of this study and wrote the manuscript; H.W., G.Z. and Y.W. carried out data collection and analysis; P.Q. and H.W. contributed valuable analysis and manuscript review. All authors approved the final manuscript. The authors are indeed informed and agree to publish. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Strategic Priority Research Program of the Chinese Academy of Science (Grant No. XDA28100105, Peng Qi), the Outstanding Young Scientist Project in Jilin Province (20230508099RC, Peng Qi), and the National Natural Science Foundation of China (42371037, Peng Qi).

Data Availability Statement: All data used in this study are publicly available, and the results are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest relevant to this study.

References

1. White, I.; Falkland, T. Management of freshwater lenses on small Pacific islands. *Hydrogeol. J.* **2010**, *18*, 227–246. [CrossRef]
2. Döll, P.; Hoffmann-Dobrev, H.; Portmann, F.; Siebert, S.; Eicker, A.; Rodell, M.; Strassberg, G.; Scanlon, B. Impact of water withdrawals from groundwater and surface water on continental water storage variations. *J. Geodyn.* **2012**, *59*, 143–156. [CrossRef]
3. Alcamo, J.; Flörke, M.; Märker, M. Future long-term changes in global water resources driven by socio-economic and climatic changes. *Hydrol. Sci. J.* **2007**, *52*, 247–275. [CrossRef]
4. Danielopol, D.L.; Griebler, C.; Gunatilaka, A.; Notenboom, J. Present state and future prospects for groundwater ecosystems. *Environ. Conserv.* **2003**, *30*, 104–130. [CrossRef]
5. Taylor, R.G.; Scanlon, B.; Döll, P.; Rodell, M.; Van Beek, R.; Wada, Y.; Longuevergne, L.; Leblanc, M.; Famiglietti, J.S.; Edmunds, M.; et al. Ground water and climate change. *Nat. Clim. Chang.* **2013**, *3*, 322–329. [CrossRef]
6. Dube, T.; Seaton, D.; Shoko, C.; Mbow, C. Advancements in earth observation for water resources monitoring and management in Africa: A comprehensive review. *J. Hydrol.* **2023**, *623*, 129738. [CrossRef]
7. Feng, W.; Zhong, M.; Lemoine, J.; Biancale, R.; Hsu, H.; Xia, J. Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements. *Water Resour. Res.* **2013**, *49*, 2110–2118. [CrossRef]
8. Chen, J.; Famigliett, J.S.; Scanlon, B.R.; Rodell, M. Groundwater storage changes: Present status from GRACE observations. *Remote Sens. Water Resour.* **2016**, *37*, 207–227. [CrossRef]
9. Frappart, F.; Ramillien, G. Monitoring groundwater storage changes using the Gravity Recovery and Climate Experiment (GRACE) satellite mission: A review. *Remote Sens.* **2018**, *10*, 829. [CrossRef]
10. Scanlon, B.R.; Rateb, A.; Anyamba, A.; Kebede, S.; MacDonald, A.M.; Shamsudduha, M.; Small, J.; Sun, A.; Taylor, R.G.; Xie, H. Linkages between GRACE water storage, hydrologic extremes, and climate teleconnections in major African aquifers. *Environ. Res. Lett.* **2022**, *17*, 014046. [CrossRef]
11. Li, B.; Rodell, M.; Kumar, S.; Beaudoing, H.K.; Getirana, A.; Zaitchik, B.F.; de Goncalves, L.G.; Cossetin, C.; Bhanja, S.; Mukherjee, A.; et al. Global GRACE Data Assimilation for Groundwater and Drought Monitoring: Advances and Challenges. *Water Resour. Res.* **2019**, *55*, 7564–7586. [CrossRef]

12. Yang, W.; Long, D.; Scanlon, B.R.; Burek, P.; Zhang, C.; Han, Z. Human intervention will stabilize groundwater storage across the North China Plain. *Water Resour. Res.* **2022**, *58*, e2021WR030884. [CrossRef]
13. Kundzewicz, Z.W.; Mata, L.J.; Arnell, N.W.; Döll, P.; Kabat, P.; Jiménez, B.; Miller, K.A.; Oki, T.; Sen, Z.; Shiklomanov, I.A. Freshwater resources and their management. In *Climate Change 2007: Impacts, Adaptation, and Vulnerability, Contributions of the Working Group II to the Fourth Assessment Report of the Inter-Governmental Panel on Climate Change*; Parry, M.L., Canziani, O.F., Paluticof, J.P., van der Linden, P.J., Hanson, C.E., Eds.; Cambridge University Press: Cambridge, UK, 2007; pp. 173–210.
14. Lo, M.; Famiglietti, J.S.; Yeh, P.J.; Syed, T.H. Improving parameter estimation and water table depth simulation in a land surface model using GRACE water storage and estimated base flow data. *Water Resour. Res.* **2010**, *46*, W05517. [CrossRef]
15. Lekula, M.; Lubczynski, M.W. Use of remote sensing and long-term in-situ time-series data in an integrated hydrological model of the Central Kalahari Basin, Southern Africa. *Hydrogeol. J.* **2019**, *27*, 1541–1562. [CrossRef]
16. Maier, H.R.; Jain, A.; Dandy, G.C.; Sudheer, K. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environ. Model. Softw.* **2010**, *25*, 891–909. [CrossRef]
17. Rajaee, T.; Ebrahimi, H.; Nourani, V. A review of the artificial intelligence methods in groundwater level modeling. *J. Hydrol.* **2019**, *572*, 336–351. [CrossRef]
18. An, L.; Hao, Y.; Yeh, T.-C.J.; Liu, Y.; Liu, W.; Zhang, B. Simulation of karst spring discharge using a combination of time–frequency analysis methods and long short-term memory neural networks. *J. Hydrol.* **2020**, *589*, 125320. [CrossRef]
19. De Filippi, F.M.; Ginesi, M.; Sappa, G. A Fully Connected Neural Network (FCNN) Model to Simulate Karst Spring Flowrates in the Umbria Region (Central Italy). *Water* **2024**, *16*, 2580. [CrossRef]
20. Pölz, A.; Blaschke, A.P.; Komma, J.; Farnleitner, A.H.; Derx, J. Transformer Versus LSTM: A Comparison of Deep Learning Models for Karst Spring Discharge Forecasting. *Water Resour. Res.* **2024**, *60*, e2022WR032602. [CrossRef]
21. Wunsch, A.; Liesch, T.; Broda, S. Groundwater level forecasting with artificial neural networks: A comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX). *Hydrol. Earth Syst. Sci.* **2021**, *25*, 1671–1687. [CrossRef]
22. Jin, Z.-Q.; Zhu, D.-W. Impacts of changes in climate and its variability on food production in Northeast China. *Acta Agron. Sin.* **2008**, *34*, 1588–1597. [CrossRef]
23. Lu, C.; Song, Z.; Wang, W.; Zhang, Y.; Si, H.; Liu, B.; Shu, L. Spatiotemporal variation and long-range correlation of groundwater depth in the Northeast China Plain and North China Plain from 2000~2019. *J. Hydrol. Reg. Stud.* **2021**, *37*, 100888. [CrossRef]
24. Wang, Z.L.; Sun, C.H.; Jiang, Q.X.; Liu, C.X.; Shan, J.X. Characterization of spatial and temporal changes in groundwater storage and analysis of its influ-encing factors in the three northeastern provinces of China. *Adv. Water Sci.* **2023**, *34*, 360–373. [CrossRef]
25. Yan, B.; Xiao, C.; Liang, X.; Fang, Z. Impacts of urban land use on nitrate contamination in groundwater, Jilin City, Northeast China. *Arab. J. Geosci.* **2016**, *9*, 105. [CrossRef]
26. Sun, Q.; Xu, C.; Gao, X.; Lu, C.; Cao, B.; Guo, H.; Yan, L.; Wu, C.; He, X. Response of groundwater to different water resource allocation patterns in the Sanjiang Plain, Northeast China. *J. Hydrol. Reg. Stud.* **2022**, *42*, 101156. [CrossRef]
27. Wang, S.; Cui, G.; Li, X.; Liu, Y.; Li, X.; Tong, S.; Zhang, M. GRACE Satellite-Based Analysis of Spatiotemporal Evolution and Driving Factors of Groundwater Storage in the Black Soil Region of Northeast China. *Remote. Sens.* **2023**, *15*, 704. [CrossRef]
28. Cheng, W.; Dong, F.; Tang, R.; Yin, H.; Shi, L.; Zhai, Y.; Li, X. Improved Combination Weighted Prediction Model of Aquifer Water Abundance Based on a Cloud Model. *ACS Omega* **2022**, *7*, 35840–35850. [CrossRef]
29. Li, B.; Beaudoing, H.; Rodell, M. GLDAS Catchment Land Surface Model L4 Daily 0.25 × 0.25 degree GRACE-DA1 V2. 2 [Dataset]. Goddard Earth Sciences Data and Information Services Center (GES DISC). 2020. Available online: https://disc.gsfc.nasa.gov/datasets/GLDAS_CLSM025_DA1_D_2.2/summary (accessed on 10 October 2023).
30. Amanambu, A.C.; Obarein, O.A.; Mossa, J.; Li, L.; Ayeni, S.S.; Balogun, O.; Oyebamiji, A.; Ochege, F.U. Groundwater system and climate change: Present status and future considerations. *J. Hydrol.* **2020**, *589*, 125163. [CrossRef]
31. Jia, X.; O'Connor, D.; Hou, D.; Jin, Y.; Li, G.; Zheng, C.; Ok, Y.S.; Tsang, D.C.; Luo, J. Groundwater depletion and contamination: Spatial distribution of groundwater resources sustainability in China. *Sci. Total Environ.* **2019**, *672*, 551–562. [CrossRef]
32. Amaranto, A.; Munoz-Arriola, F.; Solomatine, D.P.; Corzo, G. A Spatially Enhanced Data-Driven Multimodel to Improve Semiseasonal Groundwater Forecasts in the High Plains Aquifer, USA. *Water Resour. Res.* **2019**, *55*, 5941–5961. [CrossRef]
33. Kundzewicz, Z.W.; Doell, P. Will groundwater ease freshwater stress under climate change? *Hydrol. Sci. J.* **2009**, *54*, 665–675. [CrossRef]
34. Zhang, X.; Kong, X. Sustainable utilization of cultivated land resources under groundwater crisis in the Huanghuaihai Plain. *China Land Sci.* **2014**, *5*, 90–96. [CrossRef]
35. Li, X.; Li, G.; Zhang, Y. Identifying Major Factors Affecting Groundwater Change in the North China Plain with Grey Relational Analysis. *Water* **2014**, *6*, 1581–1600. [CrossRef]

36. Peng, S. 1-km Monthly Mean Temperature Dataset for China (1901–2022), National Tibetan Plateau/Third Pole Environment Data Center. 2019. Available online: https://data.tpdc.ac.cn/en/data/71ab4677-b66c-4fd1-a004-b2a541c4d5bf/ (accessed on 10 October 2023).
37. Peng, S. 1-km Monthly Precipitation Dataset for China (1901–2022). National Tibetan Plateau/Third Pole Environment Data Center. 2020. Available online: https://data.tpdc.ac.cn/zh-hans/data/faae7605-a0f2-4d18-b28f-5cee413766a2 (accessed on 10 October 2023).
38. Peng, S. 1 km Monthly Potential Evapotranspiration Dataset in China (1901–2022). National Tibetan Plateau/Third Pole Environment Data Center. 2022. Available online: https://loess.geodata.cn/data/datadetails.html?dataguid=34595274939620&docid=74 (accessed on 10 October 2023).
39. Zhang, C.; Dong, J.; Ge, Q. IrriMap_CN: Improved Annual Irrigation Maps Across China in 2000–2019 Based on Satellite Imagery and Machine-Learning Method. figshare. Dataset. 2022. Available online: https://figshare.com/articles/dataset/IrriMap_CN_Improved_annual_irrigation_maps_across_China_in_2000_2019_based_on_satellite_imagery_and_machine-learning_method/20363115/1 (accessed on 10 October 2023).
40. Jiang, D.; Hu, D.; Tian, Z.; Lang, X. Differences between CMIP6 and CMIP5 Models in Simulating Climate over China and the East Asian Monsoon. *Adv. Atmos. Sci.* **2020**, *37*, 1102–1118. [CrossRef]
41. Allen, R.G.; Pruitt, W.O.; Wright, J.L.; Howell, T.A.; Ventura, F.; Snyder, R.; Itenfisu, D.; Steduto, P.; Berengena, J.; Yrisarry, J.B.; et al. A recommendation on standardized surface resistance for hourly calculation of reference ETo by the FAO56 Penman-Monteith method. *Agric. Water Manag.* **2005**, *81*, 1–22. [CrossRef]
42. Yuan, F.; Tung, Y.-K.; Ren, L. Projection of future streamflow changes of the Pearl River basin in China using two delta-change methods. *Hydrol. Res.* **2015**, *47*, 217–238. [CrossRef]
43. Bulut, M. Hydroelectric generation forecasting with long short term memory (LSTM) based deep learning model for turkey. *arXiv* **2021**, arXiv:2109.09013.
44. Nkhonjera, G.K.; Dinka, M.O. Significance of direct and indirect impacts of climate change on groundwater re-sources in the Olifants River basin: A review. *Glob. Planet. Chang.* **2017**, *158*, 72–82. [CrossRef]
45. Haq, M.A.; Jilani, A.K.; Prabu, P. Deep Learning Based Modeling of Groundwater Storage Change. *Comput. Mater. Contin.* **2022**, *70*, 4599–4617. [CrossRef]
46. Xiong, R.; Zheng, Y.; Chen, N.; Tian, Q.; Liu, W.; Han, F.; Jiang, S.; Lu, M.; Zheng, Y. Predicting Dynamic Riverine Nitrogen Export in Unmonitored Watersheds: Leveraging Insights of AI from Data-Rich Regions. *Environ. Sci. Technol.* **2022**, *56*, 10530–10542. [CrossRef]
47. Zhu, Q.; Chen, J.; Zhu, L.; Duan, X.; Liu, Y. Wind Speed Prediction with Spatio–Temporal Correlation: A Deep Learning Approach. *Energies* **2018**, *11*, 705. [CrossRef]
48. Love, B.J.; Nejadhashemi, A.P. Water quality impact assessment of large-scale biofuel crops expansion in agricultural regions of Michigan. *Biomass Bioenergy* **2011**, *35*, 2200–2216. [CrossRef]
49. Feng, D.; Fang, K.; Shen, C. Enhancing Streamflow Forecast and Extracting Insights Using Long-Short Term Memory Networks with Data Integration at Continental Scales. *Water Resour. Res.* **2020**, *56*, e2019WR026793. [CrossRef]
50. Kim, S.; Hong, S.; Joh, M.; Song, S.K. Deeprain: Convlstm network for precipitation prediction using multichannel radar data. *arXiv* **2017**, arXiv:1711.02316.
51. Khan, M.I.; Maity, R. Hybrid Deep Learning Approach for Multi-Step-Ahead Daily Rainfall Prediction Using GCM Simulations. *IEEE Access* **2020**, *8*, 52774–52784. [CrossRef]
52. Crosbie, R.S.; Scanlon, B.R.; Mpelasoka, F.S.; Reedy, R.C.; Gates, J.B.; Zhang, L. Potential climate change effects on groundwater recharge in the High Plains Aquifer, USA. *Water Resour. Res.* **2013**, *49*, 3936–3951. [CrossRef]
53. Tang, J.; Niu, X.; Wang, S.; Gao, H.; Wang, X.; Wu, J. Statistical downscaling and dynamical downscaling of regional climate in China: Present climate evaluations and future climate projections. *J. Geophys. Res. Atmos.* **2016**, *121*, 2110–2129. [CrossRef]
54. Kurylyk, B.L.; MacQuarrie, K.T. The uncertainty associated with estimating future groundwater recharge: A summary of recent research and an example from a small unconfined aquifer in a northern humid-continental climate. *J. Hydrol.* **2013**, *492*, 244–253. [CrossRef]
55. Yuan, W.; Wu, S.-Y.; Hou, S.; Xu, Z.; Pang, H.; Lu, H. Projecting Future Vegetation Change for Northeast China Using CMIP6 Model. *Remote. Sens.* **2021**, *13*, 3531. [CrossRef]
56. Zhang, Y.-G.; Nearing, M.A.; Zhang, X.-C.; Xie, Y.; Wei, H. Projected rainfall erosivity changes under climate change from multimodel and multiscenario projections in Northeast China. *J. Hydrol.* **2010**, *384*, 97–106. [CrossRef]
57. Yang, X.; Zhou, B.; Xu, Y.; Han, Z. CMIP6 Evaluation and Projection of Temperature and Precipitation over China. *Adv. Atmos. Sci.* **2021**, *38*, 817–830. [CrossRef]

58. Feng, W.; Shum, C.K.; Zhong, M.; Pan, Y. Groundwater Storage Changes in China from Satellite Gravity: An Overview. *Remote. Sens.* **2018**, *10*, 674. [CrossRef]
59. Green, T.R. Linking Climate Change and Groundwater. In *Integrated Groundwater Management*; Jakeman, A.J., Barreteau, O., Hunt, R.J., Rinaudo, J.D., Ross, A., Eds.; Springer: Cham, Switzerland, 2016; pp. 97–141. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.