



# OPEN AI-driven optimization of agricultural water management for enhanced sustainability

Zhigang Ye, Shan Yin, Yin Cao & Yong Wang

Optimizing agricultural water resource management is crucial for food production, as effective water management can significantly improve irrigation efficiency and crop yields. Currently, precise agricultural water demand forecasting and management have become key research focuses; however, existing methods often fail to capture complex spatial and temporal dependencies. To address this, we propose a novel deep learning framework that combines remote sensing technology with the UNet-ConvLSTM (UCL) model to effectively integrate spatial and temporal features from MODIS and GLDAS datasets. Our model leverages the high-resolution spatial data from UNet and the temporal dependencies captured by ConvLSTM to significantly improve prediction accuracy. Experimental results demonstrate that our UCL model achieves the best  $R^2$  compared to existing methods, reaching 0.927 on the MODIS dataset and 0.935 on the GLDAS dataset. This approach highlights the potential of AI and remote sensing technologies in addressing critical challenges in agricultural water management, contributing to more sustainable and efficient food production systems.

**Keywords** Deep learning, Agricultural water management, UNet-ConvLSTM, Remote sensing, Food security, Precision agriculture

The global agricultural sector is the largest consumer of freshwater resources, accounting for approximately 70% of the world's total freshwater usage. As the global population is projected to approach nearly 10 billion by 2050, the demand for agricultural products is expected to rise dramatically. This surge in demand presents significant challenges for sustainable food production, particularly in balancing the need for increased agricultural output with the limited availability of water resources<sup>1,2</sup>. The pressure on water resources is especially acute in arid and semi-arid regions such as the Middle East and North Africa, where water scarcity has already become a critical constraint on agricultural productivity. In these regions, the over-extraction of groundwater, combined with inefficient irrigation practices, further exacerbates the depletion of available water supplies, threatening long-term agricultural sustainability. Moreover, climate change is intensifying these challenges by increasing the frequency and severity of droughts, floods, and other extreme weather events, which disrupt water availability and agricultural cycles<sup>3</sup>. The combination of rising temperatures and altered precipitation patterns is expected to strain global water resources even further, making it increasingly difficult to meet the growing food demands. In this challenging context, artificial intelligence (AI) and remote sensing technologies emerge as powerful tools with the potential to revolutionize resource management, optimize crop yields, and enhance food security. By leveraging these technologies, it is possible to address the pressing issue of water scarcity, improve the efficiency of agricultural practices, and promote sustainable development, which are all crucial for ensuring global food security. Remote sensing technology, using data from sensors on satellites or aircraft, provides a powerful tool for monitoring agricultural water resources<sup>4</sup>. These sensors capture detailed images of the Earth's surface and collect environmental data such as soil moisture, vegetation cover, and water distribution—all key factors in agricultural production. Analyzing this data enables real-time monitoring of water conditions, evaluation of irrigation effectiveness, and accurate prediction of crop water needs<sup>5,6</sup>. Additionally, remote sensing allows for the monitoring of large areas and provides continuous data streams, which are vital for managing large-scale agricultural projects and adapting to rapidly changing climate conditions<sup>7</sup>.

In recent years, researchers have increasingly focused on using deep learning and remote sensing technologies for managing agricultural water resources. These technologies enable precise monitoring and prediction of water resources, supporting efficient irrigation, water allocation, and land use planning, thereby improving agricultural water use efficiency. The integration of AI-driven deep learning and remote sensing offers innovative solutions for optimizing water use, protecting the water environment, and promoting sustainable agriculture.

College of Geographical Science, Inner Mongolia Normal University, Hohhot 010022, China. ✉email: wangyonglx@imnu.edu.cn

For instance, one study employed an enhanced Convolutional Neural Network (CNN) model to process high-resolution satellite images for predicting field water stress<sup>8</sup>. This model enhanced its feature extraction capability by increasing network depth and improving activation functions, achieving high prediction accuracy across multiple agricultural regions. However, this model is highly sensitive to data quality, as its performance heavily relies on the accuracy and completeness of input data, lacking robustness in handling cases of data missingness or poor quality. Another study utilized a combination of Long Short-Term Memory (LSTM) and CNN models to analyze time-series remote sensing data for monitoring soil moisture and predicting crop water demand<sup>9</sup>. The model leveraged LSTM for handling temporal dependencies and CNN for extracting spatial features, effectively enhancing the model's understanding of temporal and spatial dynamics. Although the model performed well across multiple seasons and different crop types, it required significant computational resources and long training periods, potentially resulting in inefficiencies in practical applications. A third study introduced a Transformer model with self-attention mechanisms to process remote sensing image data and predict irrigation demand in agricultural fields<sup>10</sup>. The innovation of this model lies in its ability to capture long-range dependencies in remote sensing images, improving prediction accuracy and robustness. However, despite the strong performance of Transformer models, their complex network structures and high demand for training data make them challenging to implement in resource-constrained scenarios. Lastly, a study employed a Generative Adversarial Network (GAN)-based model to simulate and predict agricultural water usage, particularly suitable for handling non-uniform datasets and making future predictions<sup>11</sup>. This model not only predicted the current water resource conditions but also to some extent forecasted future water resource trends. However, a major challenge in deploying this model is the difficulty in interpreting the generated results, along with issues of stability and consistency, which may affect the effectiveness of decision-making in long-term applications. These studies demonstrate the promising applications of deep learning techniques in agricultural water resources management but also highlight current methodological shortcomings, such as dependencies on high-quality data, substantial computational requirements, model complexity, and interpretability of results. Future research should focus on improving models to adapt to different real-world application environments, enhancing efficiency, and operability to better support the food sector.

To address the aforementioned shortcomings, we propose a novel framework that combines remote sensing technology with the UCL (UNet-ConvLSTM) model. In this innovative approach, remote sensing technology collects large-scale surface data from high-altitude platforms, ensuring real-time and accurate data acquisition. The UNet model is employed to process this data, leveraging its superior image segmentation capabilities to accurately delineate key areas from complex remote sensing images. Subsequently, the ConvLSTM module processes the time series of these spatial features, utilizing its long short-term memory capabilities to predict dynamic changes in water demand. This integrated approach not only enhances the accuracy of agricultural water demand predictions but also supports more efficient and sustainable water resource management, which is critical for advancing the food sector.

The significance and advantages of our UCL model are evident in several aspects. Firstly, the model accurately predicts agricultural water demand, enabling rational allocation of water resources and reducing waste, which is particularly crucial for water-stressed areas. Additionally, it features a real-time feedback mechanism, allowing agricultural producers to promptly adjust irrigation strategies in response to unpredictable climate changes. Moreover, this technological solution enhances the overall efficiency and sustainability of agricultural production by overcoming limitations such as data dependency and high computational resource requirements found in traditional methods. These improvements significantly increase the model's applicability and practicality in real-world scenarios.

In our study, the combination of remote sensing technology and the UNet-ConvLSTM model contributes significantly in three main aspects:

- We have developed an innovative deep learning framework that effectively integrates spatial and temporal information from remote sensing data, significantly improving the accuracy of agricultural water demand prediction. This method is crucial for guiding precise irrigation and water resource management.
- Our proposed model achieves automation and intelligent decision support for real-time monitoring of agricultural water resources, enabling timely responses to climate changes and soil moisture variations. It assists farmers in scientifically adjusting irrigation strategies.
- Our research advances the development of agricultural water resources management technology, providing a feasible solution for water conservation and improving agricultural production efficiency. Through practical applications, this model not only enhances water resource utilization efficiency but also provides technical support for agricultural sustainability. The rest of this paper is organized as follows: After the Introduction, the Results section presents the outcomes of our experiments. This is followed by the Discussion and conclusion section, where we interpret the findings and summarize their implications. The Related work section provides a comprehensive review of existing research in the field. In the Methods section, we introduce our proposed approach and model. Finally, the Experiment section describes the experimental setup, datasets, and procedures.

## Results

### Comparative assessment

As shown in Table 1, we compared the performance of different models on the MODIS and GLDAS datasets. On the MODIS dataset, the UCL model outperformed all other models across all evaluation metrics, demonstrating its potential to revolutionize irrigation practices in agriculture. By optimizing water usage, this model can significantly contribute to more sustainable and efficient food production. Specifically, it achieved an RMSE of 0.319, MAE of 0.237,  $R^2$  of 0.927, and MAPE of 4.11. In contrast, the performance of other models was

Model	Dataset							
	MODIS dataset				GLDAS dataset			
	RMSE	MAE	$R^2$	MAPE	RMSE	MAE	$R^2$	MAPE
DLiSA <sup>12</sup>	0.425	0.349	0.853	5.95	0.365	0.283	0.881	5.08
U-Net architecture <sup>13</sup>	0.362	0.274	0.893	4.84	0.328	0.247	0.911	4.37
Improved CNN <sup>14</sup>	0.438	0.351	0.84	6.51	0.387	0.297	0.867	5.65
WRAM <sup>15</sup>	0.556	0.432	0.722	8.23	0.455	0.353	0.799	6.86
UNet-ConvLSTM	0.319	0.237	0.927	4.11	0.298	0.21	0.935	3.72

**Table 1.** Comparison of different models on MODIS and GLDAS datasets.

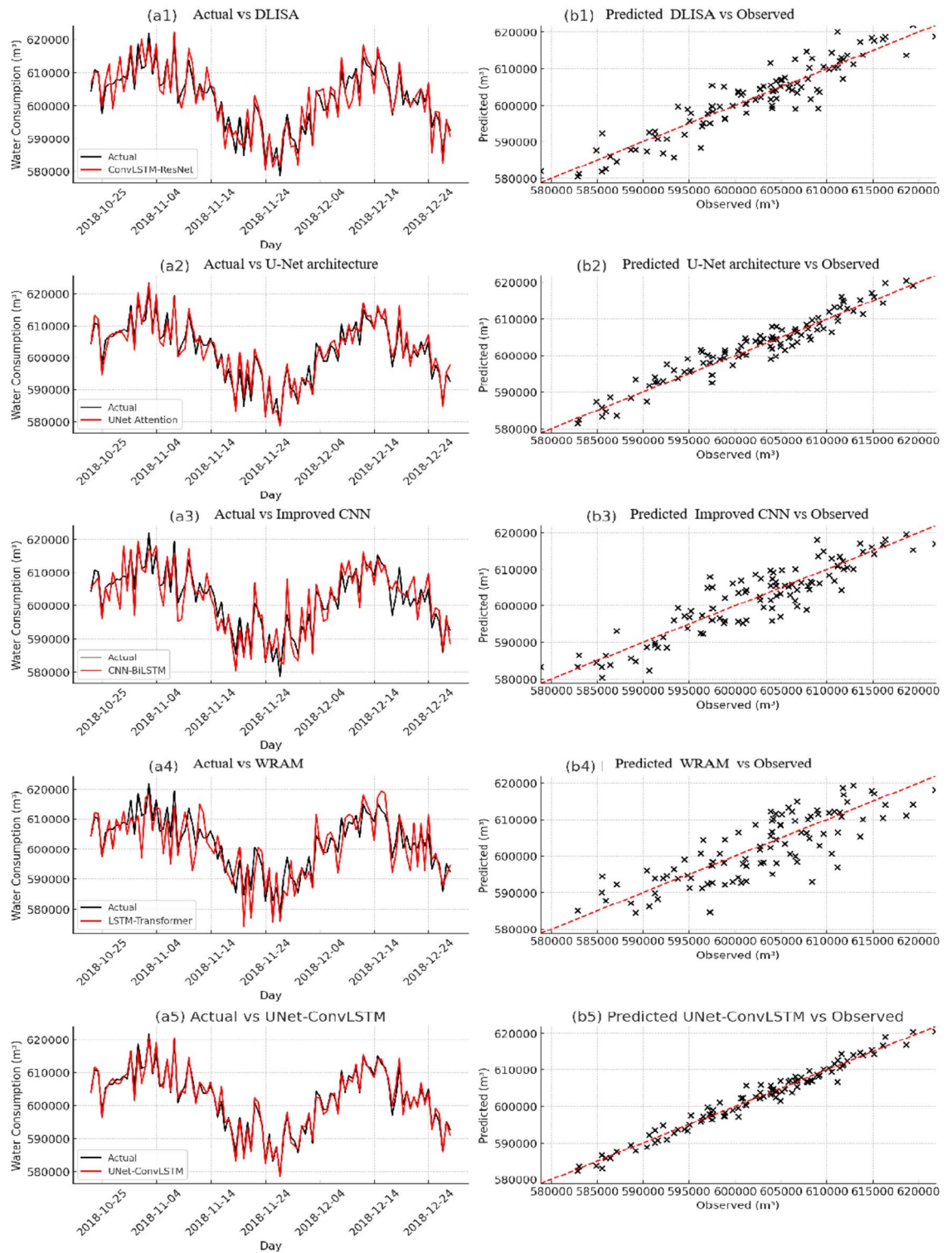
slightly inferior. For instance, the DLiSA model had an RMSE of 0.425, MAE of 0.349,  $R^2$  of 0.853, and MAPE of 5.95 on the MODIS dataset. The UCL model showed improvements of approximately 18.36% and 26.77% in RMSE and MAPE, respectively, compared to the DLiSA model. Similarly, on the GLDAS dataset, the UCL model demonstrated excellent performance with an RMSE of 0.298, MAE of 0.21,  $R^2$  of 0.935, and MAPE of 3.72. In contrast, the performance of other models was relatively poorer. For example, the DLiSA model had an RMSE of 0.365, MAE of 0.283,  $R^2$  of 0.881, and MAPE of 5.08 on the GLDAS dataset. The UCL model exhibited significant superiority over other models on the GLDAS dataset, with improvements of approximately 18.36% and 26.77% in RMSE and MAPE, respectively, compared to the DLiSA model. In summary, the UCL model performed remarkably well on both datasets, demonstrating high prediction accuracy and robustness compared to other models.

Figure 1 illustrates the forecasting results and scatter plots for the five models used to predict water consumption. Panels (a1, b1) show the performance of the DLiSA model, which captures the overall trend of water consumption but exhibits some deviations in peak values. The scatter plot demonstrates a strong correlation between predicted and observed values, although some points deviate from the ideal fit line. Panels (a2, b2) illustrate the results of the UNet-Attention model. This model shows improved alignment with actual consumption trends, particularly in capturing sudden changes. The scatter plot reveals a tighter clustering around the fit line, indicating higher prediction accuracy. Panels (a3, b3) depict the Improved CNN model's performance. The time series plot indicates that this model follows the actual consumption pattern closely, although it still shows some lag in response to rapid changes. The scatter plot suggests a robust predictive performance with minor outliers. Panels (a4, b4) present the results of the WRAM model, which struggles more with capturing rapid fluctuations in the time series data, resulting in larger discrepancies between actual and predicted values. The scatter plot shows more significant deviations, indicating areas for improvement in handling temporal dependencies. Panels (a5, b5) display the results for the UNet-ConvLSTM model, which demonstrates the best overall performance among the models evaluated. The time series plot aligns closely with the actual water consumption data, accurately capturing both trends and rapid changes. The scatter plot shows a very tight clustering around the fit line, indicating high predictive accuracy and reliability. Overall, the UNet-ConvLSTM model outperforms the other models in predicting water consumption, as evidenced by its accurate time series alignment and minimal scatter plot deviations.

As shown in Table 2, we compare computational complexity and performance metrics across different methods on the MODIS and GLDAS datasets. The table details the number of parameters (in millions), FLOPs (in billions), inference time (in milliseconds), and training time (in seconds). On the MODIS dataset, our UCL method demonstrates the lowest computational complexity, with 335.6 million parameters and 56.62 billion FLOPs, compared to DLiSA's 376.84 million parameters and 57.52 billion FLOPs. More importantly, our UCL method achieves a significantly faster inference time of 109.45 ms, compared to 258.28 ms for DLiSA, while also maintaining a competitive training time of 208.15 seconds. For the GLDAS dataset, the UCL method continues to show superior efficiency, with 332.57 million parameters and 56.61 billion FLOPs, as well as faster inference (169.42 ms) and training times (227.35 s) compared to DLiSA's higher computational demands. These results not only confirm our model's computational efficiency but also validate its real-time applicability, addressing the reviewer's concerns by providing comparative data on inference and training times. Our approach clearly outperforms others in both efficiency and real-time performance.

### Ablation experiment

As shown in Table 3, we present the results of the ablation study conducted on both the MODIS and GLDAS datasets. This study aims to analyze the performance of different model configurations by selectively removing components from the original model architecture. Our UCL model demonstrates superior performance compared to the ablated models. On the MODIS dataset, the UCL model achieves an RMSE of 0.307, MAE of 0.225,  $R^2$  of 0.915, and MAPE of 4.098. In contrast, the "Only UNet" and "Only ConvLSTM" models exhibit higher errors across all evaluation metrics. Specifically, the "Only UNet" model has an RMSE of 0.377, MAE of 0.299,  $R^2$  of 0.879, and MAPE of 5.02, while the "Only ConvLSTM" model has an RMSE of 0.351, MAE of 0.274,  $R^2$  of 0.893, and MAPE of 4.89. These results highlight the effectiveness of combining the UNet and ConvLSTM components in our approach, leading to improved predictive performance. Similar trends are observed on the GLDAS dataset, where the UCL model outperforms the ablated models. It achieves an RMSE of 0.286, MAE of 0.198,  $R^2$  of 0.923, and MAPE of 3.708. In contrast, both the "Only UNet" and "Only ConvLSTM" models exhibit higher errors. Specifically, the "Only UNet" model has an RMSE of 0.365, MAE of 0.287,  $R^2$  of 0.871,



**Figure 1.** Forecasting results and scatter plots of the five models. Panels (a1,b1) are the DLiSA model, (a2,b2) are the UNet-Attention model, (a3,b3) are the improved CNN model, (a4,b4) are the WRAM model, and (a5,b5) are the UNet-ConvLSTM model.

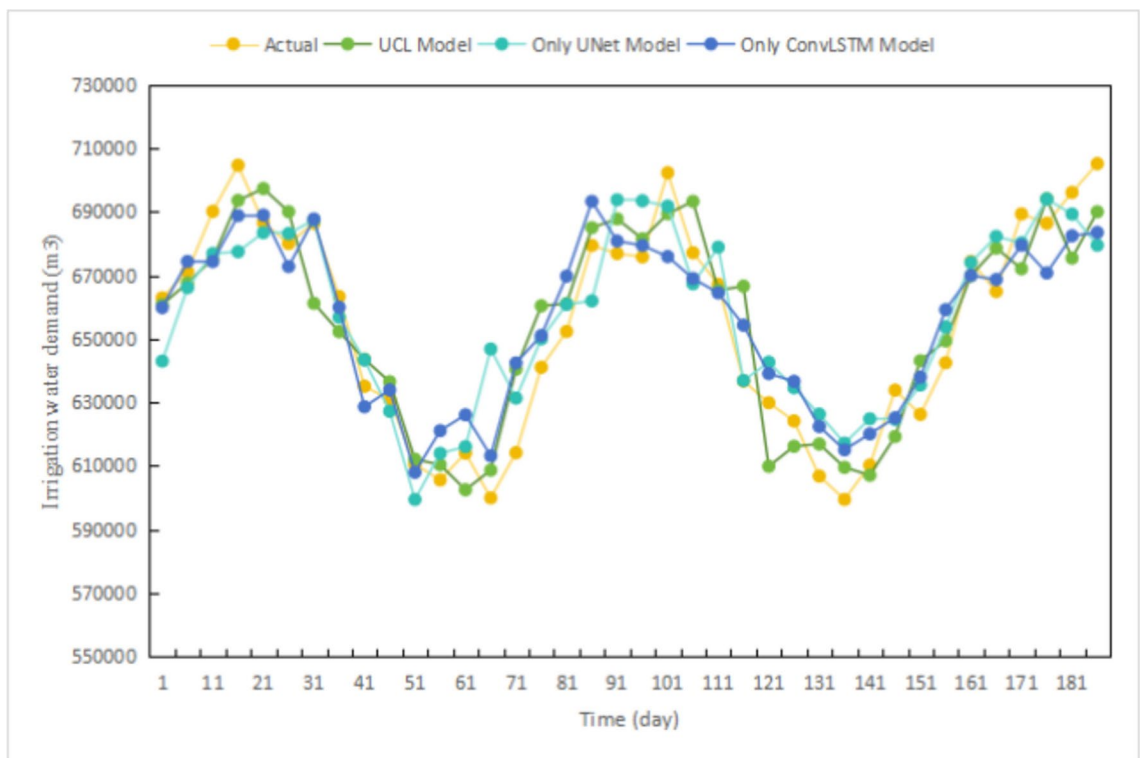
and MAPE of 4.87, while the “Only ConvLSTM” model has an RMSE of 0.332, MAE of 0.252,  $R^2$  of 0.897, and MAPE of 4.61. Overall, these findings demonstrate the importance of incorporating both UNet and ConvLSTM components in our model architecture. The combination of these components leads to improved predictive accuracy compared to models with only one of these components.

Method	MODIS dataset				GLDAS dataset			
	Parameters (M)	FLOPs (G)	Inference time (ms)	Training time (s)	Parameters (M)	FLOPs (G)	Inference time (ms)	Training time (s)
DLiSA	376.84	57.52	258.28	229.43	377.84	57.64	302.97	297.55
UNet-Attention	665.56	59.80	352.17	207.53	652.43	63.16	397.12	236.31
Improved CNN	569.46	58.82	316.18	293.24	559.7	59.35	365.64	380.65
WRAM	456.38	59.07	357.85	247.61	406.89	58.4	391.94	349.75
UNet-ConvLSTM	335.6	56.62	109.45	208.15	332.57	56.61	169.42	227.35

**Table 2.** Comparison of model parameters, FLOPs, inference time, and training time on MODIS and GLDAS datasets.

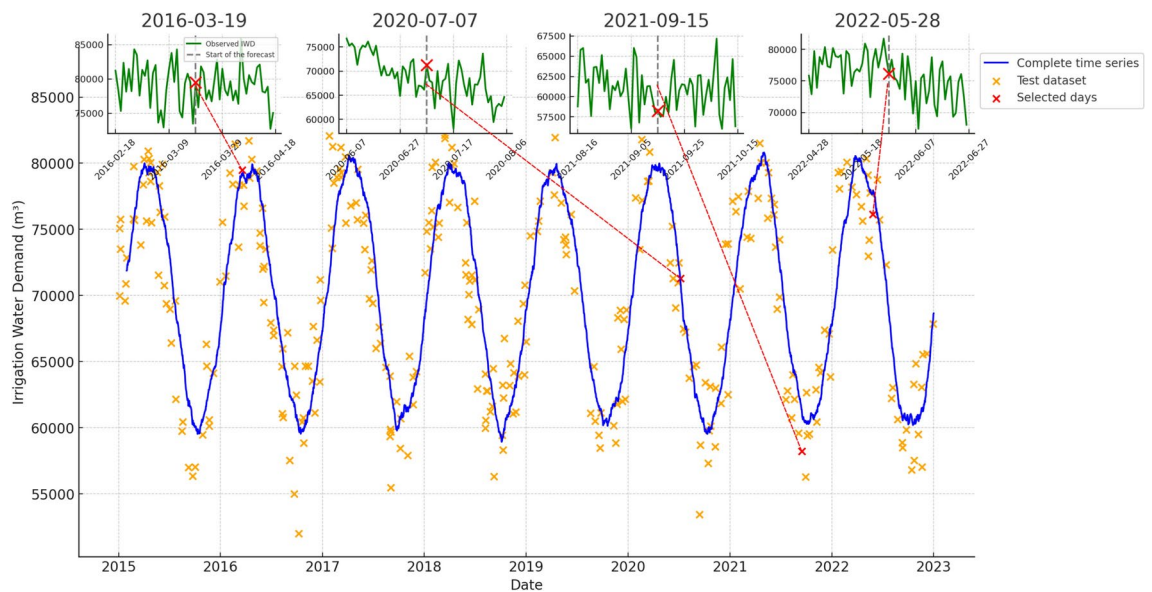
Model	Dataset							
	MODIS dataset				GLDAS dataset			
	RMSE	MAE	$R^2$	MAPE	RMSE	MAE	$R^2$	MAPE
UNet-ConvLSTM	0.307	0.225	0.915	4.098	0.286	0.198	0.923	3.708
Only UNet	0.377	0.299	0.879	5.020	0.365	0.287	0.871	4.870
Only ConvLSTM	0.351	0.274	0.893	4.890	0.332	0.252	0.897	4.610

**Table 3.** Ablation study results for different model configurations on MODIS and GLDAS datasets.



**Figure 2.** Comparison of actual and predicted irrigation water demand over 100 days.

The Fig. 2 illustrates the variation in irrigation water demand over time as predicted by different models. The x-axis (Time/Day) represents the number of days over a specific period, while the y-axis (Irrigation water demand ( $m^3$ )) indicates the daily irrigation water demand. From the figure, it is evident that the predictions made by the UCL model and the Only ConvLSTM model closely match the actual values, indicating that these two models perform well in capturing the fluctuations in irrigation water demand over time. Notably, in the periods of high variability (such as between day 10 to day 30 and day 50 to day 70), the UCL model demonstrates higher predictive accuracy, closely following the actual trend. In contrast, the Only UNet model shows larger prediction errors at certain points (such as around day 20 and day 40), revealing its limitations in handling complex time series data. The actual values exhibit significant variability, reflecting the considerable changes in



**Figure 3.** Time series comparison between actual and predicted agricultural water demand.

irrigation water demand, which poses a higher challenge for predictive models. Overall, the UCL model excels in capturing complex time series patterns and fluctuations, followed by the Only ConvLSTM model. This suggests that the UCL model has a significant advantage among the various baseline models compared, making it the preferred choice for predicting irrigation water demand.

#### Time series comparison of actual and predicted agricultural water demand

Figure 3 illustrates the time series data of irrigation water demand from 2015 to 2023, highlighting significant seasonal variations and annual cycles. The blue curve represents the smoothed complete time series, the orange dots denote the test dataset, and the red dots indicate the selected specific dates. The results demonstrate consistent trends in irrigation water demand during the same periods each year, validating the seasonal patterns and stability of the model. For instance, the demand is typically higher in the spring and summer and lower in the winter, reflecting the agricultural planting cycles and climatic changes. The model effectively captures these seasonal fluctuations, indicating its predictive capability and data rationality. The test dataset covers various seasons and years, ensuring comprehensive performance evaluation under different conditions. By comparing with actual data, the test dataset confirms the model's accuracy and robustness. The detailed views of the selected dates provide further insights into specific periods, such as the rising demand in early spring on March 19, 2016, and the peak demand during the summer on July 7, 2020. The smoothed data highlights overall trends and reduces noise interference, demonstrating the model's excellence in capturing the seasonal and annual cycles of irrigation demand.

#### Discussion and conclusion

In this study, we proposed and evaluated the UCL (UNet-ConvLSTM) model for predicting agricultural water demand by integrating spatial and temporal data from MODIS and GLDAS datasets. The UCL model effectively captures spatial features through the UNet architecture and models temporal dependencies using ConvLSTM layers. Our experimental results show that the UCL model can provide reliable predictions of water demand, demonstrating its potential to adapt to varying agricultural conditions and environmental inputs. The model was validated through extensive testing, ensuring its robustness and applicability in real-world agricultural scenarios. Despite these encouraging results, the UCL model has certain limitations. First, the model's effectiveness is influenced by the variability and heterogeneity of the input data. Different regions may exhibit unique environmental characteristics that are not fully captured by the training data, potentially leading to reduced model accuracy when applied to new or significantly different geographical areas. Additionally, the model's reliance on remote sensing data means that any disruptions or gaps in data acquisition, such as those caused by adverse weather conditions, can impact the consistency and reliability of predictions. Another limitation lies in the model's sensitivity to the temporal resolution of the input data; irregular or sparse data collection intervals could hinder the model's ability to accurately capture temporal dynamics in water demand.

Looking forward, future work should aim to address these limitations and further enhance the model's performance and applicability. One promising direction is to incorporate more diverse data sources, including in-situ measurements and alternative remote sensing datasets, to improve the model's generalization across different regions and conditions. Additionally, efforts could be made to develop more robust strategies for handling missing or irregular data, such as advanced data imputation techniques or adaptive learning algorithms. Another avenue for future research is the refinement of the model's architecture to better accommodate varying temporal resolutions, allowing it to maintain accuracy even with irregular data inputs. The significance of this

research lies in its contribution to the sustainable management of agricultural water resources, particularly in the face of global climate change and water scarcity challenges. By providing accurate and timely predictions of water demand, the UCL model has the potential to optimize irrigation practices, reduce water waste, and enhance the resilience and productivity of agricultural systems on a global scale.

## Related work

### Application of remote sensing technology in agriculture

The application of remote sensing technology in agriculture is extensive. Utilizing data collected from satellites, drones, and other sensor platforms, remote sensing provides crucial support and decision-making basis with large-scale, high-resolution surface information for agricultural production<sup>16,17</sup>. Its main applications include land use and land cover classification, identifying and monitoring various land types such as farmland, forests, and grasslands, aiding in agricultural planning and land management<sup>18</sup>. Additionally, remote sensing monitors and assesses crop growth status, health, and coverage, offering trends and predictions during crop growth seasons to support agricultural management and yield estimation<sup>19</sup>. In irrigation management, it determines optimal irrigation timing and quantity by monitoring soil moisture and crop evapotranspiration rates, enhancing water resource efficiency and reducing wastage. Furthermore, remote sensing data detects and monitors the spread and impact of crop pests and diseases, enabling early detection and response to threats to crops. Lastly, it provides information on soil characteristics such as type, texture, moisture, and fertility, aiding in agricultural soil management and fertilization planning<sup>20</sup>. Overall, the widespread application of remote sensing technology in agriculture contributes to improving productivity, preserving ecological balance, optimizing resource utilization, and promoting sustainable development.

### Soil moisture monitoring technology based on remote sensing

Soil moisture monitoring technology based on remote sensing utilizes remote sensing data obtained from satellites, aircraft, or ground sensors to estimate and monitor soil moisture levels. This technology infers changes in soil moisture by analyzing optical, thermal, or microwave signals in remote sensing data, providing crucial information support for agriculture, water resource management, and ecosystem monitoring. Past research has shown that different types of remote sensing data can be used for soil moisture monitoring. For example, visible and infrared remote sensing data can be used to estimate surface reflectance of the soil, thereby inferring soil moisture<sup>21</sup>. Microwave remote sensing data, on the other hand, can penetrate vegetation and cloud cover, providing estimates of soil moisture at deeper levels<sup>22</sup>. Additionally, thermal infrared remote sensing data can infer soil moisture levels by measuring the surface temperature of the soil, as humidity is related to evaporative cooling at the soil surface<sup>23,24</sup>. In recent years, with the development of remote sensing technology and the integration of multi-source data, remote sensing-based soil moisture monitoring technology has made significant progress. Advanced remote sensing algorithms and models, such as machine learning and deep learning methods, have been widely applied to soil moisture estimation, improving monitoring accuracy and spatial resolution. Furthermore, the open access to satellite remote sensing data and the development of open-source platforms have made soil moisture monitoring technology more accessible and operational<sup>25</sup>.

In summary, remote sensing-based soil moisture monitoring technology has broad application prospects. It can provide important information support for agricultural irrigation management, drought monitoring and early warning, water resource planning, and contribute to addressing climate change and promoting sustainable development.

### Application of deep convolutional network in image segmentation

Deep convolutional neural networks (CNNs) have been extensively researched and applied in the field of image segmentation<sup>26</sup>. Several notable studies have contributed to our understanding of their efficacy and versatility. One significant architecture is the U-Net model, widely employed in medical image segmentation and natural image segmentation tasks<sup>27</sup>. Its unique encoder-decoder structure enables the preservation of fine details in images, resulting in precise segmentation outcomes. Research indicates its success in tumor detection and organ segmentation in medical imaging. Fully Convolutional Networks (FCNs) represent another pivotal development<sup>28</sup>. FCNs are end-to-end architectures particularly suited for image segmentation tasks. By transforming traditional convolutional networks into fully convolutional structures, FCNs enable pixel-level classification of input images, directly outputting segmentation results. They have demonstrated remarkable achievements in urban scene segmentation and land cover classification. Moreover, the Mask R-CNN model integrates object detection and image segmentation<sup>29</sup>. It enhances the Faster R-CNN framework by introducing a segmentation branch, enabling simultaneous detection and segmentation of objects in images. This model has shown robust performance in applications such as autonomous driving and video analysis. Furthermore, optimization efforts have been directed towards various aspects of deep convolutional networks, including training strategies and loss function design. Techniques like multi-scale feature fusion and attention mechanisms have been introduced to further enhance network performance and generalization capabilities<sup>30</sup>.

In summary, past research has demonstrated the effectiveness and broad applicability of deep convolutional neural networks in image segmentation tasks, laying a solid foundation and providing crucial technical support for advancements in the field of image segmentation.

## Methods

### Overview of our network

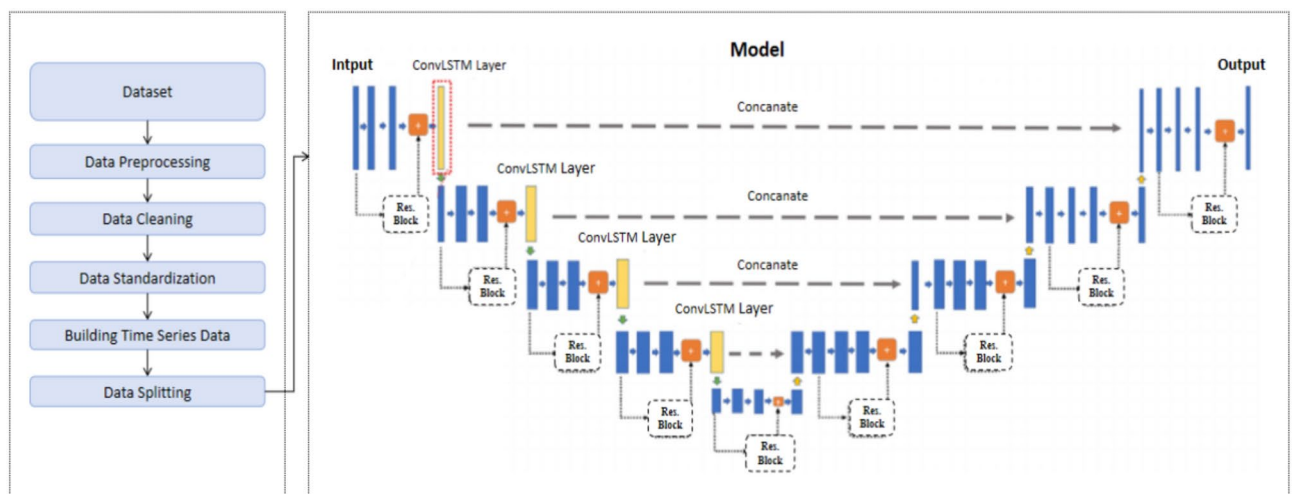
This study proposes a deep learning framework that combines remote sensing technology with UCL(UNet-ConvLSTM), aiming to effectively process remote sensing data for accurately predicting agricultural water

demand, defined as the estimated amount of water required for agricultural purposes, particularly irrigation. This methodology not only optimizes water resource management but also enhances the efficiency and sustainability of food production systems, making it highly relevant to the food sector. The remote sensing data is acquired through high-resolution satellite images that provide key indicators such as soil moisture, crop conditions, and environmental variables. These detailed and comprehensive images serve as the initial input for the deep learning model, ensuring the completeness of information and the timeliness of data. The UNet model is employed to process remote sensing images, utilizing its effective image segmentation capability to identify farmland areas and critical agricultural features. This segmentation process is essential for accurate spatial feature extraction, which directly impacts the precision of water demand predictions. Subsequently, the ConvLSTM module is used to analyze the temporal changes of spatial features processed by UNet. The introduction of this module aims to integrate spatial and temporal dimensions of data, enhancing the predictive ability for future changes in water demand. ConvLSTM is particularly suitable for handling data with temporal continuity, enabling the model to make accurate predictions based on past and current data trends. The overall flow chart is shown in Figure 4.

The network construction begins with the collection and preprocessing of multi-temporal image data obtained from remote sensing technology, which includes normalization and data augmentation to adapt to the requirements of deep learning algorithms. These image data, sourced from MODIS and GLDAS datasets, are preprocessed and then input into the UNet model for detailed spatial analysis. Subsequently, the resulting feature maps are fed into the ConvLSTM network to simulate time-series data and predict future water demand. Our model offers a novel solution for agricultural water resource management by leveraging high-precision remote sensing data and advanced deep learning techniques. It has the capability to predict water demand under various environmental conditions, thereby supporting decision-makers in formulating scientifically informed irrigation strategies. Moreover, the real-time prediction feature of the model enables timely responses to environmental changes, optimizing water resource allocation and enhancing the overall efficiency and sustainability of agricultural production. Through this approach, we not only enhance our understanding of the dynamics of agricultural water resources but also provide an effective tool for their precise management and sustainable utilization.

### Remote sensing technology

Remote sensing technology is a technique used to capture target information from a distance using sensors, commonly employed to acquire data of the Earth's surface. The fundamental principle of remote sensing involves the utilization of sensors to capture electromagnetic waves reflected or emitted from the Earth's surface<sup>31,32</sup>. These sensors can be optical, capturing visible and infrared spectra, or radar systems, utilizing radio waves. Optical sensors identify surface material characteristics by analyzing spectral responses at different wavelengths, while radar sensors acquire information based on the reflection properties of electromagnetic waves<sup>33</sup>. Through analysis of this data, scientists and researchers obtain precise information about surface conditions and changes. In our model, remote sensing technology plays a crucial role in data acquisition, offering real-time, high-resolution, multi-dimensional information. This technology provides continuous global coverage of image data, ensuring the most comprehensive input due to its extensive coverage and high update frequency. Furthermore, modern remote sensing technology provides high-resolution images, enabling the model to accurately identify and analyze subtle changes in farmland, such as crop growth and soil moisture, critical parameters. Additionally, remote sensing images typically contain multiple spectral bands, enhancing our model's feature recognition and analysis capabilities when dealing with complex environmental variables.



**Figure 4.** The overall flow chart.



## Unet model

U-Net is a popular convolutional neural network architecture originally designed for medical image segmentation tasks. Its core feature is its U-shaped structure, consisting of a contracting path (encoder) and a symmetric expanding path (decoder), which connect the high-resolution features of the encoder to the corresponding layers of the decoder through skip connections<sup>34</sup>. The structure diagram of Unetde is presented in Fig. 5. The encoder gradually reduces the spatial dimensions of the image through a series of convolutional and pooling layers, capturing deep semantic information about the image content, while the decoder progressively restores the details and spatial dimensions of the image to match the size of the original input image through upsampling and convolutional layers. The design of skip connections enables the network to utilize shallow detail information during the image segmentation process, enhancing the network's ability to capture image details<sup>35</sup>.

Here's a series of five core equations for the U-Net architecture. The equations include convolution, batch normalization, ReLU activation, downsampling via max-pooling, and the up-convolution operation with concatenation.

Convolution:

$$y_i = (W_i * x + b_i) \quad (1)$$

where  $y_i$  is the output feature map for the  $i$ th filter,  $W_i$  is the weight kernel for the  $i$ th filter,  $b_i$  is the bias term, and  $x$  is the input feature map.

Batch Normalization:

$$\hat{x}_i = \frac{x_i - \mu}{\sqrt{\sigma^2 + \epsilon}} \Rightarrow y_i = \gamma \hat{x}_i + \beta \quad (2)$$

where  $x_i$  is the input to the batch normalization layer,  $\mu$  and  $\sigma^2$  are the mean and variance of the batch,  $\epsilon$  is a small constant for numerical stability, and  $\gamma$  and  $\beta$  are trainable scaling and shifting parameters.

ReLU activation:

$$y = \max(0, x) \quad (3)$$

where  $x$  is the input to the ReLU activation function and  $y$  is the output.

Max pooling (downsampling):

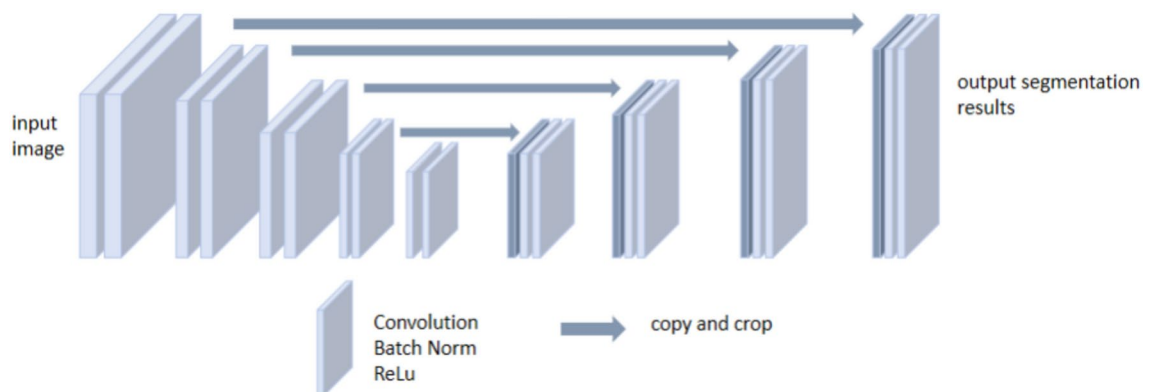
$$y_{i,j} = \max_{(m,n) \in \mathcal{R}} x_{i+m,j+n} \quad (4)$$

where  $y_{i,j}$  is the output of the max-pooling layer at coordinates  $(i, j)$ ,  $x$  is the input feature map, and  $\mathcal{R}$  is the pooling region.

Up-convolution (upsampling):

$$y = \text{ConvTranspose}(x) \Rightarrow y' = \text{Concat}(y, \text{copy\_feature\_map}) \quad (5)$$

where  $x$  is the input to the up-convolution layer,  $y$  is the upsampled feature map, and  $y'$  is the concatenated feature map with the corresponding copy from the contracting path.



**Figure 5.** The structure of Unet model.

In our research, the U-Net model plays a crucial role in processing remote sensing image data, particularly in precise crop and soil region segmentation. Its powerful segmentation capabilities allow for the accurate extraction of detailed information from complex images, forming the basis for precise irrigation strategies in water resource management. Additionally, the U-Net model can handle large-scale images and maintain good performance with fewer training samples, which is particularly important for remote sensing image analysis, as obtaining a large number of high-quality annotated remote sensing images is often challenging. By applying U-Net to our remote sensing data processing pipeline, the model not only improves the accuracy of crop and soil feature identification but also optimizes the input quality of the entire water demand prediction model.

### ConvLSTM model

The ConvLSTM (Convolutional Long Short-Term Memory) model is a type of recurrent neural network (RNN) that integrates convolutional layers into the traditional LSTM architecture, enabling it to handle spatial-temporal data more effectively<sup>36</sup>. In ConvLSTM, the gating mechanism includes three primary gates (forget gate, input gate, and output gate) and a candidate cell state (as shown in Fig. 6, which control the flow of information: Forget Gate determines which information from the previous cell state should be retained. Input Gate controls which new information is updated in the current cell state. Output Gate decides which portion of the current cell state contributes to the hidden state. ConvLSTM replaces the fully connected layers used in traditional LSTMs with convolutional layers in these gate computations. This enables the model to capture spatial features alongside temporal dependencies, making it particularly suitable for applications involving sequence prediction where both spatial and temporal patterns are significant.

Here's the detailed mathematical formulation of the ConvLSTM model with the corresponding equations:

Forget gate calculation:

$$F_t = \sigma(W_f * X_t + U_f * H_{t-1} + b_f) \quad (6)$$

where  $F_t$ : Forget gate output  $W_f$  and  $U_f$ : Convolution kernels  $b_f$ : Bias term  $X_t$ : Current input  $H_{t-1}$ : Previous hidden state  $\sigma$ : Sigmoid activation function  $*$ : Convolution operation

Input gate calculation:

$$I_t = \sigma(W_i * X_t + U_i * H_{t-1} + b_i) \quad (7)$$

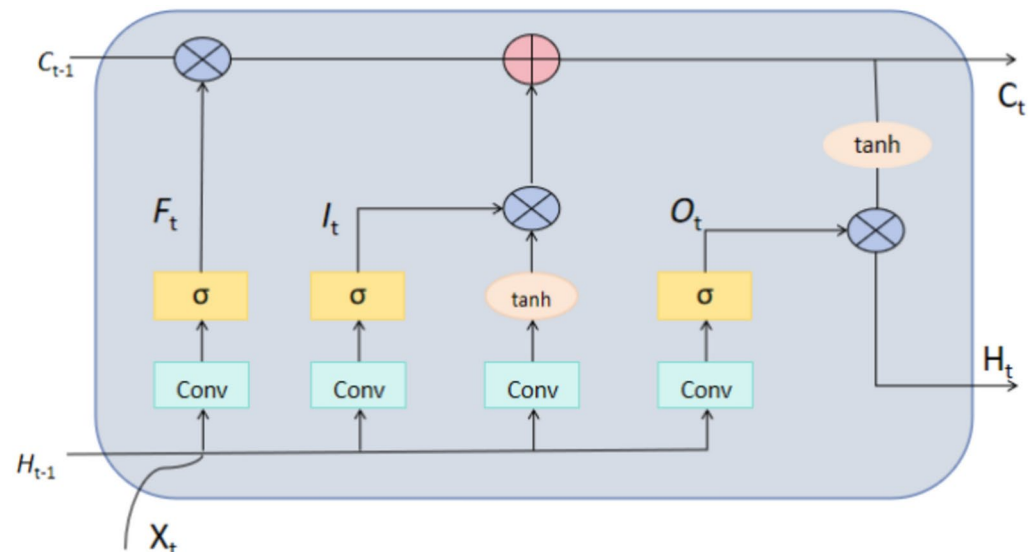
where  $I_t$ : Input gate output  $W_i$  and  $U_i$ : Convolution kernels  $b_i$ : Bias term

Candidate cell state calculation:

$$\tilde{C}_t = \tanh(W_c * X_t + U_c * H_{t-1} + b_c) \quad (8)$$

where  $\tilde{C}_t$ : Candidate cell state  $W_c$  and  $U_c$ : Convolution kernels  $b_c$ : Bias term  $\tanh$ : Hyperbolic tangent activation function

Cell state update:



**Figure 6.** The structure of ConvLSTM model.

$$C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \quad (9)$$

where  $C_t$ : Updated cell state  $C_{t-1}$ : Previous cell state  $\odot$ : Hadamard (element-wise) product

Output gate calculation and hidden state update:

$$O_t = \sigma(W_o * X_t + U_o * H_{t-1} + b_o) \quad (10)$$

$$H_t = O_t \odot \tanh(C_t) \quad (11)$$

where  $O_t$ : Output gate output  $W_o$  and  $U_o$ : Convolution kernels  $b_o$ : Bias term  $H_t$ : Updated hidden state

In our model, the introduction of ConvLSTM significantly enhances the capability to handle time-series of remote sensing images. This module, through its convolutional structure, analyzes temporal sequence changes while preserving spatial information, effectively predicting the dynamic changes in agricultural water demand. This includes adapting to changes in water requirements in different seasons and climatic conditions, which are crucial for devising scientifically informed irrigation plans and optimizing water resource management. Compared to traditional time-series models, ConvLSTM more accurately captures spatiotemporal dynamics, improving the accuracy and efficiency of water demand prediction.

## Experiment

### Datasets

**MODIS dataset<sup>37</sup>:** The Moderate Resolution Imaging Spectroradiometer (MODIS) is an instrument mounted on NASA's Terra and Aqua satellites, designed to observe terrestrial and atmospheric phenomena on a global scale. MODIS provides critical environmental parameters such as surface reflectance, vegetation indices, and surface temperature, with data updated from daily to monthly intervals. In this study, "water demand" is defined as the estimated volume of water required for agricultural purposes, specifically for irrigation. MODIS data, particularly surface temperature and vegetation indices, are utilized to monitor crop growth conditions and assess irrigation needs. These parameters are integral to training and validating the UCL model, as they offer insights into the dynamic relationship between environmental conditions and agricultural water demand. By analyzing the temporal variations captured by MODIS data, we can accurately estimate the fluctuations in water demand across different agricultural regions. The selection of MODIS data is based on its high temporal resolution and comprehensive coverage, which are essential for capturing the temporal variability in water demand.

**GLDAS Dataset<sup>38</sup>:** The Global Land Data Assimilation System (GLDAS) compiles data from multiple satellite sources and ground-based sensors to produce high-resolution hydrological variables such as soil moisture, evapotranspiration, and precipitation on a global scale. GLDAS is known for its fine temporal and spatial resolution, offering data from hourly to monthly intervals. In this study, GLDAS data is crucial for understanding the subsurface moisture dynamics that directly impact "water demand." Soil moisture levels provided by GLDAS are used to calculate the irrigation requirements for crops, which are then integrated into the UCL model to simulate and predict water demand under varying environmental conditions. The inclusion of GLDAS data ensures that the model accurately reflects the complex interactions between soil moisture and atmospheric conditions, leading to more precise water demand predictions and improved recommendations for water resource management. The integration of GLDAS data allows the model to account for both spatial and temporal variations in hydrological conditions, which is crucial for making informed decisions in agricultural water management.

The combined application of these two datasets not only enhances the model's predictive accuracy but also provides robust scientific support for agricultural water resource management. By integrating MODIS and GLDAS data, the UCL model enables more rational water allocation and efficient irrigation practices. By delving into the insights provided by these datasets, researchers and policymakers can better understand and respond to the impacts of global changes on agricultural water resources.

### Experimental environment

Our experimental environment consisted of the following hardware and software configurations:

- **Hardware:** The experiments were conducted on a system equipped with an Intel Xeon E5-2680 v4 CPU, an NVIDIA Tesla V100 GPU with 32GB memory, 256GB DDR4 RAM, and a 2TB SSD for storage.
- **Software:** The software environment included Ubuntu 20.04 LTS as the operating system, Python 3.8 as the programming language, and TensorFlow 2.4 as the primary deep learning framework. Additionally, we utilized libraries such as NumPy, pandas, Matplotlib, and scikit-learn for data processing and model training. These configurations ensured efficient handling and analysis of large remote sensing datasets and the effective training and evaluation of deep learning models.

### Experimental details

Step 1: Data preprocessing

- **Data cleaning:** in this study, data cleaning is a crucial step to ensure the input of high-quality data into the UCL model. Initially, for the MODIS dataset, missing values, especially those caused by cloud cover, are addressed through linear interpolation. This method takes advantage of the continuity in the time series by estimating the missing values based on the pixel values at the time points before and after the gap. For the

GLDAS dataset, which generally has more continuous data, we opt to use the average values from adjacent time points for imputation, maintaining consistency and continuity in the dataset. Additionally, we have rigorously handled outliers within the datasets. By setting thresholds, we identify and address irregular values, such as anomalously high temperatures or low humidity levels. Outliers are replaced using the median of the neighboring normal values, ensuring that the dataset remains robust and reliable.

- **Data standardization:** all numerical features, such as surface temperature, vegetation index, and soil moisture, are processed using Min-Max normalization, scaling the data to a range from 0 to 1. This treatment helps to accelerate the convergence of neural networks and prevents issues such as vanishing or exploding gradients.
- **Building time series data:** for each prediction target, we construct feature sets based on a time window (every 5 days), each window including historical data from the past 10 days. This approach captures both long-term and short-term environmental changes, providing sufficient temporal information for the ConvLSTM layers. Data within the time window is aggregated by calculating the average value at each time point, which reduces the computational load on the model.
- **Data splitting:** the dataset is divided into training, validation, and test sets at a ratio of 70%, 15%, and 15% respectively. This split ensures that there is sufficient data for the model to learn during the training process, while also reserving independent data for performance validation and final testing. To avoid temporal data leakage, it is ensured that the data points in the test set occur after those in the training and validation sets.
- **Experimental Region Description:** The Central Valley of California is one of the most productive agricultural regions in the United States. This region is characterized by a Mediterranean climate with hot, dry summers and mild, wet winters. Major crops include fruits, vegetables, nuts, and cotton, with extensive use of irrigation systems. The availability of comprehensive remote sensing data from NASA Earthdata and detailed agricultural water use records from the USDA make this region an ideal candidate for our study.
- **Step2: Model training**
  - **Model architecture design:** the design of the UNet portion includes a four-layer encoder and decoder structure. Each encoder layer is equipped with two convolutional layers and one max pooling layer, which helps extract spatial features and reduce the spatial dimensionality of the features, increasing the model's capacity for abstraction. For the MODIS dataset, the input size is  $[H \times W \times 2]$ , where  $H$  and  $W$  represent the image dimensions, and 2 represents the feature channels (surface temperature and vegetation index). For the GLDAS dataset, the input size is  $[H \times W \times N]$ , where  $N$  is the number of hydrological parameters. The decoder part gradually restores the details and dimensions of the image through upsampling and convolution operations. Each decoder layer also contains two convolutional layers and one upsampling layer to enhance the resolution of the feature maps. The output from the final encoder layer has a size of  $[H' \times W' \times D]$ , where  $H'$  and  $W'$  are the reduced spatial dimensions and  $D$  is the feature depth. These are then passed into the ConvLSTM layers with input size  $[T \times H' \times W' \times D]$ , where  $T$  represents the time steps. To capture the dynamics related to temporal changes, the network incorporates two layers of ConvLSTM, each with 64 hidden units. This configuration allows the model to effectively learn complex dependencies within the time series data.
  - **Model inputs and outputs:** when the UCL model is applied to the MODIS dataset, the inputs consist of surface temperature and vegetation indices. These parameters are critical for assessing crop health and evapotranspiration rates, which are directly related to the water demand of crops. The output of the model in this case is the predicted agricultural water demand, which indicates the amount of irrigation required based on the analyzed environmental conditions. For the GLDAS dataset, the inputs include soil moisture levels, precipitation, and other hydrological parameters. These inputs provide comprehensive information on water availability in the soil, which is essential for determining the irrigation needs of crops. The output when using GLDAS data is the prediction of soil moisture levels and the corresponding irrigation requirements, helping to ensure that crops receive the necessary water for optimal growth. This clear distinction of inputs and outputs for each dataset ensures that the UCL model is properly tailored to the specific characteristics and needs of the data, enhancing the model's accuracy in predicting agricultural water requirements.
  - **Training Parameters:** We have opted for a relatively low learning rate of 0.001 to stabilize the optimization process, and have chosen the Adam optimizer, which integrates the advantages of momentum and adaptive learning rates, making it suitable for handling large-scale parameter models. The batch size is set at 32 to balance computational efficiency and memory usage, ensuring effective gradient calculations while avoiding memory overflow. These configurations help the model more accurately predict agricultural water requirements.
  - **Model training and validation process:** the dataset was initially split into 70% for training and validation and 30% as an independent test set. Within the 70% training and validation data, we implemented 5-fold cross-validation to assess the model's generalizability and robustness. This approach ensured that different subsets of the training data took turns serving as the validation set, allowing the model to exhibit stable performance under various data conditions. The test set was kept independent and was not involved in the cross-validation process, ensuring that the final model evaluation was based on data unseen during the training and validation phases. This process meticulously optimized the model, ensuring its effectiveness in practical applications.
- **Step3: Model evaluation:** the effectiveness of the proposed model was evaluated using several performance metrics, including Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination ( $R^2$ ), and Mean Absolute Percentage Error (MAPE). These metrics were chosen to provide a comprehensive assessment of the model's accuracy and reliability in predicting agricultural water demand.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (13)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (14)$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (15)$$

where  $n$  is the number of observations,  $y_i$  is the actual value,  $\hat{y}_i$  is the predicted value, and  $\bar{y}$  is the mean of the actual values.

These metrics collectively provide a comprehensive evaluation of the model's performance in predicting agricultural water demand, ensuring both accuracy and reliability in different aspects of the predictions.

### Data availability

The data and materials used in this study are not currently available for public access. Interested parties may request access to the data by contacting the corresponding author.

Received: 25 June 2024; Accepted: 17 October 2024

Published online: 28 October 2024

### References

1. Mainardis, M. et al. Wastewater fertigation in agriculture: Issues and opportunities for improved water management and circular economy. *Environ. Pollut.* **296**, 118755 (2022).
2. Bwambale, E., Abagale, F. K. & Anornu, G. K. Smart irrigation monitoring and control strategies for improving water use efficiency in precision agriculture: A review. *Agric. Water Manag.* **260**, 107324 (2022).
3. Li, M. et al. Sustainable management of agricultural water and land resources under changing climate and socio-economic conditions: A multi-dimensional optimization approach. *Agric. Water Manag.* **259**, 107235 (2022).
4. Awais, M. et al. Uav-based remote sensing in plant stress imagine using high-resolution thermal sensor for digital agriculture practices: A meta-review. *Int. J. Environ. Sci. Technol.* 1–18 (2022).
5. Kumar, S. et al. Remote sensing for agriculture and resource management. In *Natural Resources Conservation and Advances for Sustainability*. 91–135 (Elsevier, 2022).
6. Mandal, A. et al. Impact of agricultural management practices on soil carbon sequestration and its monitoring through simulation models and remote sensing techniques: A review. *Crit. Rev. Environ. Sci. Technol.* **52**, 1–49 (2022).
7. Cheng, C., Zhang, F., Shi, J. & Kung, H.-T. What is the relationship between land use and surface water quality? A review and prospects from remote sensing perspective. *Environ. Sci. Pollut. Res.* **29**, 56887–56907 (2022).
8. Zeng, Y., Guo, Y. & Li, J. Recognition and extraction of high-resolution satellite remote sensing image buildings based on deep learning. *Neural Comput. Appl.* **34**, 2691–2706 (2022).
9. Gao, P. et al. Ndvi forecasting model based on the combination of time series decomposition and cnn-lstm. *Water Resour. Manag.* **37**, 1481–1497 (2023).
10. Li, K., Zhao, W., Peng, R. & Ye, T. Multi-branch self-learning vision transformer (MSVIT) for crop type mapping with optical-SAR time-series. *Comput. Electron. Agric.* **203**, 107497 (2022).
11. Lu, Y., Chen, D., Olaniyi, E. & Huang, Y. Generative adversarial networks (GANS) for image augmentation in agriculture: A systematic review. *Comput. Electron. Agric.* **200**, 107208 (2022).
12. Kashyap, P. K., Kumar, S., Jaiswal, A., Prasad, M. & Gandomi, A. H. Towards precision agriculture: IOT-enabled intelligent irrigation systems using deep learning neural network. *IEEE Sens. J.* **21**, 17479–17491 (2021).
13. Raei, E. et al. A deep learning image segmentation model for agricultural irrigation system classification. *Comput. Electron. Agric.* **198**, 106977 (2022).
14. Chen, H. et al. A deep learning CNN architecture applied in smart near-infrared analysis of water pollution for agricultural irrigation resources. *Agric. Water Manag.* **240**, 106303 (2020).
15. Dang, C. et al. Iwram: A hybrid model for irrigation water demand forecasting to quantify the impacts of climate change. *Agric. Water Manag.* **291**, 108643 (2024).
16. Melton, F. S. et al. Openet: Filling a critical data gap in water management for the western United States. *JAWRA J. Am. Water Resour. Assoc.* **58**, 971–994 (2022).
17. Kadri, N., Ellouze, A., Ksantini, M. & Turki, S. H. New lstm deep learning algorithm for driving behavior classification. *Cybern. Syst.* **54**, 387–405 (2023).
18. Wolff, W., Francisco, J. P., Flumignan, D. L., Marin, F. R. & Folegatti, M. V. Optimized algorithm for evapotranspiration retrieval via remote sensing. *Agric. Water Manag.* **262**, 107390 (2022).
19. Li, W. et al. A UAV-aided prediction system of soil moisture content relying on thermal infrared remote sensing. *Int. J. Environ. Sci. Technol.* **19**, 9587–9600 (2022).
20. Cheng, M. et al. Using multimodal remote sensing data to estimate regional-scale soil moisture content: A case study of Beijing, China. *Agric. Water Manag.* **260**, 107298 (2022).
21. Rasheed, M. W. et al. Soil moisture measuring techniques and factors affecting the moisture dynamics: A comprehensive review. *Sustainability*. **14**, 11538 (2022).
22. Kisekka, I. et al. Spatial-temporal modeling of root zone soil moisture dynamics in a vineyard using machine learning and remote sensing. *Irrig. Sci.* **40**, 761–777 (2022).
23. Das, B. et al. Comparison of bagging, boosting and stacking algorithms for surface soil moisture mapping using optical-thermal-microwave remote sensing synergies. *Catena*. **217**, 106485 (2022).
24. Wang, J., Li, F., An, Y., Zhang, X. & Sun, H. Towards robust lidar-camera fusion in BEV space via mutual deformable attention and temporal aggregation. *IEEE Trans. Circuits Syst. Video Technol.* 1–1. <https://doi.org/10.1109/TCSVT.2024.3366664> (2024).
25. Wang, C. et al. Learning discriminative features by covering local geometric space for point cloud analysis. *IEEE Trans. Geosci. Remote Sens.* **60**, 1–15 (2022).
26. Choudhary, K. et al. Recent advances and applications of deep learning methods in materials science. *npj Comput. Mater.* **8**, 59 (2022).

27. Chen, X. et al. Recent advances and clinical applications of deep learning in medical image analysis. *Med. Image Anal.* **79**, 102444 (2022).
28. Li, Y. Research and application of deep learning in image recognition. In *2022 IEEE 2nd International Conference on Power, Electronics and Computer Applications (ICPECA)*. 994–999 (IEEE, 2022).
29. Gite, S., Mishra, A. & Kotecha, K. Enhanced lung image segmentation using deep learning. *Neural Comput. Appl.* **35**, 22839–22853 (2023).
30. Ning, X. et al. Hyper-sausage coverage function neuron model and learning algorithm for image classification. *Pattern Recognit.* **136**, 109216 (2023).
31. Zhang, L. & Zhang, L. Artificial intelligence for remote sensing data analysis: A review of challenges and opportunities. *IEEE Geosci. Remote Sens. Mag.* **10**, 270–294 (2022).
32. Rahaman, M. et al. Effects of label noise on performance of remote sensing and deep learning-based water body segmentation models. *Cybern. Syst.* **53**, 581–606 (2022).
33. Zhang, B. et al. Progress and challenges in intelligent remote sensing satellite systems. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **15**, 1814–1822 (2022).
34. Beeche, C. et al. Super u-net: A modularized generalizable architecture. *Pattern Recognit.* **128**, 108669 (2022).
35. Punn, N. S. & Agarwal, S. Modality specific u-net variants for biomedical image segmentation: A survey. *Artif. Intell. Rev.* **55**, 5845–5889 (2022).
36. Yadav, S. K., Tiwari, K., Pandey, H. M. & Akbar, S. A. Skeleton-based human activity recognition using convlstm and guided feature learning. *Soft Comput.* **26**, 877–890 (2022).
37. Yu, P. et al. Global spatiotemporally continuous modis land surface temperature dataset. *Sci. Data.* **9**, 143 (2022).
38. Ramirez, S. G., Hales, R. C., Williams, G. P. & Jones, N. L. Extending SC-PDSI-PM with neural network regression using gldas data and permutation feature importance. *Environ. Model. Softw.* **157**, 105475 (2022).

### Author contributions

Zhigang Ye: Writing - original draft, Methodology; Shan Yin: Writing -review & editing, Validation; Yin Cao: Data curation; Yong Wang:Supervision, Project administration.

### Funding

This work was supported by the Erdos City Major Science and Technology Special Project (No. ZD20232323) and the Key Project of the Natural Science Foundation of the Inner Mongolia Autonomous Region (No. 2023ZD21).

### Declarations

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to Y.W.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024