首都师范大学

Integrating neural network, InSAR, and field measurements to model land subsidence evolution in the Beijing Plain

Lin Zhu

College of Resource Environment and Tourism, Capital Normal University
Key Laboratory of Land Subsidence Mechanism and Mitigation, Ministry of Education
Beijing Laboratory of Water Resource Security

February 27 CNU, Beijing

OUTLINE

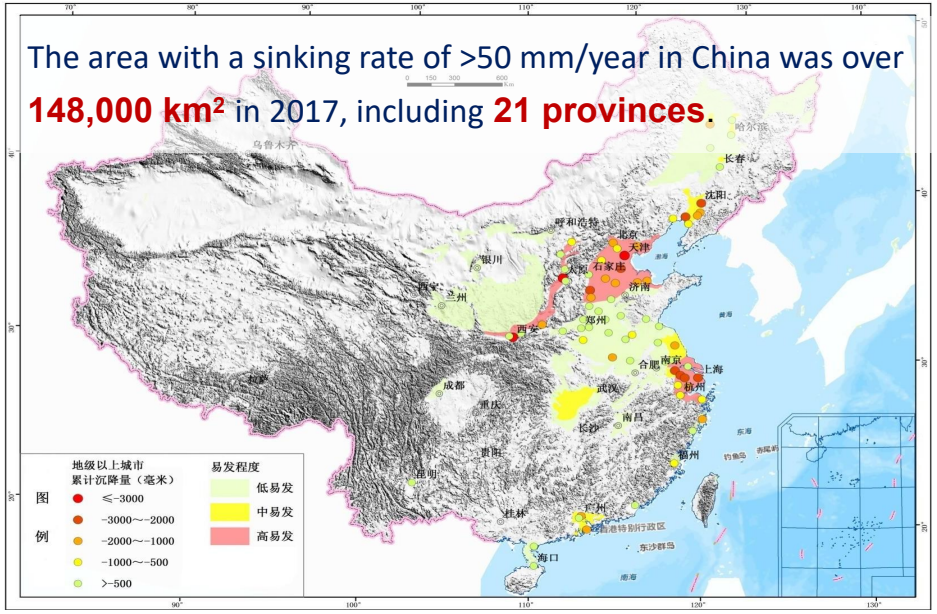
1. Background

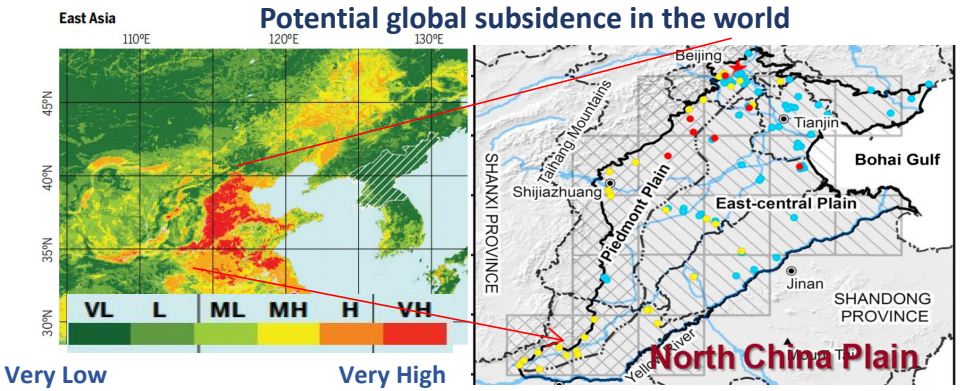
2. Methodology

3. Case Study

1.1 Land subsidence in the Beijing plain

Beijing plain lies in the northwest of the North China Plain (NCP), which has one of the world's most depleted aquifer systems and severe land subsidence due to intensive urbanization and irrigattion with groundwater overuse.





North China Plain (~ 300,000 km²)

- **Largest groundwater depression cone**

  Over-exploitation of groundwater from 2014 to 2020: **2.86 billion m³/yr**

  Area of groundwater funnel in deep aquifer: **20,000 km²**

- **Largest land subsidence region in the world**

  Area with cumulative subsidence greater than 1500 mm since 1930s and subsidence ratio greater than 50 mm in 2019: **14,000 km²**

- **Water Diversion Project**

  South-to-North Water Diversion Project operating from December 2014

  Till December 2025 the cumulative water volume: about **84 billion m³**

- **Groundwater treatment Project**

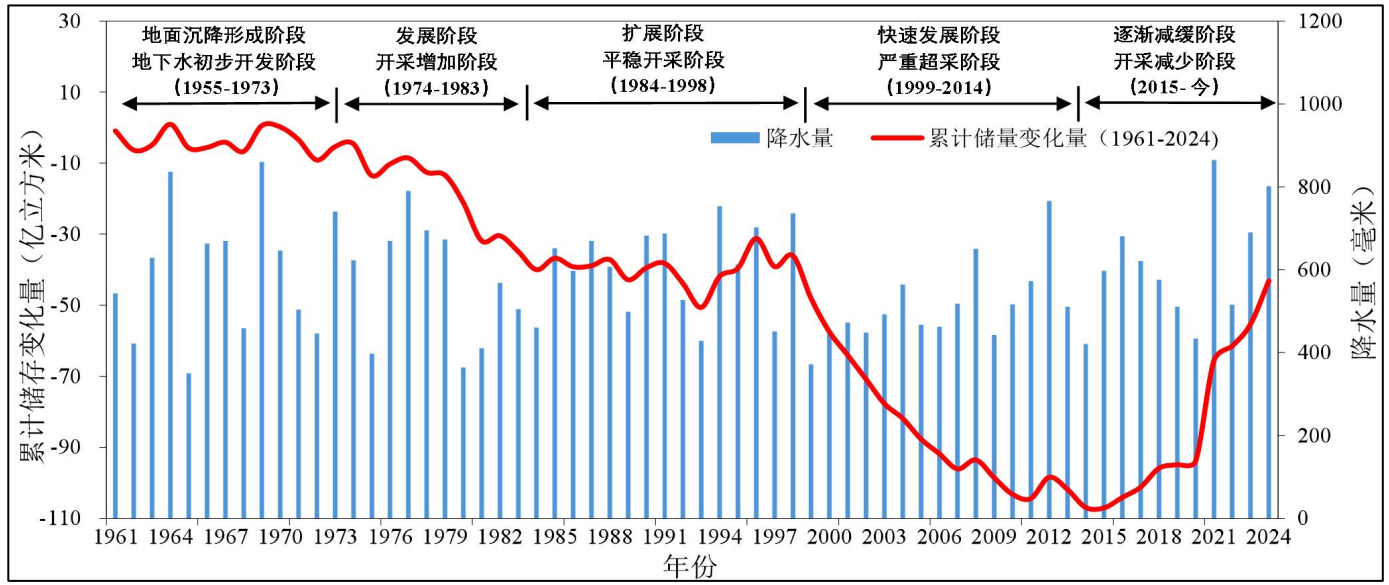
  Reduce groundwater extraction

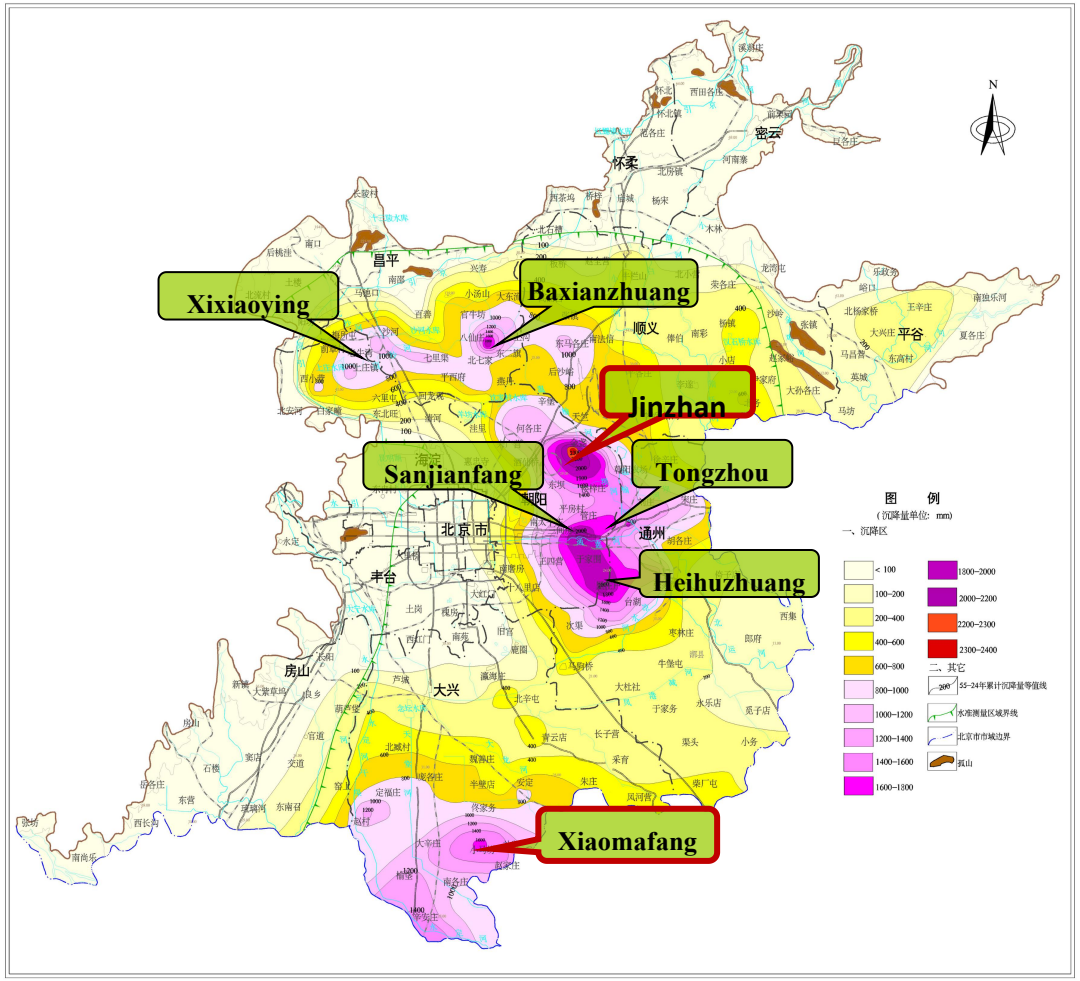
  ecological water replenishment since 2015

- Land subsidence in the Beijing Plain began in the 1930s, and has gone through ***five stages***: formation (1955-1973), development (1974-1983), expansion (1984-1998), rapid development (1999-2014), and gradual slowing down (from 2015 to present).
- The maximum annual subsidence occurred in 2012 with the value of ***159.6*** mm.

***There are two (North and south) major subsidence zones including seven subsidence centers*.**

- In "North" subsidence zone, the maximum cumulative settlement is 2348mm located in ***Chaoyang Jinzhan*.**
- In "South" subsidence zone the value is 1641mm, located in ***Daxing Lixian Xiaomafang***.

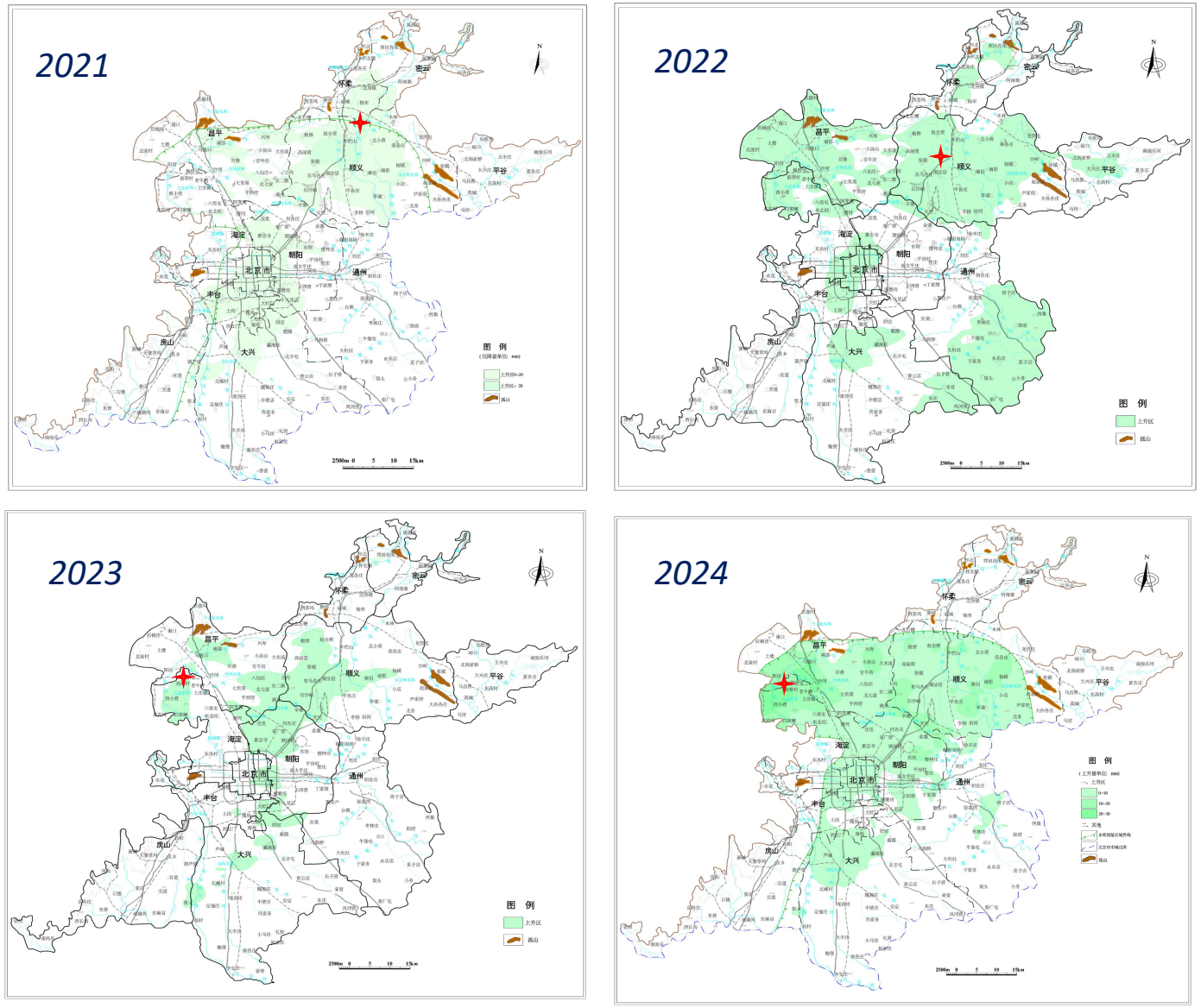




Distribution of cumulative land subsidence from 1955 to 2024

- Since 2021, due to the rapid rise of groundwater level, the Beijing plain has experienced significant ***surface uplift***.
- The area of surface uplift in 2024 is 2228 km², with the maximum annual uplift of 26 mm, located in ***Haidian Xixiaoying***.
- At present, Beijing presents a ***new pattern of coexistence of land subsidence and surface uplift***.

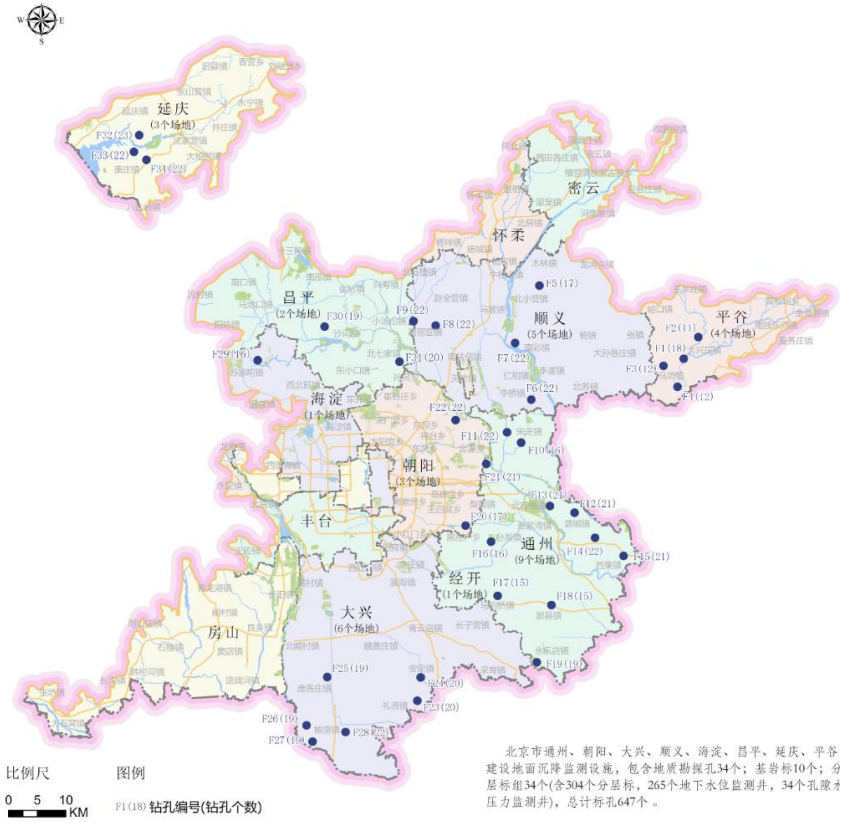
Year	Range/mm	Area/km²	Maxium value and the location
2021	0-20	1318.95	32.20mm Niulanshan, Shunyi District
	≥20	104.24	
2022	0-10	1727.18	46.90mm Nanfaxing,Shunyi District
	10-20	356.77	
	≥20	341.07	
2023	0-10	80.00	19.00mm Xixiaoying,Haidian District
	10-20	754.00	
2024	0-10	1735.39	26.80mm Xixiaoying,Haidian District
	10-20	419.68	
	20-30	73.21	



Distribution of surface uplift from 2021 to 2024

1.2 Field measurement of land subsidence

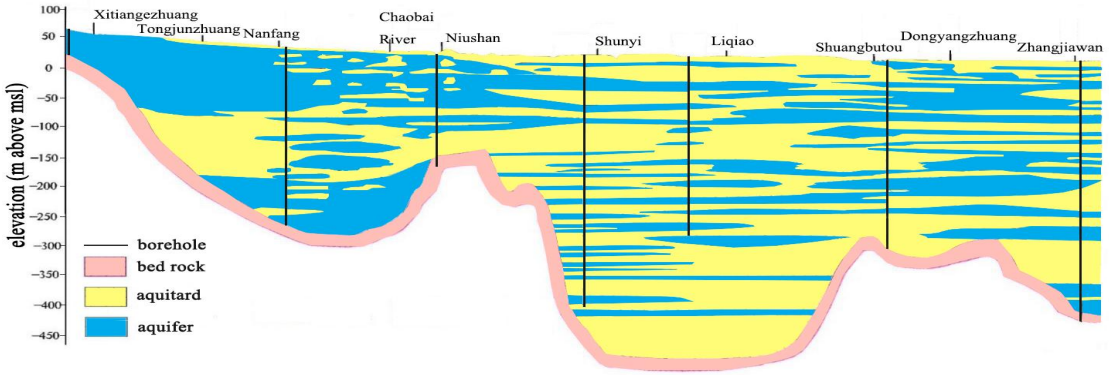
- ***Seven extensomters*** were installed in the Beijing Plain to monitor the displacement indifferent lithological layers since 2004.
- In 2024, the Beijing Institute of Geo-Environment Monitoring constructed ***10 new bedrock extensometer*** and ***34*** layered extensometer groups, and renovated ***2*** bedrock extensometer and ***2*** layered extensometer groups.



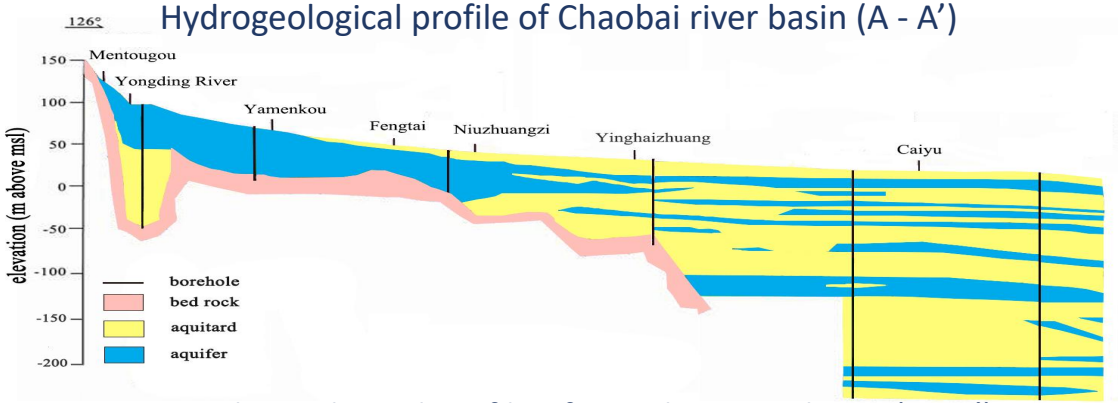
Distribution of the new built extensomter in Beijing plain

1.3 Hydrogeological condition

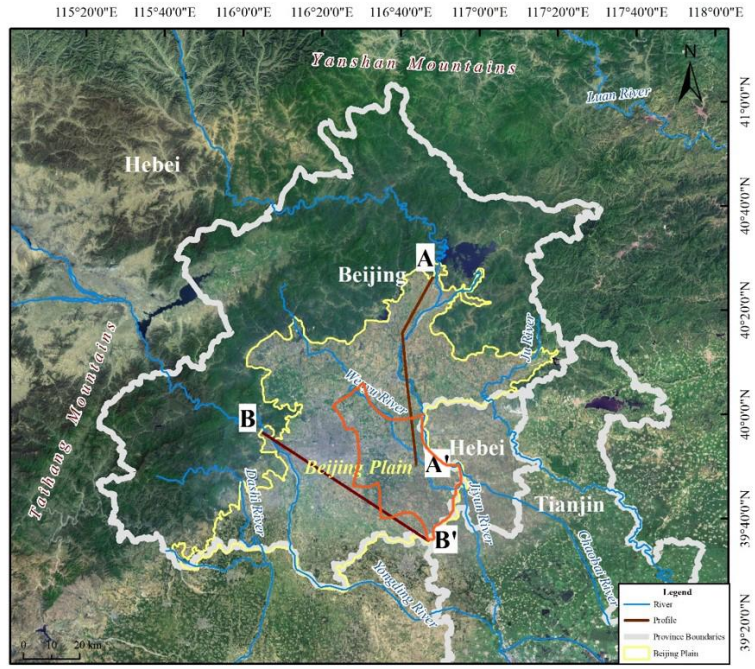
- The Beijing plain is composed of ***five alluvial plains***.
- From northwest to southeast, Quaternary sediments thicken, grain size changes from coarse to fine, and aquifers transit from a ***single sandy gravel layer to interbedded sand-clay multi-layers***.
- Vertically, the Quaternary deposits are divided into ***four aquifers*** .



Hydrogeological profile of Chaobai river basin (A - A')



Hydrogeological profile of Yongding river basin (B - B')



Beijing plain area (Orange is severe land subsidence area)

The division of aquifers in the Beijing Plain (*Chen et al., 2020*)

Aquifer	Major lithology	Depth of bottom (m)
First aquifer group (The unconfined aquifer)	Silt, silty sand, and sandy clay	0 ~ -50
Second aquifer group (The first confined aquifer)	Multiple types of gravel, sand , and clay soil	-50 ~ -100
Third aquifer group (The second confined aquifer)	Multiple types of gravel, sand , and clay soil	-100 ~ -180
Fourth aquifer group (The third confined aquifer)	Mainly sand	-180 ~ -300

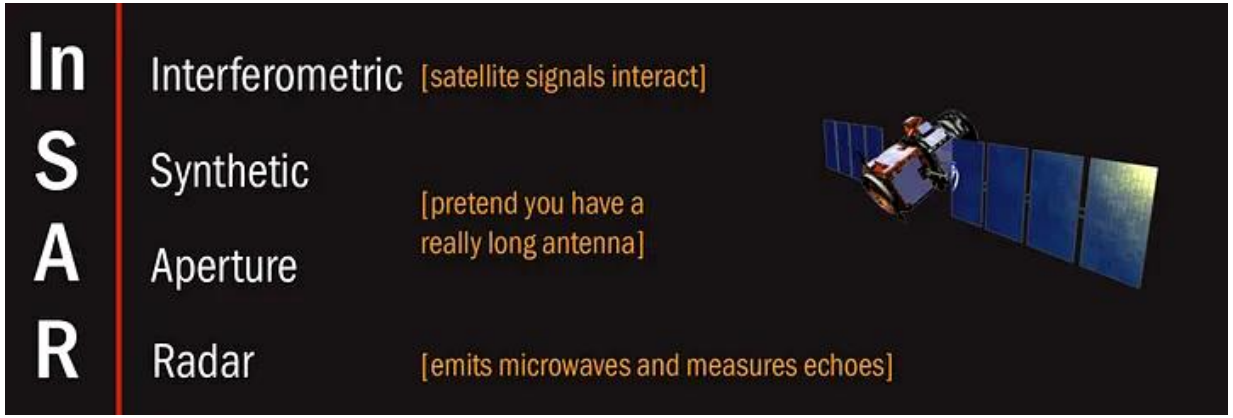
OUTLINE

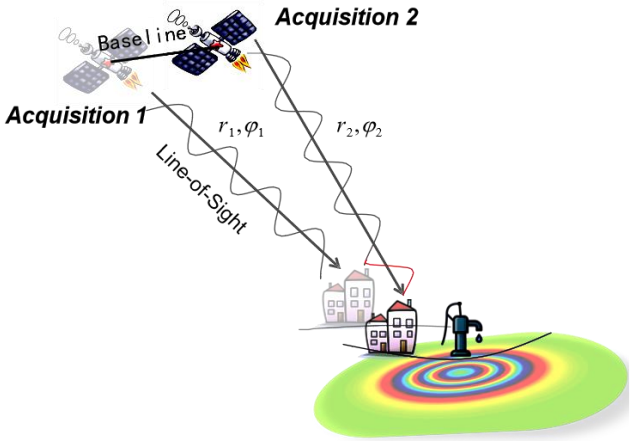
1. Background

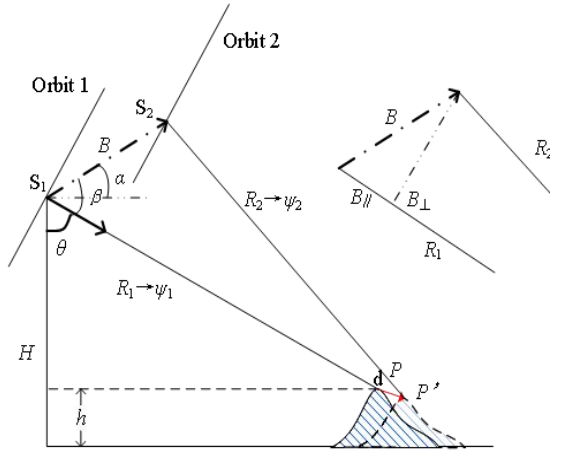
2. Methodology

3. Case Study

2.1 InSAR technology







InSAR measures surface displacement by analyzing phase differences of radar signals before and after ground movement.

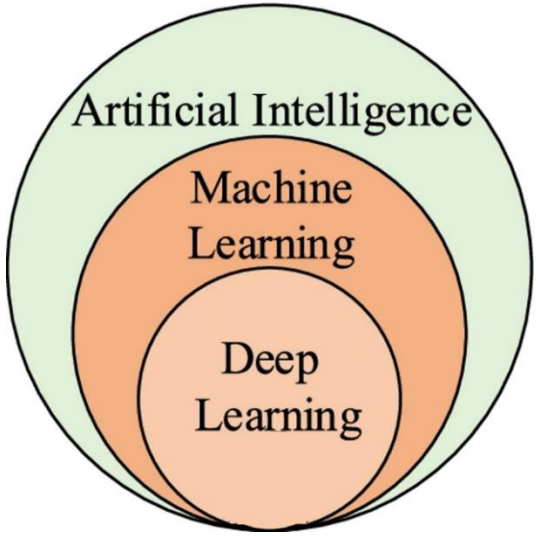
The technology can be used to obtain the regional distribution of land subsidence with high accuracy.

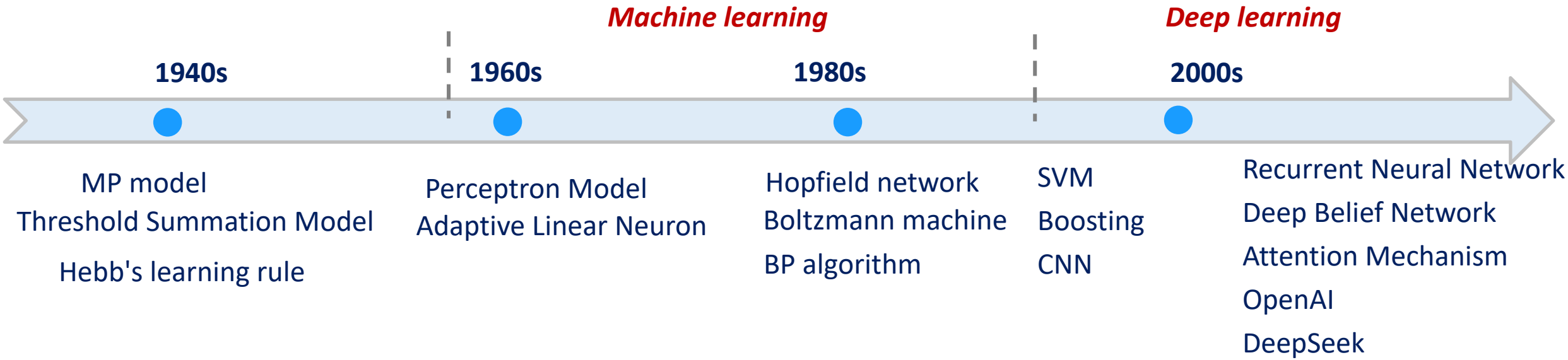
- Small baseline subset (SBAS) technology
- Persistent Scatterer Interferometric Synthetic Aperture Radar (PS-InSAR) technology

$$\Delta\varphi = \Delta\varphi_{flat} + \Delta\varphi_{elevation} + \Delta\varphi_{displacement} + \Delta\varphi_{atmosphere} + \Delta\varphi_{noise}$$

2.2 Artification Intelligence (AI) technology

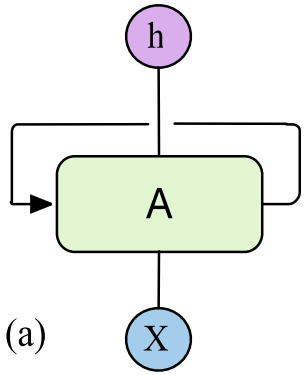
- ***Artificial Intelligence*** is a field of computer science focused on creating systems that can perform tasks typically requiring human intelligence, such as learning, problem-solving, and decision-making.
- ***Machine learning*** enables computers to identify patterns and relationships within a dataset. The main type are supervised learning, unsupervised learning and reinforcement learning.
- ***Deep learning*** uses artificial neural networks **with multiple layers** to automatically learn hierarchical representations of data.



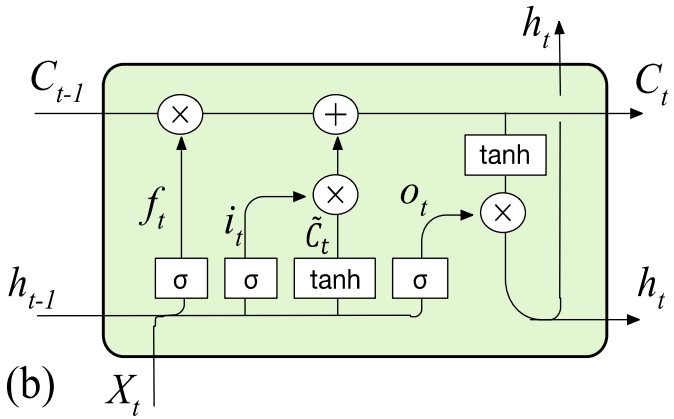


2.3 Long Short-Term Memory (LSTM)

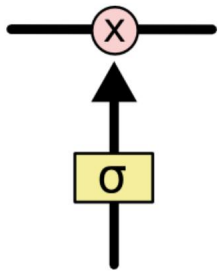
- Recurrent Neural Network (RNN, Jeffrey Elman 1990) is a type of neural network with internal loop that ***processes sequential data through recurrent connections***, enabling information from previous inputs to influence the current output.
- LSTM was proposed by Sepp Hochreiter and Jürgen Schmidhuber in 1997, which adopts gating mechanism (forget gate f_t, input gate i_t, output gate o_t) to ***overcome gradient vanishing or gradient exploding shortcomings*** that occur in RNN processing long sequential data.
- ***Forget gate*** determines ***what information is discarded*** from the cell state and how much new information is to be incorporated into the cell state.
- The gate structures can achieve information filtering through a sigmoid activation function (σ) and a pointwise multiplication operation.



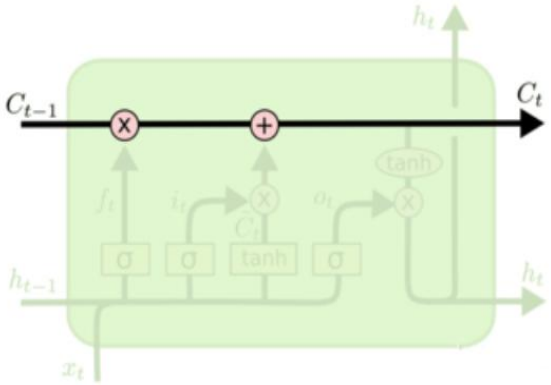
Structure of RNN



Structure of LSTM



Gate structure

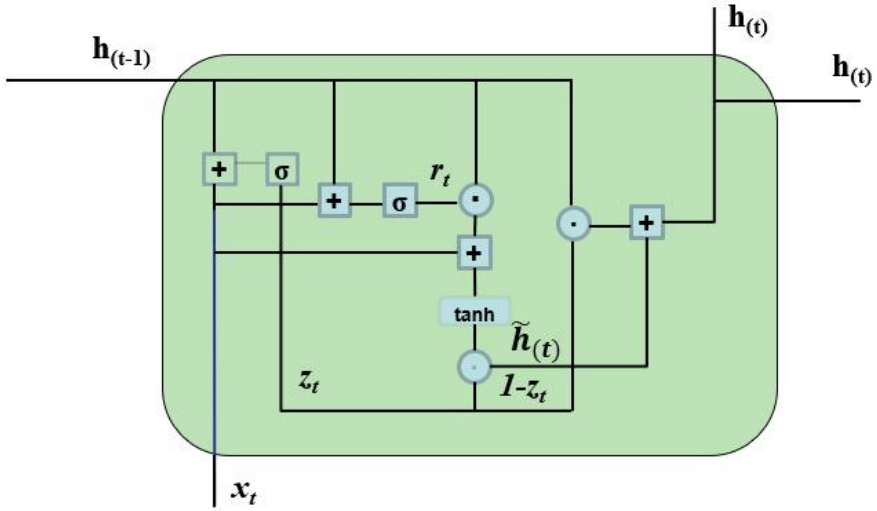


Information transmission line

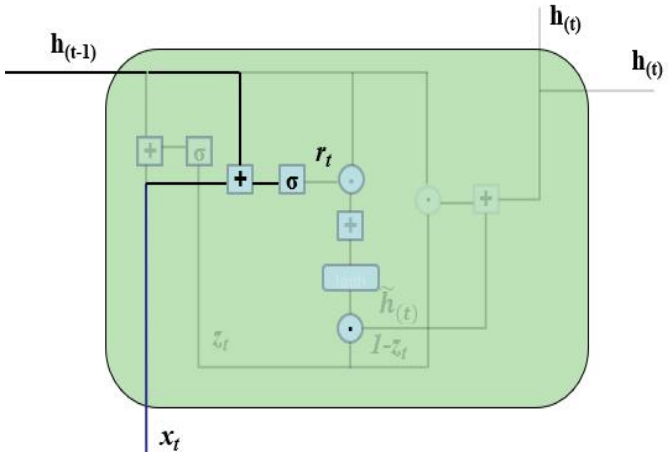
Li & Zhu et al 2020, PIAHS

2.4 Gate Recurrent Unit (GRU)

- GRU was proposed by Kyunghyun Cho et al. in 2014. Compared to LSTM, GRU structure employs ***update gate*** (z_t) and ***reset gate*** (r_t), which improves the model training efficiency.
- Update gate determines how much information **from previous hidden state should be retained and passed to the current time step**.
- Reset gate determines **how much information from previous hidden state should be ignored**, where a smaller value of the reset gate indicates a greater degree of neglect. When the input x_t is fed in, it is multiplied by the learned weight W_r and h_{t-1} is multiplied by its own weight U_r.

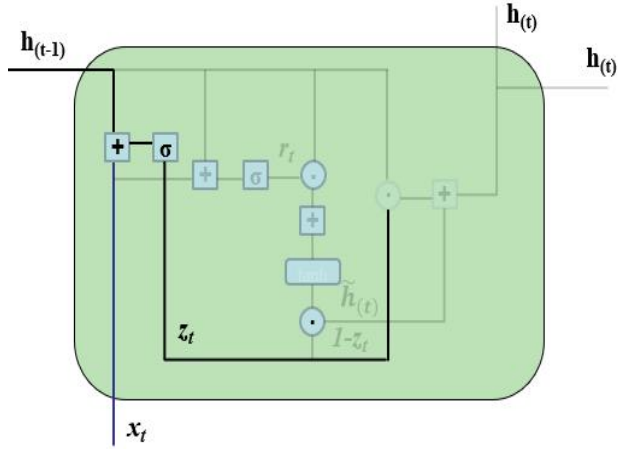


Structure of the GRU



Reset Gate

$$r_t = \sigma(W_r x_t + U_r h_{(t-1)})$$

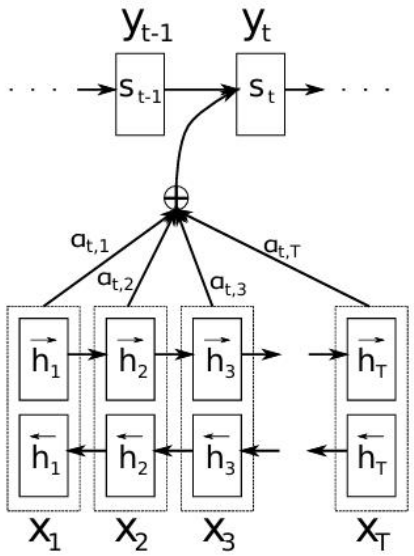


Update Gate

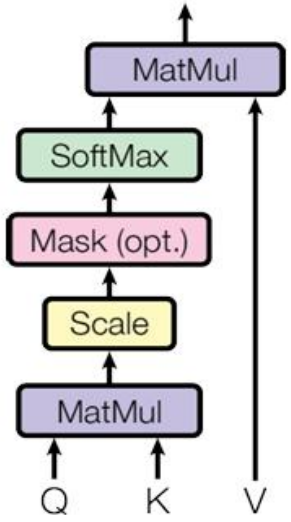
$$z_t = \sigma(W_z x_t + U_z h_{(t-1)})$$

2.5 Attention Mechanism (AM)

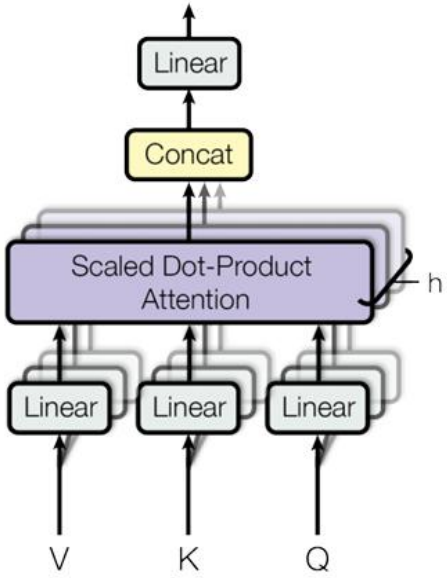
- ***AM*** (Bahdanau et al. 2014) allows the model to dynamically focus on the most relevant parts of the input data, instead of treating all input information equally. The attention weights provide a form of interpretability.
- ***Self-Attention Mechanism*** (Ashish Vaswani et al. 2017) allows each element in the sequence to attend to all other elements (including itself) to capture internal dependencies.
- ***Multi-head Self-attention mechanism:*** decompose self-attention into n independent "*sub-attention heads*" in parallel. Each head calculates different attention weights, and finally fuses the results of all heads. It has the advantage of enhancing the model's ability to understand complex sequences.



Attention mechanism



Self-Attention mechanism



Multi-head Self-Attention mechanism

2.6 Physics law contratined RNN

- To overcome the limitation of RNNs relying solely on historical data patterns and lacking physical constraints, the **effective stress principle is incorporated into RNN**.

RNN:
state variables of hidden layer
output variables

$$S_t = \sigma_S(W_S^T x_t + U_S^T S_{t-1} + b_S)$$

$$y_t = \sigma_y(W_S^T S_t + b_y)$$

Modified RNN:

$$S_t = \sigma_S(S_{t-1} + \frac{W_f^T \cdot x_t}{E} + b_f)$$

$$y_t = \sigma_y(W_g \cdot S_t + b_g)$$

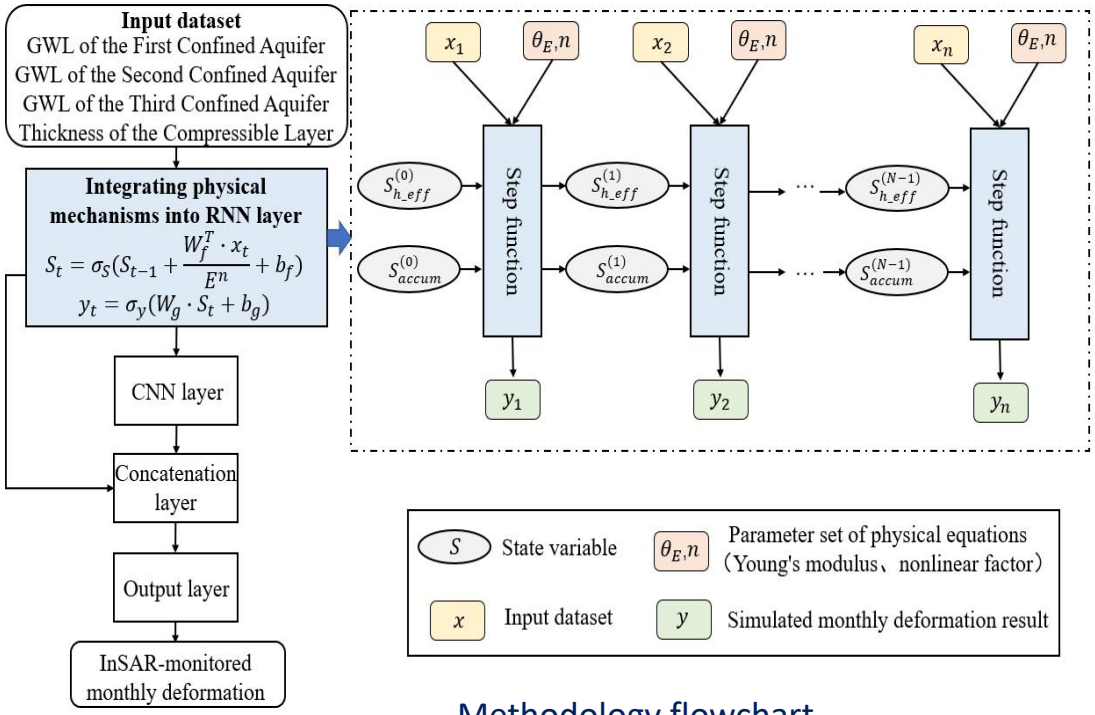
$$x_t = \left[\Delta\sigma_z^{'t}, H_t\right]^T$$

$\Delta\sigma_z^{'t}$ denoting the change in effective stress at time t, calculated as $\Delta\sigma_z^{'} = -\gamma_w \Delta h$, H is the thickness of the compressible layer. W_f^T represents the weights of the RNN hidden layer; E is a trainable parameter, corresponding to the Young's modulus

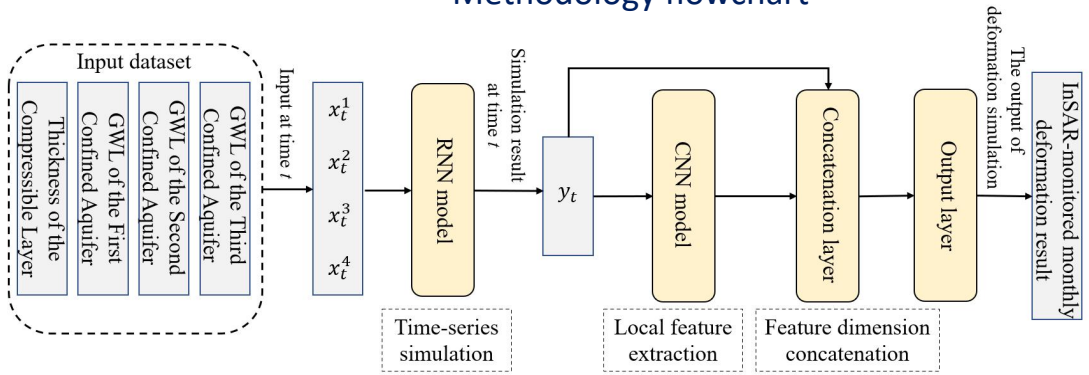
In addition, **a nonlinear factor *n*** is introduced into RNN model to characterize the nonlinear deformation.

$$S_t = \sigma_S(S_{t-1} + \frac{W_f^T \cdot x_t}{E^n} + b_f)$$

RNN is coupled with Convolutional Neural Network (CNN) to **capture long-term temporal dependencies and local temporal features of sequential data.**



Methodology flowchart



Schematic diagram of the RNN-CNN model

OUTLINE

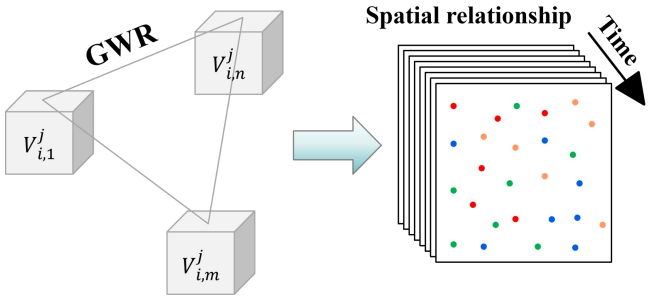
1. Background

2. Methodology

3. Case Study

CASE STUDY 1: Geographically Weighted Regression LSTM (GW-LSTM)

- Considering the temporal non-linearity and spatial heterogeneity of land subsidence, a spatiotemporal recurrent neural network model was constructed.
- ***LSTM:*** model long-term dependencies with memory cell and gating mechanisms
- ***GWR:*** model spatially varying relationships
- ***GW-LSTM:*** the **geographical weight matrix** $X_t^{w(t)}$ is constructed for each computing cell on basis of the spatial weights. Then **the spatial correlation is integrated into the network**.



Schematic diagram of spatial relationship analysis using the GWR model

Spatial weights

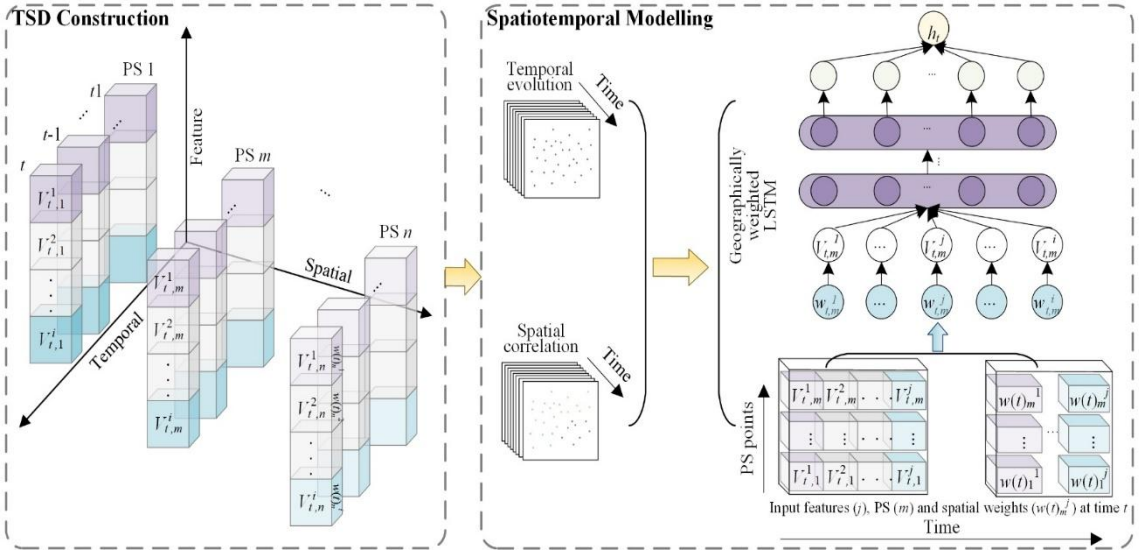
$$W(t) = \begin{vmatrix} w(t)_m^1 & w(t)_m^2 & \dots & w(t)_m^j \\ \dots & \dots & \dots & \dots \\ w(t)_1^1 & w(t)_1^2 & \dots & w(t)_1^j \end{vmatrix}$$

Forget gate $f_t = \sigma\left(W_f \cdot \left[h_{t-1}, \boldsymbol{X_t^{w(t)}}\right] + b_f\right)$

Input gate $i_t = \sigma\left(W_i \cdot \left[h_{t-1}, \boldsymbol{X_t^{w(t)}}\right] + b_i\right)$

Cell state $o_t = \sigma\left(W_o \cdot \left[h_{t-1}, X_t^{w(t)}\right] + b_o\right)$

Output gate $C_t = f_t * C_{t-1} + i_t * \tanh\left(W_C \cdot \left[h_{t-1}, X_t^{w(t)}\right] + b_C\right)$

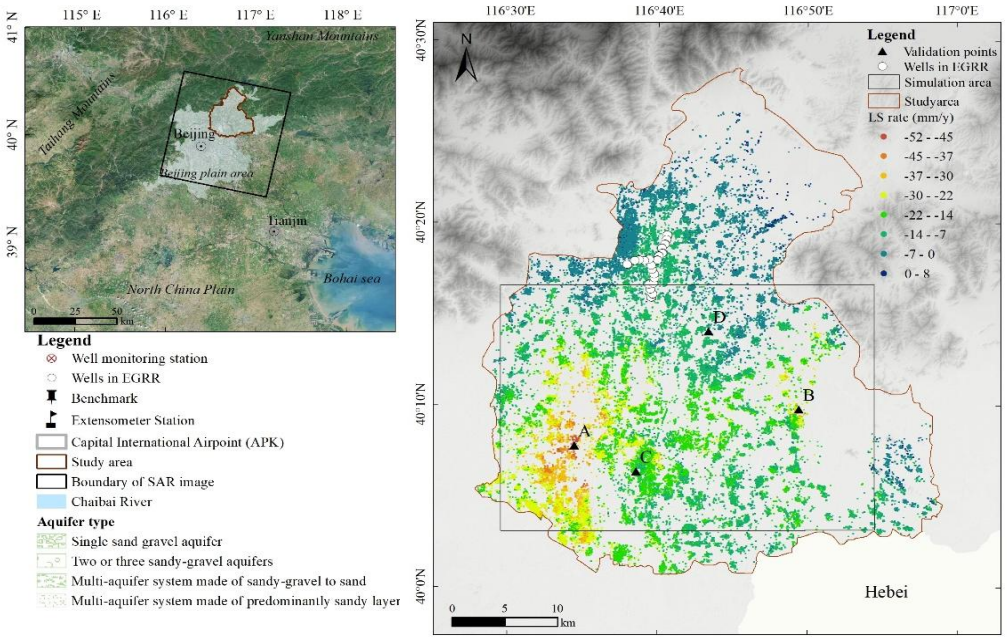


Structure of the GW-LSTM

Li & Zhu et al 2021, STOTEN

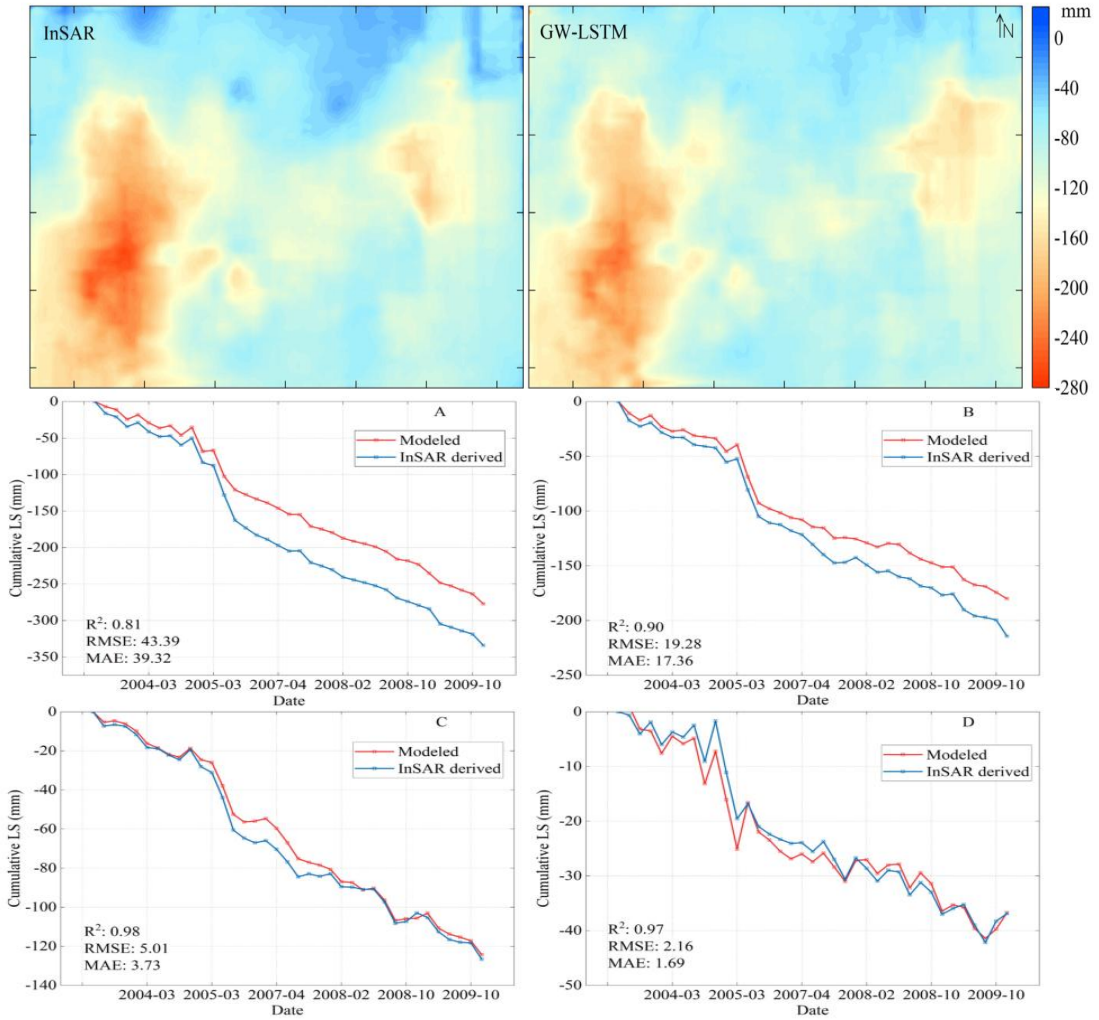
Results:

- The ***GW-LSTM model***, with the input variables including groundwater level change of the confined aquifers, inelastic specific storage, ***effectively captures the temporal evolution of subsidence***, with R^2 values consistently exceeding 0.81 and with the maximum value of 0.98.



PS distribution (with location of the four validation points in the north part of Beijing plain

Hyperparameter of the network	
Hidden layers	4
RNN unit	64
Bidirectional	False
Return_sequences	True
Time step	36



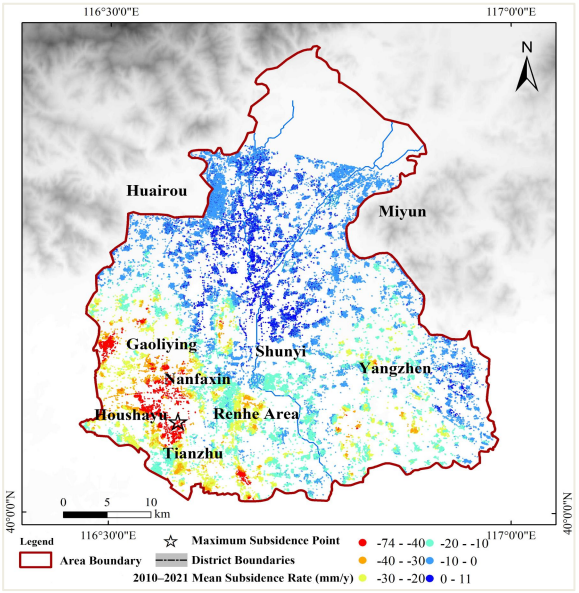
Comparison of land subsidence simulation between GW-LSTM simulation and PS-InSAR measurements (2003-2010)

CASE STUDY 2: Integration AM with GRU considering the lag time of subsidence

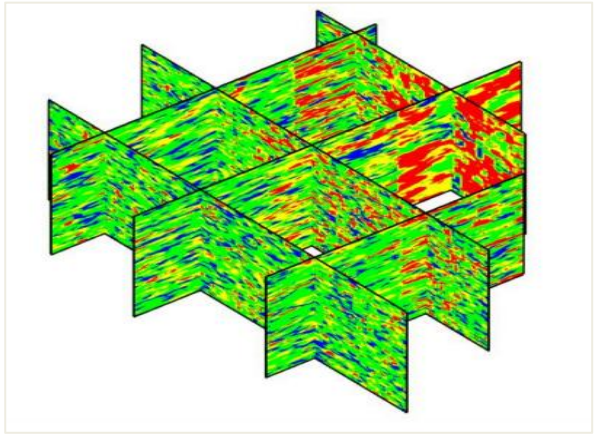
- There are multiple factors influencing land subsidence and the influence intensity of different factors is different.
- Due the hysteretic nature of land subsidence, there is the ***lag time of deformation responding to the groundwater level change*** in different aquifer system.
- Input data include groundwater level of three confined aquifers and the groundwater level considering the lag time from 2011 to 2021. The lag time is calculated based on the 3D geological model.

$$Equivalent\ interlayer\ thickness\ b_{equiv} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} b_i^2}$$

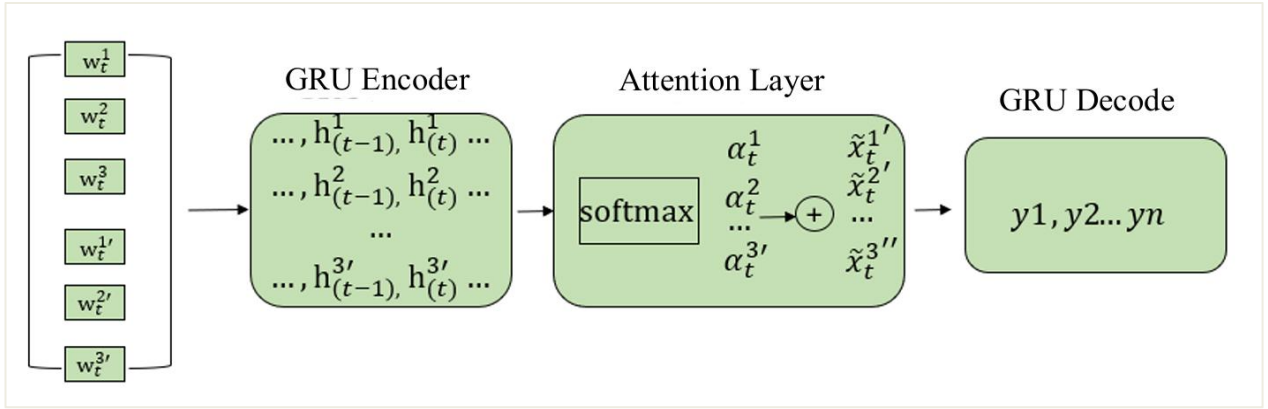
$$Lag\ time\ constant\ \tau = \frac{\left(\frac{b_0}{2}\right)^2 S_s'}{K_v'} = \frac{\left(\frac{b_0}{2}\right)^2}{D'}$$



Annual subsidence rate from 2011 to 2021



Profiles of 3D geological body



Structure of the Attention-GRU land subsidence model

Li & Zhu et al 2025, Chin. Geogra. Sci

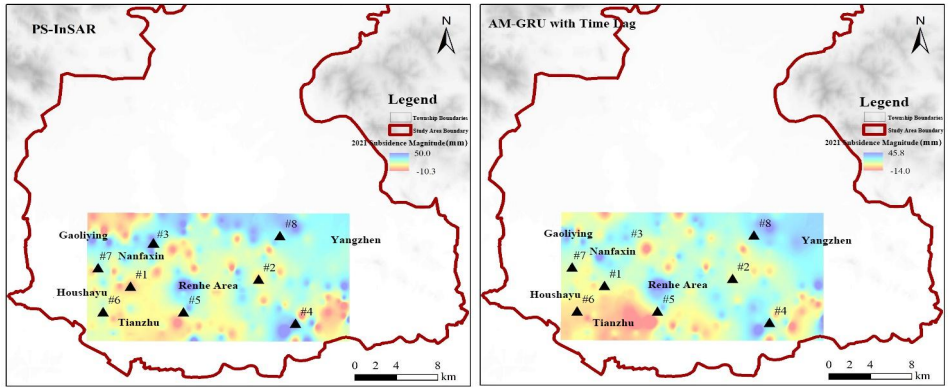
Results:

- Compared with the AM-GRU model, the lag-aware AM-GRU model achieves ***higher accuracy in predicting the monthly subsidence*** in 2021.
- Overall, the RMSE is reduced from 5.2 mm to 2.0 mm, and the coefficient of determination R^2 is improved from 0.89 to 0.93.

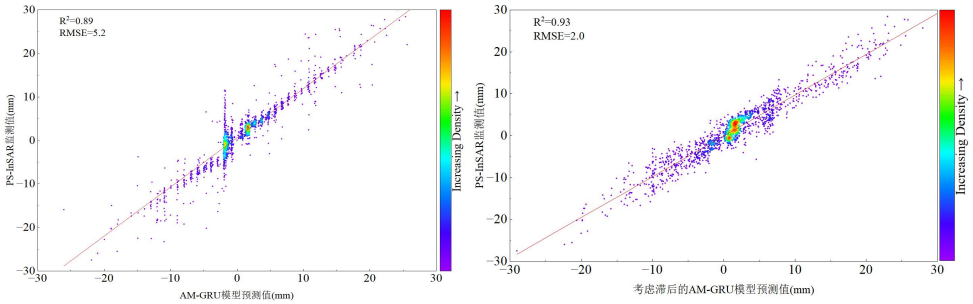
Aquifer system	Equivalent interlayer thickness (m)	lag time range (month)
The first confined aquifer	6-32	0.4-3.0
The second confined aquifer	5.6-40	0.4-4.6
The third confined aquifer	8-58	0.5-9.4

Hyperparameter setting

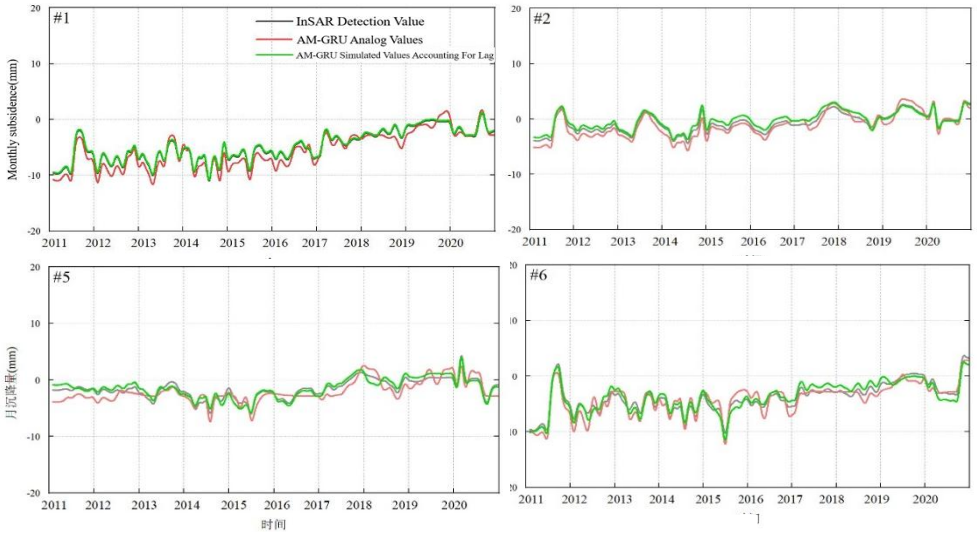
Hyperparameter	Value	Hyperparameter	Value
Learning Rate	0.001	Hidden layers	3
Batch Size	16	Training set：Validation set	7:3
Epoch	500	Optimization algorithm	Adam



Comparison between PS-InSAR monitoring data and the simulation through lag-aware AM-GRU



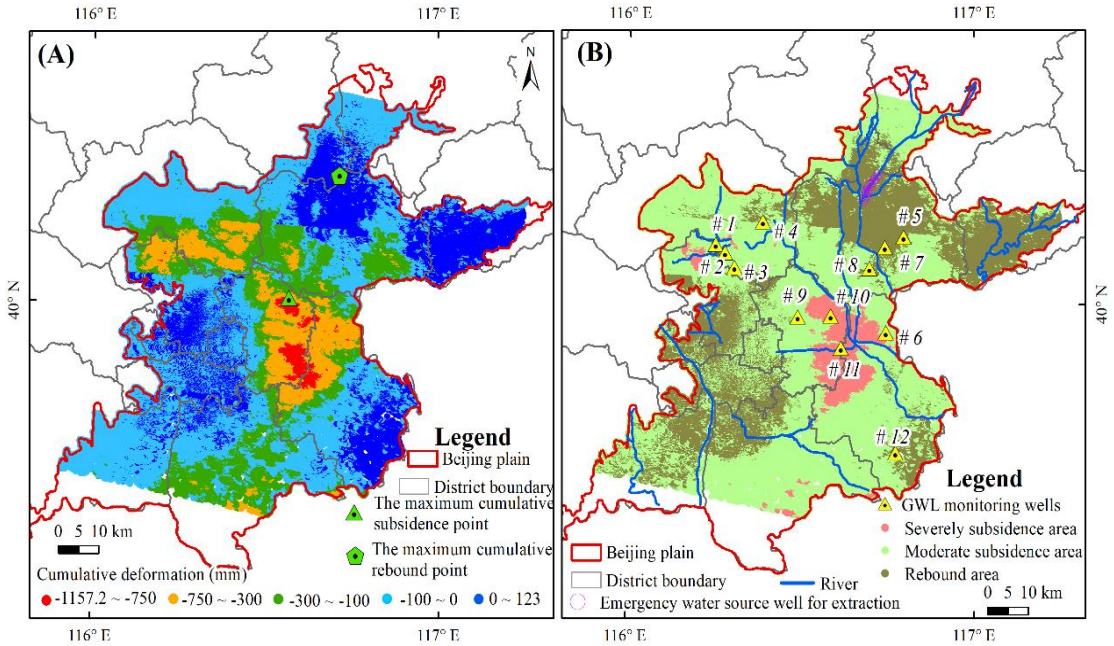
Comparison between the performance between AM-GRU and and lag-aware AM-GRU



Simulation from AM-GRU with and without thinking lag-time of subsidence

CASE STUDY 3: Coupling Physical Mechanisms with RNN

- ***In RNN layer, Terzaghi's effective stress principle is embedded and and a nonlinear factor (n) is introduced into the computation process of deformation*** at each time step.
- Input: monthly groundwater levels (first-third confined aquifers) and cumulative compressible layer thickness.
- Output: monthly InSAR-derived deformation.
- After thinning the deformation points monitored by PS, 70% were used for training, and 30% for validation.
- The training period from 2011 to 2022.



(A) Cumulative deformation over the period from 2010 to 2023.
(B) Distribution of groundwater level monitoring wells.

Hyperparameter of the network

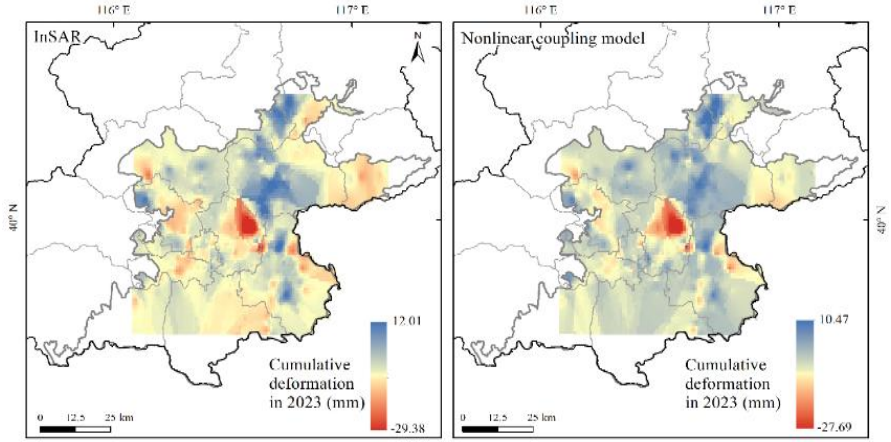
Hyperparameter	Value	Hyperparameter	Value
Learning Rate	0.001	Hidden layers	7
Batch Size	16	Training set: Validation set	8:2
Epoch	2000	Optimization algorithm	Adam

Index	Zone	Deformation rate
1	Rebound	>0 mm/yr
2	Moderate subsidence	0 ~ -30 mm/yr
3	Severely subsidence	< -30 mm/yr

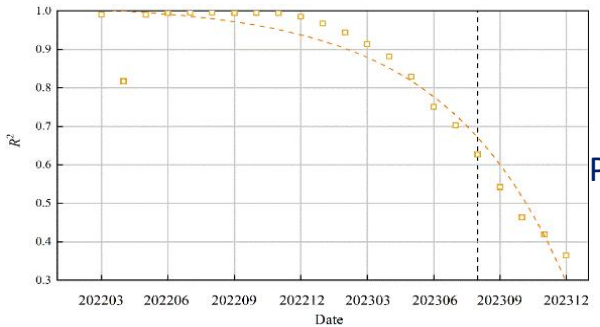
Results:

- Compared with the MODFLOW-SUB model, the RNN-CNN model, and the coupling model without **nonlinear factor *n***, the coupling model with *n* achieved **good predictive performance**, with maximum R^2 increases of 0.44, 0.31, and 0.61, respectively, and maximum RMSE reductions of 2.6 mm, 3.3 mm, and 14.6 mm, respectively.

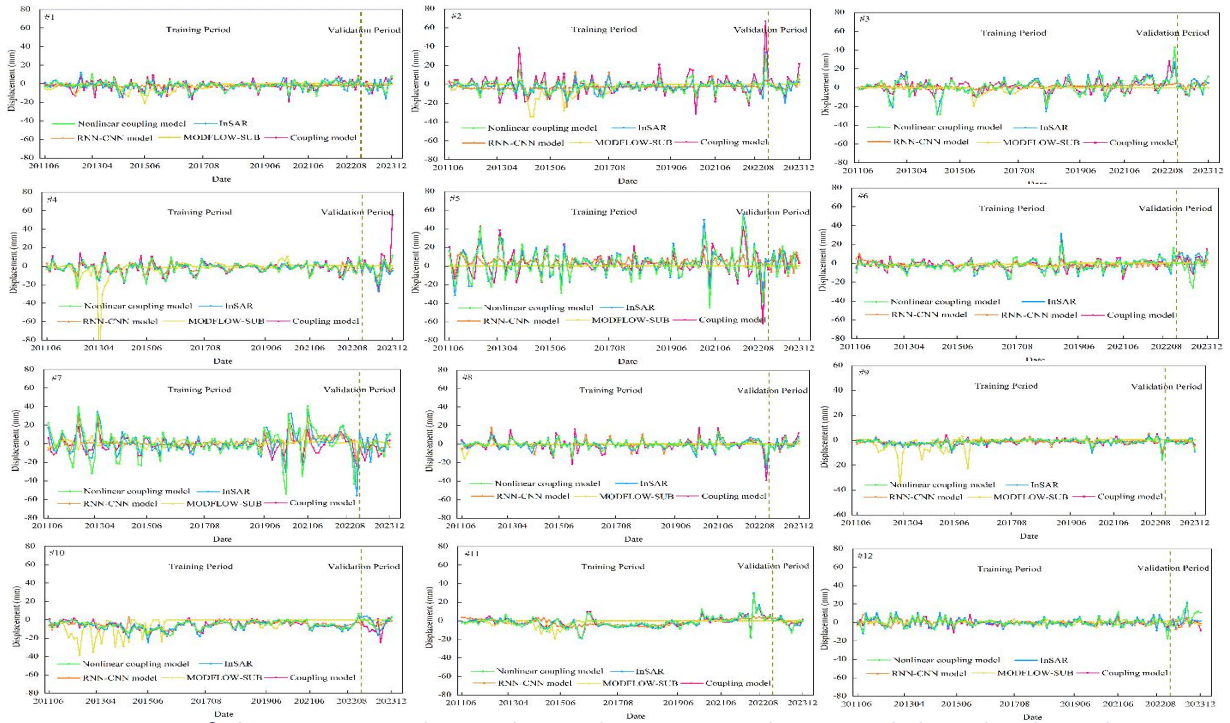
- Predictive timeliness analysis demonstrated that the nonlinear coupling model maintains **reliable predictive capability for up to 19 months**.



Distribution of cumulative deformation simulated by the nonlinear coupling model



Predictive timeliness of the nonlinear coupling model



Fitting curve of the outputs through nonlinear coupling model and other three models

Index	Zone	Correlation coefficient between simulated and monitored values
1	Rebound	0.82 ~ 0.96
2	Moderate subsidence	0.71 ~ 0.97
3	Severely subsidence	0.88 ~ 0.97

Thanks for your attention !