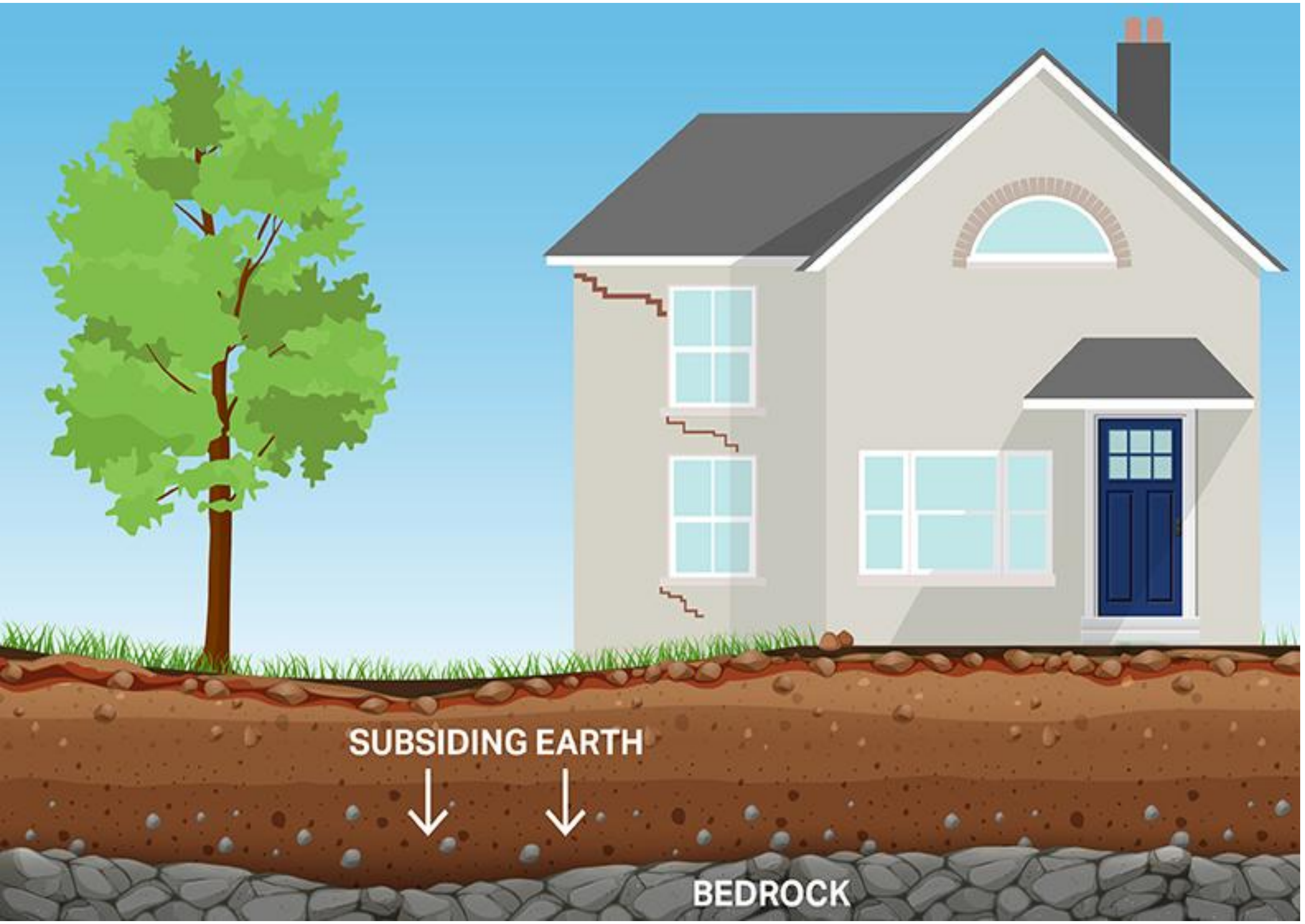
Outline

- INTRODUCTION
- METHODOLOGY
- RESULTS & DISCUSSION
- CONCLUSIONS

Impact of land subsidence



https://www.totallandlordinsurance.co.uk/knowledge-centre/subsidence

1. Land subsidence is a slow-moving global risk (about 60 countries).
2. The gradual sinking of the ground surface due to natural and human causes.

Impacts include:

- Damage to buildings and infrastructure
- Hindrance to urban development and resource use
- Seawater intrusion in coastal zones

New methods, new problems.

https://www.newsecuritybeat.org/2021/08/build-build-western-route-chinas-south-north-water-diversion-project/

1. Beijing's confirmed water shortage and over-pumping caused severe land subsidence.

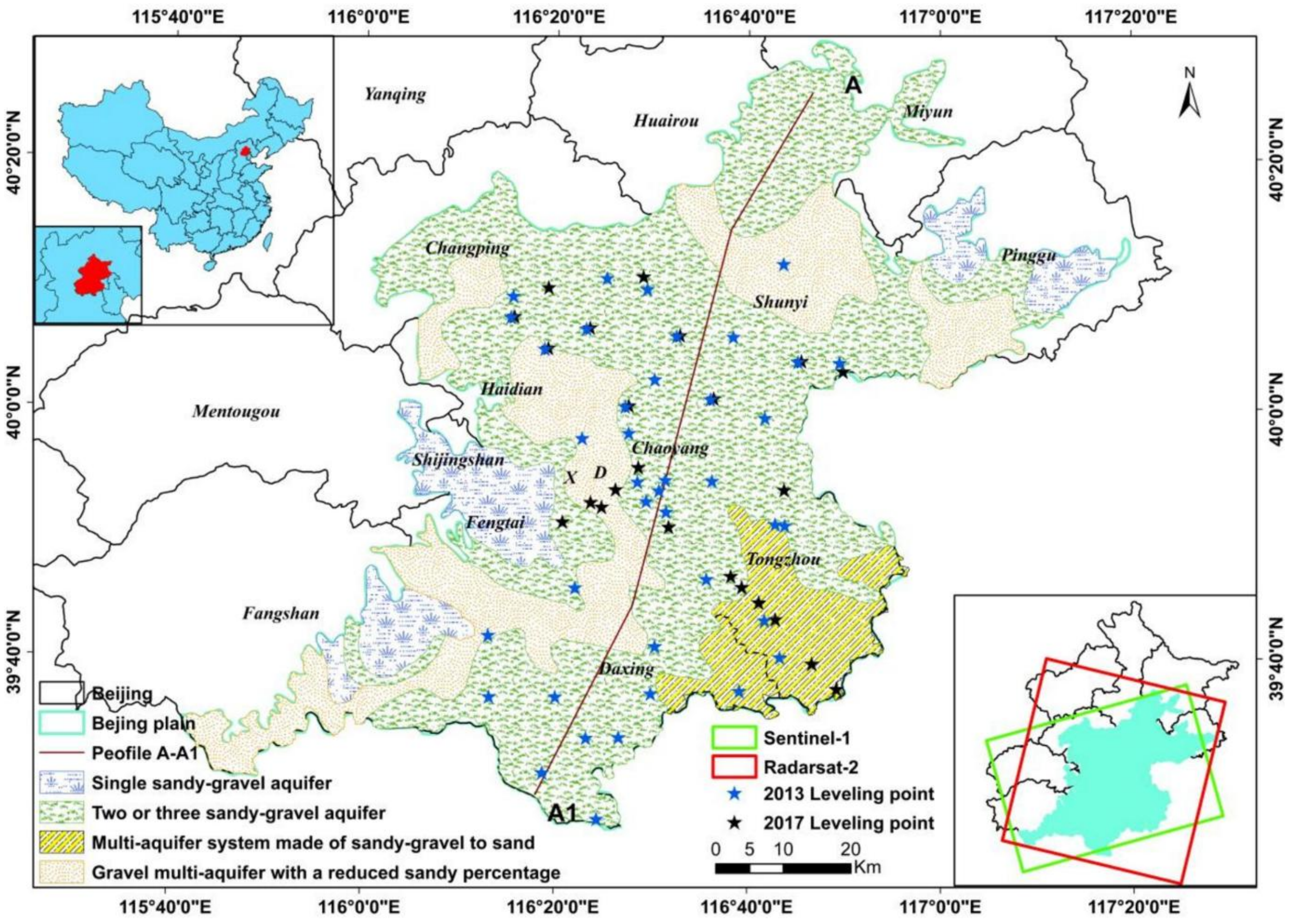
2. The 2014 South-to-North Water Diversion Project introduced external water sources and change subsidence patterns.

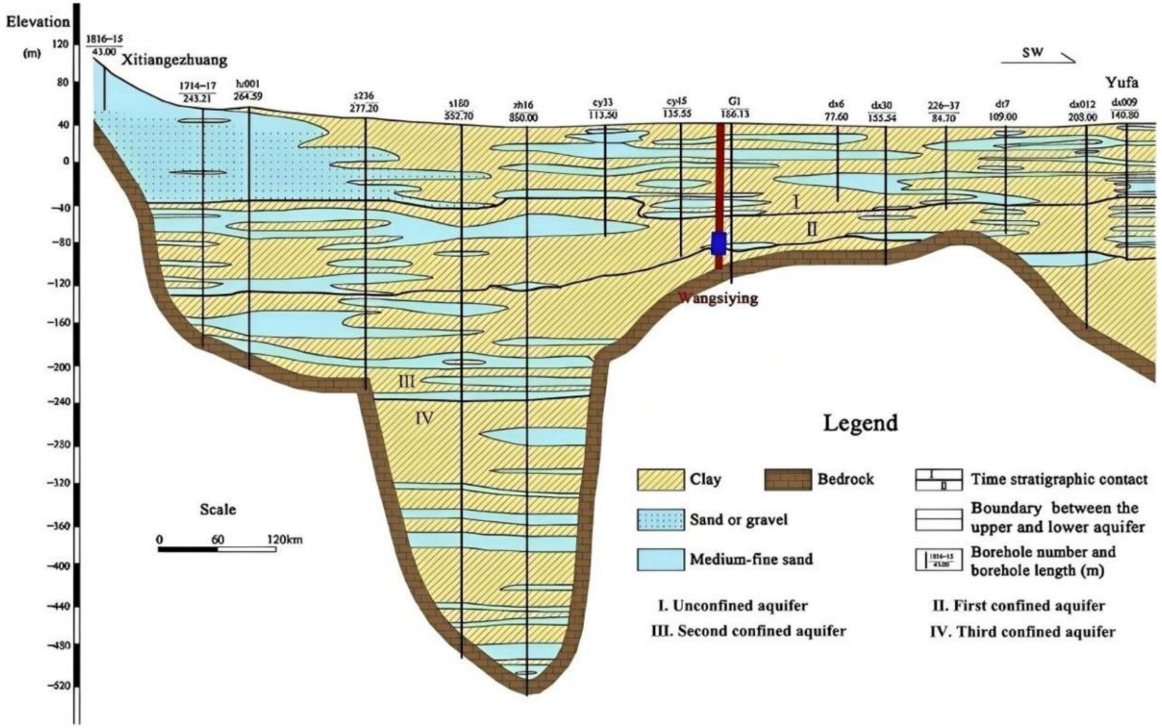
3. Reassessing key controls and trends is essential to optimize water-transfer strategies and risk management.

Objectives

1. combine WOE & LightGBM to reduce data needs and open the "black box."
2. Compare subsidence patterns before vs. after SWDP to measure impact.
3. Focus on groundwater level, clay thickness, and land use.
4. Create a risk map for planning and early warning.

Study area





The Beijing Plain in southeastern Beijing covers about 6,390 km^2, gently sloping from northwest to southeast at 100–200 m elevation.

Table 1
The division of aquifers in the Beijing Plain.

Aquifer	Time stratigraphic	Depth of bottom (m)
The unconfined aquifer (I)	Quaternary Holocene alluvial pluvium	0 ~ – 50
The first confined aquifer (II)	Quaternary Upper Pleistocene alluvial pluvium	–50 ~ – 100
The second confined aquifer (III)	Quaternary Middle Pleistocene alluvial pluvium	–100 ~ – 180
The third confined aquifer (IV)	Quaternary Lower Pleistocene alluvial pluvium	–180 ~ – 300

Data Sources

Table 2
Acquisition parameters of Radarsat-2 and Sentinel-1.

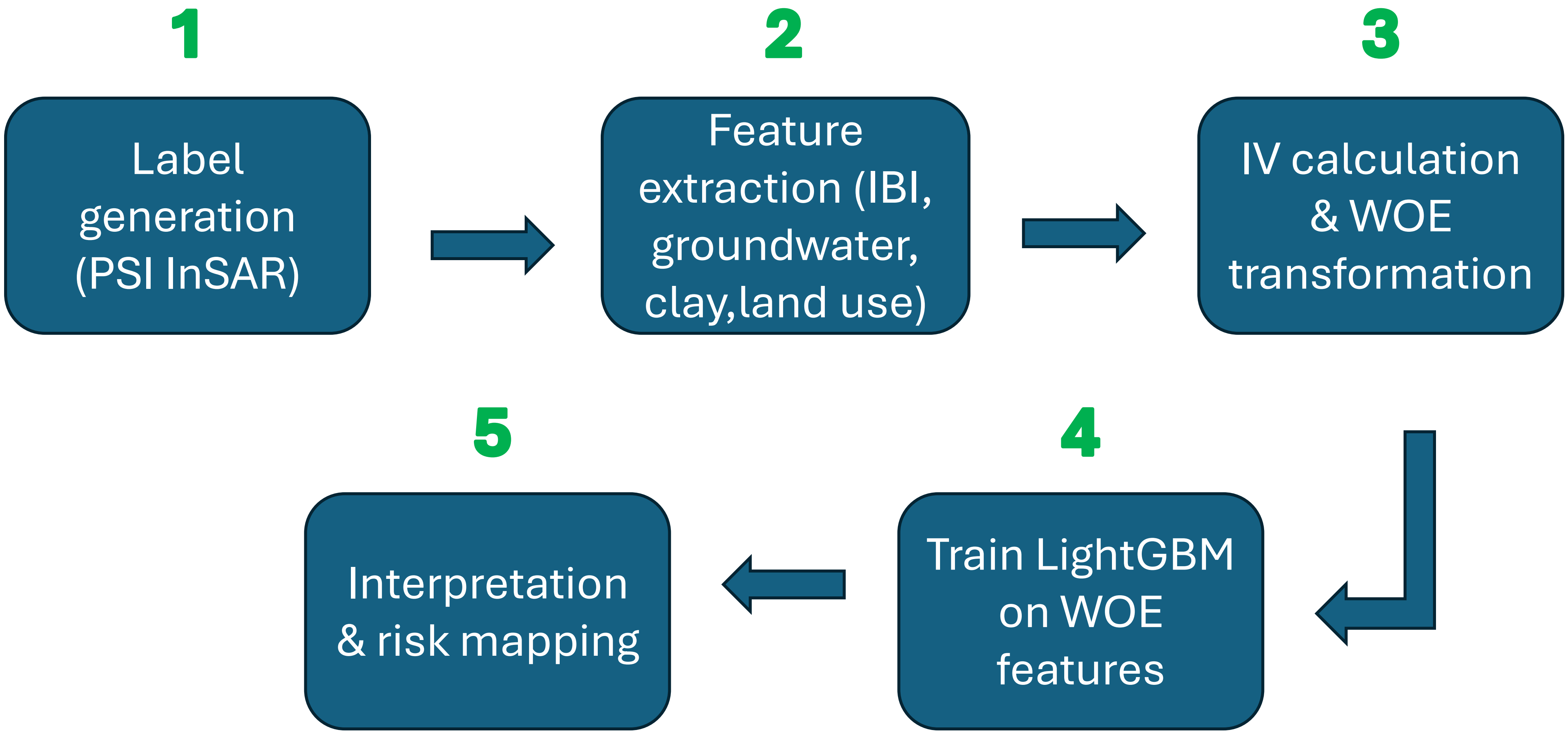
SAR Sensor	**Radarsat-2**	**Sentinel-1**
Orbit	26	142
Revisit cycle	24 days	12 days
Orbit altitude	798 km	693 km
Orbit direction	Descending	Ascending
Spatial resolution	30 *m*	20 *m*
Number of images	48	67
Band (wavelength)	C (5.6 cm)	C (5.5 cm)
Temporal coverage	2010–2015	2016–2018
Date of main image	May 22, 2013	July 18, 2017
Line of Sight Incidence Angles	43.7514°	27.0965°
Maximum temporal baseline	446.1 *m*	126.3 *m*
Maximum spatial baseline	912 days	564 days

1. PSI (SARPROZ) : Delaunay unwrap & atm. correction → mm-level subsidence (validated 2013/2017)

2. Land use : 2018 Landsat 8 IBI + POK → 30 m building-cover map

3. Formatting : All data → 30 m rasters in ArcGIS

Research Methodology Workflow

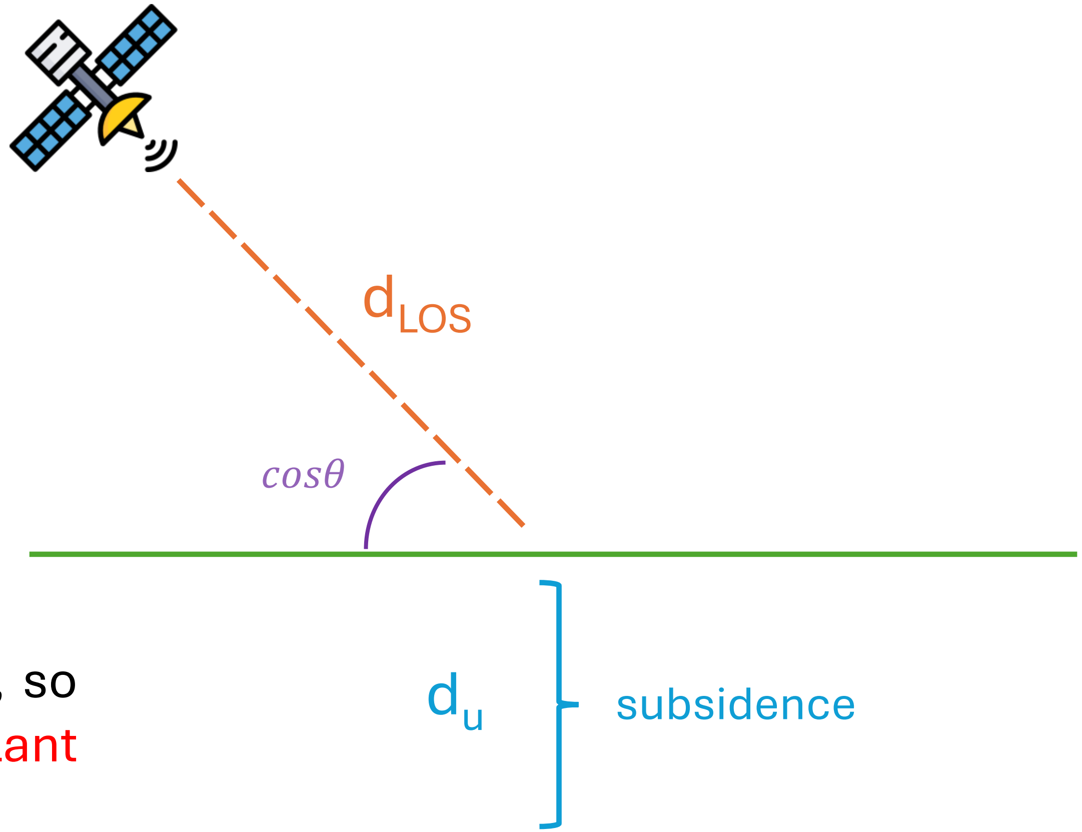


Permanent scatterers Interferometry (PSI)

→ Obtain vertical subsidence displacement

$$\Phi = \boxed{\psi_{\text{def}}} + \psi_{\text{top}} + \psi_{\text{orb}} + \psi_{\text{atm}} + \psi_{\text{noise}}$$

$$\boxed{d_{LOS}} = d_u / \cos\theta$$



1. Satellite → Observes the ground at an oblique angle, so the measured displacement (d_{LOS}) represents the slant distance, not the true vertical subsidence.

2. Subsidence (d_u)→ Multiply d_{LOS} by $\cos(\theta)$ to obtain the true vertical displacement.

Index-based built-up index (IBI)

→ An index for analyzing the degree of urbanization

RED, GREEN, NIR, MIR → NDBI 、SAVI 、MNDWI → AvgVW = (SAVI + MNDWI)/2

$$IBI = \frac{NDBI - AvgVW}{\mathrm{NDBI} + AvgVW}$$

IBI > 0 → building surface

IBI ≤ 0 → non-building surface

NIR : Near-Infrared Band

MIR : Mid-Infrared Band

NDBI : Building Surface

SAVI : soil 、vegetation

MNDWI : Water

weight of evidence (WOE), IV

→ Identifying Key Factors for Subsidence – WOE and IV.

$$W_i^+ = \ln\frac{p(B/L)}{p(B/\overline{L})}$$

positive correlation

$$W_i^- = \ln\frac{p(\overline{B}/L)}{p(\overline{B}/\overline{L})}$$

Negatively correlation

$$IV = \sum_{j=1}^{n}\left(\frac{g_j}{g} - \frac{b_j}{b}\right) \times \ln\frac{g_j/g}{b_j/b}$$

Subsidence Proportion

subsidence Strength

$$C = W_i^+ - W_i^-$$

C>0 positively correlated with subsidence.
C<0 negatively correlated with subsidence.

Table 3
Evaluation standard for IV.

IV range	Predictive effect
< 0.02	Useless for prediction
0.02 –0.1	Weak predictor
0.1 –0.3	Medium predictor
0.3 –0.5	Strong predictor
>0.5	Suspicious or too good to be true

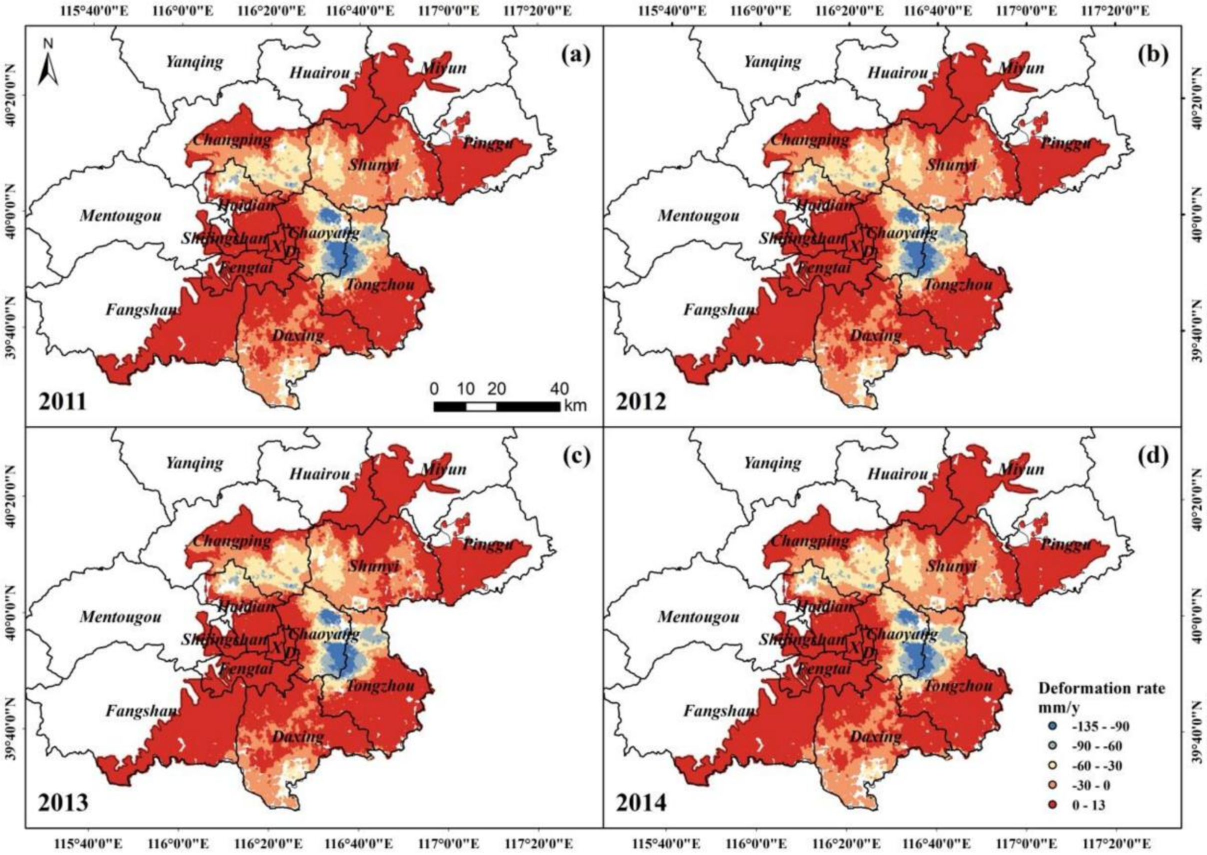
Light Gradient Boosting Machine (LightGBM)

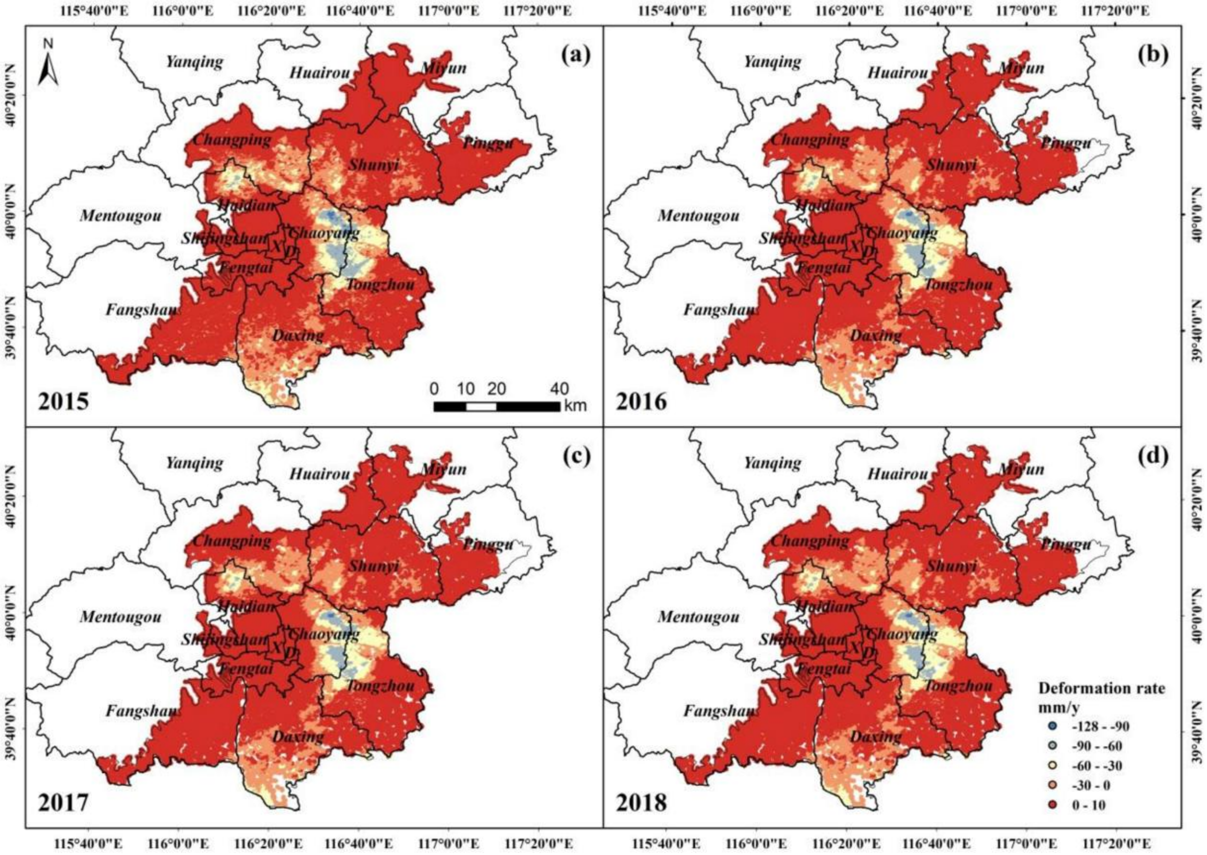
GOSS → Skips easy samples and keeps hard ones to improve learning efficiency.

EFB → Bundles exclusive features together to reduce data dimensions and speed up training.

Leaf-wise → Grows the tree by splitting the leaf with the highest information gain, making training faster and more accurate.

Changes in Land Subsidence Before and After SWDP





Years	2011-2014	2015-2018
Proportion of area with subsidence	67 %	58 %
Proportion of severe subsidence (> 30 mm/y)	14 %	7 %
Maximum annual subsidence	≈ 130 mm	< 110 mm

evaluate the impact of the SWDP on subsidence.

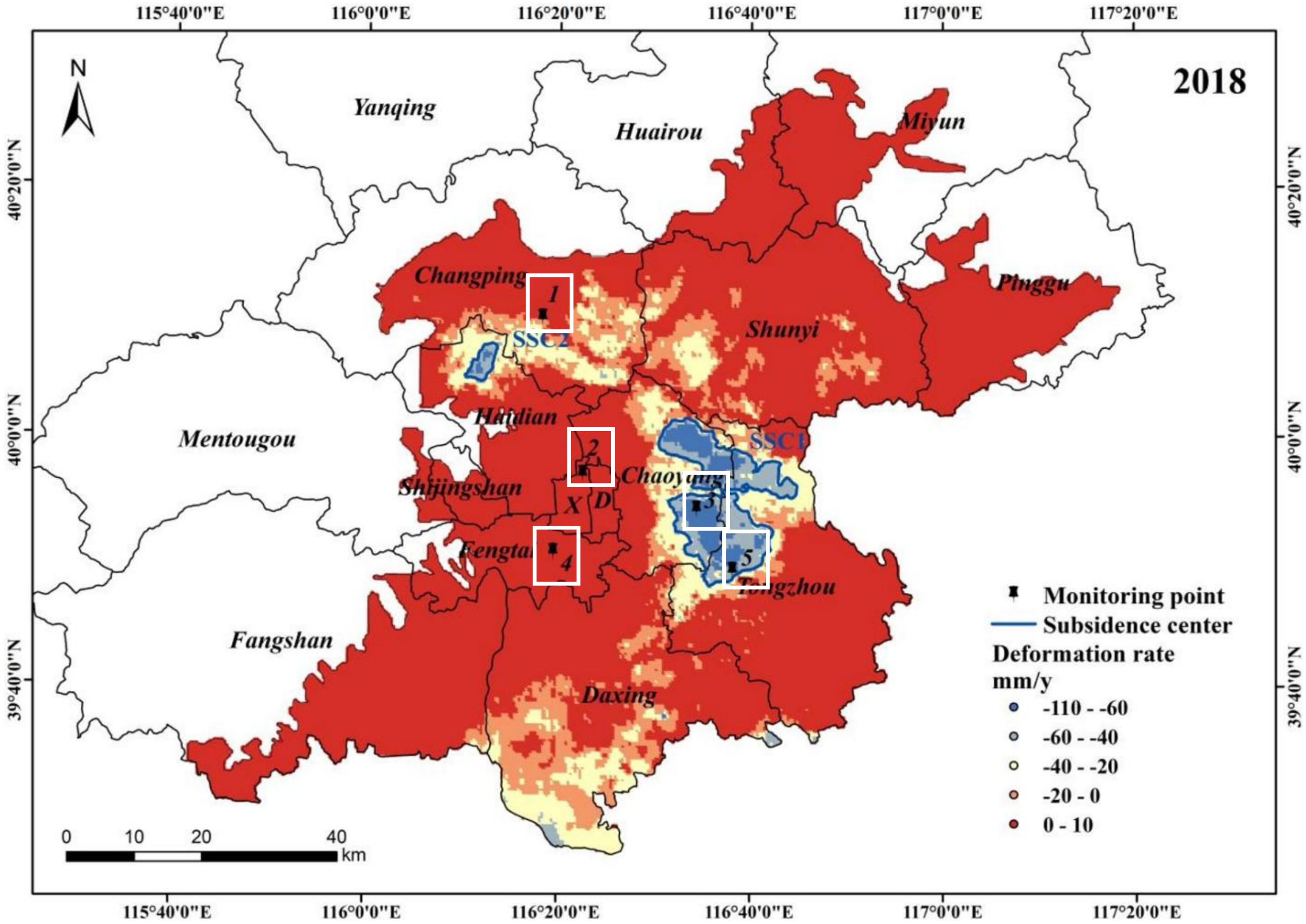
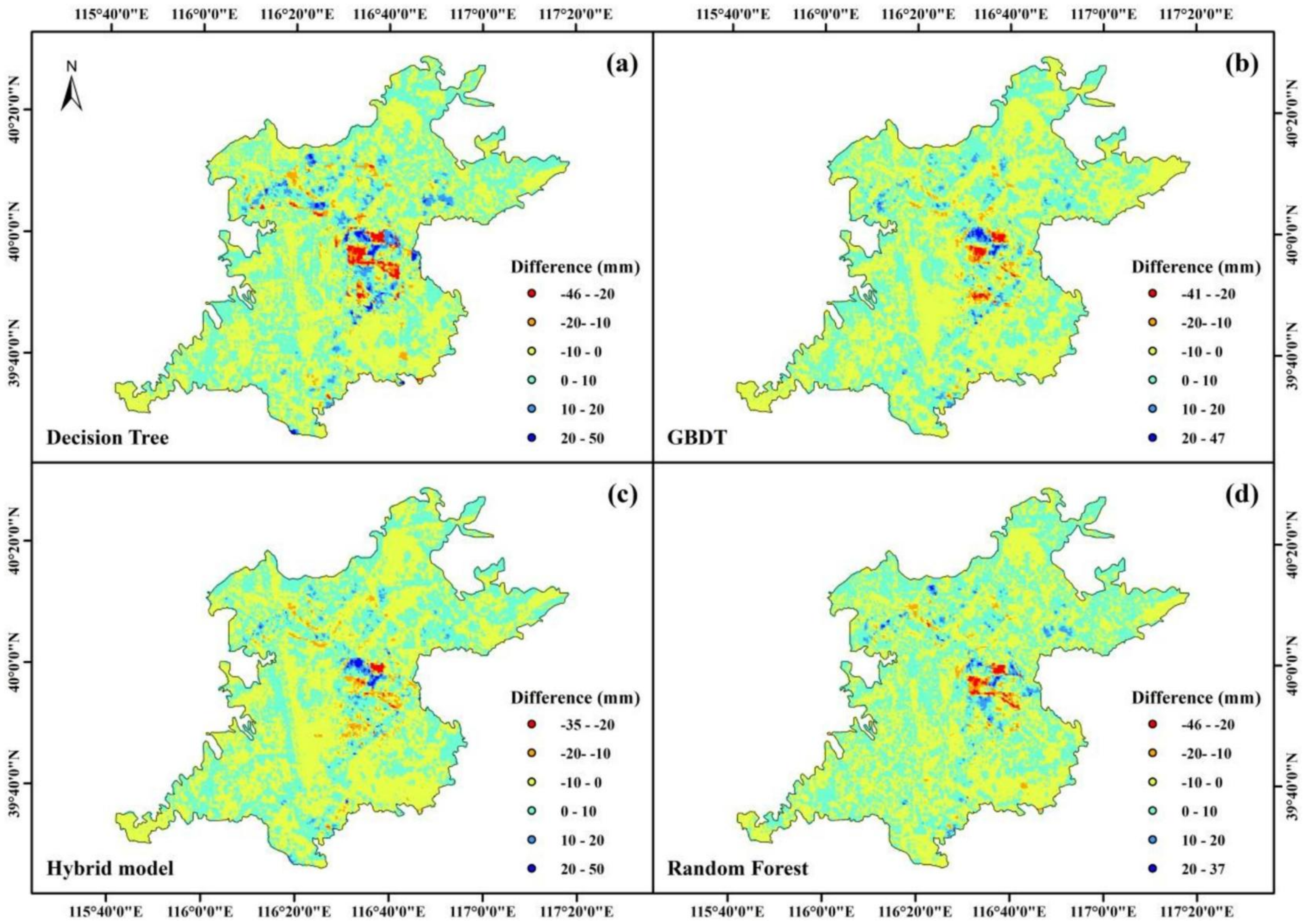


Table 4

Time series of changes in subsidence of monitoring points (The SWDP was opened on 12 December 2014).

Point	2011	2012	2013	2014	2015	2016	2017	2018
1	−8.15	−4.7506	−5.5	−0.80	6.19	9.63	9.32	9.33
2	2.14	2.16	1.11	0.42	10.09	10.63	9.47	9.47
3	−123.58	−122.19	−121.43	−118.08	−109.75	−104.16	−100.80	−100.79
4	2.01	2.21	1.96	0.45	8.79	9.63	9.69	9.64
5	−107.60	−108.04	−110.98	−111.11	−109.78	−100.54	−97.30	−97.30

Model Performance in Different Subsidence Zones



1. < 20 mm/y : All models perform good; Random Forest is slightly better
2. 20–40 mm/y : Hybrid model is the most accurate, with the most points within ±10 mm
3. > 40 mm/y : Hybrid model matches PSI best, especially at boundary regions
4. Extreme subsidence (> 90 mm/y) : All models perform poorly due to limited training samples
5. Error control : Hybrid model has the fewest points with error > ±10 mm (2,227 points)

Which Model Works Best?

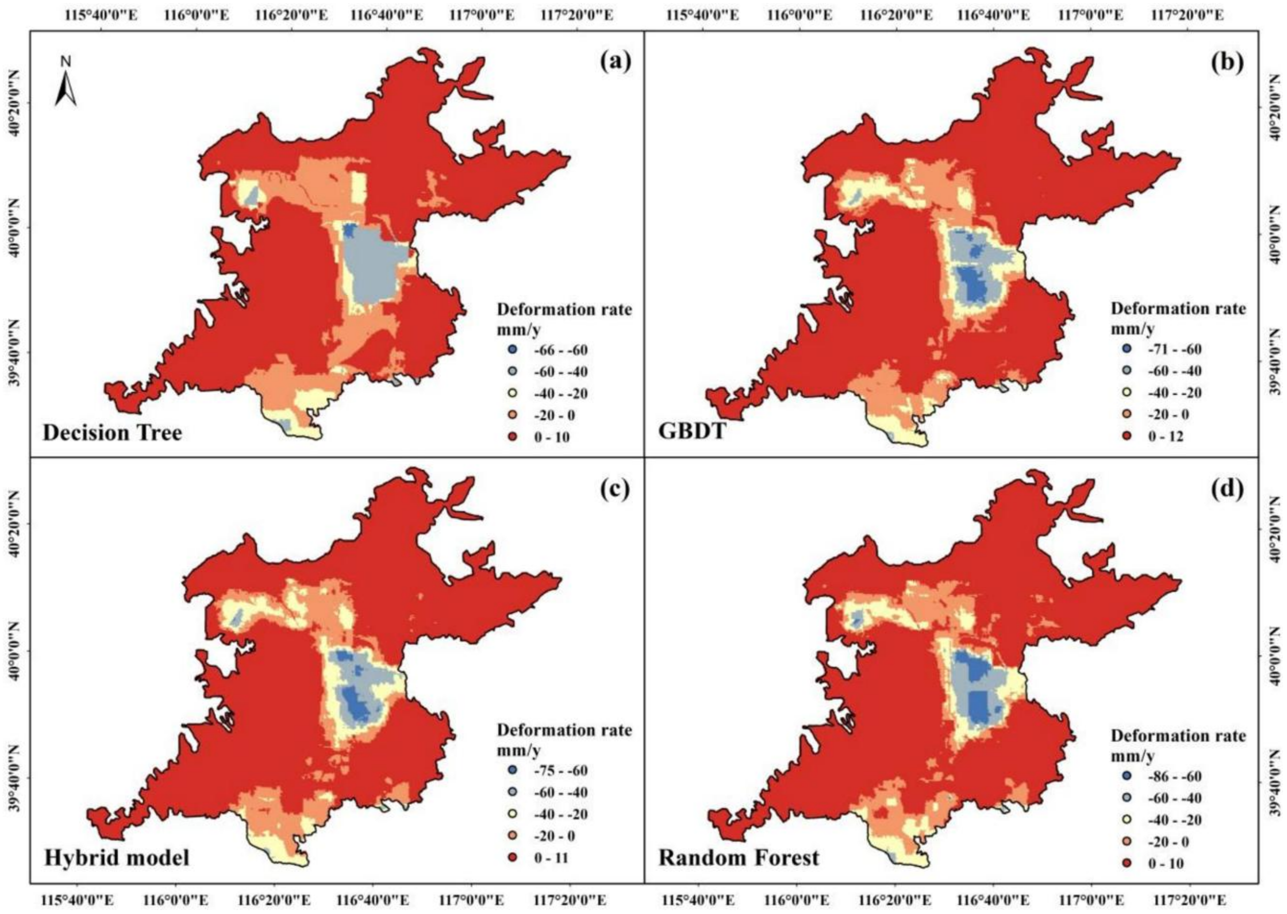


Table 5
The accuracy of five models in the validation step based on R^2, MSE and MAE criteria.

Model	Dataset	R^2	MSE	MAE
PR	validation set	0.74	59.12	4.67
	testing set	0.59	97.00	6.57
Decision Tree	validation set	0.91	22.43	2.45
	testing set	0.86	31.95	2.97
GBDT	validation set	0.96	9.38	1.83
	testing set	0.92	18.81	2.50
Hybrid model	validation set	0.97	7.84	1.72
	testing set	0.92	17.68	2.46
Random Forest	validation set	0.97	6.76	1.24
	testing set	0.93	16.35	2.02

0.92	17.68	2.46
0.93	16.35	2.02

Although Random Forest performs slightly better in numbers, the Hybrid model achieves better spatial agreement with observed subsidence and provides clearer feature interpretation for practical use.

Table 6
IV and WOE calculation results for the **multiple factors.**

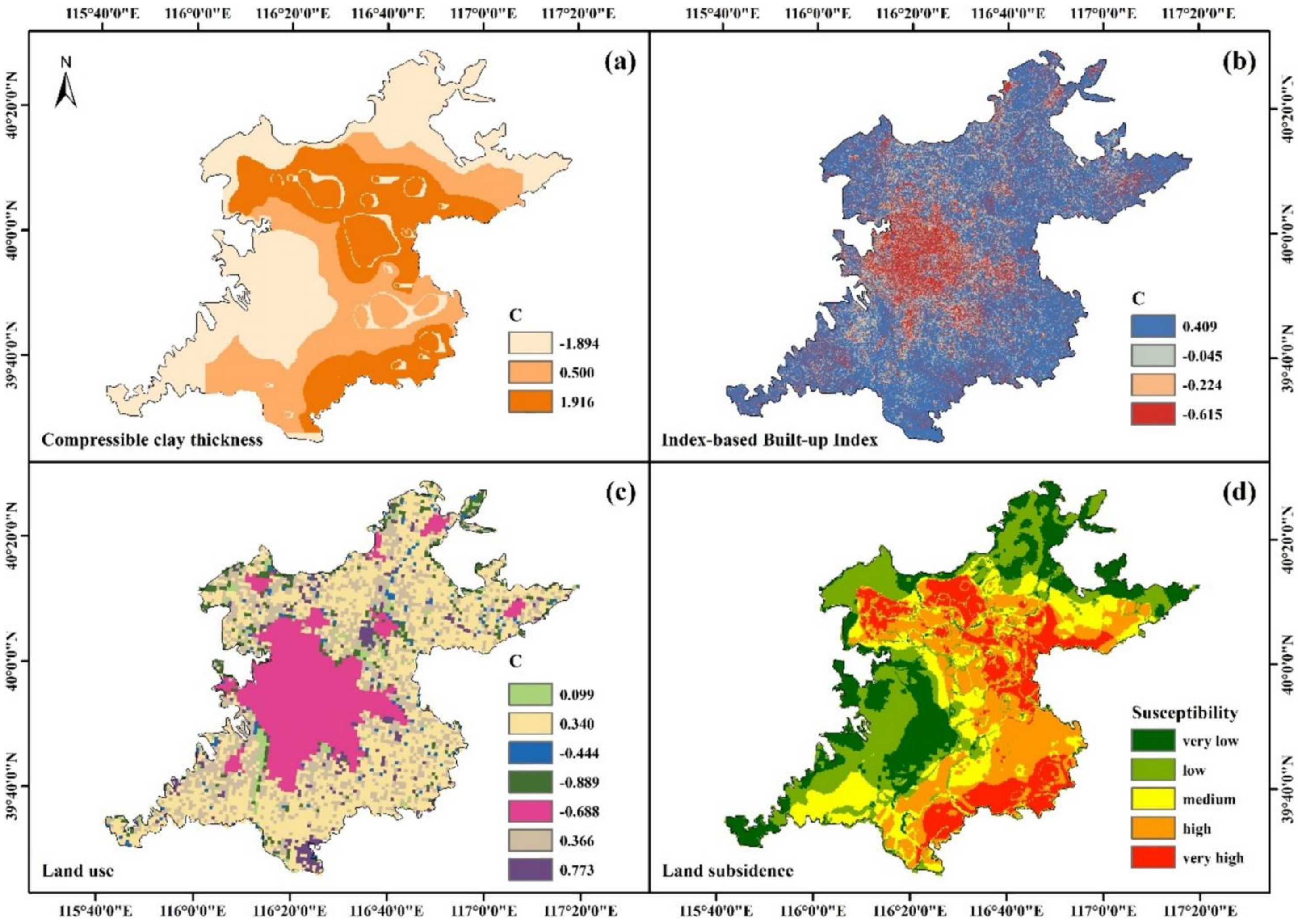
Indicator factor	Class	IV	WOE$^+$	WOE$^-$	C
Unconfined aquifer	−32.92 – −3.52 mm	0.092	−0.673	0.173	−0.846
	−3.52 – 4.86 mm	0.059	0.295	−0.661	0.956
	4.86 – 21.80 mm	0.041	−0.635	0.072	−0.707
First confined aquifer	−20.62 – −1.29 mm	0.074	0.630	−0.145	0.775
	−1.29 – 0.00 mm	0.364	−0.994	0.628	−1.622
	0.00 – 19.24 mm	0.156	0.627	−0.429	1.056
Second confined aquifer	−28.92 – −0.91 mm	0.188	0.924	−0.267	1.191
	−0.91 – 0.00 mm	0.391	−0.933	0.855	−1.788
	0.00 – 22.72 mm	0.173	0.800	−0.304	1.104
Third confined aquifer	−16.07 – −0.18 mm	0.257	1.485	−0.202	1.687
	−0.18 – 0.00 mm	0.367	−0.806	1.274	−2.080
	0.00 – 19.05 mm	0.332	1.175	−0.390	1.565
Compressible clay	0 – 80 mm	0.458	−1.050	0.844	−1.894
	80 – 140 mm	0.036	0.364	−0.136	0.500
	140 – 230 mm	0.487	1.442	−0.474	1.916
IBI	−1.00 – 0.28	0.013	0.139	−0.270	0.409
	0.28 – 0.50	0.0002	−0.039	0.006	−0.045
	0.50 – 0.62	0.003	−0.209	0.015	−0.224
	0.62 – 1.00	0.038	−0.531	0.084	−0.615
Land use	Grassland	Null	0.096	−0.003	0.099
	Arable land	Null	0.203	−0.137	0.340
	Waters	Null	−0.434	0.010	−0.444
	Woodland	Null	−0.844	0.045	−0.889
	Urban land	Null	−0.527	0.161	−0.688
	rural settlements	Null	0.287	−0.079	0.366
	other construction land	Null	0.745	−0.028	0.773

How much do groundwater level changes in different aquifers affect land subsidence?

Aquifer Type	Impact of Water Level Decline	Impact of Water Level Rise
Unconfined aquifer	Subsidence occurs when water change is between −3.52 and 4.86mm	Weak correlation
First confined aquifer	**Strong impact** – subsidence worsens	Minor impact
Second & Third confined aquifers	Minor impact	**Stronger impact** – subsidence worsens

Subsidence is mainly caused by pumping from the middle aquifer. Deeper water levels rising won't stop it, and the surface aquifer has little effect.

Key Factors Influencing Land Subsidence



(a) Thicker clay → More subsidence

(b) IBI has weak correlation with subsidence(IV < 0.1)

(c) Subsidence mainly in construction, farmland, rural areas

(d) High-risk zones align with thick clay and heavy pumping

Conclusions

1. SWDP reduce subsidence : Post-project rates and extent dropped sharply.
2. Deep aquifers control : 2nd/3rd confined aquifer levels drive most subsidence pumping still needs strict management.
3. Clay + construction risk : Major developments over thick compressible clay create high-risk zones.
4. Hybrid model strength : WOE + LightGBM delivers fast, accurate predictions with clear factor insights.

Thank you for your attention