Available online at www.jrors.yazd.iau.ir

Journal Of Radar and Optical Remote Sensing and GIS

JRORS 1 (2023) 7–20

Investigating spatial heterogeneity by Implementing the mgwr python package, a case study: southwestern of Tehran Plain

Ali, Soleymani[a*], Parham Pahlavani [a]

[a]*Geographic Information Systems, School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran*

Received 18 September 2022; Revised 6 February 2023; Accepted 20 March 2023

Abstract

Click here and insert your abstract text.Please provide an abstract of 150 to 250 words. The abstract should not contain any undefined abbreviations or unspecified Land Subsidence often causes irreversible damage to infrastructures and costs lots of expenses for governments annually; Hence, studying and monitoring subsidence in either plains or urban areas has become necessary in last decades. Studies have introduced excessive depletion of aquitards as the dominant factor in the occurrence of this hazard. In this study, the main aim was to take the impact of other spatial factors involving land subsidence into consideration. To devise a plan whether to pause or reduce the subsidence rate, we need to understand the mechanism of each factor inducing land subsidence. Here, we show the outcomes of a Geographically Weighted Regression (GWR) method with a fixed Gaussian kernel to identify the impact of each of the spatial factors inducing subsidence compared with the results from a Multi Linear Regression (MLR). In this regard, outputs of a compiled Interferometric Synthetic Aperture Radar (InSAR) time series analysis of the 15 Envisat ASAR images consumed to capture displacement from 2003 to 2005. Afterward, a kriging interpolation method is implemented to generate a surface of subsidence. The Python package, "mgwr" is used to compile both GWR and MLR models. Several statistical diagnostics are performed to assert the GWR superiority over other non-geographical methods when dealing with spatial data. Finally, the GWR results show that just six factors out of 10 tend to be the dominant factors.

Keywords: Subsidence, Multi linear regression, Geographically Weighted Regression, InSAR.

* Corresponding author. Tel:. 081-32563105

E-mail address: ali.soleymani@ut.ac.ir

1. Introduction

Land subsidence is a geo-hazard caused by the change in the effective stress among soil particles(Chen et al. 2019). Consolidation and compression of underlayers are the results of these changes(Hu et al. 2004). The movement of the surface could happen in both ways of a gentle shrinking or sudden sinking, depend on the soil materials in the area(Galloway and Burbey 2011). Mostly, the solution of carbonate rocks underground results in the sudden sinking type of subsidence(Zeitoun and Wakshal 2013). But the other type is often induced by the over-extraction of fluids underground (e.g., water)(Zeitoun and Wakshal 2013). Both types are significant threats to sustainable development, and a plan needs to be devised to control or mitigate the process. Therefore, measuring the rate and the extent of the subsided area besides studying natural and human factors are critical. A tool to measure the movement is needed. Conventional levelling and GPS methods are not capable to cover the entire subsidence area. Hence, InSAR for its capability in determining displacement in millimetre-scale precision over a wide span of land is the widely used method in land displacement monitoring (Akbari and Motagh 2011; Castellazzi et al. 2016; Hoffmann et al. 2001; Motagh et al. 2007). When it comes to spatial data where the context may vary by location of the record, numerous researches have shown that the GWR would be a reliable method to consider (Blainey and Mulley 2013; Brunsdon, Fotheringham, and Charlton 1996; Brunsdon, Fotheringham, and Charlton 1998; Chasco, García, and Vicéns 2007; Fotheringham, Brunsdon, and Charlton 2003; Sultana, Pourebrahim, and Kim 2018).Besides groundwater over-extraction, studies partially mentioned other factors provoking land subsidence. Research by (Haghighi and Motagh 2019) points out that interruption in land subsidence in Tehran Plain is either caused by Pishva Fault in Varamin via blocking the groundwater flow or change upon alluvium thickness over the extent. Hu in (Hu et al. 2019) found groundwater over-extraction as the leading factor of land subsidence and the geological faults as the controller toward its local distribution. Shemshaki in (Shemshaki, Boulourchi, and Entezam Soltani 2006) identified the condition of bedrock, the thickness of the clay layer and, the removal of sand along the rivers as the main reasons to land subsidence alongside the overdraft of groundwater.

The over-exploitation of groundwater in Beijing in response to domestic, agricultural, and industrial requirements in recent decades lead to 707 mm accumulated land subsidence in this area(Yu et al. 2020b); Based on the results of a GWR model with hydrological information (Yu et al. 2020b) Found that the second confined aquifer plays a vital role in land subsidence in most parts of the area. They also found that groundwater level change and the thickness of the compressible layer are two main factors controlling the time aspect of subsidence and the extent of deformation, respectively. (Haghighi and Motagh 2019)In long-term monitoring of Tehran plain using more than 400 interferograms and small baseline method found three discrete subsidence features in the west side of the region; Their results even showed that subsided area doesn't follow the trend of the major faults. However, the Pishva fault in Varamin controls the underground water flow and therefore shapes the subsidence bowl in this area.

Initial studies focusing on the correlation between subsidence and groundwater assume a linear relationship between those(Yu et al. 2020a). Many reliable studies point out that there is a solid positive correlation between land subsidence and groundwater level(Van Ty et al. 2021; Su et al. 2021; Zhao et al. 2021; Li et al. 2021). But as said earlier, we can consider any subject executing a change on effective stress among soil layer particles as a potential factor operating on land subsidence. In recent years, with the newly developed methods, research conducted on land subsidence and factors affecting this phenomenon has not been limited just to groundwater level change(Yong et al. 2013; Gong et al. 2018). (Chu, Ali, and Burbey 2021) propose a spatial regression model to study the nonlinear and non-stationary nature of land subsidence in the Choshui river alluvial fan (Taiwan) and identifies the zones where subsidence is inelastic/elastic based on computed coefficient. Using Artificial Neural Network (ANN) method with geology, slope, land use, and five other maps (Lee, Park, and Choi 2012) created a Ground Subsidence Susceptibility (GSS) map which accuracies were 94.84%-95.98% and found that the distance from the fault factor had the highest coefficient among others. In a similar study in Kerman, Iran, (Abdollahi et al. 2019) by employing a Supported Vector Machine (SVM) model with different kernels, proved that between 10 layers, the groundwater data, NDVI, and altitude had the most effect on land subsidence occurrence.

Tehran is the capital of Iran and is home to a more than 15 million population('Statistical Center of Iran' 2021). Many industrial infrastructures are located in this area which all of them are facing the threat of

land subsidence consequences in the following years. Many types of research are attended in Tehran Plain(Karimi, Motagh, and Entezam 2019; Bagheri et al. 2021; Haghighi and Motagh 2019). (Mahmoudpour et al. 2016) by developing a model tried to simulate the groundwater over-extraction and its designation in land subsidence using PMWIN; their results were calibrated by hydraulic and InSAR data. The model could be applied to predict land subsidence in the future. (Sharifikia 2011) by using a GIS tend to create the subsidence risk zonation and determine that how many Tehran residents are living in high-risk zones; they show that about 43 percent of the region is placed in the high-risk zone and even new constructions in this zone were continued without taking land subsidence risk into account.

Based on the groundwater level monitoring data from 2003 to 2005 and nine other spatial layers in Tehran Plain, the GWR and MLR models were employed to analyse the spatial distribution and contribution degree of each variable on the subsidence in the study area.

2. Materials and Methods

2.1 Study Area

Southwestern Tehran Plain is one of 150 plains in Iran that struggles with subsidence at the rate of 25 cm per year. Also, the subsided area expands rapidly (Motagh et al. 2007).Tehran Basin covers an area of 2250 km2 and has a semiarid-to-arid climate by an elevation decreasing trend from about 1600 m to 900 m in a north-south direction (Dehghani et al. 2013).Tehran Plain happens between Alborz Range to the north and Fashapouyeh and Arad Mountains to the south. The Plain extends from 35 49'-35 28'N and 51 6'-51 33'E. Tehran City has experienced a population boom since 1970 (7 million in 50 years) (SCI 2016) and a place for industries to grow. Therefore, the water demand for different utilization increased. Ever since groundwater exploitation has been the main water source in Tehran (Haghighi and Motagh 2019). Land subsidence was first reported by the National Cartographic Center (NCC) after performing precise levelling measurements in 2005 (Arabi 2005). As a place where population, industries, and infrastructures are concentrated, land subsidence cannot be neglected in this area. The study area shown in Figure1 has the most movement region in the vertical direction during the years 2003 to 2005, which encompasses almost 60 percent of Tehran Plain, and located at 35° 30′ N to 35° 42′ N latitude and 50° 55′ E to 51° 23′ E longitude. The aquifer's characteristics like the thickness of alluvium and the type of materials may vary by location. Tehran Plain aquitards mostly include silty clay and clayey soil, which have a suitable tendency to subside (Mahmoudpour et al. 2016).

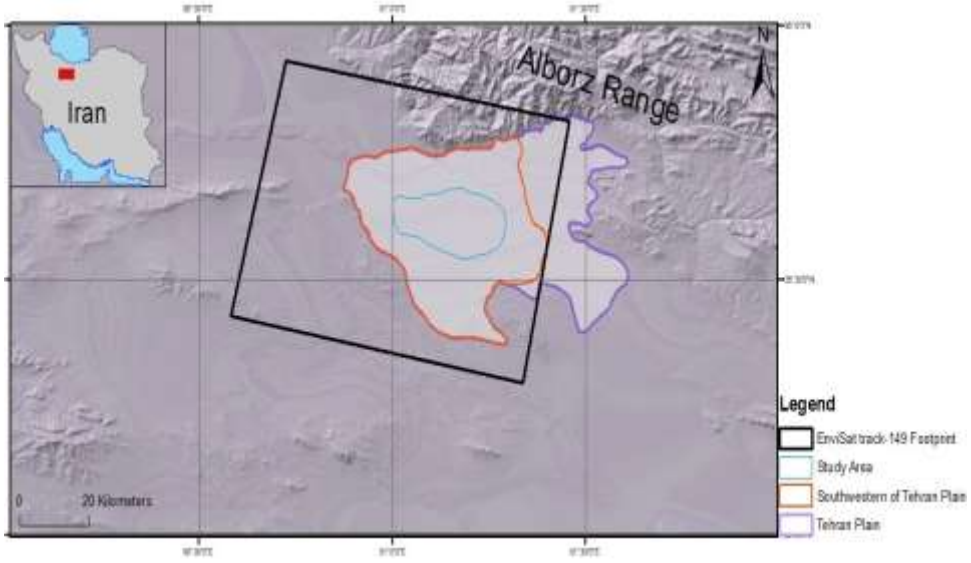


Figure1. Border of the Envisat SAR image and Study Area in western side of Tehran Plain

2.2 InSAR data

Haghighi in (Haghighi and Motagh 2019) analyzed the displacement time series by deploying the Small Baseline (SB) technique by applying 15 SAR images dataset (C-band) from the ASAR (Advanced Synthetic Aperture Radar) sensor onboard the European Space Agency (ESA) satellite Envisat in StripMap (SM) mode from July 2003 to May 2005 time interval in Descending orbit. This study used the spatially resampled of the cumulative subsidence achieved by Haghighi (Haghighi and Motagh 2019) in the 1000 m*1000 m grid. Accordingly, ASAR, an active radar on Envisat, provided information about the characteristics of the earth's surface (e.g., variation in surface height with sub-millimetre precision) from March 1st, 2002, until suddenly the satellite became unavailable in April 2012 (ESA 2002).

Figure 2 presents a set of regular resample points of InSAR data. To generate a surface of cumulative subsidence, an ordinary Kriging tool was executed over data points with a spherical model setting (see Figure 3). As shown in Figure 3, two subsidence bowls developing in the area can be seen. The displacement values for randomly sampled data are standardized to speed up the computational process.

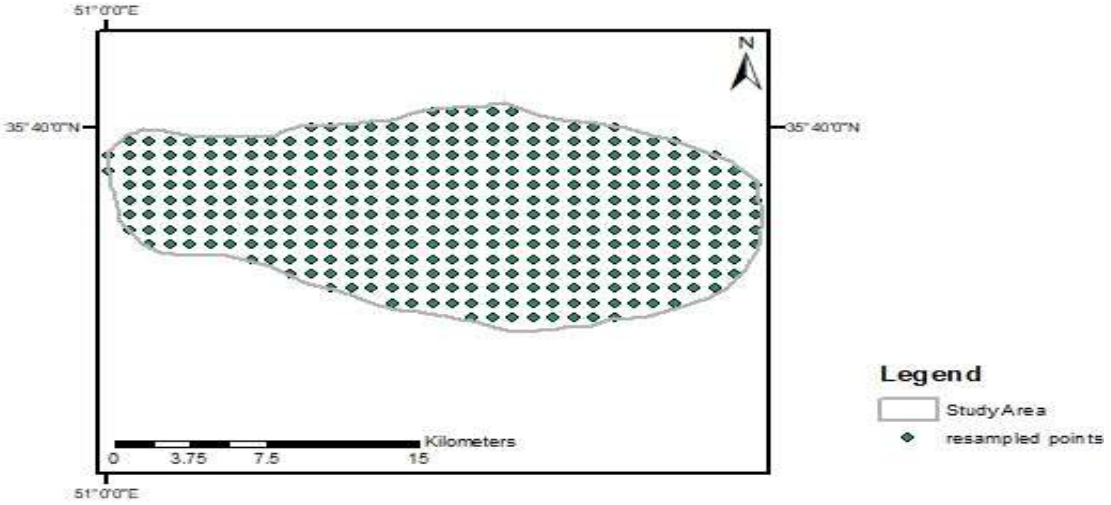


Figure 2: Spatially resampled of InSAR analysis in 1000*1000m grid

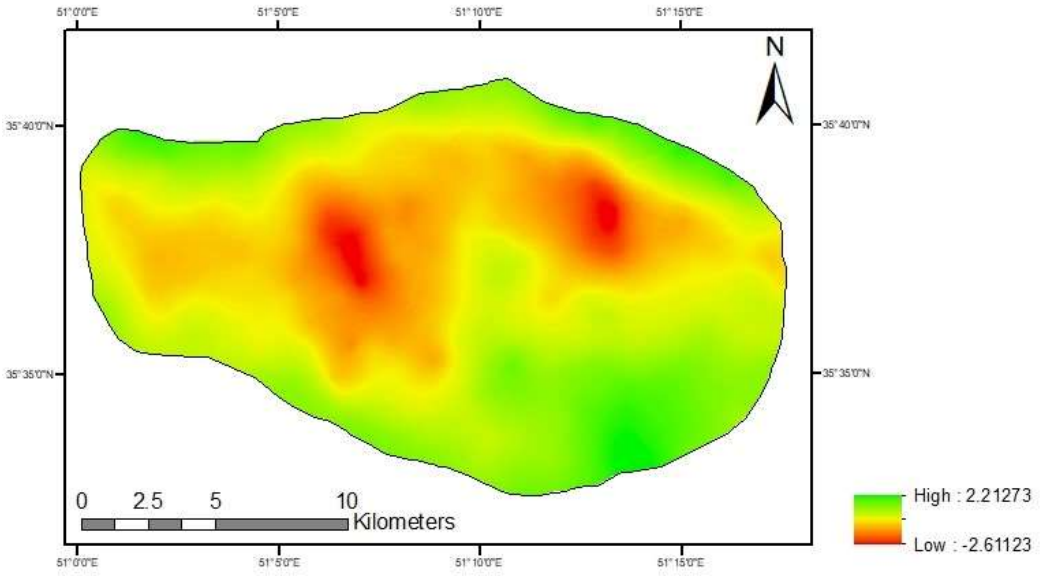


Figure 3. Land Subsidence pattern generated from 400 randomly distributed points

2.3 Ground Water Level

This study calculated the groundwater levels concerning sea mean level (SML) from the piezometric wells. Iran Water Resources Management (IWRM) collects piezometric wells data since March 1984 for Tehran. The data that be used as an independent variable in the regression is the water level change during the period of July 2003 to May 2005 that was superposed with InSAR data. The amount of variations in groundwater levels is illustrated in Figure 4. It can be derived by Figure 4 that the west and south side of the study area has been experiencing more variations than the north and east side.

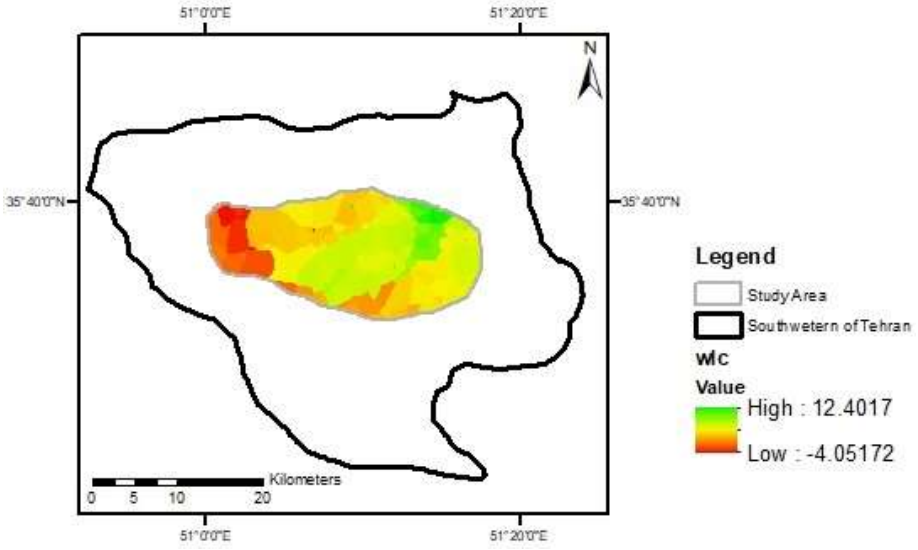
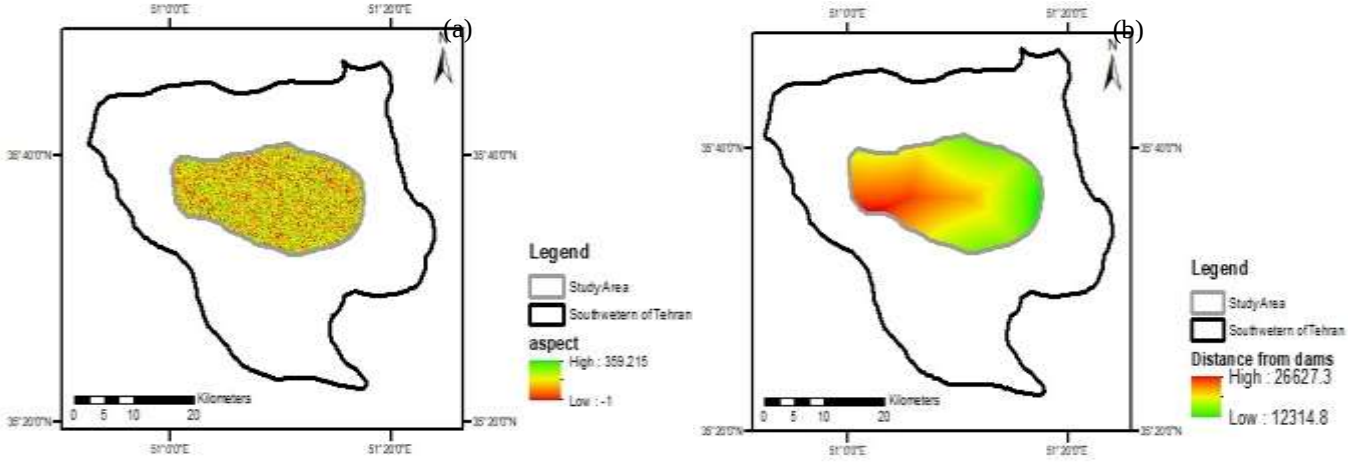
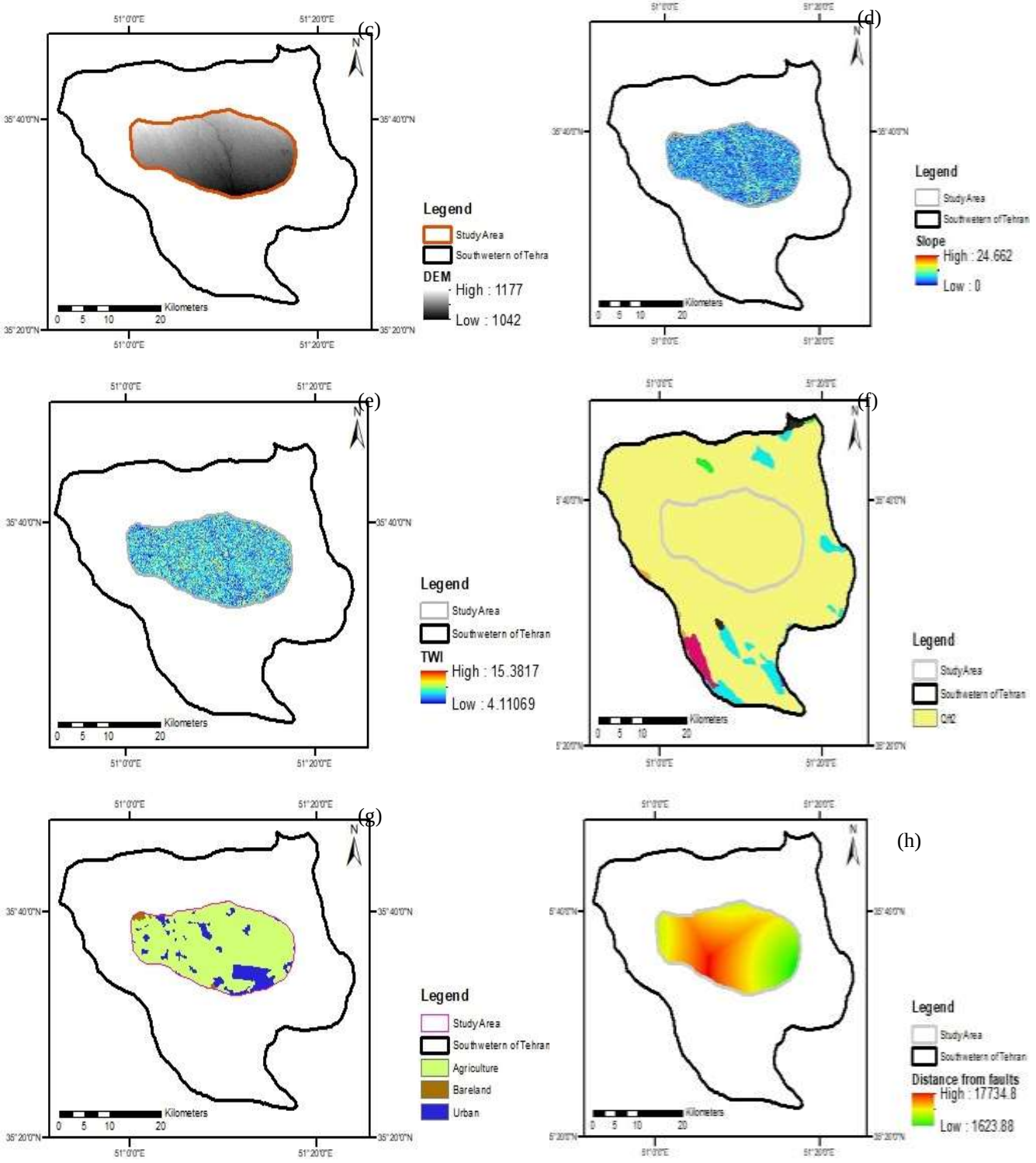


Figure 4. The surface of Water Level Change (WLC) in Study Area during 2003-2005 period gathered by IWRM

Other datasets

In this study, nine layers have been used as the explanatory variables in the regression models. All layers must be in a proper projected coordinate system, i.e., UTM. Figure 5 depicts the scope of each layer in the UTM coordinate system of zone 39. All Raw data employed in this study has been provided by the National Geoscience Database of Iran. Primary processes such as calculating slope, aspect, TWI, and Euclidean distances were provided in GIS environment.




(c)
Legend
Study Area
Southwetern of Tehra
DEM
High : 1177
Low : 1042
(d)
Legend
Study Area
Southwetern of Tehran
Slope
High : 24.662
Low : 0
(e)
Legend
Study Area
Southwetern of Tehran
TWI
High : 15.3817
Low : 4.11069
(f)
Legend
Study Area
Southwetern of Tehran
(g)
Legend
Study Area
Southwetern of Tehran
Agriculture
Bareland
Urban
(h)
Legend
Study Area
Southwetern of Tehran
Distance from faults
High : 17734.8
Low : 1623.88
Kilometers
0 5 10 20

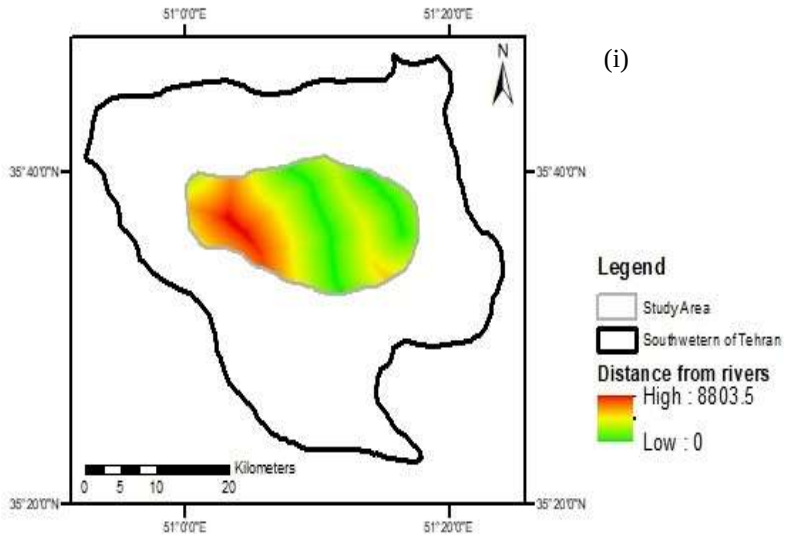


Figure 5. Explanatory layers implemented in the regression models: Aspect (a), Distance from dams (b), DEM (c), Slope (d), Topographic Wetness Index (e), Geology (f), Land use(g), Distance from Faults (h), Distance from Rivers (i)

2.4 Generating Sampled data

After assuring that there is no significant correlation among raster data used as explanatory variables, 360 random points were generated in the study area to take a sample of each raster at each location. The result is a *.dbf file, in which each column represents an independent variable sampled value. Figure 6 shows the distribute of 360 random points which covers the study area sufficiently. As shown in Figure 5, the Geology layer value is same for all locations, therefore the results of intercept term would be coupled with results for geology term too.

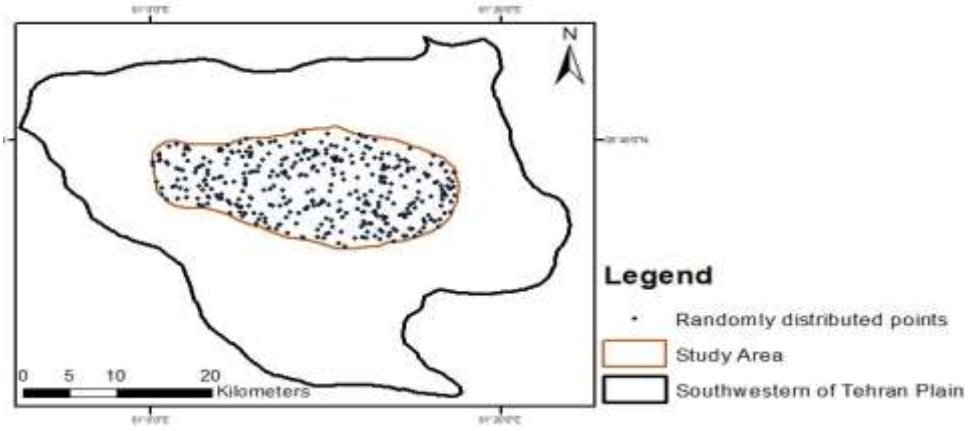


Figure 6. 400 random points used to take samples of dependent and independent variables

2.5 Multiple linear regression model

It is an extension of the ordinary least-square (OLS) model that comprises multiple explanatory variables. We express Multiple Linear Regression (MLR) as (Fotheringham, Brunsdon, and Charlton 2003):

$$y_i = \beta_0 + \sum_{k=1}^{i} \beta_k X_{ik} + \varepsilon_i \quad (1)$$

here y_i is the predicted value of subsidence by the model, β_0 stands for the intercept term, β_k as the coefficient of the kth explanatory variable, x_{ik} is the kth variable value at observation point i, and ε_i as random term associated with observation i. MLR assumes that there is a linear relationship between parameters involving in the model. This model lies in a sort of model called the global model, in which the model used considers all the data as a whole and tries to discover how the response variable Y and explanatory variables are related.

2.6 Geography Weighted Regression model

Compared to other R-based packages like spgwr and GWmodel, the mgwr is a Python-based package developed by Oshan (Oshan 2018), which provides both GWR and MGWR models along with more diagnostics for model goodness of fit for the user (Li and Fotheringham 2020). The MLR might encounter some issues while modelling spatial data. The spatial heterogeneity (process variation by spatial context) as an inherent property of spatial data causes the estimation of MLR model coefficients to be inaccurate in most cases. Unlike, the GWR model captures the process's spatial heterogeneity at each regression point by taking into account the spatial position of each coefficient. GWR model formula is as follows (Fotheringham, Brunsdon, and Charlton 2003):

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{i} \beta_{ik}(u_i, v_i) X_{ik} + \varepsilon_i \quad (2)$$

Here, y_i is the fitted subsidence value by the GWR. $\beta_0(u_i, v_i)$ is the intercept term at location (u_i, v_i), $\beta_{ik}(u_i, v_i)$ is the kth explanatory variable's coefficient at location (u_i, v_i), x_{ik} is the kth explanatory variable value at location i, and ε_i is the analogous random error term at location i. For each calibration point i, GWR borrows a set of observation points within a particularized bandwidth. Through the equation below, the coefficient at each location can be estimated as follows (Fotheringham, Brunsdon, and Charlton 2003):

$$\widehat{\beta}_i(u_i, v_i) = \left(X^T W(u_i, v_i) X\right)^{-1} X^T W(u_i, v_i) Y \quad (3)$$

where $\widehat{\beta}_i(u_i, v_i)$ is the estimated local regression coefficient, X is a matrix of explanatory variables, $W(u_i, v_i)$ is a diagonal weight matrix based on the distance between each observation and calibration point, Y is a vector of subsidence value at each observation locations intended in weight matrix bandwidth.

2.7 Kernel function

As Tobler's first law of Geography notes: "Everything is related to everything else, but near things are more related than distant things" therefore, a suitable kernel may be the one that reflects this phrase in the best shape. Figure 7 shows the three most used kernel functions in GWR and their performance when dealing with various bandwidths. In the case of Exponential and Gaussian Kernels, the problem would be that regardless of how far the observation point is from the calibration point, still, it can alter the coefficient estimating process. For scenarios of dense and sparse observations, correspond kernel function types are developed, fixed, and adaptive. Both might operate fine with dense captured data, but in the case of sparsely distributed observation points, an adaptive kernel handles that situation better than a fixed one

by selecting an optimal number of nearest neighbours for each calibration. Hence, the bandwidth would vary from location to location.

The bi-square kernel is immune to the mentioned obstacles. Therefore, it is the most common kernel. However, in this paper, a fixed Gaussian kernel is applied to data because the purpose is merely to assert the GWR model efficiency.

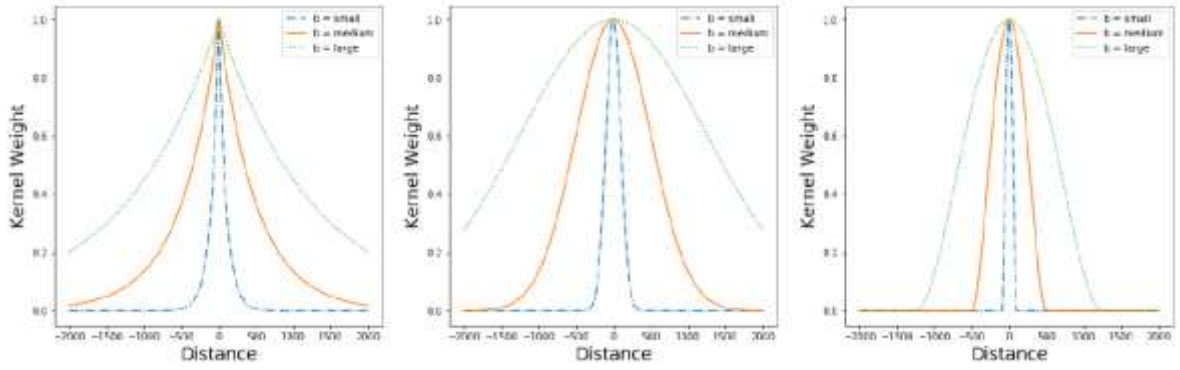


Figure 7. Examples of exponential kernels (top), Guassian kernels (middle), and bi-square kernels (bottom) for a small, medium, and large bandwidth parameter (Oshan et al. 2019)

2.8 Bandwidth

It has been proved that kernel bandwidth has more influence on outcomes accuracy than kernel function itself. Then, more attention should be paid to the bandwidth selection criteria. Typically, the bandwidth selection process proceeds by searching for an optimized value generated by a specific formula. In our study, a corrected Akaike Information Criteria (AICc), suggested by Oshan (Oshan et al. 2019), was used between other general model fit criteria (e.g., CV and BIC). The task of searching for the optimal value passes through a Golden Section procedure that regularly squeezes the range where optimal value happens and measures each step value to the last step till succeeding the lowest score.

2.9 Diagnostics

An easy way to compare two model performance would be to compare their AICc value. If the two model AICc value difference is less than three, the conclusion would be that the two models are the same. However, mapping the residual or fitted values provide the capability of interpreting the results more intuitively. In addition, the deployment of each local regression's local R2 is another option to measure model capability provided by the mgwr package. Also, statistical terms like t-value, p-value, and standard error (SE) would be calculated for every point.

3.Results and Discussion

Global model

A surface is generated using the kriging interpolation method, based on the spatially resampled data from time series analysis of 15 Envisat track 149 ASAR images in 1km*1km grids (Figure 3). Ultimately, the sampled 400 randomly distributed points as the "y" component were contributed in the regression.

Model coefficients are estimated by calibrating the MLR model, giving all layers mentioned in Figure 5. Table 1 displays the result of a global regression. According to adjusted R2, the model explains approximately 39 percent of variation corresponds to the dependent variable. Table 2 shows each variable's estimation by the MLR and its associated statistics. If we define the effective variables as those whose p-values are less than 0.01, according to the MLR statistics, we can conclude that the following terms are not influential in the occurs of land subsidence, including Intercept, TWI, Slope, Euclidean

distance from the fault, and Aspect.

Just Euclidean distance from the river and land use have a positive coefficient between remaining terms. That means these two terms and subsidence change in the same direction. Table 2 also shows the corrected AIC values for both global and local models.

Table 1. The Multi Linear Regression results (statistical diagnostics)

Residual sum of squares	245.053
AIC	959.155
AICc	961.835
R^2	0.387
Adjusted R^2	0.373

Table 2. Explanatory variables statistics

Variable	*Estimated Value*	*Standard Error*	*t-statistic*	*p-value*
Intercept	0.000	0.040	0.000	1.000
WLC	-0.281	0.049	-5.774	0.000
TWI	-0.021	0.047	-0.435	0.651
Slope	0.009	0.046	0.186	0.853
River	0.186	0.058	3.199	0.001
Land Use	0.144	0.043	3.380	0.001
Distance from faults	-0.148	0.060	-2.478	0.013
DEM	-0.325	0.044	-7.467	0.000
Distance from dams	-0.378	0.075	-5.058	0.000
Aspect	-0.014	0.041	-0.335	0.738

Local model

A fixed Gaussian kernel is used as the weighting function. To determine the optimal value of bandwidth golden section search method minimizes the model's AICc value. By comparing the AICc value of the two models, we can conclude the dominant superiority of the GWR model as a regressor for spatial data rather than usual global models such as MLR. In order to check for the goodness-of-fit, Table 3 outlines statistical diagnostics of applied models. The deviation between the model predicted value and the observed value is known as model residual at each point. Figure 8 depicts the residuals associated with GWR and MLR. The GWR model's strength can be derived by comparing the outcomes intuitively.

Table 3. The Geographically Weighted Regression results (statistical diagnostics)

Residual sum of squares	32.019
AIC	312.092
AICc	369.932o
R^2	0.920
Adjusted R^2	0.896

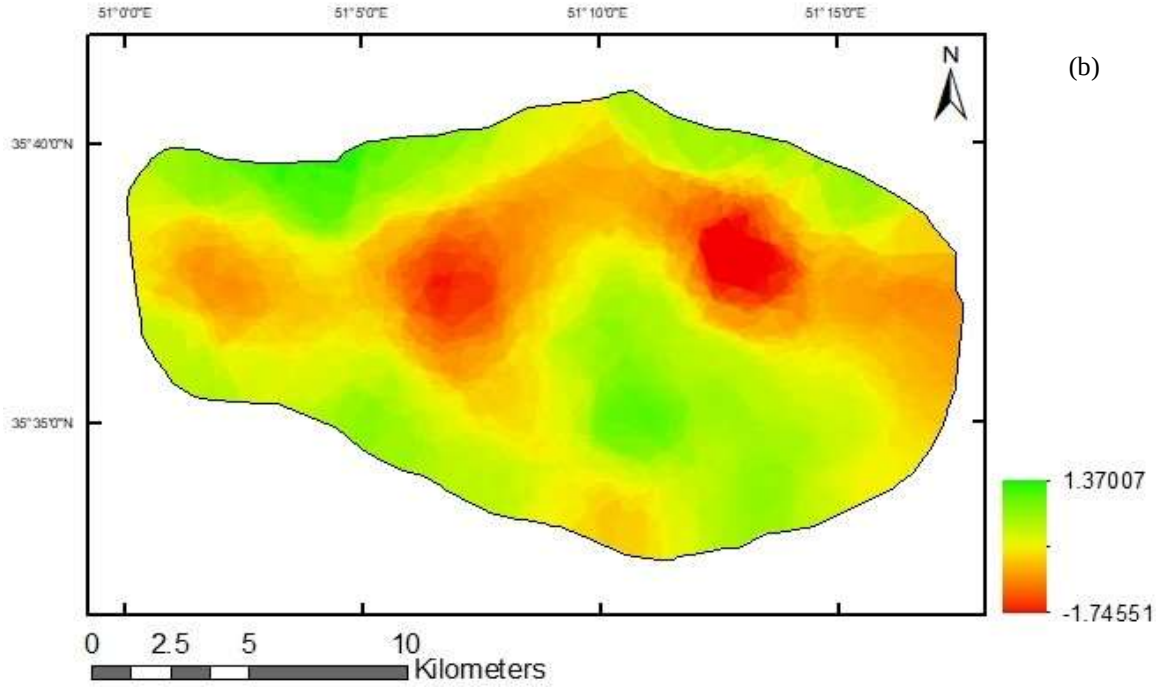


Figure 8. The difference between observed values and the model predicted values for: (a) GWR model (b) MLR model

The adjusted R2 value means if an alter in the independent value improves the dependent variable's value in the model, only in that case we can increase the R2 value. In other words, only if the new appended term enhances the model-fit more than the expected chance, it indicates that the variable can explain some of the model variations. Figure 9 maps the local R2 value of the GWR model so that we can comment on the model's ability to define the variation in each regression point.

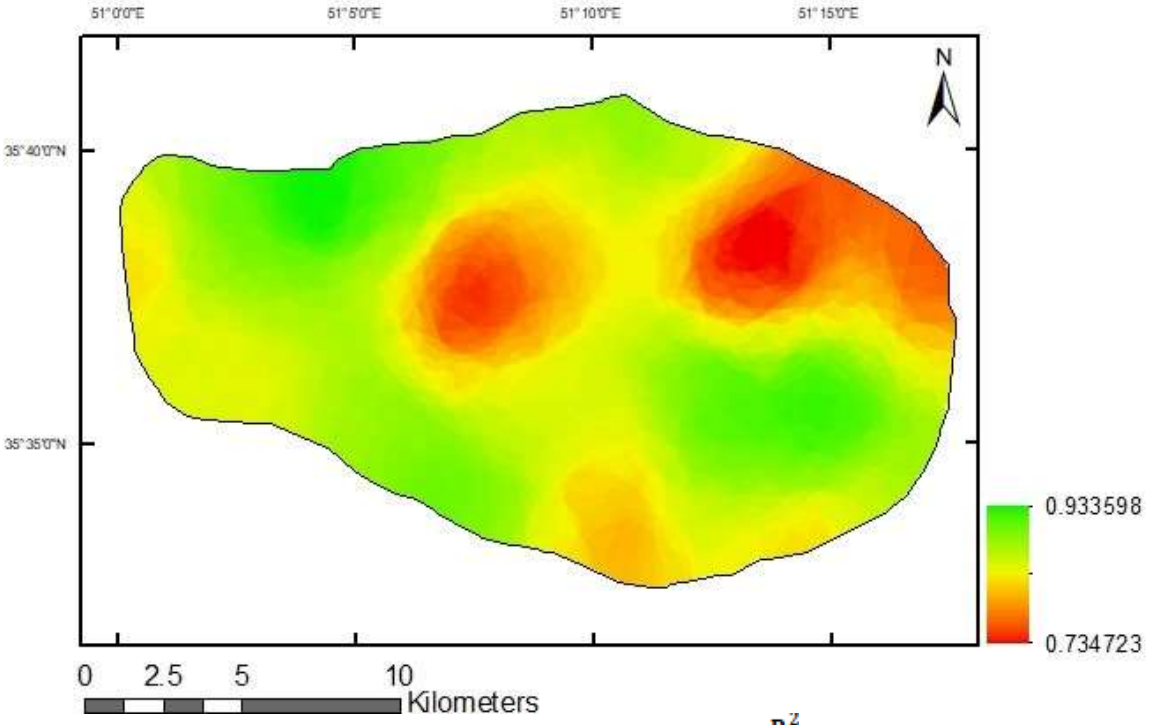


Figure 9. The computed surface of local $\mathbf{R^2}$ statistic

4.Conclusions

This study investigated the cumulative vertical movement of Southwestern Tehran Plain from 2003 through 2005 to explore the significant variables causing land subsidence. We manifest two global and local approaches for regression and explores their performance towards the spatial data.

To evaluate both the MLR and GWR model's fulfilment, we manage to deploy some statistical diagnostics by measuring the goodness-of-fit. The results introduce the GWR model as the best regressor while dealing with spatial data, that is because GWR takes the location of data into account and compiles the process with it. On the GWR model side, this study specified the maximum coefficient for each calibration point and have concluded that just six out of 10 variables get to be the dominant factor. We inferred that "Intercept" presents for 31%, "DEM" presents for 37%, "Distance from fault" presents for 17%, "Distance from dam" presents for 9%, "Distance from River" presents for 6%, and least of all "Water Level Change" in 1% of all 400 points. But on the other side, the MLR model introduces "Distance from dam" as the dominant factor and most responsible factor for land subsidence, which in terms of contribution degree is ranked four by the GWR model.

The reason for Global models (e.g., MLR) ineffectiveness in spatial data modelling is because these models consider spatial homogeneity while we are informed that spatial autocorrelation and spatial non-stationary are two prominent features of the spatial data.

References

Abdollahi, Sahar, Hamid Reza Pourghasemi, Gholam Abbas Ghanbarian, and Roja Safaeian. 2019. 'Prioritization of effective factors in the occurrence of land subsidence and its susceptibility mapping using an SVM model and their different kernel functions', *Bulletin of Engineering Geology and the environment*, 78: 4017-34.

Akbari, Vahid, and Mahdi Motagh. 2011. 'Improved ground subsidence monitoring using small baseline SAR interferograms and a weighted least squares inversion algorithm', *IEEE Geoscience and Remote Sensing Letters*, 9: 437-41.

Arabi, Montazerian, A.R., Maleki, E. , Talebi, A. . 2005. "Study of land subsidence in south-west of Tehran plain basin." In, 14-28. Natianal Cartographic Center.

Bagheri, Mehdi, Seiyed Mossa Hosseini, Behzad Ataie-Ashtiani, Yasamin Sohani, Homa Ebrahimian, Faezeh Morovat, and Shervin Ashrafi. 2021. 'Land subsidence: A global challenge', *Science of The Total Environment*: 146193.

Blainey, Simon, and Corinne Mulley. 2013. "Using Geographically Weighted Regression to forecast rail demand in the Sydney Region." In *Australasian Transport Research Forum, Brisbane*.

Brunsdon, Chris, A Stewart Fotheringham, and Martin E Charlton. 1996. 'Geographically weighted regression: a method for exploring spatial nonstationarity', *Geographical analysis*, 28: 281-98.

Brunsdon, Chris, Stewart Fotheringham, and Martin Charlton. 1998. 'Geographically weighted regression', *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47: 431-43.

Castellazzi, Pascal, Norma Arroyo-Domínguez, Richard Martel, Angus I Calderhead, Jonathan CL Normand, Jaime Gárfias, and Alfonso Rivera. 2016. 'Land subsidence in major cities of Central Mexico: Interpreting InSAR-derived land subsidence mapping with hydrogeological data', *International Journal of Applied Earth Observation and Geoinformation*, 47: 102-11.

Chasco, Coro, Isabel García, and José Vicéns. 2007. 'Modeling spatial variations in household disposable income with Geographically Weighted Regression'.

Chen, Beibei, Huili Gong, Kunchao Lei, Jiwei Li, Chaofan Zhou, Mingliang Gao, Hongliang Guan, and Wei Lv. 2019. 'Land subsidence lagging quantification in the main exploration aquifer layers in Beijing plain, China', *International Journal of Applied Earth Observation and Geoinformation*, 75: 54-67.

Chu, Hone-Jay, Muhammad Zeeshan Ali, and Thomas J Burbey. 2021. 'Development of spatially varying groundwater-drawdown functions for land subsidence estimation', *Journal of Hydrology: Regional Studies*, 35: 100808.

Dehghani, Maryam, Mohammad Javad Valadan Zoej, Andrew Hooper, Ramon F Hanssen, Iman Entezam, and Sassan Saatchi. 2013. 'Hybrid conventional and persistent scatterer SAR interferometry for land subsidence monitoring in the Tehran Basin, Iran', *ISPRS journal of photogrammetry and remote sensing*, 79: 157-70.

ESA. 2002. 'ASAR Instrument'. https://earth.esa.int/eogateway/instruments/asar.

Fotheringham, A Stewart, Chris Brunsdon, and Martin Charlton. 2003. *Geographically weighted regression: the analysis of spatially varying relationships* (John Wiley & Sons).

Galloway, Devin L, and Thomas J Burbey. 2011. 'Regional land subsidence accompanying groundwater extraction', *Hydrogeology Journal*, 19: 1459-86.

Gong, Huili, Yun Pan, Longqun Zheng, Xiaojuan Li, Lin Zhu, Chong Zhang, Zhiyong Huang, Zhiping Li, Haigang Wang, and Chaofan Zhou. 2018. 'Long-term groundwater storage changes and land subsidence development in the North China Plain (1971–2015)', *Hydrogeology Journal*, 26: 1417-27.

Haghighi, Mahmud Haghshenas, and Mahdi Motagh. 2019. 'Ground surface response to continuous compaction of aquifer system in Tehran, Iran: Results from a long-term multi-sensor InSAR analysis', *Remote sensing of environment*, 221: 534-50.

Hoffmann, Jörn, Howard A Zebker, Devin L Galloway, and Falk Amelung. 2001. 'Seasonal subsidence and rebound in Las Vegas Valley, Nevada, observed by synthetic aperture radar interferometry', *Water Resources Research*, 37: 1551-66.

Hu, Leyin, Keren Dai, Chengqi Xing, Zhenhong Li, Roberto Tomás, Beth Clark, Xianlin Shi, Mi Chen, Rui Zhang, and Qiang Qiu. 2019. 'Land subsidence in Beijing and its relationship with geological faults revealed by Sentinel-1 InSAR observations', *International Journal of Applied Earth Observation and Geoinformation*, 82: 101886.

Hu, RL, ZQ Yue, LC u Wang, and SJ Wang. 2004. 'Review on current status and challenging issues of land subsidence in China', *Engineering Geology*, 76: 65-77.

Karimi, L, M Motagh, and I Entezam. 2019. 'Modeling groundwater level fluctuations in Tehran aquifer: results from a 3D unconfined aquifer model', *Groundwater for Sustainable Development*, 8: 439-49.

Lee, Saro, Inhye Park, and Jong-Kuk Choi. 2012. 'Spatial prediction of ground subsidence susceptibility using an artificial neural network', *Environmental management*, 49: 347-58.

Li, Ming-Guang, Jin-Jian Chen, Ye-Shuang Xu, Da-Gui Tong, Wei-Wei Cao, and Yu-Jin Shi. 2021. 'Effects of groundwater exploitation and recharge on land subsidence and infrastructure settlement patterns in shanghai', *Engineering Geology*, 282: 105995.

Li, Ziqi, and A Stewart Fotheringham. 2020. 'Computational improvements to multi-scale geographically weighted regression', *International Journal of Geographical Information Science*: 1-20.

Mahmoudpour, Masoud, Mashalah Khamehchiyan, Mohammad Reza Nikudel, and Mohammad Reza Ghassemi. 2016. 'Numerical simulation and prediction of regional land subsidence caused by groundwater exploitation in the southwest plain of Tehran, Iran', *Engineering Geology*, 201: 6-28.

Motagh, Mahdi, Yahya Djamour, Thomas R Walter, Hans-Ulrich Wetzel, Jochen Zschau, and Siavash Arabi. 2007. 'Land subsidence in Mashhad Valley, northeast Iran: results from InSAR, levelling and GPS', *Geophysical Journal International*, 168: 518-26.

Oshan, Taylor. 2018. 'Multi-Scale Geographically Weighted Regression', Github. https://github.com/pysal/mgwr.

Oshan, Taylor M, Ziqi Li, Wei Kang, Levi J Wolf, and A Stewart Fotheringham. 2019. 'mgwr: A Python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale', *ISPRS International Journal of Geo-Information*, 8: 269.

SCI. 2016. 'Population and Housing Censuses', Government of Iran. https://www.amar.org.ir/english/Population-and-Housing-Censuses.

Sharifikia, Mohammad. 2011. 'Evaluation of land subsidence related disasters in plains and residential areas of Iran', *Scientific Quarterly journal of Iranian Association of Engineering Geology*, 3: 43-58.

Shemshaki, A, MJ Boulourchi, and I Entezam Soltani. 2006. "The study of land subsidence in Tehran plain and its casual factors." In *24th earth sciences meeting, Geological survey and mineral explorations of Iran. Retrieved July*, 2019.

'Statistical Center of Iran'. 2021. https://www.amar.org.ir/.

Su, Guangli, Yanqiang Wu, Wei Zhan, Zhijiang Zheng, Liu Chang, and Jiaqing Wang. 2021. 'Spatiotemporal evolution characteristics of land subsidence caused by groundwater depletion in the North China plain during the past six decades', *Journal of hydrology*, 600: 126678.

Sultana, Selima, Nastaran Pourebrahim, and Hyojin Kim. 2018. 'Household energy expenditures in North Carolina: a geographically weighted regression approach', *Sustainability*, 10: 1511.

Van Ty, Tran, Huynh Vuong Thu Minh, Ram Avtar, Pankaj Kumar, Huynh Van Hiep, and Masaaki Kurasaki. 2021. 'Spatiotemporal variations in groundwater levels and the impact on land subsidence in CanTho, Vietnam', *Groundwater for Sustainable Development*: 100680.

Yong, Yang, Zheng Fandong, Liu Licai, Wang Sufen, and Wang Rong. 2013. 'Study on the correlation between groundwater level and ground subsidence in Beijing plain areas [J]', *Geotechnical Investigation & Surveying*, 8.

Yu, Hairuo, Huili Gong, Beibei Chen, Kaisi Liu, and Mingliang Gao. 2020a. 'Analysis of the influence of groundwater on land subsidence in Beijing based on the geographical weighted regression (GWR) model', *Science of The Total Environment*, 738: 139405.

———. 2020b. 'Analysis of the influence of groundwater on land subsidence in Beijing based on the geographical weighted regression (GWR) model', *Science of The Total Environment*: 139405.

Zeitoun, David G, and Eliyahu Wakshal. 2013. *Land subsidence analysis in urban areas: the Bangkok metropolitan area case study* (Springer Science & Business Media).

Zhao, Yu, Chaolin Wang, Jinqiang Yang, and Jing Bi. 2021. 'Coupling model of groundwater and land subsidence and simulation of emergency water supply in Ningbo urban Area, China', *Journal of hydrology*, 594: 125956.