Civil Engineering Faculty Publications and Presentations

Department of Civil Engineering

A Hybrid Clustering-Fusion Methodology for Land Subsidence Estimation

Narges Taravatrooy
Yasouj University

Mohammad Reza Nikoo
Shiraz University

Mojtaba Sadegh
Boise State University

Mohammad Parvinnia
Yasouj University

A Hybrid Clustering-Fusion Methodology for Land Subsidence Estimation

Narges Taravatrooy
School of Engineering
Department of Civil and Environmental
Engineering
Shiraz University
Shiraz, Iran

Mohammad Reza Nikoo*
School of Engineering
Department of Civil and Environmental
Engineering
Shiraz University
Shiraz, Iran
nikoo@shirazu.ac.ir

Mojtaba Sadegh
Department of Civil Engineering
Boise State University
Boise, ID, U.S.

Mohammad Parvinnia
Department of Civil and Environmental
Engineering
Yasouj University
Yasouj, Iran.

Abstract

A hybrid clustering-fusion methodology is developed in this study that employs Genetic Algorithm (GA) optimization method, k-means method, and several soft computing (SC) models to better estimate land subsidence. Estimation of land subsidence is important in planning and management of groundwater resources to prevent associated catastrophic damages. Methods such as the Persistent Scatterer Interferometric Synthetic Aperture Radar (PS-InSAR) can be used to estimate the subsidence rate, but PS-InSAR does not offer the required efficiency and accuracy in noisy pixels (obtained from remote sensing). Alternatively, a fusion-based methodology can be used to estimate subsidence rate, which offers a superior accuracy as opposed to the traditionally used methods. In the proposed methodology, five SC methods are employed with hydrogeological forcing of frequency and thickness of fine-grained sediments, groundwater depth, water level decline, transmissivity and storage coefficient, and output of land subsidence rate. Results of individual SC models are then fused to render more accurate land subsidence rate in noisy pixels, for which PS-InSAR cannot be effective. We first extract 14,392 different input-output patterns from PS-InSAR technique for our study area in Tehran province, Iran. Then, k-means method is used to divide the study area to homogenous zones with similar features. The five SC models include Adaptive Neuro Fuzzy Inference System (ANFIS), Support Vector Regression (SVR), Multi-Layer Perceptron (MLP) neural network and two optimized models, namely, Radial Basis Function (RBF) and Generalized Regression Neural Network (GRNN). To fuse individual SC models, three methods including Genetic Algorithm (GA), K-Nearest Neighbors (KNN) and Ordered Weighted Average (OWA) based on ORNESS method and ORLIKE method, are developed and evaluated. Results show that the fusion-based method is significantly superior to each of the employed individual methods in predicting land subsidence rate.

Keywords: land subsidence rate, model fusion, Genetic Algorithm (GA), K-nearest neighbors algorithm (KNN), Ordered Weighted Average (OWA), persistent scatterer interferometry

1. Introduction

Metropolitan and agricultural development increase groundwater resources withdrawal, which in turn poses serious environmental challenges. Unregulated and excessive groundwater extraction for agricultural, domestic and industrial use have resulted in severe drop in groundwater table in several basins in Iran (Motagh et al. 2008; Sadegh et al. 2010; Sadegh and Kerachian, 2011). Decline in groundwater level increases the effective stress in the aquifer system that promote compaction in fine-grained sediments (Budhu and Adiyaman 2009; Dehghani et al. 2013), which in turn prompts land subsidence. In addition to groundwater level decline, other geology and hydrogeology factors can affect subsidence rate, including gas, oil and geothermal water extraction (Gambolati et al. 2005), coal mining (Jung et al. 2007) and sudden hydrogeological changes along faults (Burbey 2002).

Precise estimation of land subsidence provides helpful information to decision makers in their efforts to control and mitigate the impacts of such a grave hazard. Satellites have provided alternative land subsidence monitoring methods complementing in situe observations based on remote sensing techniques. In the previous decades, several studies have performed monitoring and analyzing land subsidence due to groundwater withdrawal based on observations from satellites and radars such as Environmental Satellite Advanced Synthetic Aperture Radar (ENVISAT ASAR) (Osmanoglu et al. 2011; Yue et al. 2011; Ng et al. 2012; Dehghani et al. 2013; Strozzi et al. 2017; Deng et al. 2017; Lu et al. 2018; Du et al. 2018). Interferometry Synthetic Aperture Radar (InSAR) is one such technique that provides accurate measurements of land subsidence (Amelung et al. 1999; Carnec and Fabriol 1999; Nakagawa et al. 2000; Ding et al. 2004; Dehghani et al. 2009; Yu et al. 2011; Calderhead et al. 2011;Cigna et al. 2012; Teatini et al. 2012; Qu et al. 2014; Strozzi et al. 2017; Lu et al. 2018; Du et al. 2018; Nadiri et al. 2018). Another such techinque is Persistant Scatterer Interferometric Synthetic Aperture Radar (PS-InSAR) that is recently developed to address the decorrelation problem in land subsidence estimation, and is widely used in the literature (Jung et al. 2007; Osmanoglu et al. 2011; Cigna et al. 2012; Teatini et al. 2012; Dehghani et al. 2013; Strozzi et al. 2013; Wu and Hu 2016; Sun et al. 2017; Deng et al. 2017; Maghsoudi et al. 2018). However, although the remotely sensed techniques offer valuable opportunities and advanteges for land subsidence estimation as opposed to traditional in situe observations, a potential drawback associated with remote sensing is lack of accurate and detailed information about subsidence in noisy pixels (Dehghani et al., 2013). In view of the fact that the PS-InSAR technique is based on persistent scatterer points, noisy behavior could be observed in many points of the study area especially in non-urban regions (Gehlot and Hanssen, 2008). In such cases, an approach based on soft computing models can be effectively utilized to estimate land subsidence. For example, Artificial Neural Networks (ANNs) have been used to monitor subsidence in various studies. In Table 1 related studies on subsidence analysis using soft computing models are sumarized.

Table 1. Summary of related studies on subsidence analysis using soft computing models based on artificial intelligence neural network.

In a closely related effort to this study, Dehghani et al. (2013) studied land subsidence due to groundwater extraction in Tehran basin, Iran. They considered six hydrogeological variables as Multi-Layer Perceptron (MLP) model's forcing to estimate subsidence rate. In their study, forcing of the MLP model is not classified, and the subsidence rate is estimated only by one soft computing model (MLP). Hence, the developed model is not sufficiently precise as evidenced by relatively high Root Mean Square Error (RMSE). To improve the accuracy of land subsidence rate estimation one can cluster forcing data and train a separate model for each cluster, and/or employ several soft computing models and fuse their estimations.

In the past years, the increasing demand for enhanced accuracy of soft computing (SC) models has stimulated researchers to develop fusion-based methods. Model fusion is the procedure of gathering data from several models such as different individual SC models' outputs, aiming to provide more precise and reliable information compared to each individual model (Dasarathy 1997). Fusion-based methods, such as Bayesian Model Averaging, have been recently used in different research areas, namely, drought index estimation (Azmi et al. 2016; Alizadeh and Nikoo, 2018), river-level forcasting (See and Abrahart 2001) and hydrological engineering (Shu and Burn 2004; Duan et al. 2007; Ajami et al. 2007; Azmi et al. 2010; Ashouri et al. 2015). But to the best of authors knowledge, there is not any study on estimation of land subsidence rate using fusion-based methods. The

importance of this phenomenon motivated the authors to develop a fusion model in order to achieve more precision in land subsidence estimation. Therefore, in this study a fusion-based methodology is developed based on five individual soft computing (SC) models, which are subsequently fused using Genetic Algorithm (GA), K-nearest neighbors (KNN) method and Ordered Weighted Average (OWA) method. Five SC models, namely, Adaptive Neuro Fuzzy Inference System (ANFIS), Support Vector Regression (SVR), Multi-Layer Perceptron (MLP), and two optimized models based on Genetic Algorithm (GA) including Radial Basis Function (RBF) and Generalized Regression Neural Network (GRNN) were employed to estimate subsidence rate. The input variables of all SC models are six effective hydrogeological variables and the output is the subsidence rate derived from PS-InSAR. Then, k-means is utilized for dividing the study area to clusters (homogenous zones) with similar features. 70% and 30% of each cluster's data are used for training and validating the individual and fusion-based models, respectively. Four fusion methods based on Genetic Algorithm (GA) optimization method, K-Nearest Neighbors (KNN) and Ordered Weighted Average (OWA) models are then developed to fuse the outputs of individual SC models. The latter (OWA) consists of two sub-models, namely, ORNESS and ORLIKE methods. The main novelty of proposed methodology is developing several fusion-based models as well as optimized soft computing (SC) models to achieve the best possible result in land subsidence estimation. In the next sections main parts of proposed methodology, case study and results are presented.

2. Methodology

Flowchart of the fusion-based methodology which proposed for land subsidence rate estimation is presented in Fig. 1. The proposed methodology consists of five main steps. In the first step, the data are prepared and derived from ENVISAT ASAR and PS-InSAR. In the next step, to obtain more precision in subsidence rate estimation, the k-means method is used to classify data. Then, each cluster's data are randomly separated as train and validation sets. In the third step, five individual Soft Computing (SC) models, namely, ANFIS, SVR, MLP and two optimized models, namely, RBF and GRNN are developed. In the fourth step, four different methods including KNN, GA and two OWA-based models, namely, ORNESS and ORLIKE are utilized as fusion methods for analysis and estimation of land subsidence rate based on PS-InSAR data. Finally, the results of SC models and model fusion methods are assessed and compared through different statistical error indices. In the next sections, the main steps of proposed methodology are briefly described.

Fig. 1. Flowchart of fusion-based methodology for land subsidence rate estimation.

2.1. K-Means Method

This algorithm classifies data into several homogenous clusters with similar features. k-means method initializes the center of k clusters by random search in each iteration and subsequently measures the distances between data points (x_{ij}) and the centers (c_j). So, by minimizing the objective function specified in Eq. 1, this algorithm assigns cluster k to data point x_{ij} (MacQueen 1967).

$$\text{Minimize: } d = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_{ij} - c_j \right\|^2 \tag{1}$$

In this study, we have tried different number of iterations to avoid converging to local optima by the k-means algorithm, and successively increased the number of iterations until the result not change anymore. We also have repeated the k-mean algorithm several times, each time setting the initial centroid point at the previous optimized points to ensure k-means reached a global optimum.

2.2. Soft Computing (SC) Models

In order to reach the optimal network architecture of individual SC models, various values of models' effective parameters (model settings) were optimized with two approaches: 1. trial-and-error analysis for MLP, SVR and ANFIS, and 2. Genetic Algorithm (GA) optimization model for RBF and GRNN. Eventually, SC model structures were selected that provided superior results according to several statistical error indices such as Scatter Index (SI), Root Mean Square Error (RMSE), Root Mean Relative Error (RMRE), Nash-Sutcliffe (NS) efficiency, Correlation Coefficient (CC) and Bias:

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(o_i - e_i) \tag{2}$$

$$SI = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}((e_i - E) - (o_i - O))^2}}{O} \tag{3}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(o_i - e_i)^2} \tag{4}$$

$$CC = \frac{\sum_{i=1}^{n}(o_i \times e_i)}{\sqrt{\sum_{i=1}^{n}o_i^2 \sum_{i=1}^{n}e_i^2}} \tag{5}$$

$$NS = 1 - \frac{\sum_{i=1}^{n}(o_i - e_i)^2}{\sum_{i=1}^{n}(o_i - O)^2} \tag{6}$$

$$RMRE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|\frac{o_i - e_i}{e_i}\right|} \tag{7}$$

Where, e_i and o_i are respectively i^{th} estimated and observed subsidence rate (SR) and n is the size of dataset. Also E and O indicate the average estimated and observed SR, respectively. The name of main parameter(s) of each individual SC model and their method of determination are shown in Fig. 2.

Fig. 2. Main parameter(s) of individual Soft Computing (SC) models and their method of determination.

2.3. Fusion-Based Models

Model fusion is a subset of data fusion technique which amalgamates different model simulations with a goal that the result of combining data from different sources become more accurate and reliable than the result of each of the primary sources (Hall and Llinas. 1997). Key role of the fusion method is to specify weights to individual models. There are various methods available to specify such weights. In this paper, in order to acquire a more

precise estimate of land subsidence rate, four fusion methods including Genetic Algorithm (GA), K-Nearest Neighbors (KNN), and Ordered Weighted Average (OWA) method based on ORNESS and ORLIKE methods are employed to derive the weights of individual SC models. For estimation purposes, suppose y_j (j =1, 2, K, k) represents observed data, n signifies each individual model, and the estimated value by i^{th} individual model is shown as $\hat{y}_{ij}$ (i =1, K, n). If weights vector is w= $[w_1, w_2, K, w_n]^T$ the estimated output of fusion model (Y_j) can be expressed as below:

$$Y_j = \sum_{i=1}^{n} w_i \hat{y}_{ij} \tag{8}$$

Subject to: $\quad \sum_{i=1}^{n} w_i = 1 \qquad\qquad w_i \in [0,1] \tag{9}$

Individual model errors are calculated through $e_{ij}= \hat{y}_{ij} - y_j$, while the corresponding error of fusion methods is specified as $E_j=Y_j - y_j$.

In order to improve estimation accuracy, the corresponding errors of fusion methods should be minimized. For this purpose, four different fusion methods including GA, KNN and OWA method (ORNESS and ORLIKE methods) are utilized (Fig. 3). A brief explanation of these approaches is given in the next sections.

Fig. 3. Schematic representation of the fusion-based methodology.

2.3.1) GA Fusion Method

Genetic Algorithm (GA) is one of the optimization methods based on evolutionary process. This method, by iteratively generating a set of possible solutions, tries to achieve a global optimum solution. In the present study, decision variables of the GA optimization model are the weights assigned to each individual SC model. The objective function is to obtain decision variables that minimize Mean Absolute Relative Error (MARE) between the weighted individual models' output (output of fusion method, $w_i\hat{y}_{ij}$) and target values of subsidence rate (y_j):

$$\text{Minimize } MARE = \frac{\sum_{j=1}^{k}\left(\dfrac{\sum_{i=1}^{n} w_i \hat{y}_{ij} - y_j}{y_j}\right)}{k} \times 100 \tag{10}$$

Subject to: $\quad \sum_{i=1}^{n} w_i = 1 \qquad\qquad w_i \in [0,1]$

Where n is the number of individual models and k is the size of dataset.

2.3.2) KNN Fusion Method

K-Nearest Neighbor (KNN) obtains the best estimate of a target variable for a specific point based on a weighted average of the target values from its k nearest samples (Altman 1992). Briefly, Euclidean distance of all available samples to the desired point are calculated according to Eq. 11, and its k nearest samples are selected to estimate

the target value at this point. Neighbors that are closer to the desired point should be weighted more heavily than more distant ones. Hence, reciprocal of squared distance of each data (Eq. 12) is used as weight. Then, using Eq. 13, the target value for the desired point is estimated (Larose 2005):

$$\forall xx \qquad d(xx, x_i) = \sqrt{\sum_{i=1}^{n}(x_i - xx)^2} \qquad\qquad (11)$$

$$w_i = \frac{1}{d(xx, x_i)^2} \qquad\qquad (12)$$

$$Y_{new} = \frac{\sum_{i=1}^{n} w_i\, Y_i}{\sum_{i=1}^{n} w_i} \qquad\qquad (13)$$

Where x_i ($i = 1, K, n$) are the sample data with known target values Y_i and xx is the desired point, for which the target value, Y_{new}, is being sought. In this study, the best results of five SC models are applied as inputs of KNN method in order to fuse them. For further information about this method refer to Altman (1992) and Larose (2005).

2.3.3) ORNESS-OWA

The Ordered Weighted Average (OWA) method is a mapping tool $F: R^n \rightarrow R$, in which n is the number of individual models. This method allocates weight of each model with the constraint:

$$\sum_{i=1}^{n} w_i = 1 \qquad\qquad w_i \in [0,1] \qquad\qquad (14)$$

Here, we first provide some background about the OWA methodology, and then discuss assigning w_i values in the subsequent sections. If the predicted subsidence rate of ith individual model is b_i, then the vector of the results of n individual models will be $B = \{b_1, b_2, \dots, b_n\}$ and the vector of corresponding weights will be $w = \{w_1, w_2, \dots w_n\}$. Since weighing individual models by the OWA method makes different combinations of weights, Yager (1988) defined the parameter *orness* and *Dispersion* to determine the dispersion of weights around the median value of the parameters:

$$orness\,(w) = \sum_{i=1}^{n} \frac{(n-i)}{(n-1)} w_i \qquad\qquad orness\,(w) \in [0,1] \qquad\qquad (15)$$

$$Dispersion(w) = -\sum_{i=1}^{n} w_i \ln(w_i) \qquad\qquad (16)$$

A 0.5 value for *orness* represents the equality of all weights, meaning weights are normally distributed around the median. Therefore, *orness* values between 0.5 and 1 indicate that the weight distribution has a positive skewness, so that larger weight is assigned to a better model.

In this approach, the fusion weights for the SC models are estimated using the Genetic Algorithm (GA) optimization model for different α values between 0 and 1 (Eq. 16). O'Hagan (1988) used the following optimization model to determine the weights so that for the specified values (α) of the *orness* parameter, the maximum value of *Dispersion* (w) is calculated:

$$\text{Maximize } Dispersion(w) = -\sum_{i=1}^{n} w_i \ln(w_i) \tag{17}$$

$$\text{Subject to: } orness(w) = \sum_{i=1}^{n} \frac{(n-i)}{(n-1)} w_i = \alpha \qquad \alpha \in [0,1] \tag{18}$$

$$\sum_{i=1}^{n} w_i = 1 \qquad\qquad w_i \in [0,1] \tag{19}$$

2.3.4) ORLIKE-OWA

Yager and Filev (1994) present a family of OWA weights named S-OWA, which include two main weighing methods (ORLIKE and ANDLIKE[1]). In this paper, ORLIKE-OWA is used. In this method, similar to ORNESS, the best model gets the highest weight and worst model is penalized in the weighting scheme. This method calculates the weights by solving the following equations. F is the ORLIKE method's operator as shown in Eq. 20:

$$w_i = \begin{cases} \dfrac{1}{n}(1-\alpha)+\alpha & i=1 \\[2mm] \dfrac{1}{n}(1-\alpha) & i=2,K,n \end{cases} \qquad \alpha \in [0,1] \tag{20}$$

$$F = \left(\frac{1-\alpha}{n}+\alpha\right)b_1 + \frac{1-\alpha}{n}\sum_{i=2}^{n} b_i \tag{21}$$

$$orness(F) = \frac{1-\alpha}{2}+\alpha+\frac{1-\alpha}{n}\sum_{i=2}^{n}\frac{n-i}{n-1} = \frac{1-\alpha}{n}+\alpha+\frac{1-\alpha}{n}\frac{n-2}{2}$$
$$\Rightarrow orness(F) = \frac{\alpha+1}{2} \tag{22}$$

In this method, for $\alpha=0$, the *orness* parameter is equal to 0.5 and for $\alpha=1$, the value of *orness* parameter is equal to 1, that means the first model will gain more weight. Since the results of individual models are sorted in the descending order of the best result to the worst one, it is necessary that the model weights are also adjusted in the descending order so that the highest weight is given to the best model output. It is noteworthy that for α values smaller than 0.5, the weights of the models are ascending, and for more than 0.5, the weights of the models are descending. To accommodate weights and variables, α values greater than 0.5 have been used to implement the optimization and weighting process. For more information about OWA methods, refer to Yager (1998), O'Hagan (1998) and Yager and Filev (1994). In this study, ORNESS and ORLIKE methods are performed for different α values and their best results are determined (Tables 6-7).

---

[1] In the ANDLIKE method, the worst model gets the highest weight. Authors considered in this study both models to assign weight in a sequence. Both (ORLIKE and ORNESS) assign the highest weight to the best.

3. Study Area

Tehran basin with a total area of 2,250 km^2 is surrounded by the Alborz and the Fashapouye Mountains (Dehghani et al., 2013). The vast majority of this basin, which is under agricultural activities, is subject to subsidence due to excessive abstraction of groundwater. The study area, depicted in Fig. 4, is in the southwest of the Tehran basin in central north part of Iran.

The data collected by Dehghani (2010) is used in this study to develop a model fusion methodology for estimating land subsidence using PS-InSAR technique. Dehghani (2010) extracted, with piezometric measurements, the effective parameters on subsidence rate, including water level decline (observed between 1968-2003), groundwater depth, storage coefficient, transmissivity, alluvial thickness and frequency of fine-grained sediments. Also, the subsidence rates were inferred from the PS-InSAR technique. As previously mentioned, to achieve more accuracy in land subsidence rate estimation, the dataset were clustered using k-means method. Statistics of hydrogeological variables, including maximum, minimum and average, for each cluster are presented in Table 2.

Table 2. Range of hydrogeological variables of each cluster.
Fig. 4. Location of the study area in Tehran province.

----

4. Results

For all Soft Computing (SC) models, available hydrogeology information, which were extracted from piezometric measurements, are utilized as forcing and subsidence rates from PS-InSAR technique is used as output data. Available data is categorized by k-means method into five clusters, and of each cluster's data, 70% are used for training and 30% for validation. To train the five SC models, namely, ANFIS, SVR, MLP, RBF and GRNN, we employed trial-and-error analysis and Genetic Algorithm (GA) optimization. It should be noted that the GA optimization method is utilized for training two SC models (RBF and GRNN), optimizing ORNESS-OWA and also as one of the fusion methods. Specifications relating to the GA optimization method used in aforementioned models are presented in Table 3. Scattered crossover function with fraction value 0.8 is considered for this approach. The TolFun 1e-10 (tolerance value) for StallGenLimit 80 (generation limits) are defined as stopping criteria for models. The results of each individual model's parameters and their method of determination are listed in Table 4.

Table. 3. Specifications of Genetic Algorithm (GA) used in different parts of the proposed methodology.
Table 4. Results of each individual models' parameters and their method of determination.

In order to evaluate the accuracy of these models, six statistical error indices including NS, CC, SI, RMSE, RMRE and Bias (Eqs. 2-7) are calculated for all individual models. As an example, the results of all SC models in estimating land subsidence rate in validation stage are presented in Table 5.

Table 5. Statistical error indices for individual Soft Computing models in estimating land subsidence rate in validation stage

----

Proximity of the NS and CC to 1, and RMSE, RMRE, SI and Bias indices to 0, indicate higher accuracy of the model. Each model result is then ranked based on superior performance and ranked, with 1 representing best model. Minimum summation of ranking in each cluster (Table 5) specifies the more accurate SC model according to all indices. Since the aim of this study is to improve accuracy of subsidence rate estimation, four fusion-based methods including Genetic Algorithm (GA) optimization model, K-Nearest Neighbors (KNN) and two Ordered Weighted Average (OWA) models, namely, ORNESS and ORLIKE methods were used to fuse the outputs of individual SC models and were compared with the best individual model in each cluster (Fig. 6). The performance

of ORNESS (Eqs. 14-19) for α values from 0.5 to 1 and ORLIKE (Eqs. 20-22) for α values from 0.1 to 1 are determined and compared based on trial-and-error analysis (Tables 6-7). As noted earlier, in ORNESS and ORLIKE methods, the best model gets the highest weight. Therefore, according to Tables 6 and 7, the results obtained from these two methods are compared based on two statistical error indices (NS and RMSE) for different α values. Best prediction and associated α are shown in bold in Tables 6 and 7 for the ORNESS and ORLIKE methods. In addition to the trial-and-error analysis, α values were also optimized using GA optimization methods, results of which didn't significantly change the findings of Tables 6 and 7.

Table 6. Comparison of results of ORNESS method for different α values based on NS[*] and RMSE[**] error indices
Table 7. Comparison of results of ORLIKE method for different α values based on NS[*] and RMSE[**] error indices
----

Table 8 presents performance evaluation of four fusion methods of this study in terms of the six statistical error indices mentioned before. Comparing the statistical error indices shown in Tables 5 and 8 shows the superior accuracy of the fusion methods compared to the individual models.

Table 8. Statistical error indices for four fusion methods in estimating land subsidence rate in validation stage.

To make the intercomparison of fusion methods more visually appealing, bar charts of fusion models performance with respect to different statistical error indices are presented in Fig. 5. This figure shows that ORNESS-OWA model has a superior performance and is more accurate as opposed to the other fusion methods in most of clusters.

Fig. 5. Fusion methods' performance with respect to several statistical error indices.

Fig. 6 compares the best fusion method in each cluster with the best individual SC model in the same cluster. The figure confirms that the fusion-based methods are more accurate in estimation of land subsidence rate.

Fig. 6. Comparison the results of best fusion method and best individual SC model of each cluster.

The average RMSE reported by Dehghani et al. (2013) is 4.055 (mm/year), while in this study, we obtained an RMSE value of 3.89 (mm/year) for the best individual SC model (SVR) in most clusters and 2.55 (mm/year) for the best fusion model (ORNESS-OWA) in most clusters. Comparing the present study results with Dehghani et al. (2013) shows that the presented methodology in this study is more accurate. Moreover, fusion-based methods are more accurate than individual soft computing methods.

5. Summary and Conclusion

Land subsidence due to excessive and unsustainable groundwater withdrawal is a paramount hazard to infrastructure safety. Estimating subsidence rate (SR) with sufficient precision is hence of particular interest to sustain human and environmental safety and well-being. In this paper, in order to increase the precision of subsidence rate estimation in the Tehran basin, Iran, a new methodology is developed based on four fusion-based methods, namely, Genetic Algorithm (GA), K-Nearest Neighbors (KNN) and Ordered Weighted Average (OWA) with two weighting methods (ORNESS and ORLIKE) to fuse five individual Soft Computing (SC) models. The approach initiates with obtaining hydrogeological information and subsidence rates estimated based on PS-InSAR technique, and employing a k-means method to categorize different station data into homogeneous groups. The cluster data are in turn used to train five Soft Computing (SC) models, namely, Adaptive Neuro Fuzzy Inference

System (ANFIS), Support Vector Regression (SVR), Multi-Layer Perceptron (MLP) neural network and two optimized models namely Radial Basis Function (RBF) and Generalized Regression Neural Network (GRNN). Fusion methods then create a weighted average of individual SC models to improve land subsidence rate accuracy. To evaluate and compare the results of all models, six statistical error indices, namely, Scatter Index (SI), Root Mean Square Error (RMSE), Root Mean Relative Error (RMRE), Nash-Sutcliffe (NS) efficiency, Correlation Coefficient (CC) and Bias, were utilized. The results show that, fusion methods are more accurate than individual SC models. Also, the result of fusion methods, reveals that ORNESS-OWA method is the superior model in most of clusters. Authors' suggestions for future studies are (i) to consider Subsidence Vulnerability Indices (SVIs) to represent subsidence potential that affect the vulnerable aquifer, and (ii) to employ the proposed methodology to determine these indices more precisely. Also, Fuzzy set theory can be utilized to address uncertainty sources in land subsidence estimation.

References

Ajami NK, Duan Q, Sorooshian S (2007) Bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologycal prediction. Water Resource Research. 43,W01403. https://doi.org/10.1029/2005WR004745.

Alizadeh MR, Nikoo MR (2018) A fusion-based methodology for meteorological drought estimation using remote sensing data. Remote Sensing of Environment. 211, 229-247.

Altman NS (1992) An introduction to kernel and nearest-neighbor nonparametric regression. The American Statistician 46 (3), 175-185.

Ambrožič T, Turk G (2003) Prediction of subsidence due to underground mining by artificial neural network. Computers and Geosciences 29, 627-637.

Amelung F, Galloway DL, Bell JW, Zebker HA, Laczniak RJ (1999) Sensing ups and downs of Las Vegas: InSAR reveals structural control of land subsidence and aquifer-system deformation. Geology 27 (6), 483-486.

Ashouri H, Hsu KL, Sorooshian S, Braithwaite DK, Knapp KR, Cecil LD, Nelson BR, Prat OP (2015) PERSIAN-CDR: Daily precipitation climate data record from multi satellite observations for hydrological and climate studies. Bulletin of the American Meteorological Society 96 (1), 69-83.

Azmi M, Araghinejad S, Kholghi M (2010) Multi model data fusion for hydrological forecasting using K-nearest neighbor method. Iranian Journal of Science and Technology 34 (B1). 81.

Azmi M, Rodiger C, Walker, JP (2016) A data fusion-based drought index. Water Resource Research 52 (3), 2222-2239.

Budhu M, Adiyaman IB (2010) Mechanics of land subsidence due to groundwater pumping. International Journal for Numerical and Analytical Methods in Geomechanics 34 (14), 1459-1478.

Burbey TJ (2000) The influence of faults in basin-fill deposits on land subsidence, Las Vegas, Valley, Nevada, USA. Hydrology Journal 10 (5), 525-538.

Calderhead AI, Therrien R, Rivera A, Martel, R, Garfias J (2011) Simulating pumping-induced regional land subsidence with the use of InSAR and field data in the Toluca Valley, Mexico. Advances in Water Resources 34 (1), 83-97.

Carnec C, Fabriol H (1999) Monitoring and Modeling land subsidence at the Cerro Prieto Geothermal field, Baja California, Mexico, using SAR interferometry. Geophysical Research Letters 26 (9), 1211-1214.

Cigna F, Osmanoglu B, Cabral-Cano E, Dixon TH, Avila-OLIVERA JA, Garduno-Monroy VH, DeMets C, Wdowiski S (2012) Monitoring land subsidence and its induced geological hazard with Synthetic Aperture Radar Interferometry: A case study in Morelia, Mexico. Remote Sensing of Environment 117, 146-161.

Dasarathy BV (1997) Sensor fusion potential Exploitation-Innovative Architectures and Illustrative Applications. Proceeding of the IEEE 85 (1), 24-38.

Dehghani M, Valadan Zoej MJ, Saatchi S, Biggs J, Parsons B, Wright T (2009) Radar interferometry time series analysis of Mashhad subsidence. Journal of Indian Society of Remote Sensing 37 (1), 147-156.

Dehghani M (2010) Estimation of deformation rate and modeling of land subsidence induced by groundwater exploitation using interferometry. Ph.D. thesis. K. N. Toosi University.

Dehghani M, Valadan Zoej MJ, Entezam I (2013) Neural Network Modeling of Tehran Land Subsidence Measured by Persistant Scaterrer Interferometry. Photogrammetrie- Fernerkundung-Geoinformation 2013(1), 5-17.

Deng Z, Ke Y, Gong H, Li X, Li Z (2017) Land subsidence prediction in Beijing based on PS-InSAR technique and improved Grey-Marcov model. GIScience and Remote Sensing 54 (6), 797-818.

Ding XL, Liu GX, Li ZL, Chen YQ (2004) Ground subsidence monitoring in Hong Kong with Satellite SAR Interferometry. Photogrammetric Engineering and Remote Sensing 10, 1151-1156.

Duan Q, Ajami NK, Gao X, Sorooshian S (2007) Multi-model ensemble hydrologic prediction using Bayesian model averaging. Advances in Water Resoureces 30 (5), 1371-1386.

Du Z, Ge L, Ng AHM, Li X, Li L (2018) Monitoring land deformation in Liulin district, China using InSAR approaches. International Journal of Digital Earth 11 (3), 264-283.

Gambolati G, Teatini P, Ferronato M (2006) Anthropogenic land subsidence. Encyclopedia of Hydrological Sciences 13:158.

Gehlot S, Hanssen RF (2008) Monitoring and Interpretation of Urban Land Subsidence Using Radar Interferometric Time Series and Multi-Source GIS Database. In: Nayak S, Zelatanova S. (eds) Remote Sensing and GIS Technologies for Monitoring and Prediction of Disasters. Environmental Science and Engineering (Environmental Science). Springer, Berlin, Heidelberg

Hall DL, Llinas J (1997) An introduction to multisensor data fusion. Proceeding of the IEEE 85 (1), 6-23.

Jung HC, Kim SW, Jung HS, Min KD, Won JS (2007) Satellite observation of coal mining subsidence by Persistant Scatterer analysis. Engineering Geology 92 (1-2), 1-13.

Kim DK, Le e S, Oh HJ (2009) Prediction of ground subsidence in Samcheok City, Korea using artificial neural network and GIS. Environmental Geology 58 (1), 61-70.

Larose DT (2005) "K-Nearest Neighbor Algorithm," Discover Knowledge in Data: An Introduction to Data Mining. 90-106.

Lee S, Park I, Choi Jk (2012) Spatial prediction of ground subsidence susceptibility using an artificial neural network. Environmental Management 49 (2), 347-358.

Lu Y, Ke CQ, Zhou X, Wang M, Lin H, Chen D, Jiang H (2018) Monitoring land deformation in Changzhou City (China) with multi-band InSAR datasets from 2006 to 2012. International Journal of Remote Sensing 39 (4), 1151-1174.

MacQueen J (1967) June. Some methods for classification and analysis of multivariate observations. In proceedings of the fifth Berkeley Symposium on mathematical statistics and probability 1(14), 281-297.

Maghsoudi Y, Meer F, Hecker C, Perissin D, Saepuloh A (2018) Using PS-InSAR to detect surface deformation in geothermal areas of West Java in Indonesia. International Journal of Applied Earth Observation and Geoinformation 64, 386-396.

Motagh M, Walter TR, Sharifi MA, Fielding E, Schenk A, Andeson, J, Zschau, J (2008) Land subsidence in Iran caused by widespread water reservoir overexploitation. Geophysical Research Letters 35. DOI: 10.1029/2008GL033814

Nadiri AA, Taheri Z, Khatibi R, Barzegari G, Dideban K (2018) Introducing a new framework for mapping subsidence vulnerability indices (SVIs): ALPRIFT. Science of the Total Environment. 628-628, 1043-1057.

Nakagawa H, Murakami M, Fujiwara S, Tobita M (2000) Land subsidence of the northern Kanto Plains caused by ground water extraction detected by JERS-1 SAR Interferometry. IEEE 2233-2235.

Ng AHM, Ge L, Li X, Zhang K (2012) Monitoring ground deformation in Beijing, China with Persistent Scatterer SAR interferometry. Journal of Geodesy 86 (6), 375-392.

Ocak I, Seker SE (2013) Calculation of surface settlements caused by EPBM tunneling using artificial neural network, SVM, and Gaussian processes. Environmental Earth Science 70 (3), 1263-1276.

O'Hagan M (1988) Aggregating template or rule antecedents in real-time expert systems with fuzzy set logic. In Signals, Systems and Computers, 1988. Twenty-Second Asilomar Conference on (Vol. 2, pp. 681-689). IEEE.

Osmanoglu B, Dixon TH, Wdowinski S, Cabral-Cano E, Jiang Y (2011) Mexico City subsidence observed with persistent scattere InSAR. International Journal of Applied Earth Observation and Geoinformation 13 (1), 1-12.

Qu F, Zhang Q, Lu Z, Zhao C, Yang C, Zhang J, (2014) Land subsidence and ground fissures in Xi'an, China 2005-2012 revealed by multi-band InSAR time series analysis. Remote Sensing of Environment 155, 366-376.

Rafie M, Samimi Namin F (2015) Prediction of subsidence risk by FMEA using artificial neural network and fuzzy inference system. International Journal of Mining Science and Technology 25 (4), 655-663.

Sadegh M, Mahjouri N, Kerachian R (2010) Optimal inter-basin water allocation using crisp and fuzzy Shapley games.Water Resources Management 24(10), 2291-2310.

Sadegh M, Kerachian R (2011) Water resources allocation using solution concepts of fuzzy cooperative games: fuzzy least core and fuzzy weak least core. Water resources management 25(10), 2543-2573.

See L, Abrahart RJ (2001) Multi-model data fusion for hydrological forecasting. Computers & geosciences 27(8), 987-994.

Shu C, Burn DH (2004) Artificial neural network ensembles and their application in pooled flood frequency analysis. Water Resources Research 40(9). DOI: 10.1029/2003WR002816

Strozzi T, Teatini P, Tosi L, Wegmuller U, Warner C (2013) Land subsidence of natural transitional environments by satellite radar interferometry on artificial reflectors. Journal of Geophysical Research: Earth surface 118, 1177-1191.

Strozzi T, Caduff R, Wegmuller U, Raetzo H, Houser M (2017) Widespread surface subsidence measured with satellite SAR interferometry in the Swiss alpine range associated with the construction of the Gotthard Base Tunnel. Remote Sensing of Environment 190, 1-12.

Sun H, Zhang Q, Zhao C, Yang C, Sun Q, Chen W (2017) Monitoring land subsidence in the southern part of the lower Liaohe Plain, China with a multi-track PS-InSAR technique. Remote Sensing of Environment 188, 73-84.

Teatini P, Tosi L, Strozzi T, Ceccini G, Rosselli R, Libardo S (2012) Resolving land subsidence within the Venice Lagoon by Persistant Scatterer SAR Interferometry. Physics and Chemistry of the Earth, Parts A/B/C 40-41, 72-79.

Wu J, Hu F (2016) Monitoring ground subsidence along the Shanghai Maglev Zone using TerraSAR-X Images. IEEE Geoscience and Remote Sensing Society 14 (1), 117-121.

Yager RR, Filev DP (1994) Parameterized AND-UKE and OR-LIKE OWA operators. International Journal of General System 22(3), 297-316.

Yager RR (1988) On ordered weighted averaging aggregation operators in multi criteria decision making. IEEE Transactions on systems, Man, and Cybernetics 18(1), 183-190.

Yue H, Liu G, Guo H, Li X, Kang Z, Wang R, Zhong X (2011) Coal mining induced land subsidence monitoring using multiband spaceborne differential interferometric synthetic aperture radar data. Journal of Applied Remote Sensing 5 (1), 053518. Doi: https://doi.org/10.1117/1.3571038.

Table 1. Summary of related studies on subsidence analysis using soft computing models based on artificial intelligence neural network.

Number	Study	Soft Computing model	Study area	Reason of subsidence
1	Ambrožič and Turk, 2003	ANN	Velenje mine (Slovenia)	Mining
2	Kim et al., 2009	ANN	Samcheok (Korea)	Mining
3	Lee et al., 2012	ANFIS	Samcheok (Korea)	Mining
4	Dehghani et al., 2013	ANN	Tehran (Iran)	Grounwater withdrawal
5	Ocak and Seker, 2013	ANN, SVM	Istanbul (Turkey)	Tunneling
6	Rafie and Samimi Namin, 2015	ANN	Tehran (Iran)	Tunneling

Table 2. Range of hydrogeological variables of each cluster

Cluster's number		Hydrogeological variables					
		Alluvial frequency* (%)	Alluvial thickness (m)	Water level decline (m)	Groundwater depth (m)	Storage coefficient (%)	Transmissivity (m²/day)
1	Maximum	85.0	83.0	-7.4	50.0	3.8	1004.0
	Minimum	60.0	20.0	-21.4	30.0	1.0	321.0
	Average	**76.9**	**46.3**	**-17.8**	**38.0**	**2.1**	**725.2**
2	Maximum	100.0	176.0	-10.4	97.0	10.2	2012.0
	Minimum	53.0	62.0	-45.1	30.0	0.4	1515.0
	Average	**84.9**	**94.8**	**-20.6**	**40.9**	**2.0**	**175.8**
3	Maximum	100.0	103.0	-7.1	45.0	4.0	1521.0
	Minimum	68.0	38.0	-22.4	30.0	1.0	1004.0
	Average	**85.8**	**71.2**	**-15.2**	**34.4**	**1.8**	**1285.1**
4	Maximum	96.0	107.0	-24.0	113.0	9.7	3000.0
	Minimum	12.0	31.0	-31.9	51.0	1.6	2560.0
	Average	**55.3**	**69.1**	**-28.9**	**86.0**	**4.7**	**2840.3**
5	Maximum	100.0	190.0	-19.6	120.0	11.1	2559.0
	Minimum	10.0	30.0	-51.8	42.0	1.0	2005.0
	Average	**55.8**	**98.2**	**-32.9**	**82.5**	**5.7**	**2267.3**

Table. 3. Specifications relating to the Genetic Algorithm (GA) used in different parts of the proposed methodology

Model	Number of variables	Population size	Number of generations
RBF	2	20	40
GRNN	1	100	20
ORNESS	5	100	20
GA as fusion model	5	50	80

Table 4. Results of each individual models' parameters and their method of determination.

Model	Method of parameters determination (training)	Parameter	Cluster's number				
			1	2	3	4	5
MLP	Trial-and-Error	Number of hidden layer	1	1	1	1	1
		Number of neurons	26	20	18	20	32
		Type of transfer function	Tangent sigmoid	Tangent sigmoid	Tangent sigmoid	Tangent sigmoid	Tangent sigmoid
SVR	Trial-and-Error	Kernel function	Gaussian	Gaussian	Gaussian	Polynomial	Gaussian
ANFIS	Trial-and-Error	Number of membership function	2	2	2	2	2
		Type of membership function	Trapezoid	Trapezoid	Trapezoid	Trapezoid	Trapezoid
		Number of epochs	20	15	15	10	25
RBF	GA optimization	Maximum number of neurons	99	100	100	88	400
		Spread	87.48	40.62	43.19	28.89	20
GRNN	GA optimization	Spread	0.11	1.7	1.1	0.3	3.6

Table 5. Several statistical error indices for individual soft computing models in estimating land subsidence rate in validation stage

Clusters' number	Individual Model		RMRE*	RMSE (mm/year)	Bias (mm/year)	CC#	NS•	SI^Δ	Summation of ranking
1	**MLP**	Value	0.880	15.070	11.130	0.728	-0.208	0.270	
		Rank	1	5	5	4	5	2	22
	SVR	Value	1.160	2.870	0.390	0.982	0.920	0.260	
		Rank	3	1	2	1	1	1	(9)
	ANFIS	Value	1.320	3.350	0.460	0.975	0.891	0.310	
		Rank	4	2	3	2	3	3	17
	RBF	Value	0.959	3.365	0.163	0.975	0.893	0.310	
		Rank	2	3	1	2	2	3	13
	GRNN	Value	1.460	4.790	0.870	0.955	0.777	0.440	
		Rank	5	4	4	3	4	4	24
2	**MLP**	Value	0.260	25.330	-0.270	0.998	0.643	0.070	
		Rank	3	5	3	2	5	2	23
	SVR	Value	0.240	5.060	0.040	0.999	0.986	1.000	
		Rank	1	1	2	1	1	4	(10)
	ANFIS	Value	0.250	5.620	0.000	0.998	0.982	0.060	
		Rank	2	2	1	2	2	1	10
	RBF	Value	0.304	8.767	1.276	0.996	0.957	0.099	
		Rank	5	3	3	3	3	3	22
	GRNN	Value	0.280	11.560	-1.120	0.993	0.926	1.010	
		Rank	4	4	4	4	4	5	25
3	**MLP**	Value	0.270	3.930	0.99	0.997	0.967	0.080	
		Rank	1	2	2	1	2	1	9
	SVR	Value	0.270	3.760	0.640	0.997	0.969	0.080	
		Rank	1	1	1	1	1	1	(5)
	ANFIS	Value	0.290	4.420	0.990	0.996	0.958	0.090	
		Rank	2	3	2	2	3	2	14
	RBF	Value	0.375	7.244	-1.625	0.990	0.887	0.153	
		Rank	4	4	3	3	4	3	21
	GRNN	Value	0.370	8.390	-2.010	0.987	0.848	1.020	
		Rank	3	5	4	4	5	4	22
4	**MLP**	Value	0.500	33.820	0.820	1.000	0.868	0.030	
		Rank	1	5	3	1	4	1	15
	SVR	Value	1.040	8.10	1.190	0.998	0.992	0.930	
		Rank	5	2	4	3	3	5	22
	ANFIS	Value	0.790	10.380	1.190	0.997	0.988	0.110	
		Rank	4	4	4	4	1	3	20
	RBF	Value	0.553	5.433	0.606	0.999	0.997	0.056	
		Rank	2	1	2	2	2	2	(11)
	GRNN	Value	0.570	10.370	0.350	0.997	0.988	0.910	
		Rank	3	3	1	4	1	4	16
5	**MLP**	Value	0.880	61.950	0.510	0.999	0.053	0.060	
		Rank	4	5	3	1	5	1	19
	SVR	Value	1.810	5.470	1.480	0.998	0.993	0.080	
		Rank	5	2	5	2	2	2	22
	ANFIS	Value	0.660	4.440	0.070	0.999	0.995	0.070	
		Rank	1	1	1	1	1	2	(7)
	RBF	Value	0.758	6.313	0.392	0.998	0.990	0.093	
		Rank	3	3	2	2	3	4	17
	GRNN	Value	0.690	7.420	-0.610	0.997	0.986	1.050	
		Rank	2	4	4	3	4	5	22

*RMRE: Root Mean Relative Error, ¤RMSE: Root Mean Square Error, #CC: Correlation Coefficient, •NS: Nash-Sutcliff, ΔSI: Scatter Index

Table 6. Comparison of results of ORNESS method for different α values based on NS[*] and RMSE[**] statistical error indices

Clusters' number	Statistical error index	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$	$\alpha = 1$
1	RMSE	2.813	2.764	2.478	2.049	1.872	**1.714**
	NS	0.923	0.925	0.940	0.959	0.965	**0.971**
2	RMSE	5.061	4.722	3.824	3.368	3.204	**3.021**
	NS	0.985	0.987	0.991	0.993	0.994	**0.994**
3	RMSE	3.309	3.160	2.631	2.251	1.806	**1.736**
	NS	0.976	0.978	0.985	0.989	0.992	**0.993**
4	RMSE	4.597	43.370	3.479	2.219	1.766	**1.639**
	NS	0.997	0.998	0.998	0.999	0.999	**0.999**
5	RMSE	3.872	3.239	2.988	2.343	2.145	**1.831**
	NS	0.996	0.997	0.997	0.998	0.998	**0.999**

[*]NS: Nash-Sutcliff   [**]RMSE: Root Mean Square Error

Table 7. Comparison of results of ORLIKE method for different α values based on NS[*] and RMSE[**] statistical error indices

Clusters' number	Statistical error index	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$	$\alpha = 1$
1	RMSE	4.041	3.784	3.576	3.425	3.338	**3.321**	3.375	3.497	3.679	3.914
	NS	0.841	0.861	0.875	0.886	0.891	**0.892**	0.889	0.881	0.868	0.851
2	RMSE	16.76	14.723	12.828	11.150	9.799	8.926	**8.677**	9.102	10.118	11.569
	NS	0.844	0.879	0.908	0.931	0.946	0.955	**0.958**	0.954	0.943	0.925
3	RMSE	8.268	7.174	6.245	5.567	**5.237**	5.321	5.800	6.588	7.591	8.734
	NS	0.852	0.888	0.915	0.933	**0.941**	0.938	0.927	0.906	0.875	0.835
4	RMSE	21.931	9.026	16.180	13.432	10.854	8.602	6.997	**6.536**	7.434	9.3067
	NS	0.944	0.958	0.969	0.979	0.986	0.991	0.994	**0.995**	0.993	0.990
5	RMSE	16.478	14.380	12.339	10.389	8.591	7.065	6.018	**5.724**	6.282	7.510
	NS	0.933	0.949	0.962	0.973	0.981	0.987	0.991	**0.992**	0.990	0.986

[*]NS: Nash-Sutcliff   [**]RMSE: Root Mean Square Error

Table 8. Statistical error indices for four fusion methods in estimating land subsidence rate in validation stage

Clusters' number	Fusion Methods		RMRE*	RMSE¤ (mm/year)	Bias (mm/year)	CC#	NS•	SIᐃ	Summation of ranking
1	**GA**	Value	0.898	2.870	0.332	0.982	0.920	0.263	
		Rank	2	3	3	2	2	2	14
	KNN	Value	1.372	4.237	1.736	0.969	0.825	0.357	
		Rank	4	4	4	3	3	4	22
	ORNESS	Value	0.658	1.714	0.151	0.993	0.971	0.158	
		Rank	1	1	1	1	1	1	(6)
	ORLIKE	Value	1.117	2.868	0.204	0.982	0.920	0.264	
		Rank	3	2	2	2	2	3	14
2	**GA**	Value	0.265	5.804	-0.359	0.998	0.981	0.066	
		Rank	3	3	3	2	3	3	17
	KNN	Value	0.282	7.542	2.685	0.997	0.968	0.081	
		Rank	4	4	4	3	4	4	23
	ORNESS	Value	0.156	3.022	0.152	1.000	0.995	0.035	
		Rank	1	1	2	1	1	1	(7)
	ORLIKE	Value	0.235	5.365	-0.034	0.998	0.984	0.061	
		Rank	2	2	1	2	2	2	11
3	**GA**	Value	0.267	3.877	0.957	0.997	0.968	0.081	
		Rank	1	2	2	1	2	2	10
	KNN	Value	0.324	5.264	2.421	0.996	0.940	0.101	
		Rank	3	3	3	2	3	3	17
	ORNESS	Value	0.269	3.634	-0.202	0.997	0.971	0.078	
		Rank	2	1	1	1	1	1	(7)
	ORLIKE	Value	0.346	7.722	-4.411	0.991	0.871	0.137	
		Rank	4	4	4	3	4	4	23
4	**GA**	Value	0.502	3.212	0.143	1.000	0.999	0.034	
		Rank	1	1	1	1	1	1	(6)
	KNN	Value	0.625	5.409	2.658	0.999	0.997	0.049	
		Rank	3	3	3	2	3	3	17
	ORNESS	Value	0.611	3.868	0.255	1.000	0.998	0.04	
		Rank	2	2	2	1	2	2	11
	ORLIKE	Value	0.754	9.323	-5.626	0.998	0.990	0.078	
		Rank	4	4	4	3	4	4	23
5	**GA**	Value	0.830	4.13	0.400	0.999	0.996	0.06	
		Rank	3	2	2	2	2	2	13
	KNN	Value	0.970	5.34	2.54	0.999	0.993	0.070	
		Rank	4	3	3	2	3	3	18
	ORNESS	Value	0.430	1.84	0.18	1.000	0.999	0.030	
		Rank	1	1	1	1	1	1	(6)
	ORLIKE	Value	0.810	7.51	-3.62	0.998	0.986	0.100	
		Rank	2	4	4	3	4	4	21

*RMRE: Root Mean Relative Error, ¤RMSE: Root Mean Square Error, #CC: Correlation Coefficient, •NS: Nash-Sutcliff, ᐃSI: Scatter Index
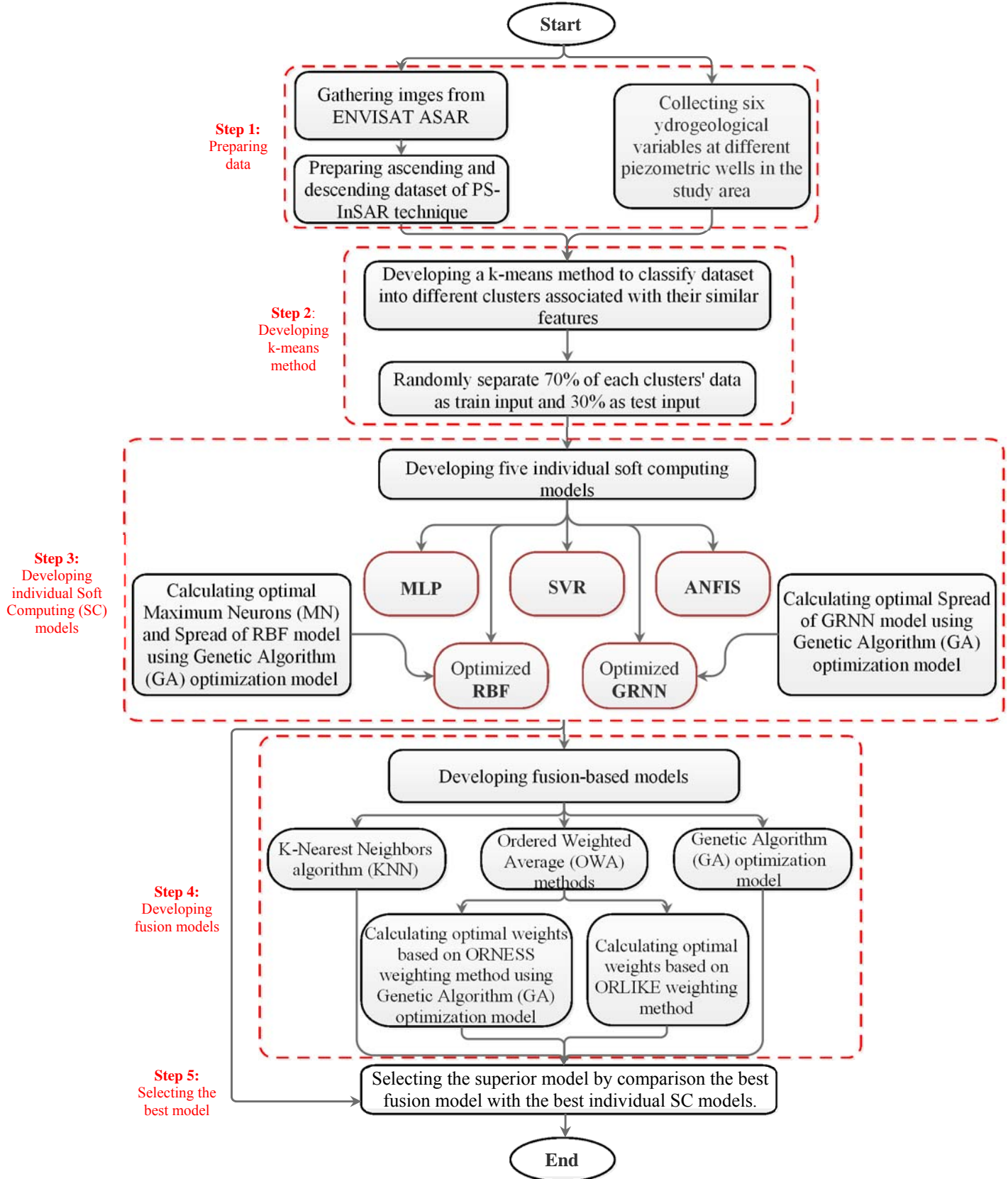
Fig. 1. Flowchart of fusion-based methodology for land subsidence rate estimation
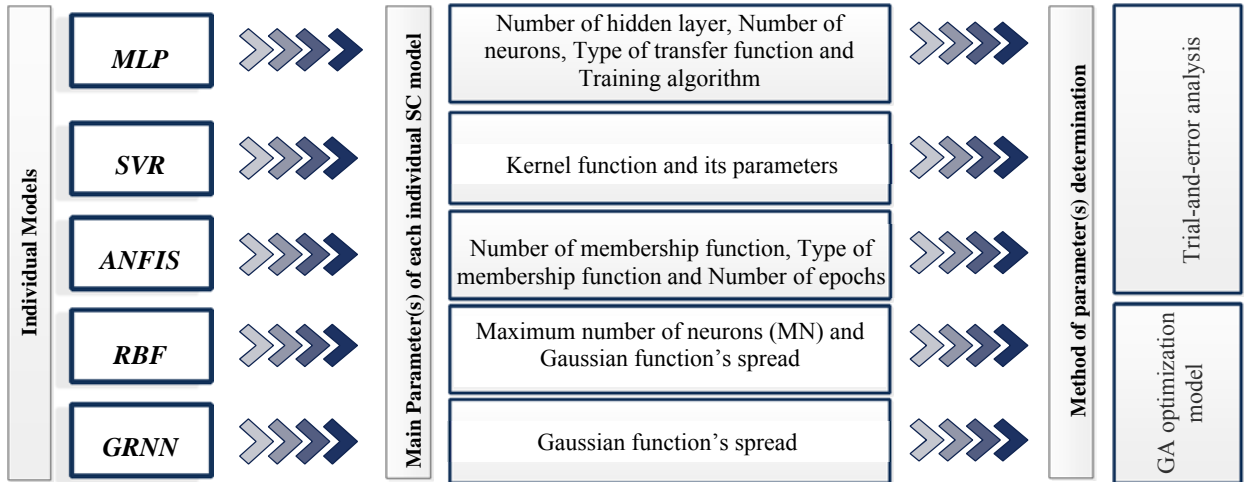
Fig. 2. Main parameter(s) of individual Soft Computing (SC) models and their method of determination
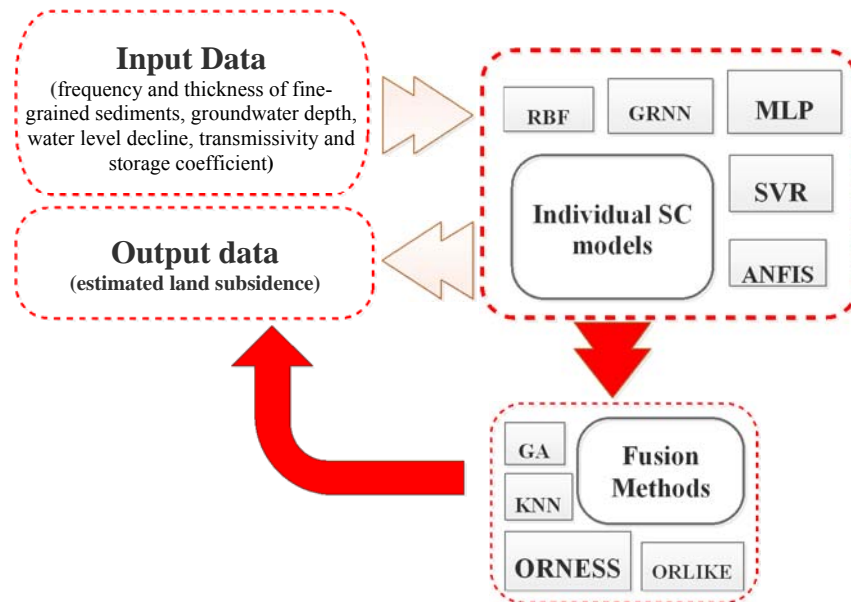

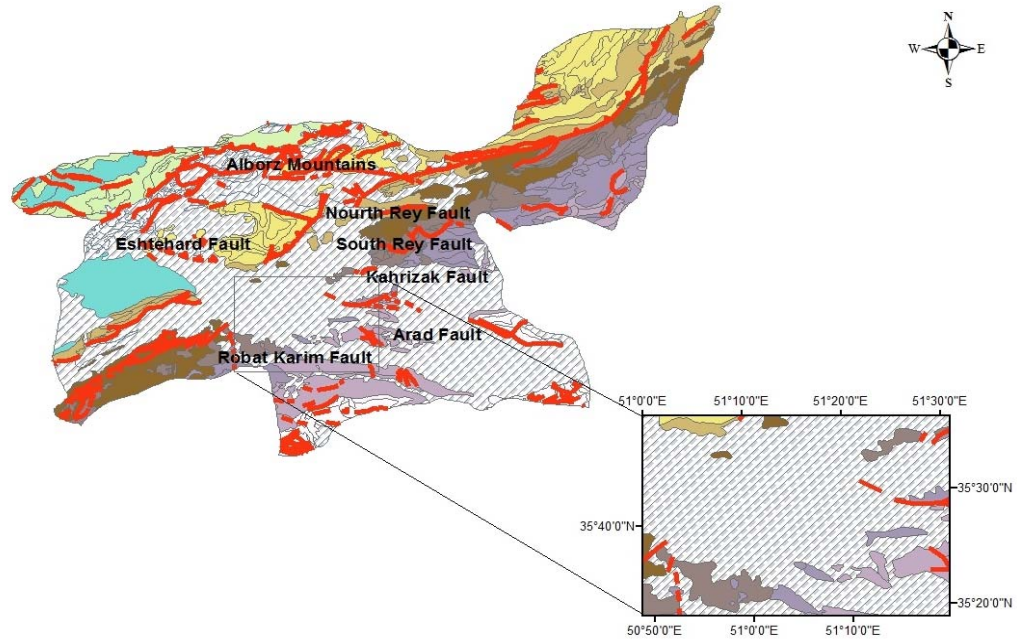**Fig. 3.** Schematic representation of the fusion-based methodology
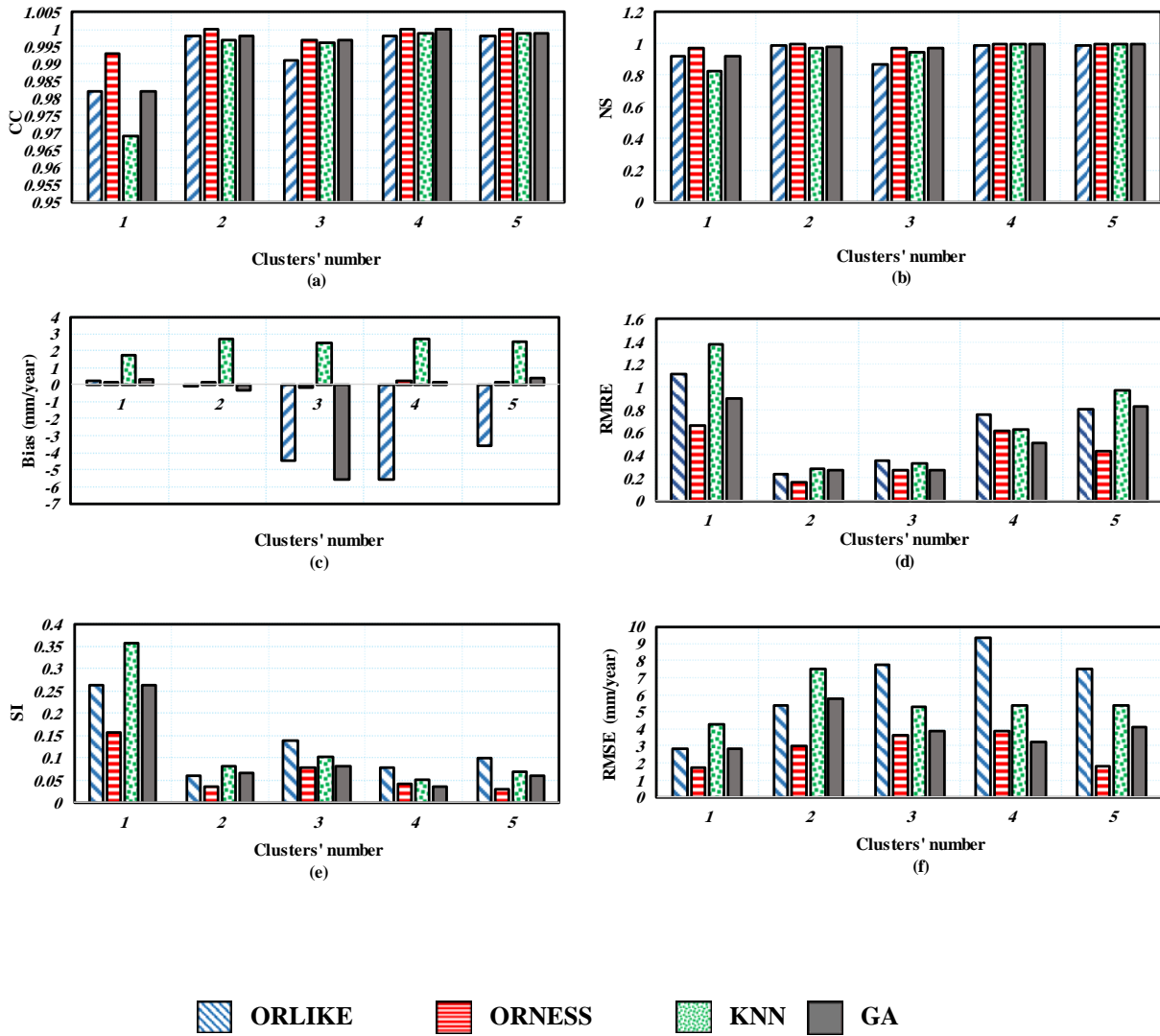
Fig. 4. Location of the study area in Tehran province

Fig. 5. Fusion methods' performance with respect to several statistical error indices.
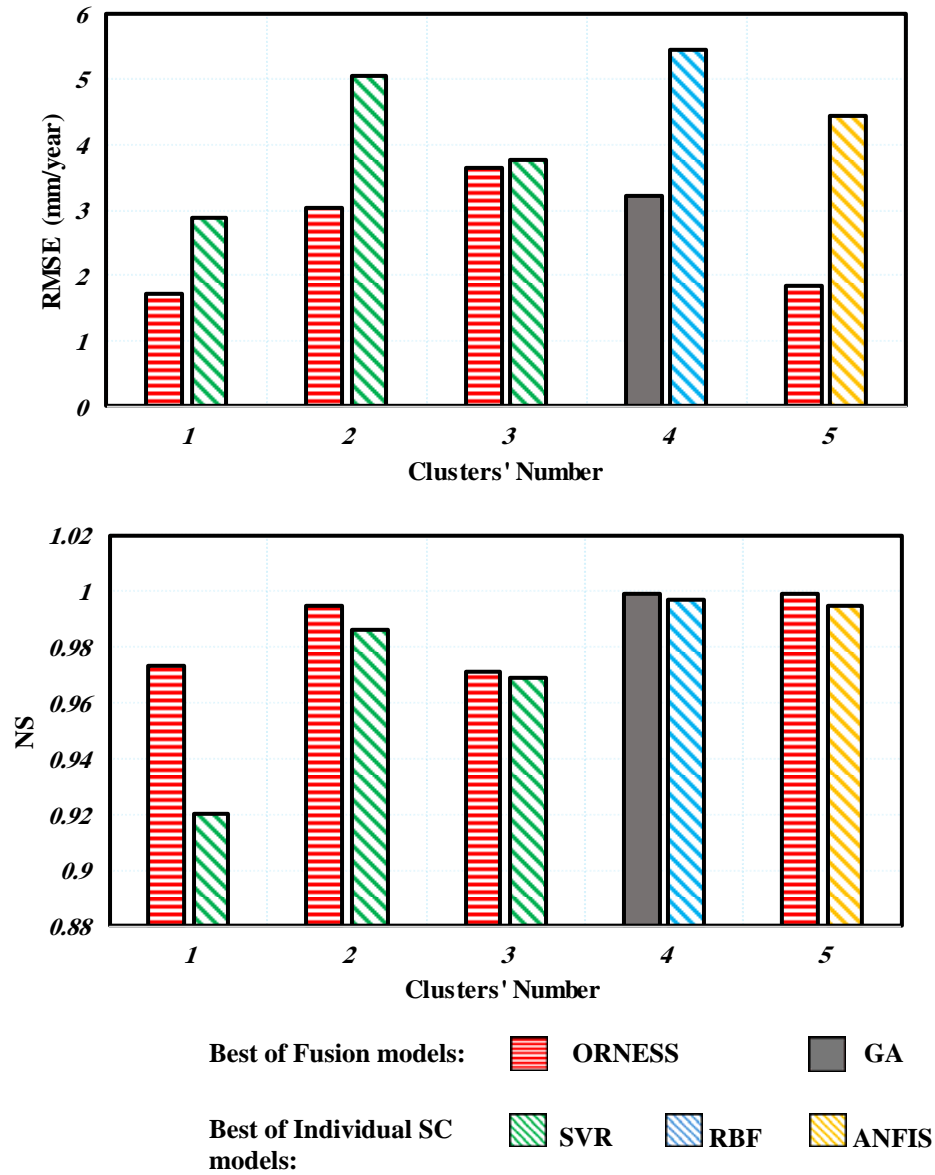
Fig. 6. Comparison the results of best fusion method and best individual SC model of each cluster.