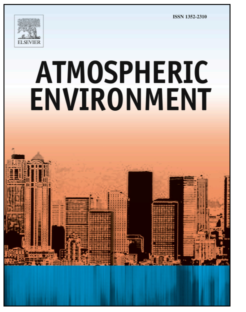
Journal Pre-proof



Hybridized neural fuzzy ensembles for dust source modeling and prediction

Omid Rahmati, Mahdi Panahi, Seid Saeid Ghiasi, Ravinesh C. Deo, John P. Tiefenbacher, Biswajeet Pradhan, Ali Jahani, Hamid Goshtasb, Aiding Kornejady, Himan Shahabi, Ataollah Shirzadi, Hassan Khosravi, Davoud Davoudi Moghaddam, Maryamsadat Mohtashamian, Dieu Tien Bui

PII: S1352-2310(20)30061-3

DOI: https://doi.org/10.1016/j.atmosenv.2020.117320

Reference: AEA 117320

To appear in: *Atmospheric Environment*

Received Date: 27 October 2019

Revised Date: 23 January 2020

Accepted Date: 1 February 2020

Please cite this article as: Rahmati, O., Panahi, M., Ghiasi, S.S., Deo, R.C., Tiefenbacher, J.P., Pradhan, B., Jahani, A., Goshtasb, H., Kornejady, A., Shahabi, H., Shirzadi, A., Khosravi, H., Moghaddam, D.D., Mohtashamian, M., Bui, D.T., Hybridized neural fuzzy ensembles for dust source modeling and prediction, *Atmospheric Environment* (2020), doi: https://doi.org/10.1016/j.atmosenv.2020.117320.

Conceptualization: Omid Rahmati, Mahdi Panahi, Seid Saeid Ghiasi, Ravinesh C. Deo, John P. Tiefenbacher. **Data curation:** Ali Jahani, Hamid Goshtasb, Aiding Kornejady, Himan Shahabi, Ataollah Shirzadi, Hassan Khosravi. **Methodology:** Omid Rahmati, Mahdi Panahi, Biswajeet Pradhan, Davoud Davoudi Moghaddam, Maryamsadat Mohtashamian, Dieu Tien Bui. **Writing—original draft preparation**: Omid Rahmati, Mahdi Panahi, Aiding Kornejady, Himan Shahabi, Ataollah Shirzadi, Hassan Khosravi. **Writing—review and editing:** Omid Rahmati, Ravinesh C. Deo, John P. Tiefenbacher, and Dieu Tien Bui

Hybridized Neural Fuzzy Ensembles for Dust Source Modeling and Prediction

Omid Rahmati[a,b], Mahdi Panahi[c], Seid Saeid Ghiasi[d], Ravinesh C. Deo[e], John P. Tiefenbacher[f], Biswajeet Pradhan[g], Ali Jahani[h], Hamid Goshtasb[h], Aiding Kornejady[i], Himan Shahabi[j], Ataollah Shirzadi[j], Hassan Khosravi[d], Davoud Davoudi Moghaddam[k], Maryamsadat Mohtashamian[l], Dieu Tien Bui [m*]

[a] Geographic Information Science Research Group, Ton Duc Thang University, Ho Chi Minh City, Viet Nam

[b] Faculty of Environment and Labour Safety, Ton Duc Thang University, Ho Chi Minh City, Viet Nam

[c] Department of Geophysics, Young Researchers and Elites Club, North Tehran Branch, Islamic Azad University, Tehran, Iran

[d] Department of Arid and Mountainous Regions Reclamation, Faculty of Natural Resources, University of Tehran, Karaj, Iran

[e] School of Agricultural, Computational and Environmental Sciences, Centre for Sustainable Agricultural Systems & Centre for Applied Climate Sciences, University of Southern Queensland, Springfield, QLD 4300, Australia

[f] Department of Geography, Texas State University, San Marcos, TX 78666, USA

[g] Center for Advanced Modeling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and IT, University of Technology Sydney, 2007 NSW, Australia

[h] Department of Natural Environment and Biodiversity, College of Environment, Karaj, Iran

[i] Department of watershed management, Gorgan University of Agricultural Sciences and Natural Resources, Gorgan, Iran

[j] Department of Geomorphology, Faculty of Natural Resources, University of Kurdistan, Sanandaj, Iran

[k] Department of Watershed Management, Faculty of Agriculture and Natural Resources, Lorestan University, Khorramabad, Iran

[l] Department of Computer Engineering, University of Qazvin, Qazvin, Iran

[m] Institute of Research and Development, Duy Tan University, Da Nang 550000, Viet Nam

* Corresponding author's email address: buitiendieu@duytan.edu.vn

Abstract

Dust storms are believed to play an essential role in many climatological, geochemical, and environmental processes. This atmospheric phenomenon can have a significant negative impact on public health and significantly disturb natural ecosystems. Identifying dust-source areas is thus a fundamental task to control the effects of this hazard. This study is the first attempt to identify dust source areas using hybridized machine-learning algorithms. Each hybridized model, designed as an intelligent system, consists of an adaptive neuro-fuzzy inference system (ANFIS), integrated with a combination of metaheuristic optimization algorithms: the bat algorithm (BA), cultural algorithm (CA), and differential evolution (DE). The data acquired from two key sources – the Moderate Resolution Imaging Spectroradiometer (MODIS) Deep Blue and the Ozone Monitoring Instrument (OMI) – are incorporated into the hybridized model, along with relevant data from field surveys and dust samples. Goodness-of-fit analyses are performed to evaluate the predictive capability of the hybridized models using different statistical criteria, including the true skill statistic (TSS) and the area under the receiver operating characteristic curve (AUC). The results demonstrate that the hybridized ANFIS-DE model (with AUC=84.1%, TSS=0.73) outperforms the other comparative hybridized models tailored for dust-storm prediction. The results provide evidence that the hybridized ANFIS-DE model should be explored as a promising, cost-effective method for efficiently identifying the dust-source areas, with benefits for both public health and natural environments where excessive dust presents significant challenges.

Keywords: Environmental modeling; dust; neural fuzzy; ensemble; Iran

1. Introduction

Dust storms are natural atmospheric events that mainly occur in arid areas, reducing air quality and visibility (Nazari, 2016). Dust is comprised of large-grained particulate matter (PM) that is light enough to be entrained by horizontal atmospheric flows. However, dust storms also carry minute and fine-grained solid matter that is small enough to be more easily elevated aloft and carried by prevailing winds. The occurrence of dust storms has increased in the Middle East in recent years, providing compelling evidence that dust particles are carried long distances (Khaniabadi et al., 2017; Soleimani et al., 2016). From a sustainable development viewpoint, dust storms present challenges to people that run counter to the sustainable development goals (SDGs) outlined the United Nations in its 2030 Agenda on Sustainable Development. For example, out of 17 goals, goal 3 (good health and well-being), goal 6 (clean water and sanitation), goal 7 (affordable and clean energy), goal 9 (industry, innovation and infrastructure), goal 11 (sustainable cities and communities), goal 13 (climate action), and goal 15 (life on land) are directly or indirectly made more challenging to achieve by dust storms. Since dust is generated by wind erosion, it seems rational that land-use practices are a cause (Lu and Shao, 2001; Keesstra et al., 2016). Land degradation and dust generation can directly affect SDGs. Therefore, prevention of land degradation by maintaining or enhancing the natural capital and associated ecosystem services of the land should dominate the land management paradigm (Keesstra et al., 2018).

Dust storms are integral to Earth's natural systems and have impacts that are numerous and wide-ranging. These include effects on-air chemistry, soil characteristics, water quality, nutrient dynamics, and biogeochemical cycling in both oceanic and terrestrial environments (Crooks et al., 2016; Khaniabadi et al., 2017; Middleton and Kang, 2017). Local and regional climates can be affected by dust storms for the scattering and absorption of solar radiation by dust particles, but the impacts can extend great distances from the sources of dust. Dust can modify the microphysical properties of clouds and change precipitation efficiency. In totality, dust storms can affect atmospheric conditions at many different scales (Wang et al., 2018; Yilbas et al., 2015).

The airborne PM is a health-damaging pollutant that adversely affects human cardiovascular systems and causes respiratory problems (Crooks et al., 2016). Inhalation of PM can also exacerbate various diseases and trigger health issues such as asthma in children and the elderly, ultimately increasing morbidity (Kanatani et al., 2010). Pathogenic and non-pathogenic microorganisms (including Coxiella Burnetii, Mycobacterium, Aspergillus, Mycobacterium, Brucella, Cladosporium, Actinomycetes, Clostridium perfingens, and Bacillus), toxins, and influenza viruses can adhere to dust particles and can be transported to great distances (Goudie, 2014; Leski et al., 2011; Soleimani et al., 2016). Moreover, metallic elements are transported as inhalable dust particles, and these could potentially affect the respiratory tracts and can cause neurological and other physiological impacts (Neisi et al., 2016; Yamaguchi et al., 2012). In addition to the health impacts, there are economic impacts from sand and dust storms. Crops and livestock have also been destroyed by dust and sandstorms (Schepanski, 2012). Recently, Gholami et al. (2020) used several data-mining models to map the provenance of storm dust in Khuzestan Province, Iran. Although they provided information about dust movement that can be used to mitigate its off-site effects, the control of wind erosion and dust production in dust-source areas was not explored.

Dust particles emitted from different sources (termed dust-source) are likely to affect a plant's life in different ways (Supe and Gawande, 2015). The largest sources of dust in Earth's atmosphere are from the Sahara and Sahel regions of North Africa (so-called "African dust"), the Gobi, Taklamakan, and Badain Juran deserts of Asia ("Asian dust"), and Australian desert environments ("Australian dust") (Griffin, 2004; Uno et al., 2009). Asian dust particles can also migrate globally, perhaps circumnavigating the Earth in as minimum as 13 days, as recorded in the French Alps (Grousset et al., 2003) and ice and snow cores from Greenland (Bory et al., 2003). Recent changes to regional climates have considerably increased the frequency of dust storm events in the Middle East (Yilbas et al., 2015). Given the hazardous effects of dust storms, new measures are needed to identify and control their genesis regionally proactively. Furthermore, it is also crucial for all sectors to mitigate the catastrophic effects of dust storms.

Although the dust has long been known to be important in weather processes and storms and can influence local weather, the prediction of dust-source areas is challenging, rudimentary and somewhat not effective in current systems. Despite the sophisticated weather models, it remains difficult to forecast the entrainment and transport of dust in the lower atmosphere. One reason for this is a limited understanding of the distribution of the sources and behavioral mechanisms of dust concerning their spatiotemporal volatility in response to various activities and processes (Feuerstein and Schepanski, 2019). For the analysis of dust sources, and the modeling of their impact on Earth's natural system, it is crucial to identify the spatial and temporal diffusion rates of sources (Feuerstein and Schepanski, 2019). In some previous studies, a diverse range of remotely operating methods have been used to identify dust source areas including, but not limited to: (1) remote sensing analysis, (2) horizontal visibility, (3) mineralogy of dust samples, and (4) Lagrangian back-trajectory (Baddock et al., 2009). The drawbacks of each have been discussed in Schepanski et al. (2012). Although these approaches provide useful information regarding the potential sources of dust and the coupling and analysis of geo-environmental and weather conditions, to recognize dust sources over large areas remains relatively difficult.

Considering this need, artificial-intelligence models that apply machine-learning techniques have been developed in the context of geo-environmental research. Adaptive neuro-fuzzy inference system (ANFIS) is a common machine-learning technique in geosciences due to its advantages, such as having the abilities to integrate information from several sources, to handle large amounts of noisy, and to find non-linear relationships between inputs and outputs (Sambariya et al., 2014). However, the main drawback of the ANFIS model is its poor generalization capability for unseen data. Another disadvantage is weak scalability with the number of membership functions and a number of inputs (Jang, 1991; Jang, 1993; Jang et al., 1997). Furthermore, it often requires relatively large data sets for calibration and validation purposes (Liška et al., 2018). In dust-source assessments, however, it is difficult to collect and/or generate adequate amounts of data, particularly over large regions (e.g., deserts), due to the constraints of time, costs, and measurement difficulties. To address this significant gap in dust-source prediction methodologies, this study aims to develop a suite of hybridized artificial-intelligence models using ANFIS where metaheuristic optimization algorithms are used to improve the resulting predictive model. To determine the accuracy of the models, field investigations were conducted, and statistical analyses were performed to identify the dust-source areas in three provinces of eastern Iran. This research promotes the SDGs by developing a modeling approach that can identify dust-source areas. Sustainable land management in dust-source regions can focus on reducing wind erosion (Cerdà et al., 2018a, b).

2. Material and methods

2.1. Study area

The study region, the provinces of Razavi Khorasan, Jonobi Khorasan, and Sisstan-Balochistan in eastern Iran (Fig. 1) covers an area of 444,904 km^2 and forms a homogenous geographical unit with specific characteristics: proximity to the eastern deserts of the Iran plateau, variability and deficiency of precipitation, desertification, high evaporation rates, frequent high wind conditions, and lack of permanent surface water bodies. The climate of this region is hot and arid. The wind is more frequent here than in other parts of the country with approximately 120 high-wind days annually. To develop a predictive model for dust storm sources, model hybridization was achieved by combining ANFIS with three metaheuristic optimization algorithms: the bat algorithm (BA), the cultural algorithm (CA), and the differential evolution (DE) approach. This framework integrated several modeling approaches and achieved a model with superior performance and efficient computing time. The method presented here can be used to distinguish source regions of dust in arid and semi-arid regions.

Fig. 1 here

2.2. Methodology

The methodology (Fig. 2) involved several steps, including conducting a dust-source inventory, identification of the factors that influence dust generation, and modeling.

Fig. 2 here

2.2.1 Dust-source inventory

This study has used two common satellite remote-sensing products to identify dust sources in the study region: the “Moderate Resolution Imaging Spectroradiometer (MODIS)” Deep Blue and the Ozone Monitoring Instrument (OMI). These have been widely applied in previous research as not only they are cost-effective and robust sources of data but also they provide the first direct characterization of the origin of individual sources possible (Baddock et al., 2009; Ginoux et al., 2010; Prospero et al., 2002). Following these studies, we have also used the frequency-of-occurrence (FOO) to localize dust sources. As one of its advantages, the use of this method is not limited to arid regions but can be used beyond. FOO is the number of days that aerosol optical thickness (τ) is greater than $\tau_{\text{threshold}}$, and Ångström exponent (α), and single scattering albedo (ω_0) satisfy criteria of freshly emitted dust particles (i.e., large particles which have not yet been omitted by gravitational settling). Therefore, the satellite-retrieved values of τ, α, and ω_0 for each day should be monitored. Simultaneous consideration of these factors provides comprehensive information on the column-averaged features of the air mass that allows the distinction of dust from aerosols (e.g., anthropogenic pollution aerosols). A detailed description of this method is given by Ginoux et al. (2010) and Prospero et al. (2002), and we provide only a brief overview. We investigated dust storms using the previously described indices during April and July (2014-2018). In the analyses, the high α values were not observed which implies that there are not fine-mode anthropogenic pollutions (smoke) in the study area (i.e., dust can be recognized by a small Ångström exponent), reducing the complexity of dust identification. After July 2018, several field surveys were conducted, and geo-environmental and terrain characteristics were identified and investigated. The most frequent (i.e., having the highest FOO) were significantly associated with the dried bed of the Jazmurian wetland, the Hirmand River, Hamun Lake and some ephemeral wetlands. A relatively similar spatial distribution of dust storm occurred in these areas during April and July in all five years. Beyond these areas, there were also other locations that are sensitive to wind erosion. Both the frequency and intensity of dust storms have increased in 2017 and 2018 compared to previous years. A total of 85 dust-source areas were detected and geolocated with a GPS receiver. These regions are quite active, and they pumped significant amounts of dust particles into the atmosphere during this period. The locations were randomly divided into two groups for training (n=56 or 70%) and for validation (n=29 or 30%) of the models (Figure 1).

2.2.2 Factors that influence dust generation

There is no predetermined set of geo-environmental and topographical factors known to be linked to dust-source areas. According to the field investigations and previous studies, a total of eight factors – wind speeds, geology, maximum air temperatures, land uses, slopes, soils, precipitation amounts, and land cover were considered to be potential predictive factors for modeling locations of dust generation (Fig. 3).

Wind speed. Wind is the primary factor for aeolian erosion (Borrelli et al., 2014). Generally, winds can transport sands and dust at various altitudes; this is dictated by wind speed and uplift. In this study, wind speed data were obtained from weather stations. Several interpolation techniques were used to generate a wind speed map and their accuracies were compared. Subsequently, kriging was selected as the technique to use, as it yielded the lowest root mean square error (RMSE). The wind speed in the study region averages between 10 to 17 m/s at the surface (Fig. 3a). Therefore, wind speed is an essential factor for mapping dust-source potential because it increases the probability of dust entrainment. The wind-speed map demonstrates that speeds are high in the eastern part of the region and are moderate in the western part. Winds tend to be lower in the northern portion of the study area.

Geology. A geological map of the study area was obtained from the Geological Survey of Iran (GSI). The study area's geology is comprised of alluvium, ophiolites, conglomerates, sandstones, acidic and basic igneous, and volcanic rocks (Fig. 3b). Dolomites, limestones, mud volcanics, recent volcanics, and some colored series are also found in the area. Some areas have not been surveyed geologically, however. The Jazmurian basin is the largest basin in the study area. However, rocks from the Cambrian to the Triassic period are found in this region (Stocklin, 1968). Pyroclasts, alluvium, limestone, sandstone, basic and ultra-basic stones, and ophiolites are easily eroded by wind and provide for abundant sources of dust production.

Air temperature. Air temperature plays a vital role in dust production. Higher air temperatures increase rock decomposition to rapidly generate significant quantities of dust particles (Kimura, 2012). Ambient air temperature measurements were obtained from weather stations in the study area. Like the wind speed map, several interpolation techniques to generate an air temperature map were compared and kriging was deemed the most appropriate because it was most accurate (i.e., had the lowest RMSE). The maximum air temperatures in the study region ranged from 49°C to 42.1°C (Fig. 3c).

Land use. Land use is also an indicator used to map dust potential (Kimura, 2012). Land use reflects the intensity of human activities and the potential for environmental degradation and disturbance of the surface. This study used a land-use map derived from a Landsat OLI image (2016) employing an object-based image-classification technique (Fig. 3d). The image was radiometrically corrected with a pre-processing technique by converting the detected radiometrics into reflectance values.

Slope. Slope is crucial to dust production and it is incorporated into dust emission and transport models. The dust sources are widely distributed in areas of lower slopes and can be identified and assessed with remotely sensed time-series data (Hahnenberger and Nicoll, 2012). The slope value is represented as a percentage; the highest slope value was 185.3 (Fig. 3e).

Soil. The characteristics of soils, directly and indirectly, affect the dust-storm initiation (Hahnenberger and Nicoll, 2012; Kimura, 2012). Eroded particles vary in size (i.e., from dust particle to boulder). Heavier materials cannot be moved very far by wind, but dust particles can be transported long distances and are deposited when they collide with obstacles in their paths or when wind speed diminishes and loses its capacity to move them. Soil type is also a primary influence on plant growth. Fig. 3f shows the distribution of the dominant soil types in the region.

Rainfall. Rainfall influences soil moisture, significantly impacting the strength of some soils against erosion and, consequently particulate production. If rainfall and or soil moisture decreases, dust increases. It, therefore, has a significant influence on the spatial distribution of dust potential. In this study, precipitation data were obtained from the Iranian Department of Water Resources Management (IDWRM). The rainfall map was also produced by using kriging. The study area is dominated by landscapes of sparse shrubs and annual plants that reflect the arid climate with low precipitation; the northern and southern parts receive more precipitation than the central region of the study area (Fig. 3g).

Land cover. Land cover is relevant to discerning dust-source potential. Land cover influences the susceptibility of the soil to erosion. Compared to forests, land degradation is more severe on land with scant vegetation. A land cover map of the study area was obtained from the Forest, Range and Watershed Organization (FRWO) of Iran (Fig. 3h).

Fig. 3 here

2.2.3 Basics and application of models

While artificial neural networks (ANNs) can model any function regardless of its complexity and are characterized by excellent learning and generalization capacities, they have drawbacks: difficulty selecting the optimal number of layers and neurons in the model and interpreting functionality (Al-Mahasneh et al., 2016; Jahani, 2019; Liška et al., 2018). Fuzzy inference systems (FIS) are based on fuzzy logic, enabling classifications that allow partial membership in multiple classes. The advantages and disadvantages of fuzzy inference systems and ANNs have explained in Al-Mahasneh et al. (2016). ANFIS, also known as the universal estimator, is the combination of artificial neural networks (ANNs) and the Takagi–Sugeno fuzzy-inference system which was first developed in the early 1990s (Jang, 1991; Jang, 1993; Jang et al., 1997). Combining neural networks and fuzzy logic is one way to overcome the disadvantages of both techniques (Singh et al., 2012). Several studies report that ANFIS integrates the advantages of neural networks in dealing with the implicit knowledge that can be acquired by learning and fuzzy systems and in dealing with the explicit knowledge that can be explained and understood (Heddam et al., 2012; et al., 2012; Wei, 2016). ANFIS also analyzes, learns, and adapts quickly (Chen et al., 2013). Furthermore, fuzzy if-then rules serve as inference-engines that enable ANFIS to approximate non-linear patterns by perpetually updating the knowledge of that system based on newly defined rules, and concurrently updating the linear and nonlinear parameters based on gradient descent and recursive least-square algorithms (Polat and Güneş, 2006). Therefore, ANFIS is an artificial intelligence approach used for solving complicated problems in several scientific fields (Premkumar and Manikandan, 2016; Wang and Elhag, 2008). A feed-forward network that includes different layers with various functions is the fundamental configuration of the ANFIS. One of the important steps in an ANFIS model is the fuzzification of input data which is implemented using fuzzy membership functions. There are different membership functions including the triangular, trapezoidal, Gaussian, and bell (Chen et al., 2017). The Gaussian function was used in this study. It is popular for specifying fuzzy sets, and its curve is smooth and never equals zero (Tzeng, 2010). It is defined as (Eq. 1):

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}({x-\mu}/{\sigma})^2} \quad (1)$$

where μ and σ are parameter sets that change the shape of the membership function. Optimizing parameters of the Gaussian function increases the accuracy of the ANFIS model. One crucial feature of ANN and ANFIS models is the activation function that decides whether a neuron should be activated or not (Yilmaz and Kaynar, 2011; Sun et al., 2015). When the activation function is not used, weights and bias simply do a linear transformation; consequently, such ANNs and ANFIS models are substantially linear regressions (Rani et al., 2019). In fact, the activation function applies a non-linear transformation to input to enable the learning and prediction of complex tasks (Toghyani et al., 2016). The most common activation functions are identity, binary step, sigmoid, tanh, ramp, ReLU, leaky ReLU, and softmax (Mishkin et al., 2017). In this study, the sigmoid activation function was used. It has been used in several subfields (Topçu and Sarıdemir, 2008; Sarıdemir, 2009; Hajduk, 2017). Sigmoid also has advantages: it is characterized as a smooth function (i.e., it ranges from zero to one and has an S shape) and it is continuously differentiable (Alçın et al., 2016). When multiple neurons include a sigmoid function, the output will be nonlinear (Da and Xiurun, 2005). To achieve the best fit between estimated and measured values, the RMSE was used as the cost function (Eq. 2):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{est}-X_{obs})^2}{n}} \quad (2)$$

where X_{est} and X_{obs} are defined as the estimated and observed (actual) dust, respectively, and n is the number of dust observations. However, it is crucial to tune the learning parameters because it requires considerable time and requires a significant amount of input data. Also, the accuracy of ANFIS significantly depends on the adequacy of training data (Chen et al., 2017). Therefore, many optimization algorithms have been developed to automatically optimize these learning-parameters with inadequate data (Polat and Güneş, 2006; Tien Bui et al., 2018a). Among metaheuristic optimization algorithms, bat, cultural, and differential evolution algorithms were adopted and fused into the ANFIS model as ANFIS-BA, ANFIS-CA, and ANFIS-DE.

In summary, the bat algorithm, as the name implies, imitates the echolocation behavior of bats (i.e. sound pulses) and was first developed by Yang (2010). It entails three main components: frequency, loudness, and pulse emission rate (See Yang (2010) for details). Flying with random velocity in a random space (i.e., randomly moving through the parameters' space) and analyzing the three variables mentioned above, bats distinguish an object from obstacles and obstacles from open space (i.e., the presence and absence of localities) (Ali, 2014; Sambariya and Prasad, 2014). With this information, the bat optimizer can tune the learning parameters of ANFIS.

The CA algorithm, on the other hand, develops with evolutionary computations. It is a mathematical representation of how societies evolve or adapt to their environments. First expounded by Reynolds (Reynolds et al., 2008), the algorithm is underpinned by a two-level computational process, termed a dual-inheritance (Soza et al., (Soza et al., 2002). The first level focuses on a population that shares a set of behavioral traits that is continuously handed down through the generations and is possibly spread to others in society by social motivators. The second level focuses self-experiences and self-forecasts that can be generalized and merged into a global belief. Thus, the circulation between the population, a belief and subcomponents therein provide an outline for a cultural-evolution framework that can be mathematically represented by various models, such as genetic algorithms (Schepanski et al., 2012).

DE, as a stochastic global-optimization method, can optimize the properties of a non-linear and non-differentiable problem in a continuous space (Wang et al., 2014; Wu et al., 2016). The DE targets an objective function (e.g., a cost function) and minimizes it under certain constraining functions with an easy-to-operate implementation process (Soleimani et al., 2016). Using a vector (or parameter) population and reliable handling of stochastic perturbations in the population enables DE to fairly quickly provide practical results. The DE has been used to contribute to evolutionary optimization and is one of the fastest and most practical optimization methods, particularly in comparison to other prominent minimization methods such as annealing and genetic algorithms. The DE algorithm contains four basic steps: initialization, mutation, crossover (also known as recombination), and selection[8]. The last three steps are reiterated until a termination criterion is satisfied. Several termination criteria can be considered in the modeling process. In this study, its iterative process was terminated when the root-mean-square error (RMSE) was minimal.

A schematic of each metaheuristic optimization algorithm (i.e., BA, CA, and DE) is given in Fig. 4. There are some notable differences between their architectures and their data analytical processes. Detailed descriptions of these popular algorithms can be found in the literature (Premkumar and Manikandan, 2016; Tien Bui et al., 2018a; Tien Bui et al., 2018b). In this study, all individual and hybrid models (i.e., ANFIS, ANFIS-BA, ANFIS-CA, and ANFIS-DE) were executed with MATLAB software.

Fig. 4 here

To apply the ANFIS model (i.e., standalone), the presence/absence of a dust source was used as the dependent variable, whereas the dust-influencing factors were independent variables. ANFIS was calibrated using the training data set (70% of dust-source locations in inventory) as explained in the previous sections. Therefore, a dust-source probability map was generated with the standalone ANFIS model. Subsequently, the training data set was also used for training the hybridized models (i.e., ANFIS-BA, ANFIS-CA, and ANFIS-DE). Three dust-source probability maps were produced for the study area using the hybridized models. It should be noted that the validation data set (30% of dust-source locations) was not used in the training stage.

2.2.4 Accuracy assessment

In this study, some standard evaluation metrics including root mean square error (RMSE), the area under the receiver operating characteristic curve (AUC), and true skill statistic (TSS) were used. The RMSE was described in the previous sections. These metrics are used to assess goodness-of-fit and predictive performance.

The AUC metric is calculated with the receiver operating characteristic (ROC) curve and measures how well a model generally performs (Pham et al., 2019; Tien Bui et al., 2018b). ROC curve plots the "1–specificity" (also known as false positive rate, FPR) on the horizontal axis against the sensitivity (also termed as true positive rates, TPR) on the vertical axis (Tien Bui et al., 2018a). The sensitivity reflects the probability of correctly predicting the positives (i.e., dust source sites) as observed, whereas the "1–specificity" shows the probability of incorrectly predicting a non-event location (i.e., non-dust source) as an event (i.e., dust source). "1–specificity" and sensitivity can be calculated using the components of the confusion matrix, including true positives (TP), false positives (FP), false negatives

(FN), and true negatives (TN). TP and TN are dust source and non-dust source locations correctly classified, respectively. FP and FN are the numbers of misclassified positives (i.e., dust source) and negatives (i.e., non-dust source). TPR and FPR can be calculated:

$$TPR = \frac{TP}{TP+FN} \quad (3)$$

$$FPR = \frac{FP}{FP+TN} \quad (4)$$

In an analytic expression of the ROC curve, it is denoted as *f*. The AUC is formulated as (Eq. 5):

$$AUC = \int_0^1 f(FPR)dFPR = 1 - \int_0^1 f(TPR)dTPR \quad (5)$$

TSS is the other metric used to check the model performance based on the TPR and FPR statistical measures. It can be expressed as follows (Eq. 6):

$$\mathrm{TSS} = \frac{TP}{TP+FN} - \frac{FP}{FP+TN} = TPR - FPR \quad (6)$$

To suggest or reject a model for other susceptible areas, its reliability and performance should be evaluated using training and validation datasets (Bahraminejad et al., 2018; He et al., 2019). Therefore, all three evaluation metrics were calculated during training (i.e., using the training data set) and validation (i.e., using the validation data set).

3. Results and discussion

3.1 Preparation of maps of potential dust-sources

The spatial distribution of potential dust sources derived from standalone ANFIS models and from the equivalent hybridized models in which optimization algorithms used are illustrated in Fig. 5. Upon initial inspection, the spatial distribution of potential dust sources seems to be clearly differentiated across the study area. Notably, all four predictive models (i.e., the standalone ANFIS, as well as the ANFIS-BA, ANFIS-CA, and ANFIS-DE hybridized models) reveal a relatively similar spatial pattern of dust potential across the study region. The northern, eastern, and southwestern parts of the region are highly active dust-production sources, while the central parts show significantly less dust-potential and are a rather low-dust zone. Visual comparison of the enlarged insets clipped from the dust-potential maps reveals the less precise classifications dust-potential produced by the standalone ANFIS model (Fig. 1a inset), particularly in areas without original source-data. The hybridized models produce a clearer and more precise differentiation of localities with and without dust storms. This is discernible in the proportional distribution of the dust-potential classes each hybridized model generates (Table 1).

The ANFIS model has classified nearly 69% of the region as highly dust storm active, which contradicts the empirical evidence of dust storms in this particular region. These predictions are of little practical value to guide pragmatic action to mitigate the impacts of dust storms. Conversely, the study areas classified as ‘high’ and ‘very high’ by the hybridized models are smaller proportions of the whole; they present more realistic representations of dust storm occurrence. This attests to the enhancement that optimized ANFIS models provide for more differentiation between classes and, therefore, perhaps, a more accurate solution. All three hybridized models indicated that the dried bed of the Hirmand River, the Jazmurian wetland, Hamun Lake and some ephemeral lakes are the most active dust sources. These findings are consistent with Rashki et al. (2015) who also investigated transport pathways and mechanisms of dust using meteorological station datasets describing the southern part of the study area (the Sistan region). Their results indicated that the dried bed of the Jazmurian wetland and Hamun Lake were the most active dust sources in the Sistan region. Specific management should be implemented in dust-source areas. For example, human activities related to water resources and land cover should be rationally controlled.

Table 1 here

Fig. 5 here

3.2 Validation and comparison of the novel hybridized- and standalone-ANFIS models

To determine the accuracy of the hybridized ANFIS models, a goodness-of-fit test assessed the models in terms of mean-square error (MSE), root-mean-square error (RMSE), mean, and standard deviation (StD) metrics from the observed and predicted data (supplementary Figures S1-S4). All performance metrics computed for the ANFIS model with the training dataset were set to 0 (Fig. S1b, c). However, the values generated by the validation dataset for MSE, RMSE, mean, and StD were about 0.072, 0.269, 0.018, and 0.271, respectively (Fig. S1e, f). This indicated that the model had over-fitted the training dataset during its learning stage. These results demonstrated the tendency of the standalone ANFIS model to over-fit, as was shown in the study of Tien Bui et al. (2018b). By contrast, in the ANFIS-BA model, the values of 0.023, 0.153, 0.06, and 0.154 were obtained for the MSE, RMSE, mean, and StD, respectively, in the training phase (Fig. S2b, c). The values for the same variables generated with validation data were 0.020, 0.143, 0.013, and 0.144, respectively (Fig. S2e, f).

Similarly, the values of MSE, RMSE, mean, and StD obtained with the training dataset as input into the hybridized ANFIS-CA model were about 0.021, 0.146, 0.016, and 0.146, respectively (Fig. S3b, c). And for the validation dataset, they were about 0.022, 0.149, 0.010, and 0.150, respectively (Figure S3e, f). For the hybridized ANFIS-DE model, the training-data generated values for MSE, RMSE, mean, and StD were 0.016, 0.126, 0.005, and 0.127, respectively (Fig. S4b, c) and the validation-data values were 0.020, 0.142, -0.016, and 0.143, respectively (Fig. S4e, f). In this regard, similar to Tien Bui et al. (2018a, 2018b), we have demonstrated that a hybridized-ANFIS model can be considered to be a more robust predictive model for dust-storm prediction, as it attained a much greater accuracy than with the standalone-ANFIS model.

Therefore, it is evident that as MSE and RMSE values diminish, goodness-of-fit increases, as does the overall performance for each optimized hybridized-ANFIS model. The ANFIS-DE model performed the best, followed by the ANFIS-BA, ANFIS-CA, and ANFIS models. There are several possible reasons for these results. DE has several advantages over the other algorithms. There are no restrictions on the regularization methods and error function (i.e., non-differentiable transfer functions may be used). Easy tuning of the algorithm parameters (mainly population size). Not only can convergence to a global minimum be expected, but the linear time and space complexity of the algorithm can also be established. Ilonen et al. (2003) confirmed that the structure of the DE algorithm influences its capabilities. Our study indicated that DE's characteristics (compact structure, reliable search capability, high convergence characteristics, and few control parameters) have made it a powerful population-based stochastic optimizer. Some researchers believe that the main reason for its strength is its design principles (simplicity, efficiency, and real coding) (Noman and Iba, 2008; Price, 2013; Das et al., 2016). As discussed by Khazraee et al. (2011), the use of the differential evolution (DE) algorithm is likely to generate a more robust and efficient optimization tool for any predictive model, given its ability to perform a direct search of data features without requiring any derivative estimation or assumptions. This explains the enhanced performance capability of the ANFIS-DE hybridized model.

To evaluate the validity of the models developed for dust-storm prediction, the resulting susceptibility maps were also evaluated spatially for their validity. We tested the accuracy of the prediction of dust storms that have occurred and those that are expected to occur using the training and validation datasets. The results showed that the AUC in the training step (i.e., a measure of the goodness-of-fit) were about 88.1%, 84.9%, 83.0%, and 85.4% for the ANFIS, ANFIS-BA, ANFIS-CA and ANFIS-DE models, respectively. These values in the validation step (i.e., predictive performance) were about 63.7%, 83.4%, 80.3%, and 84.1%, respectively (Table 2).

Another robust statistical metric applied to validate the dust-susceptibility maps is the true skill statistic (TSS). Accordingly, the training TSS value for the ANFIS, ANFIS-BA, ANFIS-CA and ANFIS-DE models was found to be about 0.78, 0.74, 0.73, and 0.75, respectively. Slightly lower values of about 0.64, 0.72, 0.70 and 0.73 were produced with the validation dataset. Although the AUC and TSS metrics produced from the training data and the ANFIS model had the highest performance, ANFIS-BA's metrics using the validation dataset indicated the highest power of prediction. Therefore, the best-hybridized models in order of performance are ANFIS-DE, ANFIS-CA, and ANFIS-BA.

A direct comparison of our results to the findings of other studies is difficult and must be done with caution because the prediction performance of these hybridized models has not been compared in other studies. As described by Das et al. (2016), the differential evolution algorithm's automatic adaptation property, used as a unique feature extraction tool, can significantly enhance the search process of the algorithm for solving multi-objective, dynamic, constrained, and large-scale optimization problems. Besides, both Wang et al. (2008) and Wu et al. (2016) have also explained that the differential evolution algorithm, when used as a population-based stochastic search technique, exhibited remarkable performance in terms of final accuracy, robustness, computational speed. On the other hand, this algorithm requires only three control parameters (i.e., crossover rate, scale factor, and population size) which can be applied to solve a different real-world problem from a diverse array of science and technology areas in practical ways.

According to the literature, the standalone ANFIS model can also have some drawbacks (Jang, 1991; Jang, 1993). The main one is the poor generalization capability for unseen data. Another disadvantage is its weaker scalability when using several membership functions and inputs required to train and execute the model. Also, this model often requires numerous recalibrations. In dust-source assessments, however, it is difficult to collect sufficient data over large regions (e.g., deserts) for various reasons such as time and cost constraints, and measurement difficulties. In this study, metaheuristic optimization algorithms, notably the DE algorithm, were intended to address these significant gaps and enabled not only efficient fitting of data to the model but also enhanced the generalizability of the final model. Consequently, metaheuristic optimization algorithms are likely to dramatically improve the predictive performance of the ANFIS model when applied for spatial prediction of dust storms.

Table 2 here

3.3 Comparison of the models' predictions

Scatter plots for the standalone ANFIS model predictions were compared to those from each hybridized ANFIS model predictions (Fig. 6). The distributions are very near and evenly distributed on both sides of the 1:1 line implying strong agreement between the two data series (i.e., the predictions of ANFIS and each hybridized model are shown accordingly). Clear patterns are not discernable in these plots, indicating that there is almost no agreement between the predictions of the ANFIS and the hybridized ANFIS model. However, two distinct point-patterns are visually discernable on the plots and they are grouped as two clusters of points using cluster analysis. Most of the high values predicted by the ANFIS model (roughly higher than 0.5 on the *x*-axis) lie below the 1:1 line, which means that they are under-predicted by the hybrid model. In contrast, most of the low values produced by the ANFIS model (values lower than 0.5 on the x-axis) are over-predicted by the hybrid models.

The ANFIS model tends to generate results that are composed of a greater number of extreme outliers, while the hybridized models seem to produce predictions with outliers that have moderate values. Although this does not prove that hybridized ANFIS models perform significantly better than a standalone ANFIS model, there is a significant difference between the prediction patterns of the standalone ANFIS and hybridized ANFIS models. Since the ANFIS model by itself has not been applied to topics in this field of study, a direct comparison to results from previous studies is not possible. However, to explore these results further, we consider that several other studies in environmental and hydrological fields have demonstrated that the hybridized ANFIS models can improve prediction of extreme observed values compared to a standalone ANFIS model. For example, the study of Yaseen et al. (2017, 2018) found that the standalone ANFIS model integrated with the firefly optimization algorithm (ANFIS-FFA) was able to capture heavy to extreme rainfall events more accurately than did a standard, non-optimized ANFIS model. In a study on streamflow forecasting, the authors demonstrated that although both standalone- and hybridized-ANFIS models were able to forecast peak streamflow data points quite successfully, the hybrid ANFIS model could forecast low flows more accurately.

Fig. 6 here

4. Conclusion

An ANFIS model was developed and hybridized with model-optimization algorithms to perform comparative analysis for the spatial identification of dust source. The state-of-the-art models developed and tested were standalone ANFIS and three equivalent hybridized models – ANFIS-BA, ANFIS-CA, and ANFIS-DE. The resulting dust-source maps were validated using actual field data and statistical metrics comparing predicted and observed dust-source datasets divided into training and validation subsets. Several model parameters – historical dust-storm data, high-speed wind event data, soil types, air temperatures, geomorphic units, slope, land use, and rainfall – were used as predictors that enabled mapping of potential dust-source areas. We can draw several conclusions from this study.

- Based upon the models developed, there is a significant potential for increasing amounts of dust in the study region because of the interactions of the factors that initiate and promote dust production;
- ANFIS hybridized models can be used to map dust-source areas at a regional scale, creating new pathways to assess dust-storm potential and to examine the effects of these storms on human health

and the environment. The four ANFIS models achieved strong predictive capacities as indicated by the AUC and TSS statistical tests: standalone ANFIS (AUC=63.7%, TSS=0.64), and the hybridized ANFIS-BA (AUC=83.4%, TSS=0.72), ANFIS-CA (AUC=80.3%, TSS=0.7), and ANFIS-DE (AUC=84.1%, TSS=0.73). These accuracy assessments demonstrate that hybridization enhances standalone algorithms, at least with models depicting dust-generation.

- This approach should be of interest to local environmental and health agencies and to governments to identify and mitigate sources of dust. They should consider methods to transfer this approach to other regions that are experiencing a similar problem. This new dust-storm potential modeling approach can be replicated to identify current and future dust sources in other regions.
- The Gaussian membership function was used in this study. Future studies should examine the influence of it and other membership functions (triangular, trapezoidal, Gaussian, and bell functions) on the performance of hybridized models.

Acknowledgments

The authors would like to thank the Iranian Department of Geological Surveys (IDGS) and the Forest, Range and Watershed Organization (FRWO) for supplying required data, reports, useful maps, and their nationwide geodatabase. The authors also would like to thank Prof. Andrew S. Goudie (University of Oxford) for improving the manuscript and providing constructive comments. This research was partially supported by the Geographic Information Science Research Group, Ton Duc Thang University, Ho Chi Minh City, Viet Nam. We greatly appreciate the assistance of the Editor, Prof. Chak K. Chan, and anonymous reviewers for their constructive comments that helped us to improve the paper.

References

Alçın, M., Pehlivan, İ., Koyuncu, İ., 2016. Hardware design and implementation of a novel ANN-based chaotic generator in FPGA. Optik 127(13), 5500-5505.

Ali, E.S., 2014. Optimization of power system stabilizers using BAT search algorithm. Int. J. Elec. Power 61, 683-690.

Al-Mahasneh, M., Aljarrah, M., Rababah, T., Alu'datt, M., 2016. Application of hybrid neural fuzzy system (ANFIS) in food processing and technology. Food Eng. Rev. 8(3), 351-366.

Baddock, M.C., Bullard, J.E., Bryant, R.G., 2009. Dust source identification using MODIS: a comparison of techniques applied to the Lake Eyre Basin, Australia. Remote Sens. Environ. 113(7), 1511-1528.

Bahraminejad, M., Rayegani, B., Jahani, A., Nezami, B., 2018. Proposing an early-warning system for optimal management of protected areas (Case study: Darmiyan protected area, Eastern Iran). J. Nat. Conserv. 46, 79-88.

Borrelli, P., Ballabio, C., Panagos, P. and Montanarella, L., 2014. Wind erosion susceptibility of European soils. Geoderma 232, 471-478.

Bory, A.J.M., Biscaye, P.E., Grousset, F.E., 2003. Two distinct seasonal Asian source regions for mineral dust deposited in Greenland (NorthGRIP). Geophys. Res. Lett. 30(4). https://doi.org/10.1029/2002GL016446

Bui, D.T., Khosravi, K., Li, S., Shahabi, H., Panahi, M., Singh, V., Chapi, K., Shirzadi, A., Panahi, S., Chen, W., Bin Ahmad, B., 2018. New hybrids of anfis with several optimization algorithms for flood susceptibility modeling. Water 10(9), p.1210.

Bui, D.T., Panahi, M., Shahabi, H., Singh, V.P., Shirzadi, A., Chapi, K., Khosravi, K., Chen, W., Panahi, S., Li, S., Ahmad, B.B., 2018. Novel hybrid evolutionary algorithms for spatial prediction of floods. Sci. Rep-UK 8(1), p.15364.

Cerdà, A., Rodrigo-Comino, J., Giménez-Morera, A., Keesstra, S.D. 2018a. Hydrological and erosional impact and farmer's perception on catch crops and weeds in citrus organic farming in Canyoles river watershed, Eastern Spain. Agr. Ecosyst. Environ. 258, 49-58.

Cerdà, A., Rodrigo-Comino, J., Novara, A., Brevik, E.C., Vaezi, A.R., Pulido, M., et al. 2018b. Long-term impact of rainfed agricultural land abandonment on soil erosion in the Western Mediterranean basin. Prog. Phys. Geog. Earth Environ. 42(2), 202-219.

Chen, B., Matthews, P.C., Tavner, P.J., 2013. Wind turbine pitch faults prognosis using a-priori knowledge-based ANFIS. Expert Syst. Appl. 40(17), 6863-6876.

Chen, W., Panahi, M., Pourghasemi, H.R., 2017. Performance evaluation of GIS-based new ensemble data mining techniques of adaptive neuro-fuzzy inference system (ANFIS) with genetic algorithm (GA), differential evolution (DE), and particle swarm optimization (PSO) for landslide spatial modelling. Catena 157, 310-324.

Crooks, J.L., Cascio, W.E., Percy, M.S., Reyes, J., Neas, L.M., Hilborn, E.D., 2016. The association between dust storms and daily non-accidental mortality in the United States, 1993–2005. Environ. Health Persp. 124(11), 1735-1743.

Da, Y., Xiurun, G., 2005. An improved PSO-based ANN with simulated annealing technique. Neurocomputing 63, 527-533.

Das, S., Mullick, S.S., Suganthan, P.N., 2016. Recent advances in differential evolution–an updated survey. Swarm Evol. Comput. 27, 1-30.

Feuerstein, S., Schepanski, K., 2019. Identification of Dust Sources in a Saharan Dust Hot-Spot and Their Implementation in a Dust-Emission Model. Remote Sens. 11(1), p.4.

Gholami, H., Mohamadifar, A., Collins, A.L., 2020. Spatial mapping of the provenance of storm dust: Application of data mining and ensemble modelling. Atmos. Res. 233, p.104716.

Ginoux, P., Garbuzov, D. and Hsu, N.C., 2010. Identification of anthropogenic and natural dust sources using Moderate Resolution Imaging Spectroradiometer (MODIS) Deep Blue level 2 data. J. Geophys. Res.: Atmospheres 115(D5).

Goudie, A.S., 2014. Desert dust and human health disorders. Environ. Int. 63, 101-113.

Griffin, D.W., Kellogg, C.A., 2004. Dust storms and their impact on ocean and human health: dust in Earth's atmosphere. EcoHealth 1(3), 284-295.

Griffin, N.L., Lewis, F.D., 1989, October. A rule-based inference engine which is optimal and VLSI implementable. In [Proceedings 1989] IEEE International Workshop on Tools for Artificial Intelligence (pp. 246-251). IEEE.

Grousset, F.E., Ginoux, P., Bory, A., Biscaye, P.E., 2003. Case study of a Chinese dust plume reaching the French Alps. Geophys. Res. Lett. 30(6). https://doi.org/10.1029/2002GL016833.

Hahnenberger, M., Nicoll, K., 2012. Meteorological characteristics of dust storm events in the eastern Great Basin of Utah, USA. Atmos. Environ. 60, 601-612.

Hajduk, Z., 2017. High accuracy FPGA activation function implementation for neural networks. Neurocomputing 247, 59-61.

He, Q., Shahabi, H., Shirzadi, A., Li, S., Chen, W., Wang, N., Chai, H., Bian, H., Ma, J., Chen, Y., Wang, X., 2019. Landslide spatial modelling using novel bivariate statistical based Naïve Bayes, RBF Classifier, and RBF Network machine learning algorithms. Sci. Total Environ. 663, 1-15.

Heddam, S., Bermad, A., Dechemi, N., 2012. ANFIS-based modelling for coagulant dosage in drinking water treatment plant: a case study. Environ. Monit. Assess. 184(4), 1953-1971.

Ilonen, J., Kamarainen, J.K., Lampinen, J., 2003. Differential evolution training algorithm for feed-forward neural networks. Neural Process. Lett. 17(1), 93-105.

Jahani, A., 2019. Forest landscape aesthetic quality model (FLAQM): A comparative study on landscape modelling using regression analysis and artificial neural networks. J. For. Sci. 65(2), 61-69.

Jang, J.S., 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE T. Syst. Man Cyb. 23(3), 665-685.

Jang, J.S.R., 1991, July. Fuzzy modeling using generalized neural networks and kalman filter algorithm. In AAAI (Vol. 91, 762-767).

Jang, J.S.R., Sun, C.T., Mizutani, E., 1997. Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence [Book Review]. IEEE T. Automat. Contr. 42(10), 1482-1484.

Kanatani, K.T., Ito, I., Al-Delaimy, W.K., Adachi, Y., Mathews, W.C., Ramsdell, J.W., 2010. Desert dust exposure is associated with increased risk of asthma hospitalization in children. Am. J. Resp. Crit. Care. 182(12), 1475-1481.

Keesstra, S. D., Bouma, J., Wallinga, J., Tittonell, P., Smith, P., et al. 2016. The significance of soils and soil science towards realization of the United Nations Sustainable Development Goals. Soil 2, 111-128

Keesstra, S., Mol, G., de Leeuw, J., Okx, J., de Cleen, M., Visser, S. 2018. Soil-related sustainable development goals: Four concepts to make land degradation neutrality and restoration work. Land 7(4), 133.

Khaniabadi, Y.O., Daryanoosh, S.M., Amrane, A., Polosa, R., Hopke, P.K., Goudarzi, G., Mohammadi, M.J., Sicard, P. and Armin, H., 2017. Impact of Middle Eastern Dust storms on human health. Atmos. Pollut. Res. 8(4), 606-613.

Khazraee, S.M., Jahanmiri, A.H., Ghorayshi, S.A., 2011. Model reduction and optimization of reactive batch distillation based on the adaptive neuro-fuzzy inference system and differential evolution. Neural Comput. Appl. 20(2), 239-248.

Kimura, R., 2012. Factors contributing to dust storms in source regions producing the yellow-sand phenomena observed in Japan from 1993 to 2002. J. Arid Environ. 80, 40-44.

Leski, T.A., Malanoski, A.P., Gregory, M.J., Lin, B., Stenger, D.A., 2011. Application of a broad-range resequencing array for detection of pathogens in desert dust samples from Kuwait and Iraq. Appl. Environ. Microbiol. 77(13), 4285-4292.

Liška, A., Kruszewski, G., Baroni, M., 2018. Memorize or generalize? searching for a compositional RNN in a haystack. arXiv preprint arXiv:1802.06467.

Lu, H. and Shao, Y., 2001. Toward quantitative prediction of dust storms: an integrated wind erosion modelling system and its applications. Environ. Modell. Softw. 16(3), 233-249.

Middleton, N., Kang, U., 2017. Sand and dust storms: impact mitigation. Sustainability 9(6), p.1053.

Mishkin, D., Sergievskiy, N., Matas, J., 2017. Systematic evaluation of convolution neural network advances on the imagenet. Comput. Vis. Image Und. 161, 11-19.

Nazari, S., Kermani, M., Fazlzadeh, M., Matboo, S.A., Yari, A.R., 2016. The origins and sources of dust particles, their effects on environment and health, and control strategies: a review. J. Air Pollut. Health 1(2), 137-152.

Neisi, A., Goudarzi, G., Akbar Babaei, A., Vosoughi, M., Hashemzadeh, H., Naimabadi, A., Mohammadi, M.J., Hashemzadeh, B., 2016. Study of heavy metal levels in indoor dust and their health risk assessment in children of Ahvaz city, Iran. Toxin Rev. 35(1-2), 16-23.

Noman, N., Iba, H., 2008. Differential evolution for economic load dispatch problems. Electr. Pow. Syst. Res. 78(8), 1322-1331.

Pham, B.T., Prakash, I., Singh, S.K., Shirzadi, A., Shahabi, H., Bui, D.T., 2019. Landslide susceptibility modeling using Reduced Error Pruning Trees and different ensemble techniques: Hybrid machine learning approaches. Catena 175, 203-218.

Polat, K., Güneş, S., 2006. A hybrid medical decision making system based on principles component analysis, k-NN based weighted pre-processing and adaptive neuro-fuzzy inference system. Digit. Signal Process. 16(6), 913-921.

Premkumar, K., Manikandan, B.V., 2016. Bat algorithm optimized fuzzy PD based speed controller for brushless direct current motor. Engineering Science and Technology, an International Journal, 19(2), 818-840.

Price, K.V., 2013. Differential evolution. In Handbook of Optimization (pp. 187-214). Springer, Berlin, Heidelberg.

Prospero, J.M., Ginoux, P., Torres, O., Nicholson, S.E., Gill, T.E., 2002. Environmental characterization of global sources of atmospheric soil dust identified with the Nimbus 7 Total Ozone Mapping Spectrometer (TOMS) absorbing aerosol product. Rev. Geophys. 40(1), 2-1.

Rani, M.L.P., Rao, G.S., Rao, B.P., 2019. ANN Application for Medical Image Denoising. In Soft Computing for Problem Solving (pp. 675-684). Springer, Singapore.

Rashki, A., Kaskaoutis, D.G., Francois, P., Kosmopoulos, P.G., Legrand, M., 2015. Dust-storm dynamics over Sistan region, Iran: Seasonality, transport characteristics and affected areas. Aeolian Res. 16, 35-48.

Reynolds, R.G., 1999, January. Cultural algorithms: Theory and applications. In New ideas in optimization (pp. 367-378). McGraw-Hill Ltd., UK.

Reynolds, R.G., Ali, M., Jayyousi, T., 2008. Mining the social fabric of archaic urban centers with cultural algorithms. Computer 41(1), 64-72.

Sambariya, D.K., Prasad, R., 2014. Robust tuning of power system stabilizer for small signal stability enhancement using metaheuristic bat algorithm. Int. J. Elec. Power 61, 229-238.

Sarıdemir, M., 2009. Predicting the compressive strength of mortars containing metakaolin by artificial neural networks and fuzzy logic. Adv. Eng. Softw. 40(9), 920-927.

Schepanski, K., Tegen, I., Macke, A., 2012. Comparison of satellite based observations of Saharan dust source areas. Remote Sens. Environ. 123, 90-97.

Singh, R., Kainthola, A., Singh, T.N., 2012. Estimation of elastic constant of rocks using an ANFIS approach. Appl. Soft Comput. 12(1), 40-45.

Soleimani, Z., Goudarzi, G., Sorooshian, A., Marzouni, M.B. and Maleki, H., 2016. Impact of Middle Eastern dust storms on indoor and outdoor composition of bioaerosol. Atmos. Environ. 138, 135-143.

Soza, C., Becerra, R.L., Riff, M.C., Coello, C.A.C., 2011. Solving timetabling problems using a cultural algorithm. Appl. Soft Comput. 11(1), 337-344.

Stefanski, R., Sivakumar, M.V.K., 2009. Impacts of sand and dust storms on agriculture and potential agricultural applications of a SDSWS. In IOP Conference Series: Earth and Environmental Science (Vol. 7, No. 1, p. 012016). IOP Publishing.

Stocklin, J., 1968. Structural history and tectonics of Iran: a review. AAPG bulletin, 52(7), 1229-1258.

Sun, W., Hu, P., Lei, F., Zhu, N., Jiang, Z., 2015. Case study of performance evaluation of ground source heat pump system based on ANN and ANFIS models. Appl. Therm. Eng. 87, 586-594.

Toghyani, S., Ahmadi, M.H., Kasaeian, A., Mohammadi, A.H., 2016. Artificial neural network, ANN-PSO and ANN-ICA for modelling the Stirling engine. Int. J. Ambient Energy 37(5), 456-468.

Topçu, İ.B., Sarıdemir, M., 2008. Prediction of mechanical properties of recycled aggregate concretes containing silica fume using artificial neural networks and fuzzy logic. Comp. Mater. Sci. 42(1), 74-82.

Tzeng, S.T., 2010. Design of fuzzy wavelet neural networks using the GA approach for function approximation and system identification. Fuzzy Sets Syst. 161(19), 2585-2596.

Uno, I., Eguchi, K., Yumimoto, K., Takemura, T., Shimizu, A., Uematsu, M., Liu, Z., Wang, Z., Hara, Y. and Sugimoto, N., 2009. Asian dust transported one full circuit around the globe. Nat. Geosci. 2(8), p.557.

Wang, X., Liu, J., Che, H., Ji, F., Liu, J., 2018. Spatial and temporal evolution of natural and anthropogenic dust events over northern China. Sci. Rep-UK 8(1), p.2141.

Wang, Y., Li, H.X., Huang, T., Li, L., 2014. Differential evolution based on covariance matrix learning and bimodal distribution parameter setting. Appl. Soft Comput. 18, 232-247.

Wang, Y.M., Elhag, T.M., 2008. An adaptive neuro-fuzzy inference system for bridge risk assessment. Expert Syst. Appl. 34(4), 3099-3106.

Wei, L.Y., 2016. A hybrid ANFIS model based on empirical mode decomposition for stock time series forecasting. Appl. Soft Comput. 42, 368-376.

Yilmaz, I., Kaynar, O., 2011. Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. Expert Syst. Appl. 38(5), 5958-5966.

Table 1 The area of dust-source potential classes assigned by the four models (in percent)

Model Type	Model Name	Very low	Low	Medium	High	Very high
Standalone	ANFIS	22.30	1.56	7.49	60.25	8.38
Hybridized Models	ANFIS-BA	5.60	17.95	29.97	33.60	12.85
	ANFIS-CA	3.28	18.47	31.10	24.41	22.71
	ANFIS-DE	2.95	15.82	29.40	36.78	15.1

Table 2. The goodness-of-fit and predictive performance of hybrid and individual models based on AUC and TSS metrics.

Model Type	Model Name	AUC (%)		TSS	
		Training	**Validation**	**Training**	**Validation**
Standalone	ANFIS	88.1	63.7	0.78	0.64
Hybridized Models	ANFIS-BA	84.9	83.4	0.74	0.72
	ANFIS-CA	83.0	80.3	0.73	0.7
	ANFIS-DE	85.4	84.1	0.75	0.73

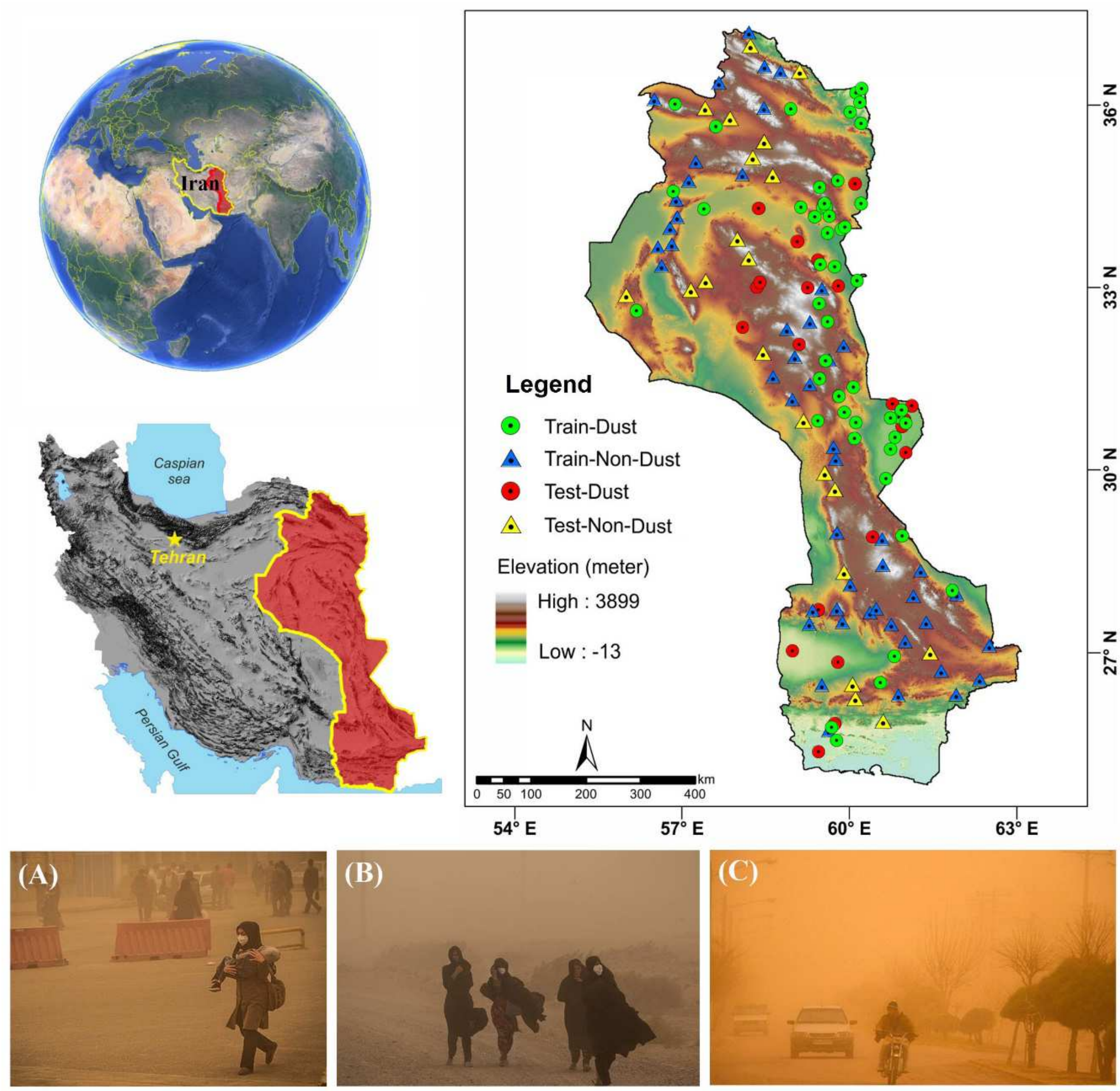


Fig. 1 A map of the study area and field photographs of some major dust storms that have occurred in (A) Zabol, (B) Zahedan, and (C) Iranshahr, Iran (field photographs were taken by the third author (S.S.G.)).

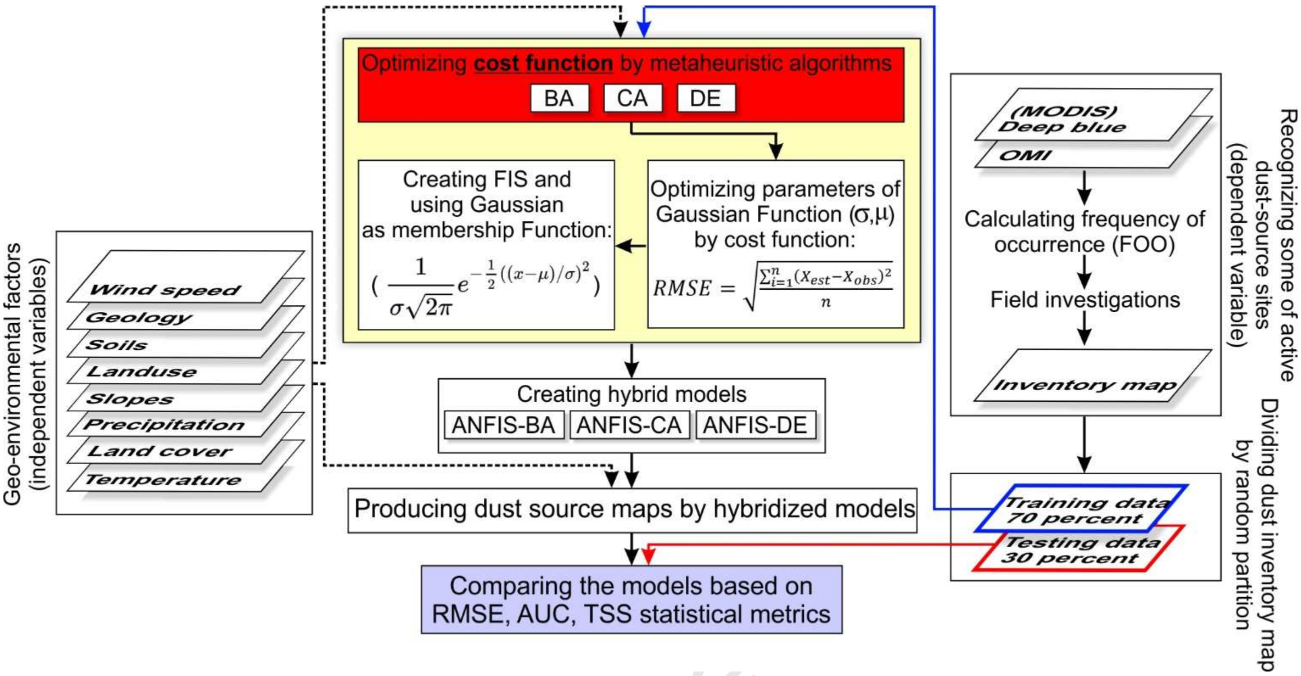


Fig. 2 Methodological flowchart of the present study.

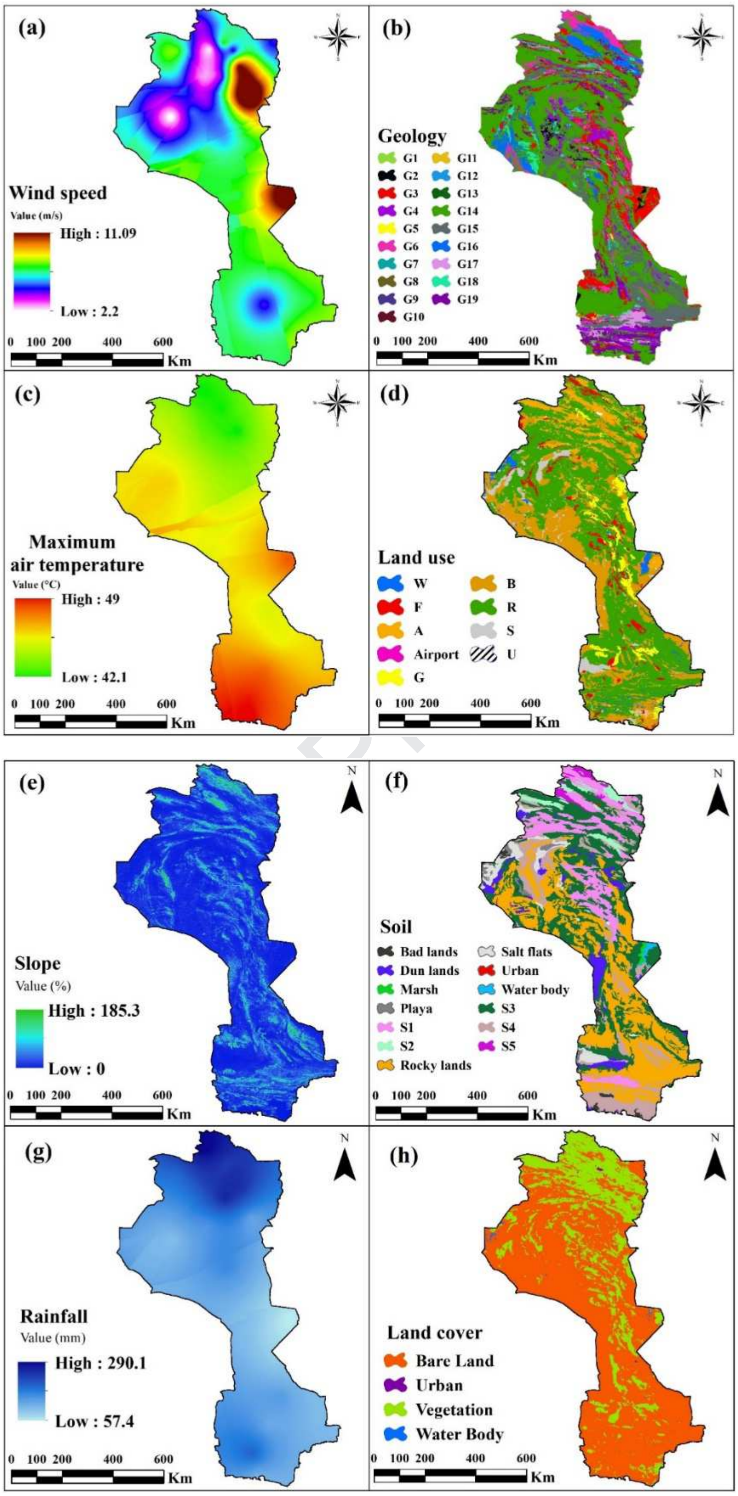


Fig. 3 Dust influencing factors: a) wind speed, b) geology, c) maximum air temperature, d) land use, e) slope, f) soil, g) rainfall, h) land cover.

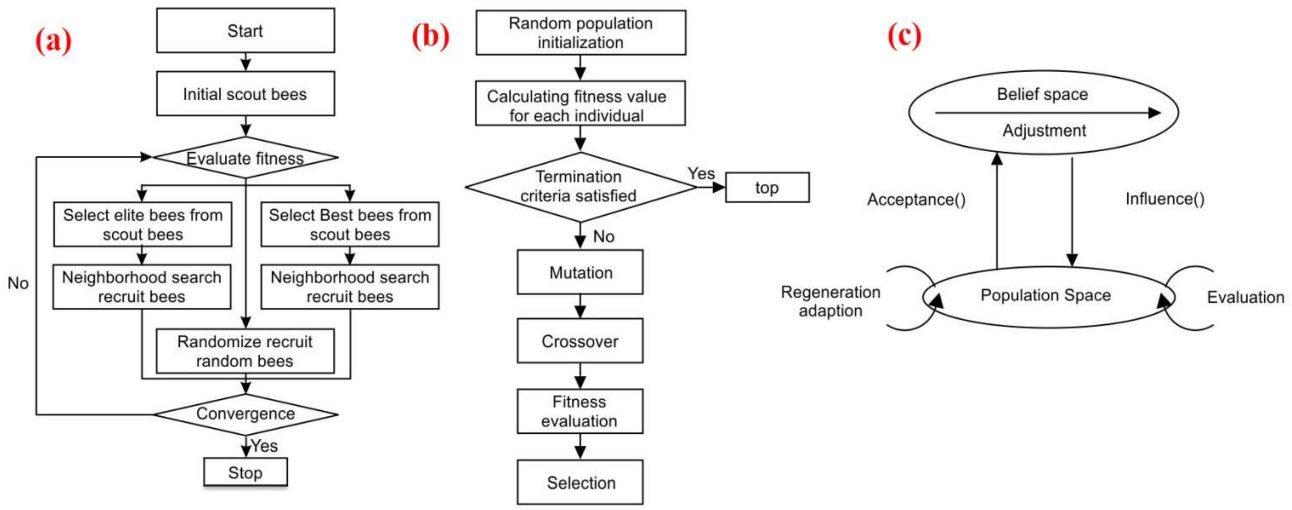


Fig. 4 Metaheuristic optimization algorithms: a) BA, b) DE, and c) CA

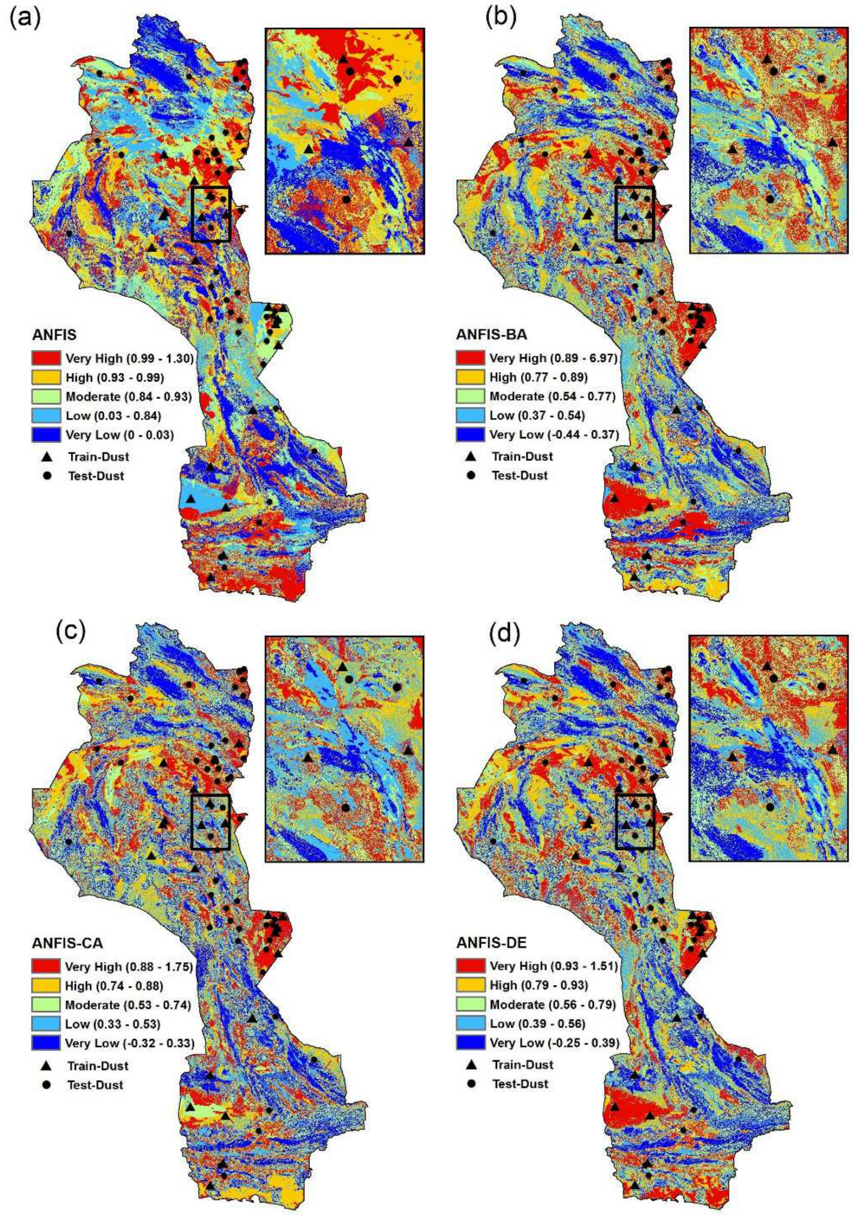


Fig. 5 Dust-source potential mapping prepared by the standalone and hybridized ANFIS models: a) ANFIS, b) ANFIS-BA, c) ANFIS-CA, and d) ANFIS-DE

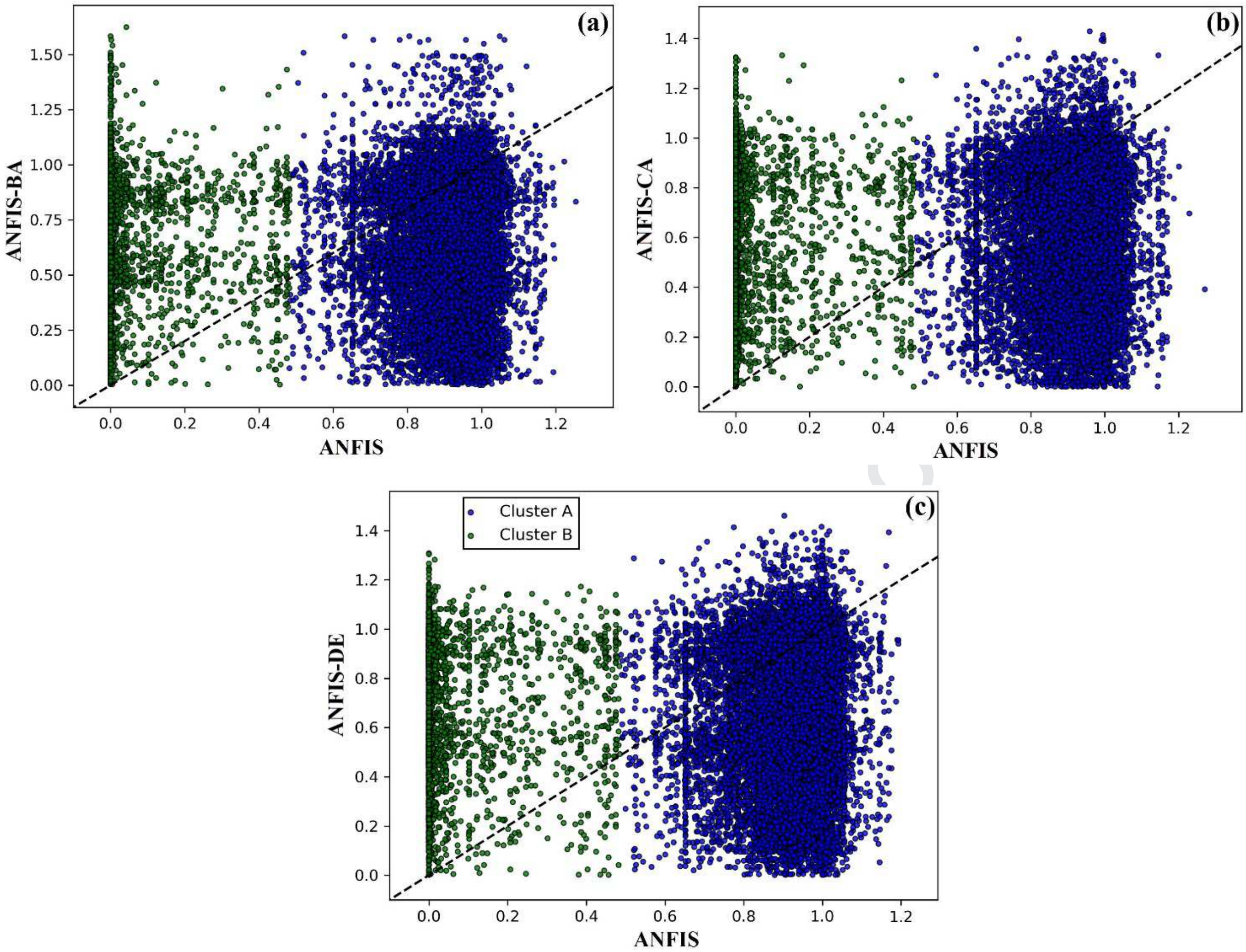


Fig. 6 The results of cluster analysis: a) ANFIS-BA versus ANFIS, b) ANFIS-CA versus ANFIS, and c) ANFIS-DE versus ANFIS

Highlights

- A new framework was developed for identification of dust-sources.
- Three novel hybridized ANFIS models were developed: ANFIS-BA, ANFIS-CA, ANFIS-DE.
- The hybridized ANFIS-DE model had the highest accuracy (AUC=84.1%, TSS=0.73).
- All hybridized models outperformed the standalone ANFIS model.

Conflicts of Interest:

The authors declare no conflict of interest