

1. 1. M.Sc. in Remote Sensing and GIS, Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran
2. **Corresponding Author**, Assistant Professor, Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran
Email: b_mirbaqheri@sbu.ac.ir **ORCID:** <https://orcid.org/0000-0002-4161-1540>
3. Associate Professor, Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran

Article Info	ABSTRACT
Article type: Research Article Article history: Received: [REDACTED] Revised: [REDACTED] Accepted: [REDACTED] Published: [REDACTED] Keywords: Spatial modeling, Desertification, Machine Learning Algorithms, Central Region of Iran.	In recent years, desertification has become one of the most significant global environmental challenges, highlighting the need to identify areas prone to desertification and the factors driving this process. In the present study, desertification was modeled using machine learning algorithms across the provinces of Isfahan, Markazi, Qom, Kerman, and Yazd during 2002–2010. The trained algorithms were then applied to predict desertification and validate their performance for a second period, 2011–2018. In this research, the Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting Machine (GBM), and Extreme Gradient Boosting (XGBoost) algorithms were employed to predict the spatial distribution of desertification. To identify desertified areas, remote sensing indices were utilized, while precipitation, air temperature, elevation, direct soil evaporation, wind speed, air humidity, rural population density, and distances from cities and roads were considered as input variables. After optimizing the hyperparameters, the Area Under the Receiver Operating Characteristic Curve (AUC) was calculated to evaluate model performance and determine the most accurate algorithm. The results for the first period indicated that RF outperformed the other algorithms, achieving the highest accuracy with an AUC value of 0.879. In addition, the AUC values for SVM, GBM, and XGBoost were 0.872, 0.872, and 0.861, respectively. RF also demonstrated strong temporal stability by maintaining the highest predictive accuracy (AUC = 0.800) during the second period. During this period, the AUC values for GBM, XGBoost, and SVM were 0.780, 0.770, and 0.564, respectively. Moreover, RF achieved the highest accuracy in Qom (0.655), Kerman (0.818), Markazi (0.898), and Yazd (0.798). Finally, wind speed was identified as the most influential variable contributing to desertification in the study area. This study demonstrated that machine learning algorithms, particularly RF, are effective tools for predicting desertification risk in central Iran.
How to Cite: Last Name, Initial., Last Name, Initial., & Last Name, Initial. (2021). Title of paper. <i>Journal of Natural Environmental Hazards</i>, -- (--), ----.	
	© The Author/Authors DOI: 00000000000000000000 Publisher: University of Sistan and Baluchestan

EXTENDED ABSTRACT

INTRODUCTION

Desertification, defined as land degradation in arid, semi-arid, and dry sub-humid regions, is primarily driven by human activities and climate change (Chen et al., 2019; Feng et al., 2022). According to the Desertification and Land Degradation Conference, this phenomenon is projected to affect more than 785 million people living in dryland areas—equivalent to 17.7% of the global population—in the coming decades (Kamali Maskooni et al., 2021). Various models have been employed to assess the severity, spatial dynamics, and future trends of desertification in Iran and worldwide, including IMDPA (Fakhrabadi et al., 2019; Geldizadeh et al., 2023), MEDALUS (Boali et al., 2020), MICD (Silakhuri et al., 2014), and CA–Markov (Halabian and Soltanian, 2017). Nevertheless, the application of machine learning algorithms in desertification studies remains relatively limited. The present study aims to model desertification during the first period (2002–2010) using machine learning algorithms and subsequently extend the trained model to predict desertification patterns for the second period (2011–2018). In addition, the study identifies the key environmental and anthropogenic drivers contributing to desertification and generates a probabilistic map of its potential occurrence.

DATA AND METHODOLOGY

The study area encompasses the provinces of Qom, Markazi, Isfahan, Yazd, and Kerman. Initially, a desertification reference dataset consisting of desertified (class 1) and non-desertified (class 0) locations was generated using NDVI, TGSi, and Albedo indices combined with visual interpretation of Google Earth imagery. Areas that showed no signs of desertification during the first period but exhibited degradation in the second period were classified as desertified points. In the second step, a set of predictor variables—including precipitation, direct soil evaporation, wind speed, near-surface humidity and relative humidity, air temperature, elevation, rural population density, distance to cities, and distance to roads—was compiled for the two time intervals, 2002–2010 and 2011–2018.

Subsequently, four machine learning algorithms (SVM, RF, GBM, and XGBoost) were trained using grid search hyperparameter tuning combined with cross-validation, and a probabilistic desertification map was produced for the first period (2002–2010). Model performance was evaluated using the Area Under the ROC Curve (AUC). The trained models were then transferred to the second period (2011–2018) to generate the corresponding desertification probability map. Finally, the importance of the contributing variables was quantified using the Mean Decrease Impurity (MDI) criterion within the RF algorithm.

RESULTS AND DISCUSSION

Table 1 presents the accuracy assessment of the machine learning algorithms based on AUC values for the first study period (2002–2010). During this period, the AUC values obtained for the test samples were 0.879 for RF, 0.872 for both SVM and GBM, and 0.861 for XGBoost. These results indicate that all four algorithms achieved an acceptable and reliable performance in modelling desertification during the first time interval.

Table 1: AUC values of machine learning algorithms (2002–2010)

Algorithm	AUC (Training Samples)	AUC (Testing Samples)
RF	0.877	0.879
SVM	0.877	0.872
GBM	0.875	0.872
XGBoost	0.873	0.861

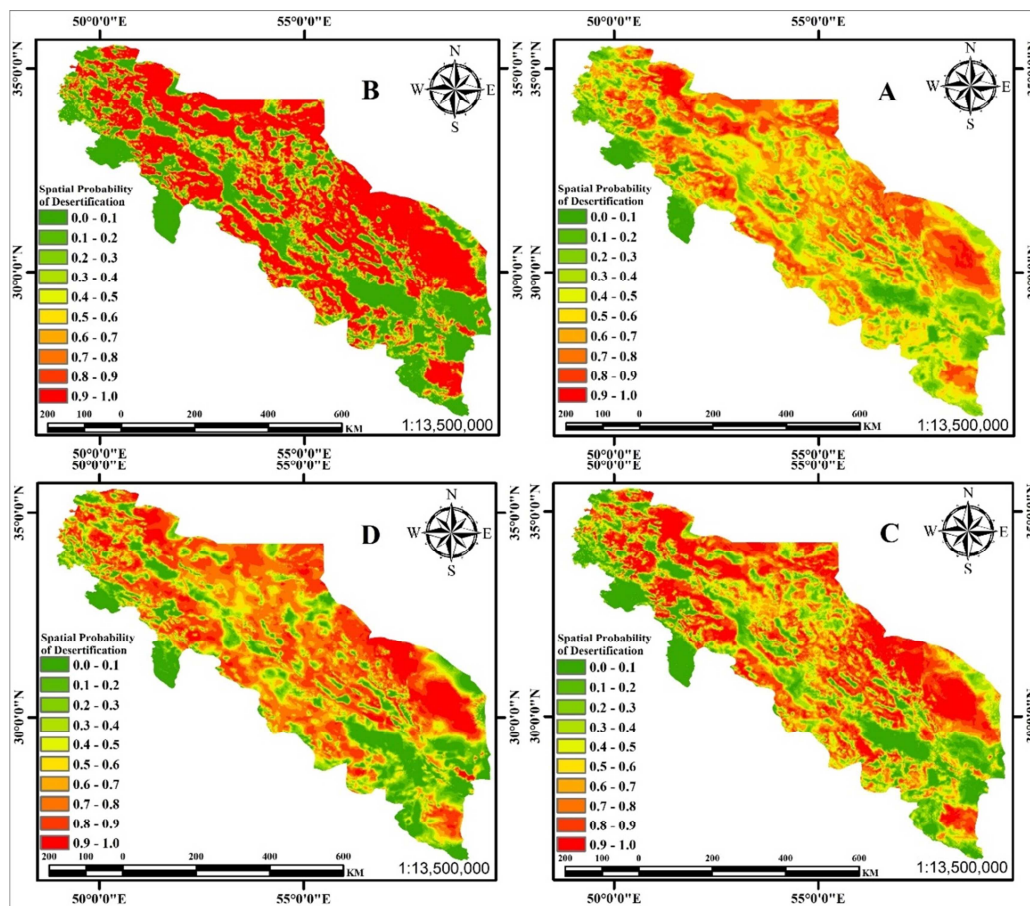


Figure 1: Desertification probability maps for the period 2002–2010: RF (a); GBM (b); XGBoost (c); and SVM (d)

Table 2 summarizes the accuracy assessment of the machine learning algorithms based on AUC values for the second study period (2011–2018).

Table 2: AUC values of machine learning algorithms (2011–2018)

Algorithm	AUC
RF	0.800
SVM	0.564
GBM	0.780
XGBoost	0.770

During this period, the RF algorithm achieved the highest performance with an AUC of 0.800, whereas the SVM algorithm yielded the lowest accuracy, with an AUC of 0.564. These results indicate that the Random Forest (RF) algorithm exhibits a higher generalization capability with respect to temporal variations and shifts in desertification patterns, making it a more reliable approach for future predictions. The superior performance of the RF algorithm can be attributed to its ensemble nature. In this method, each decision tree is constructed through bootstrap sampling and the random selection of a subset of predictor variables. This diversity in tree construction reduces inter-tree correlation and consequently decreases the generalization error. As a result, increasing the number of trees enhances model stability and predictive accuracy, ultimately minimizing the overall error of the model (Saha et al., 2020).

The importance of variables influencing desertification in the study area, as assessed by the RF algorithm, is presented in Table 3. According to these results, wind speed (58.13), direct soil evaporation (56.66), and air humidity (53.29) were identified as the most influential factors contributing to desertification. Other variables ranked lower in importance, with comparatively smaller contributions.

Table 3: Variable importance in desertification within the study area

Variable	Importance
Wind speed	58.13
Direct soil evaporation	56.66
Air humidity	53.39
Distance from cities	37.14
Population density	36.77
Daily air temperature	36.45
Distance from roads	35.85
Elevation	35.56
Near-surface humidity	33.85
Precipitation	32.30

This finding is consistent with previous research conducted in Iran, which identifies arid climatic conditions as a primary driver of desertification (Akbari et al., 2007; Halabian & Jamshidian, 2019; Heydari Alamdarloo et al., 2022; Zeaiean Firouzabadi et al., 2013). Moreover, wind speed—identified as the most influential factor in this study—intensifies wind erosion, affecting soil structure and fertility, and is widely recognized as one of the major drivers of land degradation in arid regions (Lackóová et al., 2023).

CONCLUSION

Spatial modeling using machine learning algorithms provides an effective approach for identifying areas susceptible to desertification and analyzing the key drivers influencing this process. The results of this study demonstrated that the Random Forest (RF) algorithm outperforms other methods in terms of accuracy, temporal stability, and generalization capability, making it a robust tool for predicting future desertification in arid regions of Iran. Furthermore, the findings highlighted that climatic factors, particularly wind speed, soil evaporation, and air humidity, play a critical role in the expansion of desertification. These insights can assist environmental planners in identifying high-risk areas and developing more effective strategies to mitigate and reverse the desertification process.

ETHICAL CONSIDERATIONS

Conflict of Interest Statement: The authors declare no conflict of interest.

Ethical Statement: This article does not contain any studies with human participants or animals performed by any of the authors.

REFERENCES

References [in Persian]

Akbari, Morteza, Hamid Reza Karim Zadeh, Reza Modares, and Bahareh Chakoshi. (2007). "Assessment and Classification of Desertification Using RS & GIS Techniques (Case Study: The Arid Region, in the North of Isfahan)." *Iranian Journal of Range and Desert Research*, 14(2):124–42. [In Persian]

- Boali, Abdalhossein, Hamidreza Asgri, Ali Mohammadian Behbahani, Abdolrassoul Salmanmahiny, and Babak Naimi. (2021). "Evaluation of Desertification Intensity Based on Quantitative and Qualitative Changes in Groundwater and Soil Criteria Using Madalus Model and Geostatistical Methods." *Environmental Sciences*, 19(4):85–106. doi: 10.52547/envs.2021.37277. [In Persian]
- Fakhrabadi, Amir, Mohamad Salighe, Mehri Akbari, and Mohammed Hossai Nasserzadeh. (2019). "Evaluation of Desertification Intensity in Kashan Plain." *Geographical Research*, 34(4):527–38. doi: 10.29252/geores.34.4.527. [In Persian]
- Geldizadeh, Faezeh, Roshan, Gholamreza, & Ghanghermeh, Abdolazim. (2023). Monitoring the temporal-spatial changes of desertification in Gavakhoni Basin based on climatic criteria. *Climate Change Research*, 4(13), 55–68. doi.org/10.30488/ccr.2023.396869.1133. [In Persian]
- Halabian, Amir Hossien, and Mahmood Soltanian. (2017). "Assessment and Forecasting the Desertification Changes in the East and South of Isfahan by CA-Markov Model." *Journal of Spatial Analysis Environmental Hazards*, 3(4):71–88. doi: 10.18869/acadpub.jsaeh.3.4.71. [In Persian]
- Halabian, Amirhossein, and Zohreh Jamshidian. (2019). "Analysis of Desertification Changes with Emphasis on Precipitation in Qom Province." *Emergency Management*, 7(2):105–19. [In Persian]
- Heydari Alamdarloo, Esmail, Hassan Barabadi, Dehghan Dehghan Rahimabadi, Hassan Khosravi, and Javad Rafieh Sharifiabad. (2022). "Evaluation of Control Strategies for Desertification Using SWOT Model in Yazd-Ardakan Plain." *Integrated Watershed Management*, 2(1):1–14. doi: 10.22034/iwm.2022.550355.1025. [In Persian]
- Kamali Maskooni, Ehsan, Mohammad Amin Kamali, and Ali Khanamani. (2021). "Investigation and Preparation of Desertification Map Based on Iranian Model of Desertification Potential (IMDPA) with an emphasis on two criteria of soil and vegetation (Case study: Faryab-Kerman Province)." *Journal of Environmental Science and Technology*, 12(12):163. doi: 10.22034/jest.2021.30618.3917. [In Persian]
- Silakhuri, Esmail, Majid Onagh, Amir Sadodin, and Esmail Filekesh. (2014). "Comparing Efficiency of Iranian Desert Hazard Assessment Models, Namely MICD and IMDPA (Case Study: Sabzevar Region)." *Journal of Water and Soil Conservation*, 21(4):1–28. [In Persian]
- Zeaiean Firouzabadi, Parvis, Reza Talkhabi Hamid, and Leila Hosseinijani. (2014). "CHANGE DETECTION OF MIGHAN DESERT BASIN USING ETM+, TM, MSS IMAGES AND CLIMATE DATA FROM 1973 TO 2011." *Journal of Applied Research in Geographical Sciences*, 13(31):173–89. [In Persian]

References [in English]

- Chen, W., Pourghasemi, H. R., Zhang, S., & Wang, J. (2019). 21—A Comparative Study of Functional Data Analysis and Generalized Linear Model Data-Mining Techniques for Landslide Spatial Modeling. In H. R. Pourghasemi & C. Gokceoglu (Eds.), *Spatial Modeling in GIS and R for Earth and Environmental Sciences* (pp. 467–484). Elsevier. <https://doi.org/10.1016/B978-0-12-815226-3.00021-1>
- Feng, K., Wang, T., Liu, S., Kang, W., Chen, X., Guo, Z., & Zhi, Y. (2022). Monitoring Desertification Using Machine-Learning Techniques with Multiple Indicators Derived from MODIS Images in Mu Us Sandy Land, China. *Remote Sensing*, 14(11), 2663. <https://doi.org/10.3390/rs14112663>
- Lackóová, L., Kaletová, T., & Halászová, K. (2023). Are Drought and Wind Force Driving Factors of Wind Erosion Climatic Erosivity in a Changing Climate? A Case Study in a Landlocked Country in Central Europe. *Land*, 12(4), Article 4. <https://doi.org/10.3390/land12040757>
- Saha, S., Saha, M., Mukherjee, K., Arabameri, A., Ngo, P. T. T., & Paul, G. C. (2020). Predicting the deforestation probability using binary logistic regression, random forest, ensemble rotational forest, REPTree: A case study at the Gumani River Basin, India. *Science of The Total Environment*, 730, 139197. <https://doi.org/10.1016/j.scitotenv.2020.139197>

مدل سازی مکانی احتمال رخداد بیابان زایی در نواحی مرکزی ایران بر اساس الگوریتم های یادگیری ماشین

احسان الله قلی پور^۱، بابک میرباقری^{۲*} , علیرضا شکبیا^۳

۱. کارشناس ارشد سنجش از دور و GIS، مرکز مطالعات سنجش از دور و GIS، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران

۲. استادیار مرکز مطالعات سنجش از دور و GIS، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران (نویسنده مسئول)

ایمیل: b_mirbagheri@sbu.ac.ir ORCID: <https://orcid.org/0000-0002-4161-1540>

۳. دانشیار مرکز مطالعات سنجش از دور و GIS، دانشکده علوم زمین، دانشگاه شهید بهشتی، تهران، ایران

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی تاریخ دریافت: --/--/-- تاریخ ویرایش: --/--/-- تاریخ پذیرش: --/--/-- تاریخ انتشار: --/--/-- واژه های کلیدی: مدل سازی مکانی، بیابان زایی، الگوریتم های یادگیری ماشین، نواحی مرکزی ایران.	بیابان زایی به عنوان یکی از چالش های محیط زیستی مهم قرن حاضر، با کاهش حاصلخیزی خاک، تخریب پوشش گیاهی و کاهش کارایی اکوسیستم ها همراه است. ایران به دلیل اقلیم خشک و نیمه خشک خود، به شدت در معرض این پدیده قرار دارد. از اینرو شناسایی پهنه های مستعد بیابان زایی و محرک های موثر بر آن برای مدیریت سرزمین ضروری است. در این پژوهش مدل سازی مکانی بیابان زایی در نواحی مرکزی ایران شامل استان های اصفهان، مرکزی، قم، کرمان و یزد با استفاده از الگوریتم های یادگیری ماشین شامل جنگل تصادفی (RF)، ماشین بردار پشتیبان (SVM)، ماشین تقویت گرادیان (GBM) و الگوریتم XGBoost انجام شد. دوره زمانی ۱۳۸۰-۱۳۸۹ برای آموزش مدل ها و دوره ۱۳۸۹-۱۳۹۷ جهت پیش بینی و ارزیابی پایداری نتایج در نظر گرفته شد. برای استخراج مناطق بیابان زده، از شاخص های سنجش از دوری و برای مدل سازی از متغیرهای بارندگی، دما، ارتفاع، تبخیر از سطح خاک، سرعت باد، رطوبت نسبی، تراکم جمعیت روستایی و فاصله از شهرها و راه ها استفاده شد. پس از بهینه سازی فرایلامترها، عملکرد الگوریتم ها با شاخص مساحت زیر منحنی ROC(AUC) ارزیابی شد. نتایج نشان داد در دوره اول زمانی، الگوریتم RF با مقدار AUC برابر ۰/۸۷۹ بهترین عملکرد را داشته است. علاوه بر این مقدار AUC برای الگوریتم های (SVM (872/0)، (GBM (872/0) و XGBoost (861/0) بود. ارزیابی پایداری مدل ها در دوره دوم حاکی از حفظ دقت بالای الگوریتم RF با مقدار AUC برابر ۰/۸۰۰ است. در این دوره زمانی مقدار AUC برای الگوریتم های (GBM (780/0)، (XGBoost (770/0) و (SVM (564/0) بود. ضمن اینکه الگوریتم RF در استان های قم (۰/۶۵۵)، کرمان (۰/۸۱۸)، مرکزی (۰/۸۹۸) و یزد (۰/۷۹۸) نیز بالاترین دقت را داشت. براساس نتایج تحلیل اهمیت متغیرها، سرعت باد به عنوان مهم ترین عامل موثر بر روند بیابان زایی در منطقه شناسایی شد. این پژوهش نشان داد الگوریتم های یادگیری ماشین، به ویژه RF، ابزار مناسبی برای پیش بینی و پهنه بندی خطر بیابان زایی در نواحی خشک و نیمه خشک مرکزی کشور هستند.

استناد: نام خانوادگی، نام؛ نام خانوادگی، نام؛ نام خانوادگی، نام (۱۴۰۰). عنوان مقاله. مخاطرات محیط طبیعی، -- (--)، ---

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

مقدمه

بیابان‌زایی به عنوان یکی از چالش‌های عمده زیست محیطی، اقتصادی و اجتماعی جهان، به صورت تخریب سرزمین در مناطق خشک، نیمه خشک و خشک نیمه مرطوب تعریف می شود و عمدتاً تحت تاثیر فعالیت‌های انسانی و تغییرات آب و هوایی رخ می دهد. این پدیده موجب کاهش حاصلخیزی خاک، تخریب پوشش گیاهی و تهدید امنیت اکولوژیکی و روند توسعه اقتصادی را در مقیاس‌های مختلف محدود می کند (W. Chen et al., 2019; K. Feng et al., 2022). با توجه به اثرات بلند مدت و مخرب آن بر اکوسیستم‌های طبیعی و انسانی، بررسی عوامل مؤثر در وقوع و گسترش بیابان‌زایی ضرورت دارد (مصباح‌زاده و سلیمانی ساردو، ۲۰۱۹).

بر اساس گزارش کنفرانس بیابان‌زایی و بیابان‌زدایی، بیش از ۷۸۵ میلیون نفر انسان در مناطق خشک معادل ۱۷/۷ درصد جمعیت جهان در معرض خطر بیابان‌زایی قرار دارند که در این میان، ۶۰ تا ۱۰۰ میلیون نفر به طور مستقیم تحت تأثیر کاهش حاصلخیزی اراضی و سایر پیامدهای این فرایند، قرار می گیرند (کمالی مسکونی و همکاران، ۱۳۹۹). بیابان‌زایی صرفاً در مجاورت بیابان‌ها رخ نمی دهد و حتی مناطق دور از بیابان‌های موجود نیز می توانند در معرض این خطر باشند (D'Odorico et al., 2013).

در ایران، حدود ۸۵ درصد اراضی در اقلیم خشک، نیمه خشک و فرا خشک قرار دارد و روند یک درصدی افزایش سالیانه بیابان‌زایی، ضرورت بکارگیری روش‌های علمی برای ارزیابی و پیش‌بینی این پدیده را دو چندان کرده است (بخشنده مهر و همکاران، ۲۰۱۳). بیابان‌زایی پدیده‌ای چند عاملی و پویاست که ترکیبی از محرک‌های طبیعی و انسانی، شامل تغییرات اقلیمی، کاهش بارندگی، افزایش دما، تخریب پوشش گیاهی، استفاده بی رویه از آب‌های زیرزمینی، تغییر کاربری اراضی، چرای مفرط و رشد شهرنشینی در شکل‌گیری آن نقش دارد (Olagunju, 2015؛ صالحی و کرباسی، ۱۴۰۰؛ میجانی و همکاران، ۱۴۰۱).

نخستین گام در مقابله با بیابان‌زایی، شناسایی مناطق حساس و محرک‌های مؤثر آن است، تا از طریق تخصیص بهینه منابع بتوان راهبردهای مؤثرتری اتخاذ کرد (Geist & Lambin, 2004؛ Hu et al., 2020؛ Xu et al., 2022؛ Mihi et al., 2024؛ جهان‌شاهی و همکاران، ۱۳۹۴؛ میجانی و همکاران، ۱۴۰۱).

مدل‌سازی به عنوان ابزاری کلیدی برای ساده‌سازی سیستم‌های پیچیده و استخراج دانش از داده‌های حجیم عمل می کند (شایان و شریفی، ۱۳۸۵). در سال‌های اخیر، یادگیری ماشین به ویژه به دلیل توانایی در برقراری روابط غیرخطی میان متغیرها و افزایش دقت پیش‌بینی، به یکی از روش‌های نوین و کارآمد در مدل‌سازی پدیده‌های محیطی نظیر بیابان‌زایی تبدیل شده است (Xu et al., 2024).

مدل‌های مختلفی برای ارزیابی و پیش‌بینی بیابان‌زایی توسعه یافته که از جمله این می توان به ^۱IMDPA، ^۲ESAS، ^۳MEDALUS، ^۴MICD، ^۵GLASOD، ^۶ICD، CA-Markov و الگوریتم‌های یادگیری ماشین اشاره کرد. این مدل‌ها عمدتاً معیارهای اقلیمی، خاک، زمین‌شناسی، پوشش گیاهی و شاخص‌های اجتماعی - اقتصادی را در بر می گیرند (فیضی کوشکی و همکاران، ۱۳۹۸).

¹ - Iranian Model of Desertification Potential Assessment

² - Environmentally Sensitive Areas

³ - Mediterranean Desertification and Land Use

⁴ - Modified Iranian Classification of Desertification

⁵ - Global Assessment of Soil Degradation

⁶ - Iranian classification of desertification

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

مطالعات متعدد انجام شده در ایران و جهان نشان داده اند که مدل هایی مانند IMDPA, MICD, MEDALUS و CA-Markov در ارزیابی شدت و پهنه بندی مناطق مستعد بیابان زایی کاربرد گسترده ای دارند (سیلاخوری و همکاران در سال ۱۳۹۳؛ فخرآبادی و همکاران در سال ۱۳۹۸؛ اونیق و همکاران در سال ۱۳۹۹؛ گلدی زاده و همکاران در سال ۱۴۰۲).

در سال های اخیر استفاده از الگوریتم های یادگیری ماشین در مدل سازی بیابان زایی نیز مورد توجه قرار گرفته است، مانند مطالعه Boali و همکاران (۲۰۲۴) که با بکارگیری الگوریتم های یادگیری ماشین و شاخص های سنجش از دوری به مدل سازی بیابان زایی در دو دوره زمانی در ناحیه ای محدود از استان گلستان پرداخته است.

با این وجود مرور پیشینه تحقیقاتی نشان می دهد استفاده از رویکردهای پیشرفته داده مبنای، به خصوص الگوریتم های یادگیری ماشین برای مدل سازی و پیش بینی مکانی بیابان زایی، همچنان در این حوزه محدود بوده و نیازمند انجام مطالعات بیشتر در این حوزه است.

هدف اصلی این پژوهش، مدل سازی مکانی بیابان زایی با بهره گیری از الگوریتم های یادگیری ماشین و سنجش پایداری زمانی آنها در نواحی مرکزی ایران (استانهای قم، مرکزی، اصفهان، یزد و کرمان) و تولید نقشه پیش بینی مکانی پهنه های بیابان زده و مستعد بیابان زایی است. افزون بر این، شناسایی و رتبه بندی مهم ترین عوامل و محرک های مؤثر بر بیابان زایی بر اساس رویکرد داده مبنای از دیگر اهداف این پژوهش به شمار می رود.

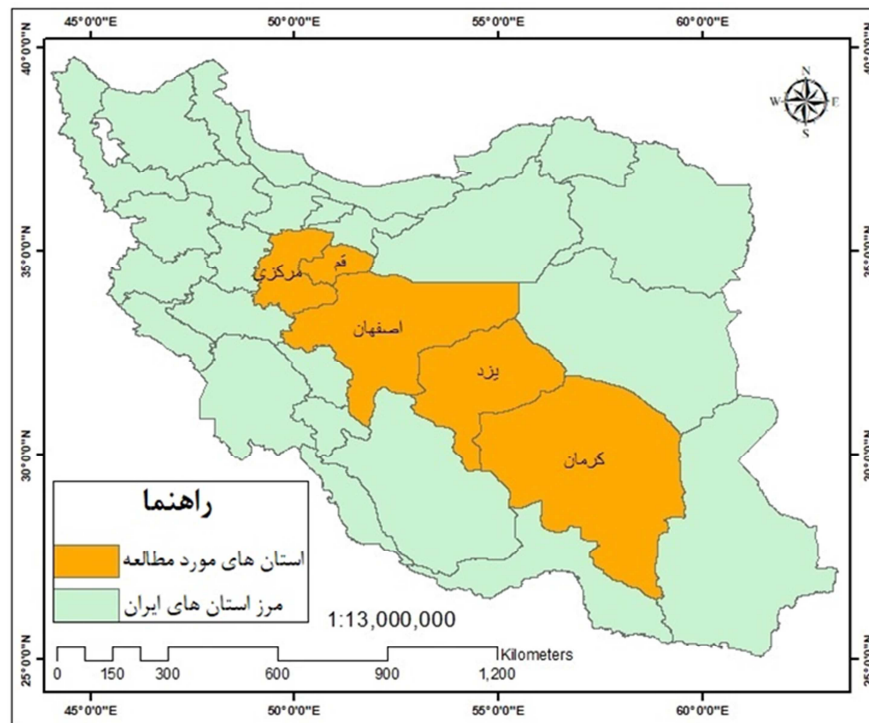
لذا نوآوری این تحقیق در ترکیب مقیاس مکانی چند استانه و استفاده از رویکرد پیشرفته یادگیری ماشین جهت مدل سازی پدیده بیابان زایی در مناطق مرکزی ایران است، ضمن اینکه ارزیابی پایداری زمانی این الگوریتم ها در مدل سازی بیابان زایی نیز از جمله مهمترین نوآوری های این تحقیق به شمار می رود.

این مقاله در نشریه علمی پژوهشی «مجله علمی پژوهشی جغرافیا» به شماره ۱۳۹، زمستان ۱۴۰۳، به چاپ رسیده است.

منطقه مورد مطالعه

منطقه مورد مطالعه در این پژوهش شامل استان های قم، مرکزی، اصفهان، یزد و کرمان در بخش مرکزی ایران است. این ناحیه بخش عمده ای از حوزه اقلیمی فلات مرکزی کشور را در بر می گیرد که از دشت های خشک تا نواحی کویری گسترده مانند دشت لوت و دشت کویر را شامل می شود. این مناطق دارای نوسانات دمایی قابل توجه، از حدود ۱۵ درجه سانتی گراد زیر صفر تا ۵۰ درجه سانتی گراد بالای صفر هستند. متوسط رطوبت نسبی در این دشت ها حدود ۳۰ تا ۴۰ درصد و متوسط بارندگی سالانه کمتر از ۱۰۰ میلی متر، غالباً نزدیک به ۵۰ میلی متر گزارش شده است (آزادی و همکاران ۱۳۹۴). شرایط آب و هوایی خشک و وجود مناطق بیابانی موجب شده است تا این مناطق از نظر محیط زیستی به شدت شکننده و حساس بوده و در صورت فراهم شدن شرایط مستعد، فرایند بیابان زایی و تخریب سرزمین با سرعت قابل توجهی در آن ها رخ دهد. شکل ۱ محدوده مورد مطالعه را نشان می دهد.

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.



شکل ۱: منطقه مورد مطالعه

بدون شماره مجله

داده‌های پژوهش

استخراج مناطق بیابان‌زایی

برای شناسایی و استخراج مناطق بیابان‌زایی، از سه شاخص سنجش از دوری شامل شاخص پوشش گیاهی، شاخص اندازه دانه خاک سطحی و آلبدو استفاده شد.

شاخص پوشش گیاهی تفاوت نرمال شده ($NDVI^1$):

برای محاسبه شاخص NDVI از محصول MOD13A1 سنجنده مادیس با قدرت تفکیک مکانی ۵۰۰ متر استفاده شد. نحوه محاسبه شاخص NDVI بر اساس رابطه (۱) است:

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (1)$$

شاخص اندازه دانه خاک سطحی ($TGSI^2$):

شاخص اندازه دانه خاک سطحی با استفاده از داده‌های روزانه محصول MCD43A4 سنجنده مادیس با قدرت تفکیک مکانی ۵۰۰ متر محاسبه شد. رابطه (۲) نحوه محاسبه این شاخص را ارائه می‌دهد (Hashem Geloogerdi et al., 2023):

¹ - Normalized Difference Vegetation Index

² - Topsoil Grain Size Index

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

$$TGSi = \frac{Red - Blue}{Red + Blue + Green} \quad (2)$$

داده آلبدو^۱:

داده های آلبدو روزانه نیز از محصول MCD43A3 سنجنده مادیس با قدرت تفکیک ۵۰۰ متر استخراج شد.

داده های مورد استفاده جهت مدل سازی بیابان‌زایی

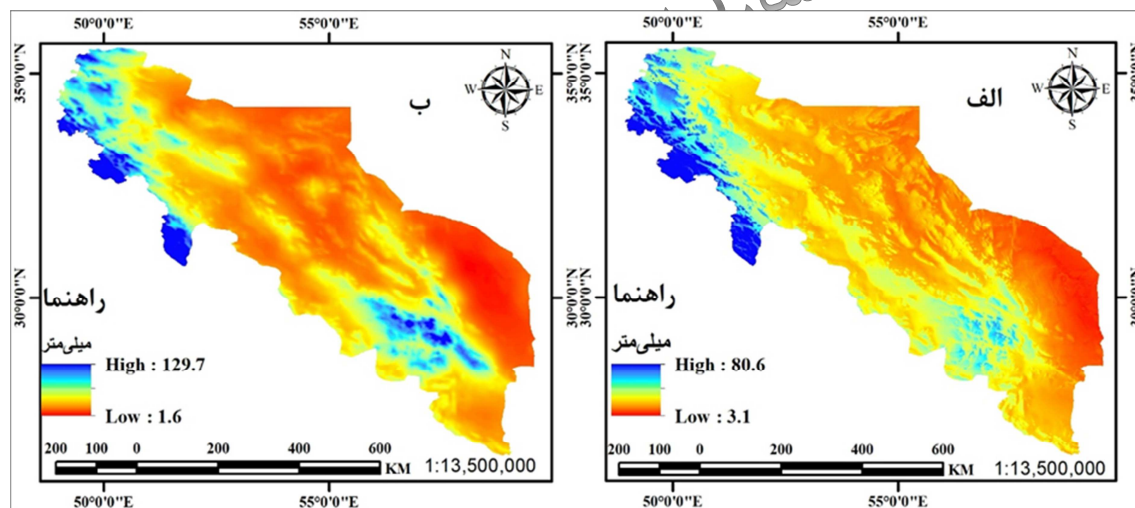
در مطالعه حاضر جهت مدل‌سازی پدیده بیابان‌زایی از داده‌هایی که تأثیر شرایط آب‌وهوایی و همچنین تأثیر فعالیت‌های انسان در منطقه را نشان می‌دهد، استفاده شد.

داده های اقلیمی

به جز داده تبخیر مستقیم از سطح خاک سایر داده های اقلیمی از پایگاه CHLSA^۲ با تفکیک مکانی حدود ۱ کیلومتر دریافت شدند. داده ها به صورت ماهانه استخراج و میانگین چند ساله برای هر دوره محاسبه شد.

بارندگی:

بارندگی از عوامل گسترش یا کاهش بیابان‌زایی است، بنابراین مشابه مطالعات Zhang et al, (2020) (از داده بارندگی به عنوان متغیر کلیدی استفاده شد (شکل ۲).



شکل ۲: بارندگی دوره ۱۳۸۰-۱۳۸۹ (الف)، بارندگی دوره ۱۳۹۷-۱۳۹۹ (ب)

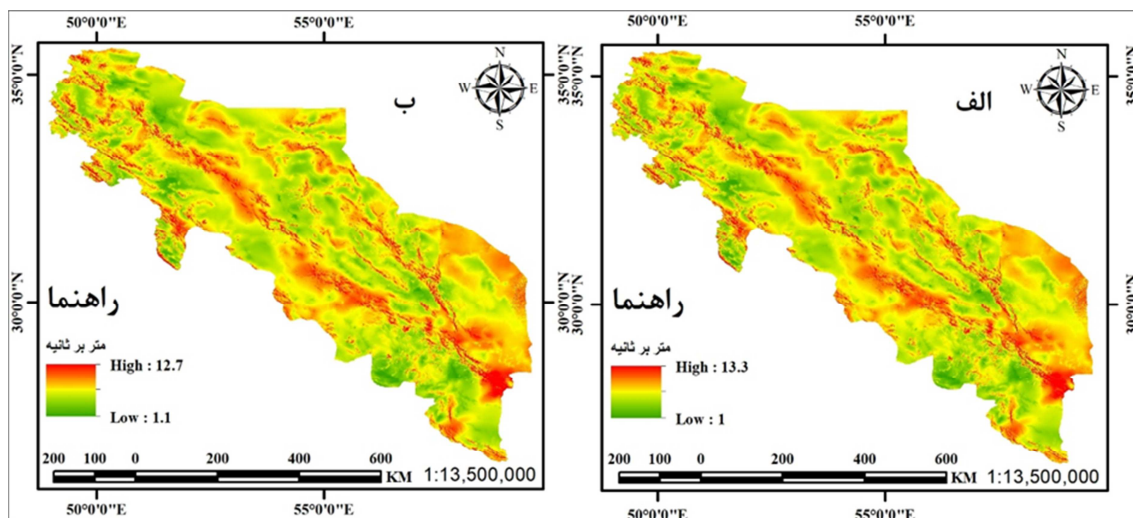
سرعت باد نزدیک به سطح:

سرعت باد از عوامل مهم بیابان‌زایی به شمار می‌رود Al-Bakri et al, (2016). در این پژوهش از داده‌های سرعت باد برای مدل‌سازی استفاده شد (شکل ۳).

¹ - Albedo

² - <https://chelsa-climate.org/>

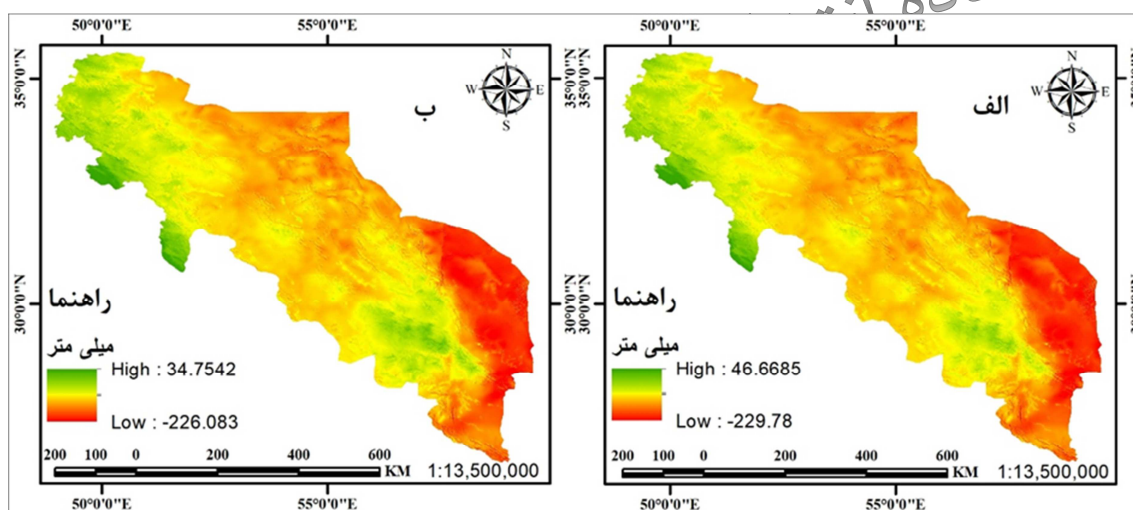
این نسخه موقت و قبل از چاپ نهایی می باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.



شکل ۳: سرعت باد دوره ۱۳۸۰-۱۳۸۹ (الف)، سرعت باد دوره ۱۳۸۹-۱۳۹۷ (ب)

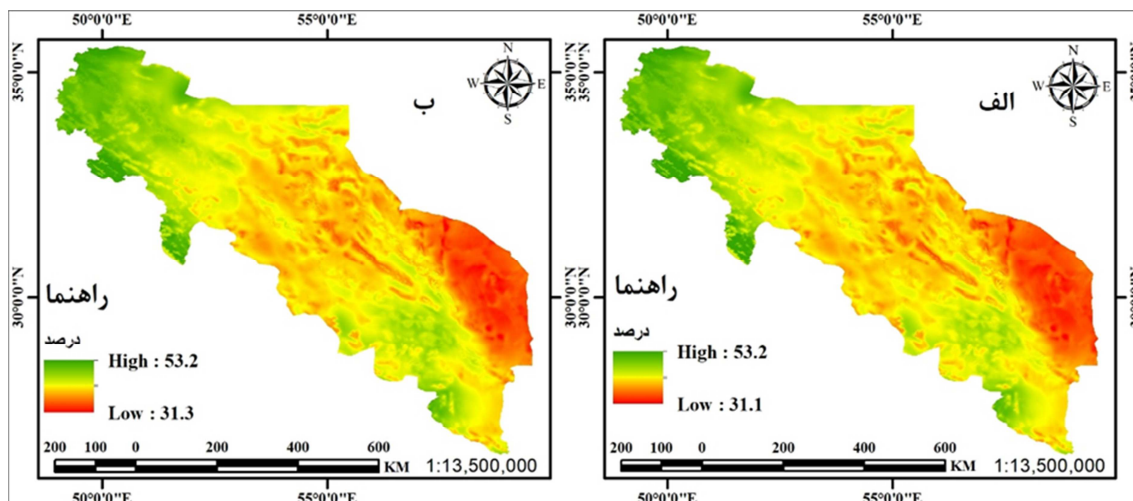
رطوبت هوا:

دو مجموعه داده شامل داده های رطوبت هوا و رطوبت نسبی نزدیک به سطح، با توجه به نقش آن ها در بیابانزایی Li et al, (2007)، مورد استفاده قرار گرفتند (شکل ۴ و شکل ۵).



شکل ۴: رطوبت هوا دوره ۱۳۸۰-۱۳۸۹ (الف)، رطوبت هوا دوره ۱۳۸۹-۱۳۹۷ (ب)

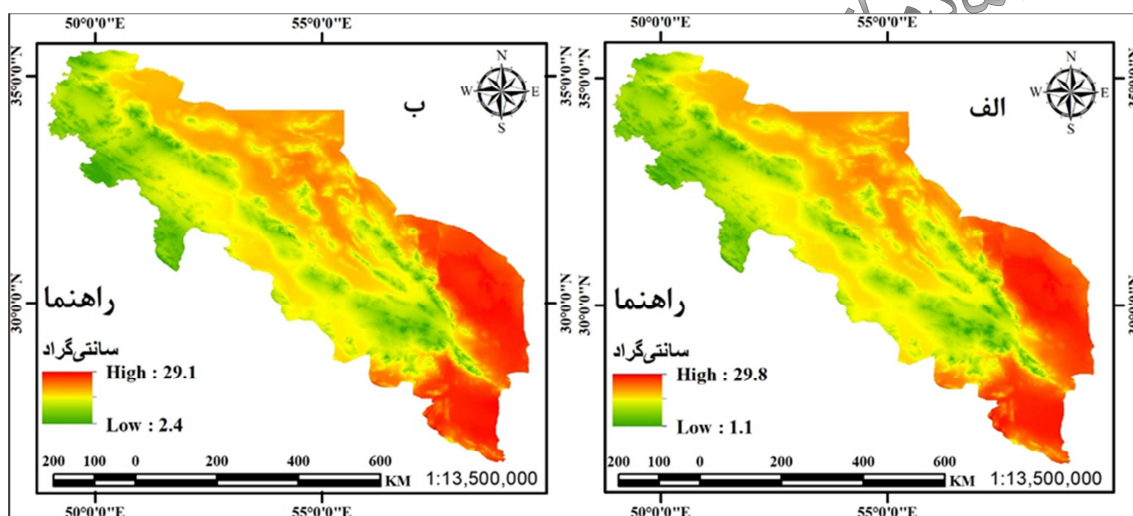
این نسخه موقت و قبل از چاپ نهایی می باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.



شکل ۵: رطوبت نسبی دوره ۱۳۸۰-۱۳۸۹ (الف)، رطوبت نسبی دوره ۱۳۸۹-۱۳۹۷ (ب)

دمای روزانه:

میانگین دمای روزانه به دلیل ارتباط با فرایندهای خشکی و تبخیر در مدل لحاظ شد. این متغیر پیش تر در پژوهش های (Feng et al., 2015 و Huang et al., 2020) استفاده شده است (شکل ۶).

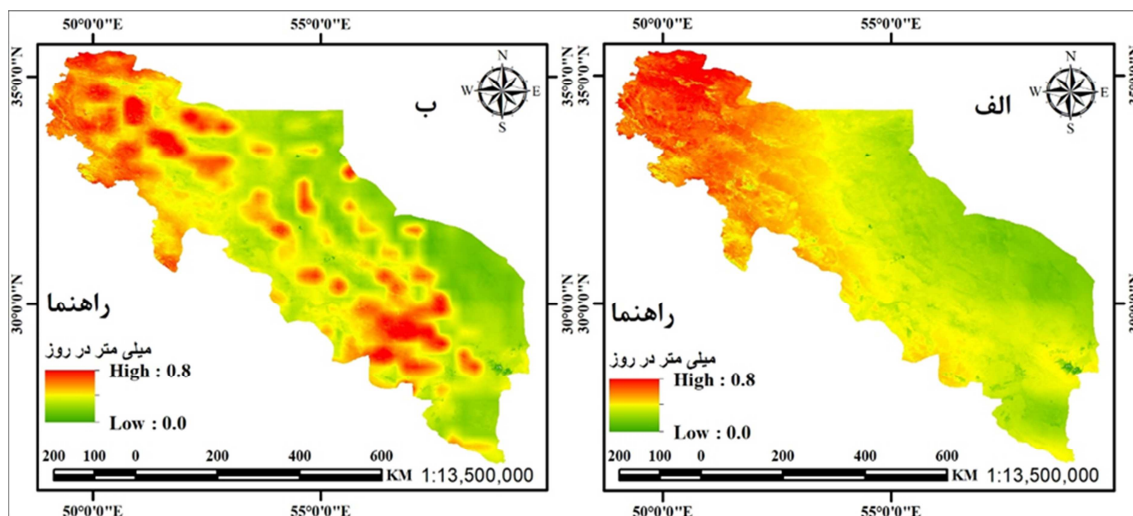


شکل ۶: دمای هوا دوره ۱۳۸۰-۱۳۸۹ (الف)، دمای هوا دوره ۱۳۸۹-۱۳۹۷ (ب)

تبخیر مستقیم از سطح خاک:

داده تبخیر از سطح خاک که نقش مهمی در خشکی سطح دارد، از مجموعه داده PML^۱ V2 در محیط Google Earth Engin استخراج شد (Han و همکاران؛ 2021). شکل ۷ نقشه تبخیر مستقیم از سطح خاک را نشان می دهد.

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.



شکل ۷: تبخیر مستقیم از خاک دوره ۱۳۸۰-۱۳۸۹ (الف)، تبخیر مستقیم از خاک دوره ۱۳۸۹-۱۳۹۷ (ب)

مدل رقومی ارتفاع^۱ (DEM):

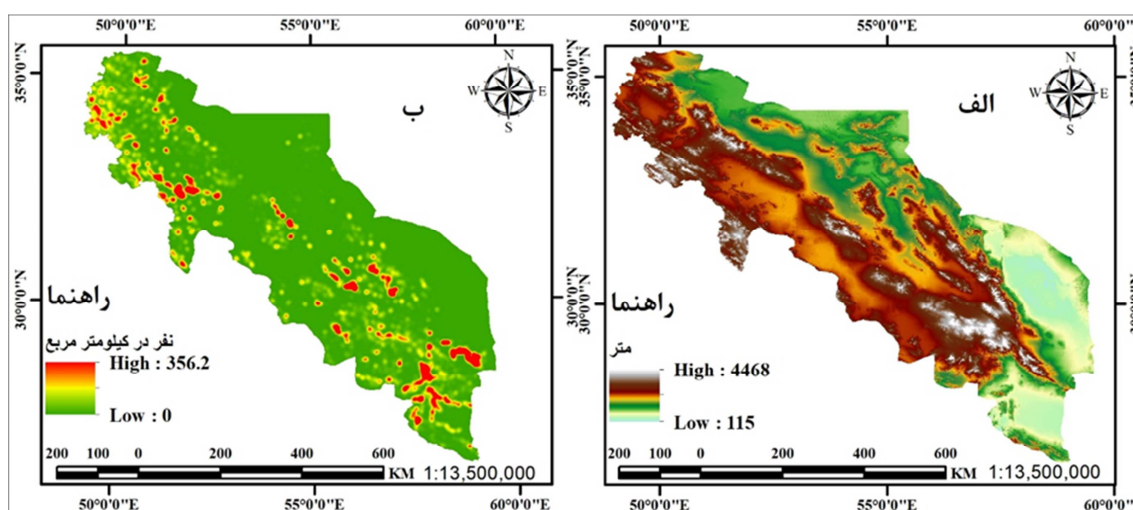
به منظور بررسی نقش توپوگرافی در بیابان‌زایی، داده ارتفاع از مجموعه داده GTOPO30 با قدرت تفکیک ۹۳۰ متر استفاده شد (شکل ۸-الف)، مشابه مطالعات (Mohamed, (2013; Al-Bakri et al, (2016).

داده های اثرگذار انسانی

با توجه به نقش عوامل انسانی در تشدید بیابان‌زایی، داده‌های مرتبط با جمعیت، دسترسی و شهرنشینی مورد استفاده قرار گرفتند مشابه پژوهش های Benassi et al, (2020; Yang et al, (2013; Egidi et al, (2022).

تراکم جمعیت:

لایه تراکم جمعیت روستایی بر اساس داده سرشماری ۱۳۹۵ مرکز آمار ایران و با استفاده از تابع Kernel Density تهیه شد (شکل ۸-ب).



شکل ۸: ارتفاع (الف)، تراکم جمعیت (ب)

1- Digital Elevation Model

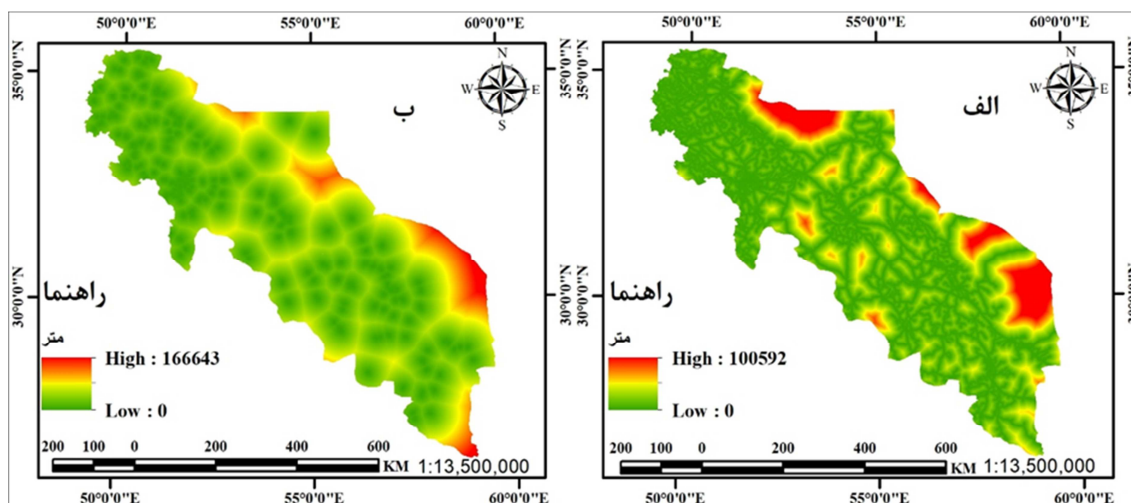
این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

شبکه راه‌ها:

داده شبکه راه‌ها از سرویس ¹overpass-turbo مرتبط با وبسایت Open Street Map استخراج شد و شامل بزرگراه‌ها، راه‌های اصلی و راه‌های درجه دوم است (شکل ۹-الف).

فاصله از شهر:

به منظور سنجش اثر شهرنشینی، فاصله از شهرهای منطقه با روش Euclidean Distance محاسبه و به‌عنوان متغیر پیش‌بینی‌کننده استفاده شد (شکل ۹-ب).



شکل ۹: فاصله از راه‌ها (الف)، فاصله از شهرها (ب)

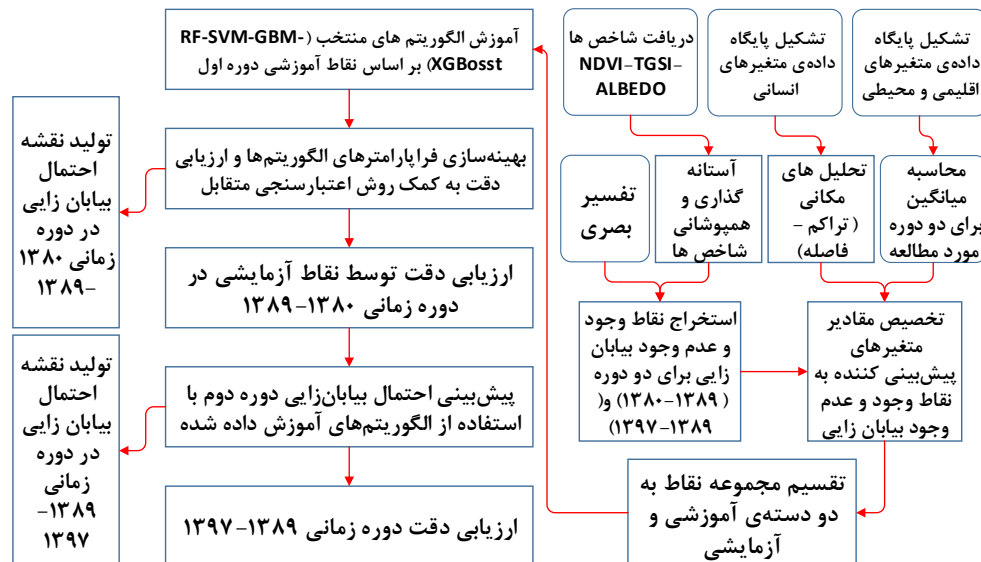
شماره مجله

روش تحقیق

شکل ۱۰ فلوچارت مراحل انجام پژوهش را نشان می‌دهد. در ادامه هر یک از مراحل تشریح می‌شود:

¹ - <https://overpass-turbo.eu/>

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.



شکل ۱۰: فلوچارت مراحل پژوهش

استخراج مناطق بیابان‌زایی

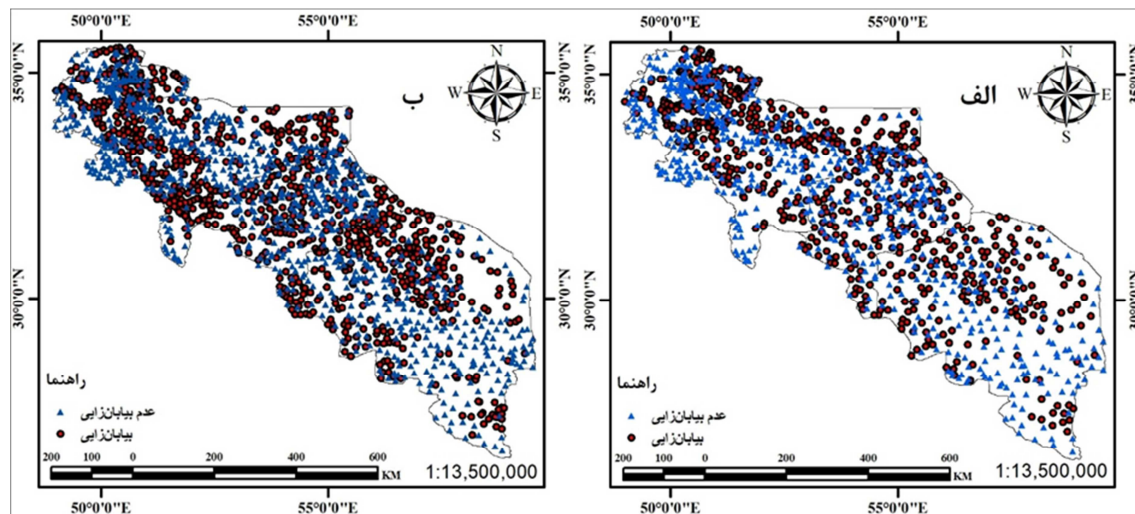
مرحله اول شامل استخراج مناطق بیابان‌زایی است که از با استفاده از سه شاخص سنجش از دوری NDVI، TGSi، ALBEDO و هم چنین تفسیر بصری تصاویر Google Earth انجام گرفت. در ابتدا میانگین سه‌ماهه این سه شاخص برای فصل تابستان در سه سال ۱۳۸۰، ۱۳۸۹ و ۱۳۹۷ محاسبه شد. انتخاب فصل تابستان به این دلیل است که در این دوره میزان بارندگی حداقل بوده و احتمال شکل‌گیری پوشش گیاهی موقت بسیار کم است. بنابراین امکان شناسایی مناطق بیابانی با دقت بیشتری فراهم می‌شود. آستانه‌گذاری برای هر شاخص به صورت جداگانه انجام شد و لایه‌هایی باینری (۰ و ۱) نشان دهنده‌ی مناطق بیابانی بر اساس هر شاخص تولید گردید. سپس سه لایه با استفاده از عملگر منطقی AND تلفیق شدند تا مناطقی که در هر سه شاخص نشانگر بیابان‌زایی بودند، استخراج شوند. مناطقی که فاقد نشانه‌های بیابان‌زایی در دوره اول بودند اما در دوره‌های بعدی به عنوان پهنه‌های بیابانی شناسایی شدند، به عنوان مناطق جدید بیابان‌زایی در نظر گرفته شدند.

آستانه‌های مورد استفاده برای شناسایی مناطق دچار بیابان‌زایی بر اساس مرور پیشینه پژوهش، به‌ویژه مطالعه Feng و همکاران (2022)، استخراج و سپس با نتایج مطالعات داخلی مرتبط از جمله Jafari و Abedi (2021) و مقصودی و پیره (۱۴۰۱) مقایسه شدند. در نهایت، آستانه‌هایی که بیشترین همخوانی را با الگوی مکانی مناطق بیابان‌زده داشتند، به عنوان مقادیر نهایی انتخاب شدند. بر این اساس، مناطق بیابان‌زده با آستانه‌های $NDVI > 0.1$ ، $ALBEDO < 0.2$ و $TGSi > 0.2$ و مناطق فاقد بیابان‌زایی نیز با آستانه‌های $NDVI < 0.2$ ، $ALBEDO > 0.2$ و $TGSi > 0.2$ شناسایی شدند. آستانه‌های بیابان‌زایی به جهت اطمینان از مناطقی که این پدیده رخ داده است، به گونه‌ای انتخاب شد که بیابان‌زایی شدید و خیلی شدید را آشکار سازد. دلیل این امر این است که بیابان‌زایی شدید و خیلی شدید نسبت به دیگر درجات بیابان‌زایی مشهودتر هستند و خطا در شناسایی آن‌ها کمتر است. همچنین آستانه‌ها برای مناطقی که بیابان‌زایی در آن‌ها رخ نداده است به گونه‌ای انتخاب شد که مناطقی که دچار بیابان‌زایی نیستند و یا بیابان‌زایی کم و متوسط دارند را آشکار سازد.

نمونه برداری

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

در نهایت برای دوره زمانی اول (۱۳۸۰-۱۳۸۹) تعداد ۱۲۰۶ نمونه برداشت شد که نیمی از آن‌ها نمونه‌های بیابان‌زایی و نیمی دیگر عدم بیابان‌زایی بودند. برای دوره زمانی دوم (۱۳۸۹-۱۳۹۷) نیز تعداد ۱۷۰۰ نمونه برداشت شد که نسبت آن‌ها مانند دوره اول مساوی است. شکل ۱۱ موقعیت نقاط نمونه بیابان‌زایی و عدم بیابان‌زایی برای سال‌های ۱۳۸۹ و ۱۳۹۷ را نشان می‌دهد.



شکل ۱۱: نقشه نقاط بیابان‌زایی و عدم بیابان‌زایی دوره زمانی ۱۳۸۰-۱۳۸۹ (الف)، دوره زمانی ۱۳۸۹-۱۳۹۷ (ب)

آماده سازی متغیرهای مستقل

متغیرهای مستقل در مطالعه حاضر شامل بارندگی، سرعت باد نزدیک به سطح، رطوبت هوا، رطوبت نسبی نزدیک به سطح، دمای روزانه، تبخیر از سطح خاک، ارتفاع، تراکم جمعیت، فاصله از شبکه راه‌ها و فاصله از شهرها است. داده های اقلیمی به صورت ماهانه دریافت و سپس میانگین چند ساله برای هر دوره محاسبه شد؛ دوره اول زمانی از سال ۱۳۸۰ تا ۱۳۸۹ و دوره دوم زمانی از سال ۱۳۸۹ تا سال ۱۳۹۷ می باشد. داده های تراکم جمعیت، فاصله از شبکه راه‌ها و فاصله از شهرها نیز همان گونه که پیش تر بیان شد، محاسبه شد.

انتخاب الگوریتم‌های یادگیری ماشین

الگوریتم‌های SVM، RF، GBM و XGBoost براساس نمونه های مناطق بیابان زده و غیر بیابانی آموزش داده شد. در این پژوهش الگوریتم SVM به دلیل عملکرد مناسب در مطالعات مرتبط با طبقه‌بندی و رگرسیون (Dabija et al., 2021؛ Nhaila et al., 2022؛ Samantaray et al., 2023؛ Tehrany et al., 2015) مورد استفاده قرار گرفت. الگوریتم RF نیز که از الگوریتم‌های یادگیری گروهی^۱ و مبتنی بر تجمع بوت استرپ^۲ است، به دلیل دقت بالاتر در مقایسه با بسیاری از الگوریتم‌های یادگیری ماشین مورد استفاده قرار گرفت. انتخاب این الگوریتم بر اساس نتایج مطالعات متعدد نظیر تحقیقات (Chen et al., 2017)، (Park & Kim, 2019) و (Sahin et al., 2018) بوده است. الگوریتم‌های GBM و XGBoost که جزء الگوریتم‌های یادگیری گروهی مبتنی بر تقویت^۳ قرار می گیرند، در سال های اخیرا بسیار مورد توجه محققین بوده‌اند (Bao et al., 2022; Chen et al., 2022; Sahin, 2020).

¹ - Ensemble Learning

² - Bootstrap Aggregating

³- Boosting

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

الگوریتم‌ها به دلیل متمرکز کردن آموزش بر روی داده‌هایی که به‌اشتباه پیش‌بینی یا طبقه‌بندی شده‌اند، دقت بالایی را ارائه می‌دهند و تحقیقاتی نظیر (Kavzoglu & Teke (2022)، Sahin (2020) و Wei et al, (2022) (نیز به برتری الگوریتم XGBoost نسبت به سایر الگوریتم‌های یادگیری ماشین اشاره کرده‌اند.

بهینه‌سازی فرایارامترها

به‌منظور انتخاب الگوریتم بهینه برای مدل‌سازی پدیده بیابان‌زایی، نیاز است الگوریتم‌های مورد بررسی در بهترین وضعیت عملکردی ارزیابی شوند. این امر از طریق تنظیم فرایارامترهای الگوریتم‌های یادگیری ماشین قابل دستیابی است. تنظیم مناسب فرایارامترها موجب بهبود دقت، جلوگیری از بیش‌برازش و افزایش پایداری مدل می‌شود. پس از تنظیم فرایارامترها، عملکرد الگوریتم‌ها با استفاده از شاخص‌های ارزیابی و مقایسه میان آن‌ها، الگوریتم بهینه انتخاب شد (Schratz et al., 2019; Weerts et al., 2020; Wu et al., 2019). بهینه‌سازی فرایارامترها با استفاده از داده‌های آموزشی دوره اول زمانی (۱۳۸۰-۱۳۸۹) و از طریق روش اعتبارسنجی متقابل^۱ با ۵-Fold و تکرار ۲۰ و جستجوی شبکه^۲ ای انجام شد.

برای الگوریتم‌های RF، SVM و GBM، به دلیل محدود بودن تعداد فرایارامترها، دامنه‌ای از مقادیر اولیه فرایارامترها تعیین و مدل‌ها آموزش داده شدند. در مقابل، با توجه به تعداد زیاد فرایارامترهای الگوریتم XGBoost، فرایند بهینه‌سازی آنها، به صورت مرحله‌ای^۳ انجام شد، به گونه‌ای که در هر مرحله تنها دو فرایارامتر بهینه‌سازی شدند و این روند تا تنظیم کامل فرایارامترها ادامه یافت.

ارزیابی دقت

برای ارزیابی دقت الگوریتم‌های مورد مطالعه، ۲۵ درصد از تعداد کل نقاط نمونه در دوره اول به صورت تصادفی انتخاب و به‌عنوان داده‌های آزمایشی در نظر گرفته شد. به منظور سنجش عملکرد مدل‌ها، از شاخص مساحت زیر منحنی ROC^۴ (AUC) استفاده شد. مقدار AUC برابر با ۱ بیانگر عملکرد کاملاً دقیق مدل بوده، و مقدار ۰/۵ بیانگر عملکرد ضعیف است (Chicco & Jurman, 2023; Hong et al., 2015; Nguyen et al., 2019).

تولید نقشه احتمال بیابان‌زایی برای دوره‌های زمانی

پس از تنظیم فرایارامترهای الگوریتم‌های مورد مطالعه، نقشه‌های احتمال بیابان‌زایی برای دوره اول زمانی (۱۳۸۰-۱۳۸۹) تولید شد. این نقشه‌ها در بخش یافته‌ها ارائه و تحلیل شده‌اند.

جهت ارزیابی پایداری زمانی عملکرد الگوریتم‌ها در مدل‌سازی بیابان‌زایی، نقشه‌های احتمال بیابان‌زایی برای دوره زمانی ۱۳۸۹-۱۳۹۷ نیز تولید و با استفاده از نمونه‌های آزمایشی به ارزیابی آنها با استفاده از شاخص AUC پرداخته شد. با انجام این مرحله می‌توان علاوه بر اعتبار سنجی مکانی که در دوره اول انجام شد، به اعتبار سنجی مکانی-زمانی الگوریتم‌ها با استفاده از داده‌های زمان آینده (در این تحقیق دوره زمانی ۱۳۸۹-۱۳۹۷) نیز پرداخت.

تعیین اهمیت متغیرهای مورد استفاده در بیابان‌زایی

1 - Cross-validation

2 - Grid Search

3 - Receiver Operating Characteristic

4 - Area Under the Curve

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

در مرحله نهایی پژوهش، اهمیت متغیرهای تأثیرگذار در مدل‌سازی بیابان‌زایی در منطقه مطالعاتی مورد ارزیابی قرار گرفت. هدف از این تحلیل، شناسایی عواملی است که بر اساس الگوریتم آموزش‌دیده، بیشترین نقش را در تبیین و پیش‌بینی بیابان‌زایی ایفا می‌کنند. پس از انتخاب الگوریتم بهینه برای مدل‌سازی، اهمیت متغیرها با استفاده از معیار کاهش میانگین ناخالصی^۱ (MDI) پرداخته شد.

نتایج و بحث

مقادیر بهینه فرآپارامترهای الگوریتم‌های یادگیری ماشین

جدول ۱ مقادیر بهینه فرآپارامترهای الگوریتم‌های یادگیری ماشین مورد استفاده در این پژوهش را نشان می‌دهد. این مقادیر بر اساس فرایند آموزش مدل‌ها و با استفاده از نمونه‌های استخراج شده از منطقه مطالعه تعیین شده‌اند.

جدول ۱: مقدار بهینه فرآپارامترهای الگوریتم‌های یادگیری ماشین

الگوریتم	فرآپارامتر	مقدار بهینه
RF	mtry	۹
	min.node.size	۱
	splitrule	extratrees
SVM	C	2^{+5}
	sigma	2^{-4}
	interaction.depth	۱۹
GBM	n.trees	۱۶۰۰
	shrinkage	۰/۰۱۲۵
	rf.minobsinnode	۳
XGBoost	nrounds	۹۰۰
	max_depth	۵
	eta	۰/۰۱۵
	gamma	۰
	colsample_bytree	۱
	min_child_weight	۱
	Subsample	۰/۷۵

ارزیابی دقت الگوریتم‌های یادگیری ماشین (دوره زمانی ۱۳۸۰-۱۳۸۹)

جدول ۲ نتایج ارزیابی دقت الگوریتم‌های یادگیری ماشین را بر اساس مقادیر AUC در دوره اول زمانی (۱۳۸۰-۱۳۸۹) ارائه می‌دهد.

جدول ۲: مقادیر شاخص AUC الگوریتم‌های یادگیری ماشین (۱۳۸۹-۱۳۸۰)

الگوریتم	مقدار AUC (نمونه‌های آموزشی)	مقدار AUC (نمونه‌های آزمایشی)
RF	۰/۸۷۷	۰/۸۷۹
SVM	۰/۸۷۷	۰/۸۷۲
GBM	۰/۸۷۵	۰/۸۷۲

^۱ - Mean Decrease Impurity

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

۰/۸۶۱	۰/۸۷۳	XGBoost
-------	-------	---------

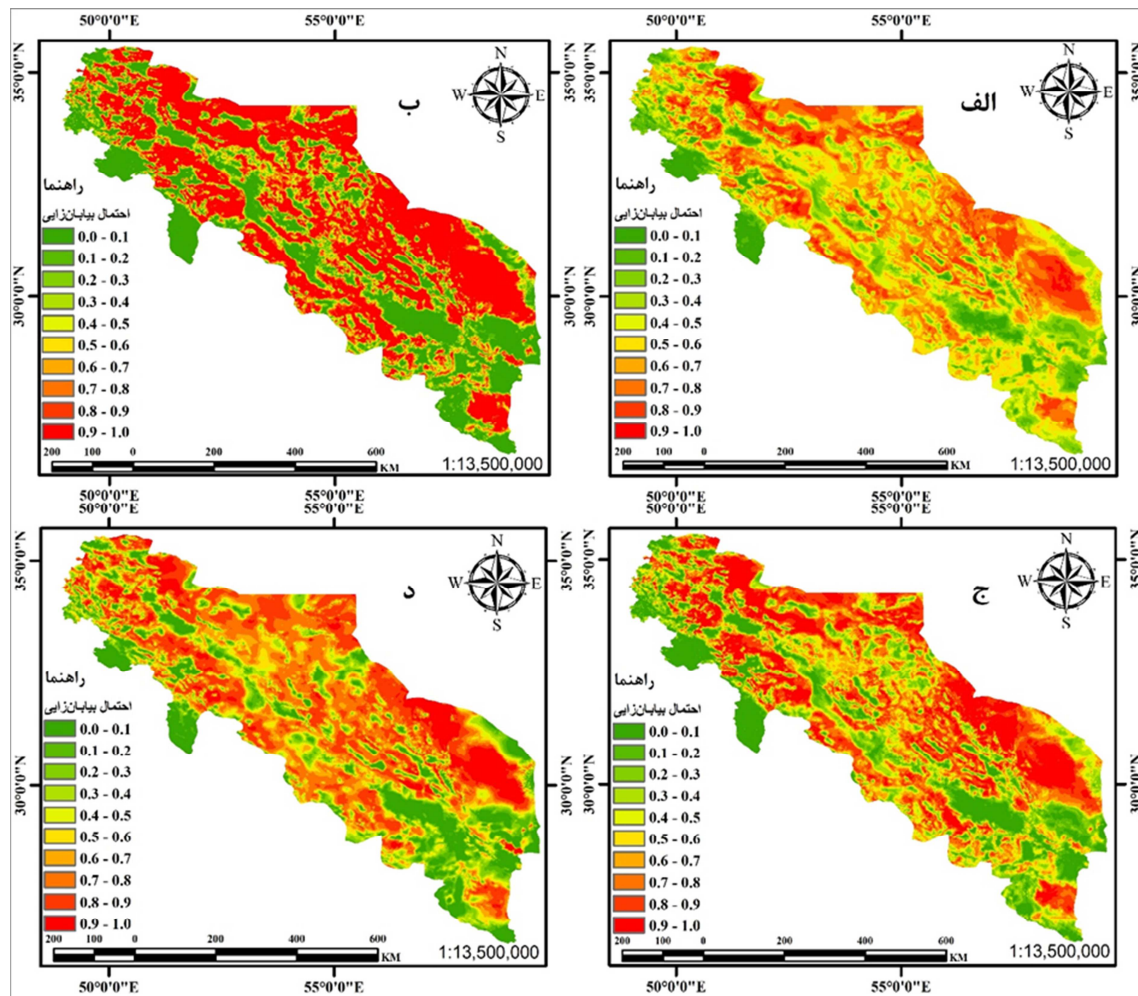
نتایج ارزیابی مدل‌ها نشان می‌دهد که الگوریتم RF در هر دو مجموعه داده آموزشی و آزمایشی عملکرد بهتری نسبت به سایر الگوریتم‌های مورد بررسی دارد. پس از آن، الگوریتم SVM در بخش داده‌های آموزشی دقت بالاتری ارائه می‌دهد، اما در ارزیابی داده‌های آزمایشی، عملکرد آن مشابه با الگوریتم GBM است. در مقابل، الگوریتم XGBoost پایین‌ترین دقت را در طبقه بندی داده‌های آزمایشی نشان داد. به طور کلی، نتایج حاکی از آن است که الگوریتم RF نسبت به سایر الگوریتم‌های یادگیری ماشین، گزینه مناسب‌تری برای مدل‌سازی بیابان‌زایی در منطقه مورد مطالعه طی دوره اول زمانی می‌باشد.

عملکرد بهتر الگوریتم جنگل تصادفی می‌تواند به دلیل آموزش گروهی درختان باشد که در نهایت دقت بالاتری نسبت به درختانی دارند که به صورت جداگانه و تکی آموزش می‌بینند. این الگوریتم در فرایند ساخت درخت، با انتخاب داده‌ها با قابلیت جایگزینی و تغییر تصادفی متغیر پیش‌بینی‌کننده، تنوع را در آموزش درختان بیشتر می‌کند و در نتیجه جنگل تصادفی خطای تعمیم بسیار کمتری ایجاد کرده و با افزایش تعداد درختان، خطای تعمیم کاهش می‌یابد. افزون بر این، جنگل تصادفی برخلاف الگوریتم‌هایی نظیر SVM یا XGBoost، به تنظیم دقیق ابرپارامترها وابستگی محدود داشته و قادر است تعاملات پیچیده و غیرخطی میان متغیرهای اقلیمی و انسانی را بدون نیاز به مفروضات اولیه استخراج کند. الگوریتم RF حساسیت کمتری به بیش‌برازش داشته و توانایی خوبی در مدیریت داده‌های پرت و مقادیر گمشده دارد (Saha et al., 2020).

نقشه احتمال بیابان‌زایی دوره اول زمانی (۱۳۸۰-۱۳۸۹)

پس از آموزش الگوریتم‌های یادگیری ماشین و بهینه‌سازی مقدار فراپارامترهای الگوریتم‌ها، به تولید نقشه احتمال بیابان‌زایی منطقه مورد مطالعه برای دوره زمانی اول (۱۳۸۰-۱۳۸۹) به کمک متغیرهای پیش‌بینی‌کننده پرداخته شد. نقشه‌های خروجی مبتنی بر احتمال بیابان‌زایی و بین مقادیر صفر و یک قرار دارند. شکل ۱۲ احتمال بیابان‌زایی برای دوره زمانی ۱۳۸۰-۱۳۸۹ را نمایش می‌دهد.

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.



شکل ۱۲: نقشه احتمال بیابان‌زایی برای دوره زمانی ۱۳۸۰-۱۳۸۹، الگوریتم RF (الف)، الگوریتم GBM (ب)، الگوریتم XGBoost (ج) و الگوریتم SVM (د)

نقشه‌های تولید شده نشان می‌دهند که مناطق دارای شرایط بیابانی با احتمال بیش از ۸۰ درصد در معرض بیابان‌زایی قرار دارند. این موضوع را می‌توان یکی از نقاط قوت الگوریتم‌های مورد استفاده دانست، زیرا این نواحی از نظر ویژگی محیطی شباهت زیادی به شرایط واقعی بیابان‌زایی داشته و مدل‌ها به درستی این الگو را شناسایی کرده‌اند. در مقابل، مناطق مجاور رشته‌کوه زاگرس به دلیل شرایط اقلیمی و توپوگرافی مناسب‌تر، احتمال کمتری از وقوع بیابان‌زایی را نشان می‌دهند.

بررسی نقشه‌های خروجی نشان می‌دهد که الگوریتم RF تغییرات احتمال بیابان‌زایی را به‌صورت تدریجی و پیوسته نمایش می‌دهد. در حالی که سایر الگوریتم‌های یادگیری ماشین، الگوی تغییرات شدیدتر و نا پیوسته‌تر را نشان می‌دهند، به‌گونه‌ای که در فاصله‌های کوتاه، اختلاف قابل‌توجهی بین نواحی با احتمال بالا و پایین مشاهده می‌شود. این رفتار در الگوریتم GBM بارزتر است، به‌طوری‌که گستره وسیعی از منطقه توسط کلاس‌های «احتمال خیلی زیاد» و «خیلی کم» پوشش یافته و بخش مربوط به احتمال بیابان‌زایی متوسط بسیار محدودتر از سایر مدل‌ها ظاهر شده است. الگوریتم‌های SVM و XGBoost نیز تغییرات را تا حدی تدریجی نشان می‌دهند، اما شدت این تغییرات

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

تدریجی کمتر از الگوریتم RF است. دلیل این تفاوت را می توان در مقاومت بالاتر الگوریتم RF در برابر بیش برآزش دانست که موجب می شود نتایج آن برای نمونه های آزمایشی از دقت پایداری بیشتری برخوردار باشد.

نتایج انحراف معیار نقشه های احتمال (جدول ۳) نیز مطالب فوق را تایید می کنند. بطوریکه الگوریتم RF کمترین میزان انحراف معیار و در نتیجه پایدارترین تخمین احتمال را داشته است. پس از آن الگوریتم های SVM و XGBoost قرار می گیرند. در مقابل الگوریتم GBM بیشترین انحراف معیار و به تبع آن، بیشترین پراکندگی مقادیر احتمال را در منطقه نشان می دهد.

جدول ۳: انحراف معیار مقادیر احتمال نقشه خروجی الگوریتم های یادگیری ماشین

الگوریتم	انحراف معیار
RF	۰/۲۳۰
SVM	۰/۲۹۳
XGBoost	۰/۳۰۶
GBM	۰/۳۹۳

ارزیابی دقت الگوریتم های یادگیری ماشین (دوره زمانی ۱۳۸۹-۱۳۹۷)

با توجه به توضیحات پیش گفته، الگوریتم های یادگیری ماشین که بر اساس داده های دوره اول (۱۳۸۰-۱۳۸۹) آموزش داده شدند، برای پیش بینی احتمال بیابان زایی در دوره دوم (۱۳۸۹-۱۳۹۷) به کار گرفته شدند. پس از تولید نقشه های احتمال بیابان زایی برای این دوره، ارزیابی دقت مدل ها انجام گرفت. جدول ۴ نتایج این ارزیابی را بر اساس مقادیر AUC برای الگوریتم های مختلف یادگیری ماشین در دوره دوم زمانی (۱۳۸۹-۱۳۹۷) ارائه می دهد.

جدول ۴: مقادیر شاخص AUC الگوریتم های یادگیری ماشین (دوره زمانی ۱۳۸۹-۱۳۹۷)

الگوریتم	مقدار AUC
RF	۰/۸۰۰
SVM	۰/۵۶۴
GBM	۰/۷۸۰
XGBoost	۰/۷۷۰

نتایج حاصل از ارزیابی الگوریتم ها در دوره دوم زمانی (۱۳۸۹-۱۳۹۷) نشان می دهد که الگوریتم RF بالاترین دقت را در میان الگوریتم های یادگیری ماشین برای مدل سازی پدیده بیابان زایی ارائه کرده است. این موضوع بیانگر آن است که، علی رغم تغییر در گستره و الگوهای بیابان زایی و در نتیجه تغییر در نقاط نمونه دوره دوم، الگوریتم RF از قدرت تعمیم پذیری بیشتری برخوردار بوده و توانسته است الگوی بیابان زایی را در طول زمان با دقت بالاتری پیش بینی کند. این ویژگی نشان دهنده قابلیت مناسب RF برای پیش بینی روند بیابان زایی در دوره های آتی است. الگوریتم GBM نیز بعد از RF بالاترین دقت را ارائه می دهد. همچنین الگوریتم XGBoost با مقدار AUC برابر با ۰/۷۷۰ با اختلاف ناچیز نسبت به الگوریتم GBM در رتبه سوم قرار دارد. در این میان الگوریتم SVM در مقایسه با سایر الگوریتم ها دقت پایین تری را ارائه می دهد. با تنظیم های پارامترهای الگوریتم SVM در دوره اول زمانی دقت به حدود ۸۹ درصد می رسد؛ با این حال دقت پیش بینی در دوره دوم به حدود ۵۰ درصد می رسد که نشان دهنده ی

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

عدم قابلیت تعمیم این الگوریتم در طول زمان است. ضمن اینکه، دقت پیش‌بینی مدل‌ها در دوره دوم که عمدتاً در محدوده نزدیک به ۸۰ درصد قرار دارند، نشان می‌دهد که تأثیر متغیرهای پیش‌بینی‌کننده نسبت به دوره اول کمی کاهش یافته است. کاهش دقت الگوریتم‌ها در دوره دوم می‌تواند نشان‌دهنده تغییر روابط میان متغیرهای مستقل و وابسته و افزایش نقش سایر عوامل مؤثر بر بیابان‌زایی در منطقه مورد مطالعه باشد.

همچنین ارزیابی دقت الگوریتم‌های یادگیری ماشین به صورت جداگانه برای هر یک از استان‌های منطقه مورد مطالعه براساس مقادیر AUC انجام گردید (جدول ۵). نتایج نشان داد الگوریتم RF در استان‌های قم، کرمان، مرکزی و یزد بالاترین مقدار AUC را به دست آورده است، در حالی که در استان اصفهان عملکرد آن ضعیف‌تر بود. در مقابل، الگوریتم GBM برای استان اصفهان بهترین عملکرد را نشان داد و پس از آن الگوریتم XGBoost نیز نتایج قابل قبولی ارائه کرد. الگوریتم SVM نیز در تمام استان‌ها کمترین مقدار AUC را به دست آورد.

جدول ۵: مقادیر شاخص AUC الگوریتم‌های یادگیری ماشین (۱۳۸۹-۱۳۹۷) به تفکیک استان‌های مورد مطالعه

الگوریتم	اصفهان	قم	کرمان	مرکزی	یزد
RF	۰/۸۰۴	۰/۶۵۵	۰/۸۱۸	۰/۸۹۸	۰/۷۹۸
SVM	۰/۵۲	۰/۴۹۱	۰/۵۷۸	۰/۵۲۹	۰/۵۳۹
GBM	۰/۸۱۱	۰/۶۰۷	۰/۷۹۲	۰/۸۹۱	۰/۷۴۶
XGBoost	۰/۸۰۶	۰/۶۱۴	۰/۸۰۱	۰/۸۸۶	۰/۷۴۰

به دلیل شباهت در هدف (مدل سازی بیابان‌زایی) و رویکرد (استفاده از الگوریتم‌های یادگیری ماشین)، نتایج بدست آمده در مطالعه حاضر با مطالعه Boali و همکاران (۲۰۲۴) مقایسه گردید. مقایسه نتایج نشان می‌دهد در مطالعه Boali و همکاران (۲۰۲۴)، دقت الگوریتم SVM در دوره زمانی اول بالاتر بوده است و این در حالیست که در دوره زمانی دوم الگوریتم RF بالاترین دقت را دارا بوده است. بنابراین نتایج حاصل در دوره دوم با نتایج تحقیق حاضر مشابه بوده است، اما با توجه به اینکه در مطالعه فوق اشاره ای به بحث پایداری زمانی نشده است، لذا نمی‌توان در خصوص مقایسه پایداری زمانی الگوریتم‌ها در مدل سازی بیابان‌زایی اظهار نظر نمود.

نقشه احتمال بیابان‌زایی دوره دوم زمانی (۱۳۸۹-۱۳۹۷)

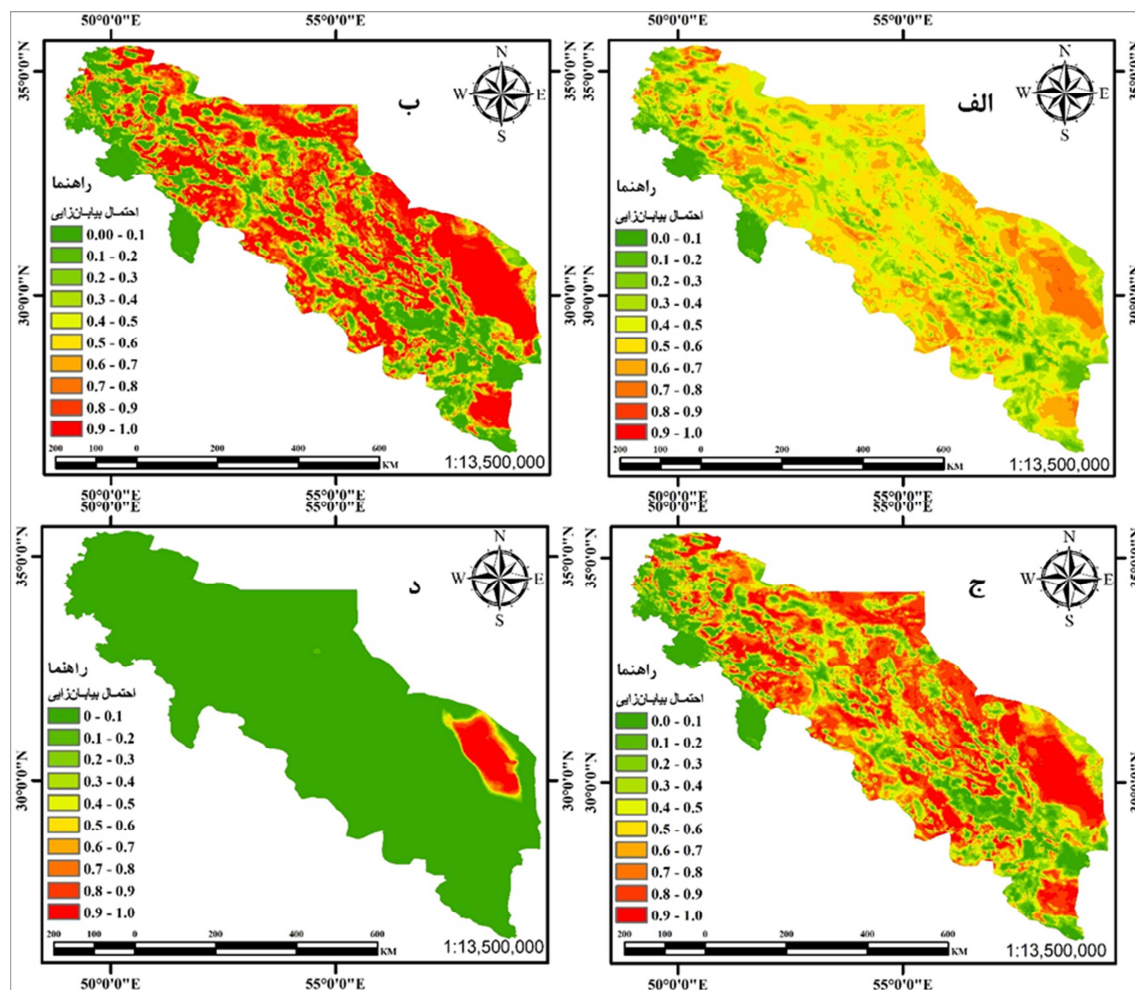
شکل ۱۳ نقشه‌های حاصل از پیش‌بینی احتمال بیابان‌زایی برای دوره دوم زمانی (۱۳۸۹-۱۳۹۷) را نمایش می‌دهد. بررسی این نقشه‌ها نشان می‌دهد که الگوریتم RF، نواحی با احتمال بالای بیابان‌زایی را در قسمت‌هایی از منطقه مورد مطالعه شناسایی کرده است. این مناطق شامل دشت لوت، بخش‌های غربی و جنوبی استان کرمان و همچنین بخش‌هایی از استان‌های مرکزی، قم و اصفهان است. این نکته می‌تواند نشان‌دهنده عملکرد مناسب الگوریتم RF در تولید نقشه احتمال بیابان‌زایی باشد.

در الگوریتم GBM نیز دشت لوت به عنوان ناحیه‌ای با احتمال بسیار بالا شناسایی شده، با این تفاوت که نسبت به الگوریتم RF مناطقی را که در حاشیه دشت لوت قرار دارند را نیز با احتمال بالا پیش‌بینی کرده است. این موضوع باعث شده است که محدوده کلاس احتمال متوسط در خروجی GBM کوچک‌تر بوده و توزیع فضایی کمتر تدریجی داشته باشد. همین ویژگی می‌تواند یکی از دلایل کاهش دقت GBM نسبت به RF باشد. الگوریتم XGBoost نیز مناطقی مانند دشت لوت را با احتمال بالا پیش‌بینی کرده است. در مقابل، بررسی نقشه تولیدشده توسط الگوریتم

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

SVM نشان می دهد که اگرچه این الگوریتم دشت لوت را به درستی تشخیص داده، ولی در سایر بخش های منطقه مورد مطالعه عملکرد مناسبی نداشته و میزان خطا در پیش بینی احتمال بیابان زایی بالاتر است.

برای تحلیل دقیق تر عملکرد الگوریتم SVM، این الگوریتم با نقاط نمونه و متغیرهای مستقل مربوط به دوره دوم نیز مجدد آموزش داده شد. در این مرحله، نیز تنظیم فرآیندها با دقت انجام شد. نتایج ارزیابی نشان داد که دقت پیش بینی این الگوریتم ۸۲ درصد رسیده است. این امر بیان می کند الگوریتم SVM به شدت نسبت به تنظیم فرآیندها حساس است به گونه ای که در صورت تغییر دوره زمانی و تغییر در الگوهای متغیر وابسته، دقت الگوریتم به طرز چشمگیری کاهش پیدا می کند.



شکل ۱۳: نقشه احتمال بیابان زایی برای دوره زمانی دوم، الگوریتم RF (الف)، الگوریتم GBM (ب)، الگوریتم XGBoost (ج) و الگوریتم

(د) SVM

تعیین اهمیت متغیرها

در این بخش، میزان اهمیت متغیرهای موثر بر فرایند بیابان زایی در منطقه مورد مطالعه بررسی می شود. از آنجا که الگوریتم RF بالاترین مقدار دقت را در بین مدل های بکار رفته ارائه داد، تحلیل اهمیت متغیرها بر مبنای خروجی این الگوریتم انجام گرفته است. نتایج نشان می دهد که به ترتیب سرعت باد، تبخیر از سطح خاک، رطوبت هوا،

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

فاصله از شهر، تراکم جمعیت، دمای هوا و فاصله از شبکه معابر مهم ترین متغیرهای اثرگذار بر بیابانزایی در منطقه هستند (جدول ۶). با توجه به اینکه سه متغیر نخست، همگی از نوع اقلیمی بوده و بیشترین نقش را در شکل گیری و تشدید بیابانزایی ایفا می کنند، بنابراین می توان نتیجه گرفت که شرایط اقلیمی خشک و نامساعد، محرک اصلی فرایند بیابانزایی در منطقه مورد مطالعه است. این یافته ها با نتایج پژوهش های پیشین مانند اکبری و همکاران (۱۳۸۶); حلبیان و جمشیدیان (۱۳۹۷); حیدری علمدارلو و همکاران (۱۴۰۱) و ضیاییان فیروزآبادی و همکاران (۱۳۹۲) نیز مطابقت دارد که در آنها حاکمیت شرایط خشک اقلیمی به عنوان عامل کلیدی در گسترش بیابانزایی شناخته می شود.

جدول ۶: میزان اهمیت متغیرها در بیابانزایی در منطقه مورد مطالعه

متغیر مورد مطالعه	میزان اهمیت متغیر
سرعت باد	۵۸/۱۳
تبخیر مستقیم از خاک	۵۶/۶۶
رطوبت هوا	۵۳/۲۹
فاصله از شهر	۳۷/۱۴
تراکم جمعیت	۳۶/۷۷
دمای هوای روزانه	۳۶/۴۵
فاصله از راه ها	۳۵/۸۵
ارتفاع	۳۵/۵۶
رطوبت نزدیک به سطح	۳۳/۸۵
بارندگی	۳۲/۳۰

بر اساس جدول فوق، سرعت باد به عنوان مهم ترین متغیر اثرگذار در مدل سازی بیابانزایی شناسایی شده است. باد یک عامل مستقل مؤثر بر خاک و سطح زمین و از مهم ترین محرک های فرسایش بادی است. فرسایش بادی یکی از عوامل اصلی تخریب خاک و گسترش بیابانزایی در بسیاری از مناطق خشک جهان به شمار می رود. این نوع فرسایش قابلیت باروری خاک را تحت تأثیر قرار داده و بر سلامت و بهره وری خاک به خصوص در زمین های خشک اثرات قابل توجهی دارد (Lackóová et al., 2023). همان گونه که نتایج پژوهش حبیبی پور و همکاران (۱۳۹۳) و Al-Bakri و همکاران (2016) نشان داده است، فرسایش بادی یکی از مهم ترین محرک های بیابانزایی است و سرعت باد از طریق افزایش شدت فرسایش بادی، نقش مؤثری در گسترش بیابانزایی در منطقه مورد مطالعه دارد (Duniway et al., 2019).

همچنین تبخیر مستقیم از سطح خاک به عنوان دومین متغیر مؤثر در مدل سازی بیابانزایی شناسایی شد. تبخیر مستقیم با کاهش رطوبت لایه های سطحی خاک، موجب ازدست دادن آب خاک و در نهایت تشدید تخریب زمین و فرایند بیابانزایی می شود. از بین رفتن پوشش گیاهی با ازدست دادن آب خاک و عدم توانایی گیاهان در استقرار مجدد همراه خواهد بود؛ بنابراین، سیستم هایی که تحت تأثیر این بازخوردها قرار می گیرند، انعطاف پذیری پایینی دارند و مستعد تغییر ناگهانی به وضعیت «بیابانی» یا تخریب هستند (D'Odorico et al., 2013). اهمیت تبخیر مستقیم از خاک در مطالعات بیابانزایی بیش از تبخیر و تعرق است، زیرا تبخیر یک اتلاف «غیرمولد آب» محسوب می شود و نقشی در فرایندهای زیستی گیاه مانند فتوسنتز و رشد ندارد، در حالی که تعرق به طور مستقیم با فتوسنتز و جذب

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

کربن مرتبط است. بر این اساس، زمانی که بیابان‌زایی موجب کاهش پوشش گیاهی و کاهش تعرق می‌شود، نسبت اتلاف غیرمولد آب (تبخیر) افزایش یافته و چرخه آب تحت تأثیر منفی قرار می‌گیرد. افزون بر این، تبخیر به انباشت نمک در سطح خاک و لایه های فوقانی خاک منجر شده و در نتیجه روند تخریب زمین را تشدید می‌کند (D'Odorico et al., 2013).

نتیجه‌گیری

مدل‌سازی مکانی یکی از رویکردهای پر کاربرد در تحلیل پدیده‌های محیطی بوده و به‌ویژه در زمینه شناسایی و پایش بیابان‌زایی، جایگاه مهمی در مطالعات نوین دارد. بیابان‌زایی به‌عنوان یکی از جدی‌ترین مخاطرات محیط زیستی، هم بر معیشت و کیفیت زندگی انسان و هم بر توان اکولوژیکی سرزمین اثرگذار است، ازاین‌رو بررسی الگوهای مکانی و عوامل مؤثر بر آن از اهمیت ویژه‌ای برخوردار است. در این پژوهش پدیده بیابان‌زایی با استفاده از الگوریتم‌های یادگیری ماشین SVM، RF، GBM و XGBoost در دو بازه‌ی زمانی مجزا مدل‌سازی شد. با توجه به مقادیر AUC محاسبه شده مشخص شد تمام الگوریتم‌ها با دقت بالای ۸۵ درصد توانسته‌اند بیابان‌زایی را در دوره زمانی اول در منطقه مورد مطالعه مدل‌سازی کنند ضمن اینکه در این دوره زمانی، الگوریتم RF در مرحله آموزش به روش اعتبارسنجی متقابل و هم در بخش ارزیابی دقت به کمک داده‌های آزمایشی بالاترین دقت را ارائه داد. همچنین با استفاده از الگوریتم‌های آموزش داده‌شده برای دوره اول، به پیش‌بینی بیابان‌زایی در دوره دوم زمانی نیز پرداخته شد که در این دوره زمانی نیز الگوریتم RF دارای بالاترین دقت بود، درحالی‌که الگوریتم SVM در پیش‌بینی بیابان‌زایی دقت پایینی داشت. مدل‌های آموزش دیده در دوره اول قادر بودند مناطق مستعد بیابان‌زایی را در دوره دوم با دقت قابل قبول شناسایی کنند، که نشان‌دهنده توان مدل‌های مکانی در پایش زمانی و پیش‌بینی روندهای آینده است. نتایج این پژوهش هم چنین نشان داد که دو الگوریتم SVM و GBM می‌توانند بعنوان الگوریتم‌های جایگزین قابل اعتماد در مدل‌سازی بیابان‌زایی مطرح شوند. این در حالیست که بر اساس همین تحقیقات، الگوریتم XGBoost نیازمند تنظیم پارامترهای بیشتر یا انتخاب ویژگی بهتر است. تحلیل اهمیت متغیرها بر اساس الگوریتم RF نیز نشان داد که سرعت باد، تبخیر از سطح خاک و رطوبت هوا مهم‌ترین عوامل مؤثر بر بیابان‌زایی در منطقه مورد مطالعه هستند. این موضوع نقش غالب عوامل اقلیمی در آغاز و گسترش بیابان‌زایی را تأیید می‌کند. نتایج این مطالعه می‌تواند به برنامه‌ریزان و مدیران محیط‌زیست در شناسایی مناطقی که بیشترین احتمال وقوع بیابان‌زایی در آن‌ها وجود دارد، کمک کند. به‌کارگیری چنین مدل‌هایی در طرح‌های مدیریت سرزمین می‌تواند مسیر مناسبی برای پیشگیری، کنترل و معکوس‌سازی روند بیابان‌زایی فراهم آورد و به تصمیم‌گیران در تدوین راهبردهای سازگار با شرایط اقلیمی و اکولوژیک منطقه یاری رساند.

منابع

آزادی، سارا؛ سعید، سلطانی؛ منیره، فرامرزی؛ علی‌رضا، سلطانی تودشکی؛ سعید، پورمنافی. (۲۰۱۵). ارزیابی شاخص خشکسالی پالمر در حوضه‌های ایران مرکزی. *علوم آب و خاک (Isfahan University of Technology)*, 19(72), 19-305. doi: 10.18869/acadpub.jstnar.19.72.26

این نسخه موقت و قبل از چاپ نهایی می باشد. لطفا برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

- اکبری، مرتضی؛ حمیدرضا، کریم زاده؛ رضا، مدرس؛ بهاره، چکشی. (۲۰۰۷). ارزیابی و طبقه بندی بیابان‌زایی با فناوری سنجش از دور و سیستم اطلاعات جغرافیایی (مطالعه موردی: منطقه خشک شمال اصفهان). *تحقیقات مرتع و بیابان ایران*، ۱۴(۲)، ۱۲۴-۴۲. https://ijrdr.areeo.ac.ir/article_105597.html
- اونق، مجید؛ نفیسه، رضائی؛ عبدالرسول، سلمان ماهینی؛ عادل، سپهر. (۲۰۲۰). ارزیابی خطر بیابان‌زایی استان خراسان شمالی با استفاده از مدل‌های IMDPA و MICD. *مهندسی اکوسیستم بیابان*، ۹(۲۶)، ۲۹-۴۲. doi: 10.22052/deej.2020.9.26.21
- بخشنده مهر، لیلا؛ سعید، سلطانی؛ عادل، سپهر. (۲۰۱۳). ارزیابی وضعیت فعلی بیابان‌زایی و اصلاح مدل مدالوس در دشت سگری اصفهان. *نشریه علمی - پژوهشی مرتع و آبخیزداری*، ۶۶(۱)، ۲۷-۴۱. doi: 10.22059/jrwm.2013.35326
- جهانشاهی، افشین؛ علیرضا، مقدم نیا؛ حسن، خسروی. (۲۰۱۵). ارزیابی شدت بیابان‌زایی با استفاده از مدل IMDPA (مطالعه موردی: دشت شهر بابک، استان کرمان). *نشریه علمی - پژوهشی مرتع و آبخیزداری*، ۶۸(۲)، ۲۴۷-۶۷. doi: 10.22059/jrwm.2015.54927
- حبیبی پور، اعظم؛ حسن، اکبری؛ علی، طالبی. (۲۰۱۴). بررسی وضعیت بیابان‌زایی با استفاده از مدل IMDPA با تأکید بر معیارهای فرسایش آبی و بادی (مطالعه موردی: منطقه بهاباد استان یزد). *جغرافیا و برنامه ریزی محیطی*، ۲۵(۲)، ۱۵۱-۶۸. https://gep.ui.ac.ir/article_18666.html
- حلیان، امیرحسین؛ زهره، جمشیدیان. (۲۰۱۹). وا کاوی تغییرات بیابان‌زایی با تأکید بر اقلیم بارش در استان قم. *مدیریت بحران*، ۷(۲)، ۱۰۵-۹۹. https://www.joem.ir/article_35226.html
- حیدری علمدارلو، اسماعیل؛ حسن، برابادی؛ پویان، دهقان رحیم آبادی؛ حسن، خسروی؛ جواد، رفیع شریف آباد. (۲۰۲۲). ارزیابی راهکارهای کنترل پدیده بیابان‌زایی با استفاده از مدل SWOT در دشت یزد- اردکان. *مدیریت جامع حوزه‌های آبخیز*، ۱۲(۱)، ۱-۱۴. doi: 10.22034/iwm.2022.550355.1025
- سیلاخوری، اسماعیل؛ مجید، اونق؛ امیر، سعدالدین؛ اسماعیل، فیله کش. (۲۰۱۴). مقایسه کارایی مدل‌های ایرانی ارزیابی خطر بیابان‌زایی IMDPA و MICD (مطالعه موردی: منطقه سبزواری). *مجله پژوهش‌های حفاظت آب و خاک*، ۲۱(۴)، ۲۸-۱. https://jwsc.gau.ac.ir/article_2023.html
- شایان، سیاوش؛ محمد، شریفی. (۱۳۸۵). مدل پیه عنوان تکنیکی در ژئومورفولوژی. *تحقیقات جغرافیایی*، ۸۰(۲۱)، ۱۰۲-۲۰. <https://www.sid.ir/paper/384320/fa>
- صالحی، اصغر؛ پریسا، کرباسی. (۲۰۲۱). نقش عوامل انسان ساخت در بیابان‌زایی شرق اصفهان. *برنامه ریزی فضایی*، ۱۱(۳)، ۲۴-۱. doi: 10.22108/sppl.2021.126373.1546
- ضیائیان فیروزآبادی، پرویز؛ لیلا، حسینجانی؛ حمیدرضا، تلخابی. (۲۰۱۴). آشکارسازی تغییرات حوضه کویر میقان با استفاده از تصاویر MSS، TM، ETM+ و داده‌های اقلیمی در دوره ۲۰۱۱-۱۹۷۳. *تحقیقات کاربردی علوم جغرافیایی*، ۱۳(۳۱)، ۱۷۳-۸۹. <http://jgs.khu.ac.ir/article-1-1683-fa.html>
- فخرآبادی، امیر؛ محمد، سلیقه؛ مهری، اکبری؛ محمد حسین، ناصرزاده. (۲۰۱۹). ارزیابی شدت بیابان‌زایی در دشت کاشان. *تحقیقات جغرافیایی*، ۳۴(۴)، ۳۸-۵۲۷. doi: 10.29252/geores.34.4.527
- فیضی کوشکی، فاطمه؛ مرتضی، اکبری؛ هادی، معماریان؛ محمود، اعظمی راد. (۲۰۱۹). شناسایی و رتبه بندی مهم‌ترین عوامل بیابان‌زایی در استان خراسان رضوی با استفاده از روش دلفی. *جغرافیا و مخاطرات محیطی*، ۸(۳)، ۲۵-۲۰۵. doi: 10.22067/geo.v0i0.84127
- کمالی مسکونی، احسان؛ محمدامین، کمالی؛ علی، خنامانی. (۱۳۹۹). بررسی و تهیه نقشه بیابان‌زایی بر اساس مدل ایرانی پتانسیل بیابان‌زایی (IMDPA) با تأکید بر دو معیار خاک و پوشش گیاهی (مطالعه موردی: فاریاب-استان کرمان). *علوم و تکنولوژی محیط زیست*، ۱۲(۱۲)، ۱۶۳. <http://sanad.iau.ir/fa/Journal/jest/Article/837862>
- گلدی زاده، فائزه؛ غلامرضا، روشن؛ عبدالعظیم، قانقرمه. (۲۰۲۳). پایش تغییرات زمانی-مکانی بیابان‌زایی در حوضه گاوخونی بر مبنای معیارهای اقلیمی. *پژوهش‌های تغییرات آب و هوایی*، ۴(۱۳)، ۵۵-۶۸. doi: 10.30488/ccr.2023.396869.1133
- مصباح‌زاده، طیبه؛ فرشاد، سلیمانی ساردو. (۲۰۱۹). پیش‌بینی بیابان‌زایی دشت کرمان با تکیه بر معیار اقلیم در سال ۱۴۰۹ با استفاده از مدل IMDPA. *ترویج و توسعه آبخیزداری*، ۶(۲۳)، ۵۰-۵۷. https://www.wmji.ir/article_254771.html
- مقصودی، مهران؛ مهین، پیره. (۱۴۰۱). شناسایی مناطق آسیب‌پذیر در مقابل بیابان‌زایی با استفاده از شاخص DVI (مطالعه موردی: استان کرمان). *پژوهش‌های فرسایش محیطی*، ۱۲(۳)، ۱-۱۷

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

میجانی، کرامت؛ رسول، مهدوی؛ حمید، غلامی؛ مرضیه، رضایی. (۲۰۲۲). پایش و مدل‌سازی شدت بیابان‌زایی با استفاده از تصاویر ماهواره لندست (مطالعه موردی حوزه آبخیز یزدان‌آباد زرنند). *نشریه علمی پژوهشی مهندسی آبیاری و آب ایران*، ۳(۲)، ۴۴۶-۴۶۲. doi: 10.22125/iwe.2022.162688

Al-Bakri, J. T., Brown, L., Gedalof, Z., Berg, A., Nickling, W., Khresat, S., Salahat, M., & Saoub, H. (2016). Modelling desertification risk in the north-west of Jordan using geospatial and remote sensing techniques. *Geomatics, Natural Hazards and Risk*, 7(2), 531-549. <https://doi.org/10.1080/19475705.2014.945102>

Bao, W., Gong, A., Zhao, Y., Chen, S., Ba, W., & He, Y. (2022). High-Precision Population Spatialization in Metropolises Based on Ensemble Learning: A Case Study of Beijing, China. *Remote Sensing*, 14(15), 3654. <https://doi.org/10.3390/rs14153654>

Benassi, F., Cividino, S., Cudlin, P., Alhuseen, A., Lamonica, G. R., & Salvati, L. (2020). Population trends and desertification risk in a Mediterranean region, 1861-2017. *Land Use Policy*, 95, 104626. <https://doi.org/10.1016/j.landusepol.2020.104626>

Boali, A., Asgari, H. R., Mohammadian Behbahani, A., Salmanmahiny, A., & Naimi, B. (2024). Remotely sensed desertification modeling using ensemble of machine learning algorithms. *Remote Sensing Applications: Society and Environment*, 34, 101149. <https://doi.org/10.1016/j.rsase.2024.101149>

Chen, T., Song, C., Fan, C., Gao, X., Liu, K., Li, Z., Cheng, J., & Zhan, P. (2022). Remote sensing modeling of environmental influences on lake fish resources by machine learning: A practice in the largest freshwater lake of China. *Frontiers in Environmental Science*, 10. <https://doi.org/10.3389/fenvs.2022.944319>

Chen, W., Pourghasemi, H. R., Zhang, S., & Wang, J. (2019). 21—A Comparative Study of Functional Data Analysis and Generalized Linear Model Data-Mining Techniques for Landslide Spatial Modeling. In H. R. Pourghasemi & C. Gokceoglu (Eds.), *Spatial Modeling in GIS and R for Earth and Environmental Sciences* (pp. 467-484). Elsevier. <https://doi.org/10.1016/B978-0-12-815226-3.00021-1>

Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D. T., Duan, Z., & Ma, J. (2017). A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *CATENA*, 151, 147-160. <https://doi.org/10.1016/j.catena.2016.11.032>

Chen, W., Zhang, S., Li, R., & Shahabi, H. (2018). Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling. *Science of The Total Environment*, 644, 1006-1018. <https://doi.org/10.1016/j.scitotenv.2018.06.389>

Chicco, D., & Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Mining*, 16(1), 4. <https://doi.org/10.1186/s13040-023-00322-4>

Dabiija, A., Kluczek, M., Zagajewski, B., Raczko, E., Kycko, M., Al-Sulttani, A. H., Tardà, A., Pineda, L., & Corbera, J. (2021). Comparison of Support Vector Machines and Random Forests for Corine Land Cover Mapping. *Remote Sensing*, 13(4), Article 4. <https://doi.org/10.3390/rs13040777>

D'Odorico, P., Bhattachan, A., Davis, K. F., Ravi, S., & Runyan, C. W. (2013). Global desertification: Drivers and feedbacks. *Advances in Water Resources*, 51, 326-344. <https://doi.org/10.1016/j.advwatres.2012.01.013>

Duniway, M. C., Pfennigwerth, A. A., Fick, S. E., Nauman, T. W., Belnap, J., & Barger, N. N. (2019). Wind erosion and dust from US drylands: A review of causes, consequences, and solutions in a changing world. *Ecosphere*, 10(3), e02650. <https://doi.org/10.1002/ecs2.2650>

Egidi, G., Quaranta, G., Salvia, R., Salvati, L., Včeláková, R., & Cudlín, P. (2022). Urban sprawl and desertification risk: Unraveling the latent nexus in a Mediterranean country. *Journal of Environmental Planning and Management*, 65(3), 441-460. <https://doi.org/10.1080/09640568.2021.1886913>

Feng, K., Wang, T., Liu, S., Kang, W., Chen, X., Guo, Z., & Zhi, Y. (2022). Monitoring Desertification Using Machine-Learning Techniques with Multiple Indicators Derived from MODIS Images in Mu Us Sandy Land, China. *Remote Sensing*, 14(11), 2663. <https://doi.org/10.3390/rs14112663>

Feng, Q., Ma, H., Jiang, X., Wang, X., & Cao, S. (2015). What Has Caused Desertification in China? *Scientific Reports*, 5(1), 15998. <https://doi.org/10.1038/srep15998>

Geist, H. J., & Lambin, E. F. (2004). Dynamic Causal Patterns of Desertification. *BioScience*, 54(9), 817-829. [https://doi.org/10.1641/0006-3568\(2004\)054\[0817:DCPOD\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2004)054[0817:DCPOD]2.0.CO;2)

Han, J., Wang, J., Chen, L., Xiang, J., Ling, Z., Li, Q., & Wang, E. (2021). Driving factors of desertification in Qaidam Basin, China: An 18-year analysis using the geographic detector model. *Ecological Indicators*, 124, 107404. <https://doi.org/10.1016/j.ecolind.2021.107404>

Hashem Geloogordi, S., Vali, A., & Sharifi, M. R. (2023). Desertification simulation using wavelet and box-jenkins time series analysis based on TGSI and albedo remote sensing indices. *Journal of Arid Environments*, 219, 105069. <https://doi.org/10.1016/j.jaridenv.2023.105069>

- Hong, H., Pradhan, B., Jebur, M. N., Bui, D. T., Xu, C., & Akgun, A. (2015). Spatial prediction of landslide hazard at the Luxi area (China) using support vector machines. *Environmental Earth Sciences*, 75(1), 40. <https://doi.org/10.1007/s12665-015-4866-9>
- Hu, Y., Han, Y., & Zhang, Y. (2020). Land desertification and its influencing factors in Kazakhstan. *Journal of Arid Environments*, 180, 104203. <https://doi.org/10.1016/j.jaridenv.2020.104203>
- Huang, J., Zhang, G., Zhang, Y., Guan, X., Wei, Y., & Guo, R. (2020). Global desertification vulnerability to climate change and human activities. *Land Degradation & Development*, 31(11), 1380–1391. <https://doi.org/10.1002/ldr.3556>
- Jafari, R., & Abedi, M. (2021). Remote sensing-based biological and nonbiological indices for evaluating desertification in Iran: Image versus field indices. *Land Degradation & Development*, 32(9), 2805–2822. <https://doi.org/10.1002/ldr.3958>
- Kavzoglu, T., & Teke, A. (2022). Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost). *Arabian Journal for Science and Engineering*, 47(6), 7367–7385. <https://doi.org/10.1007/s13369-022-06560-8>
- Lackóová, L., Kaletová, T., & Halászová, K. (2023). Are Drought and Wind Force Driving Factors of Wind Erosion Climatic Erosivity in a Changing Climate? A Case Study in a Landlocked Country in Central Europe. *Land*, 12(4), Article 4. <https://doi.org/10.3390/land12040757>
- Li, S., Zheng, Y., Luo, P., Wang, X., Li, H., & Lin, P. (2007). Desertification in western Hainan Island, China (1959 to 2003). *Land Degradation & Development*, 18(5), 473–485. <https://doi.org/10.1002/ldr.787>
- Mihi, A., Ghazela, R., & Wissal, D. (2022). Mapping potential desertification-prone areas in North-Eastern Algeria using logistic regression model, GIS, and remote sensing techniques. *Environmental Earth Sciences*, 81(15), 385. <https://doi.org/10.1007/s12665-022-10513-7>
- Mohamed, E. S. (2013). Spatial assessment of desertification in north Sinai using modified MEDLAUS model. *Arabian Journal of Geosciences*, 6(12), 4647–4659. <https://doi.org/10.1007/s12517-012-0723-2>
- Nguyen, P. T., Tuyen, T. T., Shirzadi, A., Pham, B. T., Shahabi, H., Omidvar, E., Amini, A., Entezami, H., Prakash, I., Phong, T. V., Vu, T. B., Thanh, T., Saro, L., & Bui, D. T. (2019). Development of a Novel Hybrid Intelligence Approach for Landslide Spatial Prediction. *Applied Sciences*, 9(14), Article 14. <https://doi.org/10.3390/app9142824>
- Nhaila, H., Elmaizi, A., Sarhrouni, E., & Hammouch, A. (2022). Supervised classification methods applied to airborne hyperspectral images: Comparative study using mutual information. <https://doi.org/10.1016/j.procs.2019.01.013>
- Olagunju, T. E. (2015). Drought, desertification and the Nigerian environment: A review. *Journal of Ecology and The Natural Environment*, 7(7), 196–209. <https://doi.org/10.5897/JENE2015>
- Park, S., & Kim, J. (2019). Landslide Susceptibility Mapping Based on Random Forest and Boosted Regression Tree Models, and a Comparison of Their Performance. *Applied Sciences*, 9(5), Article 5. <https://doi.org/10.3390/app9050942>
- Saha, S., Saha, M., Mukherjee, K., Arabameri, A., Ngo, P. T. T., & Paul, G. C. (2020). Predicting the deforestation probability using the binary logistic regression, random forest, ensemble rotational forest, REPTree: A case study at the Gumani River Basin, India. *Science of The Total Environment*, 730, 139197. <https://doi.org/10.1016/j.scitotenv.2020.139197>
- Sahin, E. K. (2020). Assessing the predictive capability of ensemble tree methods for landslide susceptibility mapping using XGBoost, gradient boosting machine, and random forest. *SN Applied Sciences*, 2(7), 1308. <https://doi.org/10.1007/s42452-020-3060-1>
- Sahin, E. K., Colkesen, I., & Kavzoglu, T. (2018). A comparative assessment of canonical correlation forest, random forest, rotation forest, and logistic regression methods for landslide susceptibility mapping. *Geocarto International*, 35(4), 341–363. <https://doi.org/10.1080/10106049.2018.1516248>
- Samantaray, S., Sahoo, A., & Agnihotri, A. (2023). Prediction of Flood Discharge Using Hybrid PSO-SVM Algorithm in Barak River Basin. *MethodsX*, 10, 102060. <https://doi.org/10.1016/j.mex.2023.102060>
- Schratz, P., Muenchow, J., Iturrutxa, E., Richter, J., & Brenning, A. (2019). Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data. *Ecological Modelling*, 406, 109–120. <https://doi.org/10.1016/j.ecolmodel.2019.06.002>
- Tehrany, M. S., Pradhan, B., Mansor, S., & Ahmad, N. (2015). Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. *CATENA*, 125, 91–101. <https://doi.org/10.1016/j.catena.2014.10.017>
- Weerts, H. J. P., Mueller, A. C., & Vanschoren, J. (2020). Importance of Tuning Hyperparameters of Machine Learning Algorithms (arXiv:2007.07588). arXiv. <https://doi.org/10.48550/arXiv.2007.07588>
- Wei, A., Yu, K., Dai, F., Gu, F., Zhang, W., & Liu, Y. (2022). Application of Tree-Based Ensemble Models to Landslide Susceptibility Mapping: A Comparative Study. *Sustainability*, 14(10), Article 10. <https://doi.org/10.3390/su14106330>
- Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., & Deng, S.-H. (2019). Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimization. *Journal of Electronic Science and Technology*, 17(1), 26–40. <https://doi.org/10.11989/JEST.1674-862X.80904120>
- Xiaodong, G., Jinren, N., Zhenshan, L., Ronggui, H., Xin, M., & Qing, Y. (2013). Quantifying the synergistic effect of the precipitation and land use on sandy desertification at county level: A case study in Naiman Banner, northern China. *Journal of Environmental Management*, 123, 34–41. <https://doi.org/10.1016/j.jenvman.2013.02.033>

این نسخه موقت و قبل از چاپ نهایی می‌باشد. لطفاً برای ارجاع دقت کنید این نسخه بدون تاریخ و شماره است.

Xu, D., Wang, Y., & Wang, J. (2024). A review of social-ecological system vulnerability in desertified regions: Assessment, simulation, and sustainable management. *Science of The Total Environment*, 931, 172604. <https://doi.org/10.1016/j.scitotenv.2024.172604>

Yang, Q., Jiang, Z., Ma, Z., Luo, W., Xie, Y., & Cao, J. (2013). Relationship between karst rocky desertification and its distance to roadways in a typical karst area of Southwest China. *Environmental Earth Sciences*, 70(1), 295–302. <https://doi.org/10.1007/s12665-012-2127-8>

Zhang, C., Wang, X., Li, J., & Hua, T. (2020). Identifying the effect of climate change on desertification in northern China via trend analysis of potential evapotranspiration and precipitation. *Ecological Indicators*, 112, 106141. <https://doi.org/10.1016/j.ecolind.2020.106141>.

آماده انتشار (بدون شماره مجله)