



Interpretability of Machine Learning Algorithms in Drought Susceptibility Mapping (Case Study: Khozestan Province, Iran)

Seyedeh Zeinab Shogrkhodaei¹, Amanollah Fathnia^{id}^{1*}, Seyed Vahid Razavi-Termeh^{2,3}

¹ Department of Geography, Faculty of Literature and Humanities, Razi University, Kermanshah, Iran

² Surveying of Department, Yasouj University, Yasouj, Iran

³ Researcher at Sejong University, Seoul, South Korea

ARTICLE INFO

ABSTRACT

Article History:

Received: 03 February 2026

Revised: 28 April 2026

Accepted: 30 April 2026

Available Online:

01 May 2026

Keywords:

Drought Susceptibility

Machine Learning

Random Forest (RF)

Model Interpretability /
SHAP (Shapley Additive
Explanations)

Southwest of Iran

In this study, a drought susceptibility map for Khuzestan Province, Iran, was produced using the Random Forest (RF) algorithm combined with the Shapley Additive Explanations (SHAP) interpretability approach. Drought occurrence data were derived from the Standardized Precipitation Index (SPI) for the period 2018–2022. A total of 18 environmental and climatic factors-including relative humidity, wind speed, evapotranspiration, minimum and maximum temperatures, elevation, slope, slope aspect, topographic wetness index (TWI), land cover, normalized difference vegetation index (NDVI), soil water content, river density, water level, sandy soil, soil bulk density, clay content, and soil texture-were used as input variables for modeling drought susceptibility. Model performance was evaluated using the Receiver Operating Characteristic (ROC) curve, yielding an Area Under the Curve (AUC) value of 0.987, which demonstrates the excellent predictive performance of the Random Forest algorithm. According to the drought susceptibility map, 31.03% of the study area falls in the “very low” class, 8.18% in “low”, 8.30% in “moderate”, 10.88% in “high”, and 41.58% in the “very high” susceptibility class. The highest spatial frequency was observed in the “very low” and “very high” categories. Based on SHAP results, elevation, maximum temperature, topographic wetness index, wind speed, and slope were identified as the most influential factors contributing to drought occurrence. This research highlights the advantages of interpretable machine learning approaches in drought susceptibility assessment, providing valuable insights for environmental planners and decision-makers to mitigate the adverse impacts of drought.

* Corresponding author: Amanollah Fathnia

E-mail address: a_fathnia@razi.ac.ir

How to cite this article: Shogrkhodaei, S. Z., Fathnia, A., & Razavi-Termeh, S. V. (2026). Interpretability of Machine Learning Algorithms in Drought Susceptibility Mapping (Case Study: Khozestan Province, Iran). *Journal of Geography and Environmental Hazards*, 15(2), 205-228. <https://doi.org/10.22067/geoe.2026.97359.1640>



©2026 The author(s). This is an open access article distributed under Creative Commons Attribution 4.0 International License (CC BY 4.0)

Extended Abstract

Introduction

The frequent occurrence of droughts in recent decades has posed significant challenges to the sustainability of ecosystems, agriculture, and socio-economic development. Drought is defined as a prolonged deficiency of precipitation over a specific period and represents a complex phenomenon characterized by reductions in soil moisture, declines in surface water flows, decreases in groundwater levels, and simultaneous increases in temperature. Droughts are typically classified into meteorological, agricultural, hydrological, socio-economic, and environmental types, and they exert both direct and indirect impacts, including global warming, deforestation, and urbanization. Machine learning algorithms, including SVM, BRT, RF, XGBoost, KNN, and CNN, exhibit strong capabilities in drought prediction, particularly the Random Forest (RF) algorithm, which offers high accuracy and computational efficiency. However, the “black-box” nature of these models limits physical interpretability, making the application of explainable artificial intelligence (XAI) techniques and the Shapley method essential for analyzing both the positive and negative effects of features and their interactions. Recent studies have demonstrated that integrating the Random Forest (RF) algorithm with the Shapley method enhances both the accuracy of drought prediction and the analysis of climatic, soil, topographic, and socio-economic factors. This study aims to improve the interpretability of the RF algorithm and to develop a drought sensitivity map for Khuzestan Province, with its novelty lying in the application of a spatially interpretable approach to analyze the factors influencing drought occurrence.

Material and Methods

Khuzestan Province, located in southwestern Iran, covers an area of approximately 64,236 km² with elevations ranging from 0 to 3,740 m, and experiences a climate spectrum from arid to humid. Despite its major rivers and extensive water resources, the region has been increasingly affected by environmental crises due to recurrent droughts and overexploitation of water. The development of water-intensive industries, expansion of large-scale agricultural lands, and cultivation of high-water-demand crops have intensified pressure on water resources. Additionally, declining precipitation in neighboring provinces and increased groundwater extraction have exacerbated drought conditions, impacting approximately 98.7% of the province between 2012 and 2021.

To model drought susceptibility, a total of 18 climatic, topographic, soil, and hydrological variables were considered, including relative humidity, wind speed, evapotranspiration, minimum and maximum temperatures, elevation, slope and aspect, topographic wetness index, land cover, vegetation index, soil water content, river density, water table level, sand and clay fractions, bulk density, and overall soil texture. Climatic data were derived from the annual averages of 22 meteorological stations across Khuzestan Province (2018–2022), while soil data were obtained from the USDA database via Google Earth Engine (GEE). Elevation, slope, slope direction, and watershed power index layers were generated from the SRTM digital elevation model with a resolution of 30 x 30 meters, and the vegetation index was generated from NDVI based on Landsat 8 images from 2019 to 2022. River density was calculated using the Line Density tool in ArcGIS, and the water surface layer was extracted from well water table depth data (piezometric data from the National Water Resources Management Company). The land cover layer was generated by integrating Sentinel-1 and Sentinel-2 imagery based on the 13 classes defined by Ghorban et al. (2020). Subsequently, all layers were transformed into raster format with a spatial resolution of 250 m using the Inverse Distance Weighting (IDW) method to facilitate drought modeling.

The Standardized Precipitation Index (SPI) was employed to produce the drought map, as it quantifies precipitation deficits across various time scales and is recognized as a reliable indicator for drought monitoring. Drought susceptibility was modeled using the Random Forest algorithm, which is capable of handling large, multivariate datasets while providing robust predictive performance. To enhance model interpretability, the Shapley Additive Explanations (SHAP) method was applied to assess the contribution of each input variable to drought prediction. Model accuracy was assessed using the Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) indices, while the coefficient of determination (R^2) was calculated as a dimensionless measure of model fit. The

performance of the drought susceptibility map was further evaluated using the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) index, where values approaching 1 indicate a strong capability of the model to accurately predict vulnerable areas.

Results and Discussion

In this study, the Random Forest model was employed to predict drought occurrences in Khuzestan Province based on environmental and climatic variables. Model performance was assessed using a 10-fold cross-validation approach, with 70% of the data allocated for training and the remaining 30% for testing. Based on the results, the R^2 value for the model's training data was 0.97, indicating excellent model performance. The MAE value was 0.03, and the RMSE value was 0.07. Furthermore, the RMSE, MAE, and R^2 values for the model's evaluation data were 0.18, 0.10, and 0.86, respectively. The results demonstrate the model's good accuracy for drought prediction. Interpretability analysis using the Shapley method revealed that elevation, topographic wetness index, relative humidity, and wind speed were the most influential factors, with decrease precipitation exerting a particularly significant effect. The model output was generalized in ArcMap software and using the natural Breaks method, the drought sensitivity map was classified into five classes: "very low, low, medium, high, and very high." The results showed that 31.03% of the area was in the "very low" class, 8.18% in "low," 8.30% in "medium," 10.88% in "high," and 41.58% in "very high." An AUC value of 0.987 further confirmed the model's outstanding performance. These results highlight the Random Forest model as a powerful tool for drought identification, capable of capturing complex and nonlinear relationships while processing both numerical and categorical data, thereby providing effective support for environmental management and decision-making.

Conclusion

Recent drought events have inflicted substantial impacts on both the environment and human activities. In this study, a drought susceptibility map for Khuzestan Province was developed using the Random Forest model in combination with the Shapley Additive Explanations (SHAP) method for interpretability. The model demonstrated excellent predictive performance, with an AUC of 0.987. Analysis revealed that precipitation, elevation, relative humidity, the topographic wetness index, and wind speed were the most influential factors driving drought occurrence, whereas soil texture and land cover exhibited comparatively lower impacts. The most prevalent categories in the drought susceptibility map were "very low" and "very high." Areas including Andimeshk and Dezful were classified as "very low," while parts of Mahshahr and Ahvaz fell into the "very high" category. Drought susceptibility mapping serves as an effective tool for identifying vulnerable regions and informing management strategies, thereby enhancing the resilience of both ecosystems and human communities.



تفسیرپذیری الگوریتم‌های یادگیری ماشین در تهیه نقشه‌های حساسیت به خشکسالی (مطالعه موردی: استان خوزستان، ایران)

سیده زینب شگرخدایی^۱،  امان اله فتح نیا^{*}، سید وحید رضوی ترمه^{۲ و ۳}

^۱ گروه جغرافیا، دانشکده ادبیات و علوم انسانی، دانشگاه رازی، کرمانشاه، ایران

^۲ گروه نقشه برداری، دانشگاه یاسوج، یاسوج، ایران

^۳ پژوهشگر دانشگاه Sejong، سئول، کره جنوبی

اطلاعات مقاله	چکیده
تاریخچه مقاله:	در این پژوهش، با بهره‌گیری از الگوریتم جنگل تصادفی و روش تفسیرپذیری شپلی، نقشه حساسیت خشکسالی استان خوزستان تهیه شد. برای تهیه نقشه وقوع خشکسالی، از داده‌های شاخص بارش استاندارد شده در بازه زمانی ۲۰۱۸ تا ۲۰۲۲ استفاده شد. در این مطالعه از ۱۸ فاکتور (رطوبت نسبی، سرعت باد، تبخیر و تعرق، کمینه و بیشینه دما، ارتفاع، شیب، جهت شیب، شاخص رطوبت توپوگرافی، پوشش اراضی، شاخص پوشش گیاهی، محتوای آب خاک، تراکم رودخانه، سطح آب، خاک ماسه، چگالی ظاهری خاک، خاک رس، بافت خاک) به عنوان متغیرهای ورودی برای مدل استفاده شد. ارزیابی کارایی مدل با شاخص منحنی مشخصه عملکرد سیستم (ROC) نشان داد که دقت پیش‌بینی ۰/۹۸۷ بود که بیانگر عملکرد عالی مدل در نقشه خروجی مدل جنگل تصادفی بود. بر اساس نتایج حاصل از نقشه حساسیت به خشکسالی، ۳۱/۰۳ درصد منطقه مطالعاتی در طبقه "خیلی کم"، ۸/۱۸ درصد در طبقه "کم"، ۸/۳۰ درصد منطقه در طبقه "متوسط"، ۱۰/۸۸ درصد منطقه در طبقه زیاد و ۴۱/۵۸ درصد منطقه در طبقه "خیلی زیاد" قرار گرفته است. بیشترین فراوانی مشاهده شده طبقات مرتبط با دو طبقه "خیلی کم" و "خیلی زیاد" بوده است. بر اساس روش شپلی فاکتورهایی مانند ارتفاع، دمای بیشینه، شاخص رطوبت توپوگرافی، سرعت باد و شیب بیشترین تاثیر را در وقوع خشکسالی داشته‌اند. این مطالعه مزایای استفاده از روش‌های یادگیری ماشین تفسیرپذیر را نشان می‌دهد.
کلمات کلیدی:	
حساسیت به خشکسالی	
یادگیری ماشین	
جنگل تصادفی	
توضیح‌پذیری مدل / روش شپلی	
جنوبغرب ایران	

مقدمه

وقوع مکرر خشکسالی در طی چند دهه اخیر موجب بروز چالش‌های اساسی در حوزه‌های پایداری اکوسیستم‌ها، کشاورزی و توسعه اجتماعی-اقتصادی شده است (AghaKouchak et al., 2023). امروزه خشکسالی به عنوان یک مشکل جهانی آب و هوایی ظهور کرده است و زمانی اطلاق می‌شود که منطقه‌ای خاص با بارش کمتر از حد معمول برای یک دوره زمانی خاص روبرو شود (Oyarzabal et al., 2025). علاوه بر این توصیف ساده، خشکسالی می‌تواند به عنوان فرایندی پیچیده شامل کاهش رطوبت خاک، افت جریان‌های سطحی و کاهش سطح آب زیرزمینی همراه با افزایش دما توصیف شود (Parker & Gallant, 2022). از منظر گونه‌شناسی خشکسالی، خشکسالی‌ها معمولاً به خشکسالی‌های هواشناسی، کشاورزی، هیدرولوژیکی، اجتماعی-اقتصادی و زیست محیطی طبقه‌بندی می‌شوند. خشکسالی‌های هواشناسی اغلب به عنوان منبع اصلی انواع دیگر خشکسالی‌ها در نظر گرفته می‌شوند که می‌توانند به خشکسالی‌های کشاورزی و زیست‌محیطی سرایت کنند و در نهایت باعث خسارات اجتماعی-اقتصادی شوند (Xue et al., 2020). این دسته‌بندی‌ها بر اساس کمبود بارش، رواناب، رطوبت خاک و در دسترس بودن کلی آب انجام می‌گیرد (Ozupek, 2025). خشکسالی دارای اثرات مستقیم و غیرمستقیم است. اثرات مستقیم مانند گرمایش جهانی و جنگل‌زدایی است که در نتیجه تخریب ساختار خاک پس از دوره‌های طولانی خشکسالی رخ می‌دهد. شهرنشینی از اثرات غیرمستقیم خشکسالی است (Ozupek et al., 2025). ایران به دلیل قرار گرفتن در کمربند آب و هوایی خشک و نیمه‌خشک و نوسانات شدید بارش، همواره در معرض خشکسالی‌های مکرر قرار داشته است (Torkaman Pary, Rastgoo, Opp, Zeuss & Abera, 2024; Mahmoudi & Daneshmand, 2018). در دهه‌های اخیر، کشور شاهد خشکسالی‌های شدید بوده است که بارزترین آن مربوط به سال‌های ۱۳۷۸ تا ۱۳۸۰ (۱۹۹۹ تا ۲۰۰۱) است و خسارتی بالغ بر ۳/۳ میلیارد دلار به اقتصاد کشور وارد کرده و بیش از ۳۷ میلیون نفر را تحت تأثیر قرار داده است (Liu et al., 2025; Soltani, Saboohi & Yaghmaei, 2012). کاهش بارش‌ها در استان خوزستان نیز باعث شده این منطقه دچار خشکسالی شود و تولید محصولات کشاورزی، به ویژه برنج، با چالش‌های جدی مواجه گردد (Nejadrekabi, Eslamian & Zareian, 2022).

پیش‌بینی خشکسالی همواره برای اقلیم‌شناسان و هیدرولوژیست‌ها چالشی جدی بوده است. زیرا منشاء آن پیچیده بوده و در مقیاس‌های زمانی و مکانی متفاوت رخ می‌دهد (Khan et al., 2020). هوش مصنوعی در سال‌های اخیر، به دلیل توان محاسباتی بالا و دسترسی گسترده، توانسته است بر روش‌های سنتی غلبه کند و به ابزاری مؤثر در پیشگیری از بلایای طبیعی و ارائه هشدارهای اولیه تبدیل شود. این قابلیت‌ها باعث شده است که هوش مصنوعی محبوبیت قابل توجهی در میان پژوهشگران پیدا کند (Chang, 2020; Yang, Hsieh, Wang & Yeh, 2020). روش‌های مبتنی بر هوش مصنوعی در حوزه مدل‌سازی و پیش‌بینی مکانی، بدون نیاز به آگاهی از روابط فیزیکی و ریاضی میان ورودی‌ها و خروجی‌ها عمل می‌کنند و همچنین وابستگی به اندازه‌گیری‌های میدانی و تجهیزات آزمایشگاهی پر هزینه را کاهش می‌دهند (Ghorbani, Deo, Kashani, Shahabi & Ghorbani, 2019). همچنین یادگیری ماشین با رشد چشم‌گیری که در توانایی در پردازش داده‌های پیچیده و با ابعاد بالا داشته است توانسته به طور گسترده در علوم محیطی و پیش‌بینی‌های اقلیمی به کار گرفته شود (Abbes, Inoubli, Rhif & Farah, 2023). پژوهشگران از الگوریتم‌های یادگیری ماشین در ارزیابی و پیش‌بینی خشکسالی استفاده کرده‌اند. از جمله این الگوریتم‌ها می‌توان به ماشین بردار پشتیبان (SVM) (Feng, Wang, 2019; Li Liu & Yu, 2019)، درخت رگرسیون تقویتی^۲ (BRT) (Park, Im, Jang & Rhee, 2016)، جنگل تصادفی^۳ (RF) (Wang et al., 2022)، رگرسیون آداپوست^۴ (En-Nagreg et al., 2024) تقویت گرادبان شدید^۵ (XGBoost) (Zhao et al., 2022)، کی- نزدیکترین

1-Support Vector Machines

2- Boosted Regression Trees

3- Random Forest

4- AdaBoost Regressor

5- Extreme Gradient Boosting

همسایه^۱ (KNN) (Fadaei-Kermani, Barani & Ghaeini-Hessaroeeyeh, 2017)، شبکه عصبی کانولوشن^۲ (CNN) (Mokhtar et al., 2021) اشاره کرد. این مطالعات نشان دادند که روش‌های یادگیری ماشین توانایی بالایی در ارتقاء دقت و پیش‌بینی خشکسالی دارند. همچنین الگوریتم جنگل تصادفی یکی از الگوریتم‌های یادگیری ماشین است که دارای مزایای فراوانی است. برای مثال با روشی تجمعی است که با میانگین‌گیری از چندین درخت تصمیم، از بیش‌برازش جلوگیری کرده و دقت رگرسیون را افزایش می‌دهد. این الگوریتم برای تحلیل مجموعه داده‌های بزرگ، کارآمد بوده و در نتیجه برای مطالعات مرتبط با تصاویر ماهواره‌ای مناسب است (Park, Kim & Lee, 2019). علاوه بر این، مدل جنگل تصادفی سرعت مناسب و پارامترسازی سازمان‌یافته‌ای را برای فرایند تحلیل فراهم می‌کند (Feng, Lu, Yao & Yu, 2020). این الگوریتم به طور موفقیت‌آمیزی در مطالعات مرتبط با خشکسالی مورد استفاده قرار گرفته است (Chen, Li & Wang, 2012; Wang et al., 2022).

با وجود پیشرفت‌های قابل توجه در مدل‌های یادگیری ماشین در دقت و صحت پیش‌بینی خشکسالی، ماهیت جعبه سیاه این مدل‌ها را با محدودیتی جدی در تفسیر فیزیکی نتایج و شفافیت آن روبرو می‌سازد. برای رفع این مشکل از مفهوم هوش مصنوعی تفسیرپذیر^۳ (XAI) استفاده می‌شود تا به بررسی رفتار درونی مدل‌ها بپردازد و عملکرد آن‌ها را قابل فهم سازد (Dikshit, Pradhan & Santosh, 2022). پژوهشگران در مطالعات خود با استفاده از الگوریتم‌های مختلف تلاش کرده‌اند که به آشکارسازی چگونگی اثرگذاری عوامل محرک بر خروجی مدل‌های آموزش دیده بپردازند. برای مثال گیبسون و همکاران (Gibson et al., 2021) برای پیش‌بینی خشکسالی در غرب ایالات متحده از الگوریتم جنگل تصادفی استفاده کردند. آن‌ها برای یافتن میزان اهمیت ویژگی‌ها به جابه‌جایی تصادفی مقادیر آن‌ها و بررسی تأثیر آن بر عملکرد پیش‌بینی مدل ارزیابی نمودند. کای و همکاران (Cai, Shi, Zhou, Liu & Babovic, 2024) در یک مطالعه در جنوب شرقی آمریکا، با استفاده از مدل‌های حافظه طولانی کوتاه مدت^۴ (LSTM) و روش‌های گرادیان مورد انتظار^۵ (EG) و تجزیه جمع‌پذیر^۶ (AD)، نقش عوامل محرک خشکسالی آب‌های زیرزمینی را بررسی کردند. نتایج آن‌ها نشان داد عوامل دمایی در مقیاس بزرگ و عوامل بارشی در مقیاس کوچک بیشترین تأثیر را دارند و مدل‌های تفسیرپذیر عمیق می‌توانند درک بهتری از فرآیندهای هیدرولوژیکی فراهم کنند. دیمیسستر و ویلمز (De Meeste & Willems, 2024) در مطالعه‌ای که در منطقه فلاندرز بلژیک انجام دادند با بهره‌گیری از داده‌های حوضه آبخیز و مداخلات انسان به پیش‌بینی خشکسالی با روش‌های آماری و یادگیری ماشین با استفاده از چهار شاخص خشکسالی پرداختند. نتایج آن‌ها نشان داد که مدل تقویتی گرادیان شدید عملکرد بهتری دارد. همچنین بر اساس تحلیل تفسیرپذیری مدل‌ها، عوامل آبیاری و میزان مداخلات انسانی از مهم‌ترین عوامل افزایش حساسیت رودخانه‌ها به خشکسالی بودند. جدول ۱ شماری دیگر از مطالعات داخلی و خارجی که برای بررسی خشکسالی از مدل جنگل تصادفی استفاده کرده‌اند، اشاره می‌کند. با وجود مزایای این روش‌ها که در ارتقاء تفسیرپذیری مدل‌ها انجام داده‌اند، همواره محدودیت‌هایی باقی مانده است. به عبارت بهتر روش‌های تعیین اهمیت ویژگی بر پایه جنگل تصادفی قادر نیستند به‌طور دقیق نشان دهند که یک ویژگی چه نوع اثری (مثبت یا منفی) بر نتایج پیش‌بینی دارد. روش شیلی علاوه بر تشخیص اثرات مثبت و منفی متغیرها، ثبات و دقت تفسیرها را نیز بهبود می‌بخشد (Li, Jia & Wang, 2023). همچنین شیلی با محاسبه مقادیر تعامل، قادر است اثرات متقابل بین ویژگی‌ها را کمی‌سازی کند (Chen et al., 2012). شو و همکاران (Xue et al., 2025) چارچوبی مبتنی بر یادگیری ماشین و هوش مصنوعی تفسیرپذیر (XAI) برای رشته‌کوه تاپیه^۷ چین ارائه دادند. آن‌ها چهار مدل تقویت گرادیان شدید، جنگل تصادفی، حافظه طولانی کوتاه مدت و پس انتشار^۸ را برای پیش‌بینی خشکسالی کشاورزی به کار بردند.

1- k-Nearest Neighbour

2- Convolutional Neural Network

3- Explainable Artificial Intelligence

4- Long short-term memory

5- Expected Gradient

6- Additive Decoposition

7- Ta-pieh

8- Back-propagation neural network

نتایج نشان داد که روش تقویت گرادیان شدید و جنگل تصادفی دقت بالاتری نسبت به دیگر مدل‌ها دارند. همچنین، تحلیل با روش شیلی سهم عوامل اقلیمی، خاک، توپوگرافی و اجتماعی-اقتصادی را در خشکسالی پاییز ۲۰۱۹ اندازه‌گیری و تعامل میان عوامل و الگوهای مکانی را روشن ساخت. ژنگ و همکاران (Zhang et al., 2023) در مطالعه خود که در ایالات متحده انجام دادند با بهره‌گیری از مدل تقویت گرادیان شدید همراه با روش شیلی چارچوبی تبیین‌پذیر برای پیش‌بینی اثرات چندبعدی خشکسالی ارائه کردند. این مدل نسبت به روش‌های مرجع دقت بالاتری داشته و توانست پیامدهای خشکسالی را با صحت ۰/۸۸۳ در سطح ملی و ۰/۹۴۲ در سطح ایالتی پیش‌بینی کند. نتایج نشان داد شاخص‌های بارش و دما نقش اصلی در تبیین اثرات خشکسالی دارند و روابط آن‌ها بسته به نوع و مکان اثرات متغیر است.

جدول ۱- مطالعات انجام شده در زمینه خشکسالی با استفاده از مدل جنگل تصادفی در ایران و جهان

Table 1- Review of studies on Drought using the Random Forest model in Iran and other countries

نویسندگان Authors	موضوع مقاله Research Topic
Park et al. (2019)	پیش‌بینی پهنه‌های خشکسالی شدید با استفاده از الگوریتم جنگل تصادفی: بهره‌گیری از تصاویر ماهواره‌ای و داده‌های توپوگرافی Prediction of severe drought area based on random forest: Using satellite image and topography data
Wang et al. (2022)	مدل جنگل تصادفی برای خشکسالی: پایش و اعتبارسنجی خشکسالی مراتع با تکیه بر داده‌های سنجش‌ازدور چند منبعی A random forest model for drought: Monitoring and validation for grassland drought based on multi-source remote sensing data
Hussain et al. (2025)	کاربرد الگوریتم جنگل تصادفی در شناسایی مدل مناسب برای پیش‌بینی خشکسالی هواشناسی Utilizing Logistic Regression and Random Forest to Model Meteorological Drought Persistence Across Seasonal Transitions
Elbeltagi et al. (2023)	پیش‌بینی خشکسالی هواشناسی و شاخص استاندارد شده بارش با استفاده از مدل‌های جنگل تصادفی، درخت تصادفی و رگرسیون فرایند گاوسی Prediction of meteorological drought and standardized precipitation index based on the random forest (RF), random tree (RT), and Gaussian process regression (GPR) models.
Dikshit, Pradhan & Alamri (2020)	پیش‌بینی کوتاه مدت مکانی-زمانی خشکسالی با استفاده از مدل جنگل‌های تصادفی در ایالت نیوساوتولز، استرالیا Short-term spatio-temporal drought forecasting using random forests model at New South Wales, Australia.
Zarei, Mahmoudi & Moghimi, (2023)	تعیین مناسب‌ترین شاخص خشکسالی با استفاده از الگوریتم جنگل تصادفی با تأکید بر خشکسالی کشاورزی Determining the most appropriate drought index using the random forest algorithm with an emphasis on agricultural drought
Huang et al. (2025)	پیش‌بینی خطر خشکسالی هواشناسی آینده در سرزمین اصلی چین با استفاده از مدل جنگل تصادفی Predicting future meteorological drought risk in mainland China using random forest model
Li et al. (2023)	پایش سنجش از دور خشکسالی در جنوب غرب چین با استفاده از روش‌های جنگل تصادفی و تقویت گرادیان شدید Remote sensing monitoring of drought in Southwest China using random forest and eXtreme gradient boosting methods
Lotfifard, Esmaeili-Gisavandani & Adib, (2022)	پایش و پیش‌بینی خشکسالی با استفاده از شاخص‌های SPI، SPEI و مدل جنگل تصادفی در اقلیم‌های مختلف ایران Drought monitoring and prediction using SPI, SPEI, and random forest model in various climates of Iran.
Vagheei, Torabi & Ghaemi, (2021)	ارزیابی خشکسالی به کمک شاخص بارش استاندارد و الگوریتم جنگل تصادفی Drought assessment using standardized precipitation index and random forest algorithm.
Seyedi Ghaldareh, Ahmadi & Gholamnia (2023)	ارزیابی شاخص‌های سنجش از دوری پایش خشکسالی با استفاده از رگرسیون بردار پشتیبان و مدل‌های جنگل تصادفی (مطالعه موردی: شهرستان مریوان) Evaluation of remote sensing-based drought monitoring indexes using support vector regression and random forest models (Case study: Marivan city).
Azizi and Abbasi(2025)	توسعه یک روش کارآمد مبتنی بر آموزش ماشین و داده‌های زمینی و ماهواره‌ای برای پیش‌بینی و پهنه‌بندی خشکسالی کشاورزی در ایران Developing a Robust Data-Driven Model Based on Ground and Satellite Measured Data for Agricultural Drought Prediction in Iran
Hajarian (2025)	پایش و پهنه‌بندی خطر خشکسالی با استفاده از مدل جنگل تصادفی (مورد مطالعه: استان اردبیل) Meteorological Drought Risk Monitoring and Zoning Using Random Forest Model.

این مطالعه با هدف تفسیرپذیری الگوریتم یادگیری ماشین جنگل تصادفی در تهیه نقشه‌های حساسیت به خشکسالی انجام گرفته است. برای این منظور داده‌های مرتبط با مخاطره خشکسالی در استان خوزستان تهیه و جهت مدلسازی حساسیت آماده‌سازی شده‌اند. این مطالعه با رویکردی نوین برای تفسیرپذیری مدل جنگل تصادفی، از روش شیلی استفاده کرده است. نوآوری این تحقیق استفاده از الگوریتم هوش مصنوعی جنگل تصادفی در ترکیب با یک رویکرد تفسیرپذیری مکانی به منظور تفسیر و تحلیل نقشه فاکتورهای مکانی مؤثر بر وقوع خشکسالی در استان خوزستان است.

منطقه مورد مطالعه

استان خوزستان در جنوب غربی ایران (شکل ۱) بین ۲۹ درجه و ۵۷ دقیقه تا ۳۳ درجه و ۵۷ دقیقه عرض شمالی و ۴۷ درجه و ۴۰ دقیقه تا ۵۰ درجه و ۳۳ دقیقه طول شرقی واقع شده است (Jahangir & Mohammadi, 2016). خوزستان مساحتی بالغ بر ۶۴۲۳۶ کیلومتر مربع دارد و ارتفاعی بین ۰ تا ۳۷۴۰ متر متغیر است (Zarasvandi, Carranza, Moore & Rastmanesh, 2011). این استان در طول سال میزان تبخیری برابر با ۲۰۴۴ میلی‌متر و میزان بارش آن بین ۲۵۰ تا ۴۰۰ میلی‌متر است (Baloei, 2025; Kabolizadeh & Mohammadi, 2024; Piri & Ghoghghar, 2025). استان خوزستان از نظر منابع آبی و بروز مخاطرات اقلیمی جایگاه برجسته‌ای در کشور دارد. آب و هوای این منطقه بین خشک تا مرطوب متغیر است. در بخش‌های شمالی استان هوای سرد بخش‌های جنوبی هوای گرمسیری دارد (Zarasvandi et al., 2011). رودخانه‌های بزرگی مانند کارون، کرخه، هندیجان، جراحی و دز اهمیت ذخایر آبی منطقه را دو چندان کرده است. با این حال خشکسالی‌های مکرر و بهره‌برداری بی‌رویه از منابع آب، استان خوزستان را با چالش‌ها و بحران‌های جدی زیست محیطی روبه‌رو کرده است (Baloei et al., 2024; Piri & Ghoghghar, 2025). از سوی دیگر، استقرار صنایع بالادستی و پایین‌دستی مرتبط با نفت، حضور صنایع آب بر مانند فولاد، گسترده‌ی اراضی کشاورزی و تولید محصولات استراتژیک در فصل‌های مختلف سال، همراه شکل‌گیری شرکت‌های کشت و صنعت نظیر صنایع نیشکر، بر اهمیت این منطقه افزوده است (Koochak & Behnia, 2013). با جود منابع آبی در خوزستان، این استان با چالش‌های جدی کم آبی روبرو شده است. از جمله این موارد می‌توان به مدیریت اشتباه منابع آب، کاهش بارندگی، نزدیکی جغرافیایی خوزستان به خلیج فارس و مباحث حاکمیتی، تقاضای بالای آب در بخش کشاورزی، افزایش شهرنشینی و رشد صنعتی و تغییرات اقلیمی اشاره کرد (Bahiraie, Hosseini & Hossein-Panahi, 2025). همچنین کاهش بارش در استان‌های همجوار باعث کاهش ذخایر سدها و کیفیت آب سطحی شده است. این مسائل موجب رو آوردن به آب‌های زیرزمینی و افزایش بهره‌وری از آن خصوصاً برای کشاورزی شده است که موجب بروز خشکسالی در این منطقه شده است (Bahiraie et al., 2025). بر اساس گزارش سازمان هواشناسی ایران (۲۰۲۱) و شاخص استاندارد بارش-تبخیر-تعریق (SPEI) (McKee, Doeske & Kleist, 1993)، ۹۸/۷ درصد از مساحت استان خوزستان بین سال‌های ۲۰۱۲ تا ۲۰۲۱ تحت تأثیر خشکسالی قرار گرفته است. علاوه بر این، کشت گسترده محصولات پرمصرف آب مانند برنج و نیشکر که ۸/۵ درصد از مصرف آب کشاورزی استان را به خود اختصاص داده، فشار بر منابع آب را تشدید کرده است (Bahiraie et al., 2025).

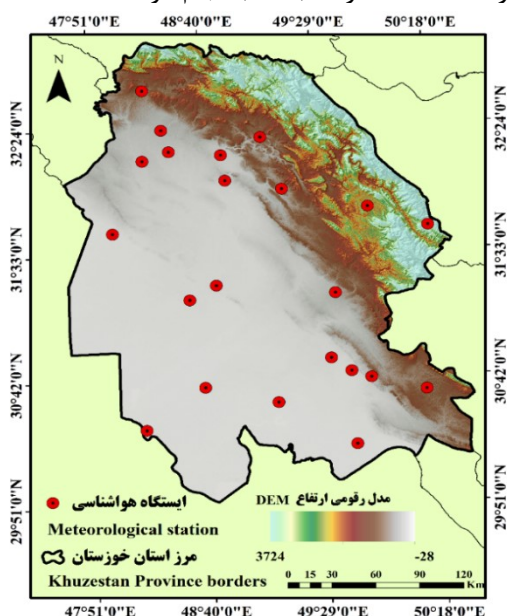
مواد و روش‌ها

این مطالعه در ۴ گام اصلی انجام گرفت (۱) مرحله اول شامل جمع‌آوری داده‌ها می‌باشد که شامل تهیه داده‌های مرتبط با معیارهای مؤثر در چهار گروه داده‌های توپوگرافی، داده‌های اقلیمی، داده‌های مرتبط با خاک و داده‌های مرتبط با پوشش زمین می‌باشد. همچنین داده‌های وقوع خشکسالی از CHIRPS^۲ تهیه شد. (۲) در مرحله دوم پیش پردازش داده‌ها انجام گرفت. برای این منظور از چند خطی بودن عوامل مؤثر استفاده شد. (۳) در مرحله بعد برای تهیه نقشه حساسیت خشکسالی (DTSM) از مدل جنگل

1- <https://www.irimo.ir/eng/index.php>

2- Climate Hazards Center InfraRed Precipitation with Station data

تصادفی استفاده شد. (۴) در مرحله آخر ارزیابی نتایج مدل‌سازی با استفاده از معیارهای ریشه میانگین مربعات خطا (RMSE)^۱، میانگین خطای مطلق (MAE)^۲ و نمودار مشخصه عملکرد (ROC)^۳ انجام گرفت.



شکل ۱- نقشه محدوده مطالعاتی

Fig.1. Study area map

تولید و آماده‌سازی نقشه‌های معیار

تولید نقشه خشکسالی

برای تولید نقشه خشکسالی هواشناسی از شاخص بارش استاندارد شده (SPI) در بازه زمانی ۲۰۱۸ تا ۲۰۲۲ استفاده شد. SPI یکی از مهمترین شاخص‌های پایش و ارزیابی خشکسالی است که براساس تعیین کمیت کسری بارش در هر مقیاس زمانی مورد استفاده قرار می‌گیرد (Rezaei, Roshni & Amirabadizadeh, 2016). این شاخص مطابق معادله ۱ محاسبه می‌شود (McKee, Doeske & Kleist, 1993):

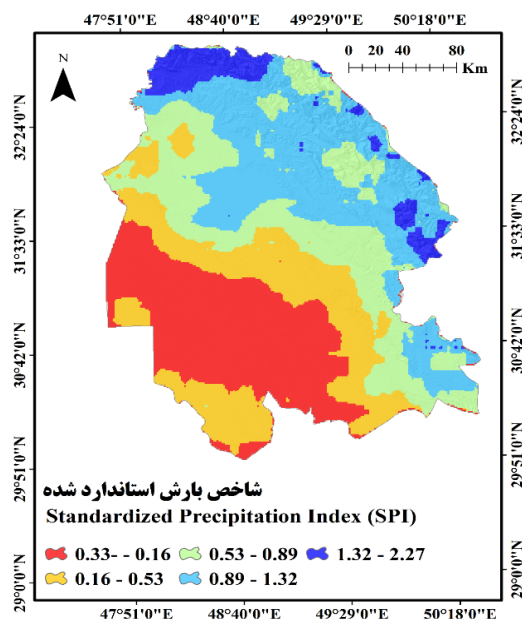
$$SPI = \frac{X_i - \bar{X}}{\sigma} \quad (1)$$

در اینجا X_i و $\bar{X}$ به ترتیب بیانگر بارش ایستگاهی و میانگین بارش است. همچنین σ انحراف استاندارد است. برای برآورد SPI از داده‌های بارش (CHIRPS) در بازه ۲۰۱۸ تا ۲۰۲۲ در GEE استفاده شد (شکل ۲). بر اساس مقادیر استاندارد مرتبط با شاخص SPI، مقادیر بین -۱ تا -۱/۴۳ در طبقه خشکسالی شدید، -۰/۵ تا -۰/۹۹ خشکسالی ملایم، ۰/۴۹ تا ۰/۹۹ تقریباً نرمال، ۰/۵ تا ۰/۹۹ خشکسالی ملایم، ۱/۴۹ تا ۱/۷۲ ترسالی متوسط، ۱/۵ تا ۱/۷۲ در طبقه ترسالی شدید و از ۲ بیشتر در طبقه ترسالی بسیار شدید قرار می‌گیرند. پس از تهیه نقشه شاخص بارش استاندارد شده مناطق با مقادیر کمتر SPI (کمتر از صفر) به صورت تصادفی به عنوان نقاط وقوع خشکسالی در نظر گرفته شده است. همچنین مقادیر بالای SPI (نقاط بالای صفر) به عنوان نقاط عدم وقوع خشکسالی در نظر گرفته شد. بنابراین ۲۰۰ نقطه وقوع خشکسالی و به صورت برابر ۲۰۰ نقطه عدم وقوع خشکسالی انتخاب شد. برای ارزیابی نسبت داده‌های وقوع به داده‌های عدم وقوع ۷۰ به ۳۰ تقسیم شد.

1- Root Mean Square Error

2- Mean Absolute Error

3- Receiver Operating Characteristic



شکل ۲- شاخص بارش استاندارد شده (SPI) برای محدوده مطالعاتی

Fig.2. Standardized Precipitation Index (SPI) for the study area

فاکتورهای مؤثر بر خشکسالی

در این مطالعه برای مدل سازی حساسیت خشکسالی از متغیرهای رطوبت نسبی، سرعت باد، تبخیر و تعرق، شیب، کمینه دما، بیشینه دما، ارتفاع، شیب، جهت شیب، شاخص رطوبت توپوگرافی، پوشش اراضی، شاخص پوشش گیاهی، محتوای آب خاک، تراکم رودخانه، سطح آب، ماسه، چگالی ظاهری خاک، خاک رس، بافت خاک استفاده شده است (شکل ۳). با توجه به اینکه شاخص SPI خود حاصل یک تبدیل غیرخطی و استانداردسازی تنها بر روی داده های بارش تجمعی است، وجود متغیر بارش خام در کنار سایر متغیرهای مؤثر در ماتریس ویژگی ها منجر به پدیده نشت داده می گردد. علاوه بر این مشخص شد که فاکتور بارش در تحلیل تست چند خطی مقادیر بالاتری نسبت به سایر پارامترها کسب کرده است. در نهایت بارش از داده های ورودی حذف شد.

برای تولید لایه های رطوبت نسبی، سرعت باد، تبخیر و تعرق، حداقل و حداکثر دما، از داده های میانگین سالانه ۲۲ ایستگاه هواشناسی استان خوزستان در بازه زمانی ۲۰۱۸ تا ۲۰۲۲ استفاده گردید. لایه های شاخص رطوبت توپوگرافی، محتوای آب خاک، خاک ماسه، چگالی ظاهری خاک، خاک رس و بافت خاک از داده های بافت خاک یواس دی^۱ (USDA) در گوگل ارث انجین^۲ (GEE) استخراج شد. برای تهیه لایه های ارتفاع، شیب و جهت شیب و شاخص توان آبراهه از یک مدل ارتفاع دیجیتال^۳ (SRTM) در رزولیشن مکانی ۳۰×۳۰ متر در گوگل ارث انجین استفاده شد. به منظور تولید لایه شاخص پوشش گیاهی، از شاخص نرمال شده تفاوت پوشش گیاهی^۴ (NDVI) بهره برداری شد. این لایه بر اساس تصاویر ماهواره ای لندست ۵^۵ در پلتفرم گوگل ارث انجین و در بازه زمانی سال ۲۰۱۸ تا ۲۰۲۲ تهیه گردید. به منظور تولید لایه تراکم رودخانه، از لایه رودخانه های استان خوزستان بهره گرفته شد و برای محاسبه میزان تراکم، ابزار تراکم خط در نرم افزار ArcGIS مورد استفاده قرار گرفت. برای تولید لایه سطح آب، از داده های مربوط به عمق سطح ایستابی چاه ها از داده های پیزومتریک که توسط شرکت مدیریت منابع آب کشور ارائه می شود، استفاده گردید. برای تهیه لایه

1- U.S. Department of Agriculture

2- Google Earth Engine

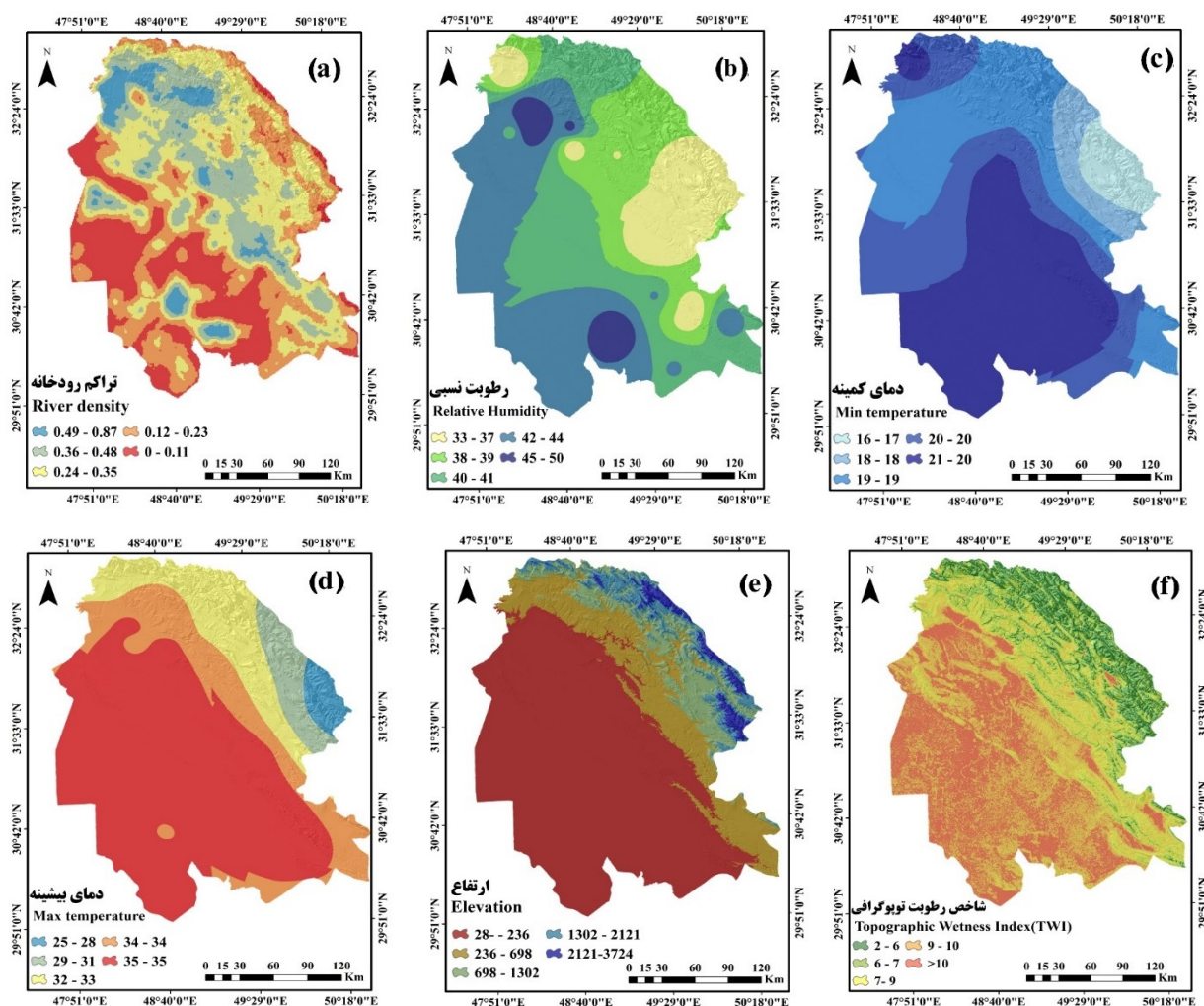
3- Shuttle radar topography mission

4- Normalized Difference Vegetation Index

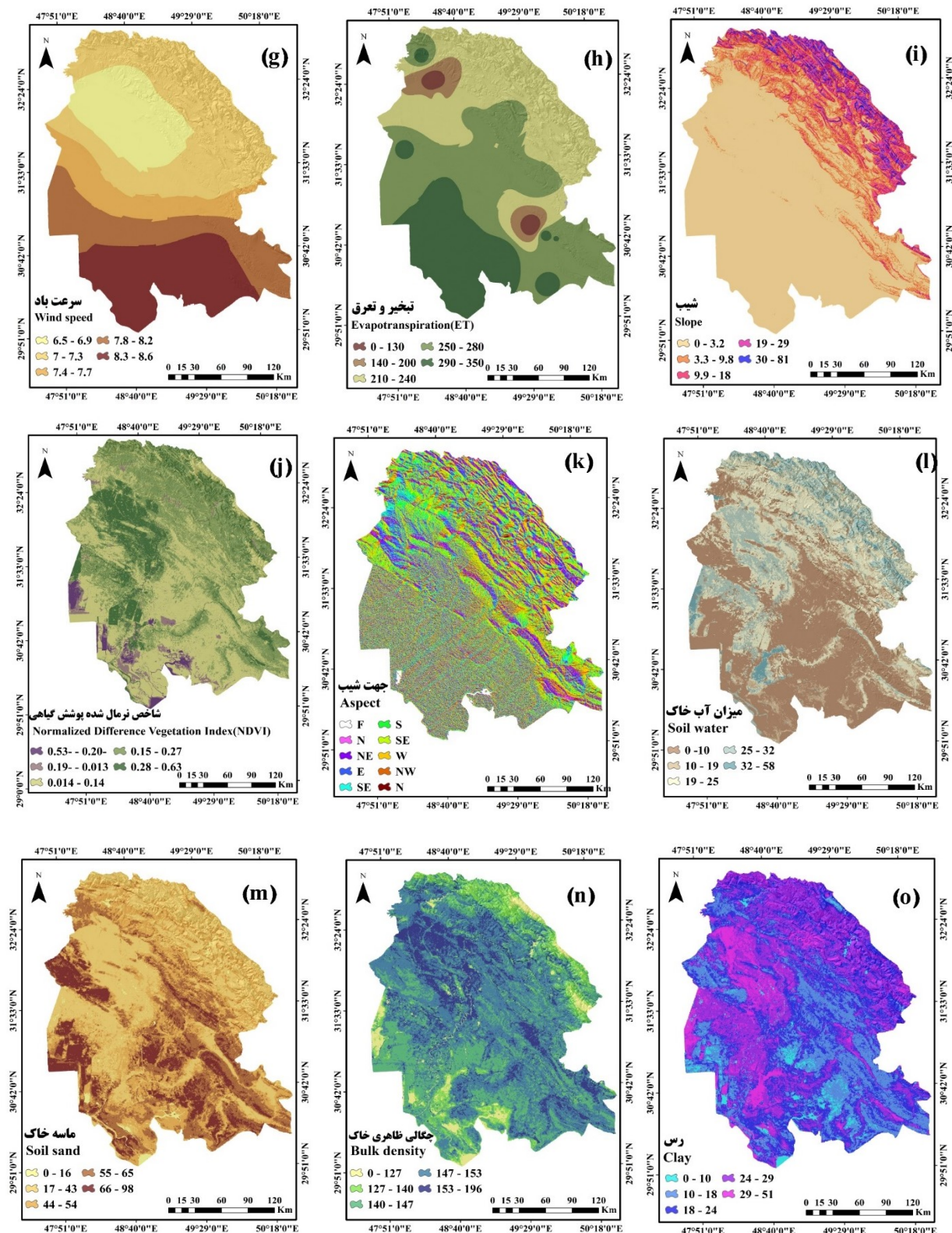
5- Landsat-8

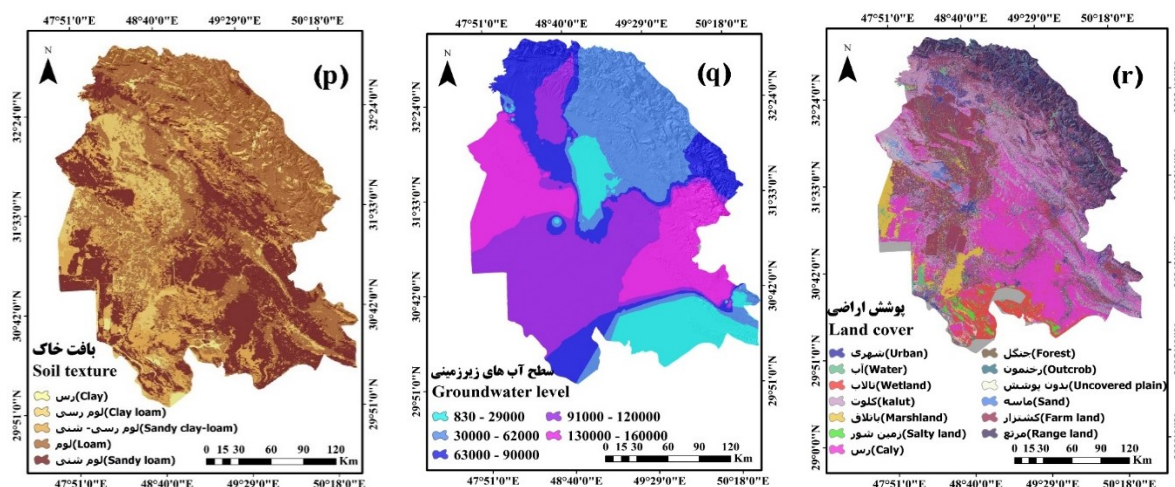
6- Line Density

پوشش اراضي از ادغام تصاویر سنتینل ۱^۱ و سنتینل ۲ که در ۱۳ کلاس توسط قربانیان و همکاران (Ghorbanian et al., 2020) در گوگل ارث انجین ارائه شده است تهیه شد. سپس تمام این لایه‌ها با به کارگیری ابزار وزن دهی معکوس فاصله^۲ (IDW) در نرم‌افزار ArcGIS به لایه‌های رستری با اندازه پیکسل ۲۵۰ متر تبدیل شدند.



- 1- Sentinel
- 2- Inverse distance weighting





شکل ۳- نقشه معیارهای مؤثر بر وقوع خشکسالی: (a) تراکم رودخانه، (b) رطوبت نسبی، (c) دمای کمینه، (d) دمای بیشینه، (e) ارتفاع، (f) TWI، (g) سرعت باد، (h) تبخیر و تعرق، (i) شیب، (j) NDVI، (k) جهت شیب، (l) میزان آب خاک، (m) ماسه خاک، (n) چگالی ظاهری خاک، (o) رس، (p) بافت خاک، (q) سطح آب‌های زیرزمینی، (r) پوشش اراضی.

Fig.3. Map of criteria affecting drought occurrence: a) River density, b) Relative humidity, c) Minimum temperature, d) Maximum temperature, e) Elevation, f) TWI, g) Wind speed, h) Evapotranspiration, i) Slope, j) NDVI, k) Aspect, l) Soil water content, m) Sand, n) Soil bulk density, o) Clay, p) Soil texture, q) Groundwater level, r) Land cover.

روش جنگل تصادفی

مدل جنگل تصادفی نخستین بار توسط بریمن (Breiman, 2001) معرفی شد و بر پایه الگوریتم درخت‌های طبقه‌بندی و رگرسیون^۱ (CART) به عنوان یکی از روش‌های درخت تصمیم بنا شده است (Li et al., 2023). مدل جنگل تصادفی ابزار قدرتمندی است که با سرعت بالا و انعطاف زیاد، روابط پیچیده و غیرخطی میان متغیرها را مدل می‌کند و همزمان داده‌های عددی و دسته‌ای را می‌تواند پردازش کند (Tramblay & Quintana Seguí, 2022). مدل جنگل تصادفی نسبت به چند هم‌خطی بین متغیرها حساسیت کمی دارد و هنگامی که داده‌ها ناقص یا نامتوازن باشند، نتایج آن نسبتاً پایدار باقی می‌ماند. با ایجاد تعداد زیادی درخت تصمیم و اعمال اصل جمع‌آوری آرای آن‌ها، طبقه‌بندی‌های مجموعه‌ای از بیش‌برازش مدل نهایی جلوگیری کرده و همزمان دقت را حفظ می‌کنند (Chen et al., 2020)، از مهم‌ترین مزایای مدل جنگل تصادفی می‌توان به جلوگیری از بیش‌برازش اشاره کرد. همچنین این مدل قادر است به طور خودکار مقادیر گمشده را مدیریت کند و تحت تأثیر داده‌های پرت قرار نمی‌گیرد (Al-Abadi, 2018). رویکرد جنگل تصادفی با بهره‌گیری از مجموعه‌ای از درخت‌های تصمیم، پیش‌بینی‌ها را انجام می‌دهد. اصطلاح «جنگل تصادفی» بیانگر استفاده از تعداد زیادی درخت تصمیم مستقل (جنگل) است که با نمونه‌گیری تصادفی از داده‌های آموزشی و انتخاب تصادفی متغیرها در هر گره ایجاد می‌شوند. این فرایند باعث کاهش مشکلات شناخته‌شده الگوریتم درخت‌های طبقه‌بندی و رگرسیون از جمله بیش‌برازش و حساسیت به داده‌های آموزشی می‌شود (Li et al., 2023). از دیگر مزیت‌های این مدل توانایی مدیریت داده‌های حجیم و همچنین قابلیت تفسیر پذیری بالا در شرایط وجود متغیرهای متعدد است. این روش علاوه بر برآورد کلی از اهمیت متغیرها، امکان بررسی اثر نهایی هر متغیر مستقل بر متغیر وابسته را فراهم می‌کند (Dikshit et al., 2020). در این روش، برای ساخت هر درخت یک زیر مجموعه تصادفی از داده‌های آموزشی انتخاب می‌شود. نتیجه نهایی نیز با ترکیب خروجی همه درخت‌ها حاصل می‌شود؛ به این

1- Classification and Regression Tree

2- Ensemble

صورت که در مسائل رگرسیون میانگین‌گیری و در مسائل طبقه‌بندی رأی‌گیری اکثریت انجام می‌گیرد. علاوه بر این، مدل جنگل تصادفی با استفاده از داده‌های خارج از کیسه^۱ و جایگزین کردن یک متغیر با مقادیر تصادفی، میزان افزایش خطای میانگین مربعات را محاسبه می‌کند (Breiman, 2001). مدل جنگل تصادفی (RF) نیاز به تنظیم سه پارامتر دارد: تعداد متغیرها (mtry)، تعداد درخت‌ها (ntree)، و حداکثر تعداد گره‌های انتهایی (nidesize). الگوریتم RF بر پایه دسته‌بندی کننده‌های درختی ساخته شده است و به صورت معادله (۲) ارائه می‌شود:

$$h(x, i_k), k = 1, 2, \dots, n \quad (2)$$

i_k نشان دهنده عوامل مؤثر بر وقوع خشکسالی است و $1, 2, \dots, n$ ردای ورودی x هستند. در مدل جنگل تصادفی، خطاهای کلی به صورت معادله (۳) تعریف می‌شوند:

$$GE = P_{x,y}(mg(x, y) < 0) \quad (3)$$

باتوجه به رابطه ۳، x و y نشان دهنده عوامل مختلف مؤثر بر وقوع خشکسالی هستند و mg تابع حاشیه را نشان می‌دهد. تابع حاشیه را می‌توان به صورت معادله (۴) توصیف کرد:

$$mg(x, y) = \alpha v_k I(h_k(x) = y) - \max_j \neq \alpha_i v_k I(h_k(x) = j) \quad (4)$$

روش تفسیرپذیری شیلی

روش شیلی که توسط لوندبرگ و لی (Lundberg & Le, 2017) معرفی شد، بر پایه مقادیر شیلی در نظریه بازی‌های تعاونی بنا شده و چارچوبی منسجم برای اندازه‌گیری سهم هر ویژگی در خروجی مدل فراهم می‌کند (Xue et al., 2025). این الگوریتم قادر است میزان اهمیت هر متغیر ورودی را در خروجی مدل‌های یادگیری ماشین مشخص کند. با شناسایی و درک نقش این متغیرها، می‌توان از دانش تجربی یا فیزیکی برای توضیح و تأیید منطق پیش‌بینی مدل استفاده کرد؛ به این ترتیب، یادگیری ماشین به عنوان ابزاری مکمل در کنار استدلال‌های مبتنی بر فیزیک انسانی عمل می‌کند (Chen, Liu, Li & Chen, 2024). شیلی خروجی مدل را به عنوان نتیجه تجمعی مقادیر شیلی مرتبط با هر ویژگی ورودی تفسیر می‌کند (Lundberg et al., 2020). این روش نه تنها امکان تبیین نتایج در سطح هر نمونه منفرد را فراهم می‌کند، بلکه تأثیر کلی ویژگی‌ها را در سطح کل داده‌ها نیز مشخص می‌سازد. از آنجا که در روش شیلی تمامی ویژگی‌ها به عنوان عوامل مشارکت‌کننده در نظر گرفته می‌شوند، خروجی مدل به صورت ترکیب خطی از ویژگی‌های ورودی تعریف می‌گردد که می‌تواند به صورت معادله (۵) بیان شود:

$$g(x') = \varphi_0 + \sum_{i=1}^m \varphi_i \quad (5)$$

g تابع توضیحی، x' بردار ائتلاف، φ_0 میانگین تمام خروجی‌های مدل، φ_i مقدار شیلی ویژگی i ام و m تعداد ویژگی‌های ورودی‌ها است (Zhang et al., 2023). در این نمودار در محور افقی مقادیر منفی و مثبت هرکدام به ترتیب نشان دهنده تأثیر در مقادیر کمتر و تأثیر در مقادیر بیشتر است. در محور عمودی، متغیرهایی که در بالای نمودار قرار دارند تأثیر بیشتر بر پیش‌بینی و مدل داشته‌اند و با حرکت به سمت پایین نمودار از اهمیت فاکتورها کاسته می‌شود.

روش‌های ارزیابی

ریشه میانگین مربعات خطا (RMSE) و میانگین خطای مطلق (MAE)

ریشه میانگین مربعات خطا (RMSE) یکی از شاخص‌های کاربرد برای سنجش میزان تطابق بین مقادیر مشاهده‌شده و مقادیر پیش‌بینی شده است. همچنین از معیار میانگین خطای مطلق (MAE) نیز برای ارزیابی دقت مدل استفاده می‌شود. معادله ۶ و ۷ به ترتیب نحوه محاسبه RMSE و MAE را نشان می‌دهد:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (X_{known} - X_{predicted})^2}{n}} \quad (۶)$$

$$MAE = \frac{\sum_{i=1}^n |X_{known} - X_{predicted}|}{n} \quad (۷)$$

n بیانگر تعداد مشاهدات، X_{known} مقادیر اندازه‌گیری شده، و $X_{predicted}$ مقادیر ارزیابی شده است (Gholami, Mohamadifar, Sorooshian & Jansen, 2020).

ضریب تعیین^۱ (R^2)

ضریب تعیین (R^2) یک شاخص بدون واحد برای ارزیابی برازش مدل‌ها است. مقدار این شاخص بین ۰ و ۱ است و هر چه به یک نزدیک‌تر باشد نشان دهنده بهتر بودن عملکرد مدل است (Nakagawa & Schielzeth, 2013). R^2 مطابق معادله (۸) محاسبه می‌شود:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (t_i - \bar{t})^2} \quad (۸)$$

n تعداد کل نمونه‌ها، y_i و t_i به ترتیب مقدار پیش‌بینی شده نمونه i و مقدار هدف نمونه i می‌باشند. همچنین $\bar{y}$ میانگین مقادیر پیش‌بینی شده توسط مدل است و $\bar{t}$ میانگین مقادیر واقعی اندازه‌گیری شده می‌باشد.

منحنی مشخصه عملکرد^۲ (ROC)

منحنی ROC ابزاری بصری برای ارزیابی عملکرد مدل به شمار می‌رود (Demissie, Rimal, Seyoum, Dutta & Rimmington, 2024). در این منحنی، حساسیت بر روی محور عمودی و ویژگی بر روی محور افقی نمایش داده می‌شود. هر نقطه روی منحنی نشان دهنده یک آستانه تصمیم مشخص است که بر اساس آن می‌توان دقت مدل را در پیش‌بینی ارزیابی کرد (Band et al., 2020). محور افقی منحنی ROC با ۱-Specificity (معادله ۹) و محور عمودی آن با Sensitivity (معادله ۱۰) شناخته می‌شود (Guo, Tian, Zhu, He & Zhang, 2024). شاخص مساحت زیر منحنی ROC که به آن AUC گفته می‌شود معیاری کمی برای سنجش کارایی مدل است؛ به طوری که هر چه مقدار AUC به عدد ۱ نزدیک‌تر باشد، عملکرد مدل بهتر است. این شاخصی برای ارزیابی توان مدل در پیش‌بینی صحیح رخدادها (خشکسالی) و عدم رخدادها (عدم خشکسالی) است. این مقدار نشان دهنده درصد پیش‌بینی درست واقعیت و همچنین درصد خطا در پیش‌بینی واقعیت بوده (Yariyan et al., 2020) و بر اساس معادله (۱۱) محاسبه می‌شود:

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (۹)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (۱۰)$$

$$\text{AUC} = \frac{\sum_{i \in \text{positive class}} \text{rank}_i - \frac{M(1+M)}{2}}{M \times N} \quad (۱۱)$$

در این روابط، FN (false negative) و FP (false positive) تعداد پیکسل‌هایی هستند که به اشتباه طبقه‌بندی شده‌اند، TN (true negative) و TP (true positive) نشان دهنده تعداد پیکسل‌هایی هستند که به درستی طبقه‌بندی شده‌اند (Shogrkhodaei, Razavi-Termeh & Fathnia, 2021).

1- Coefficient of Determination

2- Receiver Operating Characteristic

همچنین $rank_i$ رتبه نمونه i ام است و M تعداد نمونه‌های مثبت و N تعداد نمونه‌های منفی است (Guo et al., 2024). برای این منظور از ۳۰ درصد داده‌های وقوع خشکسالی و ۳۰ درصد داده‌های عدم وقوع خشکسالی برای ارزیابی ROC در نظر گرفته شد. در نهایت ۶۰ نقطه وقوع و ۶۰ نقطه عدم وقوع (در مجموع ۱۲۰ نقطه) برای ارزیابی توسط ROC در نظر گرفته شد. که ۳۰ درصد نقاط تست است و ۷۰ درصد آموزش است.

نتایج و بحث

نتایج ارزیابی مدل

برای مدل‌سازی و ارزیابی نتایج از محیط برنامه نویسی پایتون استفاده شد. ارزیابی مدل بر اساس روش 10-Fold cross validation انجام گرفت. ۷۰ درصد داده‌ها برای آموزش و ۳۰ درصد داده‌ها برای آزمایش در نظر گرفته شد. جدول ۲ نتایج ارزیابی مدل جنگل تصادفی را بر اساس معیارهای R^2 ، MAE و RMSE نشان می‌دهد. بر اساس نتایج، مقدار R^2 برای داده‌های آموزشی مدل برابر با ۰/۹۷ بود. مقدار MAE برابر با ۰/۰۳ و مقدار RMSE برابر ۰/۱۳۷ بود. همچنین مقدار RMSE، MAE و R^2 برای داده‌های ارزیابی مدل به ترتیب برابر با ۰/۱۸، ۰/۱ و ۰/۸۶ بود. نتایج نشان دهنده دقت خوب مدل برای پیش‌بینی خشکسالی بوده است.

جدول ۲- نتایج ارزیابی مدل با معیارهای R^2 ، MAE و RMSE

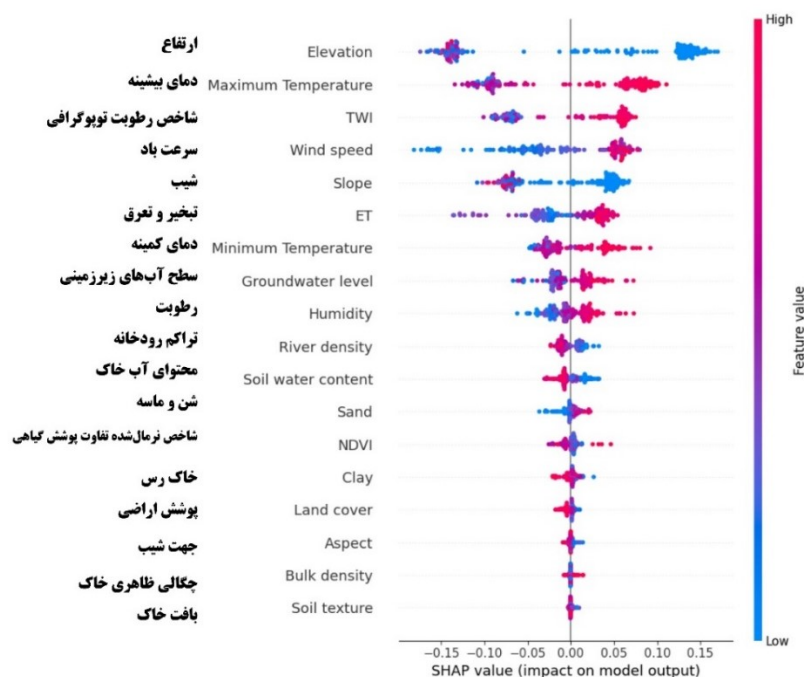
Table 2- Model evaluation results based on R^2 , MAE, and RMSE criteria

ضریب تعیین (R^2) Coefficient of Determination	میانگین خطای مطلق (MAE) Mean absolute error	خطای جذر میانگین مربعات (RMSE) Root mean square error	مدل Model
0.97	0.03	0.07	داده‌های آموزشی (Train data)
0.86	0.1	0.18	داده‌های ارزیابی (Test data)

نتایج شبلی

تحلیل نمودار پراکندگی SHAP به منظور درک جهت و ماهیت اثرگذاری متغیرها نشان داد (شکل ۴) که مقادیر پایین ارتفاع (نقاط آبی) عمدتاً با مقادیر مثبت SHAP همراه بوده‌اند، به این معنا که مناطق کم ارتفاع بیشترین نقش را در افزایش احتمال وقوع خشکسالی داشته‌اند؛ در حالی که مقادیر بالاتر ارتفاع اثر کاهنده بر خروجی مدل دارد. این موضوع نشان می‌دهد که مناطق پست تأثیر مثبتی بر پیش‌بینی مدل داشته و ارتفاعات بالاتر اثر کاهنده دارند. با افزایش ارتفاع (۴۰۰ متر به بالا) میزان رطوبت افزایش و در نتیجه خشکسالی کاهش می‌یابد (Feng et al., 2020). این نتیجه در مطالعه دویی و همکاران (Dubey, Swami, Gupt & Joshi, 2023) نیز تأیید شده است. در مقابل، مقادیر بیشتر دمای بیشینه و دمای کمینه (نقاط قرمز) عمدتاً در بخش مثبت نمودار قرار گرفته‌اند، که نشان می‌دهد افزایش دما تأثیر قابل توجهی در تشدید شرایط خشکسالی دارد. برای متغیرهای سرعت باد، تبخیر-تعرق واقعی و شاخص رطوبت توپوگرافی نیز الگوی مشابهی مشاهده شد، به طوری که مقادیر بالاتر این متغیرها سهم بیشتری در افزایش احتمال خشکسالی داشته‌اند. مقادیر کم سرعت باد باعث افزایش خروجی مدل شده و مقادیر زیاد آن اثرات نزدیک به صفر یا حتی منفی ایجاد می‌کند. این الگو نشان می‌دهد که شرایط کم‌بادتر برای مدل اهمیت مثبت بیشتری داشته است. همچنین سرعت بالای باد بر وقوع خشکسالی تأثیر گذار بوده است. سرعت باد زیاد، تعادل آبی گیاهان را مختل کرده، تبخیر و تعرق را تشدید می‌کند و در نهایت می‌تواند به آسیب به محصولات کشاورزی منجر شود. علاوه بر این، چنین شرایطی با خشکسالی‌های ناگهانی مرتبط بوده و نقش مهمی در گسترش آتش‌سوزی‌های انسانی در جنگل‌ها ایفا می‌کنند (Tavakol, Rahmani & Harrington Jr, 2020). شاخص رطوبت توپوگرافی نیز به روشنی نشان می‌دهد که مقادیر بالای شاخص رطوبت توپوگرافی (قرمز) اثر مثبت بر خروجی دارند، در حالی که مقادیر کم آن تمایل به ایجاد مقادیر منفی SHAP دارند؛ این یافته با نقش شاخص رطوبت توپوگرافی در نمایش نواحی پتانسیل رطوبت بالا همخوان است. در مورد تبخیر-تعرق نیز مقادیر زیاد تبخیر-تعرق معمولاً اثر کاهنده و مقادیر کمتر تبخیر-تعرق اثر افزایشی بر خروجی مدل داشته‌اند. در مورد متغیر شیب، مقدار اثرگذاری متعادل‌تر اما قابل تشخیص است؛ مقادیر پایین

شیب (آبی) غالباً موجب افزایش خروجی و حساسیت بیشتری نسبت به شکل‌گیری خشکسالی دارند. مقادیر بالای شیب (قرمز) بیشتر موجب کاهش خروجی می‌شوند. در مقابل، متغیرهایی نظیر رطوبت نسبی، تراکم آبراهه‌ها، محتوای آب خاک، شاخص پوشش گیاهی، شن، رس، پوشش اراضی و بافت خاک دامنه SHAP بسیار محدود و پراکندگی نزدیک به صفر داشته‌اند؛ بنابراین اثر آن‌ها بر خروجی مدل در مقایسه با متغیرهای اقلیمی و توپوگرافی ناچیز بوده است. این موضوع بیانگر آن است که مدل در تبیین وقوع خشکسالی عمدتاً بر پارامترهای حرارتی، توپوگرافی و دینامیک رطوبتی تکیه دارد. متغیرهای خاکی و کاربری زمین نقش فرعی و محدودتر داشته و اثر آنها بیشتر موضعی و کم‌دامنه باقی مانده است. این نتایج نشان می‌دهد که ساختار مدل بر پایه ویژگی‌های محیطی کلان‌مقیاس شکل گرفته و عوامل سطحی و خاکی در مرحله دوم اهمیت قرار داشته‌اند.

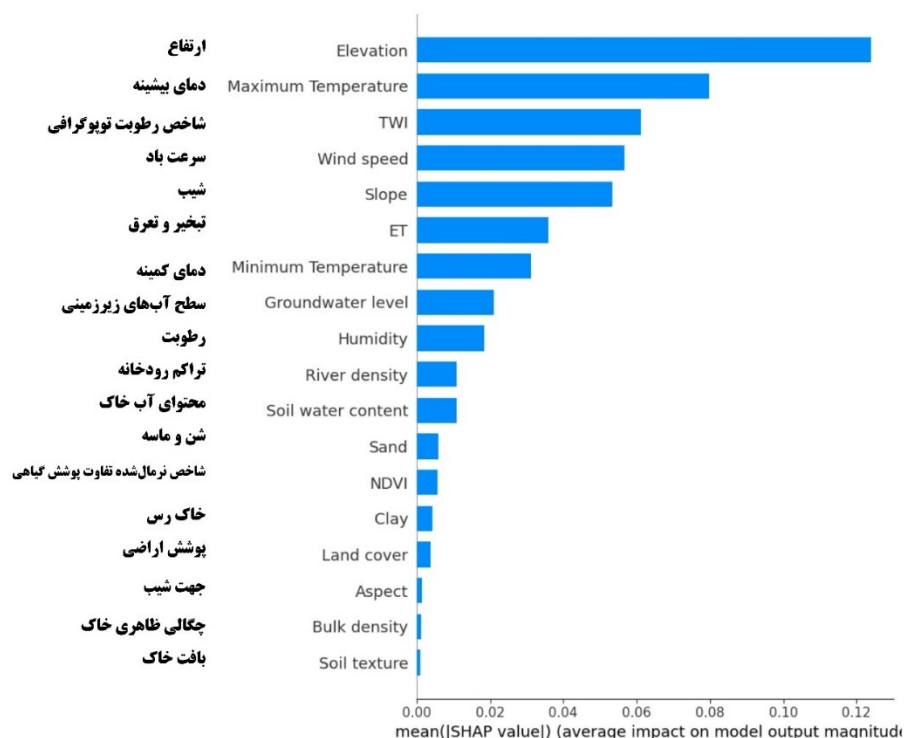


شکل ۴- نتایج مرتبط با روش شپلی

Fig.4. Results related to the Shapley method

تحلیل اهمیت

نمودار $\text{Mean}(|\text{SHAP value}|)$ (شکل ۵) اهمیت نسبی متغیرها را بر اساس میانگین قدر مطلق مقادیر SHAP نشان می‌دهد و بیانگر میزان تأثیر هر متغیر بر خروجی مدل بدون در نظر گرفتن جهت اثر است. بر اساس این نمودار، نتایج به دست آمده از تحلیل اهمیت ویژگی‌ها بر اساس میانگین قدر مطلق مقادیر SHAP نشان داد که ارتفاع مهم‌ترین متغیر تأثیرگذار در مدل بوده و بیشترین سهم را در تبیین تغییرات وقوع خشکسالی ایفا کرده است. پس از آن، متغیرهای دمای بیشینه، شاخص رطوبت توپوگرافی، سرعت باد، شیب و تبخیر - تعرق واقعی به ترتیب بیشترین میزان اهمیت را نشان دادند. این الگو بیانگر نقش غالب عوامل توپوگرافی و اقلیمی در پیش‌بینی خشکسالی است. در مقابل، متغیرهایی مانند بافت خاک، چگالی ظاهری خاک، جهت شیب و پوشش اراضی کمترین میزان اهمیت را داشته‌اند که حاکی از تأثیر محدود آن‌ها در خروجی مدل است. با توجه به این نتایج، مدل بیشترین حساسیت را نسبت به پارامترهای ارتفاعی و آب و هوایی نشان داده و ویژگی‌های مرتبط با خاک سهم بسیار کمتری در شکل‌دهی الگوی پیش‌بینی داشته‌اند.



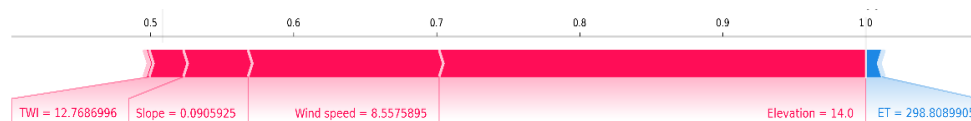
شکل ۵- نمودار Mean (|SHAP value|)

Fig. 5. Mean(|SHAP value|) chart

تحلیل محلی

نقاط وقوع خشکسالی

نمودار شکل ۶ نشان می‌دهد که برای این مورد خاص، مدل بر اساس ترکیبی از سرعت باد بالا، ارتفاع مشخص و شاخص رطوبت توپوگرافی، با اطمینان بسیار بالایی (نزدیک به ۱۰۰٪) احتمال وقوع خشکسالی را پیش‌بینی کرده است. در این سناریو، متغیر تبخیر - تعرق با اینکه به عنوان یک بازدارنده عمل کرده، اما در مقابل سیگنال‌های قدرتمند عوامل توپوگرافی و اقلیمی، تأثیر ناچیزی داشته است. این تحلیل تأکید می‌کند که در موارد شدید (احتمال نزدیک به ۱)، مدل به شدت تحت تأثیر متغیرهای اصلی (به ویژه سرعت باد و ارتفاع) قرار دارد و اثر متغیرهای دیگر در این حالت به حداقل می‌رسد.



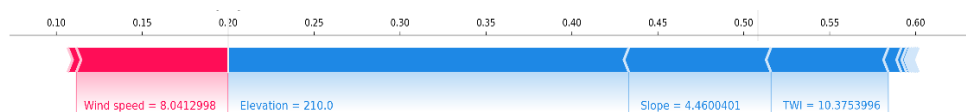
شکل ۶- تحلیل محلی برای نقاط وقوع خشکسالی

Fig. 6. Local analysis for drought occurrence points

نقاط عدم وقوع خشکسالی

به منظور درک عمیق‌تر سازوکار تصمیم‌گیری مدل در پیش‌بینی‌های فردی، از نمودار SHAP Force Plot برای یک نمونه تصادفی که مدل احتمال وقوع خشکسالی را در آن بسیار پایین (حدود ۰/۱۱) برآورد کرده است، استفاده شد. این تحلیل نشان می‌دهد که چگونه ویژگی‌های ورودی با مقادیر متفاوت خود، پیش‌بینی مدل را از «مقدار پایه» به سمت خروجی نهایی سوق می‌دهند.

بر اساس نتایج حاصل، متغیرهای ارتفاع با مقدار ۲۱۰، شیب با مقدار ۴/۴۶ و شاخص رطوبت توپوگرافی با مقدار ۱۰/۳۷، قدرتمندترین نیروهای کاهنده (بخش‌های آبی در نمودار) در این نمونه هستند. اثر تجمعی این ویژگی‌های توپوگرافی باعث شده است که مدل با اطمینان بالا، پتانسیل وقوع خشکسالی را در این نقطه کم ارزیابی کند. به ویژه، مقدار بالای ارتفاع به عنوان تعیین‌کننده‌ترین فاکتور بازدارنده عمل کرده است که نشان‌دهنده همبستگی منفی ارتفاع با ریسک خشکسالی در این منطقه خاص در ساختار یادگیری مدل است. در مقابل، متغیر سرعت باد با مقدار ۸/۰۴، تنها عامل افزایشدهنده ریسک (بخش قرمز) در این نمونه مشاهده شد که سعی در جابه‌جایی پیش‌بینی به سمت احتمال وقوع خشکسالی داشت. با این حال، با مقایسه طول بردارها در نمودار، مشخص است که شدت اثرات بازدارنده عوامل توپوگرافی (به‌طور کامل بر اثر افزایشدهنده سرعت باد غلبه کرده و در نهایت منجر به پیش‌بینی عدم وقوع خشکسالی شده است. این تحلیل محلی تأیید می‌کند که در مناطق با ارتفاع و رطوبت توپوگرافی بالاتر، مدل وزن کمتری به عوامل اقلیمی تشدیدکننده مانند سرعت باد می‌دهد، که این امر با مبانی فیزیکی و هیدرولوژیکی مناطق مورد مطالعه، سازگاری دارد.



شکل ۷- تحلیل محلی برای نقاط عدم وقوع خشکسالی

Fig. 7. Local analysis for drought non-occurrence points

تولید نقشه حساسیت خشکسالی

پس از مدلسازی با مدل جنگل تصادفی، خروجی مدل‌سازی در نرم‌افزار ArcGIS به کل منطقه تعمیم داده شد و برای هر پیکسل لحاظ گردید. سپس با بهره‌گیری از روش شکست طبیعی^۱، مخاطرات چندگانه در پنج طبقه «خیلی کم، کم، متوسط، زیاد و خیلی زیاد» دسته‌بندی شدند (شکل ۸). براساس نتایج ۳/۱۰ درصد منطقه مطالعاتی در طبقه «خیلی کم»، ۱۸/۸ درصد در طبقه «کم»، ۳۰/۸ درصد منطقه در طبقه «متوسط»، ۸۸/۱۰ درصد منطقه در طبقه زیاد و ۵۸/۴۱ درصد منطقه در طبقه «خیلی زیاد» قرار گرفته است. بیشترین فراوانی مشاهده شده طبقات مرتبط با دو طبقه «خیلی کم» و «خیلی زیاد» بوده است. بر طبق نقشه حساسیت به خشکسالی، مناطق شمالی، شرق و جنوب شرقی و مناطق مرکزی در کلاس خیلی کم قرار گرفته‌اند. مناطقی مانند اندیمشک، دزفول، مسجد سلیمان، رامهرمز، شوشتر و باغ ملک در این دسته قرار گرفته‌اند. بخشی از مناطق جنوبی، جنوب غرب و غربی نیز در دسته حساسیت «خیلی زیاد» قرار گرفته‌اند. مناطقی مانند بخش‌هایی از ماهشهر، آبادان، شادگان، اهواز، سوسنگرد، و بخش‌هایی از شوش در این طبقه قرار می‌گیرند. خرمشهر دارای تمام طبقات خشکسالی است. یک نوار که از شمال غربی شروع می‌شود و به جنوب شرقی ختم می‌شود در کلاس «متوسط» قرار گرفته است.

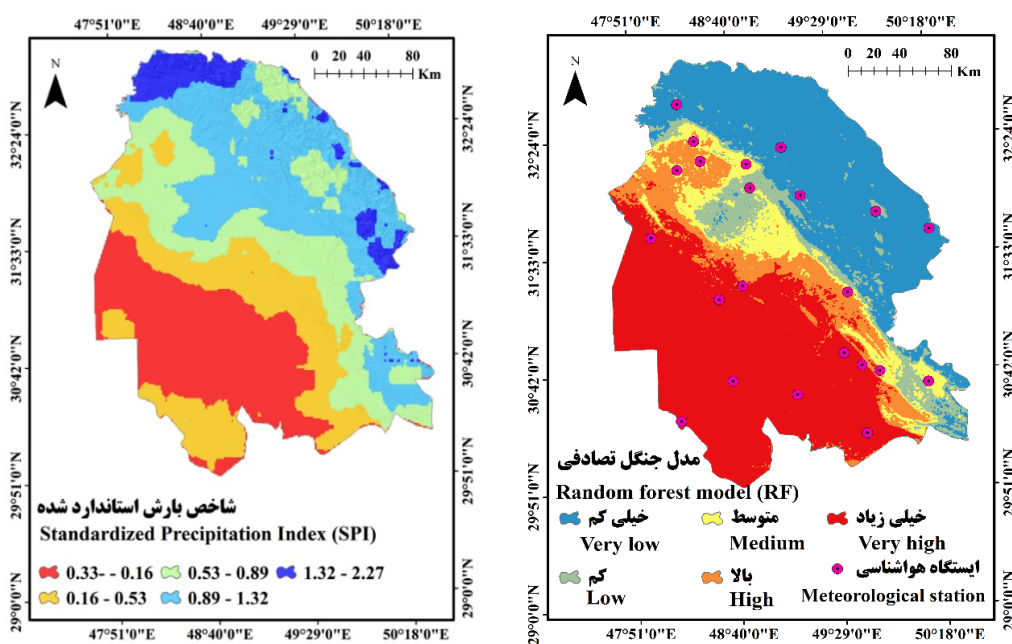
نتایج ارزیابی نقشه حساسیت

شکل ۹ منحنی ROC و نتایج ROC را برای خشکسالی نشان می‌دهد. مقدار AUC عملکرد مخاطره خشکسالی را نشان می‌دهد. بر اساس نتایج مقدار AUC برابر با ۰/۹۸۷ بوده است. این مقدار نشان‌دهنده عملکرد عالی نقشه خروجی برای خشکسالی در استان خوزستان بوده است.

این مطالعه صحت و دقت عملکرد مدل جنگل تصادفی را برای تهیه نقشه حساسیت به خشکسالی در استان خوزستان تأیید کرده است. این در راستای سایر مطالعاتی است که از جنگل تصادفی برای مدل‌سازی یا پیش‌بینی خشکسالی استفاده کرده‌اند. برای مثال حجاریان (Hajarian, 2025) به مدل‌سازی حساسیت خشکسالی با استفاده از مدل جنگل تصادفی و پهنه‌بندی خشکسالی در

1- Natural Breaks

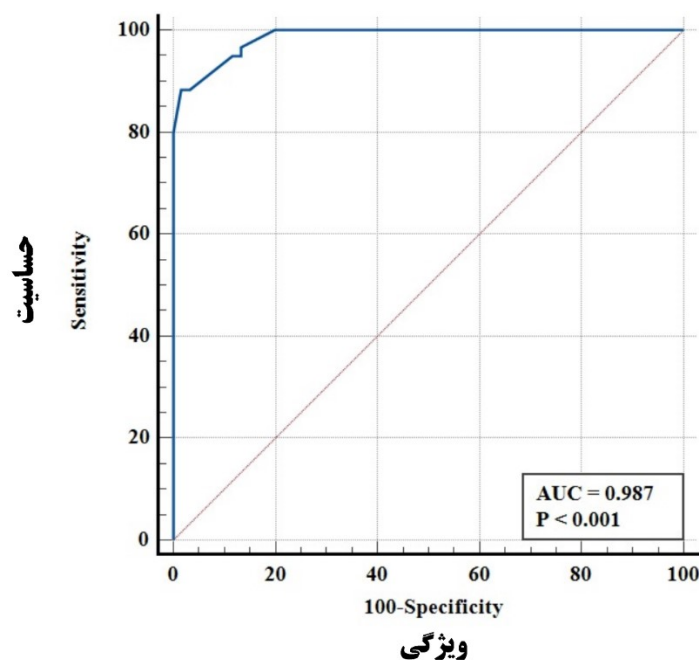
اردبیل پرداخت. نتایج او نشان داد که تقریباً ۵۰ درصد از مساحت مورد مطالعه در گروه مناطقی با حساسیت زیاد تا بسیار زیاد نسبت به وقوع خشکسالی قرار دارند. ارزیابی عملکرد مدل از طریق منحنی ROC و مقدار سطح زیر منحنی بالاتر از ۰/۹، نشان دهنده دقت بالای مدل در پیش‌بینی حساسیت به خشکسالی است. البلتاگی و همکاران (Elbeltagi et al., 2023) برای پایش و پیش‌بینی خشکسالی در هند مرکزی در ایالت ماهاراشترا، به پیاده‌سازی الگوریتم‌های یادگیری ماشین بر اساس شاخص بارش استاندارد (SPI) پرداختند. نتایج آن‌ها نشان داد که الگوریتم جنگل تصادفی نسبت به سایر الگوریتم‌های مورد استفاده در این مطالعه، عملکرد بهتری در پیش‌بینی خشکسالی در منطقه مطالعاتی داشته است. همچنین حسین و همکاران (Hussain et al., 2025) با هدف بررسی پایداری و دینامیک زمانی خشکسالی‌های فصلی، عملکرد پیش‌بینی مدل‌های جنگل تصادفی و رگرسیون لجستیک^۱ را مورد ارزیابی قرار دادند. نتایج نشان داد مدل جنگل تصادفی در شناسایی خشکسالی‌های پایدار عملکرد برتری نسبت به رگرسیون لجستیک دارد و حساسیت بالاتری در تشخیص موارد خشکسالی ارائه می‌کند. علاوه بر این، جنگل تصادفی در اغلب ایستگاه‌ها دقت و ضریب کاپا^۲ بالاتری نشان داد و توانست هماهنگی بهتری بین پیش‌بینی‌ها و مشاهدات واقعی فراهم کند. پارک و همکاران (Park et al., 2016) در مطالعه‌ای برای پایش خشکسالی هواشناسی و کشاورزی در اقلیم‌های مختلف آمریکا طی سال‌های ۲۰۰۰ تا ۲۰۱۲ از شانزده شاخص مبتنی بر داده‌های ماهواره‌ای MODIS^۳ و TRMM^۴ استفاده کردند. آن‌ها از سه روش یادگیری ماشین شامل جنگل تصادفی، درخت‌های رگرسیون تقویتی و کوبیست^۵ استفاده کردند که در این میان جنگل تصادفی بهترین نتیجه را ارائه داد. این مطالعات نتایج تحقیق حاضر را که عملکرد عالی روش جنگل تصادفی برای تهیه نقشه حساسیت به خشکسالی می‌باشد را تأیید می‌کنند.



شکل ۸- نتایج خروجی مدل جنگل تصادفی و نقشه SPI برای منطقه مطالعاتی

Fig. 8. Random Forest Model Output Results and SPI Map for the Study Area

- 1- Logistic Regression
- 2- Kappa
- 3- Moderate Resolution Imaging Spectroradiometer
- 4- Tropical Rainfall Measuring Mission
- 5- Cubist



Area under the ROC curve (AUC)	0.987
Standard Error ^a	0.00654
95% Confidence interval ^b	0.946 to 0.999
z statistic	74.392
Significance level P (Area=0.5)	<0.0001

شکل ۹- نتایج عملکرد منحنی ROC

Fig. 9. ROC curve performance results

نتیجه‌گیری

با وقوع مکرر خشکسالی در سال‌های اخیر خسارات زیادی بر محیط زیست و زندگی انسان‌ها وارد شده است. این مطالعه به ارائه چارچوبی برای تفسیرپذیری مدل جنگل تصادفی برای تهیه نقشه حساسیت خشکسالی در استان خوزستان پرداخته است. نتایج حاصل از مدل جنگل تصادفی برای استان خوزستان به خوبی نشان داد که حدود ۳۱/۰۳ درصد از منطقه در طبقه خشکسالی خیلی کم قرار گرفته است که براساس لایه تراکم رودخانه‌ها، تجمع انشعابات رودخانه‌ای در این محدوده، بالاتر بودن میزان آب خاک و کاهش در شاخص رطوبت توپوگرافی همراه با کاهش در سرعت باد (کاهش تبخیر سطحی)، قرار گرفتن در ارتفاعات بالاتر این مناطق موجب ایجاد وضعیت مطلوب شده است. این عوامل یک توزیع عملکردی همگن در محدوده را نشان می‌دهد که توانسته است با وجود تغییرات اقلیمی و رو به خشکی رفتن بسیاری از مناطق کشور ایران یک سیستم پایدار را ایجاد کند. تلاش برای حفظ منابع آبی در این نواحی و به کارگیری رویکردهای نوین مدیریتی می‌تواند به پایداری این شرایط کمک کند. با این حال حدود ۴۱/۵۸ درصد از استان خوزستان در طبقه خشکسالی بسیار زیاد قرار گرفته است که با کاهش در میزان بارش در این ناحیه همراه با افزایش در مقدار شاخص TWI و سرعت باد همراه است. همچنین کاهش ارتفاع و افزایش در بیشینه دما از عوامل مهم دیگر بوده است. شیب به عنوان عاملی مؤثر در تعیین خشکسالی عمل کرده است. با کاهش شیب شاهد افزایش در خشکسالی بوده‌ایم و در مناطقی با شیب بیشتر این روند معکوس بوده است. عوامل اقلیمی نقشی تعیین کننده در ایجاد و یا کاهش خشکسالی در استان خوزستان ایفا کرده است که بر لزوم شیوه صحیح بهره برداری از منابع آبی تأکید می‌کند. با این حال افزایش شهرنشینی و تبدیل سطوح قابل نفوذ به سطوح غیر قابل نفوذ نیز می‌تواند از عوامل مؤثر بر افزایش خشکسالی باشد.

راهکارهای عمومی برای کنترل خشکسالی نمی‌تواند پاسخ مناسبی برای شرایط حاکم باشد. زیرا خشکسالی تحت تأثیر عوامل مختلفی قرار دارد و چه بسا تأکید بر کاهش اثر یک فاکتور موجب تشدید در یک فاکتور دیگر شود. بنابراین باید رویکردی مورد استفاده قرار گیرد که همه جوانب را در نظر بگیرد و یک تضمین بلند مدت با توانایی خود تنظیمی در دسترس قرار دهد. علاوه بر این نتایج نشان داد که نمودار شیلی به خوبی توانسته است به توضیح اثرات فاکتورهای مؤثر بر خشکسالی بپردازد و آنها را اولویت بندی کند. این نشان دهنده کاربردی بودن این نمودار در جهت درک بهتر خشکسالی و عوامل تأثیر گذار بر آن در منطقه مطالعاتی است. استفاده از مدل‌های یادگیری ماشین همراه با نمودار تفسیرپذیر شیلی به عنوان یک رویکرد تحلیلی در جهت اقدامات مؤثر برای کاهش اثرات خشکسالی در منطقه عمل می‌کند.

با وجود مزایای استفاده از یادگیری ماشین و عملکرد خوب در تهیه نقشه حساسیت خشکسالی، یک روش به تنهایی نمی‌تواند همه جوانب را در نظر بگیرد. استفاده از چند روش به صورت همزمان و استفاده از روش‌های یادگیری عمیق که تحت تأثیر بیش برآزش قرار نمی‌گیرند می‌تواند دقت عملکرد مدل را افزایش دهد. محدودیت در تعداد ایستگاه‌های هواشناسی و دسترسی محدود به داده‌ها می‌تواند با تلفیقی از داده‌های سنجش از دور و ماهواره‌ای تکمیل شود و دقت عملکرد مدل افزایش یابد.

نقشه‌های حساسیت خشکسالی ابزاری مؤثر برای شناسایی نواحی آسیب‌پذیر و اولویت‌بندی اقدامات مدیریتی هستند و می‌تواند به سیاست‌گذاران و مدیران یک بینش جامع برای مقابله و یا کاهش اثرات این مخاطره ارائه دهد و با اتخاذ تصمیمات پیشگیرانه و مبتنی بر شواهد مکانی، منابع را به صورت هدفمند تخصیص دهند. استفاده از این ابزار می‌تواند پیامدهای اقتصادی و اجتماعی خشکسالی را کاهش داده و تاب‌آوری اکوسیستم‌ها و جوامع انسانی را تقویت کند.

References

- Abbes, A. B., Inoubli, R., Rhif, M., & Farah, I. R. (2023). Combining deep learning methods and multi-resolution analysis for drought forecasting modeling. *Earth Science Informatics*, 16(2), 1811-1820. <https://doi.org/10.1007/s12145-023-01009-4>
- AghaKouchak, A., Huning, L. S., Sadegh, M., Qin, Y., Markonis, Y., Vahedifard, F., ... & Kreibich, H. (2023). Toward impact-based monitoring of drought and its cascading hazards. *Nature Reviews Earth & Environment*, 4(8), 582-595. <https://doi.org/10.1038/s43017-023-00457-2>
- Al-Abadi, A. M. (2018). Mapping flood susceptibility in an arid region of southern Iraq using ensemble machine learning classifiers: a comparative study. *Arabian Journal of Geosciences*, 11(9), 218. <https://doi.org/10.1007/s12517-018-3584-5>
- Azizi, M., & Abbasi, A. (2025). Developing a Robust Data-Driven Model Based on Ground and Satellite Measured Data for Agricultural Drought Prediction in Iran. *Ferdowsi Civil Engineering*, 38(1). [In Persian] <https://doi.org/10.22067/jfpei.2024.88855.1311>
- Bahiraie, M., Hosseini, S. M., & Hossein-Panahi, B. (2025). Groundwater resources exploitation management in response to water scarcity challenges in Khuzestan Province, Iran. *Journal of Groundwater Science and Engineering*, 13(3), 268-285. doi: <https://doi.org/10.26599/JGSE.2025.9280054>
- Baloei, F., Kabolizadeh, M., & Mohammadi, S. (2024). Monitoring temporal-spatial changes of meteorological drought and trend analysis of climatic variables in Khuzestan province from 2000 to 2020. *Journal of Natural Environmental Hazards*, 13(40), 73-94. [In Persian] <https://doi.org/10.22111/JNEH.2024.46574.1987>
- Band, S. S., Janizadeh, S., Chandra Pal, S., Saha, A., Chakraborty, R., Melesse, A. M., & Mosavi, A. (2020). Flash flood susceptibility modeling using new approaches of hybrid and ensemble tree-based machine learning algorithms. *Remote Sensing*, 12(21), 3568. <https://doi.org/10.3390/rs12213568>
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Cai, H., Shi, H., Zhou, Z., Liu, S., & Babovic, V. (2024). Explaining the mechanism of multiscale groundwater drought events: A new perspective from interpretable deep learning model. *Water Resources Research*, 60(7), e2023WR035139. <https://doi.org/10.1029/2023WR035139>
- Chang, D. L., Yang, S. H., Hsieh, S. L., Wang, H. J., & Yeh, K. C. (2020). Artificial intelligence methodologies applied to prompt pluvial flood estimation and prediction. *Water*, 12(12), 3552. <https://doi.org/10.3390/w12123552>
- Chen, C., Liu, Y., Li, Y., & Chen, D. (2024). Explainable artificial intelligence framework for urban global digital elevation model correction based on the SHapley additive explanation-random forest algorithm considering

- spatial heterogeneity and factor optimization. *International Journal of Applied Earth Observation and Geoinformation*, 129, 103843. <https://doi.org/10.1016/j.jag.2024.103843>
- Chen, J., Li, M., & Wang, W. (2012). Statistical uncertainty estimation using random forests and its application to drought forecast. *Mathematical Problems in Engineering*, 2012(1), 915053. <https://doi.org/10.1155/2012/915053>
- Chen, W., Li, Y., Xue, W., Shahabi, H., Li, S., Hong, H., & Ahmad, B. B. (2020). Modeling flood susceptibility using data-driven approaches of naïve bayes tree, alternating decision tree, and random forest methods. *Science of The Total Environment*, 701, 134979. <https://doi.org/10.1016/j.scitotenv.2019.134979>
- De Meester, J., & Willems, P. (2024). Analysing spatial variability in drought sensitivity of rivers using explainable artificial intelligence. *Science of The Total Environment*, 931, 172685. <https://doi.org/10.1016/j.scitotenv.2024.172685>
- Demissie, Z., Rimal, P., Seyoum, W. M., Dutta, A., & Rimmington, G. (2024). Flood susceptibility mapping: Integrating machine learning and GIS for enhanced risk assessment. *Applied Computing and Geosciences*, 23, 100183. <https://doi.org/10.1016/j.acags.2024.100183>
- Dikshit, A., Pradhan, B., & Alamri, A. M. (2020). Short-term spatio-temporal drought forecasting using random forests model at New South Wales, Australia. *Applied Sciences*, 10(12), 4254. <https://doi.org/10.3390/app10124254>
- Dikshit, A., Pradhan, B., & Santosh, M. (2022). Artificial neural networks in drought prediction in the 21st century—A scientometric analysis. *Applied Soft Computing*, 114, 108080. <https://doi.org/10.1016/j.asoc.2021.108080>
- Dubey, A., Swami, D., Gupta, V., & Joshi, N. (2023). From the peaks to the plains: investigating the role of elevation in governing drought dynamics over the Indus river basin. *Atmospheric Research*, 291, 106824. <https://doi.org/10.1016/j.atmosres.2023.106824>
- Elbeltagi, A., Pande, C. B., Kumar, M., Tolche, A. D., Singh, S. K., Kumar, A., & Vishwakarma, D. K. (2023). Prediction of meteorological drought and standardized precipitation index based on the random forest (RF), random tree (RT), and Gaussian process regression (GPR) models. *Environmental Science and Pollution Research*, 30(15), 43183-43202. <https://doi.org/10.1007/s11356-023-25221-3>
- En-Nagre, K., Aqnouy, M., Ouarka, A., Naqvi, S. A. A., Bouizrou, I., El Messari, J. E. S., ... & El-Askary, H. (2024). Assessment and prediction of meteorological drought using machine learning algorithms and climate data. *Climate Risk Management*, 45, 100630. <https://doi.org/10.1016/j.crm.2024.100630>
- Fadaei-Kermani, E., Barani, G. A., & Ghaeini-Hessaroeiyeh, M. (2017). Drought monitoring and prediction using K-nearest neighbor algorithm. *Journal of AI and data mining*, 5(2), 319-325. <https://doi.org/10.22044/jadm.2017.881>
- Farhadi, H., & Najafzadeh, M. (2021). Flood risk mapping by remote sensing data and random forest technique. *Water*, 13(21), 3115. <https://doi.org/10.3390/w13213115>
- Feng, P., Wang, B., Li Liu, D., & Yu, Q. (2019). Machine learning-based integration of remotely sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia. *Agricultural Systems*, 173, 303-316. <https://doi.org/10.1016/j.agsy.2019.03.015>
- Feng, W., Lu, H., Yao, T., & Yu, Q. (2020). Drought characteristics and its elevation dependence in the Qinghai-Tibet plateau during the last half-century. *Scientific Reports*, 10(1), 14323. <https://doi.org/10.1038/s41598-020-71295-1>
- Gholami, H., Mohamadifar, A., Sorooshian, A., & Jansen, J. D. (2020). Machine-learning algorithms for predicting land susceptibility to dust emissions: The case of the Jazmurian Basin, Iran. *Atmospheric pollution research*, 11(8), 1303-1315. <https://doi.org/10.1016/j.apr.2020.05.009>
- Ghorbani, M. A., Deo, R. C., Kashani, M. H., Shahabi, M., & Ghorbani, S. (2019). Artificial intelligence-based fast and efficient hybrid approach for spatial modelling of soil electrical conductivity. *Soil and Tillage Research*, 186, 152-164. <https://doi.org/10.1016/j.still.2018.09.012>
- Ghorbanian, A., Kakooei, M., Amani, M., Mahdavi, S., Mohammadzadeh, A., & Hasanlou, M. (2020). Improved Land Cover Map of Iran Using Sentinel Imagery within Google Earth Engine and a Novel Automatic Workflow for Land Cover Classification Using Migrated Training Samples. *ISPRS Journal of Photogrammetry and Remote Sensing*, 167, 276-288. <https://doi.org/10.1016/j.isprsjprs.2020.07.013>
- Gibson, P. B., Chapman, W. E., Altinok, A., Delle Monache, L., DeFlorio, M. J., & Waliser, D. E. (2021). Training machine-learning models on climate model output yields skillful interpretable seasonal precipitation forecasts. *Communications Earth & Environment*, 2(1), 159. <https://doi.org/10.1038/s43247-021-00225-4>

- Guo, Z., Tian, B., Zhu, Y., He, J., & Zhang, T. (2024). How do the landslide and non-landslide sampling strategies impact landslide susceptibility assessment?—A catchment-scale case study from China. *Journal of Rock Mechanics and Geotechnical Engineering*, 16(3), 877-894. <https://doi.org/10.1016/j.jrmge.2023.07.026>
- Hajarian, A. (2025). Meteorological Drought Risk Monitoring and Zoning Using Random Forest Model. *Journal of Drought and Climate Change Research*, 3(1), 83-94. [In Persian] <https://doi.org/10.22077/jdcr.2025.8551.1093>
- Huang, C., She, D., Liu, X., Zhang, Y., Wu, Y., Shi, X., & Wei, Y. (2025). Predicting future meteorological drought risk in mainland China using random forest model. *Journal of Hydrology: Regional Studies*, 61, 102633. <https://doi.org/10.1016/j.ejrh.2025.102633>
- Hussain, A., Niaz, R., Almazah, M. M., Al-Rezami, A. Y., Cheng, H., & Tariq, A. (2025). Utilizing Logistic Regression and Random Forest to Model Meteorological Drought Persistence Across Seasonal Transitions. *Earth Systems and Environment*, 1-24. <https://doi.org/10.1007/s41748-025-00682-3>
- Hussain, A., Niaz, R., Al-Rezami, A. Y., Mohamed Omer, A., S. Al-Duais, F., & MA Almazah, M. (2025). Application of random forest for identification of an appropriate model for predicting meteorological drought. *Advances in Meteorology*, 2025(1), 7674140. <https://doi.org/10.1155/adme/7674140>
- Jahangir, M. H., & Mohammadi, A. (2016). Compare Standard Precipitation Index and TOPSIS ranking method to assess the drought adventure Condition in Khouzesan province. *Environmental Management Hazards*, 3(2), 125-139. [In Persian] <https://doi.org/10.22059/JHSCI.2016.60343>
- Khan, N., Sachindra, D. A., Shahid, S., Ahmed, K., Shiru, M. S., & Nawaz, N. (2020). Prediction of droughts over Pakistan using machine-learning algorithms. *Advances in Water Resources*, 139, 103562. <https://doi.org/10.1016/j.advwatres.2020.103562>
- Koochak, S. S., & Behnia, A. (2013). Monitoring and Prediction of Khuzestan Province, Iran Drought Using SPI drought Index and Markov Chain. *Journal of Irrigation Sciences and Engineering*, 36(3), 1-12. [In Persian]
- Li, X., Jia, H., & Wang, L. (2023). Remote sensing monitoring of drought in Southwest China using random forest and eXtreme gradient boosting methods. *Remote Sensing*, 15(19), 4840. <https://doi.org/10.3390/rs15194840>
- Liu, J., Li, M., Li, R., Shalamzari, M. J., Ren, Y., & Silakhori, E. (2025). Comprehensive Assessment of Drought Susceptibility Using Predictive Modeling, Climate Change Projections, and Land Use Dynamics for Sustainable Management. *Land*, 14(2), 337. <https://doi.org/10.3390/land14020337>
- Lotfirad, M., Esmaili-Gisavandani, H., & Adib, A. (2022). Drought monitoring and prediction using SPI, SPEI, and random forest model in various climates of Iran. *Journal of Water and Climate Change*, 13(2), 383-406. <https://doi.org/10.2166/wcc.2021.287>
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30.
- Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. *Nature Machine Intelligence*, 2(1), 56-67. <https://doi.org/10.1038/s42256-019-0138-9>
- Mahmoudi, P., Daneshmand, H., (2018). The Application of Wavelet Analysis to Identify the Periodic Behavior of Droughts in Iran, *Journal of Geography and Environmental Hazards*, 7(1), 153-168. [In Persian] <https://doi.org/10.22067/geo.v7i1.60717>
- McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. *In Proceedings of the 8th Conference on Applied Climatology*, 22, 179-183. <https://climate.colostate.edu/pdfs/relationshipofdroughtfrequency.pdf>
- Mokhtar, A., Jalali, M., He, H., Al-Ansari, N., Elbeltagi, A., Alsafadi, K., & Rodrigo-Comino, J. (2021). Estimation of SPEI meteorological drought using machine learning algorithms. *IEEE Access*, 9, 65503-65523. <https://doi.org/10.1109/ACCESS.2021.3074305>
- Nakagawa, S., & Holger S. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133-142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>
- Nejadrekabi, M., Eslamian, S., & Zareian, M. J. (2022). Spatial statistics techniques for SPEI and NDVI drought indices: a case study of Khuzestan Province. *International Journal of Environmental Science and Technology*, 19(7), 6573-6594. <https://doi.org/10.1007/s13762-021-03852-8>
- Oyarzabal, R. S., Santos, L. B., Cunningham, C., Broedel, E., de Lima, G. R., Cunha-Zeri, G., ... & Cunha, A. P. (2025). Forecasting drought using machine learning: a systematic literature review. *Natural Hazards*, 1-29. <https://doi.org/10.1007/s11069-025-07195-2>

- Ozupek, E., Teke, A., Celik, N., & Kavzoglu, T. (2025). Explainable artificial intelligence to explore the intrinsic characteristics of climatic parameters governing meteorological drought forecasting: opening the black box. *Stochastic Environmental Research and Risk Assessment*, 39, 3201-3222. <https://doi.org/10.1007/s00477-025-03007-y>
- Park, H., Kim, K., & Lee, D. K. (2019). Prediction of severe drought area based on random forest: Using satellite image and topography data. *Water*, 11(4), 705. <https://doi.org/10.3390/w11040705>
- Park, S., Im, J., Jang, E., & Rhee, J. (2016). Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions. *Agricultural and Forest Meteorology*, 216, 157-169. <https://doi.org/10.1016/j.agrformet.2015.10.011>
- Parker, T., & Gallant, A. J. (2022). The role of heavy rainfall in drought in Australia. *Weather and Climate Extremes*, 38, 100528. <https://doi.org/10.1016/j.wace.2022.100528>
- Piri, S., & Ghojghar, M. A. (2025). Development of intelligent machine learning and Jenkins box models based on the full ensemble mode decomposition method for meteorological drought modeling (Case study: Khuzestan Province). *Climate Change Research*, 6(22), 81-96. [In Persian] <https://doi.org/10.30488/CCR.2025.516420.1276>
- Rahmati, O., Darabi, H., Panahi, M., Kalantari, Z., Naghibi, S. A., Ferreira, C. S. S., ... & Haghighi, A. T. (2020). Development of novel hybridized models for urban flood susceptibility mapping. *Scientific Reports*, 10(1), 12937. <https://doi.org/10.1038/s41598-020-69703-7>
- Rezaei, P., Roshni, M., & Amirabadizadeh, S. M., (2016). The Study of Relationship between Intensity, Duration, Frequency and Involvement Level with Drought in the Northwest of Iran, *Journal of Geography and Environmental Hazards*, 3(2), 91-110. [In Persian] https://ccr.gu.ac.ir/article_220614.html?
- Seyedi Ghaldareh, J., Ahmadi, S., & Gholamnia, M. (2023). Evaluation of remote sensing-based drought monitoring indexes using support vector regression and random forest models (Case study: Marivan city). *Engineering Journal of Geospatial Information Technology*, 10(3), 121-141. [In Persian] <https://doi.org/10.52547/jgit.10.3.121>
- Shogrkhodaei, S. Z., Razavi-Termeh, S. V., & Fathnia, A. (2021). Spatio-temporal modeling of PM2.5 risk mapping using three machine learning algorithms. *Environmental Pollution*, 289, 117859. <https://doi.org/10.1016/j.envpol.2021.117859>
- Soltani, S., Saboohi, R., & Yaghmaei, L. (2012). Rainfall and rainy days trend in Iran. *Climatic Change*, 110(1), 187-213. <https://doi.org/10.1007/s10584-011-0146-1>
- Tavakol, A., Rahmani, V., & Harrington Jr, J. (2020). Temporal and spatial variations in the frequency of compound hot, dry, and windy events in the central United States. *Scientific Reports*, 10(1), 15691. <https://doi.org/10.1038/s41598-020-72624-0>
- Torkaman Pary, A., Rastgoo, P., Opp, C., Zeuss, D., & Abera, T. A. (2024). Impacts of Drought Severity and Frequency on Natural Vegetation Across Iran. *Water*, 16(22), 3334. <https://doi.org/10.3390/w16223334>
- Tramblay, Y., & Quintana Seguí, P. (2022). Estimating soil moisture conditions for drought monitoring with random forests and a simple soil moisture accounting scheme. *Natural Hazards and Earth System Sciences*, 22(4), 1325-1334. <https://doi.org/10.5194/nhess-22-1325-2022>
- Vagheei, R., Torabi, Z., & Ghaemi, A. (2021). Drought assessment using standardized precipitation index and random forest algorithm. *Journal of Meteorology and Atmospheric Science*, 3(4), 390-405. [In Persian] <https://doi.org/10.22034/jmas.2021.302576.1153>
- Wang, Q., Zhao, L., Wang, M., Wu, J., Zhou, W., Zhang, Q., & Deng, M. (2022). A random forest model for drought: Monitoring and validation for grassland drought based on multi-source remote sensing data. *Remote Sensing*, 14(19), 4981. <https://doi.org/10.3390/rs14194981>
- Xue, Z., Diao, S., Yang, F., Fei, L., Wang, W., Fang, L., & Liu, Y. (2025). Identifying Forest Drought Sensitivity Drivers in China Under Lagged and Accumulative Effects via XGBoost-SHAP. *Remote Sensing*, 17(16), 2903. <https://doi.org/10.3390/rs17162903>
- Yariyan, P., Avand, M., Abbaspour, R. A., Torabi Haghighi, A., Costache, R., Ghorbanzadeh, O., & Blaschke, T. (2020). Flood susceptibility mapping using an improved analytic network process with statistical models. *Geomatics, Natural Hazards and Risk*, 11(1), 2282-2314. <https://doi.org/10.1080/19475705.2020.1836036>
- Zarasvandi, A., Carranza, E. J. M., Moore, F., & Rastmanesh, F. (2011). Spatio-temporal occurrences and mineralogical-geochemical characteristics of airborne dusts in Khuzestan Province (southwestern Iran). *Journal of Geochemical Exploration*, 111(3), 138-151. <https://doi.org/10.1016/j.gexplo.2011.04.004>

- Zarei, A. R., Mahmoudi, M. R., & Moghimi, M. M. (2023). Determining the most appropriate drought index using the random forest algorithm with an emphasis on agricultural drought. *Natural Hazards*, 115(1), 923-946. <https://doi.org/10.1007/s11069-022-05579-2>
- Zhang, B., Salem, F. K. A., Hayes, M. J., Smith, K. H., Tadesse, T., & Wardlow, B. D. (2023). Explainable machine learning for the prediction and assessment of complex drought impacts. *Science of The Total Environment*, 898, 165509. <https://doi.org/10.1016/j.scitotenv.2023.165509>
- Zhang, J., Ma, X., Zhang, J., Sun, D., Zhou, X., Mi, C., & Wen, H. (2023). Insights into geospatial heterogeneity of landslide susceptibility based on the SHAP-XGBoost model. *Journal of Environmental Management*, 332, 117357. <https://doi.org/10.1016/j.jenvman.2023.117357>
- Zhao, Y., Zhang, J., Bai, Y., Zhang, S., Yang, S., Henchiri, M., & Nanzad, L. (2022). Drought monitoring and performance evaluation based on machine learning fusion of multi-source remote sensing drought factors. *Remote Sensing*, 14(24), 6398. <https://doi.org/10.3390/rs14246398>