

Research Paper

Species distribution modeling of Hyrcanian boxwood (*Buxus Hyrcana* Pojark) using the random forest model in the forests of northern Iran

Aref Hesabi¹, Seyed Jalil Alavi^{*2} and Omid Esmailzadeh³

1- PhD Student of Forest Management, Department of Forest Science, Faculty of Natural Resources and marine science, Tarbiat Modares University, Nur, Mazandaran, I. R. Iran. (aref.hesabi@modares.ac.ir)

2,*- (Corresponding author) Associate Professor, Department of Forest Science, Faculty of Natural Resources and marine science, Tarbiat Modares University, Nur, Mazandaran, I. R. Iran. (j.alavi@modares.ac.ir)

3- Associate Professor, Department of Forest Sciences and Engineering, Faculty of Natural Resources and Marine Sciences, Tarbiat Modares University, Nur, Mazandaran, I. R. Iran. (oesmailzadeh@modares.ac.ir)

Received: 18 July 2025

Accepted: 19 August 2025

Extended Abstract

Background and Objective: The stability and sustainability of ecosystems require understanding the relationships between the distribution of plant species and environmental factors. Species distribution modeling techniques are recognized as powerful tools to support forest management strategies in the context of climate change. Boxwood is a shade-tolerant species that is distributed in the understory of lowland and middle-altitude forests of northern Iran, where it creates a dark environment in the forest by forming a dense canopy. Due to the widespread destruction of boxwood habitats in recent decades, this species has been considered part of conservation communities to prevent its extinction. The main objective of this research is to model the distribution of this species in the Hyrcanian forests.

Material and Methods: In this study, the distribution of boxwood was modeled using 570 actual occurrence points in Hyrcanian forests. Bioclimatic variables were extracted from the WorldClim database and topographic variables from the digital elevation model. VIF test was used to reduce collinearity. 70% of the samples were assigned as training data for model development and the remaining 30% as test data for model validation. Modeling was performed using the random forest algorithm in R environment and using presence and pseudo-absence data, adjusting the model parameters (500 trees, mtry=2, min_n=5) and validating it by using the 10-fold Cross-Validation method. Model performance was evaluated with AUC, TSS, overall accuracy, and kappa coefficient.

Results: The results of the Random Forest model showed that the model has high accuracy in predicting the distribution of boxwood (AUC=0.98, Accuracy=0.95, Kappa=0.72, TSS=0.63). The random forest model, using bioclimatic and topographic variables, showed very favorable performance in predicting the distribution of Hyrcanian boxwood. Evaluation of variable importance indicated that the bioclimatic variables bio3 (isothermality), bio12 (annual precipitation), bio8 (mean temperature of the wettest quarter) and bio1 (annual mean temperature) had the greatest impact on the distribution of the boxwood species. The response curve of boxwood to the four important variables was also plotted. The boxwood habitat suitability map identified

areas with suitable ecological conditions, including large parts of Mazandaran province and the border areas of the Hyrcanian forests.

Conclusion: Analysis of variable importance showed that isothermality (Bio3) was the most effective variable, indicating the role of thermal stability in the viability of boxwood; an evergreen and shade-loving species that is sensitive to temperature fluctuations and is more present in areas with a balanced temperature range (15–20 degrees). Annual precipitation (Bio12) was also of high importance, and the response curve showed that the probability of boxwood presence is much higher in areas with precipitation above 300 mm; because it provides sufficient moisture for growth, photosynthesis, and reduction of water stress. Variable Bio8 (mean temperature of the wettest quarter) by affecting vegetative growth and resistance to fungal diseases, provides optimal conditions for this species in the temperature range of about 5 to 10 degrees. Also, Bio1 (annual mean temperature) showed that boxwood is more distributed in areas with temperate climates and an average annual temperature between 8 and 13 degrees. In contrast, topographic variables such as slope length and solar radiation played a complementary role and moderated microclimatic conditions at the local level. The habitat suitability map also showed that the central and western regions of the Hyrcanian forests, especially in the provinces of Mazandaran and Gilan, have the highest probability of species presence. Also, in some areas of Golestan province, medium to high suitability was predicted, which shows the potential for boxwood restoration in these areas. These results confirm the key role of climate, especially the combination of temperature and humidity, in explaining the distribution of boxwood and emphasize the need to protect habitats with stable climatic conditions.

Keywords: Bioclimatic variable, Habitat suitability, Machine learning, Worldclim database.

How to Cite This Article: Hesabi, A., Alavi, S. J., and Esmailzadeh, O. (2025). Species distribution modeling of Hyrcanian boxwood (*Buxus Hyrcana* Pojark) using the random forest model in the forests of northern Iran. *Forest Research and Development*, 11(3), 353-375. DOI: [10.30466/jfrd.2025.56377.1766](https://doi.org/10.30466/jfrd.2025.56377.1766)



Copyright ©2024 Hesabi et al. Published by Urmia University.

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/), which allows users to read, copy, distribute, and make derivative works for non-commercial purposes from the material, as long as the author of the original work is cited properly.

مدلسازی پراکنش گونه‌ای شمشاد هیرکانی (*Buxus Hyrcana* Pojark) با بهره‌گیری از مدل جنگل تصادفی در جنگل‌های شمال ایران

عارف حسابی^۱، سیدجلیل علوی^{۲*} و امید اسماعیل‌زاده^۳

۱- دانشجوی دکتری علوم و مهندسی جنگل، دانشکده منابع طبیعی و علوم دریایی، دانشگاه تربیت مدرس، نور، ایران. (aref.hesabi@modares.ac.ir)

۲- دانشیار، گروه علوم و مهندسی جنگل، دانشکده منابع طبیعی و علوم دریایی، دانشگاه تربیت مدرس، نور، ایران. (j.alavi@modares.ac.ir)

۳- دانشیار، گروه علوم و مهندسی جنگل، دانشکده منابع طبیعی و علوم دریایی، دانشگاه تربیت مدرس، نور، ایران. (oesmailzadeh@modares.ac.ir)

تاریخ پذیرش: ۱۴۰۴/۰۶/۲۸

تاریخ دریافت: ۱۴۰۴/۰۴/۲۷

چکیده

مقدمه و هدف: ثبات و پایداری بوم‌سازگان‌ها نیازمند شناخت روابط بین پراکنش گونه‌های گیاهی و عوامل محیطی است. تکنیک‌های مدلسازی پراکنش گونه‌ها به‌عنوان ابزار قدرتمندی برای حمایت از استراتژی‌های مدیریت جنگل در زمینه تغییرات اقلیمی شناخته می‌شوند. شمشاد گونه‌ای بردبار به سایه بوده و در زیر آشکوب جنگل‌های جلگه‌ای و میان‌بند شمال ایران پراکنش دارد که با ایجاد تاج‌پوشش انبوه و فشرده، محیطی تاریک در جنگل ایجاد می‌کند. به دلیل تخریب گسترده رویشگاه‌های شمشاد در دهه‌های اخیر، این گونه جزء جوامع ذخیره‌گاهی در نظر گرفته شده است تا از انقراض آن جلوگیری شود. هدف اصلی این پژوهش مدلسازی پراکنش این گونه در جنگل‌های هیرکانی است.

مواد و روش‌ها: در این بررسی، با استفاده از ۵۷۰ نقطه حضور واقعی شمشاد در جنگل‌های هیرکانی، پراکنش این گونه مدلسازی شد. متغیرهای زیست‌اقلیمی از پایگاه WorldClim و متغیرهای توپوگرافی از مدل رقومی ارتفاع استخراج شدند. برای کاهش هم‌خطی، از آزمون VIF استفاده شد. ۷۰ درصد نمونه‌ها به‌عنوان داده‌های آموزشی برای توسعه مدل و ۳۰ درصد باقی‌مانده به‌عنوان داده‌های آزمون برای اعتبارسنجی مدل اختصاص داده شدند. مدلسازی با الگوریتم جنگل تصادفی در محیط R و با بهره‌گیری از داده‌های حضور و شبه غیاب، تنظیم شاخص‌های مدل (۵۰۰ درخت، $mtry=2$ ، $min_n=5$) و اعتبارسنجی متقابل آن با روش (10 fold Cross-Validation) انجام شد. عملکرد مدل با شاخص‌های AUC، TSS، دقت کلی و ضریب کاپا ارزیابی شد.

یافته‌ها: نتایج حاصل از مدل جنگل تصادفی نشان داد که مدل از دقت بالایی در پیش‌بینی پراکنش شمشاد برخوردار است ($AUC=0/98$ ، $Accuracy=0/95$ ، $Kappa=0/72$ ، $TSS=0/63$). مدل جنگل تصادفی با

استفاده از متغیرهای زیست‌اقليمی و توپوگرافی، عملکرد بسیار مطلوبی در پیش‌بینی پراکنش شمشاد هیرکانی داشت. ارزیابی اهمیت متغیرها نشان داد که متغیرهای زیست‌اقليمی bio3 (هم‌دمایی)، bio12 (بارندگی سالانه)، bio8 (میانگین دما در فصل مرطوب) و bio1 (میانگین دمای سالانه) بیشترین تأثیر را در پراکنش گونه شمشاد داشتند. منحنی پاسخ گونه شمشاد نسبت به چهار متغیر مهم نیز رسم شد. نقشه مطلوبیت زیستگاه شمشاد، مناطق با شرایط بوم‌شناختی مناسب را مشخص کرد که بخش‌های گسترده‌ای از استان مازندران و نواحی مرزی جنگل‌های هیرکانی را شامل می‌شود.

نتیجه‌گیری: تحلیل اهمیت متغیرها نشان داد که هم‌دمایی (Bio3) به‌عنوان مؤثرترین متغیر، نشان‌دهنده نقش ثبات دمایی در زیست‌پذیری شمشاد است؛ گونه‌ای همیشه‌سبز و سایه‌پسند که در برابر نوسانات دمایی حساس است و در مناطق با دامنه دمایی متعادل (۱۵-۲۰ درجه) حضور بیشتری دارد. بارندگی سالانه (Bio12) نیز اهمیت بالایی داشت و منحنی پاسخ نشان داد احتمال حضور شمشاد در نواحی با بارندگی بیش از ۳۰۰ میلی‌متر به‌مراتب بیشتر است؛ زیرا رطوبت کافی برای رشد، فتوسنتز و کاهش تنش آبی فراهم می‌سازد. متغیر Bio8 (میانگین دما در فصل مرطوب) با تأثیرگذاری بر رشد رویشی و مقاومت به بیماری‌های قارچی، در بازه دمایی حدود پنج تا ۱۰ درجه شرایط بهینه برای این گونه فراهم می‌کند. همچنین Bio1 (میانگین دمای سالانه) نشان داد که شمشاد بیشتر در مناطقی با اقلیم معتدل و دمای متوسط سالانه بین هشت تا ۱۳ درجه پراکنش دارد. در مقابل، متغیرهای توپوگرافی مثل طول شیب و تابش خورشیدی نقش مکمل داشته و در سطوح محلی شرایط میکروکلیمایی را تعدیل می‌کنند. نقشه مطلوبیت رویشگاه نیز نشان داد که مناطق میانی و غربی جنگل‌های هیرکانی، به‌ویژه در استان‌های مازندران و گیلان، دارای بالاترین احتمال حضور گونه هستند. همچنین در برخی نواحی استان گلستان نیز مطلوبیت متوسط تا بالایی پیش‌بینی شد که قابلیت بالقوه برای احیای شمشاد در این مناطق را نشان می‌دهد. این نتایج نقش کلیدی اقلیم، به‌ویژه ترکیب دما و رطوبت، را در تبیین پراکنش شمشاد تأیید کرده و بر لزوم حفاظت از زیستگاه‌های با شرایط اقلیمی پایدار تأکید می‌کند.

واژه‌های کلیدی: پایگاه اطلاعاتی Worldclim، متغیر زیست‌اقليمی، مطلوبیت رویشگاه، یادگیری ماشین.

مقدمه

شناخت پراکنش جغرافیایی، ویژگی‌های بوم‌شناختی و زیستی گونه‌ها و عواملی که بر این شاخص‌ها تأثیر می‌گذارند، برای برنامه‌ریزی حفاظت و پیش‌بینی اقدام‌های آینده بسیار اساسی است (Rushton et al., 2004) و می‌تواند نقش حیاتی در بازسازی بوم‌شناختی ایفا کند (Martínez-Meyer et al., 2006). این اطلاعات همچنین برای درک عوامل تعیین‌کننده بوم‌شناختی و تکاملی الگوهای پراکنش تنوع زیستی مورد نیاز است (Graham et al., 2006). مدل‌های پراکنش گونه‌ای (Species Distribution Modeling)، که به آنها مدل‌های آشیان بوم‌شناختی نیز گفته می‌شود، روش‌های کمی هستند که داده‌های وقوع یک گونه را با متغیرهای پیشگو، مدل‌های آماری و الگوریتم‌های کامپیوتری ترکیب کرده و مقادیر هر متغیر پیشگو و تأثیر آن بر گونه مورد نظر در هر مکان شناخته شده و همچنین مناطق مناسب را بر اساس این متغیرهای پیشگو، محاسبه و شناسایی می‌کنند (Elith et al., 2009; Ren-). (Yan et al., 2014). مدلسازی پراکنش گونه‌ای یک حوزه پژوهشی و تخصصی با رشد سریع است که به بررسی درباره پیش‌بینی رویشگاه‌های بالقوه گونه‌های مختلف گیاهی و جانوری می‌پردازد، که به‌طور فزاینده‌ای برای پشتیبانی از تصمیم‌گیری در زمینه‌های مختلفی مانند بوم‌شناسی، جغرافیای زیستی، حفظ تنوع زیستی و مدیریت منابع طبیعی استفاده می‌شود (Franklin, 2013; Guisan et al., 2013). دانش بوم‌شناختی جامع از پراکنش احتمالی و رویشگاه‌های مناسب برای گونه‌هایی که جمعیت آنها در محدوده پراکنش آنها کاهش می‌یابد، برای برنامه‌ریزی حفاظت بلندمدت آنها اهمیت فراوانی دارد (Akrim et al., 2017).

تکنیک‌های مدلسازی پراکنش گونه‌ها (SDM) که

در دهه ۱۹۸۰ توسعه یافت، در سال‌های اخیر توجه زیادی را به خود جلب کرده است. این تکنیک‌ها به‌طور فزاینده‌ای به‌عنوان ابزار قدرتمندی برای حمایت از استراتژی‌های مدیریت جنگل در زمینه تغییرات اقلیمی شناخته می‌شوند. پیش‌بینی و درک پراکنش گونه‌های گیاهی برای مدیریت پایدار زمین و برنامه‌ریزی حفاظتی ضروری است (Safdar et al., 2025). این تکنیک اسامی مختلفی مانند مدل مبتنی بر آشیان، مدل رویشگاه، مطلوبیت رویشگاه، مدل آشیان محیطی (Environmental Niche Model)، پوشش اقلیمی و مدل آشیان بوم‌شناختی (Ecological Niche Model) دارد (Sillero et al., 2021, 2023). داده‌هایی که اغلب در ENM ها استفاده می‌شوند، سوابق پراکنش گونه‌های گیاهی به‌صورت موقعیت جغرافیایی همراه با مجموعه داده‌های جغرافیایی (نقشه‌های رستری) هستند که شاخص‌های محیطی و جغرافیایی را که ممکن است بر وقوع گونه‌ها تأثیر بگذارد، نمایش می‌دهند. در حالی که نمونه‌های قبلی SDM بیشتر بر روش‌های رگرسیون متمرکز بودند (Guisan et al., 2002)، در حال حاضر، توجه به استفاده از روش‌های یادگیری ماشین (Machine learning) معطوف شده است (Elith et al., 2006). این موضوع بیشتر به دلیل عملکردهای برآزش انعطاف‌پذیر، روش‌های انتخاب متغیر خودکار، توانایی مدیریت انواع مختلف داده‌ها، و گزارش‌های مکرر از عملکرد پیشگو قدرتمند است (Elith et al., 2006; Olden et al., 2008). در میان انبوهی از روش‌های یادگیری ماشین موجود برای مدلسازی و نقشه‌برداری پراکنش گونه‌ها، مدل جنگل تصادفی (Random Forest) بسیار محبوب شده است (Cutler et al., 2007; Zhang et al., 2019)، زیرا تمایل به پیش‌بینی خوب با حداقل تنظیم شاخص دارد

- (Freeman et al., 2015). مدل جنگل تصادفی (RF) همچنین مفید تلقی می شود زیرا ذاتاً شامل اثرات متقابل در مدل برازش شده است و برآوردهای با اهمیت متغیر به راحتی در دسترس هستند (Probst et al., 2019). البته، اشکالاتی نیز در ارتباط با انعطاف پذیری بالا وجود دارد (Merow et al., 2014) و از آنجایی که مدل جنگل تصادفی کنترل کمی بر شکل عملکرد برازش و سطح اثرات متقابل ارائه می دهد، پتانسیل بیش از حد برازش شدن چنین مدل هایی باید جدی گرفته شود و به طور قوی مورد ارزیابی قرار گیرد (Roberts et al., 2017). پژوهش های قبلی نشان دادند که روش های مدل سازی یادگیری ماشین ممکن است به دلیل توانایی شان در درک پیچیدگی و پاسخ های غیرخطی، عملکرد بهتری نسبت به سایر روش های SDM داشته باشند (Elith et al., 2006). جنگل تصادفی و مدل بیشینه آنتروپی (MaxEnt) رویکردهای یادگیری ماشین کم عمق هستند (Haneczok et al., 2020) که به طور گسترده ای برای پیش بینی های بوم شناختی مورد استفاده قرار گرفته اند (Elith et al., 2006; Guo et al., 2015; Kaky et al., 2017; Mi et al., 2020). جنگل تصادفی یک روش اجماعی (ensemble) است که شامل درخت های طبقه بندی و رگرسیون متعدد است؛ این مدل با موفقیت در پیش بینی پراکنش گونه ها مورد استفاده قرار گرفته است (Burns et al., 2020; Mi et al., 2017). مدل جنگل تصادفی از ساختاری بهره می برد که هزاران درخت را با مجموعه ای از متغیرهای پیشگو تصادفی انتخاب شده ایجاد می کند؛ این امر خطای تعمیم را محدود می کند (Prasad et al., 2006).
- در رابطه با شناسایی رویشگاه های مطلوب، پژوهش هایی در داخل و خارج کشور اجرا شده است. در پژوهشی بر روی گونه حرا (*Avicennia marina*) در جنگل های ماندابی جنوب ایران، Sobhani and
- (Danekar et al., 2024) با استفاده از مدل بیشینه آنتروپی به مدلسازی پراکنش این گونه پرداختند. نتایج نشان داد که این مدل با دقت بالا قادر به تفکیک رویشگاه های مطلوب و نامطلوب است. مهم ترین متغیرهای مؤثر در مدل شامل بیشینه ارتفاع موج، دمای متوسط سالانه، نوسانات جزر و مد و شوری آب بودند. در پژوهشی دیگر، Hedayati Kaliji et al. (2025) با استفاده از الگوریتم های مختلف مانند شبکه عصبی مصنوعی، مدل خطی تعمیم یافته، رگرسیون تطبیقی چندمتغیره، حداکثر آنتروپی و جنگل تصادفی، پراکنش فعلی راش شرقی (*Fagus orientalis* L.) را در جنگل های هیرکانی مدلسازی کردند. نتایج نشان داد که تمامی مدل ها قابلیت شناسایی محدوده پراکنش این گونه را داشتند، اما مدل اجماعی و پس از آن جنگل تصادفی بهترین عملکرد را ارائه دادند. همچنین ارتفاع، تغییرات فصلی دما و شیب به عنوان مهم ترین عوامل مؤثر در پراکنش راش شرقی شناسایی شدند. Habibikilak et al. (2025) با بهره گیری از مدل بیشینه آنتروپی و ترکیب های مختلفی از متغیرها، به شناسایی عوامل مؤثر بر پراکنش سرخدار (*Taxus baccata* L.) در جنگل های هیرکانی پرداختند. نتایج نشان داد مدل دارای متغیرهای اقلیمی و خاکی (مدل ۳) عملکرد ضعیف تری داشته و کمترین مقدار سطح زیر منحنی (Area Under the Curve) ($AUC=0.93$) را به خود اختصاص داد، درحالی که مدل شامل متغیرهای اقلیمی و توپوگرافی (مدل ۲) عملکرد بهتری داشت ($AUC=0.94$). بالاترین دقت مربوط به مدل ۴ با AUC برابر ۰/۹۶ بود که پس از حذف هم خطی بین متغیرها حاصل شد. بر پایه سهم نسبی متغیرها، از میان داده های زیست اقلیمی متغیرهای bio2، bio3، bio7 و bio18 و از میان متغیرهای توپوگرافی، ارتفاع و شیب، به عنوان مهم ترین عوامل تأثیرگذار در پراکنش سرخدار

شناسایی شدند. (Asadi et al., 2025) با استفاده از پنج مدل شامل جنگل تصادفی، ماشین بردار پشتیبان، K نزدیک‌ترین همسایه، مدل خطی تعمیم‌یافته و بیشینه آنتروپی، پراکنش فعلی گونه انجیلی (*Parrotia persica* C.A.Mey) را در جنگل‌های هیرکانی بررسی کردند. نتایج نشان داد مدل جنگل تصادفی با مقدار AUC برابر ۰/۸۷، بهترین عملکرد را داشت. تحلیل اهمیت متغیرها نشان داد ارتفاع از سطح دریا مؤثرترین عامل در پراکنش انجیلی است، درحالی‌که درصد شن خاک و مقدار بارش به ترتیب توسط مدل‌های جنگل تصادفی و حداکثر آنتروپی به عنوان دومین متغیر مهم شناسایی شدند. پژوهشی که بر روی گونه *Bergenia stracheyi* که بومی مناطق هیمالیا است انجام شد، نشان داد که در درجه اول میانگین دمای سالانه (Bio1) و بارندگی سالانه (Bio12) متغیرهای تاثیرگذار بر پراکنش این گونه هستند. پیش‌بینی‌های مدل گروهی نشان داد که در شرایط اقلیمی فعلی، زیستگاه‌های مناسب برای *Bergenia stracheyi* در ارتفاعات بالاتر جامو و کشمیر توزیع شده‌اند و پیش‌بینی‌های مدل گروهی آینده نشان می‌دهد که در تمام سناریوهای اقلیمی آینده، بیشتر رویشگاه‌های مناسب فعلی برای این گونه مناسب باقی خواهند ماند (Wani et al., 2025). در پژوهشی دیگر، متغیرهای تاثیرگذار بر پراکنش گونه *Bergenia ciliata* معرفی شد. نتایج نشان داد شیب (۲۰/۶ درصد)، بارش خشک‌ترین فصل (۱۳/۴ درصد)، بارش گرم‌ترین فصل (۱۱/۷ درصد) و میانگین دمای خشک‌ترین فصل (۱۰/۲ درصد) تاثیر زیادی در پراکنش گونه مذکور داشتند. مناطق با مطلوبیت بالا، متوسط و پایین تحت سناریوی اقلیمی فعلی به ترتیب در حدود ۳۹۲۷۹ کیلومتر مربع، ۳۸۱۲۴ کیلومتر مربع و ۱۱۶۸۱۱ کیلومتر مربع توزیع شدند (Singh et al., 2025). در پژوهشی که بر روی لاله واژگون

(*Fritillaria imperialis* L.) انجام شد، مدل جنگل تصادفی قوی‌ترین پیش‌بینی توزیع لاله واژگون را ارائه داد (Amindin et al., 2024). پژوهشی دیگر با بهره‌گیری از مدل‌های جنگل تصادفی و ماشین بردار پشتیبان (Support Vector Machines) بر روی گونه سیاه‌شور مصری (*Suaeda aegyptiaca*) در بوشهر انجام شد. نتایج نشان داد که منحنی مشخصه عملکرد گیرنده (Receiver Operating Characteristic) اعتبارسنجی منجر به AUC بالاتری برای مدل جنگل تصادفی با مقدار ۰/۹۶ شد، درحالی‌که AUC برای مدل ماشین بردار پشتیبان برابر با ۰/۸۸ بود. این نتیجه، دقت پیش‌بینی بالای مدل جنگل تصادفی را برای تعیین مناسب بودن رویشگاه سیاه‌شور نشان داد. تأثیرگذارترین عوامل محیطی، ارتفاع و شیب بودند (Edrisnia et al., 2025).

گونه شمشاد هیرکانی یکی از معدود درختان همیشه سبز پهن‌برگ جنگل‌های هیرکانی است که طول عمر بالایی دارد و جزء گونه‌های در حال انقراض معرفی شده است، اما امروزه وقوع بیماری قارچی سوختگی برگ شمشاد و طغیان آفت شب‌پره شمشاد وضعیت حفاظت شمشاد هیرکانی در سطح جنگل‌های شمال را در وضعیت بحرانی‌تری قرار داده است (Khabazi et al., 2019). از این رو، حفظ و حراست این گونه و شناخت مهم‌ترین ویژگی‌های بوم‌شناختی موثر در پراکنش مکانی آن، از اهمیت بالایی برخوردار است و چون شناخت دقیقی نسبت به ویژگی‌های بوم‌شناختی و دامنه بردباری این گونه وجود ندارد، هدف اصلی این پژوهش مدلسازی پراکنش این گونه در جنگل‌های هیرکانی با استفاده از مدل جنگل تصادفی است.

مواد و روش‌ها

منطقه پژوهش

هیرکانی از سطح دریا آغاز و تا ارتفاع حداکثر ۲۸۰۰ متر از سطح دریا ادامه دارد. میانگین بارندگی در این جنگل‌ها بین ۵۳۰ میلی‌متر در شرق، ۱۵۳۰ میلی‌متر در غرب است که بارندگی در غرب این جنگل‌ها در اسالم حتی به ۲۰۰۰ میلی‌متر نیز می‌رسد. در طول دهه گذشته میانگین دمای سالانه در غرب جنگل‌های هیرکانی ۱۵ درجه سانتی‌گراد و در شرق آن ۱۷/۵ درجه سانتی‌گراد بوده است (Sagheb Talebi et al., 2014).

جنگل‌های هیرکانی به صورت نوار باریک و بلندی روی دامنه‌های شمالی سلسله جبال البرز دیده می‌شوند، که از آستارا در غرب شروع و تا گلی‌داغی واقع در استان بجنورد در شرق ادامه دارند. این جنگل‌ها دارای طول حدود ۸۰۰ کیلومتر و عرض ۱۱۰ کیلومتر هستند و مساحت کلی آن‌ها ۱/۸۵ میلیون هکتار است که ۱۵ درصد از کل جنگل‌های ایران و ۱/۱ درصد از کل مساحت ایران را در بر می‌گیرد (شکل ۱) (Moghbel et al., 2023). گسترش عمودی جنگل‌های



شکل ۱- منطقه پژوهش

Figure 1. Research area

است. تمامی داده‌های حضور به وسیله دستگاه موقعیت‌یاب جهانی ثبت شد و سپس با استفاده از تابع (thindata) در بسته SDMtune در نرم افزار R به مقیاس یک کیلومتر تنک شد (R Core Team, 2024; Vignali et al., 2020) (شکل ۱). این کار برای جلوگیری از اریبی ناشی از نمونه‌برداری داده‌ها انجام شد (Kougioumoutzis et al., 2024). بسته نرم‌افزاری

داده‌های حضور

در این پژوهش از داده‌های حضور شمشاد که تمامی رویشگاه‌های این گونه را در جنگل‌های هیرکانی شامل می‌شود، استفاده شد (Khabazi et al., 2019). این داده‌ها نقاط حضور رویشگاه‌های شمشاد در سه استان شمالی کشور بوده و شامل ۱۶ رویشگاه و تعداد ۵۷۰ نقطه حضور واقعی توده شمشاد در جنگل‌های هیرکانی

SDMtune در هر پیکسل (که اندازه آن یک کیلومتر مربع است) یک نقطه حضور را در نظر گرفته و سایر نقاطی که در پیکسل باشند را حذف می‌کند. بدین صورت که در هر پیکسل فقط یک نقطه حضور باقی می‌ماند. بنابراین تعداد نقاط حضور پس از تنک کردن به تعداد ۱۱۸ نقطه حضور کاهش یافت.

داده‌های محیطی

در این پژوهش از متغیرهای زیست‌اقلیمی پایگاه اطلاعاتی WorldClim

جدول ۱- معرفی متغیرهای زیست‌اقلیمی پایگاه جهانی WorldClim

Table 1. Introduction to WorldClim global bioclimatic variables

نام اختصاری Abbreviated name	متغیر زیست‌اقلیمی Bioclimatic variable	نام اختصاری Abbreviated name	متغیر زیست‌اقلیمی Bioclimatic variable
Bio1	میانگین دمای سالانه Mean Annual Temperature	Bio11	میانگین دما در سردترین فصل Mean Temperature of Coldest Quarter
Bio2	میانگین دامنه دمای روزانه Mean Diurnal Range	Bio12	مجموع بارندگی سالانه Annual Precipitation
Bio3	هم دمایی Isothermality	Bio13	بارندگی در مرطوب‌ترین ماه Precipitation of Wettest Month
Bio4	تغییرات فصلی دما Temperature Seasonality	Bio14	بارندگی در خشک‌ترین ماه Precipitation of Driest Month
Bio5	حداکثر دما در گرم‌ترین ماه Max Temperature of Warmest Month	Bio15	تغییرات بارندگی فصلی Precipitation Seasonality
Bio6	حداقل دما در سردترین ماه Min Temperature of Coldest Month	Bio16	بارندگی در مرطوب‌ترین فصل Precipitation of Wettest Quarter
Bio7	تغییرات دمای سالانه Temperature Annual Range (BIO5-BIO6)	Bio17	بارندگی در خشک‌ترین فصل Precipitation of Driest Quarter
Bio8	میانگین دما در مرطوب‌ترین فصل Mean Temperature of Wettest Quarter	Bio18	بارندگی در گرم‌ترین فصل Precipitation of Warmest Quarter
Bio9	میانگین دما در خشک‌ترین فصل Mean Temperature of Driest Quarter	Bio19	بارندگی در سردترین فصل Precipitation of Coldest Quarter
Bio10	میانگین دما در گرم‌ترین فصل Mean Temperature of Warmest Quarter		

vifstep (2013). این پردازش آماری با استفاده از تابع usdm در بسته در نرم افزار آماری R انجام شد (Naimi et al., 2014; R Core Team, 2024) (جدول ۲).

روش مدلسازی

در این پژوهش، برای مدلسازی پراکنش گونه شمشاد (*Buxus hyrcana*) در مناطق جنگلی شمال ایران، از الگوریتم یادگیری ماشین جنگل تصادفی استفاده شد. این الگوریتم، که بر پایه مجموعه‌ای از درختان تصمیم‌گیری عمل می‌کند، به دلیل توانایی بالا در مدیریت داده‌های پیچیده و غیرخطی، برای مدلسازی پراکنش گونه‌های گیاهی مناسب است. پیاده‌سازی مدل با بهره‌گیری از بسته tidymodels در محیط نرم‌افزاری R نسخه ۴.۲.۱ انجام شد (Kuhn and Wickham, 2020). بسته tidymodels چارچوبی یکپارچه برای مدلسازی و ارزیابی فراهم می‌آورد و امکان تنظیم شاخص‌ها و اعتبارسنجی را تسهیل می‌کند.

مدل جنگل تصادفی نیازمند داده‌های حضور و عدم حضور گونه است. داده‌های حضور واقعی از منابع میدانی و پایگاه‌های داده معتبر جمع‌آوری شدند. برای تولید داده‌های عدم حضور، از روش (Environmental constraint) در بسته flexsdm استفاده شد (Velazco et al., 2022). این روش، برخلاف رویکردهای تصادفی ساده، نقاط عدم حضور را تنها در مناطقی تولید می‌کند که از نظر زیست‌محیطی با نقاط حضور واقعی تفاوت قابل توجهی دارند، که این امر به کاهش اریبی مدل و بهبود دقت پیش‌بینی کمک می‌کند. تعداد نقاط عدم حضور برابر با تعداد نقاط حضور تنظیم شد تا تعادل در داده‌ها حفظ شود. سپس، داده‌های حضور و عدم حضور ترکیب شدند تا مجموعه داده نهایی تشکیل شود. برای استخراج مقادیر متغیرهای محیطی (شامل متغیرهای زیست‌اقليمی و توپوگرافی، همان‌طور که در جدول ۲

برای تهیه متغیرهای توپوگرافی از مدل رقومی ارتفاع (Digital Elevation Model) تهیه شده از مأموریت توپوگرافی راداری شاتل (Shuttle Radar Topography Mission) استفاده شد. بدین منظور از نرم‌افزار SAGA (نسخه ۹.۲) برای استخراج متغیرهای اولیه و ثانویه توپوگرافی بهره گرفته شد. سه متغیر اصلی ارتفاع از سطح دریا، شیب و جهت جغرافیایی در نظر گرفته شد و متغیرهای ثانویه هم شامل شاخص رطوبت توپوگرافی (Topographic Wetness Index)، شاخص همگرایی (Convergence Index)، عمق دره (Valley Depth)، فاصله عمودی تا آبراه (Channel Network Length and Distance)، طول شیب و ضریب شیب (Steepness Factor) و شاخص موقعیت توپوگرافی (Topography Position Index) بود. برای تهیه متغیر شاخص تابش خورشیدی (Topographic Solar Radiation Aspect Index) از متغیر جهت جغرافیایی طبق رابطه (۱) استفاده شد. در این رابطه θ مقدار آزمون جهت برحسب درجه است. در صورت عدم وجود جهت (مناطق مسطح $= -1$)، مقدار TRASP برابر با $0/5$ قرار داده شد (Moisen et al., 2002). مقدار شاخص تابش خورشیدی از صفر تا یک متغیر است، به‌طوری که عدد یک نشان‌دهنده جهت‌های خشک و گرم (جنوب و جنوب غربی) و عدد صفر نشان‌دهنده جهت‌های خنک و مرطوب (شمال و شمال شرقی) است (Aerts et al., 2010).

$$\text{رابطه (۱)} \quad \text{TRASP} = [1 - \cos((\pi/180) (\theta - 30))] / 2$$

با توجه به چندخطی بودن بالقوه بین متغیرهای پیش‌بینی‌کننده، از تحلیل عامل تورم واریانس (Variance Inflation Factor) استفاده شد. در اینجا، پیشگوهای حفظ‌شده همگی دارای هم‌خطی کمتر از ۱۰ بودند (Ahmadi et al., 2023; Dormann et al., 2023).

فهرست شده است) برای هر نقطه حضور یا عدم حضور، از تابع sdm_extract در بسته flexsdm بهره گرفته شد (Barbet-Massin et al., 2012). این تابع، داده‌های نقطه‌ای را با لایه‌های رستری متغیرهای محیطی پیوند می‌دهد و ویژگی‌های مربوطه را استخراج می‌کند.

این مرحله حیاتی است: زیرا مدل را قادر می‌سازد تا الگوهای زیست‌محیطی مرتبط با حضور گونه را شناسایی کند و شرایطی را که احتمال حضور گونه را افزایش می‌دهد، بیاموزد.

جدول ۲- متغیرهای مورد استفاده در ساخت مدل

Table 2. Variables used in building the model

نام اختصاری	توضیح	واحد
Abbreviated name	Explanation	Unit
Bio 1	میانگین دمای سالانه	درجه سانتی‌گراد
	Mean Annual Temperature	°C
Bio 3	هم‌دمایی	-
	Isothermality	
Bio 4	موسمیت دما	-
	Temperature Seasonality	
Bio 8	میانگین دما در مرطوب‌ترین فصل	درجه سانتی‌گراد
	Mean Temperature of Wettest Quarter	°C
Bio 12	مجموع بارندگی سالانه	میلی‌متر
	Annual Precipitation	mm
Bio 15	تغییرات بارندگی فصلی	-
	Precipitation Seasonality	
TRASP	شاخص تابش خورشیدی	-
	Topographic solar radiation aspect index	
V_D	عمق دره	-
	Valley Depth	
TWI	شاخص رطوبت توپوگرافی	-
	Topographic Wetness Index	
TPI	شاخص موقعیت توپوگرافی	-
	Topography Position Index	
C_I	شاخص همگرایی	-
	Convergence Index	
LS	طول و ضریب شیب	-
	Length and Steepness Factor	
CH_N_D	فاصله عمودی تا آبراه	-
	Channel Network Distance	

پس از آماده‌سازی داده‌ها، مجموعه داده به دو بخش آموزش (۷۰ درصد) و آزمون (۳۰ درصد) تقسیم شد تا عملکرد مدل بر روی داده‌های دیده‌نشده ارزیابی شود. برای بررسی پایداری عملکرد مدل و جلوگیری از پدیده بیش‌برازش، از روش اعتبارسنجی متقابل k-fold (James et al., 2011) استفاده شد.

آموزشی به k زیرمجموعه تقسیم می‌شوند و در هر تکرار، یکی از زیرمجموعه‌ها برای اعتبارسنجی و بقیه برای آموزش مدل به کار گرفته می‌شود. این رویکرد، برآورد پایدارتری از عملکرد مدل ارائه می‌دهد و قابلیت تعمیم‌پذیری آن را بهبود می‌بخشد. در این پژوهش، از اعتبارسنجی متقابل ۱۰ تایی (10-fold cross validation) بهره گرفته شد، که تعادل مناسبی بین دقت و کارایی محاسباتی برقرار می‌کند.

در مدل جنگل تصادفی، شاخص‌های کلیدی شامل تعداد درختان، تعداد متغیرهای انتخاب‌شده در هر تقسیم (mtry) و حداقل تعداد نمونه‌ها در هر گره پایانی (min_n) هستند. برای بهینه‌سازی مدل، فرآیند تنظیم شاخص‌ها (tuning) انجام شد. این تنظیمات، تعادلی بین دقت مدل و جلوگیری از پیچیدگی بیش‌ازحد برقرار می‌کنند و عملکرد بهینه را تضمین می‌نمایند. در این پژوهش تعداد درخت‌ها، تعداد متغیرهای انتخاب‌شده در هر تقسیم و حداقل تعداد نمونه در هر گره به ترتیب ۵۰۰، دو و پنج تنظیم شد.

اعتبارسنجی مدل

برای ارزیابی عملکرد مدل جنگل تصادفی، از چندین معیار رایج و معتبر در مدلسازی استفاده شد. یکی از این معیارها منحنی مشخصه عملکرد گیرنده (ROC) و مساحت زیر این منحنی است. منحنی ROC رابطه بین نرخ مثبت‌های واقعی (sensitivity) و نرخ مثبت‌های کاذب (1-specificity) را برای آستانه‌های مختلف نشان می‌دهد و عملکرد کلی مدل را در تفکیک داده‌های حضور و عدم‌حضور بررسی می‌کند. AUC عددی بین ۰/۵ تا یک است که مقادیر نزدیک به یک نشان‌دهنده توان تفکیک بالای مدل است (Phillips et al., 2008). علاوه بر آن، از شاخص مهارت واقعی (True Skill Statistic) نیز برای ارزیابی مدل استفاده شد. این

شاخص ترکیبی از حساسیت و ویژگی است که با فرمول $TSS = Sensitivity + Specificity - 1$ تعریف می‌شود (Allouche et al., 2006). مقدار TSS بین -۱ و +۱ متغیر است که مقدار نزدیک به یک نشان‌دهنده عملکرد بسیار خوب مدل است. شاخص دیگر برای ارزیابی صحت کلی پیش‌بینی‌های مدل، شاخص دقت کلی (Accuracy) است. این شاخص، نسبت تعداد نمونه‌هایی را که مدل به درستی آن‌ها را در دسته حضور یا عدم‌حضور پیش‌بینی کرده، به کل نمونه‌ها نشان می‌دهد. مقدار آن بین صفر و یک قرار دارد که هرچه به یک نزدیک‌تر باشد، نشان‌دهنده عملکرد بهتر مدل در دسته‌بندی صحیح مشاهدات است. با این حال، دقت کلی می‌تواند در صورت وجود عدم‌توازن شدید بین دسته‌ها (مثل تعداد زیاد نمونه‌های عدم حضور نسبت به حضور) گمراه‌کننده باشد، چرا که ممکن است مدلی با پیش‌بینی همه نمونه‌ها در دسته غالب، مقدار دقت بالایی نشان دهد ولی در پیش‌بینی دسته اقلیت (مثل حضور گونه) ضعیف باشد. به همین دلیل، در این پژوهش دقت کلی به همراه شاخص‌های مکمل مانند ضریب کاپا و شاخص TSS به کار گرفته شد تا عملکرد مدل به صورت جامع‌تری ارزیابی شد (Allouche et al., 2006; Fielding et al., 1997). و تحلیل‌ها در زبان برنامه نویسی R انجام گرفت (R Core Team, 2024).

نتایج

عملکرد مدل

برای ارزیابی عملکرد مدل از شاخص‌های معتبر شامل شاخص دقت کلی، ضریب کاپا، سطح زیر منحنی مشخصه عملکرد گیرنده و شاخص TSS استفاده شد. در این پژوهش، مقدار AUC معادل ۰/۹۸، مقدار کاپا ۰/۷۲، مقدار TSS برابر ۰/۶۳ و مقدار Accuracy برابر

طبقه‌بندی صحیح داده‌ها می‌باشد. همچنین، ارزش مقدار کاپا ۰/۷۲ محاسبه شد که بیانگر توافق قابل قبول میان پیش‌بینی مدل و واقعیت میدانی است، حتی در شرایطی که احتمال پیش‌بینی تصادفی وجود دارد. در ادامه، مقدار TSS برابر ۰/۶۳ به دست آمد که نشان می‌دهد مدل دارای تعادل مناسبی میان حساسیت (Sensitivity) و ویژگی (Specificity) است. این شاخص به‌ویژه برای حذف اثرات منفی عدم توازن در داده‌های حضور و عدم حضور اهمیت دارد. در نهایت، مقدار Accuracy برابر ۰/۹۵ محاسبه شد که نمایانگر نرخ بسیار بالای طبقه‌بندی صحیح در کل مجموعه داده‌ها است.

۰/۹۵ محاسبه شد که نشان‌دهنده عملکرد مطلوب مدل در تفکیک حضور و عدم حضور گونه‌ی شمشاد هیرکانی است (جدول ۳). منحنی ROC نیز که در شکل ۲ ارائه شده، به‌وضوح نشان می‌دهد که مدل از قدرت تمایز بالایی میان نقاط حضور و عدم حضور برخوردار است. شیب تند منحنی ROC و فاصله زیاد آن از خط تشخیص تصادفی (قطر نمودار) تأییدکننده این موضوع است. در نتایج حاصل از اجرای مدل، مقدار AUC معادل ۰/۹۸، به دست آمد که نشان‌دهنده‌ی توان بسیار بالای مدل در تمایز میان نقاط حضور و عدم حضور گونه شمشاد است. مقدار AUC نزدیک به یک، به‌عنوان معیاری استاندارد، حاکی از عملکرد ممتاز مدل در

جدول ۳- عملکرد مدل به‌دست‌آمده از چهار شاخص ارزیابی

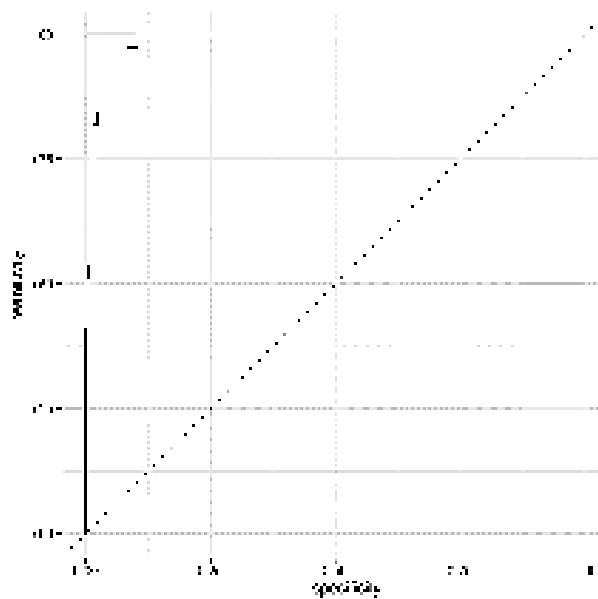
Table 3. Model performance obtained from four evaluation indicators

مدل جنگل تصادفی	معیارهای ارزیابی
Random Forest model	Evaluation criteria
0.98	سطح زیر منحنی Area Under the Curve
0.63	آماره مهارت درست True Skill Statistic
0.95	شاخص دقت کلی Accuracy
0.72	ضریب کاپا Kappa

گروه‌های درخت تصمیم‌گیری عمل می‌کند و به هر متغیر مقدار عددی اختصاص می‌دهد که بیانگر سهم آن در بهبود تصمیم‌گیری مدل است. با توجه به شکل ۳ متغیرهای اقلیمی (به‌ویژه Bio3 و Bio12) سهم عمده‌ای در تعیین پراکنش گونه داشته‌اند، در حالی‌که متغیرهای توپوگرافی مانند (LS_factor) نیز به‌صورت مکمل در بهبود عملکرد مدل مؤثر بوده‌اند.

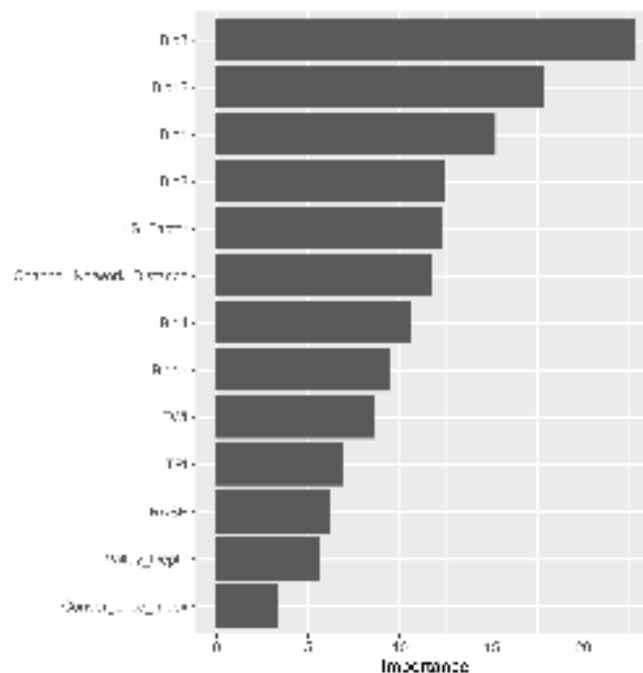
تعیین اهمیت متغیرها

شکل ۳ نتایج اهمیت نسبی هر متغیر در مدلسازی پراکنش گونه‌ای شمشاد را نشان می‌دهد. برای تعیین اهمیت نسبی متغیرهای محیطی در پیش‌بینی پراکنش شمشاد، از شاخص "کاهش میانگین ناخالصی" (Mean Decrease in Impurity) استفاده شد. این شاخص که در مدل‌های جنگل تصادفی رایج است، بر اساس مقدار کاهش ناخالصی مانند شاخص جینی (Gini index) در



شکل ۲- منحنی ROC مدل جنگل تصادفی در پیش‌بینی حضور گونه شمشاد

Figure 2. ROC curve of the random forest model in predicting the presence of boxwood species



شکل ۳- اهمیت نسبی متغیرها با بهره‌گیری از روش کاهش میانگین ناخالصی (چهار متغیر مهم، هم‌دمایی (bio3)، بارندگی سالانه (bio12)، دمای سالانه (bio1) و میانگین دما در فصل مرطوب (bio8))

Figure 3. Relative importance of variables using the Mean Decrease in Impurity method (4 important variables, isothermality (bio3), annual rainfall (bio12), annual temperature (bio1) and average temperature in the wet season (bio8))

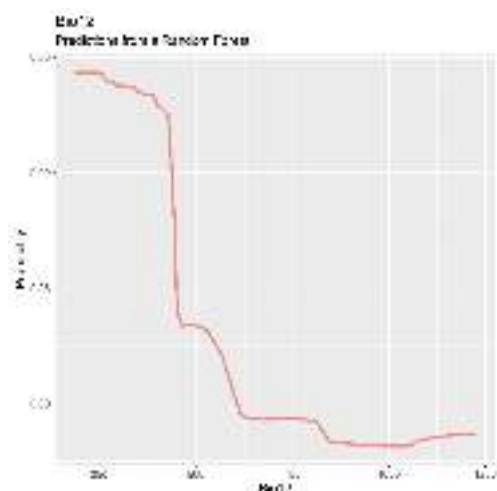
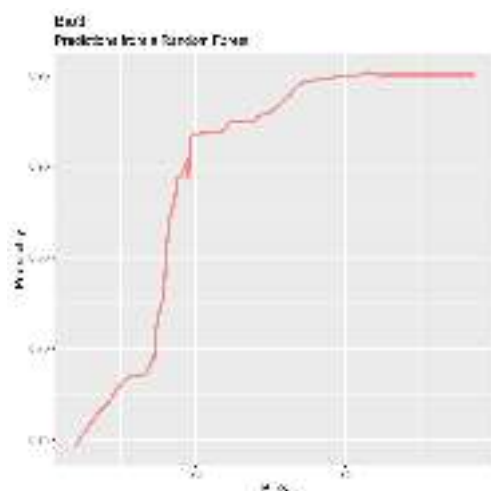
منحنی پاسخ شمشاد نسبت به متغیرها
با توجه به اهمیت نسبی متغیرها، چهار متغیر موثر بر پراکنش شمشاد مشخص شد و منحنی پاسخ گونه شمشاد نسبت به آن‌ها رسم شد (شکل ۴). منحنی واکنش گونه به بارندگی سالانه (Bio12) یک الگوی غیرخطی و به‌شدت نامتقارن را نشان داد. بالاترین احتمال وقوع گونه در مقادیر پایین‌تر بارندگی، در حدود ۲۵۰-۳۰۰ میلی‌متر ثبت شد. پس از این مقدار، با

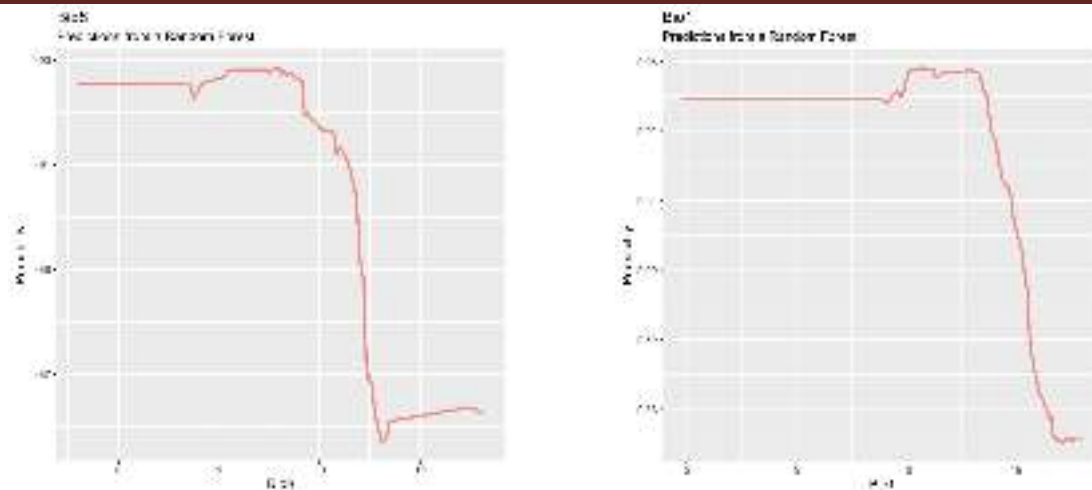
از حساسیت گونه به دماهای بالا در دوره‌های مرطوب سال است. منحنی واکنش گونه به دمای میانگین سالانه (Bio1) یک الگوی واکنش نامتقارن و یکنواخت دارد. بیشینه احتمال وقوع گونه در محدوده حدود ۱۴ تا ۱۶ درجه سانتی‌گراد ثبت شد. در مقادیر پایین‌تر از این بازه، منحنی شیب ملایم و افزایشی داشت، اما در مقادیر بالاتر از ۱۶ درجه، منحنی با کاهش سریع و ناگهانی در احتمال حضور همراه بود. این رفتار نشان‌دهنده آستانه حرارتی بالای گونه شمشاد است.

نقشه رویشگاه مطلوب گونه شمشاد

نقشه مطلوبیت رویشگاه گونه شمشاد با بهره‌گیری از متغیرهای زیست‌اقلیمی، توپوگرافی و داده‌های مکانی حضور و عدم‌حضور تهیه شد (شکل ۵). سطوح مطلوبیت رویشگاه به سه طبقه شامل مطلوبیت بالا (بیش از ۰/۷۵)، مطلوبیت متوسط (بین ۰/۵۰ تا ۰/۷۵) و مطلوبیت پایین (بین ۰/۲۵ تا ۰/۵۰) تقسیم‌بندی شد. یافته‌های حاصل از مدلسازی نشان می‌دهد که نواحی با بالاترین درجه مطلوبیت برای استقرار شمشاد بیشتر در استان مازندران و ناحیه هیرکانی مرزی واقع شده‌اند.

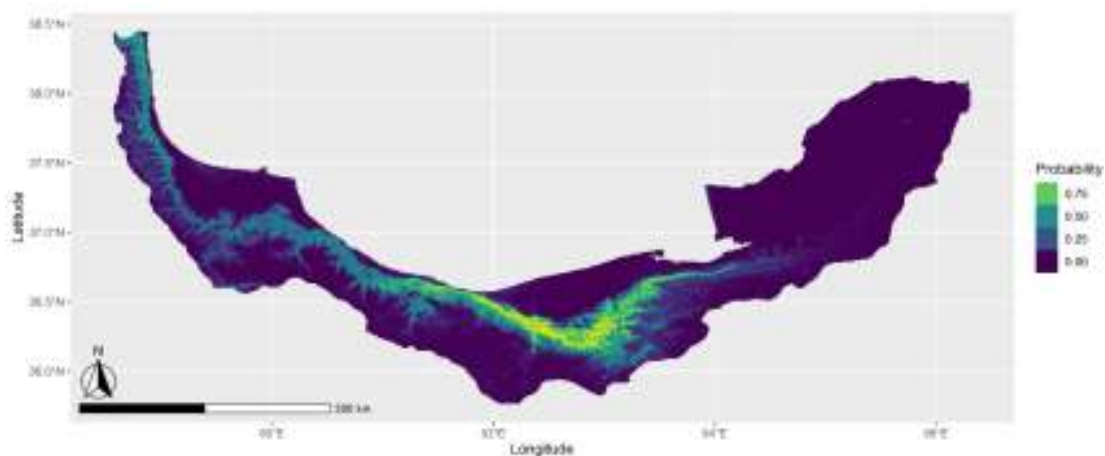
افزایش بارندگی، یک کاهش سریع و ناگهانی در احتمال حضور مشاهده شد. این کاهش در محدوده ۴۰۰ تا ۷۰۰ میلی‌متر شدیدتر بود و سپس در مقادیر بالاتر از ۸۰۰ میلی‌متر به یک ثبات نسبی در سطوح پایین احتمال رسید. منحنی واکنش به هم‌دمایی (Bio3) روندی افزایشی یکنواخت و به‌نسبت متقارن را نشان داد. در مقادیر پایین هم‌دمایی (حدود ۲۵)، احتمال وقوع به-نسبت پایین بود و با افزایش مقدار Bio3، احتمال حضور گونه نیز به‌طور مداوم افزایش یافت، تا آنکه در مقادیر بالای ۴۰ به بیشینه خود رسید. گونه شمشاد در مناطقی با ثبات دمایی بالا در طول شبانه‌روز و نوسان ملایم فصلی عملکرد بهتری داشت. منحنی واکنش گونه به متغیر Bio8 نیز الگوی یکنواخت با افت ناگهانی در آستانه مشخص را نشان داد. در مقادیر پایین‌تر (حدود پنج درجه سانتی‌گراد)، احتمال وقوع در بالاترین سطح خود باقی ماند. با افزایش تدریجی دما تا حدود ۱۰ درجه سانتی‌گراد، منحنی تقریباً پایدار بود؛ اما پس از آن یک کاهش شدید و ناگهانی در احتمال حضور گونه رخ داد که به‌وضوح نشان‌دهنده آستانه حرارتی گونه در فصل مرطوب است. به‌ویژه، در دماهای بالای ۱۲ درجه سانتی‌گراد، احتمال حضور کم می‌شود. این الگو حاکی





شکل ۴- منحنی پاسخ شمشاد نسبت به چهار متغیر محیطی موثر بر اساس کاهش میانگین ناخالصی (سمت چپ بالا bio3، سمت راست بالا bio12، سمت چپ پایین bio8، سمت راست پایین bio1)

Figure 4. Boxwood response curve to 4 effective environmental variables based on the Mean Decrease in Impurity method (upper left bio3, upper right bio12, lower left bio8, lower right bio1)



شکل ۵- نقشه مطلوبیت رویشگاه گونه شمشاد در جنگل‌های هیرکانی با مدل جنگل تصادفی

Figure 5. Habitat suitability map of boxwood species in Hyrcanian forests with the Random Forest model

بحث

تحلیل عملکرد مدل و اهمیت نسبی متغیرها

مدل جنگل تصادفی در این پژوهش توانست عملکرد بسیار مطلوبی از خود نشان دهد. چنین عملکرد قوی ناشی از توانایی این الگوریتم در مدلسازی روابط پیچیده و غیرخطی بین متغیرهای محیطی و توپوگرافی است. مقادیر شاخص‌های ارزیابی شامل $AUC = 0.98$ و $TSS = 0.63$ نشان از دقت بالا در تشخیص حضور و عدم حضور گونه شمشاد هیرکانی در جنگل‌های شمال

ایران است، به علاوه، تنوع متغیرهای استفاده شده (اقلیمی و توپوگرافی) که از منابع معتبر مانند WorldClim و DEM استخراج شده‌اند، موجب تقویت ویژگی تفکیکی مدل شده است (Hijmans et al., 2012). بررسی‌های متعددی در جنگل‌های هیرکانی نشان داده‌اند که ترکیب متغیرهای دما، بارش و توپوگرافی با مدل‌های پیچیده مانند RF عملکرد بسیار بالایی AUC بیش از ۰/۹ دارند (Alipour et al., 2023; Habibikilak et al., 2025; Sagheb Talebi et al.,

منابع آبی برای فرایندهای بوم‌شناسی مانند جوانه‌زنی، رشد رویشی و سازگاری بلندمدت گونه به‌شمار می‌رود. این پژوهش نشان داد که حضور شمشاد در نواحی با بارندگی سالانه متوسط (حدود ۳۰۰ تا ۴۰۰ میلی‌متر) بیشتر است، که با نتایج گزارش‌شده توسط Soleymanipour and Esmailzadeh (2015) هم‌راستا است. علاوه بر این، کاهش بارندگی سالانه می‌تواند سبب افت رطوبت خاک، کاهش ظرفیت تبادل گازی برگ‌ها و افزایش تنش آبی شود، که در نهایت پراکنش این گونه را محدود خواهد کرد. متغیر Bio8 نیز به‌عنوان یکی از متغیرهای کلیدی شناخته شد. این شاخص ترکیبی از عوامل دما و رطوبت را در حساس‌ترین دوره فنولوژیک گونه‌ها در بر می‌گیرد. شمشاد در دوره‌های مرطوب سال، رشد رویشی خود را آغاز می‌کند و در صورتی که دما از حدود بهینه فراتر رود، با کاهش زنده‌مانی و ریسک ابتلا به بیماری‌هایی نظیر بلایت شمشاد (*Cylindrocladium buxicola*) مواجه می‌شود (Khalique et al., 2024). در مناطق با دمای مناسب در فصل مرطوب (بین ۱۴ تا ۲۰ درجه سانتی‌گراد)، احتمال حضور و موفقیت زادآوری شمشاد بیشتر است. بنابراین، bio8 می‌تواند به‌طور غیرمستقیم نشانه‌ای از سلامت رویشگاه نیز باشد. متغیر Bio1 نمایانگر اقلیم حرارتی کلی منطقه است. شمشاد هیرکانی به‌عنوان گونه‌ای با پراکنش در ارتفاعات پایین و نواحی ساحلی خزر، نیاز به دمای متوسط سالانه به‌نسبت ملایم دارد. به‌عنوان مثال پژوهش‌های جدید نشان داده‌اند که گونه‌های جنگلی با رویشگاه پایین‌دست هیرکانی دمای میانگین سالانه حدود ۱۴-۱۵ درجه سانتی‌گراد را ترجیح می‌دهند، که با شرایط اقلیمی عمده نواحی زیرآشکوب شمشاد همخوان است (Sękiewicz et al., 2024). نکته قابل تأمل دیگر، نقش ثانویه‌ی متغیرهای

این نتایج نشان می‌دهد که ساختار پیچیده و انعطاف‌پذیر جنگل تصادفی توانسته الگوهای غیرخطی بین متغیرهای زیست‌اقلیمی، توپوگرافی و حضور گونه را به‌خوبی یاد بگیرد و تعمیم دهد (Allouche et al., 1997; Fielding et al., 2006).

تحلیل اهمیت متغیرها با استفاده از شاخص کاهش میانگین ناخالصی نشان داد که چهار متغیر اصلی شامل هم‌دمایی (Bio3)، بارندگی سالانه (Bio12)، میانگین دما در مرطوب‌ترین فصل (Bio8) و میانگین دمای سالانه (Bio1) نقش برجسته‌ای در پیش‌بینی پراکنش شمشاد ایفا کرده‌اند. متغیر Bio3، که نسبت بین نوسان دمای شبانه‌روزی به نوسان دمای سالانه را نشان می‌دهد، بیشترین اهمیت نسبی را در مدل به خود اختصاص داد. این شاخص، نشان‌دهنده ثبات حرارتی محیط است. پژوهش‌های مختلف نشان داده‌اند که شمشاد در مناطق با ثبات حرارتی بهتر رشد می‌کند (Esmailzadeh et al., 2015; Habibi kilak et al., 2019). شمشاد هیرکانی گونه‌ای سایه‌پسند، همیشه‌سبز و حساس به نوسانات شدید دمایی است و نواحی با دامنه دمایی معتدل‌تر و ثبات بالاتر حرارتی را ترجیح می‌دهد (Sagheb Talebi et al., 2014). یافته‌های این پژوهش نیز نشان داد که احتمال حضور شمشاد در مناطقی با هم‌دمایی متوسط تا بالا، به‌ویژه در نواحی میانی ارتفاعی رشته کوه البرز (با نوسانات کم بین شب و روز)، بیشتر است. مشابه این یافته، بررسی‌هایی مانند Ahmadi et al. (2023) و Velazco et al. (2022) نیز تأثیر معنی‌دار ثبات حرارتی در تعیین رویشگاه گونه‌های حساس به تغییرات اقلیمی را گزارش کرده‌اند. متغیر Bio12 یکی از اصلی‌ترین عوامل محدودکننده برای گونه‌های جنگلی رطوبت‌دوست مانند شمشاد هیرکانی است. بارش سالانه در سطح کلان، شاخصی از پایداری

توپوگرافی در کنار اقلیم است. اگرچه متغیرهایی مانند طول شیب و TRASP در مدل حضور داشتند، اما تأثیر آنها نسبت به متغیرهای اقلیمی کمتر بود. این موضوع تأیید می‌کند که شمشاد بیشتر تحت کنترل شاخص‌های بزرگ مقیاس اقلیمی است و ویژگی‌های توپوگرافی تنها در سطوح محلی‌تر نقش تنظیمی دارند.

تحلیل منحنی پاسخ شمشاد

متغیر bio3 بیانگر تفاوت میان دمای گرم‌ترین و سردترین ماه سال است و معیاری از دامنه نوسان سالانه دما محسوب می‌شود. گونه‌های حساس به تغییرات شدید اقلیمی مانند شمشاد، معمولاً در مناطقی با دامنه دمایی معتدل و پایدار پراکنش بیشتری دارند. منحنی پاسخ در این بررسی نشان داد که شمشاد بیشترین حضور را در بازه دمایی ۱۵ تا ۲۰ درجه نوسان سالانه دارد. این یافته با نتایج گزارش شده توسط Wani et al. (2024) برای گونه‌های هم‌جنس مطابقت دارد و بیانگر آن است که پایداری دمایی به‌ویژه در فصول سرد و گرم، نقش کلیدی در توان بقا و رقابت‌پذیری این گونه ایفا می‌کند. دامنه‌های نوسانی شدید می‌تواند سبب آسیب بافتی، تنش فیزیولوژیکی و کاهش زادآوری در گونه‌های همیشه‌سبز شود (Becklin et al., 2016). براساس نتایج حاصل از منحنی پاسخ، رابطه‌ای مثبت بین افزایش بارندگی سالانه و احتمال حضور شمشاد مشاهده شد. متغیر bio12، نیز از مهم‌ترین متغیرهای تاثیرگذار بر پراکنش شمشاد شناسایی شد. شمشاد هیرکانی عموماً در مناطقی با میانگین بارندگی سالانه بالاتر از ۳۰۰ میلی‌متر دیده می‌شود (Khabazi et al., 2019). آب‌دوستی این گونه سبب شده است که بیشتر در نواحی پر باران، به‌ویژه در غرب و مرکز جنگل‌های هیرکانی، گسترش یابد. این نتیجه با یافته‌های Moghbel Esfahani et al. (2023) که تأکید دارند پراکنش گونه‌های مرطوب‌پسند در جنگل‌های هیرکانی

تابعی از شیب رطوبتی غرب به شرق است، تطابق کامل دارد. متغیر Bio8 نماینده شرایط دمایی در فصل پر بارش است، فصلی که معمولاً با بیشترین رشد و فعالیت زیستی گیاهان هم‌زمان است. دمای بالا در این دوره می‌تواند سرعت فتوسنتز را افزایش دهد، مشروط بر آنکه با کمبود آب همراه نباشد. منحنی پاسخ شمشاد نسبت به متغیر bio8 نشان داد که شمشاد در دمای میانگین پنج تا ۱۰ درجه در فصل مرطوب بیشترین احتمال حضور را دارد. این بازه دمایی با محدوده بهینه رشد درختان همیشه‌سبز همخوانی دارد و نشان می‌دهد که ترکیب مناسب دما و رطوبت عامل تقویت‌کننده‌ای در عملکرد فیزیولوژیکی شمشاد است (Moore et al., 2021). متغیر bio1 یا میانگین دمای سالانه یکی از جامع‌ترین شاخص‌های اقلیمی در پژوهش پراکنش گونه‌هاست. شمشاد در بازه دمایی حدود هشت تا ۱۳ درجه حضور بیشتری نشان داده و در دماهای بالاتر از ۱۵ درجه، احتمال حضور آن کاهش یافته است. این رفتار بیانگر سقف حرارتی گونه و حساسیت آن به گرمایش اقلیمی است. دما به‌عنوان عامل کلیدی در تنظیم چرخه‌های رشد، گل‌دهی و خواب گیاهان نقش دارد و خروج از محدوده بهینه می‌تواند فرآیندهای بیوشیمیایی مانند فتوسنتز و تنفس را مختل کند (Jarvis et al., 1986). همچنین دماهای بالا با افزایش تبخیر و کاهش رطوبت خاک همراه هستند و می‌توانند مقدار تنش فیزیولوژیکی را افزایش دهند (Alavi et al., 2020).

تحلیل نقشه مطلوبیت رویشگاه شمشاد

نقشه مطلوبیت رویشگاه گونه شمشاد هیرکانی با استفاده از مدل جنگل تصادفی و بر پایه متغیرهای زیست‌اقلیمی و توپوگرافی تهیه شد (شکل ۵). نتایج مدلسازی نشان داد که مناطق با بالاترین احتمال حضور گونه بیشتر در نوار میانی جنگل‌های هیرکانی و به‌ویژه در استان‌های

و مدیریتی، این مناطق می‌توانند به‌عنوان رویشگاه‌های بالقوه برای احیاء در آینده در نظر گرفته شوند. با توجه به نتایج به‌دست آمده، پیشنهاد می‌شود که حفظ، احیا و پایش رویشگاه‌های دارای درجه مطلوبیت بالا به‌ویژه در مناطق مرطوب با ارتفاع پایین در اولویت برنامه‌های حفاظتی قرار گیرد، چراکه این نواحی نه‌تنها رویشگاه طبیعی شمشاد هستند، بلکه بیشترین پتانسیل برای بقای این گونه در معرض تهدید را دارا می‌باشند.

نتیجه‌گیری کلی

یافته‌های این پژوهش نشان می‌دهد که پراکنش شمشاد به‌شدت تحت‌تأثیر متغیرهای اقلیمی بوده و تعامل بین دما و بارندگی، به‌ویژه در سطح فصلی، نقش کلیدی در تعیین زیستگاه‌های مناسب دارد. این نتایج، ضمن تأیید وابستگی شمشاد به اقلیم مرطوب و دمای معتدل، ضرورت در نظر گرفتن تغییرات اقلیمی در آینده را برای برنامه‌ریزی حفاظتی و احیایی برجسته می‌کند. از آنجا که الگوهای اقلیمی در حال دگرگونی هستند، شناخت دقیق مرزهای بوم‌شناسی این گونه می‌تواند به مدیریت انعطاف‌پذیرتر زیستگاه‌ها کمک کند.

مازندران و گیلان قرار دارند. این مناطق از نظر اقلیمی دارای بارندگی فراوان، دمای معتدل و نوسانات حرارتی کم بوده که با نیازهای بوم‌شناسی گونه شمشاد مطابقت کامل دارد (Alipour et al., 2023; Sagheb Talebi et al., 2014). در این مناطق، رویشگاه‌های واقع در ارتفاعات پایین و با رطوبت بالا، شرایط بهینه‌ای برای رشد و زادآوری شمشاد فراهم کرده‌اند. در غرب استان گیلان، به‌ویژه در نواحی با اقلیم مرطوب خزری، نقشه مدل مناطقی را با درجه مطلوبیت بالا نشان می‌دهد که به‌دلیل بارش سالانه بالا، ثبات دمایی و پوشش جنگلی انبوه، شرایط ایده‌آلی برای این گونه فراهم آورده‌اند. در استان مازندران نیز نواحی وسیعی در منطقه میانی جنگل‌های هیرکانی از مطلوبیت بالایی برخوردار هستند، که هم‌راستا با گزارش‌های پیشین در مورد پراکنش طبیعی و توده‌های بومی این گونه است (Habibi kilak et al., 2019). با آنکه در استان گلستان به‌طور طبیعی توده‌های شمشاد مشاهده نشده است، نقشه مدل مناطقی را با مطلوبیت متوسط تا بالا در این نواحی پیش‌بینی کرده است. این موضوع نشان می‌دهد که در صورت فراهم بودن سایر شرایط زیست‌محیطی

References

- Aertsen, W., Kint, V., van Orshoven, J., Özkan, K., & Muys, B. Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. *Ecological Modelling*, **2010**, 221(8).
- Ahmadi, K., Hosseini, S. M., Tabari, M., & Nouri, Z. Modeling the potential habitat of English yew (*Taxus baccata* L.) in the Hyrcanian forests of Iran. *Forest Research and Development* **2019**, 5(4), 513-525. (In Persian)
- Alavi, S. J., Ahmadi, K., Hosseini, S. M., Tabari, M., Nouri, Z. The importance of climatic, topographic, and edaphic variables in the distribution of yew species (*Taxus baccata* L.) and prioritization of areas for conservation and restoration in the north of Iran. *Iranian Journal of Forest*, **2020**, 11(4), 477-492. (In Persian)
- Ahmadi, K., Mahmoodi, S., Pal, S. C., Saha, A., Chowdhuri, I., Nguyen, T. T., Jarvie, S., Szostak, M., Socha, J., & Thai, V. N. (). Improving species distribution models for dominant trees in climate data-poor forests using high-resolution remote sensing. *Ecological Modelling*, **2023**, 475.
- Asadi, H., Jalilvand, H., Tafazoli, M. and Hosseini, S. Modeling Suitable Habitats of *Parrotia persica* (DC.) C.A.Mey. in the Hyrcanian Forests Using Environmental Factors. *Iranian Journal of Forest and Poplar Research*, **2025**, 33(1), 50-68. (In Persian)
- Akrim, F., Mahmood, T., Hussain, R., Qasim, S., & Zangi, I. D. Distribution pattern, population estimation and threats to the

- Indian Pangolin *Manis crassicaudata* (Mammalia: Pholidota: Manidae) in and around Pir Lasura National Park, Azad Jammu & Kashmir, Pakistan. *Journal of Threatened Taxa*, **2017**, 9(3).
- Alipour, S., & Walas, Ł. The influence of climate and population density on *Buxus hyrcana* potential distribution and habitat connectivity. *Journal of Plant Research*, **2023**, 136(4).
- Allouche, O., Tsoar, A., & Kadmon, R. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **2006**, 43(6).
- Amindin, A., Pourghasemi, H. R., Safaeian, R., Rahmanian, S., Tiefenbacher, J. P., & Naimi, B. Predicting Current and Future Habitat Suitability of an Endemic Species Using Data-Fusion Approach: Responses to Climate Change. *Rangeland Ecology & Management*, **2024**, 94, 149–162.
- Barbet-Massin, M., Jiguet, F., Albert, C. H., & Thuiller, W. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in Ecology and Evolution*, **2012**, 3(2), 327–338.
- Becklin, K. M., Anderson, J. T., Gerhart, L. M., Wadgyamar, S. M., Wessinger, C. A., & Ward, J. K. Examining plant physiological responses to climate change through an evolutionary lens. *Plant Physiology*, **2016**, 172(2).
- Burns, P. A., Clemann, N., & White, M. Testing the utility of species distribution modelling using Random Forests for a species in decline. *Austral Ecology*, **2020**, 45(6).
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. Random forests for classification in ecology. *Ecology*, **2007**, 88(11).
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, **2013**, 36(1).
- Edrisnia, S., Etemadi, M., & Pourghasemi, H. R. Machine learning-driven habitat suitability modeling of *Suaeda aegyptiaca* for sustainable industrial cultivation in saline regions. *Industrial Crops and Products*, **2025**, 225, 120427.
- Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., ... E. Zimmermann, N. Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **2006**, 29(2).
- Elith, J., & Leathwick, J. R. Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, **2009**, 40.
- Esmailzadeh, O., & Soleymannipour, S. Habitat suitability and ecological requirements of *Buxus hyrcana* in the Hyrcanian forests. *Iranian Journal of Forest and Poplar Research*, **2015**, 23(1), 45–58. (In Persian)
- Fielding, A. H., & Bell, J. F. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **1997**, 24(1), 38–49. doi: 10.1017/S0376892997000088
- Franklin, J. Species distribution models in conservation biogeography: *Developments and challenges*. In *Diversity and Distributions*, **2013**, (Vol. 19, Issue 10).
- Freeman, E. A., Moisen, G. G., Coulston, J. W., & Wilson, B. T. Random forests and stochastic gradient boosting for predicting tree canopy cover: Comparing tuning processes and model performance. *Canadian Journal of Forest Research*, **2015**, 46(3).
- Graham, C. H., Moritz, C., & Williams, S. E. Habitat history improves prediction of biodiversity in rainforest fauna. *Proceedings of the National Academy of Sciences of the United States of America*, **2006**, 103(3).
- Guisan, A., Edwards, T. C., & Hastie, T. Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecological Modelling*, **2002**, 157(2–3).
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., Regan, T. J., Brotons, L., McDonald-Madden, E., Mantyka-Pringle, C., Martin, T. G., Rhodes, J. R., Maggini, R., Setterfield, S. A., Elith, J., Schwartz, M. W., Wintle, B. A., Broennimann, O., Austin, M.,

- ... Buckley, Y. M. Predicting species distributions for conservation decisions. *Ecology Letters*, **2013**, 16(12).
- Guo, C., Lek, S., Ye, S., Li, W., Liu, J., & Li, Z. Uncertainty in ensemble modelling of large-scale species distribution: Effects from species characteristics and model techniques. *Ecological Modelling*, **2015**, 306.
- Habibi kilak, S., Alavi, S. J., & Esmailzadeh, O. (2019). Analyzing the ecological niche of *Buxus hyrcana* Pojark in the northern forests of Iran. *Forest and Wood Products*, 72(1), 21–31. (In Persian)
- Habibikilak, S., Alavi, S. J., & Esmailzadeh, O. Investigating the influence of different environmental variables in modeling the distribution of yew (*Taxus baccata* L.) using the MAXENT model in Hyrcanian forests. *Forest Research and Development*, **2025**, 11(1), 25-39. (In Persian)
- Haneczok, J., & Piskorski, J. Shallow and deep learning for event relatedness classification. *Information Processing and Management*, **2020**, 57(6).
- Hedayati Kaliji, S., Hosseini, S. M., Alavi, S. J., and Amiri, M. Current and future distribution modeling of oriental beech (*Fagus orientalis* Lipsky) in Hyrcanian forests. *Forest Research and Development*, **2025**, 10(4), 527-543. (In Persian)
- Hesabi, A., Alavi, S. J., & Esmailzadeh, O. Evaluation of the accuracy of climatic data from the WorldClim and Chelsa databases in three northern provinces of Iran. *Forest Research and Development*, **2025**, 11(1), 109-132. (In Persian)
- Hijmans, R. J. Cross-validation of species distribution models: removing spatial sorting bias and calibration with a null model. *Ecology*, **2012**, 93(3), 679-688.
- James, G., Witten, D., Hastie, T., & Tibshirani, R. An Introduction to Statistical Learning: With Applications in R (2nd ed.). *Springer*. **2021**,
- Jarvis, P. G., & Mcnaughton, K. G. Stomatal Control of Transpiration: Scaling Up from Leaf to Region. *Advances in Ecological Research*, **1986**, 15(C).
- Kaky, E., Nolan, V., Alatawi, A., & Gilbert, F. A comparison between Ensemble and MaxEnt species distribution modelling approaches for conservation: A case study with Egyptian medicinal plants. *Ecological Informatics*, **2020**, 60.
- Khabazi, F., Esmailzadeh, O., & Najafi, A. Supervised classification of *Buxus hyrcana* plant communities using artificial neural network. *Iranian Journal of Forest*, **2019**, 11(3), 387–400. (In Persian)
- Khaliq, I., Avenot, H. F., Baudoin, A., Coop, L., & Hong, C. Epidemiology of boxwood blight in western North Carolina and Virginia and evaluation of the boxwood blight infection risk model. *Scientific Reports*, **2024**, 14(1), 26829.
- Kougioumoutzis, K., Tsakiri, M., Kokkoris, I. P., Trigas, P., Iatrou, G., Lamari, F. N., Tzanoudakis, D., Koumoutsou, E., Dimopoulos, P., Strid, A., & Panitsa, M. Assessing the Vulnerability of Medicinal and Aromatic Plants to Climate and Land-Use Changes in a Mediterranean Biodiversity Hotspot. *Land*, **2024**, 13(2).
- Kuhn, M., & Wickham, H. *Tidymodels: Easily install and load the 'Tidymodels' packages* (R package version 0.1.3). **2020**, Retrieved from <https://CRAN.R-project.org/package=tidymodels>.
- Martínez-Meyer, E., Peterson, A. T., Servín, J. I., & Kiff, L. F. Ecological niche modelling and prioritizing areas for species reintroductions. *ORYX*, **2006**, 40(4).
- Merow, C., Smith, M. J., Edwards, T. C., Guisan, A., McMahon, S. M., Normand, S., Thuiller, W., Wüest, R. O., Zimmermann, N. E., & Elith, J. What do we gain from simplicity versus complexity in species distribution models? *Ecography*, **2014**, 37(12).
- Mi, C., Huettmann, F., Guo, Y., Han, X., & Wen, L. Why choose Random Forest to predict rare species distribution with few samples in large undersampled areas? Three Asian crane species models provide supporting evidence. *PeerJ*, **2017**(1).
- Moghbel Esfahani, F., Alavi, S. J., Hosseini, S. M., & Tabari Kochaksarai, M. Determining the habitat suitability of *Quercus castaneifolia* C. A. Mey In order to plan restoration using species distribution modeling. *Forest Research and Development*, **2023**, 9(3), 419–436. (In Persian)
- Moisen, G. G., & Frescino, T. S. Comparing five modelling techniques for predicting forest

- characteristics. *Ecological Modelling*, **2002**, 157(2–3).
- Moore, C. E., Meacham-Hensold, K., Lemonnier, P., Slaterry, R. A., Benjamin, C., Bernacchi, C. J., Lawson, T., & Cavanagh, A. P. The effect of increasing temperature on crop photosynthesis: From enzymes to ecosystems. *Journal of Experimental Botany*, **2021**, 72(8).
- Naimi, B., Hamm, N. A. S., Groen, T. A., Skidmore, A. K., & Toxopeus, A. G. Where is positional uncertainty a problem for species distribution modelling? *Ecography*, **2014**, 37(2).
- Olden, J. D., Lawler, J. J., & Poff, N. L. Machine learning methods without tears: A primer for ecologists. In *Quarterly Review of Biology* **2008**, (Vol. 83, Issue 2).
- Phillips, S. J., & Dudik, M. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, **2008**, 31(2).
- Prasad, A. M., Iverson, L. R., & Liaw, A. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. *Ecosystems*, **2006**, 9(2).
- Probst, P., Wright, M. N., & Boulesteix, A. L. Hyperparameters and tuning strategies for random forest. In Wiley Interdisciplinary Reviews. *Data Mining and Knowledge Discovery* **2019**, (Vol. 9, Issue 3).
- R Core Team., R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. **2024**, <<https://www.R-project.org/>>.
- Ren-Yan, D., Xiao-Quan, K., Min-Yi, H., Wei-Yi, F., & Zhi-Gao, W. The predictive performance and stability of six species distribution models. *PLoS ONE*, **2014**, 9(11).
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Aroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., & Dormann, C. F. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. In *Ecography* **2017**, (Vol. 40, Issue 8).
- Rushton, S. P., Ormerod, S. J., & Kerby, G. New paradigms for modelling species distributions? In *Journal of Applied Ecology* **2004**, (Vol. 41, Issue 2).
- Safdar, S., Younes, I., Ahmad, A., & Sastry, S. A comprehensive review of spatial distribution modeling of plant species in mountainous environments: Implications for biodiversity conservation and climate change assessment. *Kuwait Journal of Science*, **2025**, 52(1), 100337.
- Sagheb Talebi, K., Sajedi, T., & Pourhashemi, M. Forests of Iran: A treasure from the past, a hope for the future **2014**, (Vol. 10). *Springer*, Dordrecht.
- Sękiewicz, K., Salvà-Catarineu, M., Walas, Ł., Romo, A., Gholizadeh, H., Naqinezhad, A., Farzaliyev, V., Mazur, M., & Boratyński, A. Consequence of habitat specificity: a rising risk of habitat loss for endemic and sub-endemic woody species under climate change in the Hyrcanian ecoregion. *Regional Environmental Change*, **2024**, 24(2), 68.
- Sillero, N., Arenas-Castro, S., Enriquez-Urzelai, U., Vale, C. G., Sousa-Guedes, D., Martínez-Freiria, F., Real, R., & Barbosa, A. M. Want to model a species niche? A step-by-step guideline on correlative ecological niche modelling. *Ecological Modelling*, **2021**, 456.
- Sillero, N., Campos, J. C., Arenas-Castro, S., & Barbosa, A. M. A curated list of R packages for ecological niche modelling. In *Ecological Modelling* **2023**, (Vol. 476).
- Singh, H. C., Maurya, A., Wairopkam, B., Tiwari, V., Tiwari, A., & Rana, T. S. Predicting current and future suitable habitats for *Bergenia ciliata* in Indian Himalayan region. *Landscape and Ecological Engineering*, **2025**, 1-14.
- Sobhani, P., & Danehkar, A. Modeling the distribution of *Avicennia marina* (Forssk.) Vierh. in the Khamir and Qeshm mangrove forests, Iran using the maximum entropy model (MaxEnt). *Iranian Journal of Forest and Poplar Research*, **2024**, 32(2), 97-111. (In Persian)
- Soleymanipour, S., Esmailzadeh, O. Flora, life form and chorology of Box trees (*Buxus hyrcana*) habitats in forests of the Farim area of Sari. *Taxonomy and Biosystematics* **2015**, 7, 39–56.
- Velazco, S.J.E., Rose, M.B., Andrade, A.F.A., Minoli, I., Franklin, J. flexsdm: An R package for supporting a comprehensive and flexible species distribution modelling workflow. *Methods in Ecology and Evolution*, **2022**, 13(8) 1661-1669.
- Vignali, S., Barras, A. G., Arlettaz, R., & Braunisch, V. SDMtune: An R package to tune and evaluate species distribution models. *Ecology and Evolution*, **2020**, 10(20).

- Wani, Z., Khan, S., Satish, K., Haq, S., Pant, S., Siddiqui, S. Ensemble modelling reveals shrinkage of suitable habitat for Himalayan Boxwood (*Buxus wallichiana* Bail.) under climate change – implications for conservation. *Phytocoenologia* **2024**, 52.
- Wani, Z. A., Dar, J. A., Lone, A. N., Pant, S., & Siddiqui, S. Habitat suitability modelling and range change dynamics of *Bergenia stracheyi* under projected climate change scenarios. *Frontiers in Ecology and Evolution*, **2025**, Volume 13-2025.
- Zhang, L., Huettmann, F., Liu, S., Sun, P., Yu, Z., Zhang, X., & Mi, C. Classification and regression with random forests as a standard method for presence-only data SDMs: A future conservation example using China tree species. *Ecological Informatics*, **2019**, 52.