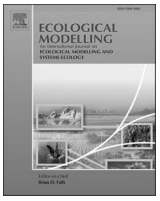
Ecological Modelling 503 (2025) 111056

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel



Bush encroachment with climate change in protected and communal areas: A species distribution modelling approach

Thabang Maphanga [a,b,*], Cletah Shoko [c], Mbulisi Sibanda [d], Blessing Kavhu [e], Corli Coetsee [f,g], Timothy Dube [a]

[a] Institute for Water Studies, Department of Earth Sciences, University of the Western Cape, South Africa
[b] Department of Occupational and Environmental Studies, Cape Peninsula University of Technology, Keizergracht,Cape Town 7925, South Africa
[c] Division of Geography, School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Johannesburg, 2050, South Africa
[d] Department of Geography, Environmental Studies, and Tourism, The University of the Western Cape, Bellville, 7535, South Africa
[e] Environmental Studies, University of California, Santa Cruz, 1156 High 5th, Santa Cruz, 95064, California, United States
[f] South African National Parks, Scientific Services, Private Bag x 402, Skukuza 1350, South Africa
[g] School of Natural Resource Management, Nelson Mandela University, George Campus, 6530, George, South Africa

ARTICLE INFO

Keywords:
Bush encroachment
Climate variability
Kruger national park
Native invasive species
Remote Sensing

ABSTRACT

Savanna rangelands have experienced widespread degradation due to bush encroachment, raising significant concerns among conservationists and rural communities. In the context of climate change, these ecosystem shifts are likely to intensify, especially in South Africa's semi-arid regions. Understanding the impacts of climate variability and change on species distribution within these rangelands is crucial for mitigating further ecosystem disruption. Environmental factors, along with climatic variables, can accelerate the process of bush encroachment, threatening both biodiversity and land use. Early identification of areas vulnerable to invasion is key to developing effective and cost-efficient management strategies. This study aims to model the distribution of invasive species across protected and communal landscapes under long-term climate change projections. A Random Forest (RF) model produced the highest accuracy metrics for Area under the curve (AUC) = 0.99 and True Skill Statistic (TSS)=0.97, while a MaxEnt model recorded the second highest AUC (0.98) and TSS (0.97). The results show a clear difference between the current and future scenarios of the spatial distribution in all the models. Applying a species distribution model (SDM) using both MaxEnt and RF produced a higher degree of prediction accuracy because RF is susceptible to overfitting training data while MaxEnt can produce predictable and complex results. Moreover, the overall predictions using the ensemble model demonstrated an increase in areas suitable for encroachment under RCP 8.5 but a decrease in the bush encroachment rate under RCP 2.6. These findings underscore the critical need for proactive management strategies to mitigate bush encroachment, particularly under high-emission scenarios, ensuring the sustainability of semi-arid savanna rangelands in the face of climate change.

1. Introduction

The increase of shrubs and trees, often referred to as bush encroachment is a process that is globally ubiquitous (Stevens et al., 2017). On the African continent, bush encroachment has significantly affected livelihoods and conservation efforts (Ward, 2005; Palmer and Bennett, 2013; Baloyi et al., 2024), through reducing access and grazing available to livestock and wildlife (O'Connor et al., 2014; Zerga, 2015). Although the causes of bush encroachment are not clear, it has been linked to altered water availability, interactions with specific soil properties, fire patterns, herbivory levels and other anthropogenic activities which may include increased CO_2 in the atmosphere (Palmer and Bennett, 2013; Buitenwerf et al., 2012; Dube et al., 2020). In addition, climate driven changes in rainfall in the southern part of the African continent (Dube and Pickup, 2001; Venter et al., 2018; Yang and Crews, 2020) are resulting in increased rainfall variability with increased drought frequencies (Ludwig et al., 2016; Ramoelo et al., 2018; Mtengwana et al., 2021). Increased rainfall variability can contribute to enhanced bush encroachment as previous studies have shown that bush encroachment accelerates during periods of higher rainfall after

* Corresponding author.
E-mail address: Maphangat@cput.ac.za (T. Maphanga).

https://doi.org/10.1016/j.ecolmodel.2025.111056
Received 31 October 2024; Received in revised form 25 January 2025; Accepted 10 February 2025
Available online 27 February 2025

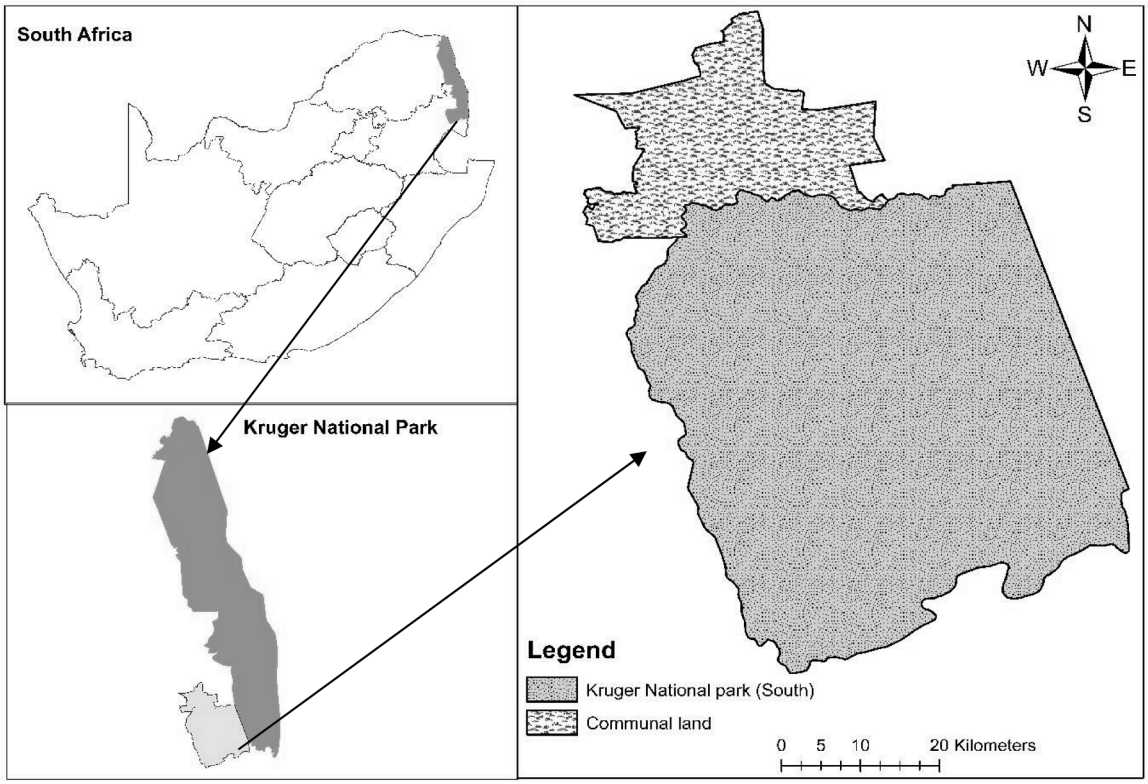


Fig. 1. Location setting of the study sites in South Africa.

droughts (O'Connor et al. 2014; Nackley et al. 2018).

In recent years, researchers have successfully used remotely sensed data to detect changes in savanna rangeland conditions and correlate these changes with climate variability (van Leeuwen et al., 2010; Buntinga et al., 2019; Mtengwana et al., 2021). The use of remotely sensed data in combination with bioclimatic data has increased in several studies (Mtengwana et al., 2021; Schwagera and Berg, 2021). Furthermore, ecologists worldwide have adopted timely and cost-effective methods to model and predict bush encroachment or the distribution of invasive species (Paliwal et al., 2024; Maphanga et al., 2024). To this end, species distribution modelling (SDM) and remote sensing have been widely applied for predictions. Ecologists have long used SDMs to (Allouche et al. 2006; Jiménez-Valverde 2014; Schwagera and Berg, 2021) to quantify the relationship between species and the environment by identifying ecological niches (Elliott et al., 2024), which enables the creation of distribution maps despite limited and abundant species distribution data (Truong et al., 2017; Moudrý et al., 2024). The performance of the models (boosted regression trees (BRT) Maximum Entropy (MaxEnt), Support Vector Machine (SVM), RF, Generalised Linear Model (GLM)) varied across species and environments, and no single best model was identified (Moudrý et al., 2024). Choosing and implementing these models is required. These modelling techniques have been applied to narrow serpentine endemics to predict species distributions under current and past climate conditions with great care (Hu et al., 2024), because inappropriate use of models can affect the accuracy of species prediction (Arenas-Castro et al., 2019; Ahmed et al., 2021). However, these modelling techniques depend on the inferred coarse-resolution environmental information, resulting in geographically overestimated predictions of habitat suitability. Research has shown that the accuracy of SDMs in predicting patterns can be improved compared to models that only use meteorological (Steiner et al., 2024), and topographical predictors gathered through field data collection (He et al., 2015; Mudereri et al., 2019; Riva et al., 2024; Wang et al., 2024).

The application of the SDM connects the current location ("presence-only" or "presence-absence" data) of species with relevant predictor variables (e.g. environmental variables), thus establishing a statistical relationship between the spatial variations of the predictor variables and the distribution of the species in the environment (Booth, 2018; Awuah et al., 2020; Riva et al., 2024). Many studies have compared machine learning algorithms with regression-based models (Shiferaw et al., 2019; Wang et al., 2024), compared machine learning algorithms (Früh et al., 2018; Steiner et al., 2024), and developed ensemble models from different SDMs based on a single model (Ng et al. 2018). However, the performance of SDMs depends not only on the application and problem type (Bhattacharya, 2013) but also on the spatial resolution of the environmental variables (Reside et al. 2019). According to the literature, there is no universally best SDM algorithm or predictor combination for species distribution (Fois et al., 2018; Shekede et al., 2018). The correct choice of modelling algorithm can identify the most pertinent predictor variables from various ecological covariates to forecast the occurrence, spread (Poirazidiset al., 2019; Wang et al., 2024; Hu et al., 2024) and distribution of encroaching woody species (Khanum et al., 2013; Früh et al., 2018). These include methods such as boosted regression trees (BRT), classification and regression trees (CART; Shen et al., 2022), MaxEnt (Mtengwana et al., 2021), generalized linear models (GLM; Khazieva et al., 2022), random forests (RF), and support vector machines (SVM) (Ponce-Fontenla et al., 2021). To apply and evaluate a model, it is essential to ensure that the requirements (such as the training dataset and accuracy) are fully met (Dai et al., 2020).

Numerous studies have demonstrated the significant performance of models in the application of SDMs. Different studies have tested the performance of difference models including RF, MaxEnt, support vector machine, GLM, Domain and Bioclim (Zhang et al. 2020, Yudaputra et al. 2019). RF and MaxEnt often outperform SVM in predicting the potential distribution of plants (Kass et al., 2018; Ncube et al., 2020), therefore the combination of bioclimatic variables and remote sensing data through machine learning has been reported as the optimal way. The combination of bioclimatic variables and remote sensing data using

machine learning is reported as optimal for reliable results, utilizing vegetation metrics and climatic variables (Pujiastuti, and Cropper, 2019; Kaky et al., 2020).

A new paradigm opens up for understanding the earth and agro-ecological systems over time and space due to the availability of free, publicly accessible earth observation "big data" from multiple sources and machine learning algorithms (Gallardo et al., 2017; Souverijns et al., 2020). To forecast how suitable habitats of bush encroachers will change in the future, effective methodologies are needed that integrate remote sensing techniques and supplementary data. This study site, in general, reflects the global pattern of increased woody encroachment but the combination of land uses and management histories provide a good range of vegetation in different states of encroachment (from no encroachers to heavily encroached). This study has two objectives: 1) To utilize three machine learning-based Species Distribution Modeling (SDM) techniques; Random Forest (RF), MaxEnt, and Boosted Regression Trees (BRT) along with their respective ensemble model, to predict the potential distribution of native woody encroachers in both protected and unprotected areas under current and future climate scenarios; and 2) To assess the impact of climate change on the distribution of invasive species within both protected and communal ecosystems.

2. Methodology

2.1. Study area

The study area is divided into two distinct sections: protected (i.e. Kruger National Park, KNP) and unprotected (community) areas. As illustrated in Fig. 1, three sub-research sites were selected in the protected area, namely Skukuza (31° 29′52.4724″ E, 25°5′30.642″ S), Pretoriuskop (31°15′49.73″ E, 25° 3′47.25″ S), and Phabeni (31°18′44.61″ E, 25° 8′13.94″ S). The selection of these locations was based on the presence of known encroaching woody species. Another two study sites were chosen in the community land, which lacks formal conservation protection, and is situated in the Bushbuckridge Municipality of the Ehlanzeni District, specifically in Somerset at coordinates 31°22′22.22″ E and 24°54′37.54″ S. The area in question encompasses an estimated 11,000 hectares of urban, cultivated and rangeland, located between the privately-owned Sabi Sands Game Reserve and communal land in close proximity to the village of Justicia. The average yearly temperature is 22 °C. Annual precipitation ranges from 235 to 1000 mm, rain falls mainly during the summer months (Munyati, 2017). For further description of the study area in terms of vegetation refer to supplementary.

2.2. Field data collection

A field survey was conducted from 10 to 30 April 2022 and 11 to 27 April 2023. The diversity and cover of woody plants were measured in 2022, at the end of the growing season (April to May). This was done for remotely sensed data to be able to classify the woody shrub species which were collected in the field during the end of the growing season when most of the native encroacher species lose their leaves. The rationale behind this is that it allows the capturing of the complex intra and intra species variability which result from natural phenological changes during the different seasons. Curtis and Mannheimer (2005) stated that woody shrubs for the remainder of the year (i.e. dry season) are devoid of leaves. During the field operation, it was observed that *Combretum apiculatum, Acacia nigrescens*, and *Dichrostachys cinerea* experienced leaf abscission and exhibited a yellowing of their foliage, starting in May. Data were collected from thirty-five plots (30 in KNP and 5 in the communal) and in each plot four (4) quadrants (50 m × 50 m) were randomly placed to capture the variability within plots. In each plot, sample points of the native non-encroacher and native encroacher woody species were recorded using a Garmin GPSMAP 65 hand-held global positioning system (GPS) at submeter accuracy. Non-encroacher species were included because some of them have similar characteristics to the native encroacher shrubs. Maphanga et al. (2024) reported that due to spectral similarities, similar ecological composition, and the fact that native vegetation is highly variable and difficult to identify with Sentinel-2 imagery, encroacher and non-encroacher species may be confused. The information derived from the quadrants served as training data for the identification and mapping of both non-encroaching and encroaching native woody species (training (70 % of data and testing 30 %). Furthermore, at a finer scale, spatial autocorrelation becomes more prominent as a measure of dependencies between spatial observations. Thus, to achieve an accurate prediction result, spatial information must be incorporated.

2.3. Sentinel-2 imagery acquisition

Sentinel-2 (S2) captures high-resolution multispectral images, on a revisit frequency of five days. The processing level-1 C Sentinel-2 data from August 24, 2018, was acquired via the USGS Earth Explorer platform (http://earthexplorer.usgs.gov) in three granules (T34HCG, T34HDG, and T34HCH). Sentinel-2 data, collected between 1 December 2021 and 1 May 2023, were acquired. The tiles were arranged into a cohesive mosaic that encompassed the entire study area. The photos obtained within the specified dates were thereafter picked and utilised. A total of 13 spectral bands are collected by the S2 Multispectral Instrument (MSI), including visible and near-infrared (NIR) bands with a spatial resolution of 10 m, red edge, shortwave infrared (SWIR) bands with a spatial resolution of 20 m, and atmospheric bands with a spatial resolution of 60 m. The initial stage of image composition involves selecting cloud free images. A filter was applied to the image collection obtained within a predetermined period in order to select Sentinel-2 granules that had the least cloud cover. The application of the compositing algorithm is limited to the four 10 m bands of Sentinel-2 satellite imagery (Vermeulen et al., 2021; Maphanga et al., 2024).

The decision to incorporate Sentinel-2 data into this study was driven by its unique characteristics, including its ability to operate at long wavelengths and its immunity to cloud interference. These attributes render Sentinel-2 well-suited for monitoring the presence and distribution of species across different time intervals, as highlighted in previous research conducted by Zhou et al. (2019) and Borges et al. (2020). Moreover, the aforementioned dataset has been effectively employed in the cartographic representation of land-cover attributes within savanna ecosystems (Baumann et al., 2018; Vermeulen et al., 2021; Shafeian et al., 2021).

The study approach involves resampling and reclassifying specific datasets, such as DEM and bioclimatic data, according to a flow chart. Unprocessed monthly climate images are retrieved from WorldClim and a Digital Elevation Model (DEM) with a sufficiently high resolution was acquired. 1) We accessed the files using the raster() or stack() functions from the 'raster' package, or the readGDAL() function from the 'rgdal' package. Another option is to utilise the getData() method from the 'raster' package. 2) These were transformed into a SpatialPointsDataFrame or SpatialGridDataFrame object using the 'sp' package. 3) Normalisation of the data, including climate, elevation, and coordinates took place. The expression [(x-mean)/sd] represents the calculation of the standard score, which is obtained by subtracting the mean from the value of x and then dividing. The experimental variogram is calculated using the same procedure, and the functions vgm(), variogram(), and fit. variogram() from the 'gstat' package are used to fit a variogram model. The beginning value of the nugget was assigned as 0 and the initial value of the psill as the mean of the gamma values in the experimental variogram. To calculate bioclimatic variables, this current study converted the data to a RasterStack and then utilised the biovars() method from the 'dismo' package.

2.3.1. Climate datasets

Bioclimatic data have been extensively incorporated into species distribution models (SDMs) to identify underlying factors that influence

Table 1
Declared indicators of bush encroachment in South Africa (The Conservation of Agricultural Resources Act, 1983 (Act No. 43 of 1983).

Botanical name	Common name
Azima tetracantha	Needle bush
Combretum apiculatum	**Red bush-willow**
Dichrostachys cinerea	**Sickle bush**
Euclea crispa	Blue guarri
Euclea divinorum	**Magic guarri**
Grewia bicolor	**Bastard raisin bush**
Leucosidea sericea	Old wood
Senegalia ataxacantha	Flame thorn
Senegalia senegal	Three-hook thorn, Three-thorned Acacia
Terminalia sericea	**Silver cluster leaf**
Vachellia karroo	**Sweet thorn/Soetdoring**

Table 2
R packages of the models used to map species distribution.

Model algorithms	'sdm' syntax	Package	Reference
Boosted regression trees	'brt'	'gbm'	Guisan et al. (2007); West et al. (2017)
Random forest	'rf'	'randomForest'	Mi et al. (2017); Früh et al. (2018); Ng et al. (2018)
MaxEnt	'maxent'	'dismo'	Phillips et al. (2006)
Ensemble	'ensemble'	'sdm'	(Naimi and Araújo 2016)

species distributions (Gallardo et al., 2017; Booth, 2018; Mtengwana et al., 2021). The climatic datasets include various parameters such as temperature, precipitation, solar radiation, and soil moisture. In addition, current and future climate datasets can be freely accessed at http://worldclim.org/. Table 1s shows the total number of bioclimatic variables representing current and future climate at 30 ′onds spatial resolution of 1 km x1km. The future climate projections were based on the Coupled Model Intercomparison Project (CMIP 6), which was downscaled and calibrated with WorldClim v21 as baseline climate (Bravo-García et al., 2024). This study considers only two of the Representative Concentration Pathways (RCPs) proposed by the Intergovernmental Panel on Climate Change (IPCC) for atmospheric carbon: RCP 2.6 (minimum) and RCP 8.5 (maximum). A series of RCPs were selected in order to gather information about uncertainties regarding impacts and adaptation planning for savanna rangeland ecosystems (Mtengwana et al., 2021). For CO_2 concentrations by 2080, RCP scenarios represent 2.6 and 8.5 W/m2 of radioactive forces (IPCC 2014). This study combined data from Southern Africa to evaluate the suitability of bush encroachment and to measure the potential scale of this future conflict under two standard concentration pathways (RCP 2.6 and 8.5; Methods), representing the best- and worst-case scenarios for regional emissions (Souverijns et al., 2020). This study analysed the two most extreme Representative Concentration paths (RCPs), RCP 2.6 and RCP 8.5, to deduce the impacts of intermediate emission paths through approximate interpolation (Shekede et al., 2018; Shen et al., 2022).

2.3.2. Topography data

The spatial distribution of bush encroachment is significantly influenced by topography due to the climate transition caused by terrain and the meso-climatic conditions at specific sites (Schwager and Berg, 2021). Furthermore, Oldeland et al. (2010) asserted that the spatial arrangement of invasive species in rangelands is impacted by the topographic depression (elevation) of a given landscape and we assumed elevation will equally be important for the distribution of bush encroachment species (Mtengwana et al., 2021). In the presence of a discernible environmental gradient within a designated study area, it is expected that both species composition and vegetation structure will undergo alterations along this gradient (Bravo-García et al., 2024). In this study, the topographic variables were obtained from the elevation model (DEM) sourced from the website (DEM: https://dwtkns.com/srtm30m). This DEM was employed to derive additional topographic attributes such as aspect, slope, Topographic Wetness Index (TWI), and Topographic Position Index (TPI) within the specified study region. Moreover, the calculation of the "saga wetness index" (SWI) was performed using the "rsaga. wetness.index" function (Ponce-Fontenla et al., 2021) in the R programming language. The SWI is a modified version of the TWI. The soil type data utilized in this study was obtained from the ISRIC data hub, accessible at http://data.isric.org.

2.3.3. The process of selecting species distribution models

Currently, there is a wide range of modeling methodologies that may be broadly classified as "regression," "profile," and "machine learning" (Ahmed et al., 2021; Ponce-Fontenla et al., 2021). This research aims to conduct a comparative analysis of three frequently used models (which were?) for invasive species distribution modeling (SDM) and one model that combined these. A range of methodologies such as machine learning, regression analysis, and profiling techniques were employed to evaluate the performance of models in the study (Table 2). The different models demonstrate differences in their strategies for managing model complexity and overfitting, specifically in how they address the bias-variance trade-off (Feilhauer et al., 2012; Braunisch et al., 2016), as well as their inclusion or exclusion of interactions. Several machine learning methods were assessed, including Bioclim, Boosted Regression Trees (BRT), and Random Forests (RF) and their respective ensemble model were used (Table 2).

The study further showed that Species Distribution Modeling techniques can effectively determine the dynamic distribution patterns of encroaching species across landscapes. The BRT model utilizes a maximum likelihood and ensemble-like method to enhance a single regression tree (Elith et al. 2018). The RF model operates by selecting the class with the greatest probability across several decision trees (Bangira et al. 2019). MaxEnt maps species presence by evaluating the highest entropy within the environmental niche, meaning the largest area. These techniques have certainly become popular for estimating the spatial distributions of different species in ecological applications. For instance, numerous research endeavours have shown the existence of RF in SDMs (Mudereri et al., 2020; Reva et al., 2024). Mtengwana et al. (2021) demonstrated the potential of utilizing BRT to improve predictive capabilities for modeling species distribution. This minor variation in accuracies typically stems from the differences in algorithmic structures and input data assumptions across the models. Along with the capability to merge strengths and minimize weaknesses in separate models, this justifies the adoption of an ensemble modeling approach, which is practical because of its high usability. Nonetheless, to conserve resources, a frequently employed method is to jointly predict using various models in a single run; however, this does not apply to the current research, which actually leverages the outcomes of the ensemble of models.

2.3.4. The collinearity tests

Prior to estimating the distribution of bush encroachment, the bioclimatic data underwent an assessment for collinearity, which refers to a significant correlation between predictor variables. This assessment was conducted using the Variance Inflation Factors (VIF) as described by Akinwande et al. (2015) and Shekede et al. (2018). According to Pradhan (2016), the presence of multicollinearity among the Bioclimatic variables used for characterizing species distribution might lead to the occurrence of over-fitting in models. The VIF is a statistical metric that quantifies the extent to which multicollinearity amplifies the variance of slope estimates (Feilhauer et al., 2012; He et al., 2015; Mudereri et al., 2020). It is calculated using a multiple regression model that incorporates paired predictor variables. In order to model the distribution while eliminating variables with high VIF, this study used the 'usdm' package in the R programming language. The VIF has been widely used

Table 3
The anticipated changes in Bioclimatic variables within both protected and unprotected regions. Positive numbers indicate an increase, whilst negative values indicate a reduction of the stated amount.

Methods	AUC	COR	TSS	Deviance
RF	0.99	0.5	0.97	0.07
Maxent	0.98	0.43	0.97	0.06
BRT	0.96	0.46	0.93	0.09

in several studies to identify significant characteristics associated with the propagation of invasive species (Zimmermann et al., 2007; Ng et al., 2018).

2.4. Model evaluation

Each model's performance was evaluated to test the accuracy and reliability of its outcomes (Fois et al., 2018; Ndlovu and Shoko, 2023). To calibrate the models, three commonly employed algorithms were employed in the R package "Biomod2″ (R Core Team, R Foundation for Statistical Computing, Vienna, AT). The study established these unique models using the 'biomod2′ package version 3.3–7 (Thuiller et al. 2009) within the R statistical programming environment. 'biomod2' provides various methods and functionalities relevant to distribution modelling, featuring an effective framework for developing ensemble SDMs. Ensemble SDM procedures can be carried out using alternative toolkits such as BioEnsembles (Diniz-Filho et al. 2009) or even without dedicated toolsets (Shekede et al., 2018). The current study chose 'biomod2′ because of its widespread use among ensemble species distribution model users, which makes it a representative example of the standard ensemble SDM methodology. Algorithms used included generalized linear models, generalized additive models (GAM), generalized boosted models, multivariate adaptive regression splines, random forest, and artificial neural networks. Each algorithm was set to its default settings. Some of these algorithms include random forest, classification, regression trees and gradient boosting. In this current study, random forest and MaxEnt model were used to determine the area under the curve (AUC). Moreover, the study used the SDM function to provide the code to fit the gradient boosting classifier, which changed the classifier and activated the gradient boosting classifier. The setOutputMode () function was used to obtain the results as probability and as classification. Through the utilization of this methodology, it became feasible to visually represent a binary outcome, specifically the predicted presence, within an interactive cartographic display within a brief timeframe. A threshold independent measure of the receiver operator characteristic (ROC) plot area under the curve (AUC) was calculated with the ROCR library (Li et al., 2022; Ndlovu and Shoko, 2023) and the omission rate in R statistical package. AUC scores exceeding 0.5 indicate performance beyond chance, whereas values approaching 1 indicate highly accurate model predictions. During each iteration, the spatial blocks were divided randomly into two subsets: 70 % of the blocks was used for model fitting, while the remaining 30 %was used for model validation.

3. Results

3.1. Presents of encroachers and predicted species distribution under current climatic conditions

Different evaluation metrics were used to evaluate conditions, such as the receiver operating characteristic curve (AUC-ROC), true skill statistic (TSS), and kappa coefficient. Based on the model's results, species distributions under current climatic conditions were accurately predicted, with high AUC-ROC values, TSS scores close to 1, and kappa coefficients showing substantial agreement between observed and predicted distributions. Table 3 shows the Random Forest model (AUC= 0.99 and TSS=0.97) produced the highest accuracy metrics for both AUC and TSS, while the MaxEnt model recorded the second highest AUC (0.98) and TSS (0.97) (Table 3). The accuracy values offered are generated only from the usage of current bioclimatic variables. The results show that kappa coefficients for MaxEnt was 0.43 and BRT was 0.46. The Receiver Operating Curves (ROC) shown in Fig. 3 demonstrate that the Random Forest (RF) and MaxEnt models exhibited a notable level of consistency in their predictive performance across the duplicate models. A MaxEnt algorithm provides spatial predictions for species distribution modeling based on presence-only data and occurrence data. MaxEnt models are used to predict the likelihood of geographical distribution for species whose distribution is closest to uniform yet influenced by environmental factors. Also, it attempts to predict a species' suitability for a particular habitat. The performance of the species distribution models ranged from 0.61 to 0.99, thus falling into the good to excellent category. The models were deemed ideal for predicting encroacher species' potential habitat based on their performance. Moreover, the future projection results show that the models had an area under the receiver operating characteristic curve (AUC) greater than 0.9, sensitivity and specificity that exceeded 0.95. With an AUC greater than 0.9, sensitivity greater than 0.9, and specificity greater than 0.99, this level of accuracy was maintained even when projecting beyond the range of data used for calibration.

3.2. Encroachment distribution pattern is affected by environmental factors

In order to identify the critical bioclimatic predictors for bush encroacher species habitat suitability under different climate scenarios, the percentage contribution of each variable to the model was considered. Fig. 4 illustrates that land use and land cover had the greatest significance in forecasting the present distribution of species in all models, followed by Bio 2 (the average monthly difference between maximum and minimum temperature) for both RF and MaxEnt. Under the current climatic conditions, Bio 2, Bio 5 (Max Temperature of Warmest Month), Bio 6 (Min Temperature of Coldest Month), Bio 13 (Precipitation of Wettest Month), Bio 17 (Precipitation of Driest Quarter) and Bio18 (Precipitation of warmest quarter) are the most important predictors of bush encroachers (species) habitat (Fig. 4a-c). Proposed drivers of the phenomenon encompass a range of factors, from localized changes in fire regimes, herbivory patterns, and direct anthropogenic impacts, to global alterations in climate or atmospheric [CO2] concentrations that may be accelerating woody growth. These potentially widespread increases, unexplained by changes in local disturbance history, suggest the possible influence of drivers beyond the scope of local control (Maphanga et al., 2024). Alterations in disturbance regimes, which may be subject to local management, may contribute to woody encroachment. Chronic disturbances by fire and herbivores play a crucial role in regulating savanna woody cover. The possibility exists that biological activity and species richness remain largely unaffected by spatial variations in climate. However, changes in plant species richness were jointly driven by the magnitude of changes in temperature and precipitation. The type of land use and land cover influences the expose to bush encroachment for example an overgrazed area is more susceptible to invasive species under dry conditions. In arid regions, vegetation structure changes in arid regions with increasing precipitation. Increased precipitation causes vegetation structure to change in arid regions due to climatic changes. In the wet season, woody shrubs are photosynthetically active, whereas during the dry season they are not. Bush encroachment is driven by climate variability, especially temperature and rainfall, especially in areas where human activity is minimal.

Moreover, the results from Fig. 4 demonstrate that RF and MaxEnt had key variable importance which are non-climatic predictors which were TPI, slope and aspect, while soil and TWI were the least important variables. This contrasts with the BRT whereby slope was the second most important variable after land cover and the least important were

TWI and aspect. Surprisingly the BRT model had the least number of bioclimatic factors when compared to RF and MaxEnt models. The combined variables explained over 90 % of the habitat suitability model under the current climate with Bio 2 being the most important variable (95 %−53.5 %) for RF and MaxEnt respectively. Upon examining the variable of relevance, it is evident that only the land cover and aspect variables significantly contributed to the model's predictive capabilities. Regarding the bioclimatic variables, the primary predictor of significance is Bio 18, specifically the precipitation during the hottest quarter. The future projections have also shown that Moisture index, soil, slope and aspect amongst some of the key climatic variables (precipitation and temperature) were good predictors for the suitable areas for encroachment for both RCP 2.6 and RCP 8.5.

3.3. *Distributions of encroaching woody species current and in the future climatic conditions*

The results depicted in Fig. 5 shows the estimated suitable habitat in kilometer square (km^2) for the occurrence and distribution of encroacher woody species. The RF model shows the lowest estimated suitable areas for encroachment which is 72 km^2 under current conditions and this will decrease to 28.39 km^2 for RCP 2.6 and 28.58 km^2 RCP 8.5. However, MaxEnt illustrated the highest suitable habitats of bush encroachment out of all the models with current being 395 km^2 while the RCP 2.6 109.91 km^2 and 173 km^2 for RCP 8.5. It was noted that the results shows that the current area of encroachment will decrease in the future scenario under RCP 2.6; while MaxEnt and BRT indicate an increase from the RCP 2.6 to the RCP 8.5. The BRT model shows that the current suitable area is 146 km^2 and 77.21 km^2 for RCP 2.6 while RCP 8.5 (81 km^2) shows a slight increase from RCP 2.6. Moreover, the overall predictions using the ensemble model demonstrated an increase in suitable areas under RCP 8.5 while there is a decrease in suitable areas under RCP 2.6. In summary there is a variation in the occurrence and distribution of bush encroachment of the estimated suitable areas across all models (Fig. 5).

Fig. 6 illustrates that the broad distribution pattern is relatively consistent across all the models. It is important to mention that the estimated likelihood of encroaching shrubs varied from 0 to 1 throughout the study area. The MaxEnt, BRT, and RF models forecast exhibit notable geographical variations in the western region of the KNP and surrounding communal areas in current scenarios. Overall, the Random Forest (RF) exhibits distinct spatial variations compared to the projected appropriate regions identified by the MaxEnt and RF Species Distribution Models (SDM). Fig. 6 Future distribution of woody encroacher species is based on the three machine learning algorithms used and their respective ensembles. The predicted future distribution of woody encroacher species differs across the models. The results show a clear difference between the current scenario and the future scenario of the spatial distribution in all the models.

But looking at the future scenario of RF, and MaxEnt illustrates a distinct spatial distribution in the northern part of the study area. MaxEnt shows a decrease in the expansion of the woody native encroacher species in the RCP 8.5 relative to the increase observed in RCP 2.6 (Fig. 6). It is worth noting that during the current climatic scenarios, a complete opposite pattern is observed whereby the woody encroacher species are spatially distributed in the southern part of the study area of the park. Contrary to what is observed in MaxEnt model the BRT shows an expansion in the RCP 8.5 from the reduction in RCP 2.6 (Fig. 6). According to these results, woody native encroacher species are likely to occur more frequently in the northern part of the study area, towards the communal areas surrounding it. Using an ensemble model, the predictions from several models are combined by taking a weighted average. By doing so, the ensemble model capitalizes on the strengths of each model while compensating for its weaknesses. By doing so, spatial uncertainty that exists in individual models is reduced.

4. Discussion

In this study, encroacher species distribution was modelled in protected and communal areas under long-term climate change predictions in the Lowveld of South Africa. This study used different machine learning algorithms which were coupled with bioclimatic data of current and future scenarios to assess the likelihood of climate change effects on the distribution of encroaching woody species under climate projections. The rationale for using different models was to identify which models can optimally predict species distribution accurately. The results showed that the application of a species distribution model (SDM) combining MaxEnt and RF produced a higher degree of prediction accuracy than either of these alone. This is likely because RF is susceptible to overfitting training data (He et al., 2011; Ahmed et al., 2021; Hu et al., 2024) while MaxEnt can produce predictable and complex results (Shekede et al., 2018; Ncube et al., 2020). The spatial distribution of the woody native encroacher species under the RCP 2.6 (minimum) and RCP 8.5 (maximum) climate projections using the machine learning approach BRT, MaxEnt, and RF can accurately predict species distribution to establish optimized models. MaxEnt, BRT, and RF models predict the vegetation to look different in the western part of the KNP and neighbouring communal lands. It is anticipated that these variables will change markedly under the RCP2.6 scenario, leading to considerable sections of the existing distribution area becoming unsuitable by 2080. Communal areas face diverse challenges, including the encroachment of residential structures into rangeland areas and harvesting of woody species for fuel by some community members. Consequently, the projections for both RCP 2.6 and RCP 8.5, are expected to differ significantly from those of Kruger National Park. This analysis forecasts a greater reduction in range under RCP8.5 by 2080, representing a notably more severe scenario. The interplay of temperature, precipitation, and various environmental factors will shape the habitat needs of the species in question. Conversely, the response curves produced for this species indicated varying variables that yielded distinct responses for each algorithm. Due to the current situations, the western parts of the KNP and parts of neighbouring communal lands have undergone significant alteration based on the MaxEnt, BRT, and RF models. In comparison with the future prospective areas forecasted by the MaxEnt and RF Species Distribution Models (SDMs), Random Forest (RF) exhibits pronounced spatial variations. Based on four machine learning methods and ensembles, depicted in Fig. 6, this paper predicts the future distribution range of woody encroacher species.

The results of this study further illustrated there was a variation in the variable of importance when it comes to the predictor variables across all the models, and this can be attributed to the estimating ability of the models and the algorithms they are based on. Additionally, the modeling used in this study benefited from the inclusion of land use data and bioclimatic variables. The combined variables explained over 90 % of the habitat suitability model under the current climate with Bio 2 (Mean of monthly (max temp - min temp)) being the most important variable (95 %−53.5 %) for both RF and MaxEnt. Additionally, the variable of relevance, it is evident that the land use and land cover and aspect variables also significantly contributed to all the models' predictive capabilities. Ye et al. (2018) and Shekede et al. (2018) reinforce the findings of this research by finding that, while present land cover was identified as a highly significant variable, the predictions generated need to be interpreted carefully because land cover may change in the future. The forecast of acceptable habitats, however, differs depending on the model used, since each model utilizes a different set of equations and algorithms. Although performance comparisons among several GCM models exist for the CMIP5 project (Sanderson, Knutti and Caldwell, 2015), such analyses are lacking for the new CMIP6 simulations. The new IPCC AR6 (IPCC, 2021) introduced CMIP6 model findings utilising a distinct framework known as "Shared Socioeconomic Pathways" (SSP). SSPs incorporated several assumptions regarding the evolution of socioeconomic elements, including population increase,

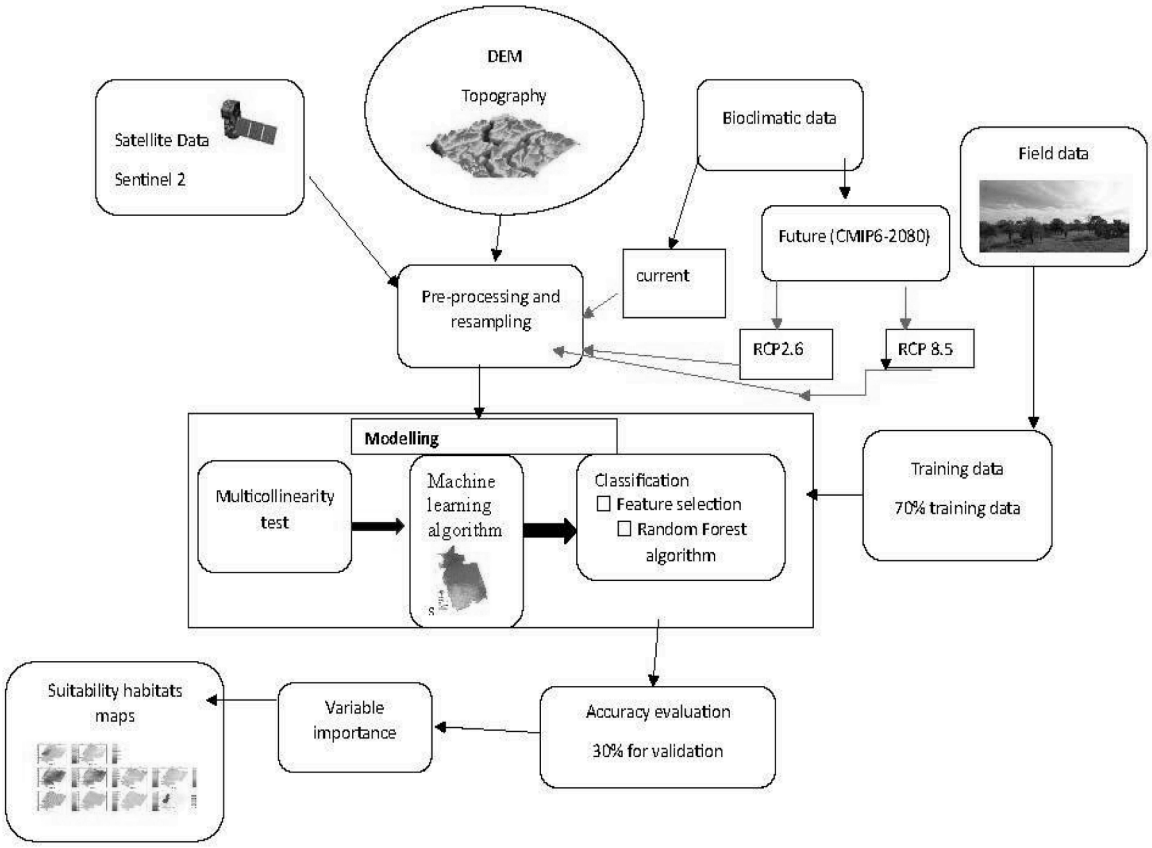


Fig. 2. Methodological overview of the processing flow.

technological progress, and climate policy, by 2100, but RCPs excluded any socioeconomic considerations in their modelling. Mudereri et al. (2020), for instance, noted that it was dependent on the kind of model used to make its predictions. However, other research has proven that land cover is a significant driver of habitat adjustment (Ndlovu et al. 2018). According to Nath et al. (2019), the MaxEnt model fundamentally indicated that climate predictors are the most important factors in projecting species distributions. Terzano et al. (2018) discovered that land cover had a limited impact on the modelling of Invasive Alien Plants (IAP) distribution on a large scale. Despite a reduction in zonal bias with CMIP6, the spatial organisation of biases has not significantly changed between CMIP5 and CMIP6 (Priestley et al., 2020). Harvey et al. (2020) shown that whereas the multimodel biases in CMIP3, CMIP5, and CMIP6 exhibit similar spatial patterns, the biases in the CMIP6 models possess markedly reduced magnitudes. The introduction of CMIP6, with its sophisticated characteristics akin to the model outputs of prior versions utilised for analysing historical climate changes in Southern Africa, offers robust projections.

Bush encroachment changes soil moisture and nutrient competition among shrubs or woody vegetation (Dube and Nhamo, 2020). Ludwig et al. (2016) and Tessema et al. (2017) observed a negative correlation between encroachment rate and initial woody cover in the Lowveld savanna rangelands of the northeastern part of the Kingdom of Eswatini, formerly known as Swaziland. Climate variability and change affect local and national parks, private nature reserves, and communal areas; however, the literature demonstrates that the impact is more pronounced in communal areas, as their livelihood depends on the health of the rangeland (Dube and Nhamo, 2020; Hu et al., 2024). Elevation has been identified as a significant variable, potentially due to topography influencing soil texture, water availability, temperature, and the general distribution of vegetation (Borges et al., 2022; Shen et al., 2022). Oldeland et al. (2010) determined that rangelands are influenced by the topographic depression (elevation) of the landscape. Variations in species composition and vegetation structure are likely to occur along an environmental gradient in a study area. Furthermore, numerous studies have indicated that encroaching plant species are typically distributed based on topographical factors (Symeonakis et al., 2016; Higginbottom et al., 2018; Souverijns et al., 2020).

The projected future range of woody native encroacher species varies across all of the models. All of the models indicate a considerable disparity between present and future distributions of woody native encroacher species. Various applications require the superior performance of RF, which makes it the preferred choice. Several studies also found that RF consistently performed better than other methods when applied to the prediction of the spread of invasive species from remotely sensed data (West et al., 2016; Jensen et al. 2020). Literature has shown that there is insufficient data to support the existence of a superior overall model, as indicated by Guo et al. (2019), Hao et al. (2019), and Mudereri et al. (2020). Therefore, the application of the ensemble analysis becomes important in all predictive modeling, particularly for generating an accurate and comprehensive projection (Mtengwana et al., 2021). With high AUC-ROC values, TSS scores close to 1, and Kappa coefficients showing substantial agreement between observed and predicted species distributions under current climate conditions, the RF model was able to accurately predict species distributions. Additionally, the future projection results indicate that the models have an area under the receiver operating characteristic curve (AUC) greater than 0.9, along with sensitivity and specificity exceeding 0.95. Ensembles have the benefit of reducing the spatial uncertainties of the models for each climatic scenario, allowing for accurate spatial estimations (Guan et al., 2020).

The use of ROC curves serves to assess the effectiveness of deterministic and probabilistic identifications, hence offering a practical approach for the application of plant species prediction models

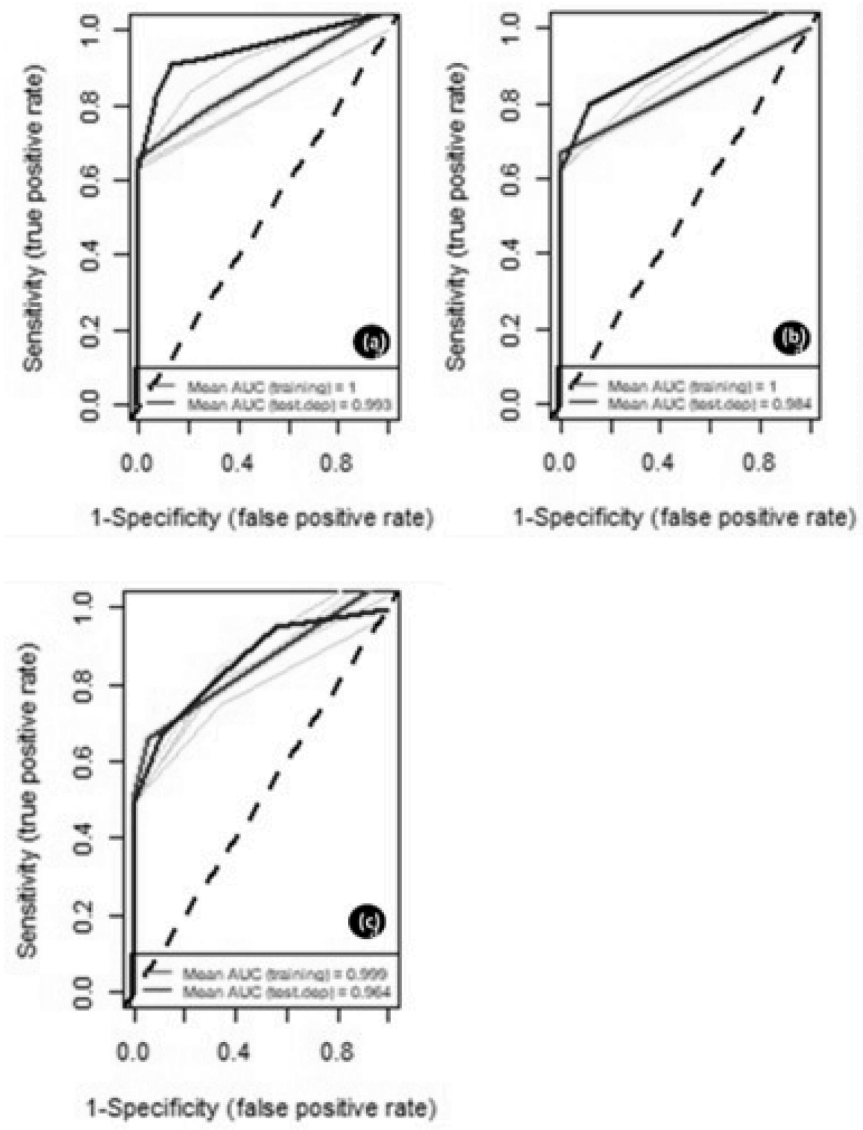


Fig. 3. The ROC curve for (a) Random Forest, (b) Maxent and (c) BRT models.

(Khanum et al., 2011; Kass et al., 2018; Awuah et al., 2020). Given the performance of high sensitivity, specificity and the AUC of the models used in this study both for the current scenario and future projections it can be noted that the suitable habitat of native invasive species-based on their efficacy. While the projections (current and future) have high accuracy levels such as sensitivity, specificity, TSS, and AUC, it is advisable to be cautious when making management decisions based on these models as real-world scenarios might change. The argument of this was also stated by Ponce-Fontenla et al., 2021 that the level of soil moisture, erosion levels and grazing pressures might change the projected species distribution within an ecosystem. Moreover, the models can also accurately represent how different soil characteristics and types, and slope strongly influence the selection of plant communities. According to Zhao et al. (2022) machine learning such as RF and MaxEnt models perform well in the prediction of potential species distributions. The guidelines which were outlined by Allouche et al. (2006), were used in this study to evaluate the models' performances by using the Area Under Curve (AUC) and TSS metrics. Furthermore, the selection of pseudo-absences from the study area outside the environmental range proved to lead to an inflation on the AUC-ROC score, hence causing an increase in the AUC-ROC value. Moreover, in cases when the species being studied is limited in its geographical range compared to the overall size of the research area, there is a possibility of overestimating the area under the receiver operating characteristic curve (AUC-ROC) (Lobo et al., 2008; Mtengwana et al., 2021). But in this study, it was not the case for the encroacher species of the KNP and surrounding communal areas.

The models used in this study demonstrated that the combination of both bioclimatic data and remotely sensed data can predict the species distribution with high accuracy and low sensitivity. Furthermore, the results show that the suitable habitat modelled with the bioclimatic variables are specific, because of the combination of the remotely sensed data which adds more nuanced spatial details to the overall pattern that is shown on the maps. Like most of the literature which conducted similar studies in different parts of the world, this study found that the inclusion of elevation in the bioclimatic predictors produced a generalized portrayals of environmental variability (Truong et al., 2017; Dai et al., 2020). The climate can also exhibit continuous variation across large regions, so climate-based distribution models alone cannot accurately describe fluctuations in species diversity across landscapes (Saatchi et al., 2008; West et al., 2016). The variation of geological variables generally represents the contrasting vegetation that is known in the invasive shrubs' species max temperature of warmest (bio5) are an essential condition for germination in some of the shrub species (e.g. Margreiter et al., 2020).

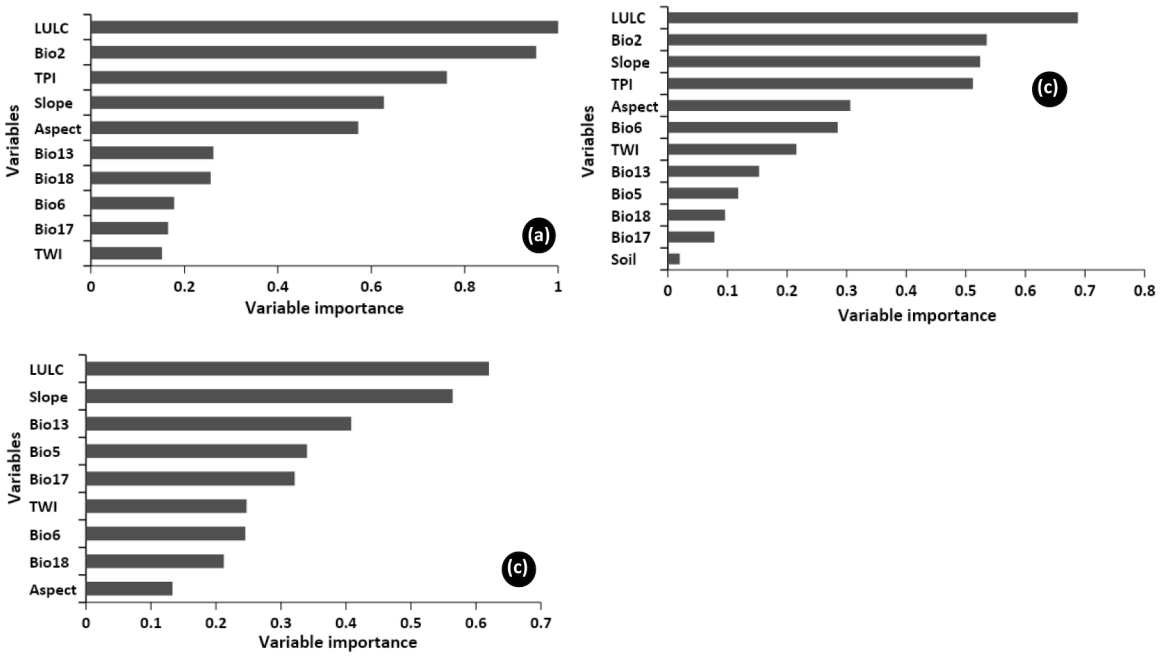


Fig. 4. Variable importance to measure for the prediction of invasive species (a) Random Forest (RF), (b) Maxent, and (c) BRT.

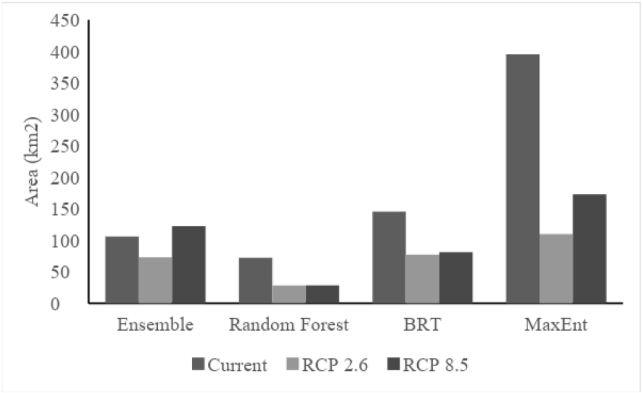


Fig. 5. Bush encroachment distributions in Kruger National Park (KNP) and neighbouring communal lands are estimated for current and future climate scenarios (RCP 2.6 and RCP 8.5).

Insufficient soil moisture due to low average rainfall hinders tree growth, while excessive rainfall causes dense woodlands with a mixed age distribution (Dai et al., 2020). By increasing soil moisture levels, young woody plant seedlings are able to survive and grow, enabling them to develop into bush coppices (Belayneh and Tessema, 2017). It is still unclear why biodiversity is unevenly distributed (Merritt et al., 2019). However, the distinctiveness of biodiversity hotspots can be attributed to their unique evolutionary history, which has been influenced by climatic and geological factors. Plant biodiversity clusters in areas with unique geology because of the diversity of habitats, as well as evolutionary pressures caused by significant changes in soil conditions, such as those found in ultramafic outcrops. The majority of countries in the region possess appropriate habitats for these species. As indicated by their lower variable importance scores, shrubs exhibit less pronounced environmental associations compared to other life forms. In contrast to other factors, climatic factors influenced shrub and native species distributions the most. This suggests shrubs can withstand a wider range of conditions. This could enhance their dissemination in response to climate change. Invasive species may experience a greater advantage as a result of global climate change (Dukes and Mooney, 1999), and rangelands serve as a notable illustration of this phenomenon. The climate projections for the region indicate that there will be higher annual temperatures and increased summertime precipitation. The climate model projections for southern Africa indicate a general decrease in rainfall, increased variability in both regional and temporal patterns, and an increase in the occurrence of extreme events such as droughts (IPCC 2014). This is achieved through the integration of presence-only occurrence records with bioclimatic and topographic characteristics. The potential for habitability rises alongside the expected levels of greenhouse gas concentrations across various scenarios. The results demonstrate that rainfall and temperature variables play a significant role in modelling the distribution of these bush encroacher species. Bravo-García et al. (2024) the findings confirm that the factors related to precipitation and temperature are crucial for modelling the spread of these bush encroacher species. The results support the model outputs of this study and highlight the importance of the variables analysed. During the model outputs, current distribution scenarios yield distribution maps that closely resemble the current maps delineating species distribution limits (Marggraff and Venter, 2020). The introduction of CMIP6, with its sophisticated characteristics akin to the model outputs of prior versions utilised for analysing historical climate changes in Southern Africa, offers robust projections. The CMIP6 models have deficiencies in accurately reproducing the observed variability in rainfall throughout Southern Africa. Precipitation exhibits significant spatial and temporal variability, often resulting in frequent occurrences of drought. Droughts exhibit variability in their duration, severity, and geographical scope. This increase in precipitation is particularly significant for the distribution of vine species, as all of them have shown positive associations with this variable.

The high degree of concordance in spatial predictions between models based solely on climatic predictors and those incorporating both climatic and remote sensing predictors, as observed in this study, along with the substantial variable importance scores attributed to climatic predictors in the combined models, further corroborates the indispensability of climate in determining the distribution of encroaching plant species. Comparable investigations have also demonstrated that the exclusive utilization of either climatic-derived or remote sensing-derived predictors frequently results in the overprediction of species

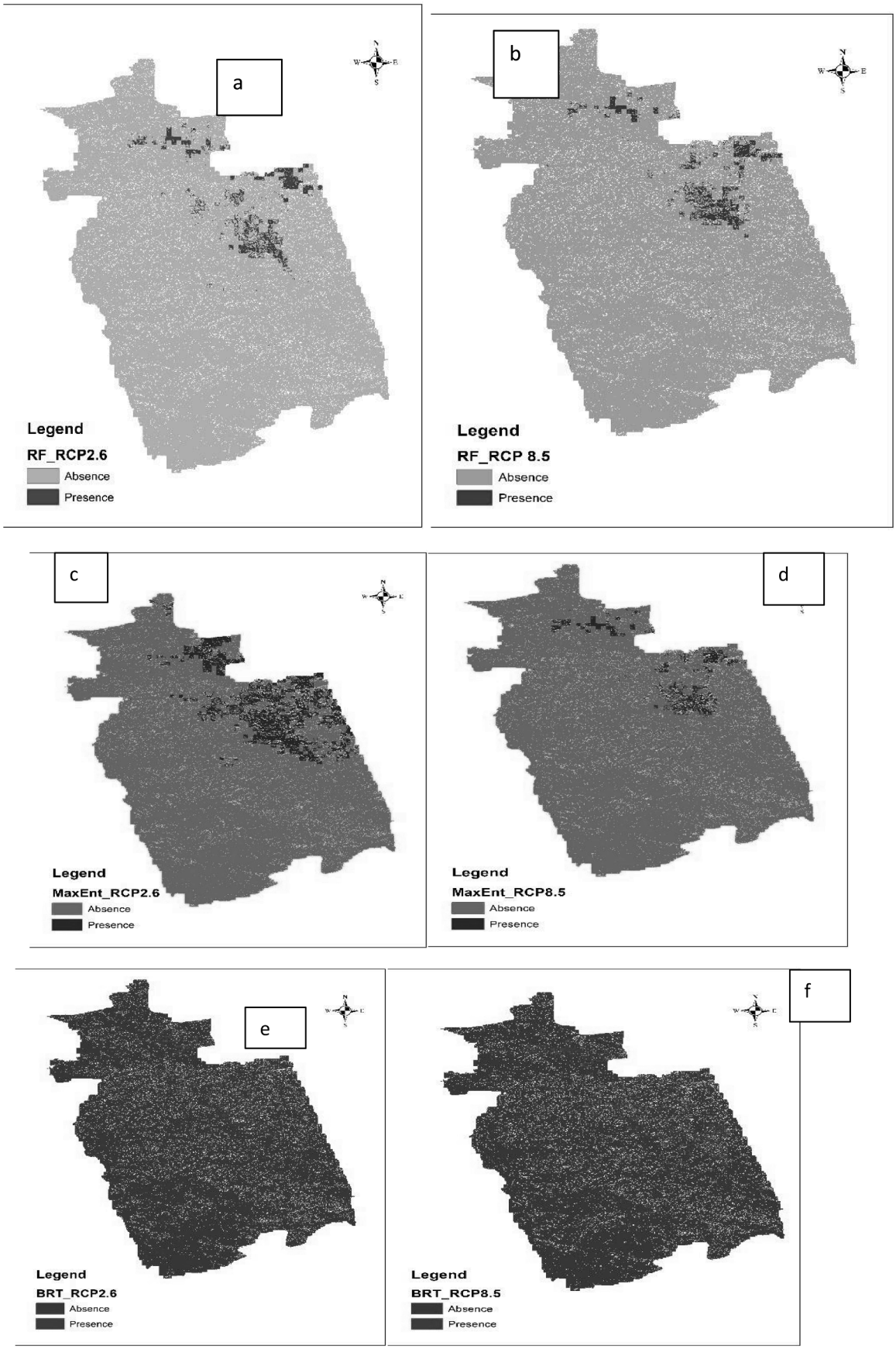


Fig. 6. illustrates the predicted suitability maps for encroacher species derived using Sentinel-2 satellite data, based on the three machine learning algorithms used. Each algorithms have a scale of 0 to 1 whereby 0 is absence and 1 represents presence.

distributions. Case et al. (2020) demonstrated that several plant species susceptible to drought and prone to encroachment into woody areas exhibit comparable traits to non-encroaching species. Although potential rooting depth decreased with increasing wet season length, mean annual rainfall had no systematic effect on rooting depth. Drought negatively impacts encroaching species, particularly those with root-suckering characteristics. The combination of this effect with increased fire frequency following intense rainfall events after drought periods suggests a potential decrease in encroacher populations (Case et al., 2020; Baloyi et al., 2024). 2024). Consequently, through Species Distribution Models (SDM), managers and local communities can implement preventive measures in anticipation of the Representative Concentration Pathway (RCP) 2.6 or RCP 8.5. This research demonstrated that climatic variables can be utilized to develop a more comprehensive understanding of bush encroachment for proactive management in these areas (protected and communal). Moreover, the study illustrated that the use of bioclimatic data can be advantageous in extracting information regarding the climatic conditions under which bush encroachment is likely to increase in semi-arid savanna rangelands. These results provide valuable insights into the effective management of bush encroachment and may be employed for prioritized monitoring.

5. Conclusion

An important factor in the spread of vegetation is elevation, possibly because terrain impacts soil texture, water availability, and temperature. It is expected that species and plant arrangements will vary across a range of environmental conditions in the research area. According to this study, bush encroachment is most likely to occur in regions with high elevations and high temperatures. The species distribution models exhibited a range of performance values, ranging from 0.61 to 0.99, which classified them as being of good to exceptional quality. The models were deemed optimal for forecasting the possible habitat of indigenous invasive species. Based on the Area Under Curve (AUC) and True Skill Statistics (TSS) metrics, the models' performance was assessed. By analysing the percentage contributions of each variable to the models for every climate scenario, the essential bioclimatic predictors of bush encroachment were identified. The results also show that RF and MaxEnt were predictive variables and that non-climatic predictors such as TPI, slope, and aspect were critical. Soil and TWI were the least influential variables. In contrast to the BRT, where land cover was the most relevant variable followed by slope, the least influential variables were TWI and aspect. Bush invasion can be considered a natural occurrence in the rocky regions of the research area. At a larger scale, the savanna system can be regarded as stable. Due to the impact of climate change, which includes reduced rainfall and higher temperatures, native invasive woody species will be able to spread and decrease in numbers easier. Using ensemble models, we are able to reduce spatial uncertainty of the models for each climate scenario, thus providing us with accurate spatial predictions. Figs. 1–6

Recommendations

- This study recommends that future studies should include factors such as soil depth and focus of one invasive species.
- The inclusion of anthropogenic activities which are site-specific and observed during field data collection.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is based on research supported in part by the National Research Foundation of South Africa [Grant Numbers: 138,170] and the authors acknowledge that findings and conclusions or recommendations expressed in this publication generated by the NRF-supported research are those of the author(s).

CRediT authorship contribution statement

Thabang Maphanga: Writing – original draft, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Cletah Shoko:** Writing – review & editing, Validation, Supervision, Conceptualization. **Mbulisi Sibanda:** Writing – original draft, Validation, Supervision, Methodology, Conceptualization. **Blessing Kavhu:** Writing – review & editing, Visualization, Software, Formal analysis, Data curation, Conceptualization. **Corli Coetsee:** Writing – review & editing, Investigation, Data curation, Conceptualization. **Timothy Dube:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to acknowledge NRF Black Academics Advancement Programme (BAAP) (grant number: 138170) for funding the research. The authors gratefully acknowledge assistance from the Scientific Services division of South African National Parks (SANParks) in the form of accommodation at the research camp at Skukuza and escorts by game guards during fieldwork and for registering the research.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.2025.111056.

Data availability

Data will be made available on request.

References

Ahmed, Z., Le, H.P., 2021. Linking information communication technology, trade globalization index, and CO 2 emissions: evidence from advanced panel techniques. Environ. Sci. Pollut. Res. 28, 8770–8781.

Akinwande, M.O., Dikko, H.G., Samson, A., 2015. Variance inflation factor: as a condition for the inclusion of suppressor variable in regression analysis. Open. J. Stat. 5 (07), 754.

Allouche, O., Tsoar, A., Kadmon, R., 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). J. Appl. Ecol. 43 (6), 1223–1232.

Arenas-Castro, S., Regos, A., Gonçalves, J.F., Alcaraz-Segura, D., Honrado, J., 2019. Remotely sensed variables of ecosystem functioning support robust predictions of abundance patterns for rare species. Remote Sens. (Basel) 11 (18), 2086.

Awuah, K.T., Aplin, P., Marston, C.G., Powell, I., Smit, I.P., 2020. Probabilistic mapping and spatial pattern analysis of grazing lawns in southern african savanna s using worldview-3 imagery and machine learning techniques. Remote Sens. (Basel) 12 (20), 3357.

Baloyi, T.M., Maphanga, T., Madonsela, B.S., Mongwe, Q.G., Malakane, K.C., Grangxabe, X.S., Gqomfa, B., 2024. Indigenous strategies for managing bush encroachment in rural areas of South Africa. Challenges. (Basel) 15 (3), 33.

Belayneh, A., Tessema, Z.K., 2017. Mechanisms of bush encroachment and its interconnection with rangeland degradation in semi-arid African ecosystems: a review. J. Arid. Land. 9, 299–312.

Booth, T., 2018. Mapping inclusion and exclusion: concepts for all? Towards Inclusive schools? Routledge, pp. 96–108.

Borges, J., Higginbottom, T.P., Cain, B., Gadiye, D.E., Kisingo, A., Jones, M., Symeonakis, E., 2022. Landsat time series reveal forest loss and woody encroachment in the Ngorongoro Conservation area, Tanzania. Remote Sens. Ecol. Conserv. 8 (6), 808–826.

Borges, J., Higginbottom, T.P., Symeonakis, E., Jones, M., 2020. Sentinel-1 and sentinel-2 data for savanna land cover mapping: optimising the combination of sensors and seasons. Remote Sens. (Basel) 12 (23), 3862.

Braunisch, V., Patthey, P., Arlettaz, R., 2016. Where to combat shrub encroachment in alpine timberline ecosystems: combining remotely sensed vegetation information with species habitat modelling. PLoS. One 11 (10), e0164318.

Bravo-García, J., Camarillo-Naranjo, J., Blanco-Velázquez, F.J., González-Peñaloza, F., Anaya-Romero, M., 2024. Mapping the potential habitat suitability and opportunities of bush encroacher species in Southern Africa: a case study of the SteamBioAfrica project. Front. Biogeogr. 17, e136222. https://doi.org/10.21425/fob.17.136222.

Buitenwerf, R., Bond, W.J., Stevens, N., Trollope, W., 2012. Increased tree densities in South African savannas:>50 years of data suggests CO 2 as a driver. Glob. Chang. Biol. 18 (2), 675–684.

Bunting, E.L., Munson, S.M., Bradford, J.B., 2019. Assessing plant production responses to climate across water-limited regions using google earth engine. Remote Sens. Environ. 233, 111379.

Curtis, B.A., Mannheimer, C.A., 2005. Tree atlas of namibia national botanic research institute, windhoek. Incl. Pic 623.

Dai, J., Roberts, D.A., Stow, D.A., An, L., Hall, S.J., Yabiku, S.T., Kyriakidis, P.C., 2020. Mapping understory invasive plant species with field and remotely sensed data in Chitwan, Nepal. Remote Sens. Environ. 250, 112037.

Diniz-Filho, J.A.F., Mauricio Bini, L., Fernando Rangel, T., Loyola, R.D., Hof, C., Nogués-Bravo, D., Araújo, M.B., 2009. Partitioning and mapping uncertainties in ensembles of forecasts of species turnover under climate change. Ecography. 32 (6), 897–906.

Dube, K., Nhamo, G., 2020. Evidence and impact of climate change on South African national parks. Potential implications for tourism in the Kruger National Park. Environ. Dev. 33, 100485.

Dube, O.P., Pickup, G., 2001. Effects of rainfall variability and communal and semi-commercial grazing on land cover in southern African rangelands. Clim. Res. 17 (2), 195–208.

Dube, T., Shoko, C., Sibanda, M., Madileng, P., Maluleke, X.G., Mokwatedi, V.R., Tibane, L., Tshebesebe, T., 2020. Remote sensing of invasive lantana camara (verbenaceae) in semiarid savanna rangeland ecosystems of south africa. Rangel. Ecol. Manag. 73 (3), 411–419.

Elliott, C.P., Tomlinson, S., Lewandrowski, W., Miller, B.P., 2024. Species distribution and habitat attributes guide translocation planning of a threatened short-range endemic plant. Glob. Ecol. Conserv. 51, e02915.

Fois, M., Cuena-Lombraña, A., Fenu, G., Bacchetta, G., 2018. Using species distribution models at local scale to guide the search of poorly known species: review, methodological issues and future directions. Ecol. Modell. 385, 124–132.

Früh, L., Kampen, H., Kerkow, A., Schaub, G.A., Walther, D., Wieland, R., 2018. Modelling the potential distribution of an invasive mosquito species: comparative evaluation of four machine learning methods and their combinations. Ecol. Modell. 388, 136–144.

Gallardo, B., Aldridge, D.C., González-Moreno, P., Pergl, J., Pizarro, M., Pyšek, P., Thuiller, W., Yesson, C., Vilà, M., 2017. Protected areas offer refuge from invasive species spreading under climate change. Glob. Change Biol. 23 (12), 5331–5343.

Hao, T., Elith, J., Guillera-Arroita, G., Lahoz-Monfort, J.J., 2019. A review of evidence about use and performance of species distribution modelling ensembles like BIOMOD. Divers. Distrib. 25 (5), 839–852.

He, Y., D'Odorico, P., De Wekker, S.F., 2015. The role of vegetation–microclimate feedback in promoting shrub encroachment in the northern Chihuahuan desert. Glob. Chang. Biol. 21 (6), 2141–2154.

Higginbottom, T.P., Field, C.D., Rosenburgh, A.E., Wright, A., Symeonakis, E., Caporn, S. J., 2018. High-resolution wetness index mapping: a useful tool for regional scale wetland management. Ecol. Inform. 48, 89–96.

Hu, Y., Si-Moussi, S., Thuiller, W., 2024. Introduction to deep learning methods for multi-species predictions. Methods Ecol. Evol.

Kass, J.M., Vilela, B., Aiello-Lammens, M.E., Muscarella, R., Merow, C., Anderson, R.P., 2018. Wallace: a flexible platform for reproducible modeling of species niches and distributions built for community expansion. Methods Ecol. Evol. 9 (4), 1151–1156.

Khanum, R., Mumtaz, A.S., Kumar, S., 2013. Predicting impacts of climate change on medicinal asclepiads of Pakistan using Maxent modeling. Acta Oecol. 49, 23–31.

Khazieva, E., Verburg, P.H., Pazúr, R., 2022. Grassland degradation by shrub encroachment: mapping patterns and drivers of encroachment in Kyrgyzstan. J. Arid. Environ. 207, 104849.

Li, J., Ravi, S., Wang, G., Van Pelt, R.S., Gill, T.E., Sankey, J.B., 2022. Woody plant encroachment of grassland and the reversibility of shrub dominance: erosion, fire, and feedback processes. Ecosphere 13 (3), e3949.

Ludwig, A., Meyer, H., Nauss, T., 2016. Automatic classification of Google Earth images for a larger scale monitoring of bush encroachment in South Africa. Int. J. Appl. Earth Obs. Geoinf. 50, 89–94.

Maphanga, T., Dube, T., Shoko, C., Sibanda, M., Gxokwe, S., 2024. Using multisource remotely sensed data and cloud computing approaches to map non-native species in the semi-arid savannah rangelands of Mpumalanga, South Africa. S. Afr. Geogr. J. 1–24.

Marggraff, P., Venter, M.P., 2020. Monitoring of Namibian encroacher bush using computer vision. Agri Eng. 2 (2), 213–225.

Moudrý, V., Bazzichetto, M., Remelgado, R., Devillers, R., Lenoir, J., Mateo, R.G., Lembrechts, J.J., Sillero, N., Lecours, V., Cord, A.F., Barták, V., 2024. Optimising occurrence data in species distribution models: sample size, positional uncertainty, and sampling bias matter. Ecography. 2024 (12), e07294.

Mtengwana, B., Dube, T., Mudereri, B.T., Shoko, C., 2021. Modeling the geographic spread and proliferation of invasive native plants (IAPs) into new ecosystems using multi-source data and multiple predictive models in the Heuningnes catchment, South Africa. GISci. Remote Sens. 58 (4), 483–500.

Mudereri, B.T., Dube, T., Adel-Rahman, E.M., Niassy, S., Kimathi, E., Khan, Z., Landmann, T., 2019. A comparative analysis of planetscope and sentinel sentinel-2 space-borne sensors in mapping striga weed using guided regularised random forest classification ensemble. *The International Archives of the Photogrammetry*. Remote Sens. Spat. Inf. Sci. 42, 701–708.

Munyati, C., 2017. The potential for integrating Sentinel 2 MSI with SPOT 5 HRG and Landsat 8 OLI imagery for monitoring semi-arid savannah woody cover. Int. J. Remote Sens. 38 (17), 4888–4913.

Nackley, L.L., Betzelberger, A., Skowno, A., West, A.G., Ripley, B.S., Bond, W.J., Midgley, G.F., 2018. CO2 enrichment does not entirely ameliorate Vachellia karroo drought inhibition: a missing mechanism explaining savanna bush encroachment. Environ. Exp. Bot. 155, 98–106.

Naimi, B., Araújo, M.B., 2016. sdm: a reproducible and extensible R platform for species distribution modelling. Ecography. 39 (4), 368–375.

Ncube, B., Shekede, M.D., Gwitira, I., Dube, T., 2020. Spatial modelling the effects of climate change on the distribution of Lantana camara in Southern Zimbabwe. Appl. Geogr. 117, 102172.

Ndlovu, H., Shoko, C., 2023. Assessing Lantana camara spatial distribution in response to rainfall and temperature variations in Inkomati catchment in Mpumalanga. . *Phys. Chem. Earth Parts A/B/C* 132, 103470.

O'connor, T.G., Puttick, J.R., Hoffman, M.T., 2014. Bush encroachment in southern Africa: changes and causes. Afr. J. Range Forage Sci. 31 (2), 67–88.

Oldeland, J., Dorigo, W., Wesuls, D., Jürgens, N., 2010. Mapping bush encroaching species by seasonal differences in hyperspectral imagery. Remote Sens. (Basel) 2 (6), 1416–1438.

Paliwal, A., Mhelezi, M., Galgallo, D., Banerjee, R., Malicha, W., Whitbread, A., 2024. Utilizing artificial intelligence and remote sensing to detect Prosopis juliflora invasion: environmental drivers and community insights in rangelands of Kenya. Plants 13 (13), 1868.

Palmer, A.R., Bennett, J.E., 2013. Degradation of communal rangelands in South Africa: towards an improved understanding to inform policy. Afr. J. Range Forage Sci. 30 (1–2), 57–63.

Poirazidis, K., Bontzorlos, V., Xofis, P., Zakkak, S., Xirouchakis, S., Grigoriadou, E., Kechagioglou, S., Gasteratos, I., Alivizatos, H., Panagiotopoulou, M., 2019. Bioclimatic and environmental suitability models for capercaillie (Tetrao urogallus) conservation: identification of optimal and marginal areas in Rodopi Mountain-Range National Park (Northern Greece). Glob. Ecol. Conserv. 17, e00526.

Ponce-Fontenla, S., Serrano, M., Carballal, R., Regos, A., 2021. Sentinel 2 images enable reliable prediction of fine-scale habitat dynamics of narrow endemic plant species in serpentine soils. Appl. Veg. Sci. 24 (4), e12614.

Ramoelo, A., Stolter, C., Joubert, D., Cho, M.A., Groengroeft, A., Madibela, O.R., Zimmermann, I. and Pringle, H., 2018. Rangeland monitoring and assessment: a review.

Riva, F., Martin, C.J., Galán Acedo, C., Bellon, E.N., Keil, P., Morán-Ordóñez, A., Fahrig, L., Guisan, A., 2024. Incorporating effects of habitat patches into species distribution models. J. Ecol. 112 (10), 2162–2182.

Schwager, P., Berg, C., 2021. Remote sensing variables improve species distribution models for alpine plant species. Basic Appl. Ecol. 54, 1–13.

Shafeian, E., Fassnacht, F.E., Latifi, H., 2021. Mapping fractional woody cover in an extensive semi-arid woodland area at different spatial grains with sentinel-2 and very high-resolution data. Int. J. Appl. Earth Obs. Geoinf. 105, 102621.

Shekede, M.D., Murwira, A., Masocha, M., Gwitira, I., 2018. Spatial distribution of Vachellia karroo in Zimbabwean savannas (southern Africa) under a changing climate. Ecol. Res. 33, 1181–1191.

Shen, X., Liu, Y., Liu, B., Zhang, J., Wang, L., Lu, X., Jiang, M., 2022. Effect of shrub encroachment on land surface temperature in semi-arid areas of temperate regions of the Northern Hemisphere. In: Agric. For. Meteorol., 320, 108943.

Souverijns, N., Buchhorn, M., Horion, S., Fensholt, R., Verbeeck, H., Verbesselt, J., Herold, M., Tsendbazar, N.E., Bernardino, P.N., Somers, B., Van De Kerchove, R., 2020. Thirty years of land cover and fraction cover changes over the Sudano-Sahel using Landsat time series. Remote Sens. (Basel) 12 (22), 3817.

Steiner, M., Huettmann, F., Bryans, N., Barker, B., 2024. With super SDMs (machine learning, open access big data, and the cloud) towards more holistic global squirrel hotspots and coldspots. Sci. Rep. 14 (1), 5204.

Symeonakis, E., Karathanasis, N., Koukoulas, S., Panagopoulos, G., 2016. Monitoring sensitivity to land degradation and desertification with the environmentally sensitive area index: the case of lesvos island. Land Degrad. Dev. 27 (6), 1562–1573.

Thuiller, W., Lafourcade, B., Engler, R., Araújo, M.B., 2009. BIOMOD–a platform for ensemble forecasting of species distributions. Ecography. 32 (3), 369–373.

Truong, T.T., Hardy, G.E.S.J., Andrew, M.E., 2017. Contemporary remotely sensed data products refine invasive plants risk mapping in data poor regions. Front. Plant Sci. 8, 770.

Van Leeuwen, W.J., Davison, J.E., Casady, G.M., Marsh, S.E., 2010. Phenological characterization of desert sky island vegetation communities with remotely sensed and climate time series data. Remote Sens. (Basel) 2 (2), 388–415.

Wang, L., Diao, C., Lu, Y., 2024. The role of remote sensing in species distribution models: a review. Int. J. Remote Sens. 1–25.

Ward, D., 2005. Do we understand the causes of bush encroachment in African savannas? Afr. J. Range Forage Sci. 22 (2), 101–105.