

PREDICTING THE SUITABLE HABITAT OF THE INVASIVE *XANTHIUM STRUMARIUM* L. IN SOUTHEASTERN ZIMBABWE

CHIKURUWO, C.* – MASOCHA, M. – MURWIRA, A. – NDAIMANI, H.

*Department of Geography and Environmental Science, University of Zimbabwe
P O Box MP 167, Harare, Zimbabwe*

**Corresponding author*

e-mail: charonchikuruwo@gmail.com; tel: +26-377-284-0285

(Received 18th Aug 2016; accepted 14th Oct 2016)

Abstract. Biological invasions have negatively affected virtually all ecosystems on earth. Thus, reducing the spread of such invasive species is important. A few broad strategies are available on how to reduce the spread of invasive species. The aim of this study is to predict the suitable habitat of the invasive species *Xanthium strumarium* and to identify the factors associated with its spread in order to improve efforts against invasive species. Presence (53 plots) and absence (52 plots) data for building the prediction were collected from Gonarezhou National Park and the adjacent Malipati communal land, Zimbabwe in April 2013. Ten modelling techniques were employed namely; Generalized linear model, General additive model, Classification tree analysis, Surface range envelope, Multivariate adaptive regression splines, Random Forest, Gradient boosting machines, Maximum entropy, Artificial neural networks and an ensemble model. Environmental factors related to the establishment success of *Xanthium strumarium* were also identified. Results of this study indicate that machine learning based techniques performed better at predicting the occurrence of *Xanthium strumarium*, although the ensemble prediction outperformed all the individual models. *Xanthium strumarium* was predicted to occur in areas that receive direct radiation from the sun, thus, aspect was identified as a critical factor in habitat selection for *Xanthium strumarium*.

Keywords: *area, climate change, plant population, bioclimatology, prediction, habitat, invasive*

Introduction

Rough cocklebur (*Xanthium strumarium* L.), is a broadleaved, tap rooted, forb whose stems are erect, ridged, rough and hairy (Hare, 1980). The species belongs to the Asteraceae family and it reproduces annually (Venodha, 2016). Rough cocklebur often appears in thick mono-specific stands usually in low-lying riparian areas and in agricultural fields (Marwat et al., 2010). The native habitat of *Xanthium strumarium* is North America and Argentina (Nel et al., 2004; Oksanen et al., 2016). Several studies have identified *Xanthium strumarium* to be invasive in southern African savannas, (Ekeleme et al., 2000; Holmes et al., 2008; Sithole et al., 2012). The success of *Xanthium strumarium* as an invasive species has been attributed to competitive genetic makeup (Gray et al., 1986) among other adaptive capabilities that the species possesses. However, the extent of *Xanthium strumarium* invasion in southeastern Zimbabwe is yet to be established.

Studies on species invasion have highlighted that the invasion process is reliant upon both the invasive capabilities of the species and the invasibility of the ecosystem. Hence it is important to consider the factors which are associated with habitat invasibility when predicting the occurrence of an invasive species. Given the fact that introduced species are more successful in disturbed ecosystems where competition has either been reduced or totally eliminated (Dukes, 2001; Eppstein and Molofsky, 2007), it is vital to consider

the natural and human factors that are associated with the occurrence of *Xanthium strumarium*.

Information on the potential distribution of invasive species is essential as it may improve the quality of control efforts against invasion. Single model species distribution models are often affected by bias (Araújo and New, 2007). Thus more recently, researchers have ventured into multi-model approaches and ensemble forecasting (Ochoa-Ochoa et al., 2016). The use of several models could remove model-based bias and improve the chances of model suitability. Although multi algorithm species distribution modelling has gained popularity among spatial geographers, more work still needs to be done in southern African savannas, especially for invasive species.

A limited number of studies exist that attempt to identify the potential distribution of invasive species, *Xanthium strumarium* included. The predictions are especially important given that savannas play a vital role in human welfare and economic development. Half of the biome's human population is directly dependant on it for their sustenance (Sankraan et al., 2005). Also, in the southeastern low-veld of Zimbabwe cattle farming is the main economic activity supporting livelihoods (Zengeya et al., 2011; Chigwenhese et al., 2016), and the area hosts a considerable number of wild mammals in the Gonarezhou national park, information on the suitable habitat of the invasive species is critical. *Xanthium strumarium* is unpalatable to both livestock and wildlife, thus, if left unmanaged the invasive species could cause major economic and biological losses. In that regard, this study aims to predict the occurrence of invasive species *Xanthium strumarium* in southeastern Zimbabwe. This prediction will be made by using an ensemble model which will incorporate the best performing algorithm among: Gradient boosting machines (GBM), Artificial neural networks (ANN), Maximum entropy (MaxEnt), Random forests (RF), General linear model (GLM), General additive machines (GAM), Multivariate adaptive regression splines (MARS) Surface range envelope (SRE) and Classification tree analysis (CTA). In this study, the ensemble model was also used to identify the environmental factors that are associated with the distribution of *Xanthium strumarium*.

Literature review

Invasive species are species that have an ability to spread on their own, causing harm to other species and their environment when introduced to new environments (Rejmánek et al., 2013; Richardson and Pyšek, 2006). These species are in most cases 'alien' in the sense that they are not native in the regions that they become invasive. Mostly they would have been transported by humans (Vitousek et al., 1997). The transportation of organisms by humans through global trade and globalization which has now increased has led to the breaking of natural biogeographic barriers (Clout and Williams, 2009). Organisms have found a way through humans to migrate from geographic areas of superior competition to areas where they can easily dominate and cause enormous damage. It is important however, to note that not all introduced species have the potential to become invasive. In the case that they become invasive, they often spread irreversibly, and damages increase over time (Keller et al., 2009) non-indigenous species (e.g., plants or animals) that adversely affect the habitats they invade economically, environmentally or ecologically (Dassonville et al., 2008). In recent years, economists are also making efforts to understand how invasive species affect

economic systems, and in how invaders can be controlled to increase societal welfare (Keller et al., 2009). Such is the extent of the harm caused by invasive species.

Ever since 1958, scholars have been active in the field of invasion science (Rejmanek et al., 2005) and management has been one of the popular topics. Managing biological invasions involves identifying passageways of transport or the methods of transport and the attributes of the species that guarantees success when faced with barriers. It also involves using appropriate strategies to prevent, eradicate, or control the species of concern (Clout and Williams, 2009). It is important to obtain accurate assessments of location and abundance of invasive species so that managers can set these priorities and have the information to quickly and effectively combat the invaders, (Clout and Williams, 2009). It is also important to identify barriers to invasion and habitats where an invasive species cannot persist or cause much harm, (Clout and Williams, 2009). Given that invasive alien species represent a global threat, it is necessary to make the best use of all information available (Capdevila-Argüelles and Zilleti, 2005) to manage their spread.

More recently, research has ventured into modeling the potential habitats of invasive species using species distribution models. A species distribution model (SDM) is a model that relates species distribution data (occurrence or abundance at known locations) with information on the environmental and/or spatial characteristics of those locations (Elith and Leathwick, 2009). SDMs provides insight not only into the mechanisms but also potential hotspots for invasion. They explore the species-environment relationship further and thus provide useful information for protecting environments against invasive species. However, if species distribution models are to be used for early detection and management of invasive species in conservation practice, their accuracy and correct interpretation is crucial to minimize the ecological impact and economic cost of biological invasions. (Václavík and Meentemeyer, 2009). Thus, it becomes crucial to assess several species distribution modeling algorithms in order to determine the most accurate.

Although studies date back to the 1950's, most studies, have however been geographically biased (Pysek et al., 2008) towards regions in Britain, the Americas and closer to home South Africa. This bias has been a setback for the development of globally representative data sets for the management of invasive species. Hence the need to address the problem of invasive species for local ecosystems, which is the aim of this study.

Methods

Study site

This study was conducted in the Gonarezhou National park (GNP) and the adjacent Malipati communal land (MCL). The study area is located in the southeastern low veld region of Zimbabwe between latitudes 21° 00' and 22° 15' south and longitudes 32° 30' and 33° 15' east (*Fig. 1*). The southeastern low veld is semi-arid, with average monthly maximum temperatures of 25.9 C in July and 36 C in January and average monthly minimum temperatures range between 9 C in June and 24 C in January (Mott et al., 1984). Characteristic vegetation in the area is savanna woodland dominated by *Colophospermum mopane*, interspersed with *Androstachys johnsonii* and *Combretum apiculatum* on ridges and in riverine areas.

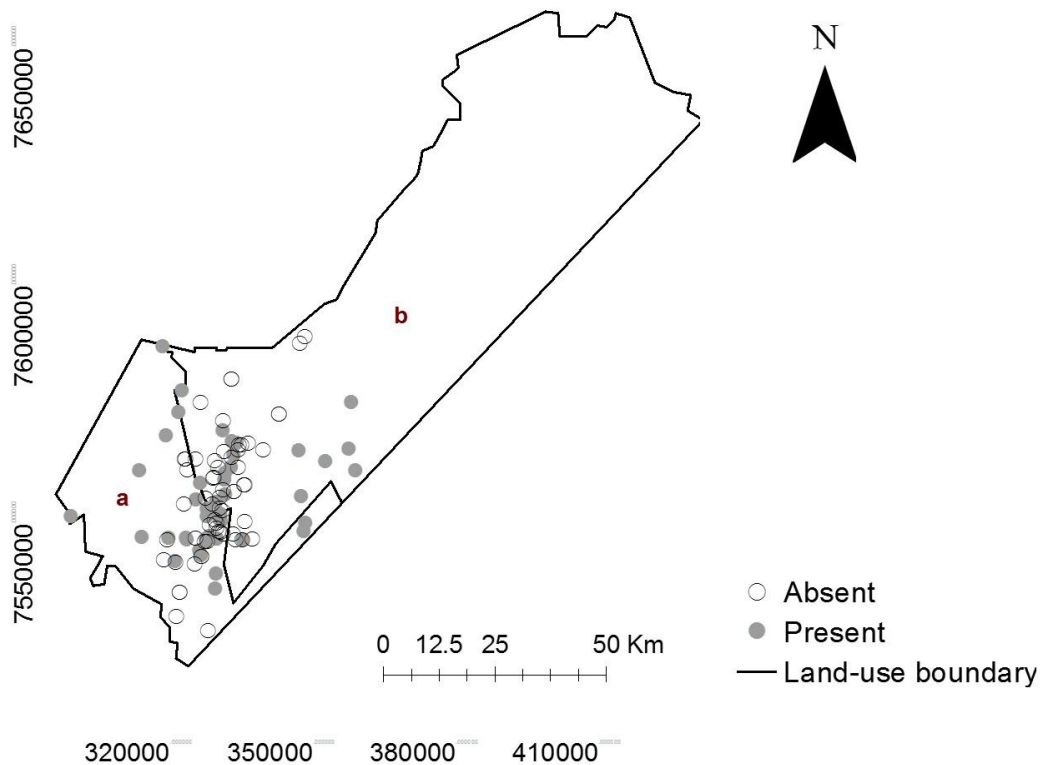


Figure 1. Distribution of presence and absence data points in (a) Malipati communla lands and (b) Gonarezhou National Park.

The study was conducted in these separate but adjacent sites so as to account for the effect of land use on the distribution of *Xanthium strumarium*. The sites, GNP and MCL have similar biophysical characteristics (Mombeshora and Bel, 2009) albeit they differ in their main land uses and, therefore, disturbance regimes. GNP is a protected area where wildlife conservation is the main land use activity while MCL is a communal land where cropping and livestock grazing are the main economic activities.

Presence and absence data

Species presence and absence data was collected in April 2013. The study area was stratified into two strata namely a protected landscape (GNP) and communal landscape (MCL). Thereafter, sampling units were randomised in the sites within a GIS environment. Vegetation surveys were conducted in 105 randomised plots. In each plot, all present rooted species were recorded. Most species were identified with expert knowledge in the field with the help of the staff from the Zimbabwe Parks and Wildlife Management authority. The remainder of the species were identified using the help of the National Botanical gardens database and experts. 53 locations were identified as present locations for *Xanthium strumarium*, and 52 were absent locations (*Fig 1*).

Environmental variables

Figure 2 presents maps of four environmental variables namely distance from homesteads, distance from roads, normalised difference vegetation index (NDVI), and

aspect selected to predict the distribution of *Xanthium strumarium*. The selection of these variables was based on the ecological niche theory and the invasive species' dependency on disturbance regimes.

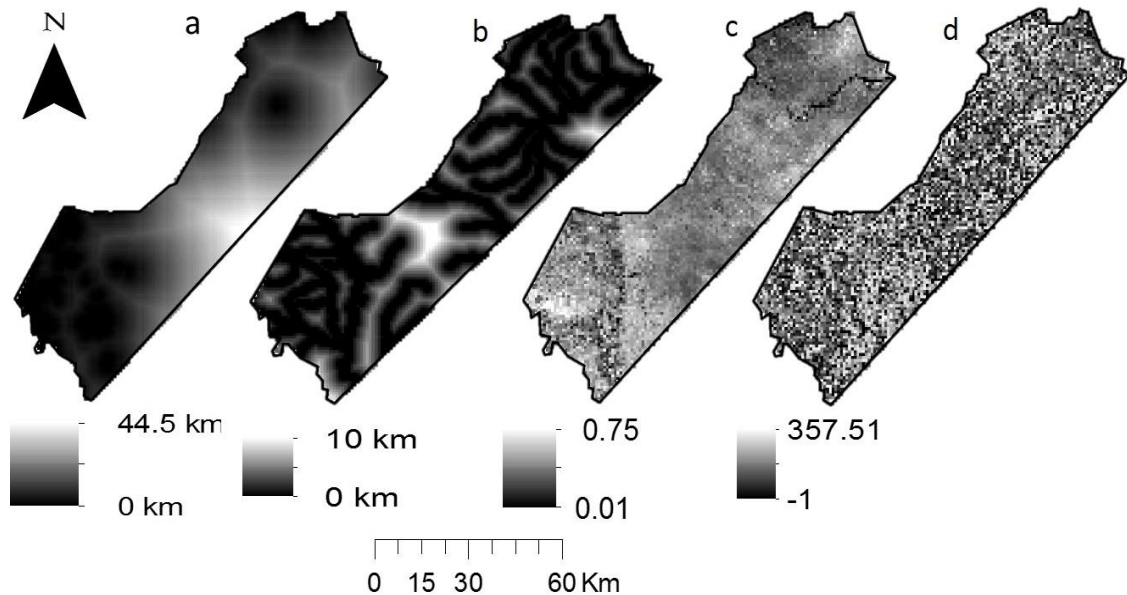


Figure 2. Environmental variables used to build a predicted location of *Xanthium strumarium*. (a) Distance from homesteads, (b) Distance from rivers, (c) NDVI and (d) Aspect.

To capture the potential role of human and livestock disturbance in the distribution of *Xanthium strumarium*, we calculated inverse the distance from settlements. Many studies identify anthropogenic disturbance as one of the major drivers of invasive species establishments in most ecosystems (Barrat-Segretain, 2001; Rejmanek and Sandlund, 1999; Foxcroft et al., 2011; Mortensen et al., 1993; Kumschick et al., 2012). The use of distance from settlements as a proxy of anthropogenic disturbance was justified on the basis that human-mediated disturbance activities such as crop production and livestock production occur in close proximity to where people live and keep their livestock. A point map of all settlements (that is dwelling units) in the study area was created by digitizing these as point features using a Landsat 8 image made available via Google Earth (www.googleearth.com) as the base layer. Next, the Euclidean distance from each dwelling unit was calculated using the Spatial Analyst tool in ArcGIS 9.2 (Wang et al., 2001). Then, an inverse distance map was derived from the Euclidean distance map using the in-built inverse distance Algorithm in Spatial Analyst tool. The weighted map of distance from homesteads was used as a proxy for human and livestock disturbance.

Distance from rivers was used as a proxy for moisture availability, which is crucial to the establishment of *Xanthium strumarium* given that the study site is a semi-arid savanna hence moisture availability is a limiting factor. Distance from rivers also served as a surrogate for propagule dispersal since it is known that the burs that carry the seeds of *Xanthium strumarium* are buoyant and can therefore be transported in flowing water, (Davies and Johnson, 2012). Also considering the fact that the burs of *Xanthium strumarium* can be dispersed by animals via attachment it makes ecological sense to

assume that the propagules of *Xanthium strumarium* are more likely to be dispersed more efficiently along riparian areas as animals both domestic and wildlife come to drink water. The inverse distance from rivers was calculated from rasterized polylines of the river network using the standard Euclidean distance tool in ESRI ArcGIS 9.2 (Wang et al., 2001).

The third environmental predictor used for modelling the distribution of *Xanthium strumarium* is NDVI, which is a widely used proxy for primary production, which in turn represents the competitive ability of species already established in the study site (Hobbs and Huenneke, 1992). Previous studies have highlighted a negative relationship between competition and invasion (Kauffman et al., 1983; Mortensen et al., 1993). After the image was processed for atmospheric corrections and calibrated NDVI was calculated from a Landsat TM image acquired in April 2013. Formally, NDVI was calculated as:

$$NDVI = \frac{(NIR - R)}{(NIR + R)} \quad (\text{Eq.1})$$

Where NIR is the Near Infrared band and R is the red band of the Landsat TM image.

Aspect was included because previous studies have highlighted that *Xanthium strumarium* establishes more successfully in areas where there is light, (especially the red band of light), as opposed to shaded areas (Sharkey and Raschke, 1981). Thus aspect was included as a proxy for light availability. Aspect was derived from a 30 m shuttle radar topography mission (SRTM) digital elevation model with Spatial Analyst-surface tool in ArcGIS 9.2 (Wang et al., 2001). The DEM was downloaded from (www.usgs.gov).

Prediction of suitable habitat

The study used a range of machine learning based, classification based, regression based and surface range envelope techniques to predict the distribution of *Xanthium strumarium* based on the presence/absence data and four environmental factors described earlier. Three regression-based modelling techniques were used in this study, namely, the General linear model (GLM), Generalized additive model (GAM) and the Multivariate additive regression splines (MARS). GLM predicts from the linear relationship of a given variable to a set of explanatory variables. GAM is almost similar to GLM except it adopts a smoothing effect which improves the output. MARS derives its predictions from recursive partitioning to derive a good set of basic functions. Thus each of these techniques were included for their individual competences.

Also, four machine learning techniques were used in this study. These were: Artificial neural networks (ANN) which were developed from the mathematical models of the human nervous system. Prediction is accomplished by adjusting to cause the overall network to output appropriate results (Abraham, 2005). The Random forests (RF) which operate with a set of trees called ‘forests’. To make a classification each tree gives a classification and the ‘forest’ votes on the classification. Random forests chooses the classification which received the most votes (Cutler et al., 2007). Maximum entropy (MaxEnt) which uses the maximum entropy theory which estimates the maximised probability distribution given a set of constraints (Phillips et al., 2006).

The above mentioned regression-based and machine-learning techniques were complemented by a classification based technique, Classification tree analysis (CTA). Classification methods consist of recursive partitions of dimensional space defined by predictions into groups that are as homogeneous as possible, as such classification have their strengths. The ninth modelling technique used was the surface range envelope modelling technique (SRE). It involved the use of interpolation of various climatic parameters to produce a model of the best fit of those parameters (Beaumont et al., 2005).

All of the nine modeling algorithms were implemented with the default settings of the BIOMOD2 package (Thuiller et al., 2014) in the R computing environment (Regnier and Harrison, 1993). BIOMOD2 was also set to split 80% of the data to be used for model calibration and 20% for model validation. Presence and absence data points were set such that they contributed equally towards the prediction. BIOMOD2 was set to run 3 permutations in estimating variable importance.

Ensemble modelling

A consensus or ensemble model (EM) which incorporated results from the best performing models was then executed, the results from which were regarded the final prediction. This was done within the Ensemble Modeling environment in the R statistical computing software. The total consensus model (Thuiller et al., 2014) uses the datasets used in building individual models, as well as the individual models. A TSS threshold of 0.6 was used to determine models that were included in building the ensemble model. Thus, models that had TSS of <0.6 were not included in the EM, thus ensuring that the consensus was robust. The weighted mean of probabilities algorithm which gives more weight to individual models with better TSS scores was adopted as the final prediction of the location of *Xanthium strumarium* in this study.

Model evaluation

Ten model evaluation metrics namely the Receiver operating curve (ROC), Cohen's Kappa (KAPPA), True skill statistic (TSS), False alarm ratio, (FAR) Success ratio (SR), Accuracy (fraction correct), Bias score (frequency bias), Probability of detection (POD), Critical success index (CSI), and Equitable threat score (ETS) are available in BIOMOD2 and can be used to determine the models that provided accurate predictions, given that a species distribution model is only useful if it is robust (Václavík and Meentemeyer, 2009). However, to strike a balance between the need for multiple test statistics which improves the robustness of assessments (Elith and Graham, 2009) and to avoid redundancy by using all of them, two of the commonly used model evaluation metrics namely ROC and TSS were used. ROC measures the ability of a model to discriminate between presence and absent points. Values of the ROC statistic range from 0 to 1, where < 0.5 represent predictions which are not better than random. The TSS statistic ranges from random (-1) to perfect agreement ($+1$) (Guisan and Thuiller, 2005). The advantage of using TSS is that it has all the good properties of KAPPA yet it is not sensitive to prevalence (Allouche et al., 2006).

Variable importance

The relative importance of each predictor variable used in building the ensemble model was also calculated using the BIOMOD2 package (Thuiller et al., 2014). In the evaluation procedure, BIOMOD first created a standard prediction by using all the four

environmental variables. Thereafter, it generated other predictions, with one independent variable randomised. Thus, each variable was evaluated using the correlation cores of the standard prediction and each of the random predictions (Capinha and Anastácio, 2011). If the standard and random predictions are highly correlated, then the variable's relative importance is low (Capinha and Anastácio, 2011). The relative importance values ranged from 0 (lowest importance) to 1 (highest importance).

Curves representing the predictive ability of each environmental variable were also used to assess the response of the model to the variable. The BIOMOD2 response.plot2 function (Thuiller et al., 2014), was used to plot the response curves. To build the response curves, all other variables except the target variable were set constant to a fixed value (the mean) while the target variable for which the curve was being plotted was the only one that was allowed to vary. Thus, the variations and the curve attained showed how the model was sensitive to that specific predictor variable.

Results

Predicted occurrence of Xanthium strumarium

Model performance varied among the nine individual modelling techniques used, (Fig 3). The random forest (RF) and GBM, performed well in predicting the occurrence of *Xanthium strumarium* in Gonarezhou National Park. These two modelling techniques had high TSS and ROC values exceeding 0.6. Consequently, these algorithms were selected for building the ensemble projection, which is shown in Figure 3 performed best with a TSS score of 0.87. The predictions made by MARS, MaxEnt, GLM, and SRE, were not better than outcomes of random predictions, given their TSS scores were less than 0.5.

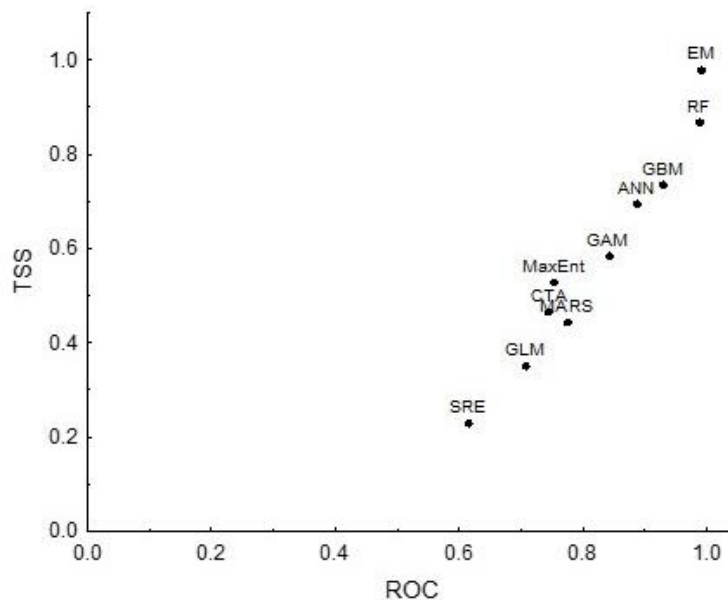


Figure 3. Scatterplot showing the performance of ten modelling techniques used to predict the occurrence of the invasive species *Xanthium strumarium* in Gonarezhou National Park of Zimbabwe. The x-axis represents True skills statistic (TSS) and the y-axis shows Receiver operating curve (ROC). Distance from homesteads, (b) Distance from rivers, (c) NDVI and (d) Aspect.

The ensemble projection illustrates that probability of occurrence of *Xanthium strumarium* was overall higher in the protected national park compared to the adjacent communal farmland (Fig.4).

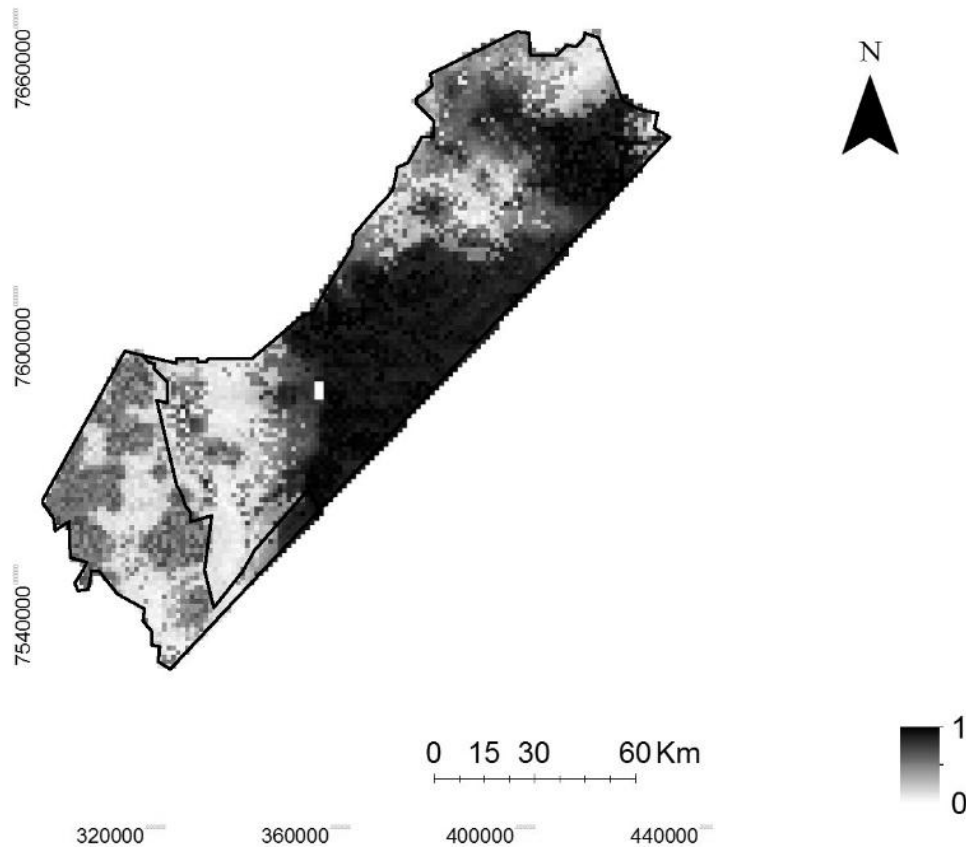


Figure 4. The predicted *Xanthium strumarium* suitable habitat in Malipati Communal lands and Gonarezhou national park derived from the Ensemble

Predicted occurrence of *Xanthium strumarium*

Results show that aspect, with a relative importance of 0.486 was the most important variable influencing the results of the ensemble prediction, followed NDVI and distance from settlements with relative importance values of 0.303 and 0.277, respectively. The predictor variable with the lowest relative importance of 0.021 was distance from rivers.

Results (Fig. 5) illustrate the response curves of *Xanthium strumarium* to the four factors considered important in predicting the species' probability of occurrence in Gonarezhou National Park. It can be observed that in general, probability of occurrence of *Xanthium strumarium* increases with aspect and there is a local maximum centered around 100°. The response curves for distance from settlements indicates that the highest peak in the probability of *Xanthium strumarium* occurred within a distance of 2-km from settlements. After this distance, the probability of *Xanthium strumarium* fluctuates but an increasing trend is discernible up to a distance of 15 km from settlements.

With regard to NDVI, results in Figure 5, illustrate that the probability of occurrence of *Xanthium strumarium* decreases with increasing NDVI.

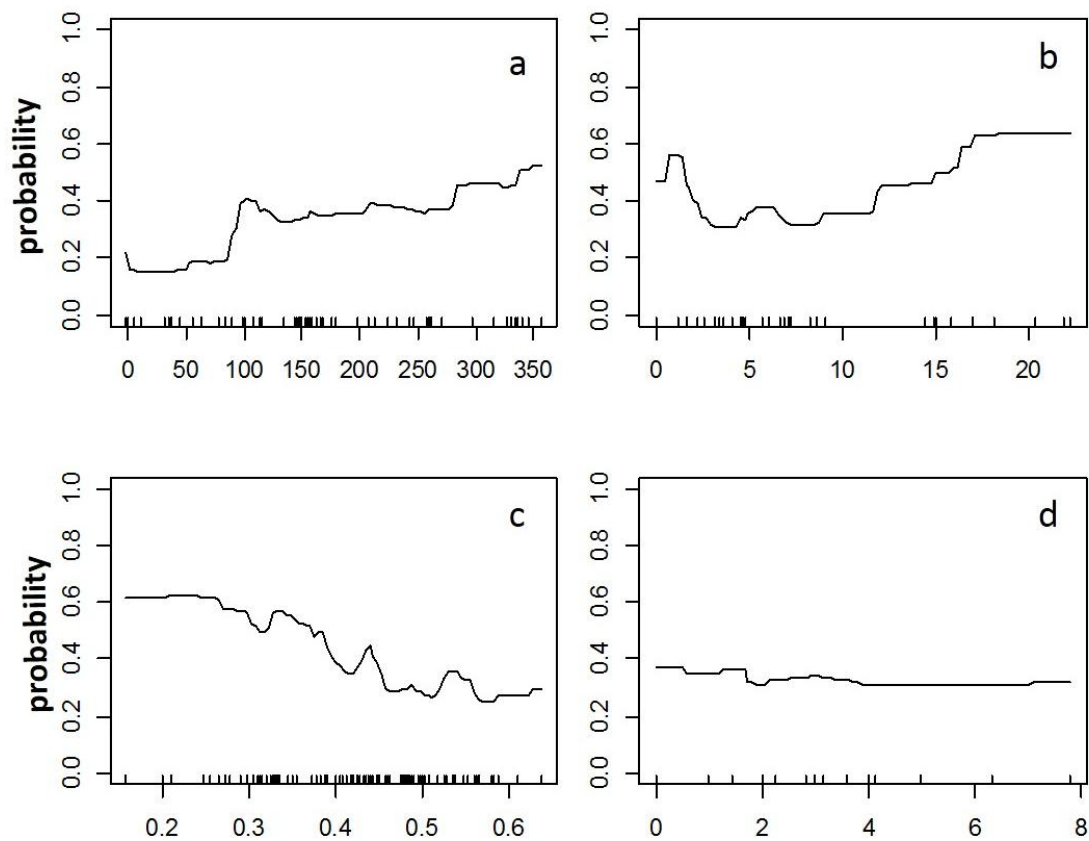


Figure 5. Response curves of *Xanthium strumarium* (a) aspect (b) distance from settlements (c) NDVI (d) Distance from rivers

Discussion

The aim of this study was to make a prediction of the locations where an invasive species *Xanthium strumarium* may occur within and around a protected area in southeastern Zimbabwe. Our hypothesis stated that since human disturbance is known to promote invasion (Chown et al., 2005; Addo-Fordjour et al., 2009; Holland et al., 2011), factors associated with human disturbance in conjunction with biotic factors improve the performance of models in predicting the occurrence of exotic species. The prediction produced by the ensemble model was fairly accurate with both the TSS and ROC more than 0.8 (Fig. 2). The prediction suggested that areas within the protected area have a higher probability of *Xanthium strumarium* occurrence than those in the communal area (Fig. 4). This result implies the need for improved efforts in managing the spread of *Xanthium strumarium* inside the protected area since conditions inside the national park is conducive to the growth of the species.

The study also intended to make a comparison of the performance of modeling techniques when predicting the location of *Xanthium strumarium* in and around a protected area in southeastern Zimbabwe. As expected, the ensemble model provided a superior prediction to all the other algorithms in this study. The superior performance of the consensus model could be attributed to the fact that it combines the outcomes of several good predictions to come up with one that has a better accuracy (Mestre et al., 2015; Ochoa-Ochoa et al., 2016). According to the accuracy assessment parameters employed in this study, RF performed better than the other algorithms. In fact, the top

two algorithms RF and GBM are machine learning techniques. The success of machine learning algorithms could be attributed to the fact that learning based models learn from the existing data thus, they do not require an a priori model (Kasabov, 1996). Hence their predictions are unbiased and mostly turn out to represent nature more accurately.

The success of machine learning techniques in this study is not unique. Other studies, (Guisan et al., 2006; Austin, 2002; Austin et al., 2006) which also made comparisons of species distribution modelling techniques have also cited the strength of machine-based algorithms. Some of the algorithms, however, produced poor predictions regardless of the fact that similar variables were used for all the models. The differences in the predictions of the ensemble model and SRE model illustrates the importance of accurate species distribution modelling of invasive species. Accurate predictions of invasive species locations allow for targeted application of limited resources; especially in African savannas where resources are constrained. Therefore, from the results, it is deducible that the use of appropriate algorithms may improve efficiency in invasive species management. The poor performance of these algorithms in this study regardless of previous success in other studies proves that no single model works best for all species, in all areas, at all spatial scales, and over time (Jarnevich et al., 2015). Thus, from these results, we deduce that a multi-model approach is important for eliminating problems associated with model bias in species distribution modelling.

Another objective of this study was to investigate the environmental variables that were most influential in predicting the occurrence of *Xanthium strumarium*. The results suggest that aspect was the most important factor in building the predicted location of *Xanthium strumarium*. This finding implies that among the factors that were considered, the direction in which the slope is facing is most critical for the success of *Xanthium strumarium*. Several studies have established that quality of light reaching a *Xanthium strumarium* plant is important for its growth (Tucker and Mansfield, 1971; Wang et al., 2001; Mott et al., 1984). Therefore, given its preference for direct sunlight, management strategies could prioritize surfaces that have direct access to sunlight to reduce the survival rate of new introductions.

It is not surprising that the results suggest distance from homesteads to be the second most important variable in building the model for predicting the occurrence of *Xanthium strumarium*. In fact, the response curve for the distance from homesteads shows a rise of probability within 1-2 km of homesteads. Other studies on invasive species have identified humans as a source of propagules (Mack and Lonsdale, 2001; Brown and Peet, 2003; Farji-Brener and Ghermandi, 2008; Thébaud and Simberloff, 2001; Wichmann et al., 2009) thus, the human environment acts as the epicentre for the spread of invasive species. Again, *Xanthium strumarium* is an agricultural weed, land tillage activities that occur in the communal area adjacent to Gonarezhou systematically creates niches for the establishment of *Xanthium strumarium*.

Also, given that the SELV of Zimbabwe is dry savanna, and animal husbandry is the main economic activity, the livestock numbers in the area may promote the establishment of *Xanthium strumarium* in two ways. The first way is through overgrazing (Roath and Krueger, 1982) which may disturb the natural vegetation and create niches for potential establishment of *Xanthium strumarium*. Secondly, is through propagule dispersal. *Xanthium strumarium* seeds are dispersed by water and by attaching to animal fur. Thus large amounts of livestock improve the transport of *Xanthium strumarium* hence it spreads over large areas quickly. It is thus critical to

consider the critical role that humans are playing in facilitating the establishment of invasive *Xanthium strumarium*.

Xanthium strumarium responded negatively to competition just as previous studies (Stachowicz et al., 1999; Currie and Paquin, 1987; Foxcroft et al., 2011) have established the competition avoidance nature of most invasive species. Studies on ecosystem invasibility have established that invasive species generally avoid competitive patches as a competition avoiding strategy (Stachowicz et al., 1999; Bartomeus et al., 2012). As such, the response of *Xanthium strumarium* to NDVI may imply that *Xanthium strumarium*'s strength is its capability to utilise niches left over by disturbance events. This is evident in the fact that the species follows an annual lifecycle (Venodha, 2016) and also one plant can produce up to several hundreds of propagule. Thus, it can regenerate quickly and invade communities in very short amounts of time. To that end, it is important to acknowledge the role that disturbance plays in the establishment of invasive *Xanthium strumarium* in order to improve efforts of managing the spread of this invasive species.

Although rivers had the least influence in building the ensemble model, the response curve shows that highest probability of occurrence is observed within 2 km of rivers. This area is commonly known as the riparian zone. Riparian zones are plant communities on or near the banks of river channels. A variety of natural disturbances usually creates a unique mosaic of vegetation in riparian zones, which may be parallel to vegetation in the larger landscape (Didham et al., 2007). Adding to that, rivers play a second significant role of propagule transportation to terrestrial plants. The riparian zone mostly consists of palatable forbs (Roath and Krueger, 1982; Kauffman et al., 1983), thus there is the effective transport of *Xanthium strumarium* to and from riparian zones when they come to for food and water. Therefore, it is important to target these riparian areas when implementing management strategies of *Xanthium strumarium*, given the role riparian areas play in the species' establishment.

This study differs from other spatial modelling studies in the southern African savanna (Colgan et al., 2012; Baldeck et al., 2014; Mashintonio et al., 2014) in that it adopted a consensus approach to species modelling targeting an invasive species. In fact, previous studies in the biome have only been limited to single models (Stevens et al., 2014). Using the consensus modelling framework, robust predictions were made to aid pinpointing of potential hotspots *Xanthium strumarium* invasion in Gonarezhou national park and the adjacent communal landscapes.

Caution should be practiced when interpreting the results of this study since the prediction is only based on one site. More study sites could be necessary to further test their validity. Also, this study focused on only one invasive species which may have unique characteristics. Focus on more than one species could lead to more generalised conclusions. In this regard, it is recommended, that future studies focus on more than one study species, as well as on multiple study sites while using multiple distribution models.

Acknowledgements. This research was funded by the CIRAD Research Platform "Production and Conservation in Partnership" (RP-PCP) program, project Eco#6. The Parks and Wildlife management authority of Zimbabwe and the National Herbarium and Botanical gardens of Zimbabwe provided support in the form of manpower and resources. Of special mention are Mr. Mupunga and Mr. C. Chapano from the respective departments. It would be unfair not to extend gratitude to the academic staff of the Department of Geography and Environmental science for their input. To colleagues Ms. Chigwenhese, Mrs. Mudyawabikwa, Mr. Pachavo, Mr. Ndaimani, Mr. Gara, Mr. Zvidzai, Dr Dube thank you. Atheneum the tremendous support is appreciated.

REFERENCES

- [1] Abraham, A. (2005): Artificial neural networks. Handbook of measuring system design.– John Wiley and Sons, New York City.
- [2] Addo-Fordjour, P., Obeng, S., Addo, M. G., Akyeampong, S. (2009): Effects of human disturbances and plant invasion on liana community structure and relationship with trees in the Tinte Bepo forest reserve, Ghana.–Forest Ecology and Management 258:728-734.
- [3] Allouche, O., Tsoar, A., Kadmon, R. (2006): Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS).–Journal of Applied Ecology 43:1223-1232.
- [4] Araújo, M. B., New, M. (2007): Ensemble forecasting of species distributions.–Trends in Ecology & Evolution 22:42-47.
- [5] Austin, M. P. (2002): Spatial prediction of species distribution : an interface between ecological theory and statistical modelling.–Ecological Modelling 157(2):101-118.
- [6] Austin, M. P., Belbin, L., Meyers, J. a., Doherty, M. D., Luoto, M. (2006): Evaluation of statistical models used for predicting plant species distributions: Role of artificial data and theory.–Ecological Modelling 199:197-216.
- [7] Baldeck, C., Colgan, M., Féret, J.-B., Levick, S. R., Martin, R., Asner, G. (2014): Landscape-scale variation in plant community composition of an African savanna from airborne species mapping.–Ecological Applications 24:84-93.
- [8] Barrat-Segretain, M.-H. (2001): Invasive species in the Rhône River floodplain (France): replacement of *Elodea canadensis* Michaux by *E. nuttallii* St. John in two former river channels.–Archiv für Hydrobiologie 152:237-251.
- [9] Bartomeus, I., Sol, D., Pino, J., Vicente, P., Font, X. (2012): Deconstructing the native exotic richness relationship in plants.–Global Ecology and Biogeography 21:524-533.
- [10] Beaumont, L. J., Hughes, L., Poulsen, M. (2005): Predicting species distributions: use of climatic parameters in BIOCLIM and its impact on predictions of species' current and future distributions.–Ecological Modelling 186:251-270.
- [11] Capdevila-Argüelles, L., Zilleti, B. (2005): Issues in bioinvasion science.–Springer.
- [12] Capinha, C., Anastácio, P. (2011): Assessing the environmental requirements of invaders using ensembles of distribution models.–Diversity and Distributions 17:13-24.
- [13] Chigwenhese, L., Murwira, A., Zengeya, F. M., Masocha, M., Garine Wichtatitsky, M., Caron, A. (2016): Monitoring African buffalo (*Syncerus caffer*) and cattle (*Bos taurus*) movement across a damaged veterinary control fence at a Southern African wildlife/livestock interface.–African Journal of Ecology (online).
- [14] Chown, S. L., Hull, B., Gaston, K. J. (2005): Human impacts, energy availability and invasion across Southern Ocean Islands.–Global Ecology and Biogeography 14:521-528.
- [15] Clout, M. N., Williams, P. A. (2009): Invasive species management: a handbook of principles and techniques.–Oxford University Press.
- [16] Colgan, M. S., Baldeck, C. A., Féret, J.-B., Asner, G. P. (2012): Mapping savanna tree species at ecosystem scales using support vector machine classification and BRDF correction on airborne hyperspectral and LiDAR data.–Remote Sensing 4:3462-3480.
- [17] Currie, D. J., Paquin, V. (1987): Large-scale biogeographical patterns of species richness of trees.–Nature 329:326-327.
- [18] Cutler, D. R., Edwards, T. C., Beard, K.H., Cutler, A., Hess, A. T., Gibson, J., Lawler, J.J. (2007): Random forests for classification in ecology.–Ecology 88(11):2783-2792.
- [19] Dasonville, N., Vanderhoeven, S., Domken, S., Meerts, P., Chapuis-Lardy, L. (2008): Invasive Species: detection, Impact and Control.–Nova Science Publishers, New York City.
- [20] Davies, K. W., Johnson, D. D. (2012): Are We "Missing the Boat" on Preventing the Spread of Invasive Plants in Rangelands?–Invasive Plant Science and Management 4:166-171.

- [21] Didham, R. K., Tylianakis, J. M., Gemmill, N. J., Rand, T. A., Ewers, R. M. (2007): Interactive effects of habitat modification and species invasion on native species decline.– *Trends in Ecology & Evolution* 22:489-496.
- [22] Dukes, J. S. (2001): Biodiversity and invasibility in grassland microcosms.– *Oecologia* 126:563-568.
- [23] Ekeleme, F., Akobundu, I. O., Isichei, A. O., Chikoye, D. (2000): Influence of fallow type and land-use intensity on weed seed rain in a forest/savanna transition zone.– *Weed Science* 48:604-612.
- [24] Elith, J., Graham, C. H. (2009): Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models.– *Ecography* 32:66-77.
- [25] Elith, J., Leathwick, J. R. (2009): Species distribution models: ecological explanation and prediction across space and time.– *Annual Review of Ecology, Evolution, and Systematics* 40:677.
- [26] Eppstein, M. J., Molofsky, J. (2007): Invasiveness in plant communities with feedbacks.– *Ecology letters* 10:253-63.
- [27] Farji-Brener, A. G., Ghermandi, L. (2008): Leaf-cutting ant nests near roads increase fitness of exotic plant species in natural protected areas.– *Proceedings of the Royal Society B: Biological Sciences* 275:1431-40.
- [28] Foxcroft, L. C., Jarošík, V., Pyšek, P., Richardson, D. M., Rouget, M. (2011): Protected-Area Boundaries as Filters of Plant Invasions.– *Conservation* 25:400-405.
- [29] Límites de Áreas Protegidas como Filtros para la Invasión de Plantas.– *Conservation Biology* 25:400-405.
- [30] Gray, A. J., Mack, R. N., Harper, J. L., Usher, M. B., Joysey, K., Kornberg, H. (1986): Do Invading Species have Definable Genetic Characteristics? [and Discussion].– *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 314:655-674.
- [31] Guisan, A., Lehmann, A., Ferrier, S., Austin, M., Overton, J. M. C., Aspinall, R., Hastie, T. (2006): Making better biogeographical predictions of species' distributions.– *Journal of Applied Ecology* 43:386-392.
- [32] Guisan, A., Thuiller, W. (2005): Predicting species distribution: offering more than simple habitat models.– *Ecology Letters* 8:993-1009.
- [33] Hare, J. D. (1980): Variation in fruit size and susceptibility to seed predation among and within populations of the cocklebur, *Xanthium strumarium* L.– *Oecologia* 46:217-222.
- [34] Hobbs, R. J., Huenneke, L. F. (1992): Disturbance, Diversity, and Invasion: Implications for Conservation.– *Conservation Biology* 6:324-337.
- [35] Holland, P., Woodland, D. W., Hetherington, J. (2011): Impact of Plants Introduced by People During a Time of Rapid Environmental Change.– *Geography Compass* 4:545-560.
- [36] Holmes, P. M., Esler, K. J., Richardson, D. M., Witkowski, E. T. F. (2008): Guidelines for improved management of riparian zones invaded by alien plants in South Africa.– *South African Journal of Botany* 74:538-552.
- [37] Jarnevich, C. S., Stohlgren, T. J., Kumar, S., Morissette, J. T., Holcombe, T. R. (2015): Caveats for correlative species distribution modeling.– *Ecological Informatics* 29:6-15.
- [38] Kasabov, N. K. (1996): Foundations of neural networks, fuzzy systems, and knowledge engineering.– Marcel Alencar.
- [39] Kauffman, J. B., Krueger, W. C., Vavra, M. (1983): Effects of late season cattle grazing on riparian plant communities.– *Journal of Range Management* 36(6):685-691.
- [40] Keller, R. P., Lodge, D. M., Lewis, M. A., Shogren, J. F. (2009): Bioeconomics of invasive species: integrating ecology, economics, policy, and management.– Oxford University Press.
- [41] Kumschick, S., Bacher, S., Dawson, W., Heikkilä, J., Sendek, A., Pluess, T., Robinson, T., Kühn, I. (2012): A conceptual framework for prioritization of invasive alien species for management according to their impact.– *NeoBiota* 15:69-100.

- [42] Mack, R. N., Lonsdale, W. M. (2001): Humans as Global Plant Dispersers: Getting More Than We Bargained For: Current introductions of species for aesthetic purposes present the largest single challenge for predicting which plant immigrants will become future pests.–*BioScience* 51:95-102.
- [43] Marwat, K. B., Hashim, S., Ali, H. (2010): Weed management: a case study from North-West Pakistan.–*Pakistan Journal of Botany* 42:341-353.
- [44] Mashintonio, A. F., Pimm, S. L., Harris, G. M., Van Aarde, R. J., Russell, G. J. (2014): Data-driven discovery of the spatial scales of habitat choice by elephants.–*PeerJ* 2:e504.
- [45] Mestre, F., Pita, R., Paupério, J., Martins, F. M., Alves, P. C., Mira, A., Beja, P. (2015): Combining distribution modelling and non-invasive genetics to improve range shift forecasting.–*Ecological Modelling* 297:171-179.
- [46] Mombeshora, S., Bel, S. (2009): Parks-people conflicts: the case of Gonarezhou National Park and the Chitsa community in south-east Zimbabwe.–*Biodiversity and Conservation* 18:2601-2623.
- [47] Mortensen, D. A., Johnson, G. A., Young, L. J., Robert, P., Rust, R., Larson, W. (1993): Weed distribution in agricultural fields.–In: Proceedings of soil specific crop management: a workshop on research and development issues, April 14-16, 1992, Minneapolis, MN, USA. American Society of Agronomy (ASA), pp. 113-124.
- [48] Mott, K. A., Jensen, R. G., O'Leary, J. W., Berry, J. A. (1984): Photosynthesis and ribulose 1, 5-bisphosphate concentrations in intact leaves of *Xanthium strumarium* L.–*Plant Physiology* 76:968-971.
- [49] Nel, J. L., Richardson, D. M., Rouget, M., Mgidi, T. N., Mdzeke, N., Maitre, D. C. L., Wilgen, B. W. V., Schonegevel, L., Henderson, L., Neser, S. (2004): A proposed classification of invasive alien plant species in South Africa : towards prioritizing species and areas for management action.–*South African Journal of Science* 100:53-64.
- [50] Ochoa-Ochoa, L. M., Flores-Villela, O. A., Bezaury-Creel, J. E. (2016): Using one vs. many, sensitivity and uncertainty analyses of species distribution models with focus on conservation area networks.–*Ecological Modelling* 320:372-382.
- [51] Brown, R.L., Peet, R. K. (2003): Diversity and invasibility of southern Appalachian plant communities.–*Ecology* 84:32-39.
- [52] Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Henry, M., Stevens, H., Wagner, H. (2016). *Vegan: Community Ecology Package*. <https://cran.r-project.org/web/packages/vegan/vegan.pdf>
- [53] Phillips, S. J., Anderson, R. P., Schapire, R. E. (2006): Maximum entropy modeling of species geographic distributions.–*Ecological Modelling* 190:231-259.
- [54] Pyšek, P., Richardson, D. M., Pergl, J., Jarosik, V., Sixtova, Z., Weber, E. (2008): Geographic and taxonomic biases in invasion ecology and evolution –*Trends in Ecology and Evolution* 23:237-244.
- [55] Regnier, E. E., Harrison, S. K. (1993): Compensatory responses of common cocklebur (*Xanthium strumarium*) and velvetleaf (*Abutilon theophrasti*) to partial shading.–*Weed Science* 41(4): 541-547.
- [56] Rejmanek, M., Richardson, D. M., Higgins, S. I., Pitcairn, M. J., Grotkopp, E. (2005): Ecology of invasive plants: state of the art.– In: Mooney, H.A., Mack, R.N., Mc Neely, J.A., Neville, L., Schei, P.J., Waage, J. (eds) *Invasive alien species: A new synthesis* , pp. 104-161. Island Press, Washington D.C.
- [57] Rejmanek, M., Richardson, D. M., Pyšek, P. (2013): Plant Invasions and Invasibility of Plant Communities. –In: van der Maarel, E., Franklin, J. (eds) *Vegetation Ecology*, 2nd Edition, John Wiley & Sons, Ltd, Oxford, UK. doi: 10.1002/9781118452592.ch13
- [58] Rejmanek, M., Sandlund, O. (1999): Invasive plant species and invulnerable ecosystems.–In: Sundland, O. T., Schei, P. J., Viken, A. (ed) *Invasive species and biodiversity management*, Kluwer academic publishers, Netherlands.
- [59] Richardson, D. M., Pyšek, P. (2006): Plant invasions: merging the concepts of species invasiveness and community invasibility.–*Progress in Physical Geography* 30:409-431.

- [60] Roath, L. R., Krueger, W. C. (1982): Cattle grazing influence on a mountain riparian zone.–*Journal of Range Management* 35(1): 100-103.
- [61] Sankraan, M., Hanan, N. P., Scholes, R. J. (2005): Determinants of woody cover in African savannas.–*Nature* 438:846-849.
- [62] Sharkey, T. D., Raschke, K. (1981): Effect of light quality on stomatal opening in leaves of *Xanthium strumarium* L.–*Plant Physiology* 68:1170-1174.
- [63] Sithole, D., Zisadza-Gandiwa, P., Gandiwa, E. (2012): An assessment of alien plant species in Gonarezhou National Park, Zimbabwe.–*Greener Journal of Agricultural Sciences* 2:334-339.
- [64] Stachowicz, J. J., Whitlatch, R. B., Osman, R. W. (1999): Species diversity and invasion resistance in a marine ecosystem.–*Science* 286:1577-1579.
- [65] Stevens, N., Swemmer, A. M., Ezzy, L., Erasmus, B. F. N. (2014): Investigating potential determinants of the distribution limits of a savanna woody plant: *Colophospermum mopane*.–*Journal of Vegetation Science* 25:363-373.
- [66] Thuiller, W., Georges, D., Engler, R. (2014). BIOMOD- a platform for ensemble forecasting of species distributions. –*Ecography* 32:369-373.
- [67] Thêbaud, C., Simberloff, D. (2001): Are Plants Really Larger in Their Introduced Ranges?– *The American Naturalist* 157:231-236.
- [68] Tucker, D. J., Mansfield, T. (1971): Effects of light quality on apical dominance in *Xanthium strumarium* and the associated changes in endogenous levels of abscisic acid and cytokinins.–*Planta* 102:140-151.
- [69] Václavík, T., Meentemeyer, R. K. (2009): Invasive species distribution modeling (iSDM): Are absence data and dispersal constraints needed to predict actual distributions?– *Ecological Modelling* 220:3248-3258.
- [70] Venodha, P. M. (2016): Landscape Degradation and Restoration–A Planning Approach.– *International Journal of Environmental Science and Development* 7(3):229-233.
- [71] Vitousek, P. M., Mooney, H. A., Lubchenco, J., Melillo, J. M. (1997): Human domination of Earth's ecosystems.–*Science* 277:494-499.
- [72] Wang, X., Lewis, J. D., Tissue, D. T., Seemann, J. R., Griffin, K. L. (2001): Effects of elevated atmospheric CO₂ concentration on leaf dark respiration of *Xanthium strumarium* in light and in darkness.–*Proceedings of the National Academy of Sciences* 98:2479-2484.
- [73] Wichmann, M. C., Alexander, M. J., Soons, M. B., Galsworthy, S., Dunne, L., Gould, R., Fairfax, C., Niggemann, M., Hails, R. S., Bullock, J. M. (2009): Human-mediated dispersal of seeds over long distances.– *Proceedings of the Royal Society B: Biological Sciences* 276:523-32.
- [74] Zengeya, F., Murwira, A., de Garine-Wichatitsky, M. (2011): An IKONOS-based comparison of methods to estimate cattle home ranges in a semi-arid landscape of southern Africa.–*International Journal of Remote Sensing* 32:7805-7826.