


Environmental and anthropogenic determinants of the spread of alien plant species: insights from South Africa's quaternary catchments

Dilva Terzano · Ian Kotzé · Christo Marais · Silvio Cianciullo · Alessio Farcomeni · Paolo Caroli · Luca Malatesta · Fabio Attorre 

Received: 30 March 2017 / Accepted: 8 January 2018
© Springer Science+Business Media B.V., part of Springer Nature 2018

Abstract Alien plants invasion has negative impacts on the structure and functionality of ecosystems. Understanding the determinants of this process is fundamental for addressing environmental issues, such as the water availability in South Africa's catchments. Both environmental and anthropogenic factors determine the invasion of alien species; however, their relative importance has to be quantified. The aim of this paper was to estimate the importance of 32 explanatory variables in predicting the distribution of the major invasive alien plant species (IAPS) of South Africa, through the use of Species Distribution Models. We used data from the National Invasive Alien Plants Survey, delineated at a

quaternary catchment level, coupled with climatic, land cover, edaphic, and anthropogenic variables. Using two-part generalized linear models, we compared the accuracy of two different sets of variables in predicting the spatial distribution of IAPS; the first included environmental correlates alone, and the second included both environmental and anthropogenic variables. Using Random Forest, we explored the relative importance of the variables in producing a map of potential distribution of IAPS. Results showed that the inclusion of anthropogenic variables did not significantly improve model predictions. The most important variables influencing the distribution of IAPS appeared to be the climatic ones. The modeled potential distribution was analyzed in relation to provinces, biomes, and species' minimum residence time.

Communicated by Peter Minchin.

D. Terzano (✉) · S. Cianciullo · P. Caroli · L. Malatesta · F. Attorre
Department of Environmental Biology, Sapienza University of Rome, Rome, Italy
e-mail: dilva.terzano@gmail.com

I. Kotzé
ARC-Institute for Soil, Climate and Water, Stellenbosch, South Africa

C. Marais
Department of Environmental Affairs, Natural Resource Management Programmes, Cape Town, South Africa

A. Farcomeni
Department of Public Health and Infectious Diseases, Sapienza University of Rome, Rome, Italy

Keywords Alien Species · Generalized Linear Models · Random Forest · Species Distribution Models

Introduction

Invasive alien plant species are recognized as a major component of global change (Higgins and Richardson 1996; Higgins et al. 2000) with effects on biodiversity, ecosystem services, disturbance regimes,

biogeochemical cycles, and human societies (Van Wilgen et al. 2001; Pauchard and Alaback 2004; Sharma et al. 2005; Thuiller et al. 2007; Gassó et al. 2009; Rödder et al. 2009; Pejchar and Mooney 2009).

Species Distribution Models (SDMs) have been used for identifying suitable areas for species to invade and thus prioritizing IAPS surveys and management actions. One of the factors influencing model robustness is the selection of meaningful predictor variables in relation to the spatial scale and geographical extent (Austin 2002; Araújo and Guisan 2006; Elith and Leathwick 2009; Franklin 2010). Robust SDMs are useful to improve our understanding of the importance of environmental and human-related variables in influencing the ranges of species (Luoto et al. 2006; Ohlemuller et al. 2006).

The selection of variables should be based on the ecological and physiological characteristics of the IAPS, data resolution, quality, and availability (Thuiller et al. 2003; Dormann 2007; Rödder et al. 2009; Austin and Van Niel 2011). Many studies have used SDMs to predict the invasion suitability of areas based on climatic variables (Peterson and Vieglais 2001; Rödder et al. 2009), which at large biogeographical scales, set the physiological niche of plant species for survival and reproduction (Rouget and Richardson 2003; Thuiller et al. 2003). However, opinions differ on whether climatic variables are sufficient for explaining species distributions (Pearson and Dawson 2003; Thuiller et al. 2004; Pearson et al. 2004). There is a need for more evidence in support of the idea that purely climate-based modeling proves sufficient to predict the distribution of IAPS (Araújo and Guisan 2006; Araújo and Luoto 2007; Austin and Van Niel 2011), compared to using other environmental and human-related variables (e.g., Pearson et al. 2004; Thuiller et al. 2004; Pauchard and Alaback 2004; del Barrio et al. 2006; Coudun et al. 2006; Luoto et al. 2006; Ficetola et al. 2007).

Other factors such as land cover, soils, fire frequency, and anthropogenic variables determine the presence or absence of a species in a particular area (Rouget and Richardson 2003; Pearson and Dawson 2003; Staver et al. 2011). Unfortunately, these variables are rarely taken into account due to the scarcity of adequate data sets (Thuiller et al. 2003, 2005). SDMs that incorporate these factors could have an increased predictive power, and unveil specific factors responsible for detailed distributional

phenomena (Roura-Pascual et al. 2004). Only a few studies have been conducted comparing the importance of variables as well as the variation of this importance at different scales (Ohlemuller et al. 2006; Luoto et al. 2006; Funk et al. 2016). Consequently, recent studies have started to incorporate human activities and disturbances variables (e.g., Pino et al. 2005; Thuiller et al. 2006; Chytrý et al. 2008; Gassó et al. 2009). Roura-Pascual et al. (2004) stated that SDMs that do not take into account anthropogenic disturbance could underestimate the predicted species distribution. Hence, the need to assess the importance of human-related predictor variables for a successful modeling of IAPS (Dormann 2007; Broennimann et al. 2007).

South Africa represents an interesting case study since it has been invaded by numerous species with significant ecological and economic implications (Richardson et al. 2005; Sharma et al. 2005). The number of IAPS and their impacts have increased since the first concerns regarding plant invasion emerged in the 1770s with the introduction of *Opuntia ficus-indica*, and *Prosopis glandulosa*, *P. juliflora* and *P. velutina* in the 1800s (Milton and Dean 2010). IAPS, especially large trees, have a greater water usage compared to native species (Calder and Dye 2001; Mallory et al. 2011) resulting in the annual runoff decline registered in the past decades, adversely affecting water supplies (Le Maitre et al. 1996, 2000). According to Cullis et al. (2007), the impact of IAPS in South Africa is approximately 4% of current registered water use, and could increase to 16% in the future. These adverse impacts of IAPS on water flows have been the prime motivation for the establishment, in 1995, of South Africa's national Working for Water programme (WfW) (Le Maitre et al. 2016). The rationale behind it being that the most cost-effective way to increase water supplies is to remove IAPS from catchments—more water could be available at a lower cost where IAPS control operations are in place compared to developing additional water supply schemes (Le Maitre et al. 1998). Although this public program is one of the largest ecosystem restoration programs in the world (Hobbs 2004; Richardson et al. 2005), its effectiveness in minimizing the spread has recently been proven to be questionable (McConachie et al. 2012; van Wilgen et al. 2016). The use of robust and reliable SDMs for IAPS can be a useful tool for stakeholders to prioritize interventions by control

programs and monitor their progress, in order to reach successful and efficient implementation.

This study tested the appropriateness of including human-related variables in SDMs for predicting the distribution of IAPS, and assessed the variables' importance on the presence and abundance of these species. Using a newly available dataset on IAPS distribution and density, the National Invasive Alien Plant Survey (NIAPS), conducted at quaternary catchment (QC) level, we addressed the following questions: (1) what are the main factors influencing the spatial distribution of the main IAPS of South Africa? (2) are there differences in the influence that these variables have on the potential presence and abundance of these IAPS? (3) does the inclusion of anthropogenic variables in the SDMs result in a better prediction and successful modeling of these invasive plants? Considering that IAPS present a growing challenge for environmental managers and policy-makers (Higgins and Richardson 1996; Rouget et al. 2002), we analyzed the predicted potential spread of the main South African IAPS, in relation to provinces, biomes, and species' residence time with the intent of providing to decision makers a clearer understanding of the impact of each factor.

Methods

Study area

South Africa is the southernmost country in Africa and due to a varied topography and oceanic influence, a great variety of climatic zones exist, ranging from the extreme desert of the southern Namib in the northwest to the lush subtropical climate along eastern coastal areas. As a result of the wide range of climatic and geomorphological differences, South Africa hosts a variety of different distinct communities of plants and animals (Rutherford and Westfall 1986) that have common characteristics for the environment they exist in formed in response to a shared physical climate (Biomes). The biomes identified in South Africa are eight. They include the grassy dwarf shrublands of the Nama and Succulent Karoo, and the annual plant formations of the environmentally harsh Desert in the western part of the Country (Mucina and Rutherford 2006); the Mediterranean like shrubland and heathland vegetation of the biodiversity hotspot Cape

Fynbos; the Grasslands of the central plateaux where the combination of frost, fire, and grazing prevent the establishment of trees; the eastern subtropical Albany Thicket, which is a closed shrubland/low "forest" dominated by evergreen, sclerophyllous, or succulent trees, shrubs, and vines; the eastern Indian Ocean Coastal Belt dominated by evergreen temperate multilayer forests; the northern Savanna, which is characterized by the dynamic coexistence of grasses (mainly C4 type) and woody vegetation.

Response variable

Plant surveys provide essential information that can determine the geographic extent of an IAPS, define its ecological requirements, indicate the potential for further spread, and provide a historical account of the IAPS's introduction and expansion (Henderson 1999). The NIAPS was implemented to establish a cost-effective and statistically sound IAPS monitoring system for South Africa at a quaternary catchment level. It used a stratified systematic sampling approach, which consists of 72,682 sample points distributed proportionally among environmental strata representing the steepest gradient contributing the most to the IAPS occurrence within the country. An aerial field survey was conducted and at each 100 m² survey plot the following attributes were captured; (1) overall density of IAPS; (2) the three dominant IAPS; (3) density per dominant IAPS, and (4) size class per dominant IAPS according to the WfW mapping standards (Table 1)(Working for Water Mapping Standards 2003; Kotzé et al. 2010).

Data were grouped into quaternary catchments (QC), which is the basic unit for water resource management in South Africa, and the operational delineation for WfW clearing projects. Total species included in the survey amounted to 215 and was obtained from three referenced species classifications (Nel et al. 2004; Robertson et al. 2003; Marais et al. 2004). A final list of 27 major plant invaders was produced, and species were mapped and included in the NIAPS database (Table 2). They were recorded as species (e.g., *Arundo donax*) or genera when referring to multiple species not easily differentiable but with similar ecological requirements (e.g., *Opuntia* spp. includes *Opuntis ficus-indica* and *Opuntia stricta*). The majority of species are alien phanerophytes likely reflecting the high introduction pressure of trees since

Table 1 List of explanatory variables divided into 25 environmental (E) and 5 anthropogenic variables (A)

Explanatory variables (unit)	Acronyms	Variable type
Annual precipitation (mm)	Ptot	E
Winter precipitation (mm)	Pw	E
Summer precipitation (mm)	Psm	E
Spring precipitation (mm)	Pa	E
Autumn precipitation (mm)	Psp	E
Mean annual temperature (°C)	Tm	E
Maximum temperature of January (hottest month) (°C)	Tmx	E
Minimum temperature of July (coldest month) (°C)	Tmn	E
Duration of frost period (days)	Frost	E
Growing season duration (days)	Msg	E
Total profile plant available water (mm)	Paw	E
Slope (%)	Slope	E
Soil pH	pH	E
Nitrogen content	N	E
Podzolic soils	Soil1	E
Red-yellow well-drained, massive or weakly structured soils	Soil2	E
Rocky areas	Soil3	E
Sandy soils with little or no profile development	Soil4	E
Soils with a plinthic horizon	Soil5	E
Soils with a strong texture contrast	Soil6	E
Soils with limited pedological development	Soil7	E
Strongly saline soils	Soil8	E
Well-structured soils generally with a high clay content	Soil9	E
Natural areas	Natur	E
Waterbodies	Water	E
Rivers	River	E
Cultivations and plantations	Cult	A
Degraded areas	Degr	A
Urban areas	Urb	A
Railways	Rail	A
Roads	Roads	A
Fire frequency	Fire	A

and the facilitation of alien tree spread through deliberate and massive planting (Richardson 1998).

As response variable we used the condensed area value for each taxon, which expresses the equivalent of the total invaded area with the canopy cover rescaled to 100% (Versfeld et al. 1998). For example, 100 ha that was covered by 10% with IAPS was expressed as a condensed area of 10 ha with 100% cover. The condensed area was calculated by multiplying the average IAPS density (%) by the invaded area (ha) and dividing by 100. We then normalized the

condensed area of each IAPS within the QCs (Kotzé et al. 2010).

Explanatory variables

The selection of variables was based on the understanding of the ecological and biophysical processes influencing the dispersal of the species, the availability of data and the purpose of the model (Thuiller et al. 2003, 2005). The choice of the climatic and land use variables follows Rouget et al. (2002, 2004), and Thuiller et al. (2007). The use of other environmental

Table 2 List of NIAPS species/genus referred to as IAPS from here on (Kotzé et al. 2010)

IAPS	Family
<i>Acacia Cyclops</i>	Fabaceae
<i>Acacia melanoxyton</i>	Fabaceae
<i>Acacia saligna</i>	Fabaceae
<i>Acacia</i> spp.	Fabaceae
<i>Agave</i> spp.	Agavaceae
<i>Arundo donax</i>	Poaceae
<i>Atriplex nummularia</i>	Chenopodiaceae
<i>Caesalpinia decapetala</i>	Fabaceae
<i>Cereus jamacaru</i>	Cactaceae
<i>Cestrum</i> spp.	Solanaceae
<i>Chromolaena odorata</i>	Asteraceae
<i>Eucalyptus</i> spp.	Myrtaceae
<i>Hakea</i> spp.	Proteaceae
<i>Jacaranda mimosifolia</i>	Bignoniaceae
<i>Lantana camara</i>	Verbenaceae
<i>Melia azedarach</i>	Meliaceae
<i>Opuntia</i> spp.	Cactaceae
<i>Pinus</i> spp.	Pinaceae
<i>Populus</i> spp.	Salicaceae
<i>Prosopis</i> spp.	Fabaceae
<i>Psidium gaujva</i>	Myrtaceae
<i>Rosa rubiginosa</i>	Rosaceae
<i>Salix babylonica</i>	Salicaceae
<i>Senna didymobotrya</i>	Fabaceae
<i>Sesbania punicea</i>	Fabaceae
<i>Solanum mauritianum</i>	Solanaceae
<i>Tamarix chinensis</i>	Tamaricaceae

variables, such as soil water, as well as some human-related variables, such as roads and urban areas was based on Richardson et al. (2005).

- We estimated the values of environmental (climate, soil, land use, etc.) and human-related variables (roads, railways, urban areas) from two databases with spatially explicit GIS data layers (Table 1). We identified nine climatic variables (Schulze 2008), slope, and the ten soil variables from the ARC's Land Type Survey database (Land Type Survey Staff 1972–2006 (2006)), while land use and other human-related variables were obtained from the National Land Cover database (CSIR and ARC 2005). Considering also the

importance of fire history and soil nitrification on the spread of alien plant species in limiting environments as South Africa (e.g., Funk et al. 2016) we have included fire frequency fire frequency (F) from a fire count map for the period 2000–2013 and based on MODIS MCD64 burnt area product (Giglio et al. 2009), as well as the nitrogen content (N) of soils at a depth of 20 cm obtained from the AfSIS Africa Soil Map at spatial resolution of 250 m (Hengl et al. 2015).

Land cover, soil type, and human-related variables were standardized for the area of each QC. The rest of the variables were calculated by averaging the data within each QC (e.g., climatic variables, soil pH, etc.).

Model

Zero-inflated generalized linear models (GLMs) were used to examine the variation in the occurrence and abundance of the IAPS for each QC across the study area (Segurado and Araújo 2004). We fitted two-part models to analyze the relationship and identify factors explaining occurrence and abundance. The first part of the model is a logistic regression model to analyze the influence of the explanatory variables on the presence of IAPS. The second part is a multivariate linear regression model for the log-abundance, which highlights the influence of the explanatory variables on the abundance of IAPS (Di Lorenzo et al. 2011). Since the NIAPS database had both presence and absence data, and presence-only models tend to overpredict the actual invasion distribution, we applied a model that uses presence/absence data (Senan et al. 2012). The log-likelihood of the two-part model was expressed as the sum of the log-likelihood of each part, and maximum likelihood estimation was then performed simultaneously on the two parts to avoid bias in parameter and standard error estimates.

Model selection was performed using a forward stepwise algorithm based on minimization of the Akaike Information Criterion (AIC), which gives guarantees against multicollinearity issues and leads to parsimonious models with good predictive properties. Collinearity was checked for the final multivariate model through variance inflation factors.

The two-part model was run twice: first, using only the environmental variables, and secondly using both environmental and anthropogenic variables. To

understand which set of variables is more appropriate we compared the out-of-bag prediction error values of random forest models. The prediction error is the squared difference between observed and predicted values, and this can be estimated repeatedly on subsets of observations that are not used for estimation (out-of-bag observations). The model with the lowest error-rate was considered the best.

We also analyzed if the difference between the error-rates for each species were statistically significant, by applying the Wald test (Agresti 2002) after estimation of the standard errors via resampling. The null hypothesis for the selection of the explanatory variables was that the difference between error-rates of the two models was 0, in other words, that the use of either model does not result in significant differences in the response predictions. If p value < 0.05 the difference between the models is significant, hence the inclusion of anthropogenic variables results in a more adequate prediction. Similarly, the null hypothesis for the selection of the species was that all the coefficients of the regression model were 0 for each species.

We then used a Random Forest (RF) model to assess the importance of the explanatory variables in influencing the distribution and abundance of IAPS (Breiman 2001). In an RF model, bootstrap samples are used to construct multiple trees with a subset of randomized variables. The measure of importance was derived from the contribution of each explanatory variable accumulated along all nodes and trees (Senan et al. 2012). We analyzed the model accuracy by computing the Gini Impurity Measure. We then compared the current IAPS distribution with the predictions from the RF model, taking into account the relationship with the variables. In order to have maps for this comparison, we converted the probability data into binary presence/absence data, using a probability threshold value above which the IAPS were deemed to be present. We used the method of the least predicted value: the lowest value of the presence probability associated with the areas where the species was actually observed, as the minimum value beyond which a species was considered present. All the statistical analyses were performed in R (R Core Team 2014) Fig. 1. We finally analyzed the potential distribution of the IAPS in relation to the South African provinces and biomes, and we also assessed the relative occupancy as a function of the species' residence time.

Results

Assessment of variables

We assessed whether the use of anthropogenic variables could improve the accuracy of the predictive model. Table 3 shows the values for each species of the error-rate related to the model run using only environmental variables (OOB error-rate ENV) and to the model run including the human-related variables (OOB error-rate TOT). For each taxon, the best model is expressed with the lowest error-rate value. The “ Δ error-rate (ENV-TOT)” shows the differences between the error-rate of the two models: if Δ error-rate is negative, the best model is the one with the environmental set of variables alone, if Δ error-rate is positive, the model with both environmental and anthropogenic variables is the one with a better prediction.

For 13 IAPS, the model built with environmental and anthropogenic variables (TOT) seemed to be more suitable for the prediction. For 10 IAPS, the model built only with environmental variables (ENV) was more suitable. For the remaining 4 IAPS, no difference was observed. We then analyzed the statistical significance of these differences, through the use of the Wald test. This test indicated that for nearly all the IAPS, excluding *Caesalpinia decapetala* and *Senna didymobotrya*, there are no significant differences in the predictions given by the two models (p value > 0.05).

For the rest of the analysis, we decided to use the ENV model, by excluding the human-related variables, which were not found to be instrumental in improving the accuracy of the model.

Potential spread

The average decrease in the Gini impurity measure showed the species for which the RF model provided a statistically significant prediction. Out of the 27 IAPS considered, only 14 had significant results (p value > 0.05), the remaining IAPS were found to be those with the least number of QC currently invaded (generally less than 50—Table 4). The latter IAPS were consequently excluded from further analysis. Analyzed species showed a differential behavior with species characterized by a small current and predicted invaded areas, e.g., *Ceresus jamacaru* and *Acacia*

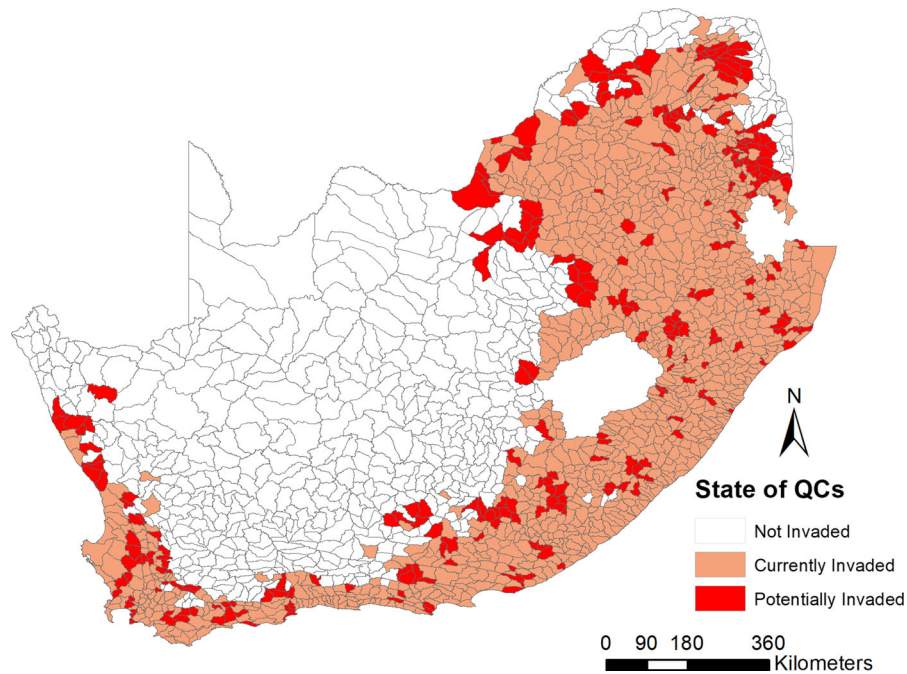


Fig. 1 Map of South Africa showing the currently invaded QCs and potentially new QCs suitable for invasion. In white are the QCs for which the two-part model showed invasion potential values that were lower than the set presence threshold. (Color figure online)

cyclops, and species with a large invaded areas, such as *Acacia* spp. (*A. baileyana*, *A. dealbata* and *A. mearnsii*) and *Eucalyptus* spp. (*E. camaldulensis* and *E. conferruminata*). Some species have a low relative occupancy and a great potential invasion increase, such as *Lantana camara* and *Melia azederach*, while others have already occupied the potential spread (Table 5). We assessed if this was related to the difference in residence time. However, the association between relative occupancy and minimum residence time was weak and no association was identified.

The biomes differed markedly in their invasibility. According to the predictions obtained by the model no species seemed to be able to invade the Desert, only two IAPS (*Agave* spp., *Opuntia* spp.) could potentially invade the Nama-Karoo and four IAPS (*Acacia cyclops*, *Eucalyptus* spp., *Pinus* spp. and *Acacia* spp.) could potentially invade the Succulent Karoo. While the other biomes are more prone to invasion, with the Savanna and Grasslands suitable for almost all the IAPS considered (Table 6).

The areas most affected by the potential spread of the IAPS were found to be those adjacent to the major urban areas, coastal settlements, and the northwestern plains. With the exception of the Northern Cape

Province, all South African provinces were found to have a percentage of potentially invaded QCs higher than 50%. The eastern coastal and central-eastern provinces (KwaZulu-Natal, Mpumalanga, and Gauteng) were those that had the highest percentage of potentially invaded area. Gauteng and KwaZulu-Natal, according to the predictive model, would be completely suitable for the spread of the IAPS (Table 7, Fig. 2).

Determinants of invasion

The two-part model highlighted both positive and negative correlations between the explanatory variables and the species distribution. First, the influence of the explanatory variables on the presence of the IAPS was highlighted with the logistic part (Table 8 in Appendix), and second, the influence of the explanatory variables on the richness of the IAPS was highlighted with the regression part (Table 9 in Appendix).

The presence of IAPS was found to be positively correlated with the mean precipitation of autumn season, mean annual temperature, mean growing season duration, density of natural areas, mean soil

Table 3 Out-of-bag (OOB) error-rate for each species for the two RF models built with only environmental variables (ENV) and with the inclusion of anthropogenic variables (TOT)

IAPS	OOB error-rate ENV	OOB error-rate TOT	Δ error-rate (ENV-TOT)	<i>p</i> value (Wald test)	Sign.
<i>Acacia cyclops</i>	0.071	0.071	0	1	
<i>Acacia melanoxydon</i>	0.084	0.084	0	1	
<i>Acacia saligna</i>	0.084	0.084	0	1	
<i>Acacia</i> spp.	0.163	0.163	0	1	
<i>Agave</i> spp.	0.294	0.296	− 0.002	0.841	
<i>Arundo donax</i>	0.236	0.294	− 0.058	0.724	
<i>Atriplex nummularia</i>	0.311	0.311	0	1	
<i>Caesalpinia decapetala</i>	0.166	0.166	0	1	
<i>Cereus jamacaru</i>	0.235	0.133	0.102	0.035	*
<i>Cestrum</i> spp.	0.153	0.153	0	1	
<i>Chromolaena odorata</i>	0.102	0.102	0	1	
<i>Eucalyptus</i> spp.	0.297	0.282	0.015	0.898	
<i>Hakea</i> spp.	0.076	0.076	0	1	
<i>Jacaranda mimosifolia</i>	0.214	0.285	− 0.071	0.699	
<i>Lantana camara</i>	0.198	0.241	− 0.043	0.580	
<i>Melia azedarach</i>	0.241	0.222	0.019	0.567	
<i>Opuntia</i> spp.	0.190	0.157	0.033	0.198	
<i>Pinus</i> spp.	0.240	0.230	0.010	0.897	
<i>Populus</i> spp.	0.209	0.215	− 0.006	0.878	
<i>Prosopis</i> spp.	0.115	0.096	0.019	0.688	
<i>Psidium gaujava</i>	0.152	0.152	0	1	
<i>Rosa rubignosa</i>	0.001	0.001	0	1	
<i>Salix babylonica</i>	0.203	0.190	0.013	0.512	
<i>Senna didymobotrya</i>	0.276	0.166	0.110	0.041	*
<i>Sesbania punicea</i>	0.250	0.250	0	1	
<i>Solanum mauritianum</i>	0.093	0.091	0.002	1	
<i>Tamarix chinensis</i>	0.100	0.298	− 0.198	0.272	

Significance of the difference between the two values (Δ error-rate) was tested with the Wald test (*p* value). Significance: $p < 0.01$ (**); $p < 0.05$ (*)

water content, and most of the soil variables. IAPS presence was found to be negatively correlated with the mean annual precipitation, mean precipitation of winter and summer season, mean maximum temperature of hottest month; mean minimum temperature of coldest month, and mean duration of frost period.

IAPS abundance was positively correlated with the mean precipitation of spring and autumn season, mean moisture growing season duration, density of rivers, and two soil variables (index of rocky areas and soils with a plinthic horizon). IAPS abundance was found to be negatively correlated with the mean precipitation of

summer season, mean annual temperature, mean duration of frost period, density of water bodies, and two soil variables (red-yellow well-drained, massive or weakly structured soils and well-structured soils, generally with a high clay content).

Importance of variables

The RF model provided a measure of the importance of the explanatory variables influencing the distribution of the IAPS. These importance values were species-specific, but we identified similarities when

Table 4 The number of QCs currently invaded is reported for each species

IAPS	Currently invaded QCs	<i>p</i> value (Gini index)	Sign.
<i>Acacia cyclops</i>	55	0.120	
<i>Acacia melanoxylon</i>	12	0.001	***
<i>Acacia saligna</i>	48	0.001	***
<i>Acacia</i> spp.	515	0.931	
<i>Agave</i> spp.	61	0.941	
<i>Arundo donax</i>	17	0.001	***
<i>Atriplex nummularia</i>	16	0.001	***
<i>Caesalpinia decapetala</i>	24	0.001	***
<i>Cereus jamacaru</i>	38	0.619	
<i>Cestrum</i> spp.	13	0.001	***
<i>Chromolaena odorata</i>	113	0.999	
<i>Eucalyptus</i> spp.	409	0.991	
<i>Hakea</i> spp.	39	0.001	**
<i>Jacaranda mimosifolia</i>	14	0.001	***
<i>Lantana camara</i>	45	0.992	
<i>Melia azedarach</i>	109	0.999	
<i>Opuntia</i> spp.	173	0.912	
<i>Pinus</i> spp.	307	0.933	
<i>Populus</i> spp.	154	0.967	
<i>Prosopis</i> spp.	18	0.001	***
<i>Psidium gaujava</i>	26	0.000	***
<i>Rosa rubignosa</i>	7	0.000	***
<i>Salix babylonica</i>	139	0.987	
<i>Senna didymobotrya</i>	51	0.456	
<i>Sesbania punicea</i>	4	0.001	***
<i>Solanum mauritianum</i>	97	0.998	
<i>Tamarix chinensis</i>	10	0.001	***

The Gini impurity measure highlighted the species for which the predictor variables significantly explained their distribution. A lower value of the index represents the susceptibility of the model to the exclusion of a variable. Significance: $p < 0.01$ (**); $p < 0.001$ (***)

comparing the 14 IAPS. Climatic variables have an overwhelming importance on the distribution of IAPS, particularly both annual and summer precipitation. While land cover, pedological, and anthropogenic variables are similarly of minor importance, the only exception being pH for *Lantana camara*, *Agave* spp. and *Opuntia* spp. and density of waterbodies for *Acacia cyclops* (Table 10 in appendix).

Discussion

Modeling approach

We have modeled the distribution of the major South African IAPS, on the assumption that these species are at equilibrium with their environment (Rouget et al.

2004; Guisan and Thuiller 2005). Although most of our IAPS have a long history in the South African region (Table 5), it is likely that they have not yet reached equilibrium due to dispersal limitations (Rouget et al., 2004; Pearson, 2010).

Five major factors could affect the accuracy and the predictive capacity of our models. First, SDMs are based on the assumption that the current distribution of IAPS and the environmental characteristics of their current range provide a good indication of their potential range. However, the predicted potential range could have been overestimated for IAPS occurring in few scattered locations, or underestimated for IAPS currently occurring in a small range (Rouget et al. 2004). Second, spatial bias in the NIAPS database, related to the two overlapping layers (landscape and riparian), may have led in the riparian areas

Table 5 For each IAPS, we reported the number of QCs currently invaded and potentially invaded (in numbers and %, calculated as a proportion of the total 1840 QCs), the relative occupancy (RO), being the proportion of potential distribution range currently occupied by each species, the percentage of invasion increase, and the residence time (RT)

IAPS	Currently invaded QCs	Potentially invaded QCs	Currently invaded QCs (%)	Potentially invaded QCs (%)	Relative occupancy (RO) (%)	Invasion increase (%)	Residence time (RT) (years)
<i>Acacia cyclops</i>	55	143	3.0	7.8	38.5	159.9	181
<i>Acacia</i> spp.	515	617	28.0	33.5	83.5	19.8	97
<i>A. baileyana</i>							158
<i>A. dealbata</i>							158
<i>A. mearnsii</i>							158
<i>Agave</i> spp.	61	182	3.3	9.9	33.6	197.9	158
<i>A. americana</i>							158
<i>Cereus jamaicaru</i>	38	108	2.1	5.9	35.3	183.6	91
<i>Chromolaena odorata</i>	113	228	6.1	12.4	49.6	101.8	158
<i>Eucalyptus</i> spp.	409	552	22.2	30.0	74.1	35.0	120
<i>E. camaldulensis</i>							116
<i>E. conferruminata</i>							158
<i>Lantana camara</i>	45	240	2.4	13.0	18.8	432.2	216
<i>Melia azedarach</i>	109	359	5.9	19.5	30.3	229.6	360
<i>Opuntia</i> spp.	173	312	9.4	17.0	55.4	80.4	79
<i>O. ficus-indica</i>							331
<i>O. stricta</i>							109
<i>Pinus</i> spp.	307	540	16.7	29.3	56.8	76.1	158
<i>Pinus pinaster</i>							331
<i>Pinus patula</i>							109
<i>Populus</i> spp.	154	314	8.4	17.1	49.1	103.8	158
<i>P. nigra var italica</i>							337
<i>Salix babylonica</i>	139	265	7.5	14.4	52.4	90.7	107
<i>Senna didymobotrya</i>	51	150	2.8	8.1	34.1	193.1	154
<i>Solanum mauritianum</i>	97	251	5.3	13.6	38.6	158.8	

The RT was calculated from the date of earliest record stated in Wilson et al. (2007). The RO was calculated subtracting the currently invaded QCs (%) to the potentially invaded QCs (%)

Table 6 For each biome, we reported the number of QCs within the biome, the number of QCs currently and potentially invaded, the percentage area suitable for invasion, the invasion

increase, the percentage of new QCs invadable compared to the total, and the number of IAPS for which the biome is suitable for invasion

Biomes	Total QCs	Currently invaded QCs	Potentially invaded QCs	Area suitable for invasion (%)	Invasion increase (%)	New QCs (%)	No. of IAPS
Albany thicket	110	83	102	92.7	22.9	17.3	10
Desert	9	0	0	0.0	0.0	0.0	0
Fynbos	236	150	200	84.7	33.3	21.2	9
Grasslands	601	428	517	86.0	20.8	14.8	13
Indian Ocean Coastal belt	50	46	50	100.0	8.7	8.0	10
Nama-Karoo	219	10	16	7.3	60.0	2.7	2
Savanna	489	294	393	80.4	33.7	20.2	14
Succulent Karoo	126	4	21	16.7	425.0	13.5	4

Table 7 For each province we reported the number of QCs within the province, the number of QCs currently and potentially invaded, the percentage area suitable for invasion,

the invasion increase, the percentage of new QCs invadable compared to the total, and the number of IAPS for which the province is suitable for invasion

Provinces	Total QCs	Currently invaded QCs	Potentially invaded QCs	Area suitable for invasion (%)	Invasion increase (%)	New QCs (%)	No. of IAPS
Western cape	310	138	192	61.9	39.1	17.4	8
Northern cape	203	4	11	5.4	175.0	3.4	2
Eastern cape	411	277	337	82.0	21.7	14.6	13
Free State	161	68	83	51.5	22.1	9.3	8
Northwest	75	36	51	68.0	41.7	20.0	9
KwaZulu-Natal	263	223	263	100.0	17.9	15.2	13
Gauteng	32	30	32	100.0	6.7	6.2	10
Mpumalanga	208	159	197	94.7	23.9	18.3	13
Limpopo	177	83	133	75.1	60.2	28.2	13

to underestimation of the current and potential distribution of IAPS that are underrepresented in the landscape layer. Third, the NIAPS may have underestimated or overestimated the current distribution of those species that are not very conspicuous or underrepresented for other reasons (e.g., taxonomic identification difficulties). Fourth, averaging the values of the explanatory variables for each QCs is based on the assumption that the mean values represent the point of occurrence of the IAPS. The likelihood of this assumption being erroneous depends on the level of variability of the explanatory variables in the QCs, and will be higher in complex topographic areas (Le Maitre et al. 2004). Fifth, the modeled invasion could

have been underestimated due to the fact that after the Gini impurity measure, we excluded 13 response variables from our SDM, which did not result in having statistically significant results.

However, this study is intended to provide guidance for decision-makers, and none of the factors mentioned above significantly affect the overall accuracy or usefulness of our results.

Invasion determinants

Our study showed that the addition of the human-related variables to the environmental set of variables did not result in statistically significant improvements

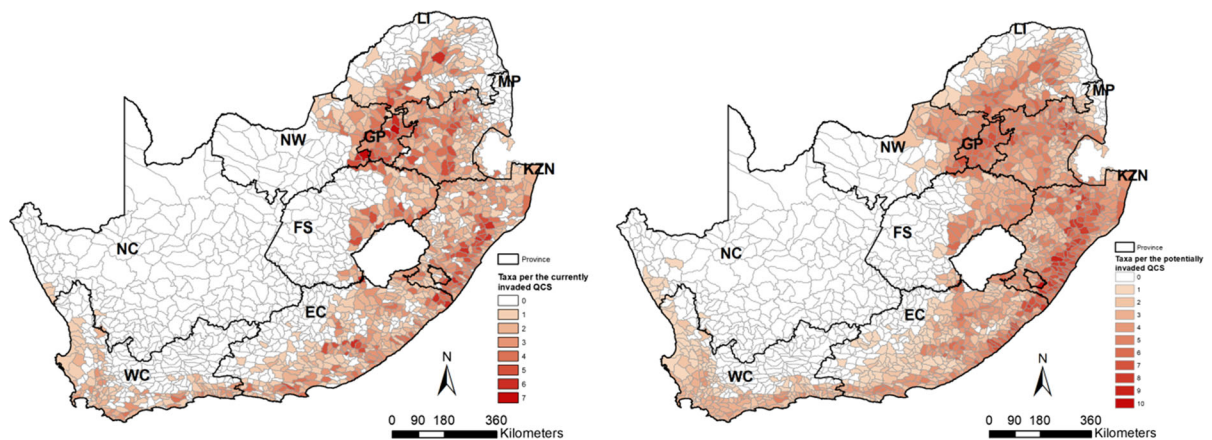


Fig. 2 Maps of South Africa showing the currently invaded QCs (left) and potentially invaded QCs (right) with the related number of IAPS currently/potentially found in each QCs.

Provinces: WC Western Cape, EC Eastern Cape, NC Northern Cape, NW North West, FS Free State, KZN KwaZulu-Natal, GP Gauteng, MP Mpumalanga, LI Limpopo. (Color figure online)

in the accuracy of SDM for 24 out of the 27 IAPS (Table 3). Improvements were obtained only for *Caesalpinia decapetala* and *Senna didymobotrya*, which are located in highly urbanized and degraded areas. The reason that the difference between the two GLMs was not significant could be related to the extension of the study area and the rather coarse scale of the study. This scale could undermine the effect of some human-related variables that require higher resolution to be influential.

Moreover, it was found that for only 14 out of the 27 IAPS considered, the variables were statistically significant in explaining the potential distributions (Table 4, Fig. 1). Consequently, we excluded the other 13 IAPS from subsequent analysis, which were found to be those with the least number of QC currently invaded (generally less than 50).

We observed some general trends when analyzing the results obtained from the two-part model and the RF model. The importance values obtained from the RF model showed that at this scale of analysis climatic variables are overall more important than land use, soil and anthropogenic variables in predicting the potential distribution of the IAPS, in line with Ohlemuller et al. (2006)'s findings (Table 10 in Appendix).

The two-part model analysis confirmed the importance of the climatic variables, as they proved to be influential for many of the IAPS. The logistic part of the model also highlighted some influence of the soil variables (Fig. 2, Table 8 in Appendix). We observed general trends of both positive and negative

correlations. The fact that the growing period and the mean autumn precipitation had an influence on the presence of plant species was in line with Rouget et al. (2004) and Coudun et al. (2006). We analyzed the maps of the soil types and noticed that the distribution of some soil variables matched the potential distribution patterns of the IAPS. For example, the plinthic horizon soils are usually associated with humid and subhumid warm climates with a distinct dry season (Fey 2010); such climatic conditions are characteristic of the Grasslands (Rutherford and Westfall 1986) where *Eucalyptus* spp., *Populus* spp. and *Salix babylonica* are major invaders (Henderson 1999; Van Wilgen 2009). The logistic part outputs also showed a negative correlation between the presence of most of the IAPS and the following climatic variables (Table 8 in Appendix): mean maximum temperature of hottest month, mean minimum temperature of coldest month (as in Rouget et al. 2002, 2004), frost period, and winter and summer precipitation (as in Randin et al. 2009).

In the regression part of the model (Table 9 in Appendix), which analyzed the influence of the environmental variables on species abundance, the fact that the growing days and the frost period had an influence on the abundance of IAPS was in line with Rouget et al. (2004), whereas the influence of the summer precipitation was in line with Pearman et al. (2008) and Randin et al. (2009). The influence of the mean annual temperature on IAPS was already demonstrated in many studies (e.g., Leathwick 1998;

Pearman et al. 2008; Rickebusch et al. 2008; Randin et al. 2009).

Aside from the climatic variables, this study seems to confirm that the invasion process is highly species-specific as well as spatially and temporally specific (Mack 1996; Wilson et al. 2007; Theoharides and Dukes 2007). As a result, the spread of the IAPS can be explained only through a combination of environmental factors, species characteristics and anthropic use for each species. Nonetheless, we are confident that by including land cover and soil variables in our models, the predictions are more accurate than by only including climatic variables (Ibáñez et al. 2009; Gassó et al. 2012). Moreover, progress in using SDMs will only be made by increasing the understanding of the ecological drivers of IAPS' distributions, and the extent of their relationship with determinant variables (Hijmans and Graham 2006). Failure to incorporate an influential predictor decreases model performance and outcome relevance (Austin and Van Niel 2011). Hence great attention should be given to the relative weight, causality, and estimation of each SDM predictor (Araújo and Guisan 2006).

Potential distribution

Once we selected the variables and the species for which the model provided statistically significant predictions, we analyzed the differences between the current and potential distributions, in order to assess the overall invasion risk for each species and the invasion susceptibility of the QCs. The potentially invaded QCs predicted by the model for the 14 IAPS considered, amounted to 1299 QCs out of the 1840 QCs, i.e., 568089.8 Km² (46, 5% of South Africa). Compared to the 1018 QCs currently invaded by our IAPS, the model predicted that 181 new QCs have the potential to be invaded, corresponding to 120972.0 Km², which would represent a 21.3% increase of the current invasion (Fig. 1). The 172 new QCs are in the proximity of the currently invaded QCs: along the western and eastern coasts, and across the northern plains. The QCs that appeared to be suitable to a higher number of IAPS were located along the eastern coast of KwaZulu-Natal and the plains, grasslands and savanna of Limpopo, Gauteng, and Mpumalanga (Fig. 2).

The invasion process resulted highly dependent on the biome types. The results showed that the Desert

had not been and was not predicted to be invaded; the Nama-Karoo and the Succulent Karoo have, respectively, only 2.7 and 13.5% of their QCs potentially to be invaded (Table 6). Conversely, the other biomes (Albany Thicket, Fynbos, Grasslands, Savanna, and the Indian Ocean Coastal Belt) were found to be suitable for invasion from IAPS with over 80% of their area suitable for invasion. This result may be explained considering the higher level of human disturbance, propagule pressure and nutrients that differentiates these biomes from those of the arid and hyperarid zones.

From a catchment and water stress perspective, the suitability of the Grasslands to invasion, particularly by a high number of woody IAPS including riparian *Salix babylonica* and *Populus* spp., is of great concern (Table 7). In fact, in this area, catchments have generally high water yields and these invaders can significantly reduce catchment runoff and cause great water scarcity (Le Maitre et al. 2000; Rouget et al. 2004).

Richardson and Rejmánek (2011) have shown that the Savanna is one of the biomes that most recently experienced a high increase in plant invasion due to propagule pressure from other savannas outside South Africa. Earlier literature indicated that in South African savanna plant invasion was not considered as a major issue (Parsons, 1972), however, subsequent studies highlighted the increased invasion in humid and riparian areas of this biome (Henderson and Wells 1986; Turpie 2004). The major invaders in the Savanna biome were found to be *Cereus jamacaru*, *Melia azederach* and species of *Eucalyptus*, *Acacia*, and *Agave*.

The model showed the suitability of South African provinces to invasion. The Northern Cape was the only province that had a very low suitability for IAPS; 5.4% of its areas could potentially be invaded (Table 7). In fact, the Northern Cape is mostly occupied by biomes with a very low suitability: Desert, Succulent- and Nama-Karoo. The lack of potential IAPS in the mountain areas of the Western Cape depends mainly from the fact that the Fynbos is invaded by alien species, fire dependent and well specialized to survive nutrient-poor soils (Richardson et al. 1997) that were excluded from the modeling due to the low number of sample points in the NIAPS. The other provinces were found to be highly suitable for IAPS invasion: 50 (Free state province) to 100% (Gauteng province) of the

total province area could potentially be suitable (Table 7, Fig. 2).

The KwaZulu-Natal was suitable to all the 14 IAPS considered, except *Acacia cyclops*. The area that could potentially be completely invaded amounts, and according to the low invasion increase value (17.9%), it seems that the future invasion will be related to an increase in the abundance of currently present IAPS, rather than to an increase in the number of invaded QCs. In this case, control actions should focus on eradicating the most affected areas rather than preventing further spread in other QCs. The Gauteng province resulted as the most threatened by our IAPS, in line with the fact that two of the most suitable biomes are found within it, namely, Grasslands and Savanna, and that it is characterized by large urban areas, such as Johannesburg. The invasion increase value is also very low (6.7%) since the current IAPS distribution is already consistent (Table 7).

Table 5 shows the difference between the invasion increase percentages among the IAPS considered. Some species had a rather limited current and potential invasion area, e.g., *Acacia cyclops* and *Senna didymobotrya*. These species tend to occupy specific and often very limited biomes, such as the Fynbos in the case of *Acacia cyclops*, and the Indian Ocean coastline in the case of *Senna didymobotrya*. Other species are characterized by a wider potential distribution: species of the genera *Acacia*, *Eucalyptus*, and *Pinus* are able to potentially invade several biomes, namely Savanna and Grasslands. Moreover, these last species are characterized by the highest relative occupancy *Eucalyptus*, *Pinus*, and *Acacia*, with more than 80% of the suitable area occupied, and thus appear to have nearly reached the full extent of their potential distribution range. Other species seem to be still in the full expansion phase, e.g., *Lantana camara* has an invasion increase of 432.2%, and *Melia azederach* of 229.6%. This difference in invasion dynamics is due to the following factors: climate, characteristics of the species, characteristics of the invaded biome, and anthropic plant use. In South Africa, the largest biomes are Savanna, Grasslands, and the Karoo, therefore, the species more suited to invading these biomes will have a greater chance to invade a broader area, compared to species adapted to other biomes (Wilson et al. 2007). Moreover, Maestre and Cortina (2004) showed that arid environments are more likely to be less competitive than humid environments that,

being more open and exposed, are more susceptible to invasion.

Residence time is often identified as having a positive correlation with the extent of occurrence of IAPS, i.e., species introduced earlier, on average occupy a higher proportion of their potential distribution range (Rejmánek 2000; Castro et al. 2005; Gassó et al. 2012). However, in South Africa, Thuiller et al. (2006) showed that minimum residence time did not explain the distribution patterns of invaders, even after removing the confounding effect of the environment, and that minimum residence time is a limited value when considering distribution patterns at regional scale after a century of residence. Similarly, in this study, no significant relationship was observed, as species introduced earlier (e.g., *Salix babylonica*, RO = 52.4%, RT = 337 years) showed a relative occupancy that was comparable to that of other species introduced more recently (e.g., *Chromolaena odorata*, RO = 49.6%, RT = 158 years). Some species introduced a long time ago had still a restricted distribution with respect to their modeled potential distribution range. Species introduced more recently (*Eucalyptus* spp.) had a much higher relative occupancy (RO = 74.1%) compared to species introduced only few years earlier (*Lantana camara* with a RO = 13%). The link between relative occupancy and minimum residence time could be weak for species introduced many centuries ago (*Salix babylonica*, *Pinus pinaster* and *Opuntia ficus-indica*) due to the high uncertainty related to old first records (Gassó et al. 2012). Furthermore, as mentioned, SDMs are based on the assumption that organisms are at equilibrium with their environment, however, this might not be the case for recently introduced species (*Opuntia stricta* and *Cereus jamacaru* introduced less than a century ago) which could undermine predictions (Gassó et al. 2012; Thuiller et al. 2006).

Conclusions

IAPS pose a significant threat to biodiversity and functioning of the ecosystems, and billions of dollars are spent to control them (de Lange and van Wilgen 2010). The most cost-effective approach is prevention (McConnachie et al. 2012). Hence, the success of biological invasion management depends on the ability to predict the potential distribution of IAPS

and identify the invasion determinants (Sharma et al. 2005). Van Wilgen et al. (2012) suggested that South African plant control programs should prioritize both the species and the areas: IAPS control operations begin by identifying the priority species and catchments, and then consulting with the decision-makers and key stakeholders (Balmford 2003). SDMs are powerful tools that can be used to plan IAPS management programs through: (i) the classification of invaded areas for differentiated management actions, (ii) the support of control initiatives for preventing IAPS' spread, (iii) the information of funds reallocation away from controlling IAPS in areas where suitability is expected to decrease in the future, and (iv) the identification of opportunities for relatively inexpensive invasion prevention (Kriticos et al. 2011).

In this paper, we modeled the distribution of IAPS using as sampling units the quaternary catchments that are also the basic hydrological unit for water management. This approach provides us with both scientific insights on IAPS spread processes as well as some indications for decision-making to plan control operations, for which efficiency and prioritization are

fundamental (Rouget et al. 2004). At the national scale, we confirmed the prevailing importance of climatic variables with respect to land cover and anthropogenic factors. Moreover, we found out that overall the newly potentially invaded QCs are much less than the currently invaded ones, therefore control operations should focus on managing the density of priority IAPS within their current range (see Van Wilgen et al. 2007), rather than preventing the expansion of the distribution (Rouget et al. 2004). This type of approach should necessarily be accompanied with finer scale analysis to identify local components of the problem so that the control operations can be structured as efficiently as possible.

Acknowledgements This study was supported by SECOSUD II Project funded by the Italian Agency for Development Cooperation. We are also grateful to Laura Clarke for revising the text.

Appendix

See Tables 8, 9 and 10.

Table 8 Log odds or estimate values of the Logistic part of the two-part model

	Ptot	Pw	Psm	Psp	Pa	TM	TMX	TMN	Frost	MGS	pH	PAW	Natur	Water	River
<i>Acacia cyclops</i>	- 0.247	- 0.118	0.345	3.948	- 1.915	- 3.548									- 30.852
<i>Agave</i> spp.	- 0.043	- 0.021		- 0.3									- 3.772		
<i>Cereus jamacaru</i>	- 0.084	0.087	10.035	- 4.354	- 4.023	0.148	7.333								
<i>Chromolaena odorata</i>		2.809				0.039	1.937								
<i>Eucalyptus</i> spp.	- 0.004	- 0.005		- 0.175	- 0.126	- 0.015	0.015			0.021				4.901	
<i>Lantana camara</i>						- 0.044						5.18			
<i>Melia azedarach</i>			1.214	- 0.77	- 0.044							3.271			
<i>Opuntia</i> spp.	- 0.017	0.064	0.025	1.282	- 0.924	- 0.802	- 0.023	- 0.013						- 24.605	
<i>Pinus</i> spp.	- 0.006		- 0.658			- 0.018	0.011							8.191	
<i>Populus</i> spp.	- 0.006	- 0.023	- 0.299			0.028									
<i>Salix babylonica</i>	- 0.012	- 0.052	- 0.536			0.047									
<i>Senna didymobotrya</i>	0.022	- 0.096	2.246	- 2.594						0.069					
<i>Solanum mauritianum</i>	- 0.034	- 0.074	0.042												
<i>Acacia</i> spp.	- 0.012	- 0.019	0.02	- 0.608	- 0.171	- 0.031	0.012	- 0.705				0.992			
	Slope	Soil1	Soil2	Soil3	Soil4	Soil5	Soil6	Soil7	Soil8	Soil9					
<i>Acacia cyclops</i>		- 25.182											- 5.115		
<i>Agave</i> spp.			15690.35	15686.5	15679.94	15693.09	15682.43	15686.87						15692.1	
<i>Cereus jamacaru</i>	0.377		6.758										2.481		
<i>Chromolaena odorata</i>	0.279		1.845	3.453		4.053	3.364	2.765						3.714	
<i>Eucalyptus</i> spp.	- 0.079	3.463													
<i>Lantana camara</i>			- 2.652												
<i>Melia azedarach</i>	- 0.134												1.16		1.188
<i>Opuntia</i> spp.															
<i>Pinus</i> spp.						1.13									
<i>Populus</i> spp.						2.236	1.699							2.798	
<i>Salix babylonica</i>						2.145	2.717							4.458	
<i>Senna didymobotrya</i>															
<i>Solanum mauritianum</i>		146.612	148.227	144.167	140.78	143.199	142.36	137.222							
<i>Acacia</i> spp.					1.044	0.645								- 2.115	

Table 9 Coefficient or estimate values of the regression part of the two-part model

	Ptot	Pw	Psm	Psp	Pa	TM	TMX	TMN	Frost	MGS	pH	PAW	Natur	Water	River
<i>Acacia cyclops</i>			- 0.037	0.147		- 0.463			- 0.022						
<i>Agave</i> spp.	- 0.014			0.036	0.909	- 0.816				0.014				- 32.631	
<i>Cereus jamacaru</i>										0.029					
<i>Chromolaena odorata</i>										- 0.006					
<i>Eucalyptus</i> spp.	0.008	- 0.009	- 0.007	- 0.014	- 0.101										3.331
<i>Lantana camara</i>				0.032	- 3.059	2.084	1.112			- 0.024					15.066
<i>Melia azedarach</i>									- 0.009	0.01		- 0.019			- 4.903
<i>Opuntia</i> spp.	0.004				- 0.359		- 0.332		- 0.027	- 0.013	0.604	0.033			
<i>Pinus</i> spp.	- 0.003		0.002	0.017		- 0.195			- 0.005	0.01					
<i>Populus</i> spp.					0.017	- 0.584	0.403								
<i>Salix babylonica</i>			- 0.026		0.021	- 0.632	0.435	0.302							
<i>Senna didymobotrya</i>			- 0.056												8.551
<i>Solanum mauritianum</i>					0.223					0.015					
<i>Acacia</i> spp.			- 0.006		- 0.658	0.141	- 0.013	0.006	- 0.386		1.153			- 3.857	
	Slope	Soil1	Soil2	Soil3	Soil4	Soil5	Soil6	Soil7	Soil8	Soil9					
<i>Acacia cyclops</i>	- 0.111	- 7.297									- 3.314				
<i>Agave</i> spp.															
<i>Cereus jamacaru</i>	- 0.180			4.083		2.428									
<i>Chromolaena odorata</i>	0.041														2.335
<i>Eucalyptus</i> spp.		4.868	4.947	5.152		4.646	4.331	4.321							
<i>Lantana camara</i>															
<i>Melia azedarach</i>															
<i>Opuntia</i> spp.			- 1.549												- 1.726
<i>Pinus</i> spp.															- 2.025
<i>Populus</i> spp.			- 1.083												- 1.379
<i>Salix babylonica</i>			- 1.739												- 1.039
<i>Senna didymobotrya</i>				33.499											
<i>Solanum mauritianum</i>					- 4.941										
<i>Acacia</i> spp.															- 2.826

Table 10 Classification of explanatory variables importance predicted with the Random Forest (RF) model

	Ptot	Pw	Psm	Psr	Pa	Tm	TMX1	TMN7	Frost	MGS	pH	PAW	Natur	Water	River	Slope
<i>Acacia cyclops</i>	2	3	5	8	10	1	16	13	7	9	27	6	20	4	11	31
<i>Acacia</i> spp.	5	10	1	7	2	11	4	6	16	32	12	3	8	9	27	13
<i>Agave</i> spp.	1	4	5	2	10	11	31	3	29	13	6	7	9	8	27	18
<i>Cereus jamacaru</i>	3	4	1	10	2	5	6	9	8	11	7	13	29	21	16	31
<i>Chromolaena odorata</i>	6	5	1	8	10	9	2	4	7	3	29	31	16	32	28	11
<i>Eucalyptus</i> spp.	10	5	1	2	4	3	11	7	27	13	16	21	6	29	32	31
<i>Lanata camara</i>	1	4	2	5	11	10	6	8	7	9	3	12	29	13	16	22
<i>Melia azedarach</i>	1	2	6	5	7	10	11	3	4	9	8	31	13	21	27	16
<i>Opuntia</i> spp.	1	5	2	10	11	31	4	3	7	13	6	8	9	27	23	16
<i>Pinus</i> spp.	10	7	1	3	5	2	4	6	11	16	27	19	13	8	9	29
<i>Populus</i> spp.	1	5	4	10	8	2	9	3	6	13	11	29	21	27	16	7
<i>Salix babylonica</i>	1	10	2	5	6	9	4	8	11	3	7	13	16	22	25	12
<i>Senna didymobotrya</i>	5	1	10	6	4	2	11	3	7	8	9	29	23	12	32	27
<i>Solanum mauritianum</i>	1	2	5	10	4	6	3	8	9	7	23	16	32	11	13	29
	Soil1	Soil2	Soil3	Soil4	Soil5	Soil6	Soil7	Soil8	Soil9	N	Fire	Cult	Degr	Urb	Rail	Road
<i>Acacia Cyclops</i>	12	19	15	32	26	28	23	22	17	29	24	30	14	18	21	25
<i>Acacia</i> spp.	29	23	19	18	28	15	21	31	22	25	24	26	14	20	30	17
<i>Agave</i> spp.	16	23	12	32	22	28	21	19	30	15	24	14	26	25	20	17
<i>Cereus jamacaru</i>	19	27	18	12	28	14	30	32	15	22	24	23	25	20	26	17
<i>Chromolaena odorata</i>	12	23	13	27	18	26	14	15	25	19	24	22	21	30	20	17
<i>Eucalyptus</i> spp.	30	8	9	28	14	12	19	26	18	23	24	22	15	25	20	17
<i>Lanata camara</i>	28	32	18	27	19	23	21	26	14	15	31	30	20	25	17	24
<i>Melia azedarach</i>	29	23	22	12	32	18	15	26	14	28	24	30	25	19	20	17
<i>Opuntia</i> spp.	26	32	12	28	29	15	19	14	18	22	24	21	25	30	20	17
<i>Pinus</i> spp.	28	15	31	32	23	17	14	12	30	22	24	21	26	18	25	20
<i>Populus</i> spp.	19	28	12	25	23	22	31	18	26	32	24	14	15	17	30	20
<i>Salix babylonica</i>	28	27	29	21	23	18	19	32	31	14	24	30	15	26	20	17
<i>Senna didymobotrya</i>	28	26	31	16	18	13	14	22	15	19	24	20	21	30	17	25
<i>Solanum mauritianum</i>	26	27	31	12	28	18	19	30	22	14	24	15	20	25	21	17

References

- Agresti A (2002) Categorical Data Analysis. Wiley, New York
- Araújo MB, Guisan A (2006) Five (or so) challenges for species distribution modelling. *J Biogeogr* 33(10):1677–1688
- Araújo MB, Luoto M (2007) The importance of biotic interactions for modelling species distributions under climate change. *Glob Ecol Biogeog* 16(6):743–753
- Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecol Model* 157(2–3):101–118

- Austin MP, Van Niel KP (2011) Improving species distribution models for climate change studies: variable selection and scale. *J Biogeogr* 38(1):1–8
- Balmford A (2003) Conservation planning in the real world: South Africa shows the way. *Trends Ecol Evol* 18(9):435–438
- Breiman L (2001) Random Forests. *Mach Learn* 45:5–32
- Broennimann O, Treier UA, Müller-Schärer H, Thuiller W, Peterson AT, Guisan A (2007) Evidence of climatic niche shift during biological invasion. *Ecol Lett* 10(8):701–709
- Calder I, Dye P (2001) Hydrological impacts of invasive alien plants. *Land Use W Resour Res* 1:1–12
- Castro SA, Figueroa JA, Muñoz-Schick M, Jaksic FM (2005) Minimum residence time, biogeographical origin, and life cycle as determinants of the geographical extent of naturalized plants in continental Chile. *Divers Distrib* 11(3):183–191
- Chytrý M, Jarošík V, Pyšek P, Hájek O, Knollová I, Tichý L, Danihelka J (2008) Separating habitat invasibility by alien plants from the actual level of invasion. *Ecology* 89(6):1541–1553
- Coudun C, Gégout JC, Piedallu C, Rameau JC (2006) Soil nutritional factors improve models of plant species distribution: an illustration with *Acer campestre* (L.) in France. *J Biogeogr* 33(10):1750–1763
- Council for Scientific and Industrial Research and Agricultural Research Council (2005) National land cover database (2000) Council for Scientific and Industrial Research and Agricultural Research Council. Pretoria, South Africa
- Cullis J, Gorgens A, Marais C (2007) A strategic study of the impact of invasive alien plants in the high rainfall catchments and riparian zones of South Africa on total surface water yield. *Water SA* 33(1):35–42
- de Lange WJ, van Wilgen BW (2010) An economic assessment of the contribution of biological control to the management of invasive alien plants and to the protection of ecosystem services in South Africa. *Biol Invasions* 12:4113–4124
- del Barrio G, Harrison PA, Berry PM, Butt N, Sanjuan ME, Pearson RG, Dawson T (2006) Integrating multiple modelling approaches to predict the potential impacts of climate change on species' distributions in contrasting regions: comparison and implications for policy. *Environ Sci Policy* 9(2):129–147
- Di Lorenzo B, Farcomeni A, Golini N (2011) A Bayesian model for presence-only semicontinuous data, with application to prediction of abundance of *Taxus Baccata* in two Italian regions. *J Agric, Biol Environ Stat* 16(3):339–356
- Dormann CF (2007) Promising the future? Global change projections of species distributions. *Basic Appl Ecol* 8(5):387–397
- Elith J, Leathwick JR (2009) Species distribution models: ecological explanation and prediction across space and time. *Annu Rev Ecol, Evol Systematics* 40(1):677–697
- Fey M (2010) Soils of South Africa. Cambridge University Press, Cambridge, p 287
- Ficetola GF, Thuiller W, Miaud C (2007) Prediction and validation of the potential global distribution of a problematic alien invasive species - the American bullfrog. *Divers Distrib* 13(4):476–485
- Franklin J (2010) Moving beyond static species distribution models in support of conservation biogeography. *Divers Distrib* 16(3):321–330
- Funk JL, Sandish RJ, Stock W, Valladares F (2016) Plant functional traits of dominant native and invasive species in mediterranean- climate ecosystems. *Ecology* 97(1):75–83
- Gassó N, Sol D, Pino J, Dana ED, Lloret F, Sanz-Elorza M, Sobrino E, Vilà M (2009) Exploring species attributes and site characteristics to assess plant invasions in Spain. *Divers Distrib* 15(1):50–58
- Gassó N, Thuiller W, Pino J, Vilà M (2012) Potential distribution range of invasive plant species in Spain. *NeoBiota* 12:25–40
- Giglio L, Loboda T, Roy DP, Quayle B, Justice CO (2009) An active-fire based burned area mapping algorithm for the MODIS sensor. *Remote Sens Environ* 113:408–420
- Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecol Lett* 8(9):993–1009
- Henderson L (1999) The Southern African Plant Invaders Atlas (SAPIA) and its contribution to biological weed control. *Afr Entomol Memoir* 1:159–163
- Henderson L, Wells MJ (1986) Alien plant invasions in the grassland and savanna biomes. In: Macdonald IAW, Kruger FJ, Ferrar AA (eds) The ecology and management of biological invasions in Southern Africa. Oxford University Press, Cape Town, pp 109–117
- Hengl T, Heuvelink GBM, Kempen B, Leenaars JGB, Walsh MG, Shepherd KD, Sila A, MacMillan RA, Mendes de Jesus J, Tamene L, Tondoh JE (2015) Mapping soil properties of Africa at 250 m resolution: random forests significantly improve current predictions. *PLoS ONE* 10:e0125814
- Higgins SI, Richardson DM (1996) A review of models of alien plant spread. *Ecol Model* 87:249–265
- Higgins SI, Richardson DM, Cowling RM (2000) Using a dynamic landscape model for planning the management of alien plant invasions. *Ecol Appl* 10(6):1833–1848
- Hijmans RJ, Graham CH (2006) The ability of climate envelope models to predict the effect of climate change on species distributions. *Glob Change Biol* 12(12):2272–2281
- Hobbs RJ (2004) The Working for water programme in South Africa: the science behind the success. *Divers Distrib* 10(5–6):501–503
- Ibáñez I, Silander JA, Wilson AM, La Fleur N, Tanaka N, Tsuyama I (2009) Multivariate forecasts of potential distributions of invasive plant species. *Ecol Appl* 19(2):359–375
- Kotzé JDF, Beukes BH, Van den Berg EC, Newby TS (2010) National Invasive Alien Plant Survey. Report Number: GW/A/(2010)/21. Agricultural Research Council: Institute for Soil, Climate and Water, Pretoria
- Kriticos DJ, Watt MS, Potter KJB, Manning LK, Alexander NS, Tallent-Halsell N (2011) Managing invasive weeds under climate change: considering the current and potential future distribution of *Buddleja davidii*. *Weed Res* 51(1):85–96
- Land Type Survey Staff 1972–2006 (2006) Land Types of South Africa: Maps (69 sheets) and Memoirs (39 books). ARC-Institute for Soil, Climate and Water, Pretoria

- Le Maitre DC, van Wilgen BW, Chapman RA, McKelly DH (1996) Invasive plants and water resources in the Western Cape Province, South Africa: modelling the consequences of a lack of management. *J Appl Ecol* 33:161–172
- Le Maitre DC, van Wilgen BW, Cowling RM (1998) Ecosystem services, efficiency, sustainability and equity: South Africa's Working for Water programme. *Trends Ecol Evol* 13(9):378
- Le Maitre DC, Versfeld DB, Chapman R (2000) The impact of invading alien plants on surface water resources in South Africa: a preliminary assessment. *Water SA* 26(3):397–408
- Le Maitre DC, Richardson DM, Chapman RA (2004) Alien plant invasions in South Africa: driving forces and the human dimension. *S Afr J Sci* 100:103–112
- Le Maitre DC, Forsyth GG, Dzikiti S, Gush MB (2016) Estimates of the impacts of invasive alien plants on water flows in South Africa. *Water SA* 42(4):659–672
- Leathwick JR (1998) Are New Zealand's Nothofagus species in equilibrium with their environment? *J Veg Sci* 9:719–732
- Luoto M, Heikkinen RK, Pöyry J, Saarinen K (2006) Determinants of the biogeographical distribution of butterflies in boreal regions. *J Biogeogr* 33(10):1764–1778
- Mack RN (1996) Predicting the identity and fate of plant invaders: emergent and emerging approaches. *Biol Conserv* 78(1–2):107–121
- Maestre FT, Cortina J (2004) Are *Pinus halepensis* plantations useful as a restoration tool in semiarid Mediterranean areas? *For Ecol Manag* 198(1–3):303–317
- Mallory S, Versfeld D, Nditwani T (2011) Estimating streamflow reduction due to invasive alien plants: Are we getting it right? Techniques used to estimate streamflow reduction due to invasive alien plants. *Proceedings of the 15th SAN-CIAHS Symposium, Grahamstown*, pp. 1–11, 12–14 Sept 2011
- Marais C, van Wilgen BW, Stevens D (2004) The clearing of invasive alien plants in South Africa: a preliminary assessment of costs and progress. *S Afr J Sci* 100:97–103
- McConnachie MM, Cowling RM, van Wilgen BW, McConnachie DA (2012) Evaluating the cost-effectiveness of invasive alien plant clearing: a case study from South Africa. *Biol Conserv* 155:128–135
- Milton SJ, Dean WR (2010) Plant invasions in arid areas: special problems and solutions: a South African perspective. *Biol Invasions* 12:3935–3948
- Mucina L, Rutherford MC (2006) 'The vegetation of South Africa, Lesotho and Swaziland', *Sterlitzia* 19. S Afr Natl Biodivers Inst, Pretoria
- Nel JL, Richardson DM, Rouget M, Mgidi TN, Mdzeke N, Le Maitre DC, van Wilgen BW, Schonegevel L, Henderson L, Naser S (2004) Proposed classification of invasive alien plant species in South Africa: towards prioritizing species and areas for management action. *S Afr J Sci* 100:53–64
- Ohlemuller R, Walker S, Bastow Wilson J (2006) Local vs regional factors as determinants of the invasibility of indigenous forest fragments by alien plant species. *Oikos* 112(3):493–501
- Parsons JJ (1972) Spread of African pasture grasses to the American tropics. *J Rang Manage* 25:12–17
- Pauchard A, Alaback BP (2004) Influences of evaluation, land use and landscape context on patterns of alien plant invasions along roadsides in protected areas of south-central Chile. *Conserv Biol* 18(1):238–248
- Pearman PB, Randin CF, Broennimann O, Vittoz P, van der Knaap WO, Engler E, Le Lay G, Zimmermann NE, Guisan A (2008) Prediction of plant species distributions across six millennia. *Ecol Lett* 11(4):357–369
- Pearson RG (2010) Species' distribution modeling for conservation educators and practitioners. *Lessons Conserv* 3:54–89
- Pearson RG, Dawson TP (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob Ecol Biogeogr* 12:361–371
- Pearson RG, Dawson TP, Liu C (2004) Modelling species distributions in Britain: a hierarchical integration of climate and land-cover data. *Ecography* 27(3):285–298
- Pejchar L, Mooney H (2009) Invasive species, ecosystem services and human well-being. *Trend Ecol Evol* 24(9):497–504
- Peterson AT, Vieglais DA (2001) Predicting species invasions using ecological niche modeling: new approaches from bioinformatics attack a pressing problem. *Bioscience* 51(5):363–371
- Pino J, Font X, Carbó J, Jové M, Pallarès L (2005) Large-scale correlates of alien plant invasion in Catalonia (NE of Spain). *Biol Conserv* 122(2):339–350
- R Core Team (2014). R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- Randin CF, Engler R, Normand S, Zappa M, Zimmermann NE, Pearman PB, Vittoz P, Thuiller W, Guisan A (2009) Climate change and plant distribution: local models predict high-elevation persistence. *Glob Change Biol* 15(6):1557–1569
- Rejmánek M (2000) Invasive plants: approaches and predictions. *Austral Ecol* 25(5):497–506
- Richardson DM (1998) Forestry trees as invasive aliens. *Conserv Biol* 12:18–26
- Richardson DM, Rejmánek M (2011) Trees and shrubs as invasive alien species - a global review. *Divers Distrib* 17(5):788–809
- Richardson DM, Macdonald IAW, Hoffmann JH, Henderson L (1997) Alien plant invasions. In: Cowling RM, Richardson DM, Pierce SM (eds) *Vegetation of Southern Africa*. Cambridge University Press, Cambridge (United Kingdom), pp 534–570
- Richardson DM, Rouget M, Ralston SJ, Cowling RM, Van Rensburg BJ, Thuiller W (2005) Species richness of alien plants in South Africa: environmental correlates and the relationship with indigenous plant species richness. *Ecology* 12(3):391–402
- Rickebusch S, Thuiller W, Hickler T, Araújo MB, Sykes MT, Schweiger O, Lafourcade B (2008) Incorporating the effects of changes in vegetation functioning and CO₂ on water availability in plant habitat models. *Biol Lett* 4(5):556–559
- Robertson MP, Villet MH, Fairbanks DHK, Henderson L, Higgins SI, Hoffmann JH, Le Maitre DC, Palmer AR, Riggs I, Shackleton CM, Zimmermann HG (2003) A proposed prioritization system for the management of invasive alien plants in South Africa. *S Afr J Sci* 99:37–43

- Rödger D, Schmidtlein S, Veith M, Lötters S (2009) Alien invasive slider turtle in unpredicted habitat: a matter of niche shift or of predictors studied? *PLoS ONE* 4(11):e7843
- Rouget M, Richardson DM (2003) Inferring Process from Pattern in Plant Invasions: a semimechanistic model incorporating propagule pressure and environmental factors. *Am Nat* 162(6):713–724
- Rouget M, Richardson DM, Nel JL, van Wilgen BW (2002) Commercially important trees as invasive aliens – towards spatially explicit risk assessment at a national scale. *Biol Invasions* 4(1):397–412
- Rouget M, Richardson DM, Nel JL, Le Maitre DC, Egoh B, Mgidi T (2004) Mapping the potential ranges of major plant invaders in South Africa, Lesotho and Swaziland using climatic suitability. *Divers Distrib* 10(5–6):475–484
- Roura-Pascual N, Suarez AV, Gómez C, Pons P, Touyama Y, Wild AL, Peterson AT (2004) Geographical potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. *Proc R Soc B* 271(1557):2527–2535
- Rutherford MC, Westfall RH (1986) The Biomes of Southern Africa - an objective categorization. *Mem Bot Surv S Afr* 54:1–98
- Schulze RD (2008) South African atlas of climatology and agrohydrology. Water Research Commission, Pretoria, RSA, WRC Report 1489/1/08, Section 18.3
- Segurado P, Araújo MB (2004) An evaluation of methods for modelling species distributions. *J Biogeogr* 31(10):1555–1568
- Senan AS, Tomasetto F, Farcomeni A, Somashekar RK, Attorre F (2012) Determinants of plant species invasions in an arid island: evidence from Socotra Island (Yemen). *Plant Ecol* 213(9):1381–1392
- Sharma GP, Singh JS, Raghubanshi AS (2005) Plant invasions: emerging trends and future implications. *Curr Sci* 88(5):726–734
- Staver AC, Archibald S, Levin SA (2011) The Global Extent and Determinants of Savanna and Forest as Alternative Biome States. *Science* 334(6053):230–232
- Theoharides K, Dukes JS (2007) Plant invasion across space and time: factors affecting nonindigenous species success during four stage of invasion. *N Phytol* 176(2):256–273
- Thuiller W, Araújo MB, Lavorel S (2003) Generalized models vs. classification tree analysis: predicting spatial distributions of plant species at different scales. *J Veg Sci* 14(5):669–680
- Thuiller W, Araújo MB, Lavorel S (2004) Do we need land-cover data to model species distributions in Europe? *J Biogeogr* 31(3):353–361
- Thuiller W, Richardson DM, Pysek P, Midgley GF, Hughes GO, Rouget M (2005) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Glob Change Biol* 11:2234–2250
- Thuiller W, Richardson DM, Rouget M, Procheş S, Wilson JRU (2006) Interactions between environment, species traits, and human uses describe patterns of plant invasions. *Ecology* 87(7):1755–1769
- Thuiller W, Richardson DM, Midgley GF (2007) Will climate change promote alien plant invasions? *Biol Invasions* 193:197–211
- Turpie J (2004) The role of resource economics in the control of invasive alien plants in South Africa. *S Afr J Sci* 100:87–93
- van Wilgen BW (2009) The evolution of fire and invasive alien plant management practices in fynbos. *S Afr J Sci* 105:335–342
- van Wilgen BW, Richardson DM, Le Maitre DC, Marais C, Magadla D (2001) The economic consequences of alien plant invasion: examples of impacts and approaches to sustainable management in South Africa. *Environ Dev Sustain* 3:145–168
- van Wilgen BW, Nel JL, Rouget M (2007) Invasive alien plants and South African rivers: a proposed approach to the prioritization of control operations. *Freshw Biol* 52(4):711–723
- van Wilgen BW, Forsyth GG, Le Maitre DC, Wannenburgh A, Kotzé JDF, van den Berg E, Henderson L (2012) An assessment of the effectiveness of a large, national-scale invasive alien plant control strategy in South Africa. *Biol Conserv* 148(1):28–38
- van Wilgen BW, Fill J, Baard J, Kraaij T (2016) Historical costs and projected future scenarios for the management of invasive alien plants in protected areas in the Floristic Region. *Biol Conserv* 200:168–177
- Versfeld DB, Le Maitre DC, Chapman RA (1998) Alien invading plants and water resources in South Africa: a preliminary assessment. WRC Report No. TT99/98. Water Research Commission, Pretoria
- Wilson JRU, Richardson DM, Rouget M, Procheş S, Amis MA, Henderson L, Thuiller W (2007) Residence time and potential range: crucial considerations in modelling plant invasions. *Divers Distrib* 13:11–22
- Working for Water Mapping Standards (2003) Standards for mapping and management of alien vegetation and operational data, Version 5. Department of Water Affairs and Forestry, Working for Water Programme, Cape Town