

# Inferring Process from Pattern in Plant Invasions: A Semimechanistic Model Incorporating Propagule Pressure and Environmental Factors

Mathieu Rouget\* and David M. Richardson†

Institute for Plant Conservation, Botany Department, University of  
Cape Town, Rondebosch 7701, South Africa

Submitted February 17, 2003; Accepted July 1, 2003;  
Electronically published October 22, 2003

Online enhancement: figure.

---

**ABSTRACT:** Propagule pressure is intuitively a key factor in biological invasions: increased availability of propagules increases the chances of establishment, persistence, naturalization, and invasion. The role of propagule pressure relative to disturbance and various environmental factors is, however, difficult to quantify. We explored the relative importance of factors driving invasions using detailed data on the distribution and percentage cover of alien tree species on South Africa's Agulhas Plain (2,160 km<sup>2</sup>). Classification trees based on geology, climate, land use, and topography adequately explained distribution but not abundance (canopy cover) of three widespread invasive species (*Acacia cyclops*, *Acacia saligna*, and *Pinus pinaster*). A semimechanistic model was then developed to quantify the roles of propagule pressure and environmental heterogeneity in structuring invasion patterns. The intensity of propagule pressure (approximated by the distance from putative invasion foci) was a much better predictor of canopy cover than any environmental factor that was considered. The influence of environmental factors was then assessed on the residuals of the first model to determine how propagule pressure interacts with environmental factors. The mediating effect of environmental factors was species specific. Models combining propagule pressure and environmental factors successfully predicted more than 70% of the variation in canopy cover for each species.

**Keywords:** determinants of spread, exotic species, invasion barriers, mass effect, plant-environment interactions, tree invasions.

---

\* Corresponding author. Present address: Kirstenbosch Research Centre, National Botanical Institute, Private Bag X7, Claremont 7735, South Africa; e-mail: rouget@nbict.nbi.ac.za.

† E-mail: rich@botzoo.uct.ac.za.

Many interacting factors determine the range and abundance of an organism at a given locality. Any attempt to model range or abundance makes various assumptions depending on the aims of the study and the nature of available data (e.g., spatial and temporal scale, precision and accuracy of data). Over large spatial scales (regions to continents), static approaches have been quite successful in modeling species distribution (Franklin 1995; Guisan and Zimmermann 2000 and references therein). Such approaches rely on observed correlations between the environment and distribution; they assume that the organism is in equilibrium with the environment (Higgins and Richardson 1996). Invasive alien species, especially when the invasion is still at an early stage, are generally not at equilibrium with the environment. At this stage, invasions are frequently more constrained by propagule availability than by habitat requirements, and correlative approaches should therefore be used with caution. Over small spatial scales (stands to landscapes), mechanistic or process-based models such as individual-based cellular automata models can integrate biological attributes, space, and environmental stochasticity to model distribution (Higgins and Richardson 1996). Such models have clear theoretical advantages but are difficult to parameterize (Higgins and Richardson 1996) and validate (Higgins et al. 2001). In a few cases, individual-based models have been successfully applied for modeling spread of invasive alien plants (Higgins et al. 2000).

Biological invasions provide useful natural experiments that capture the combined effects of many interacting factors (Richardson et al. 2004). Considerable understanding of the ecology of a given invasion episode can be gained by correlating observed spatial patterns (presence or absence, abundance, age distribution, etc.) with environmental data (e.g., Higgins et al. 1999; Rouget et al. 2001). A fundamental understanding of the processes driving invasions demands a more mechanistic approach, adding biological processes to the modeling procedure. Many recent studies have addressed key issues in invasion ecology

using formal experiments, for example, to elucidate the determinants of invasibility. These studies may provide robust, nontrivial generalizations in the future, but currently available evidence from such experiments offers little practical assistance to managers (reviewed in Rejmánek et al. 2004). Also, such studies can only practically address the invasion ecology of short-lived species. For invasions of trees, shrubs, and other long-lived plants, distribution and population structure at any time reflects the cumulative effects of numerous factors, some of them rare events. More attention needs to be given to finding ways of extracting more information from existing spatial patterns and whatever can be derived from available data on the processes that produced these patterns.

The invasion process can be conceptualized as a series of barriers that an invading organism must overcome to become fully integrated in an ecosystem (Richardson et al. 2000b). For many invasions, the trajectory of population development (population growth or decline; spatial spread or contraction) can be reconstructed to describe the eventual outcome of many interacting processes that mediate the fate of immigrants. The strength of different barriers depends, to varying degrees in different systems, on the number of propagules: the greater the number of propagules, the greater the chance of a barrier being overcome and therefore invasion. The role of propagule pressure or mass effect is clearly observable in many invasions (Newsome and Noble 1986; Williamson 1996; Green 1997; Hutchinson and Vankat 1997; Lonsdale 1999). D'Antonio et al. (2001) provided a useful conceptualization of the role of propagule pressure in mediating plant invasions. They suggest that high propagule pressure can overcome biotic resistance (to some extent) but that abiotic controls on invasibility are less open to amelioration by high numbers of propagules. Few studies have attempted to quantify the role of propagule pressure in plant invasions (Williamson 1996; but see Rejmánek et al. 2004). Clearly, propagule pressure must be built into spatially explicit invasion models. As Williamson (1996, p. 55) wrote, "Looking for real differences in invasibility requires looking at the residuals from the relationship between invasion success and propagule pressure." A major challenge is to parameterize models in this regard, particularly for long-lived species.

This article explores the relative importance of factors driving plant invasions using extremely detailed data on the distribution and age structure of alien trees and shrubs on the Agulhas Plain in South Africa's Cape Floristic Region. Our aim was to explore the role of propagule pressure over the full invasion history in landscapes. The temporal scale under consideration is thus the interval between establishment of what can now be recognized as invasion foci and the present. We are thus primarily interested in exploring the roles of the oldest and now densest stands

in driving subsequent invasions (while fully appreciating that newer foci and satellite populations themselves act as secondary foci). We develop spatially explicit models of the distribution and spread of invasive alien plants that are tractable, ecologically sound, and accurate over relatively large areas ( $>1,000 \text{ km}^2$ ) and at fine-scale resolution ( $\leq 250 \text{ m}$ ). To be useful in management, such models must accurately estimate species canopy cover (abundance). For invasive alien plant species, percentage cover is largely influenced by propagule pressure and environmental suitability (Rejmánek et al. 2004). We thus need a simple model that retains key biological processes (e.g., related to propagule dispersal) and accounts for environmental heterogeneity. We also aimed to tease out the interactions between invasion and environment using the current pattern of invasion in the Agulhas Plain as a natural experiment. The Agulhas Plain is a superb locality to address these issues. Many woody alien species have been introduced and subsequently spread over the last century. Because of the short natural vegetation, invasive woody plants could be mapped very accurately. Although the spread of these species is driven by fire, the original invaded stands remain in the landscape (Richardson et al. 1992). Furthermore, good environmental data were available for this area.

The specific objectives of this article were to test the ability of correlative approaches to model species occurrence and cover at fine spatial scales, to develop a new approach for modeling cover of long-lived invasive plants in a heterogeneous environment, and to explore the interactions between the spread of invasive species and the environment.

## Methods

### *Study Area*

The Agulhas Plain lies at the southern tip of the Cape Floristic Region (South Africa) and covers approximately  $2,160 \text{ km}^2$ . It is one of the centers of plant biodiversity within the Cape Floristic Province hot spot of plant diversity and endemism (Myers et al. 2000). The predominant natural vegetation types are two types of fire-prone shrublands: fynbos, on sandy, infertile soils, and renosterveld, on more fertile soils (Cowling et al. 1988). About 40% of the area has been transformed by cultivation (Lombard et al. 1997). Almost the entire Agulhas Plain is invaded by woody alien plants to some extent. Alien species occur in 72% of the total area and in 96% of the remaining vegetation (cultivated and urbanized areas excluded). The three most widespread invasive species are (total area indicated in parentheses) *Acacia cyclops* (113,817 ha), *Acacia*

*saligna* (89,162 ha), and *Pinus pinaster* (78,654 ha). These occur mostly at low percentage cover.

#### Data Collection

Very detailed field mapping was conducted to map native vegetation as well as the distribution, percentage canopy cover, and height (age class) structure of all invasive alien trees and shrubs for the entire Agulhas Plain. All areas, whether invaded or not, were surveyed over 6 mo in 1998. Within each stand of alien trees or shrubs (relatively uniform clumps larger than 0.5 ha), the percentage cover and average height (to the nearest meter) were recorded separately for each species. Five percentage cover classes were recognized on the basis of canopy cover: very scattered (1%–5%), scattered (5%–25%), medium (25%–50%), dense (50%–75%), and closed (>75%). Up to 13 alien species were recorded in a single stand. This survey produced probably the most detailed data set on the distribution of invasive plants at the regional scale for any part of South Africa (and probably the world).

The following environmental factors were available on a Geographic Information System (GIS): geology, vegetation types, land use, topography, and a range of climate variables (table 1). Vegetation types and land use (scale 1 : 10,000) were mapped during the field survey. Thirty-six vegetation types, regrouped into 12 vegetation groups, were identified. These vegetation types were also used to derive 13 geological types based on the predominant geology type known to be associated with each vegetation type (R. M. Cowling, unpublished data). The scale at which the environmental variables were available determined the modeling resolution and, to some extent, our ability to model species distribution. Climatic layers were only available at a coarse scale, but because of the relatively low topographical heterogeneity, this was not considered a potential problem.

#### Building Predictive Models of the Distribution and Percentage Cover of Alien Invasive Species

Two models were used to quantify the role of propagule pressure and environment in explaining the distribution and canopy cover for three of the most widespread invasive species (*A. saligna*, *A. cyclops*, and *P. pinaster*). First, we used a correlative model (classification and regression trees) for predicting presence/absence and percentage cover. The ability of such a model to predict invasive species distribution has been demonstrated at larger spatial scales in South Africa (Rouget et al. 2002, 2003). Second, we used a semimechanistic model for predicting species cover. This new approach simulates, albeit rather simplistically, propagule pressure and incorporates species-environment interactions. We investigated the model's ability to accurately predict canopy cover. This also provided insights on how each species perceived the natural environment.

#### Correlative Model

##### Sampling Design

The study area was partitioned into cells of 250 m × 250 m (i.e., minimum mapping unit of 6.25 ha) using the GIS program ArcInfo. A training data set was created by randomly selecting 10,000 cells. For each observation (i.e., cell), the values of the environmental variables were known, as were the response variables (distribution and percentage cover category of each invasive species).

##### Statistical Analysis

All the observations of the training data set were entered into classification and regression tree analysis using S-Plus (Venables and Ripley 1999). These techniques have proved to be successful in similar studies that attempted to predict the distribution of plant species (Iverson and Prasad 1998; De'ath and Fabricius 2000). Recursive partitioning tech-

**Table 1:** Environmental variables used for modeling species distribution and cover

Factors	Mean	<i>Acacia cyclops</i>	<i>Acacia saligna</i>	<i>Pinus pinaster</i>
Altitude (m)	105	3	2	1
Land use <sup>a</sup>	7	1	1	
Proximity to cultivated fields (m)	516	1	2	2
Mean annual precipitation (mm)	430		1	1
Annual minimum temperature (°C)	11.1		1	2
Growth days	135.5	1		1
Geology <sup>a</sup>	13	1		1
Vegetation groups <sup>a</sup>	12	2	2	2

Note: The number of times each factor was used for modeling species distribution using classification trees is indicated.

<sup>a</sup> The number of categories is indicated for each species.

**Table 2:** Definition and characteristics of invasion foci (i.e., sites where invasion presumably started)

Type	Height (m)	% cover	Number of sites	Area invaded (ha)
<i>Acacia cyclops</i> :				
1	≥8	>75	12	200 (.17%)
2	6–8	>75	54	2,843 (2.5%)
<i>Acacia saligna</i> :				
1	≥8	>50	54	1,227 (1.4%)
<i>Pinus pinaster</i> :				
1	≥10	>75	76	473 (.6%)
2	≥10	50–75	46	923 (1.2%)

Note: See text for description of type 1 and type 2 foci. Data shown in the table relate only to the area occupied by these foci. For each species, the percentage of the total area invaded that is occupied by the foci is indicated in parentheses.

niques make no assumptions of linearity and accommodate both categorical and continuous data. They can also reveal complex interactions among variables (for a detailed discussion of the properties of classification and regression trees, see Breinam et al. 1984).

The factors identified in the regression tree analysis were used to derive potential distribution maps. The significance of each factor was evaluated by cross-validation (Breinam et al. 1984; De'ath and Fabricius 2000). Environmental conditions identified as suitable for the establishment of each species were selected in ArcView to produce habitat suitability maps of species distribution. The predictive accuracy of the model was tested using Kappa statistics on the full data set ( $n = 33,000$ ).

Such correlative models assume that species have had

sufficient time to sample all potentially suitable habitats and have reached pseudoequilibrium. They therefore ignore the fact that the availability of propagules almost certainly limits invasion. For invasive aliens, it is unlikely that all habitats have been adequately sampled, so we developed an additional approach in which the role of propagule pressure is crudely modeled.

### Semimechanistic Model

#### *Conceptual Framework*

This new approach models species canopy cover at the landscape scale (25 × 25-m cell resolution). Canopy cover for each species was modeled as a function of propagule pres-

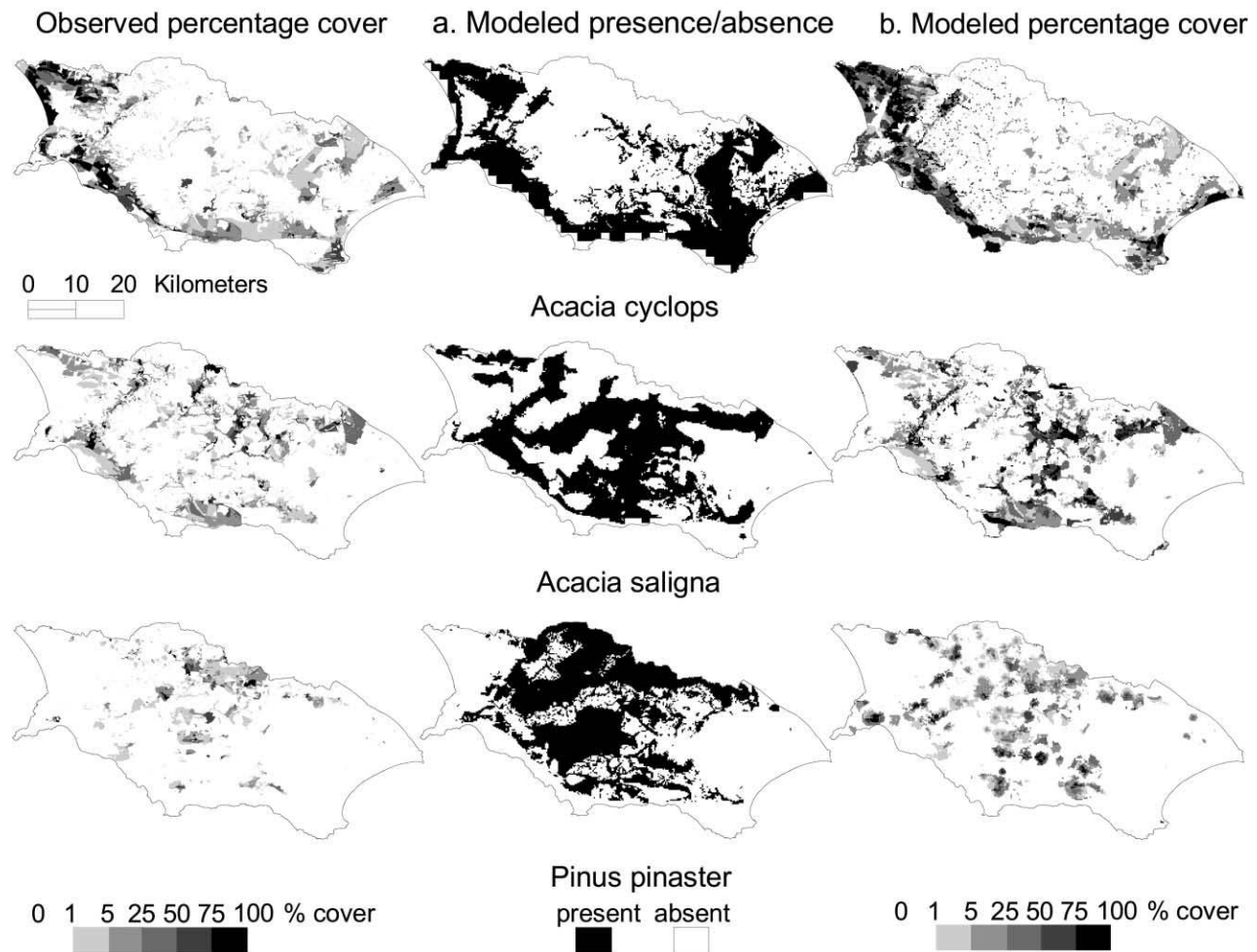
**Table 3:** Predictive accuracy of modeling invasive species distribution and cover using a correlative model (based on environmental correlates) and a homogenous spread model (based on propagule pressure)

Accuracy	<i>Acacia cyclops</i>	<i>Acacia saligna</i>	<i>Pinus pinaster</i>
Correlative model:			
Presence/absence	77.2	71.9	62.1
Kappa <sup>a</sup>	.5	.4	.2
Cover only	35.1	41.0	29.2
Total	12.2	12.2	60.8
Kappa <sup>b</sup>	0	0	.1
Homogenous spread model:			
Presence/absence	41.3	35.8	62.9
Cover only	32.4	33.6	32.1
Total	24.3	23.3	55.8
Kappa <sup>b</sup>	.1	.1	.2

Note: Percentage of correctly classified cells is indicated for each category. Cover only refers to percentage cover accuracy for the five categories of percentage cover. Total refers to all correctly classified cells (absence and five canopy cover categories). Kappa values range from 0 (random process) to 1 (deterministic process). For each species, the gain in predictive accuracy between the correlative and the homogenous spread was significantly different (McNemar's test,  $P < .001$ ).

<sup>a</sup> Refers to presence/absence model only.

<sup>b</sup> Refers to percentage cover model.



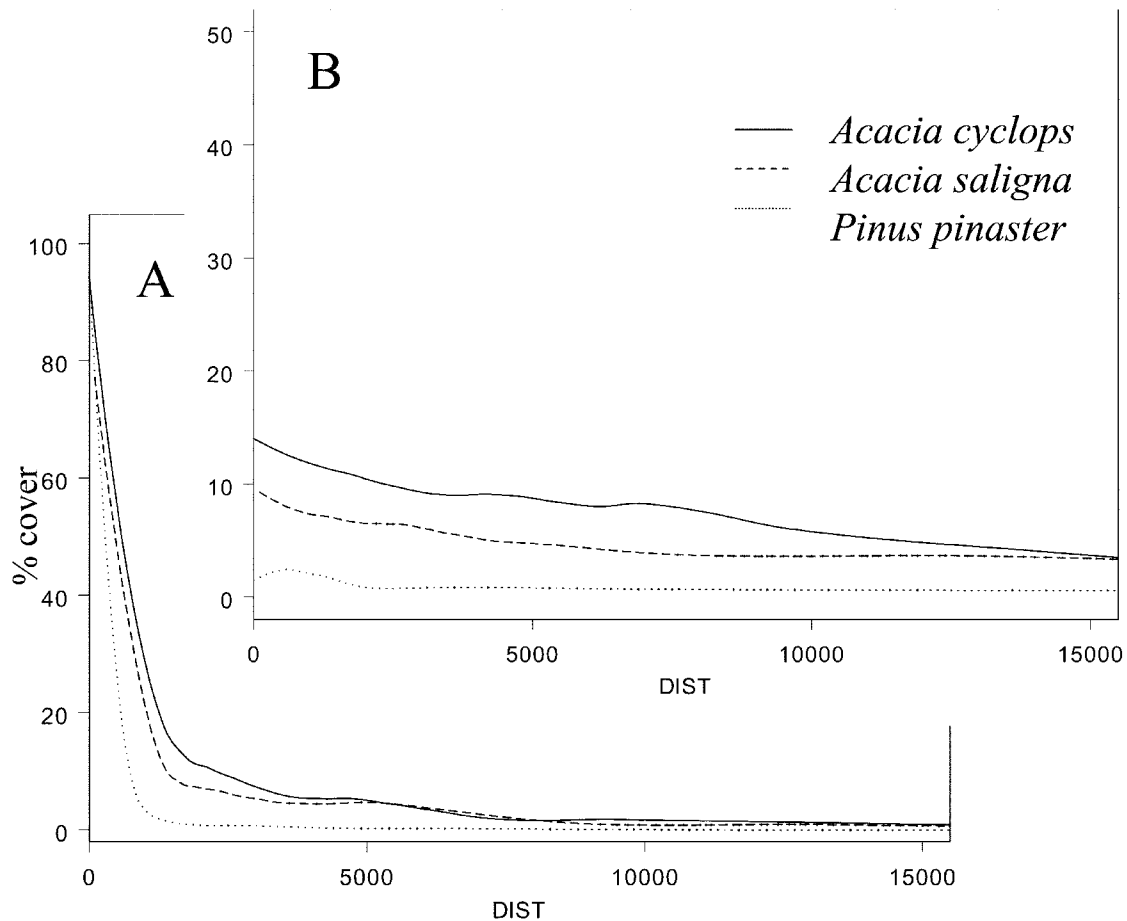
**Figure 1:** Observed and modeled distribution maps for the three most widespread invasive species on the Agulhas Plain. Models were based on (a) classification trees (based on environmental correlates) and (b) semimechanistic models (based on propagule pressure and environment).

sure (distance to the closest invasion focus) and environmental suitability (fig. A1 in the online edition of the *American Naturalist*). We aimed to tease out the effect of the two factors by using a two-stage model. In the first stage, canopy cover is modeled as a homogenous spread from the invasion focus. This stage considers the effect of propagule pressure (approximated by the distance from the invasion focus) as the only driver of invasion. We call this the homogenous spread model; all invader-environment interactions are ignored. In the next stage, the influence of environmental factors is assessed on the residuals of the first model to determine how propagule pressure interacts with environmental factors. The homogenous spread model is thus adjusted by introducing the mediating (facilitating or limiting) effects of the environment (fig. A1). We call the complete model semimechanistic because it integrates biological processes (propagule pressure) but captures invader-environment

interactions in a static manner using classification trees.

#### Identification of Invasion Foci

For each species, probable source populations (invasion foci) were identified on the basis of stand density and height (as a surrogate for stand age). Centroids of tall and dense clumps were assumed to be the foci of invasive populations. Population growth for each species is driven primarily by fires that increase local density and stimulate stratified diffusion (Richardson et al. 1992). Invasion foci thus remain evident in the landscape. We distinguished two types of sites of initial introduction on the basis of height and percentage cover (table 2). Type 1 refers to the tallest and densest category; type 2 refers to the second



**Figure 2:** Local regression models for the spread of *Acacia cyclops*, *Acacia saligna*, and *Pinus pinaster* on the Agulhas Plain. *A*, Percentage cover of each species was regressed against the distance to invasion foci (m). *B*, Predicted cover of each species (based on regression trees; see “Methods”) was also regressed against the distance to invasion foci (m).

tallest or densest category. For each species, invasion foci represent <3% of the total invaded area (table 2).

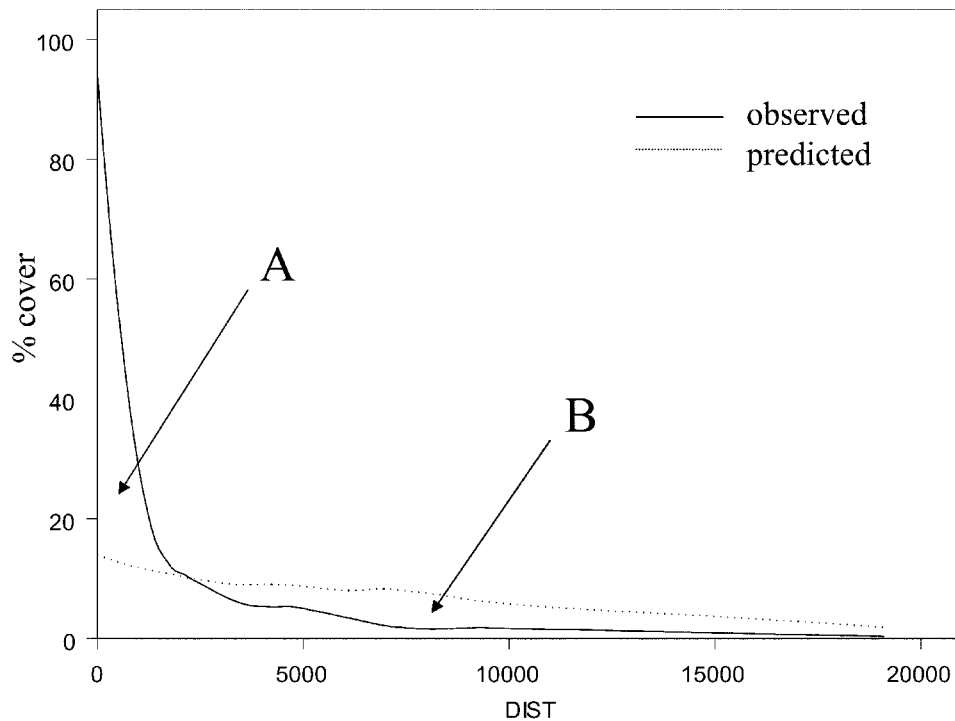
#### *Stage 1: Fitting the Homogenous Spread Model*

Locally weighted regression smoothing was applied to model the percentage cover of invasive species on the basis of the distance to the nearest site of initial introduction (fig. A1, stage 1). Local regression is a nonparametric regression technique that relies on the data to specify the form of the model. A model is fitted to the data points locally so that at any point the model depends only on the observations at that point and some specified neighboring points (Chambers and Hastie 1991). This approach was deemed appropriate for this study because of the nature (semiquantitative cover) and the nonlinear relationship of the data. Local regression has the advantage of

modeling interaction between variables, which generalized additive models cannot do (Chambers and Hastie 1991).

To fit the local regression model of each species, we used a stratified random sample of 10,000 random points (1,000 points for each cover category and 5,000 points where the species does not occur). For each of these 10,000 points, the Euclidean distance from the nearest invasion foci was derived on the basis of a 25 × 25-m grid. We modeled cover as a conditional function of two distance variables (table 2) to consider the interactive effect of neighboring sites on propagule pressure.

The fitted curve was used as a measure of the intensity of propagule pressure. This first stage generated a homogeneous spread model for the entire area, where environmental factors are assumed to have no effect (i.e., invasion is driven exclusively by propagule pressure). However, this model did not capture the full effect of propagule pressure for two reasons: not all propagule



**Figure 3:** Role of propagule pressure in determining percentage canopy cover of *Acacia cyclops*. The dotted line shows the predicted percentage cover based on environmental factors, and the solid line shows the predicted percentage cover based on distance to invasion foci (m; used as a surrogate for propagule pressure). *A*, Substantial effect of propagule pressure where canopy cover is higher than would be predicted by environmental factors. *B*, Smaller effect of propagule pressure far from invasion foci.

sources were considered, and the model was not fitted using all available propagules at a site but only those that were successful in establishment.

#### *Testing Whether Model Outputs Equate to Propagule Pressure*

We tested the ability of our approach to capture the mass effect due to concentrated sources of propagules at the landscape scale. We assumed that the intensity of propagule pressure was related to the distance from invasion foci. But, could this not simply be a spurious measure of habitat suitability? We regressed the predicted canopy cover obtained from the correlative approach against distance to invasion foci using local regression as above. Since these predictions were based on environmental factors, they should be independent of the effects of propagule pressure. We expect the positive effect of propagule pressure to facilitate invasion spread near invasion foci (i.e., canopy cover higher than would be predicted by environmental factors). Similarly, we expect the declining influence of propagule pressure to limit invasion spread far

from invasion foci (i.e., canopy cover lower than would be predicted by environmental factors).

#### *Stage 2: Modeling Interactions between Propagule Pressure and Environment*

To determine how environmental factors interacted with propagule pressure and to determine the percentage cover of invading plants, we analyzed the relationship between residuals of the homogenous spread model and environmental factors (fig. A1, stage 2). We aimed to explore situations where a homogeneous spread model fails to accurately predict species cover. The analysis of residuals therefore focused on the misclassified cells (i.e., where the homogeneous spread model is inadequate) with the assumption that environmental heterogeneity would affect spread (positively or negatively). Regression trees were used to identify environmental conditions where residuals are positive (species canopy cover is lower than expected by homogeneous spread model) or negative (canopy cover is higher than expected by homogeneous spread model). The predicted residual values were used to adjust the original spread model for the misclassified cells (see fig. A1).

**Table 4:** Variation in spread rate explained by environmental factors

Factor	Positive effect	Negative effect
<i>Acacia cyclops:</i>		
Geology	Colluvium, wetlands	Bokkeveld shale, quaternary sand, Table Mountain, granite, sandstone
Land use	Natural areas	Cultivated areas
Altitude (m)	<72	>72
<i>Acacia saligna:</i>		
Altitude (m)	<68	>68
Geology	Alluvium, Bokkeveld shale, Quaternary sand, wetland	Bokkeveld shale with ferricrete, Table Mountain, granite, sandstone
Land use	Natural areas	Cultivated areas
<i>Pinus pinaster:</i>		
Growth days (yr <sup>-1</sup> )	>119	<119
Distance to roads (m)	<1,250	>1,250
Mean annual precipitation (mm)	>392 and <483	<392 or >483

Note: For each species, residuals of the spread model (based on propagule pressure) were regressed against environmental factors (one factor at a time) using regression trees. The three most significant factors are presented. Positive effect means that the spread rate was higher than expected based on the homogenous spread model. Negative effect means that the spread rate was lower than expected.

The analysis was performed in a stepwise manner to generate predicted distribution and percentage cover for each species. The first adjustment to the original homogenous spread model generated new residuals, which were again analyzed with regression trees in search of environmental correlation. A maximum of three adjustments were performed for each species. Model performance was evaluated on the full data set ( $n = 33,000$ ) using the Kappa index. This index compares the agreement against that which might be expected by chance (Fielding and Bell 1997; Pontius 2000). Kappa can be thought of as the chance-corrected proportional agreement, and possible values range from +1 (perfect agreement) via 0 (no agreement above that expected by chance) to -1 (complete disagreement). For each species, we used McNemar's test to establish whether the accuracy gain between two models was significant (Fielding and Bell 1997).

## Results

### *Correlative Models to Predict Distribution and Percentage Cover of Invasive Alien Species*

Classification trees generated models of species distribution (presence or absence) with intermediate predictive power. Total accuracy was around 70% (table 3). Model accuracy and species range appear to be positively correlated. The most widespread species (*Acacia cyclops*) was the best-modeled species (Kappa coefficient of 0.5).

Of all environmental factors considered, vegetation groups were the best predictors of the distribution of invasive species for the three species considered. Next-best factors for explaining species distribution were climate (for *Pinus pinaster*), land use (*Acacia saligna*), and altitude (*A.*

*cyclops*). Potential distribution maps (based on the environmental profile of each species) were compared with the observed distribution for the three species in figure 1. According to these environmental profiles, 29% of the Agulhas Plain could still be invaded by *P. pinaster* (i.e., the area is suitable), 20% by *A. saligna*, and 14% by *A. cyclops*. However, between 12% and 31% of the area currently invaded by each species occurred outside the potential distribution predicted by the model.

The correlative approach failed to adequately model canopy cover (table 3). Cover was very poorly related to environmental heterogeneity. Model accuracy was in all cases lower than 20% except for *P. pinaster*. Regression trees could not identify specific environmental conditions for each cover class. Each model failed to accurately predict the distribution of dense stands (>75% cover). Regression trees were generally better at predicting low percentage cover values (percentage cover between 1% and 25%). Sites where invasive species are absent could not be identified using regression trees for both *Acacia* species, which resulted in Kappa statistics of 0 (table 3).

### *Semimechanistic Models of Invasive Alien Plant Spread*

*Testing Whether Model Outputs Equate to Propagule Pressure.* Preliminary analysis using regression trees indicated that distance to invasion foci was the best predictor of percentage cover and outperformed environmental factors. The modeled percentage cover for each species as a function of the distance to invasion foci is shown in figure 2A. Predicted percentage cover based on environmental factors (see correlative approach above) is shown in figure 2B. Although the latter decreased as distance to invasion foci



**Table 5:** Classification accuracy of the semimechanistic model for all three invasive species

Accuracy	Loess	Factor		
		First	Second	Third
<i>Acacia cyclops:</i>				
Total	24.3	63.9	75.8	81.1
Presence/absence	41.3	73.4	80.9	88.0
Percentage cover	32.4	49.4	57.5	70.6
Kappa	.1	.4	.6	.7
<i>Acacia saligna:</i>				
Total	23.4	59.5	80.4	85.0
Presence/absence	35.8	66.3	83.8	90.0
Percentage cover	33.7	46.9	60.3	70.8
Kappa	.1	.3	.6	.7
<i>Pinus pinaster:</i>				
Total	55.8	60.5	81.6	88.4
Presence/absence	62.9	67.8	84.5	90.5
Percentage cover	32.1	42.6	61.9	79.0
Kappa	.2	.3	.5	.7

Note: Species distribution and percentage cover is first modeled based on the homogenous spread model (local regression model: loess) and then adjusted by incorporating interactions between spread rate and environment (first to third factors). This is the validation on the full data set ( $n = 33,000$ ) of the spread model derived from the training data set ( $n = 10,000$ ). Kappa values range from 0 (random process) to 1 (deterministic process). For each species, the gain in predictive accuracy between two successive models was significantly different (McNemar's test,  $P < .001$ ).

increased, environmental factors failed to model high percentage cover near invasion foci (fig. 2B). The role of propagule pressure is highlighted in figure 3. Near invasion foci, percentage cover was higher than would be predicted from environmental factors (fig. 3A). After 2,200 m, percentage cover was lower than would be predicted from environmental factors (fig. 3B). This illustrates the mediating effect of propagule pressure (declining with distance from invasion foci). This also validates our assumption that the effect of propagule pressure is effectively modeled as the distance from invasion foci.

*Modeling Invasion as Homogenous Spread.* According to our homogenous spread model, the two *Acacia* species have the potential to spread much further than *P. pinaster* (fig. 2A). A predicted cover of 50% was reached at 300 m from the sites of initial introduction for *P. pinaster*, 450 m for *A. saligna*, and 600 m for *A. cyclops*. This model predicted a cover of 5% for *A. cyclops* at 5,000 m from any sites of initial introduction (fig. 2A).

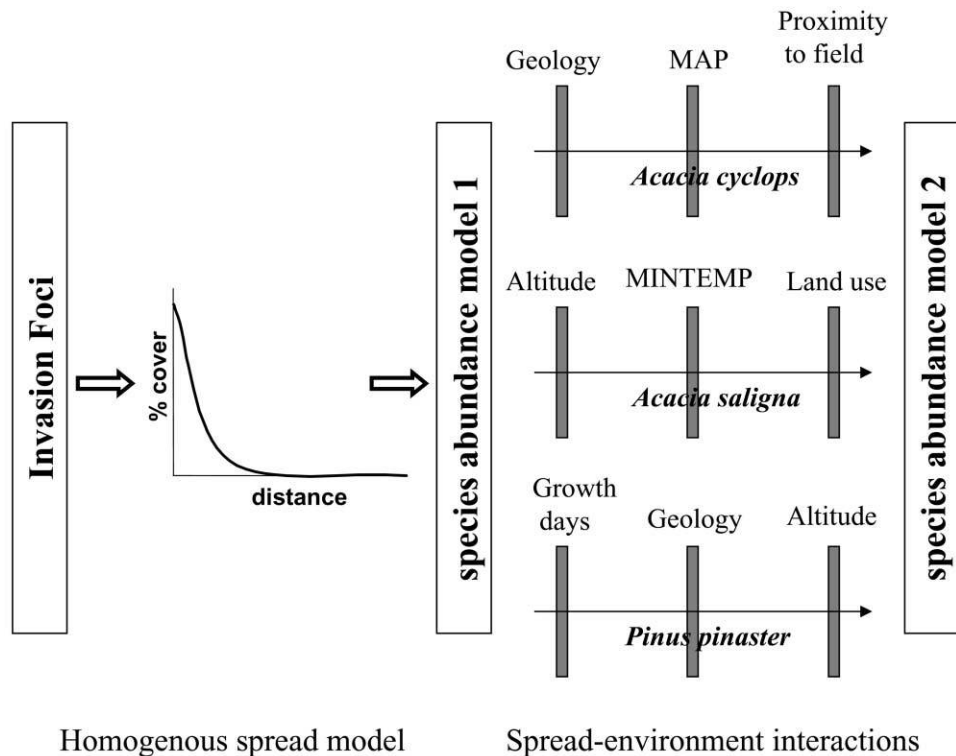
For *A. cyclops* and *P. pinaster*, the model could be improved by using a conditional function (table 2). The homogenous spread model, using distance to invasion foci as the unique predictor, produced a reasonable fit of the species distribution and some indications of the canopy

cover (table 3). The best fit was obtained for *P. pinaster*, for which the accuracy of the spread model was 63% for predicting distribution and 32% for canopy cover. Surprisingly, the level of accuracy of the spread model for predicting distribution of *P. pinaster* was similar to the classification tree based on environmental factors (see table 3). For both *Acacia* species, the total accuracy of the homogenous spread model was higher than that of the correlative model. Spread models for *Acacia* species produced low fit of species cover (table 3) and generally failed to accurately model species absence (both species are predicted to occur at low density [between 1% and 5% cover throughout most of the Agulhas Plain]).

*Modeling Invasion as a Heterogeneous Process (Incorporating Environmental Factors).* The homogenous spread models partly explained the spatial patterns of distribution and percentage cover, and a substantial amount of the residual variance remained unexplained. The analysis of the residual variance in relation to environmental variables revealed strong interactions between the environment and the spread of the three invasive species. Table 4 shows the three most significant factors explaining variation in residuals of spread model. Each species reacted differently to environmental heterogeneity. For the two *Acacia* species, the residual variance of species cover was correlated with altitude, geology, and land use. For example, percentage cover was higher than predicted (from the homogenous spread model) at low altitudes (<75 m for *A. cyclops* and <68 m for *A. saligna*) and in natural vegetation.

Stepwise analysis on the residuals of the homogenous spread model considerably improved the prediction of species cover. Table 5 shows the increase in predictive power for modeling species percentage cover. Predictive accuracy obtained from the homogenous spread model by adjusting one factor was in all cases higher than those obtained from the correlative approach (see table 3). Semimechanistic models with three factors resulted in high predictive accuracy (Kappa statistics >0.7).

Each of these factors can be viewed as an environmental barrier or facilitator (fig. 4). The factors identified are species specific, and no hierarchy was found among environmental factors for the three species studied. Geology, altitude, and the annual number of growth days were found to have the strongest mediating effect on the spread of *A. cyclops*, *A. saligna*, and *P. pinaster*, respectively. Potential distribution maps of the three species are shown in figure 1B. These are based on the homogenous spread models adjusted for the environmental factors identified on the basis of classification trees on the residual variance. The percentage cover of each invasive species was predicted with greater accuracy than was achieved in simple correlative models (table 5).



**Figure 4:** Semimechanistic model for *Acacia cyclops*, *Acacia saligna*, and *Pinus pinaster*. Homogenous spread models were first derived to predict species abundance (percentage cover) based on propagule pressure only (model 1); environmental effects were then considered on the residual variance to model spread-environment interactions (model 2). Each environment factor acts as a successive barrier or facilitator for spread (in decreasing order of importance from left to right). MAP = mean annual precipitation; MINTEMP = annual minimum temperature.

### Discussion

The results demonstrate the value of natural experiments for elucidating the mechanisms of alien tree invasions. Spatial patterns of invasion result from many different interacting factors, including biological attributes of the invader, species response to the abiotic environment, biotic interactions, and human activities (success of plant establishment; Richardson et al. 2004). In most cases, it is difficult to determine the relative roles of biological attributes and environmental factors. This study attempted to include biological processes such as propagule pressure and plant-environment interactions. Previous studies that have focused on local seed dispersal using spatial patterns of juveniles did not always consider the effects of environment (Schupp and Fuentes 1995). However, statistical analyses of plant distribution in relation to environmental factors generally ignore ecological processes such as dispersal (Guisan and Zimmermann 2000). Including propagule pressure and environmental determinants provided more accurate predictions of species cover than when environmental factors were solely considered.

Our modeling approach has some limitations. Propagule pressure was simply modeled in relation to a limited number of invasion foci; the correct identification of these foci is thus crucial. We assumed that the oldest (largest and densest) stands of trees were the origin of self-sown populations (table 2). We ignored the role of individual trees as invasion foci. Also, past alien plant management and wood harvesting have removed some source populations. Since substantial clearing operations have only affected a small part of the large study area, the resultant error due to this factor is probably very small. Finally, our surrogates for propagule pressure (distance to invasion foci) and environmental suitability were not completely independent. The former could not be measured where invaders did not establish because of limited environmental suitability, and the latter could not be measured where invaders did not establish because of lack of propagule. These factors contribute to some unavoidable and unexplained variance in the models. Despite these limitations, this study brings a deeper understanding of the main drivers of invasion process at a regional scale and provides the

first quantitative approximation of the role of propagule pressure in mediating environmental barriers for an invasive species at the regional scale.

Species distribution was modeled with satisfactory accuracy using classification trees (table 3). However, distance to sites of initial introduction—a surrogate of propagule pressure—was a much better indicator of species distribution and cover than any environmental factor considered. Few previous studies have assessed the relative importance of ecological interactions and propagule availability in producing observed patterns of species distribution at the local scale (Foster 2001). Propagule pressure is clearly crucial for accurate modeling of distribution and cover of invasive alien trees on the Agulhas Plain. We noted differences among species, and the effect of propagule pressure was more significant for *Pinus pinaster* than for the two *Acacia* species (table 3). This is difficult to relate to the natural dispersal regimes of the species (Richardson et al. 1992). It is possibly due to the fact that both *Acacia* species have hard-coated seeds that are readily dispersed in soil.

Several authors have suggested visualizing the invasion process as a series of barriers that an introduced species needs to overcome before becoming naturalized or invasive (Richardson et al. 2000b). We have provided a methodology for identifying some of these barriers and quantifying their effects on spread at a regional scale. Environmental barriers affecting patterns of spread (i.e., once propagule pressure has been accounted for) are species specific. Barriers could be geology, topography, climate, or land use without any consistent hierarchy among the three species (fig. 4). Surprisingly, land use was not identified as the major barrier to invasion. There were some indications that spread was reduced in transformed habitats (notably cultivated fields), but climatic or topographic factors were always more significant (table 3). The Agulhas Plain is, however, heavily fragmented, so one would expect invasive species to respond strongly to land use changes. Other studies that focused on local dispersal have shown that landscape structure can affect dispersal success (Higgins and Richardson 1999; King and With 2002). Modeling studies also suggest that invasive species might perform better than natives in fragmented ecosystems (Richardson et al. 2000a). Disturbance, natural or human induced, is often mentioned as a factor promoting invasion because it exposes bare ground, permitting seedling establishment (Hobbs and Huenneke 1992; Kolar and Lodge 2001). At a local scale, the spatial configuration of bare ground (e.g., large or small gaps, randomly or regularly dispersed gaps) greatly affects the spread of invasive species (Bergelson et al. 1993). Predictions at the landscape or regional scale, such as this study, could be improved by quantifying more rigorously aspects of disturbance re-

gime and intensity and the spatial structure of landscape, factors that are generally not accounted for in such studies.

#### Acknowledgments

D. Euston-Brown undertook the mammoth task of mapping the distribution of invasive trees and shrubs on the Agulhas Plain. We acknowledge financial support of the Global Environmental Facility—through the World Wide Fund for Nature South Africa as part of the Cape Action Plan for the Environment—and the University of Cape Town's research committee. Thanks are due to R. Cowling, S. Higgins, H. Mooney, P. Pyšek, M. Rejmánek, M. Williamson, and two anonymous referees for their comments on earlier drafts.

#### Literature Cited

- Bergelson, J., J. A. Newman, and E. M. Floresroux. 1993. Rates of weed spread in spatially heterogeneous environments. *Ecology* 74:999–1011.
- Breinan, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. Classification and regression trees. Wadsworth, Belmont, Calif.
- Chambers, J. M., and T. J. Hastie. 1991. Statistical models. Wadsworth & Brooks/Cole, Pacific Grove, Calif.
- Cowling, R. M., B. M. Campbell, P. J. Mustart, D. J. McDonald, M. L. Jarman, and E. J. Moll. 1988. Vegetation classification in a floristically complex area: the Agulhas Plain. *South African Journal of Botany* 54:290–300.
- D'Antonio, C. M., J. Levine, and M. Thomsen. 2001. Ecosystem resistance and the role of propagule supply: a California perspective. *Journal of Mediterranean Ecology* 2:233–245.
- De'ath, G., and K. E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81:3178–3192.
- Fielding, A. H., and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24:38–49.
- Foster, B. L. 2001. Constraints on colonization and species richness along a grassland productivity gradient: the role of propagule availability. *Ecology Letters* 4:530–535.
- Franklin, J. 1995. Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography* 19:474–499.
- Green, R. E. 1997. The influence of numbers released on the outcome of attempts to introduce exotic bird species to New Zealand. *Journal of Animal Ecology* 66:25–35.
- Guisan, A., and N. E. Zimmermann. 2000. Predictive hab-

- itat distribution models in ecology. *Ecological Modelling* 135:147–186.
- Higgins, S. I., and D. M. Richardson. 1996. A review of models of alien plant spread. *Ecological Modelling* 87: 249–265.
- . 1999. Predicting plant migration rates in a changing world: the role of long-distance dispersal. *American Naturalist* 153:464–475.
- Higgins, S. I., D. M. Richardson, R. M. Cowling, and T. H. Trinder-Smith. 1999. Predicting the landscape scale distribution of alien plants and their threats to biodiversity. *Conservation Biology* 13:303–313.
- Higgins, S. I., D. M. Richardson, and R. M. Cowling. 2000. A dynamic spatial model for managing alien plant invasions at the landscape extent. *Ecological Applications* 10:1833–1848.
- . 2001. Validation of a spatial simulation model of a spreading alien plant population. *Journal of Applied Ecology* 38:571–584.
- Hobbs, R. J., and L. F. Huenneke. 1992. Disturbance, diversity, and invasion: implications for conservation. *Conservation Biology* 6:324–337.
- Hutchinson, T. F., and J. L. Vankat. 1997. Invasibility and effects of Amur honeysuckle in southwestern Ohio forests. *Conservation Biology* 11:1117–1124.
- Iverson, L. R., and A. M. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecological Monographs* 68:465–485.
- King, A. W., and K. A. With. 2002. Dispersal success on spatially structured landscapes: when do spatial pattern and dispersal behavior really matter? *Ecological Modelling* 147:23–39.
- Kolar, C. S., and D. M. Lodge. 2001. Progress in invasion biology: predicting invaders. *Trends in Ecology & Evolution* 16:199–204.
- Lombard, A. T., R. M. Cowling, R. L. Pressey, and P. J. Mustart. 1997. Reserve selection in a species-rich and fragmented landscape on the Agulhas Plain, South Africa. *Conservation Biology* 11:1101–1116.
- Lonsdale, W. M. 1999. Global patterns of plant invasion and the concept of invasibility. *Ecology* 80:1522–1536.
- Myers, N., R. A. Mittermeier, C. G. Mittermeier, G. A. B. da Fonseca, and J. Kent. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403:853–858.
- Newsome, A. E., and I. R. Noble. 1986. Ecological and physiological characters of invading species. Pages 1–20 in R. H. Groves and J. J. Burdon, eds. *Ecology of biological invasions: an Australian perspective*. Australian Academy of Science, Canberra.
- Pontius, R. G., Jr. 2000. Quantification error versus location error in comparison of categorical maps. *Photogrammetric Engineering and Remote Sensing* 66: 1011–1016.
- Rejmánek, M., D. M. Richardson, S. I. Higgins, M. J. Pitcairn, and E. Grotkopp. 2004. Ecology of invasive plants: state of the art. In H. A. Mooney, J. A. McNeely, L. Neville, P. J. Schei, and J. Waage, eds. *Invasive alien species: searching for solutions*. Island, Washington, D.C. (in press).
- Richardson, D. M., I. A. W. Macdonald, and R. M. Cowling. 1992. Plants and animal invasions. Pages 271–308 in R. M. Cowling, ed. *The ecology of fynbos: nutrients, fire and biodiversity*. Oxford University Press, Cape Town.
- Richardson, D. M., W. J. Bond, W. R. J. Dean, S. I. Higgins, G. F. Midgley, S. Milton, L. W. Powrie, M. C. Rutherford, M. Samways, and R. E. Schulze. 2000a. Invasive aliens organisms and global change: a South African perspective. Pages 303–349 in H. A. Mooney and R. J. Hobbs, eds. *The impact of global change on alien species*. Island, Washington, D.C.
- Richardson, D. M., P. Pyšek, M. Rejmánek, M. G. Barbour, F. D. Panetta, and C. J. West. 2000b. Naturalization and invasion of alien plants: concepts and definitions. *Diversity and Distributions* 6:93–107.
- Richardson, D. M., M. Rejmánek, and M. Rouget. 2004. Using natural experiments in the study of alien tree invasions: opportunities and limitations. In M. Gordon and S. Bartol, eds. *Experimental approaches to conservation biology*. University of California Press, Berkeley (in press).
- Rouget, M., D. M. Richardson, S. J. M. Milton, and D. Polakow. 2001. Predicting the dynamics of four invasive *Pinus* species in a fragmented semi-arid shrubland in South Africa. *Plant Ecology* 152:79–92.
- Rouget, M., D. M. Richardson, J. A. Nel, and B. W. Van Wilgen. 2002. Commercially important trees as invasive aliens: towards spatially explicit risk assessment at a national scale. *Biological Invasions* 4:397–412.
- Rouget, M., D. M. Richardson, R. M. Cowling, J. W. Lloyd, and A. T. Lombard. 2003. Current patterns of habitat transformation and future threats to biodiversity in terrestrial ecosystems of the Cape Floristic Region, South Africa. *Biological Conservation* 112:63–85.
- Schupp, E. W., and M. Fuentes. 1995. Spatial patterns of seed dispersal and the unification of plant population ecology. *Ecoscience* 2:267–275.
- Venables, W. M., and B. D. Ripley. 1999. *Modern applied statistics with S-Plus*. Springer, New York.
- Williamson, M. H. 1996. *Biological invasions: population and community biology series*. Chapman & Hall, London.