

Spatial segregation of congeneric invaders in central Pennsylvania, USA

Michael R. Allen^{1,2} & Katriona Shea^{1,*}

¹Department of Biology, The Pennsylvania State University, 208 Mueller Laboratory, University Park, PA 16802, USA; ²Program in Ecology and Evolutionary Biology, The University of Illinois at Urbana-Champaign, 515 Morrill Hall, 505 S. Goodwin Ave., Urbana, IL 61801, USA; *Author for correspondence (e-mail: k-shea@psu.edu; fax: +1-814-865-9131)

Received 9 May 2004; accepted in revised form 23 August 2004

Key words: autologistic model, biological monitoring, *Carduus*, spatial distribution, spatial scale

Abstract

Carduus acanthoides and *Carduus nutans* (plumeless and musk thistles) are among the most noxious weeds in the United States of America, presenting a serious challenge in cropping and pasture systems. Unfortunately, a lack of detailed spatial distribution information hampers both our ability to understand the factors affecting their invasive success, and the effectiveness of monitoring and management efforts. To examine patterns of distribution and co-occurrence at a local level, we sampled a 5000 km² area of central Pennsylvania that cut a transect across known areas of *C. acanthoides* and *C. nutans* infestation. A number of potential environmental explanatory variables were recorded and analyzed to examine whether they correlated with observed species distribution patterns. Patterns of forest density and spatial aggregation of the thistles were the primary covariates that significantly impacted both species' distributions. The survey established that the frequency of sightings for each species diminished as the ranges converged, with only brief overlap: the two species are strongly negatively correlated in space. Understanding environmental correlates of infestation and the pattern of spatial dissociation of these two invasive species is an important step towards an improved understanding of the mechanisms underlying their invasive potential, and hence towards effective weed control.

Introduction

The exotic invasive plants, *Carduus nutans* L. and *Carduus acanthoides* L. (musk and plumeless thistles, Asteraceae) are both widely distributed in the United States of America (USA). They frequently co-occur within states (NatureServe 2002); for example, both are present in the state of Pennsylvania (PA). However, evidence for niche differences between these species is small, and their ecological similarity raises the question of whether they co-occur at finer spatial scales, and whether there are particular habitats that one or the other is more likely to invade.

Noxious weeds cause economic damage on a grand scale; crop losses and control measures are

estimated to cost approximately \$26 billion annually in the USA alone (Pimentel et al. 2000). *C. nutans* and *C. acanthoides* have been cited as two of the top noxious weeds, appearing on weed lists across the continental USA and southern Canada (Skinner et al. 2000). Much effort has been expended in the search for control or eradication methods for these species, most specifically through the use of herbicides and biological control (see review: Kok 2001). In part, such control efforts are impeded by the lack of information on two main aspects. First, there is little known about the current spatial distribution of the two species at a local scale. This hampers our ability to target management appropriately. For example, more detailed spatial information would

allow us to optimize distribution strategies for biological control agents (Shea and Possingham 2000). Second, we have a relatively poor understanding of what constitutes suitable habitat for these species. This limits our ability to monitor potentially invasible, yet still unaffected, areas for nascent populations. Information on these two aspects, taken together, would allow us to direct monitoring and management efforts more efficiently.

The current state of knowledge on the distribution of these *Carduus* species within Pennsylvania is very limited. Herbarium records offer the only legitimate distributional database for plants in the state, providing information dating to the 19th century for these two species (Rhoads and Klein 1993; A. Rhoads, personal communication; B. Isaac, personal communication). While this often means that the records can provide historical patterns of distribution and spread (e.g. Mihaluk and Pysek 2001), *Carduus* records are too sparse to provide meaningful insights about fine-scale spatial distributions. Furthermore, effort is not evenly distributed across space – some areas have been more intensively sampled than others. Thus, such records can only be used as a starting point to direct survey initiatives.

In this study, we carried out a detailed spatial survey of current *C. acanthoides* and *C. nutans* thistle distribution patterns in central Pennsylvania to assess co-occurrence of the two species in nature, and to identify predictors of species occurrence. A deeper understanding of their spatial distributions and the factors affecting those distributions will both illuminate our understanding of the mechanisms governing their co-occurrence and improve our ability to manage them.

Materials and methods

Species description and invasion history

Carduus nutans and *Carduus acanthoides* are native to Europe and western and northern Asia (Clapham et al. 1962). *C. nutans* is also native to northern Africa (Clapham et al. 1962). Both species have escaped their natural ranges and one or both have invaded the USA, Canada, South

America, southern Africa, New Zealand and Australia (Julien and Griffiths 1998). *C. nutans* and *C. acanthoides* are short-lived monocarpic perennials. After germination, individuals overwinter for one or more years as rosettes. Both species bolt during late spring and start to flower during early to mid-summer. Plants die after flowering, and seeds are wind dispersed. Both species invade pastures, cropland, roadsides, open fields and disturbed areas in patches ranging from the individual to stands of thousands. *C. nutans* was first recorded in North America around 1853 near Harrisburg, PA, in the northeastern USA (Desrochers et al. 1988) and has since spread to 45 of the 48 continental American states (Figure 1) and 9 of 14 Canadian provinces (NatureServe 2002). *C. acanthoides* is currently known to exist in 31 states (Figure 1) and 5 provinces (NatureServe 2002) and was first recorded in New Jersey in 1879 (Desrochers et al. 1988).

Field survey

A field survey was carried out from June to August 2002 by sampling within an area of 5000 km² in central Pennsylvania (bounding coordinates 77.751W, 40.942N; 76.734W, 40.470N; 77.045W, 40.087N; 78.048W, 40.557N; Figure 2). The area spans a section of the northern Appalachian ridge and valley system, including a portion of a densely populated area around Harrisburg, PA, where *C. nutans* was first recorded in North America. The ridge and valley system runs on a southwest to northeast orientation with elevation differences of up to 500 m between peak and valley. Much of the area is covered by cropland, and deciduous and evergreen forests typical to central PA. During the sampling period, both species were flowering, and their purple inflorescences made them conspicuous to distances of at least 100 m from the roadside.

The survey area was divided into 200 blocks, each 5 km × 5 km. A predetermined route was driven to search for *Carduus* presence with a maximum effort of approximately 20 km per block. Thus, if the maximum distance was driven, and we assume a conservative 50 m on each side of the road was surveyed, about 8% of the block total area would be observed. In our

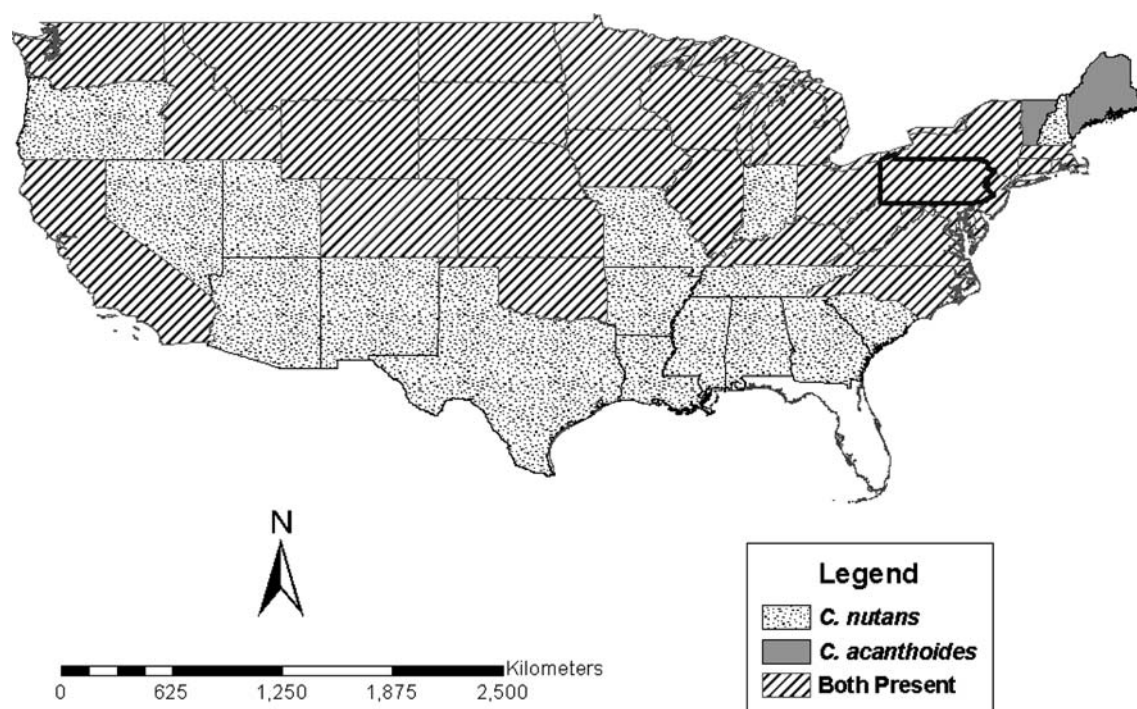


Figure 1. Presence-absence of *Carduus acanthoides* and *Carduus nutans* in the continental USA. The state of Pennsylvania is outlined in bold. State presence data were obtained from the online database NatureServe (2002).

survey, 3439 km of roads were traveled with a median of 18.9 km traveled per block (min 2.1, max 26.5 – see stopping rule below in regards to minimum). Road types included city and county roads, village roads and forest roads; only interstates were excluded from potential survey travel. This road-based survey technique is very appropriate because of the classification of these species as ‘roadside weeds’ and agricultural field pests. The distinctive flowers, and plant height and shape made both species easily identifiable from the roadside throughout the survey period. Rew et al. (2006) found that this type of method resulted in the best presence-absence survey data when species distributions were correlated with traveled routes.

The distinctive flowers, and plant height and shape made both species easily identifiable from the roadside. When either of the species was first observed, that site was surveyed for a variety of environmental and other site specific parameters. Binoculars on site and post-survey review of digital photography of sites confirmed species identifications. Once the driving recommenced, further

observations of the same species within the next 2 km were noted but not surveyed. If both species were recorded within a block, further observations in the block ceased (the ‘stopping rule’). The driving survey was very time consuming and both the 2 km and stopping rules allowed the entire area to be surveyed during the flowering season of the thistles. Thus, the survey of a block was considered complete either if both species had been recorded, or if 20 km had been surveyed. If a species was not recorded within 20 km, it was assumed to be absent from that block (i.e. the survey underestimates occurrence). It was clearly unfeasible to conduct a comprehensive search of the entire area and the controlled survey distance equalised survey effort across blocks. GPS coordinates for individual sightings were recorded using a Garmin eTrex Legend GPS unit (Garmin International, Kansas, USA). The entire survey took 10 weeks during the summer of 2002. During July 2003, a follow-up survey at half of the maximum effort of 2002 (10 km per block) was performed to examine any changes in the region of overlap.

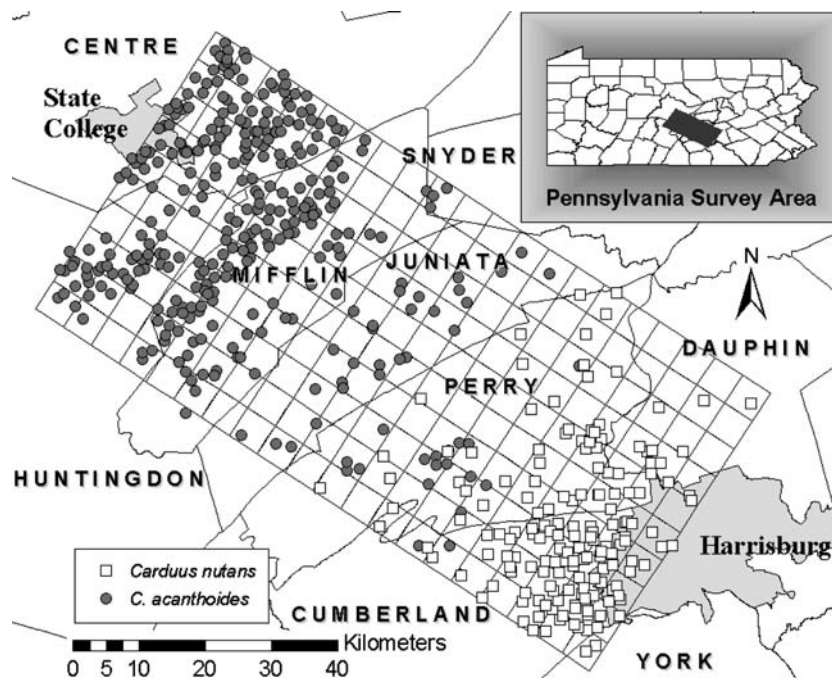


Figure 2. Survey results for *Carduus nutans* and *Carduus acanthoides*. Presence or absence within each 5 km \times 5 km block was calculated from these point data. County names are in uppercase; cities are in title case. The map and data were generated in the North American Datum 1983, Universal Transverse Mercator 18N projection.

Additional spatial data

Individual thistle site locations and all survey data were entered into a geographic database and plotted spatially using ArcGIS 8.1 (ESRI 1999–2001) digital mapping and analysis software. ArcGIS selection tools converted individual survey sites into a presence–absence map and dataset. In addition to presence or absence of each species within the block, the distance surveyed per block was added to the database as a covariate.

Using environmental datasets obtained from the Pennsylvania Geospatial Data Clearinghouse (<http://www.pasda.psu.edu>) and the National Climatic Data Center (<http://www.ncdc.noaa.gov/oa/ncdc.html>), a variety of variables characterizing each survey block were obtained (Table 1). Mean forest density was derived from a 1 km resolution national dataset (US Dept. of Agriculture 1992) using the ArcGIS 8.1 Spatial Analyst extension zonal statistics function. Similarly, vegetative land cover was derived from the Vegetative Land Cover for Pennsylvania map

Table 1. Covariates analyzed in the logistic regression models. Variables with asterisks were significant in the non-spatial analysis.

Category	Description	Variable type
Geologic	Percent cover of limestone geology*	Continuous
	Detrended measure of elevation	Continuous
Habitat	Vegetative land cover	Categorical: water, evergreen forest, mixed forest, deciduous forest, transitional (mixed vegetation), perennial herbaceous, annual herbaceous, barren/hard-surface/etc
Other	Mean forest density*	Continuous
	Distance surveyed	Continuous
	Autologistic variable for <i>Carduus nutans</i>	Continuous
	Autologistic variable for <i>Carduus acanthoides</i>	Continuous

(Myers and Bishop 1999), a 30 km² resolution dataset which breaks cells into 8 basic environmental categories for locations (see Table 1). Anecdotal evidence suggested that limestone bedrock could be an important correlate, so the percent cover of limestone bedrock was estimated from The Areas of Carbonate Lithology dataset (Environmental Resources Research Institute 1996). For time constraint reasons it was not possible to analyze soil from the different areas, but bedrock is a rough proxy for soil type except where soil has been moved (e.g. by road construction). Also, using 1:24000, 30 m resolution National Elevation Datasets (NEDs) from survey area counties (US Geological Survey 1999), the mean block elevation was calculated. A trend line was created from the NEDs by taking the mean elevation of the 10 blocks at each end of the survey area and running a gradient plane between them (trend from ~185 m to 350 m). The mean elevation was then subtracted from this gradient resulting in a detrended elevation measure for each survey block. This procedure effectively removed the natural slope from State College to Harrisburg from the elevation data, allowing the new parameter to account for differences in elevation due to ridges and valleys (ranging from ~150 m to 700 m).

Statistical methods

Species-environment survey data such as those gathered in our survey can be analyzed in a number of ways depending on the question to be asked and the nature of the data collected. Authors often use species abundance and environmental variables with such statistical methods as canonical correspondence analysis (Dieleman et al. 2000) and multivariate regression trees (De'Ath 2002), to model and predict species distribution patterns. Simple presence-absence data, however, can also get to the root of species occurrence patterns. Kunin (1998), for example, used presence-absence patterns of British plants to estimate abundance across multiple spatial scales, comparing negative binomial distribution (He and Gaston 2000) and fractal models. Logistic regression has been a frequently employed analytic technique for presence-absence data modelling (Crawley 2002), ranging from congruence

of field data versus remotely sensed patterns (Aspinall 2002) to predicting species occurrence or density patterns from environmental factors (Collingham et al. 2000; Cowley et al. 2000). Large scale surveys incorporating site-level characteristics and remotely sensed data (such as those undertaken in this study) are well suited to logistic regression techniques. Here, logistic regression (using the block dataset) and chi-square contingency table analysis (with individual survey points) allowed us to isolate variables that had a significant impact on the spatial distribution patterns of these two species.

The spatial arrangement of the species distributions suggested spatial autocorrelation among the data might be an important component of any predictive statistical analysis. Species distributions were spatially clustered, suggesting that presence in one block might increase the probability of presence in adjacent blocks (Smith 1994). To examine this, a spatial autocorrelation analysis was performed using distance intervals of 7.5 km–1.5 times the block width. This resulted in a total of 14 distance classes with variable total point pairs per class. Moran's I correlation coefficients were calculated at each distance and tested for significance with a permutation test (using 10,000 permutations; e.g. Manly 1997). Sequential Bonferroni correction was used with $\alpha = 0.05$ (e.g. Legendre and Legendre 1998). Correlograms were plotted to visualise the range of spatial autocorrelation in the dataset. The distributions of the two species were also compared with a cross-correlogram to examine spatial cross-correlation between them. The cross-correlogram equivalent of the Moran I was used to quantify the association/dissociation between the two species. Analyses were performed using the 'correlog' function in the NCF package (<http://asi23.ent.psu.edu/onb1/software.html>) (Bjørnstad 2002) for R 1.6.2 (<http://www.r-project.org>).

We used the autologistic model (Augustin et al. 1996) to allow for the autocorrelation in further data analyses: neighbourhood presence covariates were created for each of the species in every block as the proportion of the eight (or fewer for edge blocks) neighbouring blocks with the species present. This variable was added as an additional dataset covariate.

We used logistic regression analysis to estimate the predictive ability of our dataset variables with R 1.6.2. The presence–absence response variables were assessed using generalized linear models (GLMs) with binomial errors and a logit link function (Crawley 2002). For each species, backward stepwise selection was performed, starting with the models containing variables analyzed in the non-spatial analysis (Table 1) and the autologistic covariates. The autologistic model was fitted according to Method 2 of Augustin et al. (1996). Each co-variate and all pairwise interaction terms were included in the initial model. The final model was selected on the basis of likelihood ratio backwards elimination at a nominal 5% level. Models were then analyzed visually by comparing the predicted distribution of each species to its actual distribution using ArcGIS 8.1. Additionally, the fits of minimum adequate models were examined with the simple matching coefficient and the kappa statistic (Prentice et al. 1992), both of which compare the assignment of presence or absence by the model to that of the original data. The kappa statistic ranges from no agreement (0) to perfect agreement (1.0) and can be evaluated on a subjective scale (e.g. 0.4–0.55 is ‘fair agreement’, 0.55–0.7 is ‘good agreement’, etc.), as discussed in Prentice et al. (1992). The importance of the individual variables in the final model was assessed using likelihood ratio tests.

Results

While the two *Carduus* species appear to coexist at the national scale (Figure 1), segregation of the two species at the scale of this survey was quite pronounced. Within the survey area, the highest concentration of *C. acanthoides* sightings was closest to State College, Centre County. *C. nutans* sightings were concentrated in the southeast corner of the survey grid near Harrisburg, Dauphin County (Figure 2). Across the 491 individual sites in the survey area, we found 32.0% had *C. nutans* only, 66.6% had *C. acanthoides* only, and 1.4% had both species present. Despite this, the distribution of habitat types (field, forest, etc.) utilized by each species did not differ between species ($\chi^2=2.99$, $df=2$, $P=0.22$). Together, the species

delineate two zones of presence, with a small zone of co-occurrence separating the primary ranges of the two species through Perry County, PA (Figure 2). Overlap between the two species exists within a range limited to 15 of the 200 survey blocks or approximately 375 km² of the 5000 km² area (7.5%) (striped area in Figure 3a). The border remained unchanged in 2003 (data not shown). Extended driving surveys (unpublished results) show further geographic segregation south of the primary survey range, with an overlap zone similar to the surveyed area. Again, *C. nutans* dominates to the east, and *C. acanthoides* dominates to the west.

Given the spatially structured nature of each distribution, spatial autocorrelation was highly significant. Correlogram analyses suggested that both *C. acanthoides* and *C. nutans* showed significant Bonferroni corrected positive spatial autocorrelation to approximately 30 km (Moran’s I, $P < 0.001$, $n = 10,000$; Figure 4). Additionally, the two species showed a significant negative overall cross-correlation (Pearson’s $\rho = -0.33$, $P < 0.001$) and significant negative pair-wise cross-correlations to ~38 km (Moran’s cross-species I, Bonferroni corrected $P < 0.001$, $n = 10,000$, Figure 5).

Autologistic regressions of the dataset indicated that survey distance, forest density, and the autologistic covariate were significant indicators of *C. nutans* presence (Table 2a). Both survey distance and forest density were negatively correlated with presence of *C. nutans*. The negative impact of forest density on presence reflected observations in the field. While we would expect longer distances traveled to result in a higher probability of presence, the reverse was found statistically. Partially contributing to the negative relationship was the stopping rule (i.e. if both species were located in a block). A few low distance blocks reported *C. nutans* presence for this reason. More substantially, the negative correlation was a result of the number of highly surveyed blocks where *C. acanthoides*, but not *C. nutans*, was found. It appears that incomplete effort in blocks (i.e. those blocks with distance less than 20 km) did not result in a shift of the correlation direction, although it likely impacted the magnitude of the end result. The autologistic covariate accounted for the most deviance of any factor

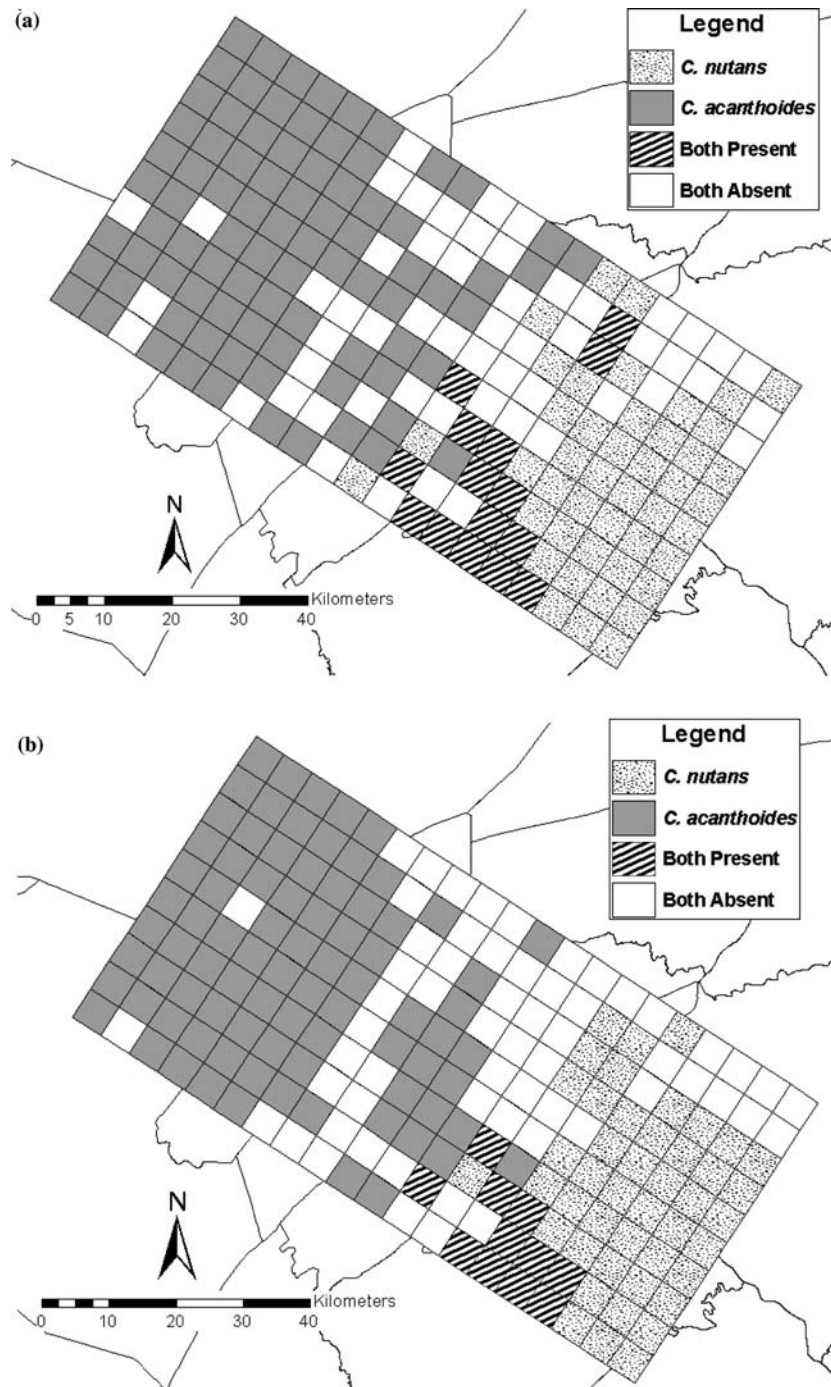


Figure 3. Presence-absence of *Carduus acanthoides* and *Carduus nutans* across the survey range (a) according to field survey (observed) and (b) according to the multiple logistic regression models (estimated).

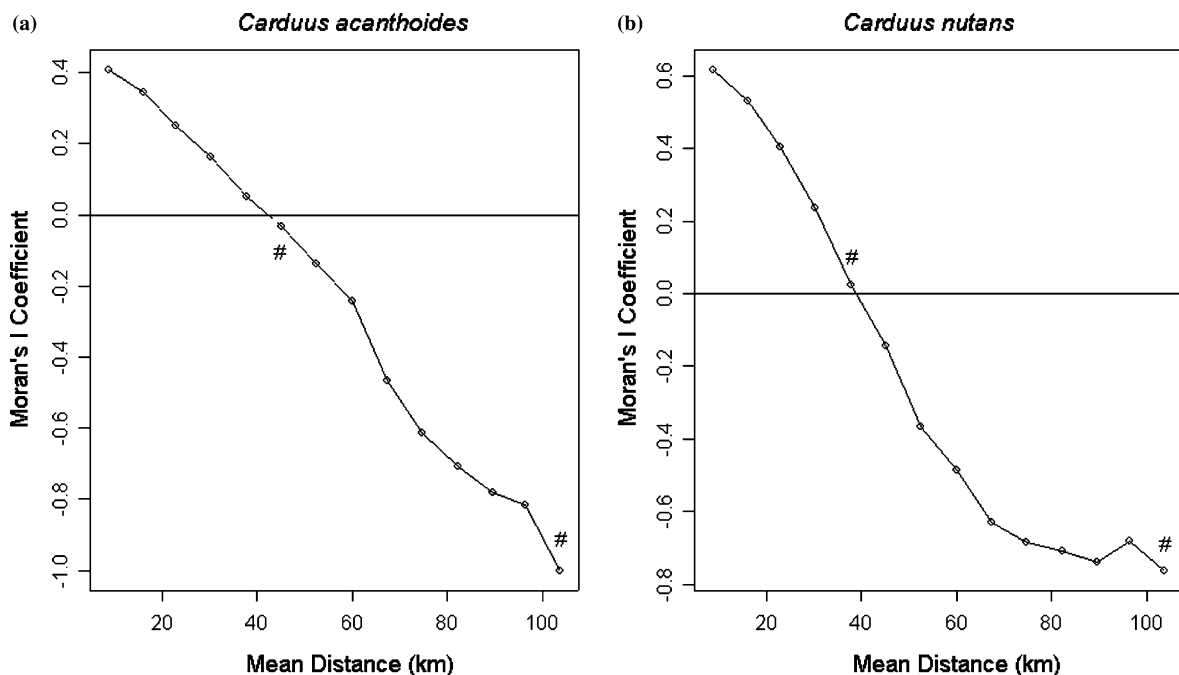


Figure 4. Spatial correlograms of (a) *Carduus acanthoides* and (b) *Carduus nutans*. *C. acanthoides* is positively autocorrelated to 40.6 km with significant autocorrelation to 30.3 km. *C. nutans* is positively autocorrelated to 44.0 km with significant autocorrelation to 37.9 km. All correlation distances are significant at the sequential Bonferroni correction ($P < 0.001$; $n = 10,000$) except for those marked by #.

($F_{1,198} = 265.117$, $P < 0.001$), suggesting that the aggregated nature of the thistle distribution is an important feature at the $5 \text{ km} \times 5 \text{ km}$ spatial scale. The model gave a fairly good fit to the data, accounting for 65% of the null deviance and correctly assigning presence or absence to 90% of the blocks. The model had a kappa statistic of 0.76, which was a 'very good' fit to the data according to the scale cited in Prentice et al. (1992).

The *C. acanthoides* logistic regression (Table 2b) suggested that forest density, survey distance, elevation and the autologistic covariate were important factors impacting presence. The model indicated that the main effect of forest density and its interaction with the autologistic covariate negatively correlated with thistle presence. Survey distance was a negative correlate for the same reasons as in the *C. nutans* model. Increased elevation had a slight positive impact on presence. The autologistic covariate again had the strongest effect of any variable in the model, reinforcing the importance of the aggregated distributions of these species. This model accounted for a modest 41% of the null deviance, but still scored 78.5% of

the blocks correctly. The resulting kappa statistic was 0.57: a 'good' fit of the data. The predictions of both models visually agree well with the observed survey data (Figure 3).

Discussion

Ecological correlates of invasive species' distributions

There are notable studies of species distributions at different spatial scales for native species (e.g. Erickson 1945; Bullock et al. 2000), however, multiple-scale distribution patterns of invasive species tend to be less well documented, even for major pest species. One notable exception lies in the extensive records of the British flora, from which are derived such works as Collingham et al.'s (2000) study on invasive weeds. Few comparable studies exist in the USA (but see, for example, Pauchard et al. 2003), hampering documentation of invasions and control efforts.

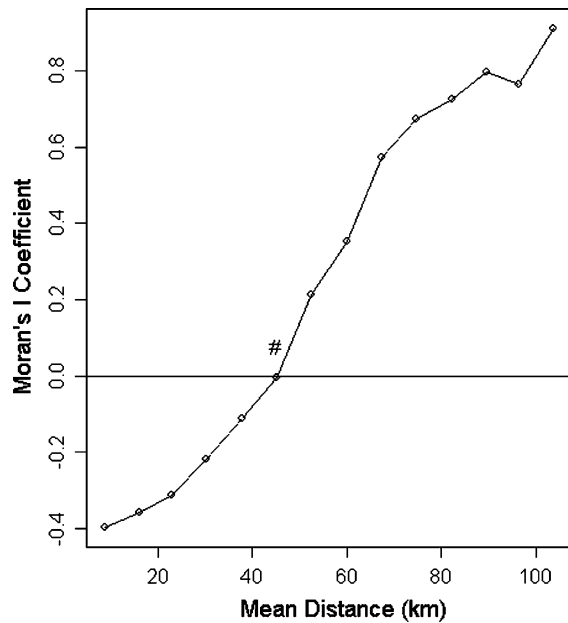


Figure 5. Spatial cross-correlation between *Carduus nutans* and *Carduus acanthoides*. Negative spatial cross-correlation coefficients are significant to 37.9 km. All correlation distances are significant at sequential Bonferroni corrected ($P < 0.001$; $n = 10,000$) except for those marked by #.

Our detailed survey clearly shows that these two congeneric invaders are strongly spatially segregated within Pennsylvania. This is despite the fact that the species are very similar ecologically, generally occupying similar habitats (Desrochers et al. 1988), and have been present in PA since the 19th century (Rhoads and Klein 1993).

Despite assessing several environmental variables, no major differences between species in habitat preference were observed in our study (e.g. habitat type in the field survey), so other hypotheses for the pattern must be considered (see below). Survey observations indicated that both species were very prolific in pastures, grassy fields and roadsides, and sparse in forested areas. It is interesting to note, however, that *C. acanthoides* was sometimes found in forested areas adjacent to gravel or dirt state forest roads in which light penetrated the canopy. Thistles in the forest were not as dense or tall as in open-field patches, but were nonetheless common in forests adjacent to high thistle density areas. *C. nutans* was almost never found in forested areas.

Initial statistical models indicated that elevation was a highly significant indicator of species presence. However, given the sloping incline of the survey area (which ranges from ~ 100 to ~ 700 m) and the distinctly grouped thistle distribution over that area, this result is possibly a trivial artefact of the survey region. Observations by Jessep (1989) of *C. nutans* presence at elevations ranging from sea level to 1200 m in New Zealand support this hypothesis. When we removed the natural slope from the elevation parameters to instead account for the ridge and valley heights, the new parameter was significant in the *C. acanthoides* model (but not the *C. nutans* model) but had a relatively modest effect on the model outcome (Table 2b). This trend reflects the fact that *C. acanthoides* was

Table 2. Summary of the minimum adequate model results for (a) *Carduus nutans* and (b) *Carduus acanthoides*.

Coefficient	Estimate	F-value	Pr(F)	Pr(Z)
(a) <i>C. nutans</i>				
(Intercept)	15.604	–	–	0.017
Mean forest density	–0.310	31.709	<0.001	–
Distance surveyed	–0.941	25.121	<0.001	–
Autologistic covariate	7.372	265.117	$\ll 0.001$	–
Mean forest density \times distance surveyed	0.015	21.529	<0.001	–
(b) <i>C. acanthoides</i>				
(Intercept)	4.828	–	–	0.232
Mean forest density	–0.124	4.478	0.030	–
Distance surveyed	–0.446	7.352	0.007	–
Elevation	0.009	5.244	0.023	–
Autologistic covariate	11.952	33.615	<0.001	–
Mean forest density \times distance surveyed	0.008	8.558	0.004	–
Mean forest density \times autologistic covariate	–0.135	9.960	0.002	–

Pr(F) is the corresponding P value for each analysis of deviance F test, each with 1, 198 degrees of freedom. The significance of the intercepts were assessed with a Wald test with 1 degree of freedom (a) $z = 2.395$ and (b) $z = 1.196$.

present in many blocks containing forested mountain ridges, while it was less prevalent in the relatively flat low lying areas of the southeastern part of the survey area.

Limestone bedrock is widely considered one of the common properties of the soil, where both thistle species are found (Batra 1978; Desrochers et al. 1988). It should be noted that without the autologistic covariate in the models, the mean cover of limestone bedrock acted as a significant indicator of presence for both species. It suggested positive correlation with *C. acanthoides* presence and a negative correlation with *C. nutans* presence. The positive correlation was anticipated, but the negative correlation for *C. nutans* was unexpected.

Possible mechanisms underlying the spatial segregation pattern

There are many factors and mechanisms that may have generated the results we observed (Shea and Chesson 2002). These include environmentally imposed constraints acting on each species, resource and apparent competition between the species, the differential effects of mutualists (e.g. pollinators), as well as priority effects and the possible impact of dispersal limitation. Given the ecological similarity of the species, it may be that this system closely approximates an unstable equilibrium Lotka–Volterra competition scenario and that one or the other species dominates locally depending on initial conditions. Most of these different hypotheses are being examined in ongoing research.

The role of climate in creating the pattern can largely be ruled out. Both *C. nutans* and *C. acanthoides* are present further to the north and south of the surveyed region and their range limits are far wider than observed here (Figure 1). Their Pennsylvania distributions cannot directly be accounted for by their intolerance of climatic aspects such as precipitation and temperature, as both species are found across a wide spectrum of conditions (Desrochers et al. 1988). Clines in climatic conditions, in association with other factors such as competition, may contribute to similar distribution patterns (Bull 1991; Bull and Possingham 1995). However, climatic parameters (including precipitation and temperature) have limited variability, and show no obvious patterns, across

the survey area (data not shown). The general trend in elevation across the region may play a role in such an interaction, but the scale of the present study does not allow this hypothesis to be tested directly. Parameters likely of great importance, but not measured directly, were soil characteristics. Soil core sampling was beyond the logistical scope of this initial survey, as was using national soils databases such as SSURGO or STATSGO; instead mean block cover with limestone bedrock was used as a surrogate. The negative correlation with *C. nutans* presence may in part result from the coarse nature of this measure.

Both species have similar resource requirements, but direct competition for suitable habitat cannot easily explain the observed large scale distribution pattern in the absence of other, larger scale, mechanisms. The two species have been found to co-occur at a few sites (only 1.4% of the sites surveyed), but usually one species has far higher relative abundances at the patch level. Ongoing studies address whether one species will eventually fully dominate such patches, i.e. whether such sites involve temporary co-occurrence, or longer term coexistence.

The receptacle weevil (*Rhinocyllus conicus* (Froelich) (Coleoptera: Curculionidae)), introduced for the control of *C. nutans*, attacks both species (Desrochers et al. 1988) and it is possible that apparent competition impacts the observed distribution. For example, weevil attack may more strongly affect the species that reaches an area later, maintaining the earlier invader. Such priority effects may also arise through Allee effects generated by pollinators if pollen from one species swamps the other. In the same vein, another possible explanation for the distinct species separation is the fact that these two species can interbreed to form hybrids. Warwick et al. (1989; 1992) noted that *C. nutans* and *C. acanthoides* successfully produce F₁ hybrids with varying degrees of morphological intermediacy; however, the hybrids are virtually sterile. A zone of mixed species and sterile hybrid individuals may slow the spread of both species and contribute to the narrow overlap region west of Harrisburg.

A final possibility arises when we ask, ‘Given the first record of *C. nutans* at Harrisburg, PA in 1853, why has this species not spread very far

immediately west of Harrisburg during the last 150 years? Central Pennsylvania has a ridge-valley system that lies in the NE–SW line, and it seems likely that dispersal and establishment along valleys will occur more rapidly than across the ridges between valleys, especially as the ridges are heavily forested. In fact, past the first forested ridge west of Harrisburg, *C. nutans* becomes more sparsely distributed (data not shown). The interplay of elevation and forest density may have created a barrier to seed dispersal and *C. nutans* spread. Thus, the spread and hence the spatial distribution of the two species may be determined by initial conditions and local population densities, combined with dispersal limitation.

Additionally, since *C. acanthoides* seems more tolerant of forested areas, one possible hypothesis to explain the overlap zone is that this species is spreading into the *C. nutans* ‘original’ range. While herbarium records are sketchy, those available suggest that *C. acanthoides* has been present in Centre, Mifflin and Huntingdon counties since at least the early 1980’s (Rhoads and Klein 1993; A. Rhoads, personal communication; B. Isaac, personal communication). However, no previous records of either species exist for Perry or Juniata counties. The survey revealed that *C. acanthoides* is now very prolific in Juniata County and has areas of concentration in Perry County, while *C. nutans* remains concentrated in Dauphin, Cumberland and the eastern part of Perry counties (Figure 2). It is possible that the ranges of these two species have been converging; however, the historic data are not sufficient to analyze these spread histories. Current studies are examining this overlap zone more thoroughly to see if the range boundary is shifting, and we are also investigating differences in dispersal abilities of both species in different habitats.

This research provides a base of information to underpin further studies of the demographics of these two noxious weeds. Taken by itself, it is evidence of regional co-occurrence, but distinct local separation, of two very similar invasive species. Interestingly, this would not be recognized from distributional data collected and presented at larger spatial scales: spatial scale matters a great deal. If a species has been identified at two locations in a country or in a state, it is often assumed that its distribution includes the area in

between: our survey shows that this is patently not the case. Thus, our work is an important stepping stone on the road toward investigating species dispersal and spread, and understanding the temporal dynamics and co-occurrence of these congeneric weeds. Furthermore, distributional survey work is an essential requirement for improved monitoring and management practices. By understanding spatial distribution patterns, we can identify risk areas for long term monitoring programs, target nascent regions with preventative maintenance and improve the effectiveness of infestation management.

Acknowledgements

Special thanks to Ottar Bjørnstad and Olav Skarpaas for substantial help with the statistics, Dave Mortensen, Robert Humston and lab, Loke Kok, Carl Keener, Art Gover, Dan Laughlin, Carrie Schwarz, Sarah Goslee, the Curran lab and Wil Mountain for their insights throughout the project, Tim Block, Ann Rhoades and Bonnie Isaac for herbarium data, Amy Griffin for help with GIS applications and Kurt Grieb for use of his camera and computer. This project would not have been possible without the time and efforts of a number of dedicated undergraduate lab assistants. Thanks to Teresina Bailey, Emily Phillips, Emily Rauschert and Zeynep Sezen for comments and physical assistance throughout the project. We are grateful to Ottar Bjørnstad, Carla Cáceres, Yvonne Collingham, Fangliang He, Dave Mortensen, Aníbal Pauchard, Olav Skarpaas, two anonymous reviewers and the Shea lab group for comments on the manuscript. We wish to thank the Schreyer Honors College and the Eberly College of Science at the Pennsylvania State University for providing funding to M.R.A. Part of this work was also supported by USDA-CSREES (Biology of Weedy and Invasive Plants) NRI Grant #2002-35320-12289 and by NSF Grant #DEB-0315860 to K.S.

References

- Aspinall RJ (2002) Use of logistic regression for validation of maps of the spatial distribution of vegetation species derived

- from high spatial resolution hyperspectral remotely sensed data. *Ecological Modelling* 157: 301–312
- Augustin NH, Muggleston MA and Buckland ST (1996) An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology* 33: 339–347
- Batra ST (1978) *Carduus* thistle distribution and biological control in the northeastern states. In: Frick KE (ed) *Biological Control of Thistles in the Genus Carduus in the United States: A progress report*, pp 18–22. Science and Education Administration, US Dept. of Agriculture, Stoneville, MS
- Bjørnstad ON (2002) NCF package. Retrieved from <http://asi23.ent.psu.edu/onb1/software.html>
- Bull CM (1991) Ecology of parapatric distributions. *Annual Review of Ecology and Systematics* 22: 19–36
- Bull CM and Possingham H (1995) A model to explain ecological parapatry. *American Naturalist* 145: 935–947
- Bullock JM, Edwards RJ, Carey PD and Rose RJ (2000) Geographical separation of two *Ulex* species at three spatial scales: does competition limit species' ranges? *Ecography* 23: 257–271
- Clapham AR, Tutin TG and Warburg DM (1962) *Flora of the British Isles* 2nd edn, Cambridge University Press, Cambridge
- Collingham YC, Wadsworth RA, Huntley B and Hulme PE (2000) Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent. *Journal of Applied Ecology* 37: 13–27
- Cowley MR, Wilson RJ, Leon-Cortez JL, Gutierrez D, Bulman CR and Thomas CD (2000) Habitat-based statistical models for predicting the spatial distribution of butterflies and day-flying moths in a fragmented landscape. *Journal of Applied Ecology* 37: 60–72
- Crawley MJ (2002) *Statistical Computing: an introduction to data analysis using S-Plus*. John Wiley & Sons, Chichester
- De'Ath G (2002) Multivariate regression trees: a new technique for modeling species–environment relationships. *Ecology* 83: 1105–1117
- Desrochers AM, Bain JF and Warwick SI (1988) The biology of Canadian weeds 89. *Carduus nutans* L. and *Carduus acanthoides* L. *Canadian Journal of Plant Science* 68: 1053–1068
- Dieleman JA, Mortensen DA, Buhler DD, Cambardella CA and Moorman TB (2000) Identifying associations among site properties and weed species abundance. I. Multivariate analysis. *Weed Science* 48: 567–575
- Environmental Resources Research Institute (1996) *Areas of Carbonate Lithology*. Pennsylvania Department of Environmental Protection, Harrisburg, PA
- Erickson RO (1945) The *Clematis fremontii* var. *riehlii* population in the Ozarks. *Annals of the Missouri Botanical Garden* 32: 413–460
- ESRI (1999–2001) ArcGIS Desktop, Version 8.1. Redlands, CA
- He F and Gaston KJ (2000) Estimating species abundance from occurrence. *American Naturalist* 156: 553–559
- Jessep CT (1989) *Carduus nutans* L., nodding thistle (Asteraceae). In: Cameron PJ, Hill RL, Bain J and Thomas WP (eds) *A Review of Biological Control of Invertebrate Pests and Weeds in New Zealand (1974 to 1987)*, pp 339–342. CAB International, Wallingford
- Julien MH and Griffiths MW (eds) (1998) *Biological Control of Weeds: A World Catalogue of Agents and their Target weeds*. 4 edn, CABI Publishing, Oxon, UK, pp 223
- Kok LT (2001) Classical biological control of nodding and plumeless thistles. *Biological Control* 21: 206–213
- Kunin WE (1998) Extrapolating species abundance across spatial scales. *Science* 281: 1513–1515
- Legendre P and Legendre L (1998) *Numerical Ecology*. 2nd edn, Elsevier, Amsterdam
- Manly BF (1997) A method for the estimation of parameters for natural stage-structured populations. *Researches on Population Ecology* 39: 101–111
- Mihulka S and Pysek P (2001) Invasion history of *Oenothera* congeners in Europe: a comparative study of spreading rates in the last 200 years. *Journal of Biogeography* 28: 597–609
- Myers WL and Bishop JA (1999) *Vegetative Land Cover for Pennsylvania, 30 meter Resolution*. Pennsylvania Gap Analysis Program. The Pennsylvania State University, University Park, PA
- NatureServe (2002) *NatureServe Explorer: An online encyclopedia of life [web application]*. Version 1.8, Nature Serve, Arlington, VA Retrieved from <http://www.natureserve.org/explorer>
- Pauchard A, Alaback PB and Edlund EG (2003) Plant invasions in protected areas at multiple scales: *Linaria vulgaris* (Scrophulariaceae) in the west Yellowstone area. *Western North American Naturalist* 63: 416–428
- Pimentel D, Lach L, Zuniga R and Morrison D (2000) Environmental and economic costs of nonindigenous species in the United States. *Bioscience* 50: 53–65
- Prentice IC, Cramer W, Harrison SP, Leemans R, Monserud RA and Solomon AM (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography* 19: 117–134
- Rew LJ, Maxwell BD, Dougher FL and Aspinall R. Searching for a needle in a haystack: evaluating survey methods for sessile species. *Biological Invasions this issue*
- Rhoads AF and Klein WM (1993) *The vascular Flora of Pennsylvania: Annotated Checklist and Atlas*. American Philosophical Society, Philadelphia
- Shea K and Chesson P (2002) Community ecology theory as a framework for biological invasions. *Trends in Ecology and Evolution* 17: 170–176
- Shea K and Possingham HP (2000) Optimal release strategies for biological control agents: an application of stochastic dynamic programming to population management. *Journal of Applied Ecology* 37: 77–86
- Skinner K, Smith L and Rice P (2000) Using noxious weed lists to prioritize targets for developing weed management strategies. *Weed Science* 48: 640–644
- Smith PA (1994) Autocorrelation in logistic-regression modeling of species distributions. *Global Ecology and Biogeography Letters* 4: 47–61
- US Dept. of Agriculture (1992) *Forest density (FRSTDEN)*. Southern Forest Experiment Station, Forest Inventory and Analysis
- US Geological Survey (1999) *National elevation dataset*. EROS Data Center, Sioux Falls, SD

Warwick SI, Bain J, Wheatcroft R and Thompson BK (1989) Hybridization and introgression in *Carduus nutans* and *C. acanthoides* reexamined. *Systematic Botany* 14: 476–494

Warwick SI, Thompson BK and Black LD (1992) Hybridization of *Carduus nutans* and *Carduus acanthoides* (compositae): morphological variation in F1 hybrids and backcrosses. *Canadian Journal of Botany* 70: 2303–2309