

Assessing plant-nutrient relationships in highly invaded Californian grasslands using non-normal probability distributions

Gea-Izquierdo, Guillermo^{1,2*}; Gennet, Sasha^{1,3} & Bartolome, James W.^{1,4}

¹Ecosystem Sciences, University of California, Berkeley, CA 94720 USA; ²Departamento Sistemas y Recursos Forestales, CIFOR-INIA, Crta. La Coruña km 7.5 28040 Madrid, Spain; ³E-mail asgennet@nature.berkeley.edu;

⁴E-mail jwbart@nature.berkeley.edu; *Corresponding author; E-mail guigeiz@inia.es; Fax +34913572293

Abstract

Question: Is native species occurrence related to soil nutrients in highly invaded Californian annual grasslands? What is the best method to analyze this relationship, given that native species occur in very low numbers and are absent from many locations?

Location: California, USA.

Methods: We investigated the effects of soil characteristics and livestock grazing on native plant occurrence at 40 plots from six sites during the period 2003–2005. Low absolute cover (< 5.8%) of native species resulted in strongly skewed, zero-inflated data sets. To overcome problems in the analysis created by non-normality and correlations within plots, we used GLMs and GLMMs, either with a Poisson or a negative binomial distribution, to analyse native species richness and *Nassella pulchra* cover.

Results: *N. pulchra* cover was strongly associated with low phosphorus in sandy soils, whereas native species richness was highest in soils with low available nitrogen (high C:N).

Conclusion: Under current conditions, phosphorus seems to be a critical factor influencing abundance of *N. pulchra*. Low fertility soils may provide refugia for native species in highly invaded California grasslands as they may be below a threshold required for non-native annuals to completely dominate. By using non-normal distributions in linear models with random components, we report well fitted models with more accurately tested significant covariates.

Keywords: GLM; GLMM; Invasive species; *Nassella pulchra*; Negative binomial distribution; Skewed distribution; Soil phosphorus.

Abbreviations: CV = Coefficient of variation; GLM = Generalized linear model; GLMM = Generalized linear mixed model; PDF = Probability density function.

Introduction

Invasion by exotic species is a biological problem worldwide. California's grasslands are among the most heavily invaded terrestrial communities (e.g. Huenneke & Mooney 1989; Keeley et al. 2002; Eviner & Chapin 2003; Seabloom et al. 2003). The ecological mechanisms and environmental variables that enable invasion and dominance of introduced annual grasses in Californian grasslands have been analysed in many studies (e.g. Huenneke et al. 1990; Dyer & Rice 1999; Seabloom et al. 2003), but there are few 'field' (i.e. without implementing any manipulation to the grassland) analyses of native species persistence and coexistence in annual grasslands already invaded (Hamilton et al. 1999). The mechanisms that enable non-natives to dominate are scale dependent and complex to analyse as they seem to be a result of the interaction of several ecological factors (Davies et al. 2005), but disturbances such as grazing management, drought and nutrient additions tend to facilitate a competitive shift to dominance by non-natives in these grasslands (e.g. Huenneke et al. 1990; Dyer & Rice 1999; Harrison et al. 2003). Dominance may be related to soil fertility (Huenneke et al. 1990; Gram et al. 2004) and to a selective advantage of non-native annuals to abiotic factors in California. Annual grass propagules germinate soon after the first rains in the fall and grow quickly to complete their life cycle while there is still water available (Huenneke & Mooney 1989). As opposed to native perennials, they allocate most of their growth to shoots. Rather than utilizing nutrients more efficiently, annual introduced species seem to dominate as a result of recruitment, dispersal and shading (Holmes & Rice 1996; Dyer & Rice 1999; Seabloom et al. 2003).

Identifying driving factors and quantifying the environmental variables that enable the native species to persist at low levels in this ecosystem is challenging (Huenneke & Mooney 1989; Eviner & Chapin 2003) because: (1) plant species composition is diverse over both space and time; (2) grassland species are morpho-

logically and phenologically plastic; (3) plant species' responses vary independently and (4) the relative rarity of natives results in strongly skewed, zero-inflated data sets (i.e. data sets with many zeroes) when sampling is random, hence, appropriate non-Gaussian statistics are needed. We examined the distribution of native species in annual, non-native species dominated grasslands as a function of soil properties and livestock grazing. Based on previous studies conducted on serpentine grasslands in California and on the lower fertility ecological conditions in which the mediterranean species thrive (Marañón 1986; Puerto & Rico 1988), we expected that presence of native species at the study sites would be negatively correlated with fertility, particularly nitrogen (Huenneke et al. 1990; Gram et al. 2004).

Generalized linear models: GLM and GLMM

Generalized linear models (GLM) are an extension of general linear models. While general linear models only handle normally distributed data, GLMs can handle both qualitative and quantitative data, with flexible mean-variance relationships. Thus, they can be used to analyse non-Gaussian, skewed data sets, where normal models are not appropriate, without data transformation or non-parametric analysis (McCullagh & Nelder 1989; Venables & Ripley 2002). Generalized linear mixed models (GLMM) are similar in structure to GLMs but can also handle non-independent observations by including both random effects and correlated errors. Their simplicity and flexibility make generalized linear models suitable for many biological applications (Vincent & Haworth 1983; White & Bennetts 1996; Guisan & Zimmerman 2000).

Material and Methods

Study area

We studied the distribution of native plants in grasslands located in the Diablo Range east of San Francisco (California). The original vegetation in the study area was probably a matrix dominated by perennial bunchgrasses, notably *Nassella pulchra*, with diverse annual forbs in the interstices, mixed with *Quercus* woodlands (Huenneke & Mooney 1989). Soils in the study area are derived primarily from a complex mosaic dominated by

sedimentary shale, greywacke sandstone and chert. Most soils are well drained clay or silt loams (Harden 1998). Most patches of grassland are grazed by either cattle or sheep at light to moderate intensities.

The central coast range of California experiences a mild mediterranean climate with warm, dry summers and cool, rainy winters. Annual precipitation (July-June mean at four local weather stations) in the study area was 561 mm in 2003, 464 mm in 2004, and 663 mm in 2005.

Experimental design

We selected 40 grassland study plots across six different protected regional parks, including nine ungrazed, 13 sheep grazed and 18 cattle grazed plots. Plot centroids were located a minimum of 200 m apart, and at least 150 m from the nearest woody vegetation. Each plot included four 18-m transects radiating in the cardinal directions from the centroid. Data points were located every 10 cm along the transects between one and 5.5 m; and every 50 cm from 5.5 m to 18 m for a total of 70 hits per transect. The first hit on canopy species (including litter and bare ground) was recorded at each point. Vegetation data were collected during March-May each year from 2003-2005. Soil samples (10 cm deep) were collected at 5.5 m from the centroid along each transect. Samples were analysed for total Carlo-Erba nitrogen (N) and carbon (C), Bray phosphorus (P) (Troeh & Thompson 1993), cation exchange capacity (CEC) and particle size at the UC Davis Analytical Laboratory (<http://groups.ucanr.org/danranlab>). Soil density and pH were measured at the UC Berkeley Range Ecology Laboratory (Table 1). Soil chemistry samples were collected during March-May 2003 and soil density samples were collected during March-May 2005. We assume that the soil variables studied are unlikely to have changed significantly between years during the study (Jones & Woodmansee 1979; Troeh & Thompson 1993).

Statistical analysis

Values for each of the soil variables were strongly correlated among transects in each plot. Therefore, values for each variable from the four transects were pooled to obtain a 'per plot' value to assure independence among samples. In addition, values for the variables carbon and nitrogen were strongly correlated (Pearson $\rho = 0.955$),

Table 1. Summary of soil plot attributes.

Per plot	Ntotal (%)	C:N	Phosphorus (ppm)	CEC (meq/100g)	pH	Sand (%)	Silt (%)	Clay (%)	Soil density (g/cm ³)
Mean	0.21	9.97	22.40	33.92	6.18	31.04	41.08	27.88	1.22
CV	20.0	6.5	108.2	29.9	9.7	46.7	18.3	38.7	15.9

as expected (Troeh & Thompson 1993). Carbon was, therefore, excluded as a separate covariate. Silt was also excluded because it was complementary to sand and clay. The full models included nine covariates: % nitrogen, carbon-nitrogen ratio (C:N), phosphorus (ppm), cation exchange capacity (meq/100g), % sand, % clay, soil density (g/cm³), pH and grazing type (GT; coded as either 'sheep grazed', 'cattle grazed', or 'ungrazed').

We used GLMs and GLMMs to run two separate analyses. In the first analysis the dependent variable was native species richness (i.e. number of native species per plot). In the second we analysed the absolute cover by plot of *N. pulchra*. GLMs have the generic expression: $E(Y_i) = \mu_i = g^{-1}(\eta_i)$, where: Y_i is either native richness or *N. pulchra* cover;

$$\eta_i = \sum_{j=1}^p x_{ij} \beta_j \quad (1)$$

with x_{ij} the value of the j th covariate for observation i , out of p different explanatory variables, β_j the unknown parameter to be estimated; and $g(\cdot)$ is the link function. GLMMs have the general expression:

$\mu_{ij} = E(Y_{ij} | b_i) = g^{-1}(x'_{ij} \beta + z'_{ij} b_i)$; x_{ij} and z_{ij} are matrices of known covariates; β is a p -dimensional vector of fixed effects, and b_i a q -dimensional vector of random effects, $b_i \sim N(0, G)$; $g(\cdot)$ and Y_{ij} have the same meaning as above (Diggle et al. 2002; Molenberghs & Verbeke 2005, see App. 1 for more details).

Selecting the most appropriate probability density function (PDF) is a critical step in generalized linear models. With count data, the most common PDF used is the one-parametric Poisson distribution (Vincent & Haworth 1983), which assumes a mean equal to the variance. This assumption makes it inappropriate for many ecological variables, which exhibit higher variances than the mean (overdispersion). For overdispersed variables, the two-parametric negative binomial distribution is superior to the Poisson because its second parameter can fit a wide spectrum of mean-variance relationships (White & Bennetts 1996; Venables & Ripley 2002). In our study we first analysed each year separately, then the three years (2003-2005) together (120 observations), using either a Poisson or a negative binomial GLM fitted

with iterative weighted least squares (we used SPLUS 7.0). The most parsimonious models with the most appropriate PDFs were selected using nested analysis of deviance tested against the χ^2 distribution (McCullagh & Nelder 1989). We used the percentage of explained deviance [$100 \cdot (\text{null deviance} - \text{residual deviance}) / \text{null deviance}$] as an index of goodness of fit. As we had three years of repeated plot measurements, the variance-covariance structure was unlikely to be constant diagonal. We used the residuals obtained from the GLM fit for the three years to graphically study the covariance structure (McCullagh & Nelder 1989; Diggle et al. 2002, see App. 2). Finally we fitted GLMMs with the selected PDFs in the GLMs to evaluate the significance of covariates. GLMMs were fitted using SAS 9.1 PROC GLIMMIX. Maximum likelihood estimates were obtained using linearization methods (pseudo-likelihood estimation), therefore no likelihood ratio tests could be performed. The significance of the fixed effects was analysed using F -tests with Satterthwaite type degrees of freedom (Molenberghs & Verbeke 2005). We used log-links in all cases.

Results

The soil variables are summarized in Table 1. The mean (\pm SD) plot species richness was 2.18 ± 2.10 for 2003, 1.75 ± 2.28 for 2004 and 2.38 ± 2.44 for 2005, whereas plot absolute cover for *N. pulchra* was $3.06 \pm 6.96\%$, $2.59 \pm 6.01\%$ and $2.17 \pm 4.65\%$ respectively. Among natives, *N. pulchra* was the species accounting for most of the cover. The rest of the native component was made up from a high number of species (up to 42 in 2005) each accounting for very few hits (Table 2, App. 2). The absolute dominance of non-native species resulted in native species richness and *N. pulchra* presenting overskewed non-normal distributions as a result of the many zeroes encountered when sampling (Fig. 1). This suggests using Poisson or negative binomial probability distributions in the models instead of the Gaussian.

Table 2. Summary of plot vegetation features. The results are % of total cover.

Year	2003	2004	2005
Relative vegetation cover (%)	85.7	79.8	83.4
Species richness (n)	95	87	106
Native species richness (n)	34	34	42
Native species relative cover (% of total vegetation hits)	5.7	4.9	4.8
<i>Nassella pulchra</i> (% of total)	3.6	3.2	2.6
(% of native)	62.3	65.6	53.4

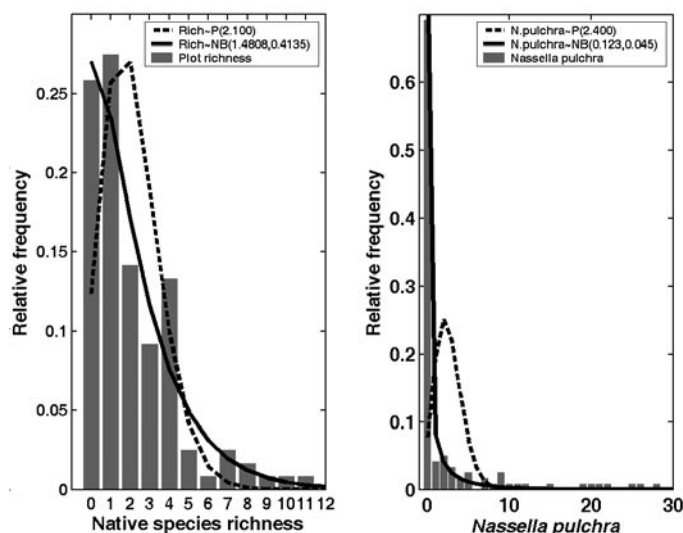


Fig. 1. Negative binomial (NB) and Poisson (P) fits to (a) *Nassella pulchra* plot relative cover and (b) Plot native species richness (Rich) distributions.

Native species richness and soil fertility

When comparing different PDFs for native species richness, the negative binomial distribution (log-likelihood $ll = -232.5$) fit was slightly better than the Poisson ($ll = -260.5$, Fig. 1). A fit of a GLM to the 120 observations for native species richness, the residuals were better displaced in the Poisson model (App. 4) compared to the negative binomial. The Poisson model showed little overdispersion ($\rho = \text{Residual deviance} / df = 1.50$; Venables & Ripley 2002); both distributions selected the same covariates, explaining similar deviance (App. 3). The expression of the Poisson GLM was (full model against reduced model: $\chi^2_6 = 6.03$, $p = 0.420$):

$$\text{Richness} = \exp(-7.574 - 0.009 \text{ Phosphorus} + 0.811 \text{ CN} - 8.228 \text{ Nitrogen} + 1.631 \text{ Soil density}) \quad (2)$$

The results for the individual Poisson GLMs fitted by year were similar to Eq. 2 in terms of the deviance explained, which in all cases was over 38%. However, P was not selected, while C:N was the variable most strongly selected (data not shown). The Poisson GLMM results are shown in Table 4 (see also App. 4). No ‘park-level’ random significant effects was found ($\sigma^2_j = 0.078 \pm 0.083$) while a plot random effect was significant (Table 4). Hence the GLMM finally fitted is a Poisson random intercept model with independent errors: $g(E[Y|y]) = x'\beta + z'y$ (Table 4). The model finally fitted was:

$$\text{Richness} = \exp(-10.027 + 1.012 \text{ CN} - 9.504 \text{ Nitrogen} + 1.972 \text{ Soil density} + \gamma_i) \quad (3)$$

with $\gamma_i \sim N(0, 0.343)$. This model included fewer covariates than the GLM, reflecting the increase in the significance tests accuracy acquired when including the random effect. Nitrogen availability (i.e. high C:N and low total nitrogen) was negatively correlated with native richness. However it

was significant that this relationship was not very strong as reflected by the percentage of deviance explained in the GLM and Fig. 2c, d.

Nassella pulchra cover and soil fertility

The negative binomial PDF maximized the log-likelihood for *N. pulchra*, ($ll = -189.9$; Poisson $ll = -557.4$; Fig. 1). Additionally, among GLM’s, the negative binomial model was the most conservative one, selecting two covariates, and its residuals were the least dispersed (data not shown). The Poisson model showed overdispersion ($\rho = 2.46$), and selected five variables (Table 3). The most parsimonious negative binomial GLM for the 120 observations explained 76.7% of the deviance (Table 3):

$$\%N. \text{ pulchra} = \exp(-0.237 \text{ Phosphorus} + 0.088 \text{ Sand}) \quad (4)$$

($\theta = 1.434 \pm 0.355$)

Phosphorus was the most important covariate, accounting for most of the deviance explained (52.9% deviance; $\chi^2_1 = 86.93$, $p < 0.001$). The same two variables (P and sand) were selected when studying the years separately (deviance > 71.8%) and when fitting the GLMM: *N. pulchra* is clearly present only in soils with the lowest phosphorus content and higher sand percentage (Fig. 2a, b). In the GLMM a ‘park-level’ random effect was again non significant ($\sigma^2_j = 0.149 \pm 0.208$), however a plot random effect was significant (Table 4). When including a normal plot random effect, which was significant (Table 4), the negative binomial shape parameter was not significant for *N. pulchra* (i.e. $\theta = 0$). Therefore these results are identical to the fitted when specifying the Poisson distribution (see App. 1) in the GLMM (Table 4). The model for *N. pulchra* was:

$$\%N. \text{ pulchra} = \exp(-0.247 \text{ Phosphorus} + 0.071 \text{ Sand} + \gamma_i), \quad (5)$$

with $\gamma_i \sim N(0, 2.321)$

Table 3. GLM Poisson and Negative binomial fitted to *N. pulchra* cover (120 observations). Dres = deviance residuals; RDR = residual deviance reduction of model K compared to the null model of no significant covariates. The *p*-value is computed comparing model K with model K-1 (of same PDF GLM), using the χ^2 -distribution. * = *p*-value for model K compared to model K-2. The most parsimonious models with each PDF are in bold type.

Error type	Model (K)	Dres (Max-Min)Median	RDR (%)	<i>p</i>
Poisson	<i>N. pulchra</i> = f (GT, N, C:N, P, CEC, Sand, Clay, pH, Soil density)	(-3.2, 3.8) -0.3	72.4	-
Poisson	<i>N. pulchra</i> = f (GT, P, CEC, Sand, Soil density)	(-3.1, 3.9) -0.3	71.5	0.136
Poisson	<i>N. pulchra</i> = f (P, Sand)	(-4.4, 4.1) -0.4	60.7	< 0.001
Poisson	<i>N. pulchra</i> = f (GT, CEC, Sand, Soil density)	(-3.1, 6.1) -0.9	35.0	< 0.001*
Negative binomial	<i>N. pulchra</i> = f (GT, N, C:N, P, CEC, Sand, Clay, pH, Soil density)	(-1.9, 1.4) -0.2	72.3	-
Negative binomial	<i>N. pulchra</i> = f (GT, N, C:N, P, Sand)	(-2.3, 2.2) -2.2	70.3	0.304
Negative binomial	<i>N. pulchra</i> = f (P, Sand)	(-2.4, 1.8) -0.3	69.9	0.677
Negative binomial	<i>N. pulchra</i> = f (Sand)	(-1.3, 1.5) -0.6	17.0	< 0.001

Discussion

Native species occurrence and soil nutrients

In our results, soils with low phosphorus are providing refugia for *Nassella pulchra* while soils with low nitrogen mineralization rates may be providing a refuge for other native plants in highly invaded Californian annual grasslands. Soil properties determine, to a great extent, plant distribution and this is particularly true of limiting nutrients such as N and P (e.g. Robertson et al. 1988; Huenneke et al. 1990; Koerselman & Meuleman 1996). Past research has also suggested that species diversity and richness decline following nitrogen enrichment (Vitousek et al. 1997) and that low nutrient availability may impede dominance by a single species (Huenneke et al. 1990). Our research corroborates these results for highly invaded Californian annual grasslands, especially for the species *N. pulchra*.

Phosphorus and nitrogen are the most important plant

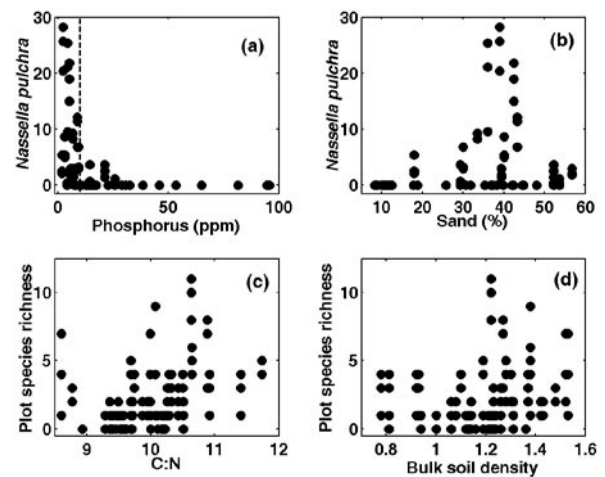


Fig. 2. Relationships between (a) *N. pulchra* absolute cover (%) and phosphorus (dashed line in phosphorus = 10 ppm); (b) *N. pulchra* and sand; (c) native species richness and C:N; (d) native richness and soil density.

Table 4. GLMM results. ¹Confidence interval estimated in GLMM: when it does not include zero, we consider it significant. ²Estimates for the most parsimonious model. **t* test for the intercept. SE = Standard error. Significant covariates are in bold type.

Model		<i>Nassella pulchra</i> Negative binomial GLMM		Native species Richness Poisson GLMM	
		Estimate ² (SE)	<i>p</i> -value* (<i>F</i> value)	Estimate ² (SE)	<i>p</i> -value (<i>F</i> value)*
Fixed effects	Intercept	-	0.439 (-0.79)	-10.027 (2.477)	0.003 (-4.05)
	Phosphorus	-0.247 (0.050)	< 0.001 (24.75)	-	0.162 (2.04)
	Sand	0.071 (0.015)	< 0.001 (21.25)	-	0.947 (0.00)
	C:N	-	0.970 (0.00)	1.012 (0.200)	< 0.001 (25.67)
	N	-	0.769 (0.09)	-9.504 (3.208)	0.005 (8.78)
	Soil density	-	0.121 (2.61)	1.972 (0.694)	0.008 (8.09)
	GT	-	0.783 (0.25)	-	0.678 (0.39)
	CEC	-	0.267 (1.30)	-	0.454 (0.58)
	pH	-	0.427 (0.64)	-	0.631 (0.24)
	Clay	-	0.533 (0.40)	-	0.963 (0.00)
Plot random effect	σ^2_i (intercept) ¹	2.321 (0.865)	-	0.343 (0.130)	-
Shape parameter	θ^1	0.008 (0.027)	-	-	-

macronutrients (Troeh & Thompson 1993). The C:N ratio is a measure of soil mineralization capacity, and is therefore a better measurement of nitrogen availability than total nitrogen (Troeh & Thompson 1993; Eviner & Chapin 2003). Our study design and the complexity of the nitrogen cycle prohibit us from identifying the precise physiological mechanism leading the relationship between native richness and nitrogen (Robertson et al. 1988; Troeh & Thompson 1993; Herman et al. 2003). The mechanism driving the relationship between soil phosphorus and vegetation in our study is easier to interpret due to the simpler, closed nature of the phosphorus cycle and the spatial and temporal homogeneity of its distribution (Jones & Woodmansee 1979; Woodmansee & Duncan 1980; Center et al. 1988; Eviner & Chapin 2003). This simplicity and the very high negative correlation between phosphorus and *N. pulchra*, increases the ecological significance of our findings. *N. pulchra* abundance was strongly associated with low levels of soil phosphorus, and exhibited a release threshold at 10 ppm (Fig. 2c). In our study, sand content above 30% and below 45% was also associated with higher occurrence of *N. pulchra* (Fig. 2b). Higher sand content decreases water retention in soils and therefore leads to a decrease in the phosphorus available to plants. Conversely, higher clay content in soils increases the surface area available for adsorption of minerals (Robinson 1971; Troeh & Thompson 1993). Higher sand content enables faster percolation of water to deeper soil layers, which perennial native species (e.g. *N. pulchra*) can access more readily than the shallow rooted, introduced annual grasses (Holmes & Rice 1996; Dyer & Rice 1999; Hamilton et al. 1999).

Some previous studies in Californian grasslands have reported results that seem to contradict our findings (Robinson 1971; Seabloom et al. 2003). However, the plant communities and abiotic conditions were not comparable (Corbin & D'Antonio 2004): the previous studies were undertaken in more mesic, and more fertile grasslands with higher cover of perennial natives. Additionally, in Robinson (1971) other nutrients might be more limiting: he points out the low nitrogen content of his soils. In this way, the effect of phosphorus or nitrogen might be reflecting the balance between N:P, with P as the limiting factor in our study (Koerselman & Meuleman 1996). Our findings support the conclusions of Huenneke et al. (1990) and Weiss (1999) in serpentine, who reported a clear response of non-native annual species to fertilization, and lower phosphorus levels in forbs plots compared to introduced annual plots. Most studies analysing the relationship between native species persistence and soil fertility have been focused on serpentine. Generally, the low fertility and toxicity of serpentine soils provide a spatial refuge for a number of native and endemic spe-

cies (e.g. Harrison 1999; Gram et al. 2004). *N. pulchra*, as a grass, is likely to demand high levels of nutrients, particularly nitrogen, but also phosphorus, in the same way as introduced grasses. However, as a perennial, it has deeper roots, and hence is able to utilize deeper soil resources unavailable to introduced dominant annuals (Holmes & Rice 1996; Dyer & Rice 1999).

Many sites in California are reported to be deficient in N, P and S (Center et al. 1988). However, the soils in the study site are appreciably more fertile than soils in the western Iberian peninsula, within the native range of the introduced species, particularly in open grasslands beyond the influence of tree canopies, where grasses do not dominate (e.g. Puerto & Rico 1988; Gallardo et al. 2000). The soils in our study have a much heavier texture, and higher nitrogen content (> 0.2 %) under good mineralization conditions (C:N < 10) (Table 1). In Iberia, most species considered invasive in California are categorized as ruderal, colonizing nutrient-enriched sites in generally poor siliceous soils and exhibit fast growth (Chapin 1980).

Management implications

Our findings can be translated to management and restoration recommendations. Although the native species in our study grassland are likely to have different ecological characteristics (Eviner & Chapin 2003), analysing them together is a practical strategy for identifying management level patterns. In the present plant community, dominated by a small number of non-native species (App. 2), the natives may be finding refuge in the few locations where non-natives are not as dominant, not because those sites are ecologically optimum. Under low nutrient (P and C:N) conditions the non-native species would dominate less. Fertility has been anthropogenically altered in the last two centuries through agricultural fertilization, burning, pollution deposition and livestock grazing (Jones & Woodmansee 1979; Center et al. 1988; Weiss 1999; Keeley 2002). Phosphorus can be related to burning, to some extent, and burning plays an extraordinary role in these grasslands (Keeley 2002; Dyer 2003; Harrison et al. 2003). Decreasing phosphorus system wide may be feasible. Annual species do not store nutrients below-ground between years, therefore larger quantities of nutrients enter and leave annual dominated grassland systems each year (Jones & Woodmansee 1979). Because the primary source of phosphorus is litter mineralization (Troeh & Thompson, 1993), it should be possible to reduce its levels to below the apparent threshold of 10 ppm, by mowing or clipping and seeding native species (at least *N. pulchra*), as suggested by Seabloom et al. (2003). From our study we cannot assess the importance of seed dispersal limitation

in native species and any lack of native seeds in seed banks. Low seed production or resource competition from non-native annual grasses may overwhelm natives in the seedling stage, and prevent them from returning to dominance once displaced (Seabloom et al. 2003). Our results cannot exclude other factors that might limit the distribution of species, including nitrogen under changed grazing conditions (Eviner & Chapin 2003). The effect of grazing on the distribution and occurrence of native species is complex and not yet well understood (e.g. Huenneke & Mooney 1989; Harrison et al. 2003).

Application of generalized linear models with non-Gaussian ecological data

Non-normally distributed dependent variables are common in ecology. Generalized linear models are not yet very common in ecology (particularly negative binomial models) whereas they are quite common in other biological sciences. We applied non-Gaussian generalized linear models to appropriately fit non-normal data and make better inferences regarding the covariates. Selecting the PDF that best fits the data greatly improved the accuracy of the models. As seen with the GLM approach for *N. pulchra*, using an inappropriate PDF make the models select too many variables, overinflating the deviance explained, and resulting in bigger residuals. In addition, and as a result of the within-plot correlation included in our data, plot residuals improved with the inclusion of a random effect in both cases (App. 4) and the accuracy of covariate selection increased as seen when comparing (1) and (2) (Table 3; App. 3). In the *N. pulchra* negative binomial GLMM, when including the normal random effect, the dispersion parameter turned out to be non-significant, as opposed to the negative binomial GLM. As $\theta = 0$, the GLMM followed a more parsimonious one parametric Poisson distribution. The influence of the 'zeroes' was partly removed by including the random effect in the model (which has mean zero).

The applicability of GLMs to non-normal data sets, common in many ecological studies, and our success in identifying patterns in our highly skewed, zero-inflated data set, strongly points towards increased application of generalized linear models in ecology. These models are easy to apply and can fit different structures of correlation within observations. In our study we focused on the community level. However, managers sometimes aim at individual species. Therefore, it would be interesting to study more single species distributions, in the way we looked at *N. pulchra*. This would require a modified sampling design, as single native species rather than *N. pulchra* were scarcely represented in our results.

Conclusions

The use of generalized linear models with non-Gaussian probability distributions and random effects enabled us to successfully fit the distribution of our zero-inflated, non-normal dependent variables, and was crucial to address the covariates really influencing the dependent variables. Our results suggest that in semi-arid, Californian grasslands dominated by introduced mediterranean annual grasses, 'islands' of low soil fertility may be providing refuges for scarce native species to compete successfully with introduced annual species. *N. pulchra* occurrence was strongly associated with sandy soils that have extremely low phosphorus, and exhibited a release threshold at $P = 10$ ppm. Native species richness seemed to increase in soils with higher C:N ratio (i.e. soils with low N mineralization rates). These findings would support what some authors have already shown in field manipulations: that above-ground biomass removal by clipping – which would decrease mineralization and lower phosphorus – can increase the richness and cover of native plants, at least *N. pulchra*. Our results are particularly interesting because they analyse at the landscape scale what other research has shown through plant and soil manipulation. For these reasons, we believe that our hypothesis should be addressed at a larger scale.

Acknowledgements. We gratefully acknowledge Peter Hopkinson for enriching discussion, and Michele Hammond and all the students and staff that helped with data collection. We thank Barbara Allen-Díaz, Cherie Briggs, Jeffrey Fehmi, Wayne Getz, Robert Hijmans, Mitch McClaran, Kevin Rice and two anonymous reviewers for helpful comments.

References

- Center, D.M., Vaughn, C.E. & Jones, M.B. 1988. Effects of management on plant production and nutrient cycling on two annual grassland sites. *Hilgardia* 57: 1-40.
- Chapin, III F.S. 1980. The mineral nutrition of wild plants. *Annu. Rev. Ecol. Syst.* 11: 233-260.
- Corbin, J.D. & D'Antonio, C.M. 2004. Competition between native perennial and exotic annual grasses: implications for an historical invasion. *Ecology* 85: 1273-1283.
- Davies, K.F., Chesson, P.L., Harrison, S., Inouye, B.D., Melbourne, B.A. & Rice, K.J. 2005. Spatial heterogeneity explains the scale dependence of the native-exotic diversity relationship. *Ecology* 86: 1602-1610.
- Diggle, P.J., Heagerty, P., Liang, K.Y. & Zeger, S.L. 2002. *Analysis of longitudinal data*, 2nd. ed. Oxford University Press, New York, NY, US.
- Dyer, A.R. 2003. Burning and grazing management in a California grassland: growth, mortality, and recruitment on *Nassella pulchra*. *Restor. Ecol.* 11: 291-296.
- Dyer, A.R. & Rice, K.J. 1999. Effects of competition on re-

- source availability and growth of a California bunchgrass. *Ecology* 80: 2697-2710.
- Eviner, V.T. & Chapin, III F.S. 2003. Functional matrix: a conceptual framework for predicting multiple plant effects on ecosystem processes. *Annu. Rev. Ecol. Syst.* 34: 455-85.
- Gallardo, A., Rodríguez, J.J., Covelo, F. & Fernández, R. 2000. Soil nitrogen in a Dehesa ecosystem. *Plant Soil* 222:71-82.
- Gram, W.K., Borer, E.T., Cottingham, K.L., Seabloom, E.W., Boucher, V.L., Goldwasser, L., Micheli, F., Kendall, B.E. & Burton, R. 2004. Distribution of plants in a California serpentine grassland: are rocky hummocks spatial refuges for native species? *Plant Ecol.* 172:159-171.
- Guisan, A. & Zimmermann, N.E. 2000. Predictive habitat distribution models in ecology. *Ecol. Model.* 135: 147-186.
- Hamilton, J.G., Holzapfel, C. & Mahall, B.E. 1999. Coexistence and interference between a native perennial grass and non-native annual grasses in California. *Oecologia* 121: 518-526.
- Harden, D.R. 1998. The Coast Ranges: Mountains of Complexity. In: Harden, D.R. (ed.) *California geology*, pp. 252-288. Prentice-Hall, Upper Saddle River, NJ, US.
- Harrison, S. 1999. Native and alien species diversity at the local and regional scales in a grazed California grassland. *Oecologia* 121: 99-106.
- Harrison, S., Inouye, B.D. & Safford, H.D. 2003. Ecological heterogeneity in the effects of grazing and fire on grassland diversity. *Conserv. Biol.* 17: 837-845.
- Herman, D.J., Halverson, L.J. & Firestone, M.K. 2003. Nitrogen dynamics in an annual grassland: oak canopy, climate, and microbial population effects. *Ecol. Appl.* 13: 593-604.
- Holmes, T.H. & Rice, K.J. 1996. Patterns of growth and soil-water utilization in some exotic annuals and native perennial bunchgrasses of California. *Ann. Bot.* 78: 233-243.
- Huenneke, L.F. & Mooney, H.A. (eds.) 1989. *Grassland structure and function: California annual grassland*, 1st ed. Kluwer, London, UK.
- Huenneke, L.F., Hamburg, S.P., Koide, R., Mooney, H.A. & Vitousek, P.M. 1990. Effects of soil resources on plant invasion and community structure in Californian serpentine grassland. *Ecology* 71: 478-491.
- Jones, M.B. & Woodmansee, R.G. 1979. Biogeochemical cycling in annual grassland ecosystems. *Bot. Rev.* 45: 111-144.
- Keeley, J.E. 2002. Native American impacts on fire regimes of the California coastal ranges. *J. Biogeogr.* 29: 303-320.
- Koerselman, W. & Meuleman, A.F.M. 1996. The vegetation N:P ratio: a new tool to detect the nature of nutrient limitation. *J. Appl. Ecol.* 33: 1441-1450.
- Marañón, T. 1986. Plant species richness and canopy effect in the savanna-like "dehesa" of SW-Spain. *Ecol. Mediter.* 12: 131-141.
- McCullagh, P. & Nelder, J.A. 1989. *Generalized linear models*. 2nd ed. Chapman & Hall, London, UK.
- Molenberghs, G. & Verbeke, G. 2005. *Models for discrete longitudinal data*, 1st ed. Springer Berlin Heidelberg, New York, NY, US.
- Puerto, A. & Rico, M. 1988. Influence of tree canopy (*Q. rotundifolia* and *Q. pyrenaica*) on old field succession in marginal areas of Central-Western Spain. *Acta. Oecol.* 9: 337-358.
- Robertson, G.P., Huston, M.A. & Evans, F.C. 1988. Spatial variability in a successional plant community: patterns of nitrogen availability. *Ecology* 69: 1517-1524.
- Robinson, R.H. 1971. An analysis of ecological factors limiting the distribution of a group of *Stipa pulchra* associations. *Korean J. Bot.* 14: 1-20.
- Seabloom, E.W., Harpole, W.S., Reichmann, O.J. & Tilman, D. 2003. Invasion, competitive dominance, and resource use by exotic and native California grassland species. *Proc. Natl. Acad. Sci.* 100: 13384-13389.
- Troeh, F.R. & Thompson, L.M. 1993. *Soils and soil fertility*, 5th ed. Oxford University Press, New York, NY, US.
- Venables, W.N. & Ripley, B.D. 2002. *Modern applied statistics with S*, 4th ed. Springer, Berlin Heidelberg, New York, NY, US.
- Vincent, P.J. & Haworth, J.M. 1983. Poisson regression models of species abundance. *J. Biogeogr.* 10: 153-160.
- Vitousek, P.M., Aber, J.D., Howarth, R.W., Likens, G.E., Matson, P.A., Schindler, D.W., Schlesinger, W.H. & Tilman, D.G. 1997. Human alteration of the global nitrogen cycle: sources and consequences. *Ecol. Appl.* 7: 737-750.
- Weiss, S.B. 1999. Cars, cows, and checkerspot butterflies: nitrogen deposition and management of nutrient-poor grasslands for a threatened species. *Conserv. Biol.* 13: 1476-1486.
- White, G.C. & Bennetts, R.E. 1996. Analysis of frequency count data using the negative binomial distribution. *Ecology* 77: 2549-2557.
- Woodmansee, R.G. & Duncan, D.A. 1980. Nitrogen and phosphorus dynamics and budgets in annual grasslands. *Ecology* 61: 893-904.

Received 24 May 2006;

Accepted 28 November 2006;

Co-ordinating Editor: R. Pakeman.

For App. 1-4, see JVS/AVS Electronic Archives;
www.opuluspress.se/