Ecology Letters, (2007) **10**: xxx–xxx doi: 10.1111/j.1461-0248.2007.01036.x

IDEA AND PERSPECTIVE

Scales of association: hierarchical linear models and the measurement of ecological systems

Sean M. McMahon[1]* and Jeffrey M. Diez[2]
[1]*Nicholas School of the Environment, Duke University, Durham, NC 27708, USA*
[2]*National Centre for Advanced Bio-Protection Technologies, Lincoln University, Canterbury, New Zealand*
**Correspondence: E-mail: seanmcm@duke.edu*

Abstract

A fundamental challenge to understanding patterns in ecological systems lies in employing methods that can analyse, test and draw inference from measured associations between variables across scales. Hierarchical linear models (HLM) use advanced estimation algorithms to measure regression relationships and variance–covariance parameters in hierarchically structured data. Although hierarchical models have occasionally been used in the analysis of ecological data, their full potential to describe scales of association, diagnose variance explained, and to partition uncertainty has not been employed. In this paper we argue that the use of the HLM framework can enable significantly improved inference about ecological processes across levels of organization. After briefly describing the principals behind HLM, we give two examples that demonstrate a protocol for building hierarchical models and answering questions about the relationships between variables at multiple scales. The first example employs maximum likelihood methods to construct a two-level linear model predicting herbivore damage to a perennial plant at the individual- and patch-scale; the second example uses Bayesian estimation techniques to develop a three-level logistic model of plant flowering probability across individual plants, microsites and populations. HLM model development and diagnostics illustrate the importance of incorporating scale when modelling associations in ecological systems and offer a sophisticated yet accessible method for studies of populations, communities and ecosystems. We suggest that a greater coupling of hierarchical study designs and hierarchical analysis will yield significant insights on how ecological processes operate across scales.

Keywords

Bayesian statistics, hierarchical linear models, inference, maximum likelihood, multilevel models, regression, scale, variance components.

Ecology Letters (2007) 10: 1–16

INTRODUCTION

Scale is essential to the analysis of ecological systems. The relationship between two variables in a natural system can be obscured by other variables at other scales (Wiens 1991; Maurer 1999), and the inferences drawn from an observed relationship can be distorted or even reversed depending on the scale at which that relationship is measured (Wiens 1991; Denny *et al.* 2004; Cadotte & Fukami 2005). For this reason, there have long been calls to incorporate scale explicitly in designing, analysing and drawing inference from ecological studies (Allen & Starr 1982; O'Neill *et al.* 1986; Rahel 1990. Wiens 1991; Holling 1992; Levin 1992, 2000). Although a great deal of work has addressed quantitative methods for measuring scale (Borcard *et al.* 1992; Thrush *et al.* 1997; Dale 1999; Dungan *et al.* 2002; He & Legendre 2002; Borcard *et al.* 2004; Harte *et al.* 2005), it is remarkable that so few ecological studies incorporate scale in the analysis of observed natural patterns or experiments. As the importance of scale in determining ecological patterns has become more apparent (Levin 1992, 2000; Harte *et al.* 2005) techniques explicitly designed to measure and interpret interactions and associations at different scales will better enable the generalization of these analyses to other systems and the predictive application of the results to future system behaviours (Underwood & Chapman 1996; Noda 2004).

Scale in general and hierarchical approaches to scale in particular have rich histories in ecological theory, observa-

© 2007 Blackwell Publishing Ltd/CNRS

tion and experimentation. Ecological data are often hierarchically structured, a fact that arises both from common sampling designs (e.g. quadrats on transects, plots within blocks) as well as biological phenomena (e.g. species within genera, clonal stems attached to rhizomes, behaviours over time, fish within reaches within watersheds). Hierarchical structure in ecology has over the years inspired treatises on proper experimental design (Hurlbert 1984; Oksanen 2001), statistical analysis (Wu & David 2002; Clark & Gelfand 2006), and broader theoretical and philosophical explorations (Allen & Starr 1982; O'Neill *et al.* 1986; Rahel 1990; Levin 1992, 2000; Whittaker *et al.* 2001; Noda 2004). Over the years, ecologists generally agree scale is important, have offered methods to quantify scale, and have implemented a number of studies that show scale to be important. Unfortunately, most ecological research that is not specifically focused on the issue of scale fails to account for scale in analysis and inference. It is this oversight, the vast gap between the agreed importance of scale and the failure to include scale in analysis, that we hope to address in this paper.

There is no consensus approach to quantifying scale in ecological studies, in part due to different philosophical approaches to scale and in part due to different sampling regimes necessitating different analyses. The methods that are applied generally fall into three categories. The first category consists of methods that primarily determine the scale at which a pattern is evident [e.g. principle coordinates of neighbour matrices (Borcard *et al.* 2004), wavelet analysis (Keitt & Urban 2005), fractal dimensions (Sugihara & May 1990; Keitt *et al.* 1997)]. Although these methods effectively designate the scale at which a response variable shows distinct patterns and are quite effective at capturing scale-dependent patterns along a continuously scaled variable (e.g. a fine-grained time-series, or detailed spatial measurements such as GIS), most require an indexing (ordination) of the variable of interest and generally cannot take into account correlation between the same measurement of predictor variables across scales (Keitt & Urban 2005). They often, therefore, serve as an initial, exploratory or detective approach to data with unknown scale dependencies.

The second category of methods relates to the classical design of experiments. Nested analysis of variance (ANOVA) and mixed models estimated with ordinary least squares (OLS) fall into this category (Benedetti-Cecchi 2001; Chase & Leibold 2002; Benedetti-Cecchi *et al.* 2005; Cadotte & Fukami 2005). We include here studies that use any basic statistical techniques to examine potential effects but apply them at two scales to perform a qualitative analysis (Tolimieri 1995; Gotelli & Ellison 2002). These methods have the benefit of ease of use, ease of interpretability and clarity of result. They fail, however, to be flexible in design (unbalanced data, more complex model constructions and missing data are difficult to contend with). Demonstrating significant differences between response to discrete treatments at a given spatial scale (e.g. using nested ANOVA) does not easily translate into generating predictions (Clark 2003b, 2005; Moran 2003). Analyses that compare independent studies done at two different scales cannot control for latent differences in compared systems, fail to account for correlations shared by scales, and fail to estimate variation at each scale given variation at the others (Underwood & Petraitis 1993).

The third category includes a suite of modelling methods which are built upon classical statistical approaches (such as likelihood), but which have advanced further in recent years because new computational power allows investigators to analyse more complex, flexible and robust models. In this category we would place more general models that include variance component estimation (Searle *et al.* 1992; Edwards 2004), multilevel models (Buckley *et al.* 2003; van de Pol & Verhulst 2006) and hierarchical Bayesian models (Clark 2003a; Hooten *et al.* 2003; Gelman *et al.* 2004; Helser & Lai 2004; Clark *et al.* 2005). Hierarchical linear models (HLM), the focus of this paper, relate to all three of these methods as they offer a specific model structure within the hierarchical Bayesian context, a generalization of the mixed models, and specialize in estimating variance components. Although HLM can be estimated using maximum likelihood or Bayesian approaches, iterative computational techniques are required for either estimation method [expectation-maximization (EM) algorithm (Dempster *et al.* 1977) or Gibbs sampler (Gelfand & Smith 1990) respectively]. Further, although estimated with sophisticated algorithms, the structure, lexicon and analysis of HLM use the common language of regression analysis. Results and predictions can be communicated across systems and research programmes. HLM has been applied to ecological problems related to community interactions (Vazquez & Simberloff 2004), species–area relationships (Storch *et al.* 2005), habitat covariates in species count data (Thogmartin *et al.* 2006), age-dependent reproduction (van de Pol & Verhulst 2006) and spatial covariance (Gering & Crist 2002; Berk & de Leeuw 2006). Applications of HLM in ecology, however, can benefit from a protocol of analysis that builds models towards a clear concept of the role of scale in a system, and incorporate into the analysis important diagnostic measures of variation and association.

In this paper we demonstrate how HLM both identifies important scales of information and measures associations that explain the information at those scales, developing this in the conceptual and mathematical framework of linear and generalized linear regression, and then demonstrate a variety of models that can be built within this framework. This method is conceptually accessible to a wide range of ecologists with a wide range of statistical experience. Our

© 2007 Blackwell Publishing Ltd/CNRS

goal, therefore, is not merely to offer another approach to ecologists interested in scale issues. We hope to show that ecological studies with no explicit treatment of scale must justify that omission. We also hope to offer a method that is readily applicable to studies with classic sampling and experimental designs. This method can be applied to any nested question, whether temporal, spatial, phylogenetic, geographical, ecological or physiological in nature. In the following sections, we introduce the basic mathematical structure of hierarchical models and use two examples to show how HLM can be fully exploited to draw inference beyond that possible using other approaches. To firmly establish the core application of HLM, we apply maximum likelihood methods to estimate the parameters in a two-level linear model that describes the association between the amount of herbivore damage to plant leaves, plant size and the species richness of the patches in which the plants grow. In order to demonstrate the more flexible approaches of hierarchical generalized linear models as well as some of the strengths of using Bayesian techniques to estimate variance–covariance parameters, we present a three-level model exploring how biotic and abiotic factors at the individual plant-level, the microsite level, and the population-level influence the probability that an individual plant will flower. We conclude with a discussion of these analyses focusing on the role of scale in inference and a call for expanded incorporation of scale into quantitative analyses of natural systems.

HIERARCHICAL LINEAR MODELS

Hierarchical linear models use nested regression equations to investigate associations between variables at different scales. This accounts for how observations can be related in groups within a hierarchy. HLM can apply hypothesis tests and diagnostic reports that address not only the significance of the relationships between variables at different scales, but also the strength of those relationships and their explanatory power across scales. Although the equations describing HLMs can be generalized to contain multiple predictors and link functions, a basic two-level linear model serves to introduce the core structure of HLM, the parameters that need to be estimated, and the inferences that can be drawn from an estimated model. Further, Raudenbush & Bryk (2002) demonstrate a protocol for model building that incorporates variation at different scales into the analysis. We begin with a description of HLM model structure and then explain the protocol for model building in our first example.

Fundamentally, HLM describes relationships between variables within a hierarchical data set. Variables can be incorporated into the model at the level they were collected or any higher level. This is an important feature of HLM, as the same observations can reflect different mechanisms at different scales. For example, variation in soil moisture measured at a fine scale (1 m^2) might reflect outflow of water through leeching due to soil texture or evaporation due to light characteristics, while at larger scales soil moisture differences (gathered from the same sampling protocol) might reflect input differences such as different weather patterns. One can imagine that soil moisture might negatively correlate with plant growth at a fine scale if higher moisture reflected low-light conditions, but positively correlate with growth at the population level if it reflected higher average precipitation.

The first level (the individual level in our examples) of an HLM in its linear form is a simple regression equation [all notation in this paper follows Raudenbush & Bryk (2002)],

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}, \tag{1}$$

where Y_{ij} is a measured response variable i that is observed in a group j. This response has a group-level intercept β_{0j} and is related to an individual-level predictor variable X_{ij} by the group-level regression coefficient β_{1j}. Why are these β terms 'group' variables? This stems from the fact that the residual error of the estimated relationships between the Y_{ij} response variable to the X_{ij} predictor variables, r_{ij}, is assumed in a simple linear regression model to be distributed independently as normal random variables with a mean of zero and variance σ^2. Because the response variable Y is associated not only with the individual i observations, but is nested within the j groups, the residuals are correlated and cannot be assumed to be independent [to assume so would constitute pseudo-replication (Hurlbert 1984)]. To correct this aggregation in HLM, the first-level relationships are modelled not around an overall intercept and slope, but around the intercept and slope of each of the $j = 1,\ldots,J$ level-2 groups. This corrects for non-independence of the errors generated by the correlation of variables within groups. Doing this however, results not in a single regression, but in j different regression equations. To obtain an overall estimate of the relationships between the response variable and the predictors, we use the j first-level regression coefficients to form two, higher-level regressions:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j}, \tag{2a}$$

and

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{1j}, \tag{2b}$$

where γ_{00} and γ_{01} are the level-2 coefficients for the intercept and slope, respectively, of these level-2 regression models (in other words, the γ parameters are group-level equivalents of the β parameters at the individual level). W_j is a level-2 predictor, and behaves as the X_{ij} does in equation. The level-2 random effects u_{0j} and u_{1j} are assumed to be distributed as multivariate normal with means of zeros and

© 2007 Blackwell Publishing Ltd/CNRS

Table 1 Parameters in the combined multilevel model

Parameter	Description
The model: $Y_{ij} = \gamma_{00} + \gamma_{01}W_j + \gamma_{10}X_{ij} + \gamma_{11}W_jX_{ij} + u_{0j} + u_{1j}X_{ij} + r_{ij}$	
Y_{ij}	The estimated percentage of leaf damage for individual plant i in patch j
X_{ij}	Initial height of individual plant i in patch j
W_j	Species richness in each patch j
γ_{00}	The grand mean of leaf herbivory
γ_{01}	The mean effect of patch species richness on leaf herbivory
γ_{10}	The average slope of the relationship between initial plant height and herbivore damage
γ_{11}	The average effect of patch species richness on the relationship between plant height and herbivore damage
u_{0j}	The effect of patch j on leaf herbivory, holding species richness (W) constant
u_{1j}	The effect of patch j on the relationship between herbivore damage and plant size, holding species richness (W) constant
r_{ij}	The random effects on individual leaf damage

variances of τ_{00} and τ_{11} respectively. The covariance between these random effects is τ_{01}. If we substitute eqns (2a) and (2b) into eqn (1) we get the combined model that simultaneously describes the relationships between all predictors and response variables including their error terms at the two levels (see Table 1 for a detailed explanation of this model):

$$Y_{ij} = \gamma_{00} + \gamma_{01}W_j + \gamma_{10}X_{ij} + \gamma_{11}W_jX_{ij} + u_{0j} + u_{1j}X_{ij} + r_{ij}. \quad (3)$$

The fundamental difference between this combined model and models typical of single-level models is that instead of having independent random errors with constant variance (the r_{ij} term), the error term of eqn (3) takes the form $u_{0j} + u_{1j}X_{ij} + r_{ij}$. We assume now that $r_{ij} \sim N(0,\sigma^2)$ and that $u_j \sim N(0,\tau_{..})$, where τ is the variance–covariance matrix of the u_j terms, whose diagonal elements describe the variance of each u parameter. The τ variance–covariance matrix of the second level models becomes an important set of parameters as it describes between-group variance and determines whether higher-level relationships between variables are needed, significant or explanatory (this will be clarified in the examples below).

The uncertainty ascribed to this modelled system contains random error at the individual level and the group level. The error estimation that partitions uncertainty across groups for both the mean and slope of the level-one model does so by estimating the group level variance of the mean (u_{0j}) and slope (u_{1j}). This model provides a great deal of information about the relationships between predictor variables and the response variable and the scales at which those relationships are found. These error terms mark an essential difference between the OLS approach which requires the deviations from the grand mean to be independent, normally distributed and with constant variance. Because the terms u_{0j} and u_{1j} can differ between groups, their variances are not assumed equal. When these terms are null (there is no group-error variance), this model reduces to an analogue of the OLS regression model. To estimate whether these terms are null, however, we must employ iterative maximum likelihood methods, such as the EM algorithm. Furthermore, by setting various parameters of this combined model to zero, a variety of more specific questions that incorporate the scale components of the system can be tested (Table 2).

EXAMPLES

The great advantage of HLM lies in its ability to estimate parameters in complex models that incorporate scale explicitly in the analysis. Determining environmental influences on population dynamics presents a challenge ideally suited for analysis using HLM. In our examples we explore the possible biotic and abiotic mechanisms that influence plant populations at different scales. These examples use data sets not specifically designed for the purpose of illustrating HLM, yet successfully demonstrate how both naturally occurring biological hierarchies (such as populations of clonal plants) and experimental hierarchies (such as nested sampling designs) can take advantage of HLM analysis.

The two-level and three-level models we constructed were estimated using two different approaches on two different data sets. The two-level model assumes maximum likelihood parameters which were estimated using EM algorithms (Dempster *et al.* 1977; Raudenbush & Bryk 2002) written in MATLAB (The Mathworks, Inc. 2003) and employs traditional hypothesis tests for model diagnosis and interpretation. In the three-level model, we applied a hierarchical Bayes approach that uses a Markov chain Monte Carlo (MCMC) sampling procedure in WinBugs (Spiegelhalter & Best 2000) to estimate model parameters (see Appendix B for code). It is worth noting that tools for implementing HLM are progressing quickly in a number of software packages. For example, PROC MIXED in SAS uses a similar EM algorithm [Singer (1998) provides a clear tutorial on

© 2007 Blackwell Publishing Ltd/CNRS

Table 2 Various models described by hierarchical equations

Model	Description
Full regression model $Y_{ij} = \gamma_{00} + \gamma_{01} W_j + \gamma_{10} X_{ij} + \gamma_{11} W_j X_{ij} + u_{0j} + u_{1j} X_{ij} + r_{ij}$	Describes relationship between the individual leaf herbivory, initial plant height and patch-level species richness
One-way ANOVA with random effects $Y_{ij} = \gamma_{00} + u_{0j} + r_{ij}$	Describes the grand mean of leaf herbivory (γ_{00}), the effects of patch on individual leaf herbivory (u_{0j}), taking into account individual variation in leaf herbivory (r_{ij})
Means-as-outcomes regression $Y_{ij} = \gamma_{00} + \gamma_{01} W_j + u_{0j} + r_{ij}$	Estimates how the mean leaf herbivory for each patch of plants can be predicted by species richness (W_j) taking into account the difference between patch variation in leaf herbivory (u_{0j}) and individual variation in leaf herbivory (r_{ij})
One-way ANCOVA with random effects $Y_{ij} = \gamma_{00} + \gamma_{10}(X_{ij} - \bar{X}_{..}) + u_{0j} + r_{ij}$	Estimates the average patch leaf herbivory, accounting for how the level-1 covariate (initial plant height (X_{ij})) influences herbivore damage within each patch
Random-coefficients regression model $Y_{ij} = \gamma_{00} + \gamma_{10}(X_{ij} - \bar{X}_{..}) + u_{0j} + u_{1j}(X_{ij} - \bar{X}_{..}) + r_{ij}$	Describes leaf herbivory as a function of the average slope of the regression between leaf herbivory and initial plant size ($\gamma_{00} + \gamma_{10}(X_{ij} - \bar{X}_{..})$) with estimates of three error terms: the effect of patch j on the mean level of leaf herbivory (u_{0j}), the effect of patch j on the slope of the regression relationship between leaf herbivory and initial plant size β_{1j} ($u_{1j}(X_{ij} - \bar{X}_{..})$, and the individual variation in leaf herbivory (r_{ij})

using SAS for HLM]. Also, there are several packages in R (R Development Core Team 2006) (most notably, lme4) that estimate HLM parameters [see package details and Gelman & Hill (2007) for description of using R and Bugs for fitting multilevel models]. The program 'HLM' is a standalone application (Raudenbush *et al.* 2004) for HLM.

We choose not to debate the relative merits of Bayesian and frequentist methods in this paper, although the distinction can be important, especially as the hierarchical Bayesian approach accurately distinguishes between error in models and biological variation in models (Raudenbush & Bryk 2002; Clark 2005). We refer the interested reader to work that explicitly and effectively tackles this issue in statistical ecology (Ellison 1996, 2004; Clark 2005; Clark & Gelfand 2006). This paper instead focuses on the importance of analysing hierarchically structured data in general. We include both estimation methods to show how either approach offers insight into the scale of ecological processes, while acknowledging a growing interest in development and estimation of ecological data with Bayesian approaches.

TWO-LEVEL MAXIMUM LIKELIHOOD MODEL

The data used in this example are a portion of a larger demographic study. The models constructed here are designed only to advance an understanding of HLM and not address issues in plant community ecology. Here we explore the possible relationship between several characteristics of an understory forest herb, *Eurybia chlorolepis* (Asteraceae) (Burgess) Nesom (1994), and its microenvironment with the interest of identifying associations with patterns of leaf herbivory. Why would this problem merit a hierarchical approach? First, the plant of interest has a natural hierarchical structure. A genetically distinct plant (a genet) is comprised of many stems (ramets) often connected by rhizomes. Ramets show variation in size, stage, herbivore damage, phenology, etc. Second, a protocol that includes the collection of such fundamental environmental variables as species richness, soil moisture and canopy openness will necessarily require sampling at a scale above the individual ramet of the plant. Plots in which these variables were sampled were designed to correspond to patches of the stems of the plant. In this way, initial study questions incorporating questions of scale are linked to sampling design and subsequently hierarchical analysis. Any questions that implicitly or explicitly address genotype, phenotype, habitat traits or demographics should distinguish between patterns and associations at the scale of the ramet and the genet. Reproductive stems vary significantly in size, even within genets; they can grow between 10 and 50 cm and have between 3 and 25 leaves.

© 2007 Blackwell Publishing Ltd/CNRS

Herbivory on the leaves of *E. chlorolepis* during the summer by a host of arthropods and molluscs can influence both reproductive output in any one year and age-class structure in the following year (S. McMahon unpublished data). The relationship between stem height and herbivore damage, measured on the ramet scale, indicates two potential processes. If the relationship is positive, it would indicate that herbivores might key in on healthy, large stems. If the relationship is negative, it might indicate that plants can outgrow herbivore loads and therefore taller stems would display proportionally less herbivore damage. Soil moisture at the patch level can increase mollusc activity (S. McMahon personal observation). In the dark understory (average canopy cover > 90%), light levels raise air temperature and therefore can increase arthropod activity as well as promote plant growth. Plant species richness at sites indicate overall microsite soil quality (generally less acidic soils in an otherwise highly acidic soil profile correlate with higher species richness).

Species, study site and protocol

Eurybia chlorolepis (Asteracea) is an understory perennial herb that grows in cove forests in the southern Appalachian mountains of the United States. Ramets emerge from rhizomes in the early spring as either a juvenile form (rosettes) or reproductive form (with internodal stems). The reproductive ramets grow through the summer and, if conditions allow, flower, are pollinated, and set seed in the fall (September through November). Leaf herbivores can cue on a range of plant features, such as palatability and biomass, and also respond physiologically and demographically to environmental variables (Price 1997). Thus, endogenous and exogenous factors can influence herbivore activity. Yet these processes often operate at distinct scales. In order to detect how herbivores respond to and are driven by these cues, a hierarchical design can organize the many potential mechanisms that influence herbivore activity.

This study tagged 10 stems in each of 20 plots (designed to correspond to the patch scale) and measured the herbivore damage to leaves on each stem by visually estimating a per cent damage to every leaf on every stem. Damage was then averaged for each stem. Stem height and leaf length are good proxies for plant biomass therefore reflect available resources for folivores. These variables were measured in late-summer (at peak growth). Abiotic factors can influence a plant's palatability and therefore herbivore load, but abiotic variables can often not easily (or accurately) be collected at the scale of the plant stem. Therefore, at the patch-scale, soil moisture was measured using gravimetric water content methods, a basic soil chemistry assay was conducted through an agricultural service, and light transmittance was estimated from the analysis of hemispherical photographs using Gap Light Analyzer software (Frazer *et al.* 1999). Species richness was also recorded for all patches.

Model building, parameter estimation and hypothesis testing

The percentage of leaf herbivory in September was chosen as a response variable to see if damage to leaves differed between patches and was predicted by biomass of the plants or environmental variables. Constructing a hierarchical model, unlike a simple linear model, explains variation in the response variable differently at different scales and therefore requires an assessment of the scale at which variation in the response variable occurs. By fitting what Raudenbush & Bryk (2002) term an 'unconditional model', which is effectively a one-way ANOVA model where the levels of the data hierarchy are the treatments of the single factor, we can establish this baseline of variation. The 'combined model', the analogue of eqn (3) is:

$$\mathrm{HERB}_{ij} = \gamma_{00} + u_{0j} + r_{ij}. \quad (4)$$

Here, the per cent of herbivore damage for an individual plant i in a specific patch j can be modelled as an overall average of the damage to every plant in every patch (the 'grand mean' γ_{00}) plus some difference between the average herbivore damage to plants in that jth patch from that overall mean (u_{0j}) plus the difference between the damage to that individual plant and its patch mean (r_{ij}). In the terminology of HLM, patches in this study correspond to the group level of the models. Thus, the variance component of every plant has two parts, the individual variance (taking into account group-variance) and group variance (taking into account individual variance). This simple formulation offers a base understanding of variation in a hierarchical system. Although rudimentary in the context of this problem, this basic understanding of the variation of a simple response variable is almost universally overlooked in ecological studies with hierarchical designs.

After assessing the scale of variation in the response, a model can be built to explain that variation in the two scales of the response variable (in this case, that of the individual stems and that of the patches). How this is done should depend directly on the distribution of the variance components discovered in the first model. For example, two distinct but not mutually exclusive additions to this 'unconditional model' could include a 'random effects' formulation or a 'means-as-outcome' formulation (Table 2). A random effects model includes covariates at the individual level and would not explain group level variance, while a means-as-outcome model would include covariates at the group level to explain the intercepts among the groups, but not within them. We begin with the random effects model. Using plant height as a predictor of leaf herbivore damage,

© 2007 Blackwell Publishing Ltd/CNRS

and remembering that modelling the relationship between plant height and herbivore damage within groups ($\beta_{1j}X_{ij}$) becomes $\beta_{1j} = \gamma_{10} + u_{1j}$ at the second level. The complete random effects model is

$$\mathrm{HERB}_{ij} = \gamma_{00} + \gamma_{10}(\mathrm{HEIGHT}_{ij}) + u_{0j} + u_{1j}\mathrm{HEIGHT}_{ij} + r_{ij}. \quad (5)$$

This model posits a series of relationships that combine to describe the herbivore damage to an individual stem given the height of that stem. Across-patch characteristics of herbivore damage are captured by the overall average damage to all plants γ_{00} (given the new regression relationship included in the model) and the deviation of this plant's patch intercept from that grand intercept (u_{0j}). The relationship between an individual's height and the amount of herbivore damage exhibited is partitioned into two components: first is the across-patch slope relating herbivore damage to plant height multiplied by the individual's height ((γ_{10}*HEIGHT$_{ij}$); the second is the difference between the slope of the within-patch relationship between plant height and herbivore damage multiplied by that plant's height (u_{1j}*HEIGHT$_{ij}$). Finally r_{ij}, which is now the residual error, takes into account deviation from the expectation of this individual given all of the above model components.

The second basic expansion on the unconditional model estimates effects of predictor variables at the second level (the patch level). It is termed the 'means-as-outcomes' model (Table 2) because the explanatory variables are set up to explain variation in group means of the response variable and not variation in individual observations of the response. If we regress herbivore damage of individual plants on patch-level soil moisture, for example, we get this combined model:

$$\mathrm{HERB}_{ij} = \gamma_{00} + \gamma_{01}\mathrm{SM}_j + u_{0j} + r_{ij}. \quad (6)$$

Here, we again have an individual's herbivore damage explained first by the grand intercept of herbivore damage (γ_{00}). The regression term ($\gamma_{01}\mathrm{SM}_j$) contains the relationship between patch-average herbivore damage and the soil moisture at each site. The u_{0j} term is the residual difference between average site herbivore damage and the overall across-site damage, taking into account site soil moisture. The r_{ij} term is the difference between the herbivore damage to an individual stem and the average damage within that stem's site. Its variance should not have changed from the unconditional model as no predictors were set up to explain that variance.

These two models, the random effects and means-as-outcome models, are easily combined or expanded to construct more sophisticated models (Table 2). Indirect effects can be modelled as second-level predictors of first-level slopes (in other words, second-level predictors can be used to predict the relationship between first-level variables and not the response group averages as in the means-as-outcome variables). Variance components can also be modelled with diagnostics of covariances.

For this example, we applied a straightforward model-building design and estimated parameters in all models using an EM algorithm (Raudenbush & Bryk 2002) run in MATLAB (The Mathworks, Inc., 2003). For the unconditional model, the parameters of interest were σ^2 and τ_{00}, the first- and second-level variance components, respectively, as these describe how herbivore damage Y_{ij} varies both from plant to plant and patch to patch. After estimating the unconditional model, it was determined that herbivore damage did show differences across the patch (quadrat) scale (see Results and interpretation below). In order to explain those differences, two separate models were estimated. First, to determine whether the size of the plant (a stem-level variable) predicted differences in herbivore damage, a random-coefficients regression model (Table 2) was estimated regressing herbivore damage against mid-summer plant height. The main parameter of interest in this level-1 model was β_{1j}, which describes the relationship between plant height and herbivore damage in each group j. At the group level, the parameter γ_{10} describes the average slope of this relationship and u_{1j} describes how the slope of this relationship varies from group to group around that average (the residuals after accounting for the overall average slope). The level-2 predictors were plant species richness, canopy openness, soil moisture, soil pH and cation exchange capacity (CEC). Because there was no significant relationship between the plant height and herbivore damage (see Results and interpretation below), plant height was removed from the subsequent models, and the level-2 predictors were included in a means-as-outcome model (Table 2).

Results and interpretation

In every model, the EM algorithm successfully converged in under 400 iterations. The unconditional model estimated the variance at the first level, σ^2, to be 327.85. The variance at the second level τ_{00} was 364.19, which was significantly different from 0 ($\chi^2 = 478$, df = 16, $P < 0.001$). This indicates patch level variation in herbivore damage to plants. To better quantify this variation, we determined the proportion of variance in the system that is described by the patch level as the interclass correlation coefficient: $\rho = \tau_{00}/(\tau_{00} + \sigma^2)$. In this model $\rho = 0.565$, indicating that over 56% of the total variation in herbivory exists between patches of plants and not within them. From this starting point, we can try to explain this variation at each level.

Late-summer plant height was not related to herbivore damage ($t = -0.658$, $P = 0.320$) (see Fig. 1 for confidence intervals of all models). The longest leaf of the plants was

© 2007 Blackwell Publishing Ltd/CNRS

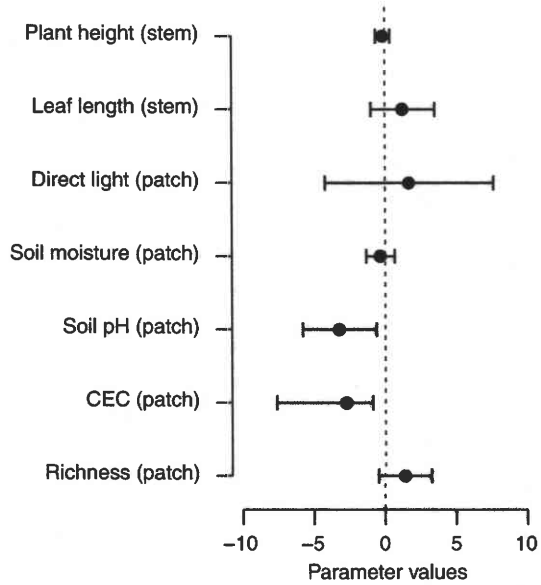

Figure 1 Predictors of leaf herbivore damage. Solid lines represent 95% confidence intervals for estimated effects of variables at three levels of the model. Dots represent point estimate of parameter values.

also not associated with herbivore damage ($t = -1.187$, $P = 0.191$). In the means-as-outcome model, direct solar irradiance and soil moisture were not associated with patch level herbivore damage ($t = 0.612$, $P = 0.321$ and $t = -1.359$, $P = 0.155$ respectively). Species richness also showed no influence on herbivore damage ($t = -0.658$, $P = 0.312$). Two variables from the soil chemistry assay did show a significant influence. Soil pH showed an overall negative association with herbivore damage ($\gamma_{10} = -32.831$, $t = -2.602$, $P = 0.020$). CEC also showed a negative relationship ($\gamma_{10} = -31.628$, $t = -2.673$, $P = 0.018$). To assess how well these patch-level variables explain patch-level variation in herbivore damage, we used the difference between calculations of $\hat{\tau}_{00}$ done in the unconditional model (UNCON) and in the means-as-outcome design with the significant variables. This is done by calculating $(\widehat{\tau_{00}}(\text{UNCON}) - \widehat{\tau_{00}}(\text{SOIL pH}))/\widehat{\tau_{00}}(\text{UNCON})$ or $364.190 - 122.158/364.190 = 0.664$. Over 65% of inter-patch variation in herbivore damage is explained by the sampled soil pH of the patch. Using a similar calculation we find that CEC reduced the patch variance to 128.624, or a 64.7% explanation.

What this hierarchical approach to these community relationships offers that simple linear models do not is the estimation of regression coefficients across scales, in addition to partitioning variation. Every relationship identified takes into account the scale of the pattern explained. In this example, we find that variation in herbivore damage is roughly split equally within and between patches. On the stem-level, no physical features of the plant predict herbivore damage. This is not surprising as at that fine scale, herbivores likely respond to cues that are not reflected in coarse measurements like height and leaf length. At the patch scale, two variables associated with soil chemistry and nutrient uptake, soil pH and CEC, both showed negative influences on herbivore damage. Soils in the study site are generally very acidic (*c.* 4.0 pH) and nutrient poor (as indicated by low CEC). The benefit of higher pH and CEC is greater nutrient retention. A more nutrient-rich site might enable plants to produce compounds that reduce herbivore damage. We now look at how more complex models can be organized using HLM.

THREE-LEVEL HIERARCHICAL BAYES MODEL

In the second example we are interested in understanding the scales of variability of the probability of flowering of a terrestrial orchid, *Tipularia discolor* (Orchidaceae), in the south-eastern United States. Factors at multiple scales can influence plant flowering, from individual level traits, to microsite variation in abiotic resources and biotic interactions, to larger population level canopy, soil or topographic effects. Although there is often significant variation from plant to plant in the likelihood of flowering, flowering synchronicity at different scales can be observed in many plant populations, suggesting the need to explicitly explore possible mechanisms at a range of scales (Crone & Lesica 2004; Satake 2004).

While sharing the same basic multilevel structure as the two-level example, this three-level model builds on the two-level in important ways. First, as a Bayesian model, all parameters of the model are considered random variables to be estimated (Gelman *et al.* 2004), and as such they are given prior distributions that are updated by the data to yield full posterior probability distributions. Second, as with many ecological data sets, flowering is a discrete response, benefiting from a generalized linear model framework.

To limit confusion about terminology, it is important to be clear that the term 'hierarchical' in 'hierarchical Bayes' refers to the assignment of probability models to model parameters. These parameters that describe the distribution of the random parameters we estimate (e.g. β,σ) are termed hyper-parameters. The 'hierarchical' in 'hierarchical linear models' refers to the structure of the data used in the model. These distinctions can be seen clearly in Fig. 2, a conceptual description of the three-level model developed in this example. The data structure in such models need not always be strictly nested (e.g. measurements of individuals may be nested within both populations and years, but population and year are not themselves nested hierarchically), suggest-

© 2007 Blackwell Publishing Ltd/CNRS

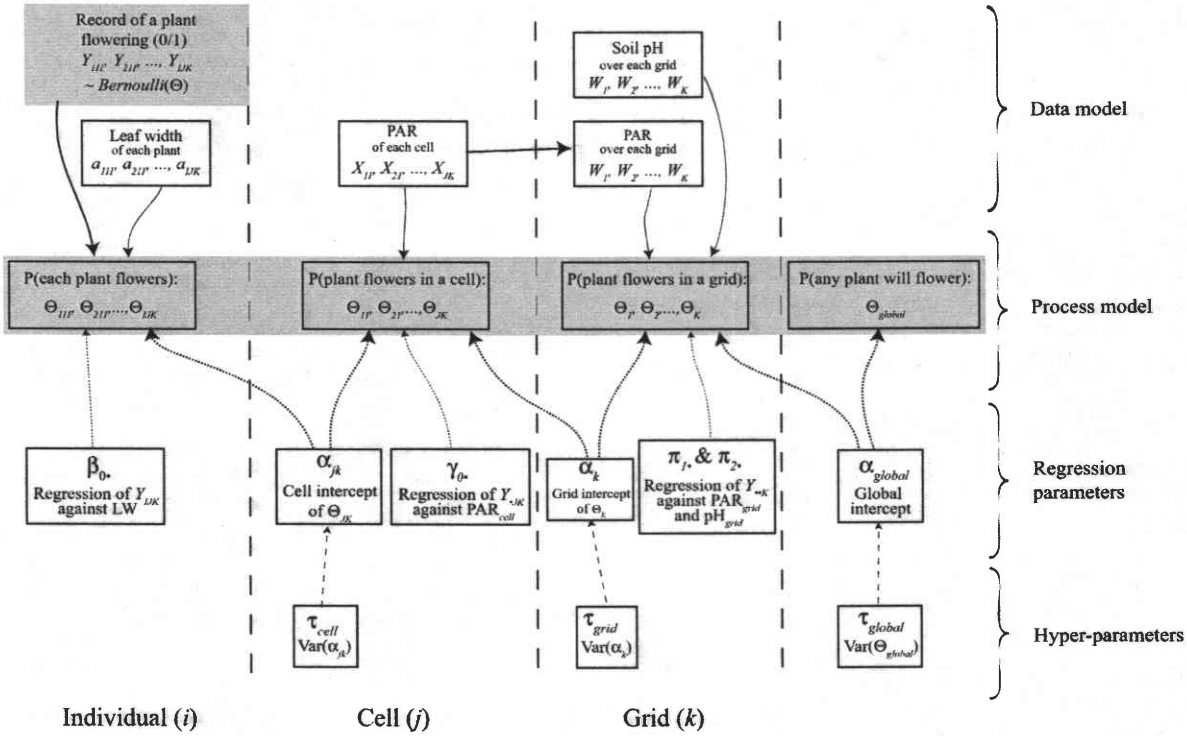

Figure 2 This conceptual map of the model roughly follows Clark & LaDeau (2006), but illustrates how the model spans levels of both parameter organization (vertical in figure) and ecological levels of organization (horizontal in figure). In order to make more clear the meaning of the model structure, not all estimated parameters and hyper-parameters are included. Further, merely the scale notation is applied to parameters and the numerical identification of the predictors (the P, Q, and S designation). For the complete model, see WinBugs code in Appendix B.

ing the more general use of the term 'multilevel' for models without strictly nested designs (Gelman & Hill 2007).

Species, study site and protocol

Tipularia discolor is a wintergreen terrestrial orchid found in mixed deciduous forests of eastern North America. In *T. discolor*'s southern range, a plant's single green leaf emerges above ground in late fall (end of September) and remains until spring (March–April). Flowers, if produced, are found on a single flowering stalk that emerges in August, before the leaf emerges. Study grids, designed to capture population level processes, range from 250 to 480 m^2 in size, each divided into 4 m^2 cells within which all plants were individually marked. Thus, the levels of the experiment, designed to reflect *a priori* ideas about interesting levels of organization in the system, were the i individual plants (flowering or not), the j cells that reflect microsite characteristics, and the k population-level grids that contain a range of microsites. For this study, we use floral surveys from the late fall of 2004. We test the overall hypothesis that biotic and abiotic variables at different scales help to explain the probability of flowering. Specifically, plant size at the level of individuals, microsite soil moisture and understory light levels, and soil pH and texture at the grid level are hypothesized to be important for plant flowering based on our biological understanding of the species (see Appendix A for a detailed explanation of sampling protocol). The predictor variables at the individual, cell and population levels are designated by p, q and s, respectively, and the intercepts are designated 0.

Model building and parameter estimation

The binary flowering data for this study are modelled at the individual level (each plant is observed to be flowering or not), so we use a Bernoulli sampling distribution, such that

$$Y_{ijk} \sim \text{Bernoulli}(\phi_{ijk}), \tag{7}$$

where Y_{ijk} is the flowering status of each plant i in cell j and population k, and ϕ_{ijk} represents the probability of flowering for each plant.

We use a logit transformation to obtain linearity of parameters. The first-level model is

© 2007 Blackwell Publishing Ltd/CNRS

$$\eta_{ijk} = \log\left(\frac{\phi_{ijk}}{1-\phi_{ijk}}\right) = \beta_{0jk} + \sum_{p=1}^{P} \beta_{(p)jk} X_{(p)ijk}. \quad (8)$$

The η_{ijk} represents the log-odds of flowering for individual plants and is related directly to a cell-level intercept β_{0jk} and P individual-level predictor variables, $X_{(p)ijk}$, in this case just the size of the plant. We chose a varying-intercept model here, with intercept terms β_{0jk} that vary among cells, but a single regression coefficient, $\beta_{1..}$, describing the effect of plant size on flowering across sites because there is no *a priori* reason to suspect a different inherent relationship between size and flowering across sites.

At level-2, the cell-level intercepts are modelled as a function of population-level intercepts and cell-level covariates. This gives a second-level model

$$\beta_{0jk} = \gamma_{00k} + \sum_{q=1}^{Q_0} \gamma_{(0q)k} W_{(q)jk} + u_{(p)jk}, \quad (9)$$

where γ_{00k} is a population-level intercept, with a vector of Q predictor variables, $W_{(q)jk}$ for every p β (in this case, we are only modelling intercepts, so Q_p is really just the Q predictors for the intercept). These Q predictors include moisture and light availability, with coefficients $\gamma_{(0q)k}$, and a normally distributed random cell effect, u_{0jk}, with a mean of zero and a variance τ_β estimated by the data.

Finally, the population-level model (level 3) regresses the second-level $\gamma_{(pq)}$ parameters against population-level variables. Because all γ parameters are associated with the first-level intercepts (β_{0jk}), they all begin with '0', followed by their 'q' designation:

$$\gamma_{(0q)k} = \pi_{(0q)0} + \sum_{s=1}^{S_{pq}} \pi_{(0qs)} Z_{(s)k} + e_{(0q)k}, \quad (10)$$

where the population intercept, γ_{0qk}, is determined by a global intercept $\pi_{(0q)0}$. The vector of S population-level covariates, $Z_{(s)k}$, include pH and per cent sand. Their relationship to the γ parameters is described by the regression coefficients $\pi_{(0qs)}$, or in this example, π_{0qs}.

The residuals from these relationships are found in the level-3 random effect $e_{(0q)k}$. As before, the random effect is considered normally distributed with mean 0 and variance τ_γ, representing the population level portion of the model variance. We have 16 measurements of pH and per cent sand per population, and thus only assume that we capture population average pH values with some variance: $pH_{(s)k} \sim$ Normal(pH_{0k},τ_{pH}). All regression coefficients were given non-informative prior distributions.

The level-1 variance allowed by the Bernoulli sampling distribution is $\tau_{ijk} = \phi_{ijk}(1-\phi_{ijk})$. Because with logistic regressions, this individual-level variance is heteroscedastic (varying across parameter values), comparison with higher level variances is not as straight-forward as with Normally distributed error models. We adopt the approach of Snijders & Bosker (1999) to assume that the level-1 random effect is logistically distributed with mean 0 and variance $\pi^2/3$. This value can then be used in calculations of variance partitioning across different levels, as demonstrated in Table 3. To explore the predictive effects of the important higher-level explanatory variables, we partition the variance explained by these predictor variables at each level of the model (following Gelman *et al.* (2004). As with a classical R^2 the proportion of variance explained by each level of a model is defined as

Table 3 Variance components of three-level model

Calculation	Description
$\frac{\pi^2/3}{(\pi^2/3+\tau_\beta+\tau_\gamma)}$	Proportion of variance at level 1
$\frac{\tau_\beta}{(\pi^2/3+\tau_\beta+\tau_\gamma)}$	Proportion of variance at level 2
$\frac{\tau_\gamma}{(\pi^2/3+\tau_\beta+\tau_\gamma)}$	Proportion of variance at level 3

$$R^2 = 1 - \frac{E(V_{k=1}^{K}\epsilon_k)}{E(V_{k=1}^{K}\theta_k)} \quad (11)$$

where ε_k are the residual errors of the K groups at the given level, θ_k are the intercept terms, v is the variance of the K groups, E is the posterior mean or expectation (Gelman & Pardoe 2005). Performed at the cell and population levels of the model, this measure gives us a per cent of the variation at each higher level of the model that is explained by the predictors at that level. All level-3 model parameters are described in Table 4.

The Bayesian models were fit using an MCMC sampling method run in WinBugs 1.4 (Spiegelhalter & Best 2000), and we used the R computing package (R Development Core Team 2006) for calculating R^2 following Gelman & Pardoe (2005).

A note on model selection. Although we built these models based on parameter estimates, scoring models for selection may also be used. The deviance information criteria (DIC) was developed to estimate the penalty term in hierarchical models (Spiegelhalter *et al.* 2002; Gelman & Pardoe 2005). This is complex because the total number of parameters estimated can equal (P*Q*S) + 2 (regression parameters and variance components) in addition to hyper-parameters for these parameters, yet all of these parameters share a great deal of information and so cannot be so easily captured in a simple penalty term (as in AIC or BIC). The DIC seeks to account for this paradox, but its implementation proves challenging and remains somewhat controversial (see Spiegelhalter *et al.* 2002 and subsequent discussion). Because of the mixed reviews of DIC and our belief that building models with parameters and variables instead of reducing

© 2007 Blackwell Publishing Ltd/CNRS

Table 4 Parameter estimates for three-level logistic models

Parameter	Mean estimate	Lower interval	Upper interval
Unconditional model			
$\sigma^2_{individual}$	$\pi^2/3$		Constant
σ^2_{cell}	1.58	0.37	3.86
σ^2_{grid}	4.80	0.64	23.32
$\rho_{individual}$	0.34		
ρ_{cell}	0.16		
ρ_{grid}	0.50		
Conditional model: leaf width, light, pH			
$\sigma^2_{individual}$	$\pi^2/3$		Constant
σ^2_{cell}	15.16	4.37	45.34
σ^2_{grid}	2.77	0.0042	16.58
Regression coefficients			
β_{lw}	1.55	1.11	2.06
γ_{par}	1.71	−0.636	4.74
π_{ph}	−2.94	−5.61	−1.12
Per cent variation explained (R^2)			
Cell level	0.31		
Grid level	0.78		

complex models to single numbers better appreciates biological processes operating in modelled systems, we do not include DIC scores in these results.

Bayesian results and inference

All inference from the models comes from the posterior distributions for the parameters and the variance diagnostics. The posterior parameter distributions are summarized in Table 4 for the unconditional and fully conditional models. The size of individual plants had a positive effect on flowering probability, with a mean effect of 1.89 and over 95% of the mass of the posterior of β_1 located above 0 (Fig. 3). Given the logit link function, this is an effect on the log-odds of flowering. At the cell level, light availability had a positive effect on flowering, with a mean of 1.23 and its 95% probability interval slightly overlapping zero. At the population level, soil pH was significantly negatively related to flowering probability, with a mean of −42.32 and 95% interval from −4.15 to −0.82. The 95% intervals of all other explanatory variables substantially overlap zero and thus are not considered likely to differ from zero (Fig. 3).

The variance analysis of the unconditional model showed 34% of the variation among individuals, 16% among cells and 50% at the population level. Using the R^2 approach of Gelman & Pardoe (2005) for estimating the explanatory power of the covariates, we find that light levels explain 31% of the variation at the cell level, and pH explains 78% of the variation at the grid level. Leaf width is clearly an important explanatory variable at the individual level, given

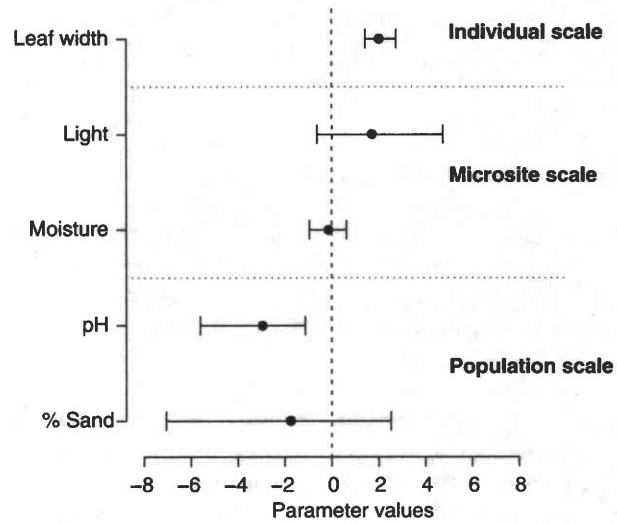

Figure 3 Scale-explicit coefficient estimates. Solid lines represent 95% posterior credible intervals for estimated effects of variables at three levels of the model. Those intervals not overlapping the zero line may be considered significantly different from zero. Leaf width of individual plants, microsite availability of light and moisture, and population level soil pH and % sand content are considered. Light refers to winter PAR readings.

its posterior distribution significantly different from zero, but because individual level variance is the constant $\pi^2/3$ (Snijders & Bosker 1999), the per cent variation explained by individual-level covariates in GLM models cannot be well estimated.

Model interpretation

As with the two-level model, our inference about the questions of ecological interest in the three-level model benefit from a hierarchical approach. From our initial calculations of variance partitioning we learn that as much of the variability in flowering resides among populations as within (50% of the variation in flowering probability exists at the population level), and most of the variation within populations is found among individuals (34%). At this individual level, plant leaf size has a strong influence on the probability of flowering. Variation in light availability within populations helps explain 31% of the variation from cell to cell, but differences in light availability among populations do not contribute much to explaining overall differences in flowering. This supports the ecological hypothesis that within-population variation in canopy light transmittance, such as light gaps, is more influential than average light transmittance differences among populations, which is plausible as these populations were all located in full canopy forest sites. However, soil pH differences among the populations explained 71% of the variation that we see in

© 2007 Blackwell Publishing Ltd/CNRS

flowering at this larger scale. Given the correlative nature of this study, we cannot attribute the effects to soil pH in any mechanistic sense, but it is an interesting finding nonetheless and suggestive that edaphic factors that vary on the scale of populations (e.g. litter layer composition) are important in determining reproductive behaviour of these orchids. It is also important to recognize that without the explicit incorporation of scale into this analysis of flowering probability, a researcher measuring light and flowering probability from a microsite perspective (quadrat to quadrat) might overemphasize the importance of local light availability for flowering probability.

These inferences, taken together, illustrate how important it is to measure the scale at which a life-history trait varies and further record the scale at which biotic and abiotic components of the system explain that variation. One might suspect that with greater future deployment of microsensor technologies for measuring the environment at finer spatial and temporal scales, the ability to explore a range of scales may increase dramatically, and using a statistical framework that can accommodate multilevel analysis will be critical. In this study, for example, our exploration of light effects were constrained to the patch level at which we could make measurements. Were we able to measure light availability at the scale of individual plants, it would be interesting to explore the fine-scale importance of light relative to leaf size in a combined model incorporating that same information at the grid and population scales.

TOWARDS A BROADER STUDY OF SCALE IN ECOLOGY

Two paradigms have traditionally guided the discussion of scale and ecological systems: one describes ecological patterns as fundamentally scale-invariant (Harte *et al.* 2005; Marquet *et al.* 2005), while the other focuses on ecological patterns as hierarchical and distinct among scales (Wu & David 2002; Leibold *et al.* 2004; Noda 2004; Takada & Miyashita 2004). Clearly both of these conceptualizations of ecological systems are appropriate depending on the question being asked (O'Neill *et al.* 1986). Regardless of the paradigm, however, ecologists need to explain mechanisms that influence patterns at different scales (Huston 1999). In the case of scale-invariant systems, power law relationships, whether derived from a single process (Marquet *et al.* 2005) or multiple processes across scales (Allen *et al.* 2001), will remain only an intriguing mathematical artefact until specific mechanisms can be identified that explain why the association between two variables does not differ across scales. Investigating patterns of species diversity, for example, will entail measuring associations with species richness at different scales in order to develop a common description of the number of species observed and the area considered and potential explanatory variables (Gotelli & Ellison 2002; Lyons & Willing 2002). Because HLM can explore the same variable at different scales, interactions between variables, and describe the uncertainty in these relationships, it is well suited for such an inquiry.

The hierarchical paradigm of ecological systems requires even closer attention to associations across scales, as not only might the mechanisms change with scale, the inferences drawn from those relationships might change (Menge 1992; Fukami 2004; Cadotte & Fukami 2005). HLM and the protocol for designing multilevel models we feature offers a powerful tool with which ecologists can explore the associations between environmental and biotic variables at different scales, the strengths of those associations, the covariance between those associations, and the propagation of uncertainty in those relationships across scales. In addition to experimental designs that often structure data in a hierarchical manner, many sub-disciplines are explicitly interested in biological hierarchies. In order to address both the importance of scale in ecological systems and account for unique mechanisms defining patterns at different scales, we believe that researchers should design observational and experimental studies in hierarchical structures. By implementing hierarchical designs, and coupling these with hierarchical analysis, ecologists can better account for and justify the scale of the relationships they discover.

Population ecology, ecological genetics and demography inherently deal with associations among individuals, within and among populations, and the scale of inference about key variables can be crucial to the ecological and evolutionary inferences (Doak *et al.* 1992; Scott *et al.* 2002; Buckley *et al.* 2003; van de Pol & Verhulst 2006). As our understanding of genetic population structure increases across a wide variety of taxa, hierarchically designed studies that connect measured genetic structures to environmental and physiological variables could offer new insights into the way in which evolutionary and ecological processes generate genetic patterns. The mechanisms driving species distributions unfold across environmental gradients at a range of spatial and temporal scales, from individual generations within microsites and populations to longer-term community level shifts over the course of decades. Accounting for scale in such analysis will be essential to any fundamental understanding of the role of ecological niches in structuring biodiversity patterns (Menge & Olson 1990; Pulliam 2000; Chase & Leibold 2003). The study of metapopulations and metacommunities are based fundamentally on a hierarchical approaches to populations (Hanski 1999; Leibold *et al.* 2004). Studies in these fields could benefit greatly from a more explicit incorporation of predictive relationships between variables at the sub- and meta-population scales. The importance of spatial correlation in ecological studies has become manifestly important (Tilman & Kareiva 1997),

© 2007 Blackwell Publishing Ltd/CNRS

and new methods to detect relationships between variables across space are more powerful (Borcard *et al.* 2004; Griffth & Peres-Neto 2006). Output from these methods could provide strong starting points for hierarchically designed studies and explicit predictive models.

The development of ecological theory, inference drawn from empirical studies, and the confrontation of one by the other will be well served by a more expanded use of tools for explicitly analysing scale. As computational power increases and data collection begins to reflect the potential for high-dimensional models, HLM can serve to integrate sub-disciplines, which are often focused around specific levels of organization.

ACKNOWLEDGEMENTS

The authors thank Aaron Ellison, Marti Anderson, Nigel G. Yoccoz and two anonymous referees for valuable comments on drafts of this manuscript. The Department of Ecology and Evolutionary Biology at the University of Tennessee helped support S.M.'s work. The *Tipularia* research was supported by NSF grants DEB-0235371 to H. Ronald Pulliam and to the Coweeta LTER program (DEB-9632854 and DEB-0218001).

REFERENCES

Allen, T.F.H. & Starr, T.B. (1982). *Hierarchy: perspectives for ecological complexity*. The University of Chicago Press, Chicago.

Allen, A.P., Li, B.L. & Charnov, E.L. (2001). Population fluctuations, power laws and mixtures of lognormal distributions. *Ecol. Lett.*, 4, 1–3.

Benedetti-Cecchi, L. (2001). Variability in abundance of algae and invertebrates at different spatial scales on rocky sea shores. *Mar. Ecol. Prog. Ser.*, 215, 79–92.

Benedetti-Cecchi, L., Bertocci, L., Vaselli, S. & Maggi, E. (2005). Determinants of spatial pattern at different scales in two populations of the marine alga *Rissoella verruculosa*. *Mar. Ecol. Prog. Ser.* 293, 37–47.

Berk, R. & de Leeuw, J. (2006). Multilevel statistical models and ecological scaling. In: *Scaling and Uncertainty Analysis in Ecology* (eds Wu, J., Jones, B., Li, H. & Loucks, O.L.). Springer, 67–88.

Borcard, D., Legendre, P. & Drapeau, P. (1992). Partialling out the spatial component of ecological variation. *Ecology*, 73, 1045–1055.

Borcard, D., Legendre, P., Avois-Jacquet, C. & Tuomisto, H. (2004). Dissecting the spatial structure of ecological data at multiple scales. *Ecology*, 85, 1826–1832.

Buckley, Y.M., Briese, D.T. & Rees, M. (2003). Demography and management of the invasive plant species *Hypericum perforatum* i. using multi-level mixed-effects models for characterizing growth, survival and fecundity in a long-term data set. *J. Appl. Ecol.*, 40, 481–493.

Cadotte, M.W. & Fukami, T. (2005). Dispersal, spatial scale, and species diversity in a hierarchically structured experimental landscape. *Ecol. Lett.*, 8, 548–557.

Chase, J.M. & Leibold, M.A. (2002). Spatial scale dictates the productivity-biodiversity relationship. *Nature*, 416, 427–430.

Chase, J.M. & Leibold, M.A. (2003). *Ecological Niches: Linking Classical and Contemporary Approaches. Interspecific interactions*. The University of Chicago Press, Chicago.

Clark, J.S. (2003a). Uncertainty and variability in demography and population growth: a hierarchical approach. *Ecology*, 84, 1370–1381.

Clark, J.S. (2003b). Uncertainty in ecological inference and forecasting. *Ecology*, 84, 1349–1350.

Clark, J.S. (2005). Why environmental scientists are becoming Bayesians. *Ecol. Lett.*, 8, 2–14.

Clark, J.S. & Gelfand, A.E. (2006). *Hierarchical Modeling for the Environmental Sciences: Statistical Methods and Applications*. Oxford University Press, Oxford.

Clark, J.S. & LaDeau, S.L. (2006). Synthesizing ecological experiments and observational data with hierarchical Bayes. In: *Hierarchical Modeling for the Environmental Sciences: Statistical Methods and Applications* (eds Clark, J.S. & Gelfand, A.). Oxford University Press, New York, pp. 41–57.

Clark, J.S., Ferraz, G.A., Oguge, N., Hays, H. & DiCostanzo, J. (2005). Hierarchical Bayes for structured, variable populations: from recapture data to life-history prediction. *Ecology*, 86, 2232–2244.

Crone, E.E. & Lesica, P. (2004). Causes of synchronous flowering in *Astragalus scaphoides*, an iteroparous perennial plant. *Ecology*, 85, 1944–1954.

Dale, M.R.T. (1999). *Spatial Pattern Analysis in Plant Ecology*. Cambridge University Press, Cambridge.

Dempster, A.P., Laird, N.M. & Rubin, D.B. (1977). Maximum likelihood from incomplete data via EM algorithm. *J. Statist. Soc. Ser. B Methodol.*, 39, 1–38.

Denny, M.W., Helmuth, B., Leonard, G.H., Harley, C.D.G., Hunt, L.J.H. & Nelson, E.K. (2004). Quantifying scale in ecology: lessons from a wave-swept shore. *Ecol. Monogr.*, 74, 513–532.

Doak, D.F., Marino, P.C. & Kareiva, P.M. (1992). Spatial scale mediates the influence of habitat fragmentation on dispersal success implications for conservation. *Theor. Popul. Biol.*, 41, 315–336.

Dungan, J.L., Perry, J.N., Dale, M.R.T., Legendre, P., Citron-Pousty, S., Fortin, M.J. *et al.* (2002). A balanced view of scale in spatial statistical analysis. *Ecography*, 25, 626–640.

Edwards, M.S. (2004). Estimating scale-dependency in disturbance impacts: El ninos and giant kelp forests in the northeast pacific. *Oecologia*, 138, 436–447.

Ellison, A.M. (1996). An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecol. Appl.*, 6, 1036–1046.

Ellison, A.M. (2004) Bayesian inference in ecology. *Ecol. Lett.*, 7, 509–520.

Frazer, G., Canham, C. & Lertzman, K. (1999). *Gap Light Analyzer*. Simon Fraser University, Burnaby, BC, and the Institute of Ecosystem Studies, Millbrook, NY, version 2.0 edition.

Fukami, T. (2004). Assembly history interacts with ecosystem size to influence species diversity. *Ecology*, 85, 3234–3242.

Gelfand, A.E. & Smith, F.M. (1990). Sampling-based approaches to calculating marginal densities. *J. Am. Statist. Assoc.*, 85, 398–409.

© 2007 Blackwell Publishing Ltd/CNRS

Gelman, A. & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge.

Gelman, A. & Pardoe, I. (2005). Bayesian measures of explained variance and pooling in multilevel(hierarchical) models. *Technometrics*, 48, 241–251.

Gelman, A., Carlin, J.B. & Rubin, H.S.S.B. (2004). *Bayesian Data Analysis*, 2nd edn. Chapman and Hall CRC, New York.

Gering, J.C. & Crist, T.O. (2002). The alpha-beta-regional relationship: providing new insights into local-regional patterns of species richness and scale dependence of diversity components. *Ecol. Lett.*, 5, 433–444.

Gilks, W.R., Richardson, S. & Spiegelhalter, D.J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London, UK.

Gotelli, N.J.A. & Ellison, A.M. (2002). Biogeography at a regional scale: determinants of ant species density in New England bogs and forests. *Ecology*, 83, 1604–1609.

Griffth, D.A. & Peres-Neto, P. (2006). Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. *Ecology*, 87, 2603–2613.

Hanski, I. (1999). *Metapopulation Ecology*. Oxford University Press, New York.

Harte, J., Conlisk, E., Ostling, A., Green, J.L. & Smith, A.B. (2005). A theory of spatial structure in ecological communities at multiple spatial scales. *Ecol. Monogr.*, 75, 179–197.

He, F.L. & Legendre, P. (2002). Species diversity patterns derived from species-area models. *Ecology*, 83, 1185–1198.

Helser, T.E. & Lai, H.L. (2004). A Bayesian hierarchical meta-analysis of fish growth: with an example for north American largemouth bass, *Micropterus salmoides*. *Ecol. Modell.*, 178, 399–416.

Holling, C.S. (1992). Cross-scale morphology, geometry, and dynamics of ecosystems. *Ecol. Monogr.*, 62, 447–502.

Hooten, M.B., Larsen, D.R. & Wikle, C.K. (2003). Predicting the spatial distribution of ground flora on large domains using a hierarchical Bayesian model. *Landsc. Ecol.*, 18, 487–502.

Hurlbert, S.H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecol. Monogr.*, 54, 187–211.

Huston, M.A. (1999). Local processes and regional patterns: appropriate scales for understanding variation in the diversity of plants and animals. *Oikos*, 86, 393–401.

Keitt, T.H. & Urban, D.L. (2005). Scale-specific inference using wavelets. *Ecology*, 86, 2497–2504.

Keitt, T.H., Urban, D.L. & Milne, B.T. (1997). Detecting critical scales in fragmented landscapes. *Ecol. Soc.*, 1, 4.

Leibold, M.A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J.M., Hoopes, M.F. *et al.* (2004). The metacommunity concept: a framework for multi-scale community ecology. *Ecol. Lett.*, 7, 601–613.

Levin, S.A. (1992). The problem of pattern and scale in ecology. *Ecology*, 73, 1943–1967.

Levin, S.A. (2000). Multiple scales and maintenance of biodiversity. *Ecosystems* 3, 498–506.

Lyons, S.K. & Willing, M.R. (2002). Species richness, latitude, and scale-sensitivity. *Ecology*, 83, 47–58.

Marquet, P.A., Quinones, R.A., Abades, S., Labra, F., Tognelli, M.F., Arim, M. *et al.* (2005). Scaling and power-laws in ecological systems. *J. Exp. Biol.*, 208, 1749–1769.

Maurer, B. (1999). *Untangling Ecological Complexity: the Macroscopic Perspective*. University of Chicago Press, Chicago.

Menge, B.A. (1992). Community regulation -under what conditions are bottom-up factors important on rocky shores. *Ecology*, 73, 755–765.

Menge, B.A. & Olson, A.M. (1990). Role of scale and environmental-factors in regulation of community structure. *Trends Ecol. Evol.*, 5, 52–57.

Moran, M.D. (2003). Arguments for rejecting the sequential Bonferroni in ecological studies. *Oikos*, 100, 403–405.

Nesom, G.L. (1994). Review of the taxonomy of *Astersensulato*(Asteraceae: Astereae), emphasizing the new world species. *Phytologia*, 77, 141–297.

Noda, T. (2004). Spatial hierarchical approach in community ecology: a way beyond high context-dependency and low predictability in local phenomena. *Popul. Ecol.*, 46, 105–117.

O'Neill, R.V., Deangelis, D.L., Waide, J.B. & Allen, T.F.H. (1986). *A Hierarchical Concept of Ecosystems*. Princeton University Press, Princeton, NJ.

Oksanen, L. (2001). Logic of experiments in ecology: is pseudoreplication a pseudoissue? *Oikos*, 94, 27–38.

van de Pol, M. & Verhulst, S. (2006). Age-dependent traits: a new statistical model to separate within-and between-individual effects. *Am. Nat.*, 167, 766–773.

Price, P.W. (1997). *Insect Ecology*, 3rd edn. Wiley & Sons, New York.

Pulliam, H.R. (2000). On the relationship between niche and distribution. *Ecol. Lett.*, 3, 349–361.

R Development Core Team. (2006). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org.

Rahel, F.J. (1990). The hierarchical nature of community persistence: a problem of scale. *Am. Nat.*, 136, 328–344.

Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods. Advanced Quantitative Techniques in the Social Sciences*, Sage Publications, Thousand Oaks, CA.

Raudenbush, S.W., Bryk, A.S. & Congdon, R. (2004). *HLM: Hierarchical and Nonlinear Modeling (Version 6.0) [computer software]*, 6.0 edition. Scientific Software International, Lincolnwood, IL.

Satake, A. (2004). Modeling spatial dynamics of episodic and synchronous reproduction by plant population: the effect of small-scale pollen coupling and large-scale climate. *Popul. Ecol.*, 46, 119–128.

Scott, M.J., Heglund, P.J., Morrison, M.L., Haufler, J.B., Raphael, M.G., Wall, W.A. *et al.* (2002). *Predicting Species Occurrences: Issues of Accuracy and Scale*. Island Press, Washington.

Searle, S.R., Casella, G. & McCulloch, C.E. (1992). *Variance Components*. John Wiley and Sons, New York.

Singer, J.D. (1998). Using SAS PROC MIXED to fit multilevel models hierarchical models and individuals growth models. *J. Educ. Behav. Statist.*, 23, 323–355.

Snijders, T. & Bosker, R. (1999). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Sage Publications, London.

Spiegelhalter, D.T.A. & Best, N. (2000). *WinBUGS User Manual*. MRC Biostatistics Unit, Cambridge.

Spiegelhalter, D.J., Best, N.G., Carlin, B.R. & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *J. R. Statist. Soc. Ser. B Statist. Methodol.*, 64, 583–616.

Storch, D., Evans, K.L. & Gaston, K.J. (2005). The species–area–energy relationship. *Ecol. Lett.* 8, 487–492.

Sugihara, G. & May, R.M. (1990). Application of fractals in ecology. *Trends Ecol. Evol.*, 5, 79–86.

© 2007 Blackwell Publishing Ltd/CNRS

Takada, M. & Miyashita, T. (2004). Additive and non-additive effects from a larger spatial scale determine small-scale densities in a web spider neriene brongersmai. *Popul. Ecol.*, 46, 129–135.

The Mathworks, Inc. (2003). *Matlab. Version 6.5*. Natick, MA.

Thogmartin, W.E., Knutson, M.G. & Sauer, J.R. (2006). Predicting regional abundance of rare grassland birds with a hierarchical spatial count model. *Condor*, 108, 25–46.

Thrush, S.F., Cummings, V.J., Dayton, P.K., Ford, R., Grant, J., Hewitt, J.E. *et al.* (1997). Matching the outcome of small-scale density manipulation experiments with larger scale patterns and example of bivalve adult/juvenile interactions. *J. Exp. Mar. Biol. Ecol.*, 216, 153–169.

Tilman, D. & Kareiva, P.M. (1997). *Spatial Ecology*. Princeton University Press, Princeton, NJ.

Tolimieri, N. (1995). Effects of microhabitat characteristics on the settlement and recruitment of a coral reef fish at two spatial scales. *Oecologia*, 102, 52–63.

Underwood, A.J. & Chapman, M.G. (1996). Scales of spatial patterns of distribution of intertidal invertebrates. *Oecologia*, 107, 212–224.

Underwood, A.J. & Petraitis, P. (1993). Structure of intertidal assemblages in different locations: how can local processes be compared? In: *Species Diversity in Ecological Communities* (eds Ricklefs, R. & Schluter, D.). University of Chicago Press, Chicago, pp. 38–51.

Vazquez, D.P. & Simberloff, D. (2004). Indirect effects of an introduced ungulate on pollination and plant reproduction. *Ecol. Monogr.*, 74, 281–308.

Whittaker, R.J., Willis, K.J. & Field, R. (2001). Scale and species richness: towards a general, hierarchical theory of species diversity. *J. Biogeogr.*, 28, 453–470.

Wiens, J.A. 1991. Ecological similarity of shrub-desert avifaunas of Australia and North-America. *Ecology*, 72, 479–495.

Wu, J.G. & David, J.L. (2002). A spatially explicit hierarchical approach to modeling Complex ecological systems: theory and applications. *Ecol. Modell.*, 153, 7–26.

APPENDIX A. THREE-LEVEL *TIPULARIA* FLOWERING STUDY DETAILS

Study design and sampling regime

Tipularia discolor is a wintergreen terrestrial orchid found in mixed deciduous forests of eastern North America. In *T. discolor*'s southern range, a plant's single green leaf emerges above ground in late fall (end of September) and remains until spring (March–April). Flowers, if produced, are found on a single flowering stalk that emerges in August, before the leaf emerges. Seed pods, containing thousands of seeds each, become dry and dehisce over the course of the fall and winter, with the seeds being predominately wind/gravity dispersed. Individuals within the 10 populations, ranging from Whitehall Forest in Athens-Clarke County, GA (33°92′N latitude, 150–240 m elevation) to the Nancytown area (34°31′N latitude, 315–450 m elevation) in Habersham County, GA have been visited yearly since 1999 to measure growth and reproduction. All populations are found in relatively mature (*c.* 80 years old) deciduous forests. Study populations range from 250 to 480 m^2 in size, each divided into 2 m^2 cells within which all plants were individually marked. For this study, we use floral surveys from the late fall of 2004.

A hand-held unit from Hydrosense was used to measure soil moisture in the top 12 cm of soil at 80 points within each of the study populations during the summer preceding flowering. Understory light availability was measured above the forest litter layer using a hand-held AccuPAR ceptometer wand at the same 80 points as soil moisture, and after leaf fall due to the wintergreen phenology of *Tipularia*. Concurrent light measurements taken in nearby open clearings allowed calculation of per cent photosynthetic photon flux densities at the forest floor. The 80 abiotic sampling points were selected to allow geostatistical predictions of light and moisture for all cells in the study grids. Soil cores were taken from 16 cells in order to characterize population-level pH and soil texture.

APPENDIX B. MODEL FITTING DETAILS AND WINBUGS CODE FOR THREE-LEVEL BAYESIAN MODEL

Model building, estimation and interpretation

The three-level Bayesian models were fit using WinBugs 1.4, which uses a variety of Markov chain Monte Carlo sampling methods, depending on the demands of the model, to describe the posterior distributions of model parameters (Gilks *et al.* 1996; Spiegelhalter & Best 2000). For hierarchical logistic regressions such as this model, Metropolis–Hastings algorithms are used for sampling, as conditional posterior distributions cannot often be directly sampled. After a burn-in period of 5000 iterations, used to avoid any relic effects of starting points, we simulated three independent chains for 50 000 iterations, thinning to every fifth sample. Convergence was assessed via the Gelman–Rubin statistic and examination of iteration histories. Non-informative priors were used in all cases to allow data to dominate posterior estimation. Both *Gamma* and *Uniform* priors were used for variance terms, and no detectable difference was observed, but *Uniform* priors on the standard deviations were used in accordance with recommendation in Gelman & Pardoe (2005).

WinBugs code

```
# Conditional 3-level with Grid level Abiotic
  model;
{
for (i in 1:716){  # Individual-level
flr04[i] ~ dbern(p[i])
```

© 2007 Blackwell Publishing Ltd/CNRS

```
logit(p[i]) <- cell.int[tipdata[i,5]] + lw*lw04[i]
lw04[i] ~ dnorm(lw.mean, lw.prec)
}
lw.mean~ dnorm(0,.01)
lw.prec~ dgamma(.01,.01)

for (c in 1:180){  # Cell-level
cell.int[c] ~ dnorm(cell.int1[c], cell.prec)
cell.int1[c] <- grid.int[gridc[c,3]] + par*PARWtr04[c] +
 tdr*TDRSum04[c]
logit(cellpred[c])<- cell.int[c]
e.cell[c] <- cell.int[c] - cell.int1[c]
}

# Coefficient Priors
tdr  ~ dnorm(0, .001)
par ~ dnorm(0, .001)
tdr2 ~ dnorm(0, .001)
par2 ~ dnorm(0, .001)
lw ~ dnorm(0,.001)
tdr.g ~ dnorm(0, .001)
par.g ~ dnorm(0, .001)

for (g in 1:10){   # Grid-level
grid.int[g] ~ dnorm(grid.int1[g], grid.prec)
grid.int1[g] <- global.int + ph.g*ph.mu[gr[g]] + sand.g*
 sand.mu[gr[g]]
ph.mu[g] ~ dnorm(ph.global, ph.prec2)
sand.mu[g] ~ dnorm(sand.global, sand.prec2)
e.grid[g] <- grid.int[g] - grid.int1[g]
}

for (s in 1:138){
ph[s,2] ~ dnorm(ph.mu[grid[s]], ph.prec)
sand[s,2] ~ dnorm(sand.mu[grid[s]], sand.prec)
}
ph.global~ dnorm(0,.001)
ph.g~dnorm(0, .001)
sand.global~ dnorm(0,.001)
sand.g~dnorm(0, .001)
global.int  ~ dnorm(0,.001)

sigma.g ~ dunif(0,10)   #### Using Uniform priors on
 StDev, following Gelman  2006
sigma2.g <-sigma.g*sigma.g
grid.prec<- 1/ sigma2.g
sigma.c ~ dunif(0,10)
sigma2.c <-sigma.c*sigma.c
cell.prec <-1/sigma2.c

sigma.ph ~ dunif(0,10)
sigma2.ph <-sigma.ph*sigma.ph
ph.prec <-1/sigma2.ph

sigma.ph2 ~ dunif(0,10)
sigma2.ph2 <-sigma.ph2*sigma.ph2
ph.prec2 <-1/sigma2.ph2

sigma.sand ~ dunif(0,10)
sigma2.sand <-sigma.ph*sigma.ph
sand.prec <-1/sigma2.ph

sigma.sand2 ~ dunif(0,10)
sigma2.sand2 <-sigma.sand2*sigma.sand2
sand.prec2 <-1/sigma2.sand2
} # end model
```

Editor, Marti Anderson
Manuscript received 7 December 2006
First decision made 11 January 2007
Second decision made 19 February 2007
Manuscript accepted 1 March 2007

© 2007 Blackwell Publishing Ltd/CNRS