Hierarchical dynamic models for multivariate times series of counts

Nalini Ravishanker*, Volodymyr Serhiyenko, and Michael R. Willig

In several application areas, we see the need for accurate statistical modeling of multivariate time series of counts as a function of relevant covariates. In ecology, count responses on species abundance are observed over several time periods at several locations, and the covariates that influence the abundance may be location-specific and/or time-varying. This paper describes a Bayesian framework for estimation and prediction by assuming a multivariate Poisson sampling distribution for the count responses and by fitting a hierarchical dynamic model. Our modeling incorporates the temporal dependence as well as dependence between the components of the response vector.

Keywords and phrases: Bayesian modeling, Ecology, Gastropod abundance, Nonlinear state space model.

1. INTRODUCTION

In several application areas, we increasingly see the need for developing accurate statistical modeling approaches for time series of multivariate count responses. The response consists of an m-dimensional vector of counts that is observed at each of N locations (or for each of N subjects) over T regularly spaced times. The objective of the statistical analysis is to understand stochastic temporal patterns in the response as a function of observed location (or subject)-specific and/or time-varying covariates. For instance, in ecology, understanding the causes and consequences of variation in the abundance of organisms as a function of topographical and environmental covariates has been a long-standing goal (Krebs, 1972, Scheiner and Willig, 2011). In business, a pharmaceutical firm may be interested in estimating and predicting the number of new prescriptions written by physicians of drugs from the firm and its competitors, as a function of the firm's promotional activities (Venkatesan *et al.*, 2012). In a problem in transportation engineering, the state of Connecticut is interested in understanding stochastic patterns in the temporal behavior of crash counts categorized by injury severity across a set of highway segments, as a function of roadway geometry,

*Corresponding author.

traffic volume, etc. (Serhiyenko *et al.*, 2014). For such applications, the modeling described in this article enable us to adequately incorporate dependence in the response over time as well as the dependence between the components of the response vector. We illustrate the approach for the ecology example.

The literature on count time series modeling includes observation driven models and parameter driven models. One approach to develop models for count time series is based on the thinning operator (Steutel and van Harn, 1979), where the thinning operator is generated by counting series of Bernoulli-distributed random variables. McKenzie (1985) and Al-Osh and Alzaid (1987) independently developed the first-order integer-valued autoregressive, INAR(1) model. McKenzie (2003) and Jung and Tremayne (2006) present a good review of subsequent developments. Regression modeling for count time series using quasi-likelihood methods were discussed in Zeger (1988). Jorgensen *et al.* (1999) described analysis of longitudinal multivariate count data driven by a latent gamma Markov process using a state space approach. Davis *et al.* (2003) described maximum likelihood estimation for generalized autoregressive moving average (GLARMA) models for count time series. Bayesian modeling of panel count data was discussed in Chib *et al.* (1998), while Chib and Winkelmann (2001) discussed models with latent effects for correlated count data. Count data models in the state space approach was discussed in Gamerman (1998), Durbin and Koopman (2000), Fruhwirth-Schnatter and Wagner (2006), and Gamerman *et al.* (2013), among others. Multivariate INAR models for counts were discussed in Pedeli and Karlis (2011, 2012) and references therein.

For Gaussian dynamic linear models (DLMs), also often referred to as Gaussian state space models, Kalman (1960) and Kalman and Bucy (1961) popularized a recursive algorithm for optimal estimation and prediction of the state vector, which then enables prediction of the observation vector; see West and Harrison (1989) for details. Carlin *et al.* (1992) described the use of Markov Chain Monte Carlo methods (Chen *et al.*, 2000) for non-Gaussian and non-linear state space models. Hierarchical dynamic linear models (HDLMs) combine the stratified parametric linear models (Lindley and Smith, 1972) and the DLMs into a general framework, which have been particularly useful in

econometric, education, and health-care applications. The Gaussian HDLM includes a set of one or more dimension reducing structural equations along with the observation equation and state (system) equation of the DLM (Gamerman and Migon, 1993). Landim and Gamerman (2000) further extended the Gaussian HDLM to a more general class of models where the response vector has a matrix-valued normal distribution.

For situations where the time series of responses consists of counts, DLMs have been generalized to dynamic generalized linear models (DGLMs) or exponential family state space models, which assume that the sampling distribution is a member of the exponential family of distributions, such as the Poisson or negative binomial distributions (Gamerman, 1998). The DGLMs may be viewed as dynamic versions of the generalized linear models (McCullagh and Nelder, 1989), and Bayesian inference is facilitated through a Metropolis-Hastings algorithm combined with the Gibbs sampler in repeated use of an adjusted version of Gaussian DLM. In many applications, the response consists of a vector-valued time series of counts, and there is a need to develop statistical modeling approaches for estimation and prediction. In this article, we describe a hierarchical multivariate dynamic model (HMDM), with a multivariate Poisson distribution (MVP) as the sampling distribution for the response vector time series of counts, and incorporating covariates that may vary over location and/or time. The use of the MVP distribution enables us to model associations between the components of the count response vector, while the dynamic framework allows us to model the temporal behavior. The hierarchical structure enables us to capture the location (or subject) specific effects over time.

The format of the paper follows. Section 2 gives a description of the ecological application, including a description of the data. Section 3 reviews the MVP distribution and describes fast computation of its probability mass function (pmf). Section 4 describes the HMDM model and gives details of the Bayesian inference. Section 5 discusses model selection and prediction based on the fitted HMDM. Section 6 presents results for simulated data, while Section 7 presents an analysis of the ecological data on gastropod abundance. Section 8 provides a summary and discussion.

2. GASTROPOD ABUNDANCE IN THE LUQUILLO EXPERIMENTAL FOREST IN PUERTO RICO – DATA DESCRIPTION

Understanding the causes and consequences of variation in the abundance of organisms has been a long-standing goal in ecology (Scheiner and Willig, 2011). Nonetheless, few long-term analyses spanning over 20 years of spatiotemporal variation in abundance exist, especially for invertebrate populations in tropical habitats that are subject to high intensity but infrequent disturbances such as hurricanes (e.g., Willig *et al.*, 2012). Terrestrial gastropods are of considerable ecological importance because of their abundance,

diversity, and trophic position. Moreover, terrestrial gastropods, like non-marine mollusks in general, are suffering from global declines and are in need of scientifically informed conservation action and management (Lydeard *et al.*, 2004). As such, there is considerable urgency to understand variation in gastropod abundance and the factors that affect it. In some ecosystems, gastropods respond to environmental gradients (Willig *et al.*, 1998a, 2011, 2012) and to disturbances (Bloch and Willig 2006), including those induced by human activities, and do so at a variety of spatial scales (Willig *et al.*, 1998b, 2007). Because gastropods are ectothermic and not particularly vagile, they are constrained in distribution and behavior by desiccation stress (Cook 2001). This fauna evinces a suite of attributes that suggests differential responses to spatial variation in habitat or micro-climate that might arise as a consequence of global change. Finally, effective management or conservation of populations threatened by altered disturbance regimes requires species-specific understanding of the particular environmental aspects of change that are associated with alterations in abundance.

Long-term censuses of terrestrial gastropods were accomplished on the Luquillo Forest Dynamics Plot (LFDP; 18º20 N, 65º49 W), a16-ha grid in the northwest of the Luquillo Experimental Forest (LEF) in the Luquillo Mountains of northeastern Puerto Rico (McDowell *et al.*, 2012). Although a modestly drier period typically extends from January to April (hereafter, the dry season), rainfall generally remains at least 20 cm in all months (Brown *et al.*, 1983). The basic census design includes sampling during the wet and dry season of each year on each of forty circular sites (3-m radius) that are spaced evenly within a rectilinear grid such that 60-m separated adjacent points along a row or column (see Bloch and Willig, 2006, Willig *et al.*, 1998b, 2007). The abundance of gastropods is sampled at each of those 40 sites during the dry and wet seasons for each year from 1991 to 2012. The number of replicate samples per season differed among years: 1 replicate in the dry season of 1991; 2 replicates in the dry season of 2003 and in the dry and wet seasons of 1992 and 1993; 3 replicates in the dry season of 1995 and in the wet and dry season of 1994; and 4 replicates in all other seasons and years between 1995 and 2012.

Although 17 species of gastropods are known to live in the Luquillo Forest Dynamics Plot, we focus on *Caracolus caracolla* and *Gaeotis nigrolineata*, the most abundant and widely distributed terrestrial gastropods in the tabonuco forest. We estimate abundance based on the minimum number known to be alive (MKNA) at each site during each season (i.e., the maximum number of individuals captured within a season at a site). All individuals were identified to species in the field and returned as closely as possible to the point of capture and always within the site of capture. There is no ecological reason to support an assumption of negative association between the counts over time for these two species (Bloch and Willig, 2012).

In addition to species counts, environmental characteristics that are invariant over the course of the study were

determined for each site. Elevation and slope (unitless) are continuous variables, ranging from 333 m asl to 428 m asl, and 0.7 to 65.1, respectively. Aspect and soil type are categorical variables that are estimated based on four 20 m by 20 m quadrants whose vertex was coincident with the center of each survey plot. The aspect of each quadrant was characterized into categories that represent the angular equivalents of the four cardinal (N, E, S, and W) or four intermediate (NE, SE, SW, and NW) compass directions. Because these categories arise from an underlying circular distribution, we quantify the central tendency after converting the angles to radians and applying a cosine transformation, doing so only when the observed aspect is no more than 135 degrees (in order to ensure environmentally informative characterizations). Aspect categories were combined into four distinct levels to increase sample sizes, viz., 1, 5, 7, and 8, of which level 1 is used as the baseline in the model shown in section 7. To ensure environmentally informative characterizations of soil type, we characterized each survey plot by the dominant soil type based on consideration of its four associated quadrants, when the most pervasive soil type was at least twice as frequent as the second most common soil type, and was dominant in at least two of the four quadrants. The soil type levels are Zarzal (1), Cristal (2), or Prieto (3).

Based on percent canopy cover (CC) evident in aerial photographs taken in 1936, US Forest Service records, and other sources, the LFDP can be subdivided into four canopy cover classes (Thompson *et al.*, 2002). CC classes 1 and 2 were combined to increase sample size, resulting in three areas of historic land use in 1936. The redefined CC level 1 (0–49% cover) experienced the most intensive logging and agriculture prior to 1934; CC level 2 (50–80% cover) was used for shade-coffee cultivation and other small scale mixed agriculture before 1934; and CC level 3 (80–100% cover) was lightly and selectively logged up to the 1950s.

As a consequence of disturbance and secondary succession, some habitat characteristics (e.g., canopy openness and plant apparency) vary over time at each site. Canopy openness (CO) was measured using a spherical densitometer, higher numbers representing greater canopy openness. Empirically, it equals the average number of grid cells that are not occluded by vegetation from measurements by a densitometer at the mid-point of the 4 cardinal radii of each sampling plot. Plant apparency is the volume of space in the understory that was occupied by plants, and is estimated using a plant apparency device at each of the mid-points along the cardinal radii. We estimated plant apparency (PA) via a method (Cook and Stubbendiek, 1986) that quantifies the aerial density of all living vegetation at heights up to 3 m above the forest floor (see Secrest *et al.*, 1996). Using a plant apparency device (Secrest, 1995), we determine the cumulative number of foliar intercepts, defined as the sum of species-specific counts of living vegetation touching a wooden dowel at each of seven heights (0, 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 m). The device comprises a set of four 0.5 m

long dowels positioned at 90° angles at each height. The device is positioned 1.5 m from the center of each plot in each of the four cardinal directions. Apparency is estimated separately for *Prestoea acuminata*, the sierra palm (PAsp) and for all other plant species (PAothers). Canopy openness and plant apparency are measured only during wet seasons. In order to investigate the gastropod abundance, we propose hierarchical dynamic modeling of the bivariate count time series, which enables us to study the effect of site-specific and time-dependent covariates, as described in the following sections.

3. MULTIVARIATE POISSON DISTRIBUTION

The definition of an m-variate Poisson distribution is based on a mapping $g : \mathbb{N}^q \rightarrow \mathbb{N}^m$, $q \geq m$, such that $\boldsymbol{Y} = g(\boldsymbol{X}) = \mathbf{A}\boldsymbol{X}$ (Mahamunulu, 1967; Johnson *et al.*, 1997). Here, $\boldsymbol{X} = (X_1, \ldots, X_q)'$ is a vector of unobserved independent Poisson random variables, i.e., $X_r \sim \text{Poisson}(\lambda_r)$ for $r = 1, \ldots, q$; and $\mathbf{A}$ is an arbitrary $m \times q$ matrix which determines the properties of the multivariate Poisson distribution. The m-dimensional vector $\boldsymbol{Y} = (Y_1, \ldots, Y_m)' = \mathbf{A}\boldsymbol{X}$ follows a multivariate Poisson distribution with parameters $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_q)'$ and pmf

$$MP_m(\boldsymbol{y}|\boldsymbol{\lambda}) = P(\boldsymbol{Y} = \boldsymbol{y}|\boldsymbol{\lambda}) = \sum_{\boldsymbol{x} \in g^{-1}(\boldsymbol{y})} P(\boldsymbol{X} = \boldsymbol{x}|\boldsymbol{\lambda})$$

$$(1) \qquad = \sum_{\boldsymbol{x} \in g^{-1}(\boldsymbol{y})} \prod_{r=1}^{q} P(X_r = x_r|\lambda_r)$$

where $g^{-1}(\boldsymbol{Y})$ denotes the inverse image of $\boldsymbol{Y} \in \mathbb{N}^m$ and for $r = 1, \ldots, q$, the pmf of the univariate Poisson distribution is $P(X_r = x_r|\lambda_r) = \exp(-\lambda_r)\lambda_r^{x_r}/x_r!$.

The literature on the use of the multivariate Poisson distribution for modeling applications was sparse until recently, possibly due to the complicated form of the pmf (1). Karlis and Meligkotsidou (2005) proposed the two-way covariance structured multivariate Poisson distribution which permits a more realistic modeling of multivariate counts for several practical applications. This distribution is constructed by setting $\boldsymbol{A} = [\boldsymbol{A}_1 \ \boldsymbol{A}_2]$, where $\boldsymbol{A}_1 = \boldsymbol{I}_m$ captures the main effects; $\boldsymbol{A}_2$ captures the two-way covariance effects; $\boldsymbol{A}_2$ is an $m \times (\frac{m(m-1)}{2})$ binary matrix; each column of $\boldsymbol{A}_2$ has exactly 2 ones and $(m-2)$ zeros and no duplicate columns exist; and $q = m + [m(m-1)]/2$. We correspondingly split the parameter $\boldsymbol{\lambda}$ into two parts, viz., $\boldsymbol{\lambda}^{(1)} = (\lambda_1, \cdots, \lambda_m)'$, which corresponds to the m main effects, and $\boldsymbol{\lambda}^{(2)} = (\lambda_{m+1}, \cdots, \lambda_q)'$ which corresponds to the $m(m-1)/2$ pairwise covariance effects. For example, when $m = 2$, the bivariate Poisson distribution with two-way covariance structure for $\boldsymbol{Y} = (Y_1, Y_2)'$ is expressed via $q = 3$ independent Poisson random vari-

ables as:

$$Y_1 = X_1 + X_3$$
$$(2) \qquad Y_2 = X_2 + X_3$$

where $X_i \sim \text{Poisson}(\lambda_i)$, $i = 1, 2, 3$. The joint pmf of Y_1 and Y_2 is:

$$P(Y_1 = y_1, Y_2 = y_2|\boldsymbol{\lambda}) = \exp\{-(\lambda_1 + \lambda_2 + \lambda_3)\}$$
$$(3) \qquad \times \frac{\lambda_1^{y_1}}{y_1!}\frac{\lambda_2^{y_2}}{y_2!}\sum_{i=0}^{s}\binom{y_1}{i}\binom{y_2}{i}i!\left(\frac{\lambda_3}{\lambda_1\lambda_2}\right)^i$$

where $s = \min(y_1, y_2)$.

Similarly, when $m = 3$, we write $\boldsymbol{Y} = (Y_1, Y_2, Y_3)'$ as:

$$Y_1 = X_1 + X_4 + X_5$$
$$Y_2 = X_2 + X_4 + X_6$$
$$(4) \qquad Y_3 = X_3 + X_5 + X_6$$

where $X_i \sim \text{Poisson}(\lambda_i)$ for $i = 1, \cdots, 6$. The joint probability mass function of Y_1, Y_2 and Y_3 is:

$$P(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3|\boldsymbol{\lambda}) = \exp\{-\sum_{i=1}^{6}\lambda_i\}$$
$$\times \sum_{(X_4, X_5, X_6)\in C} \frac{\lambda_1^{y_1-X_4-X_5}\lambda_2^{y_2-X_4-X_6}}{(y_1 - X_4 - X_5)!(y_2 - X_4 - X_6)!}$$
$$(5) \qquad \times \frac{\lambda_3^{y_3-X_5-X_6}\lambda_4^{X_4}\lambda_5^{X_5}\lambda_6^{X_6}}{(y_3 - X_5 - X_6)!X_4!X_5!X_6!}$$

where the summation is over the set C such that $C = [(X_4, X_5, X_6) \in \mathbb{N}^3 : (X_4 + X_5 \le y_1) \cap (X_4 + X_6 \le y_2) \cap (X_5 + X_6 \le y_3)] \ne \emptyset]$

It is easy to see that the matrix $\boldsymbol{A}$ has the respective forms shown below for $m = 2$ and $m = 3$:

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

The mean vector and variance-covariance matrix of $\boldsymbol{Y}$ are given by $E(\boldsymbol{Y}|\boldsymbol{\lambda}) = \boldsymbol{A}\boldsymbol{\lambda}$ and $Cov(\boldsymbol{Y}|\boldsymbol{\lambda}) = \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}'$, where $\boldsymbol{\Sigma} = \text{diag}(\lambda_1, \dots, \lambda_q)$. As pointed out by Karlis and Meligkotsidou (2005), this covariance cannot accommodate negative associations among the components of $\boldsymbol{Y}$. When $m = 1$, the $MP_m(\boldsymbol{y}|\boldsymbol{\lambda})$ in (1) reduces to the univariate Poisson pmf $P(Y = y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}$.

A recursive scheme was proposed by Tsiamyrtzis and Karlis (2004) for computing the multivariate Poisson pmf when $m = 2$ or $m = 3$. We use the following faster approach for the calculation of the multivariate Poisson pmf proposed in Hu (2012). When $m = 2$, let y_1 and y_2 denote the observed counts, and without loss of generality, assume that $y_1 \le y_2$, so that $\min(y_1, y_2) = y_1$. Since X_3 is the common term in both equations in (2), it is straightforward to

obtain the set of possible values that X_3 can assume, viz., $x_3 = 0, \dots, \min(y_1, y_2)$, and obtain the corresponding values assumed by X_1 and X_2 to be respectively $X_1 = y_1 - x_3$ and $X_2 = y_2 - x_3$. We have solved for all possible sets of values for the inverse image of $\boldsymbol{y}$, i.e., $\boldsymbol{x} \in g^{-1}(\boldsymbol{y})$. The pmf for the bivariate Poisson distribution can be calculated using (3).

When $m = 3$, without loss of generality, we assume that $y_1 \le y_2 \le y_3$. The possible values for x_4 and x_5 are in the set $C_1 = (0, \dots, y_1)$, and the possible values for x_6 are in the set $C_2 = (0, \dots, y_2)$. We have in total K different combinations for (x_4, x_5, x_6), where $K = $ (length of set C_1)^2 × (length of set C_2) $= (y_1 + 1)^2(y_2 + 1)$. The corresponding values for X_1, X_2, X_3 can be calculated from (4). Let C^* denote the set of K different combinations of possible values for all $q = 6$ independent Poisson variables. Since it is possible that in the set C^*, X_1, X_2, or X_3 may assume negative values, a subset of C^*, which only contains non-negative values of X_1, X_2, and X_3, is the inverse image of $\boldsymbol{y}$. The pmf of the trivariate Poisson distribution is then obtained using (5).

The computation of the multivariate Poisson likelihood directly depends on the magnitude of the counts, and a sizable portion of the computational effort in the Bayesian modeling is for the evaluation of the likelihood. In Table 1, we present the CPU times (seconds) for the likelihood calculation of 10,000 simulated counts under three scenarios corresponding to different λ's. In these computations, we have assumed all pairwise covariance mean effects to be equal to 1, and have chosen different main effect means for each scenario. Specifically, in the univariate case when $m = 1$, we simulate Poisson random variables Y's under the three scenarios with means $\lambda = 3$, $\lambda = 5$ and $\lambda = 10$. In the bivariate case (2), we simulate X_1, X_2, X_3 under scenario 1 as univariate Poisson with respective means $\lambda_1 = 2, \lambda_2 = 2, \lambda_3 = 1$. Under scenario 2, $\lambda_1 = 4, \lambda_2 = 4, \lambda_3 = 1$, while under scenario 3, $\lambda_1 = 9, \lambda_2 = 9, \lambda_3 = 1$. In the trivariate case (4), we simulate $X_1, \cdots, X_6$ under scenario 1 as univariate Poisson with respective means $\lambda_1 = \cdots = \lambda_6 = 1$. Under scenario 2, $\lambda_1 = \lambda_2 = \lambda_3 = 2$, and $\lambda_4 = \lambda_5 = \lambda_6 = 1$. Under scenario 3, $\lambda_1 = \lambda_2 = \lambda_3 = 8$, and $\lambda_4 = \lambda_5 = \lambda_6 = 1$. Simulations are run on Dell Optiplex 990 (Intel Core i7-2600 CPU a quad core processor with 3.4Ghz) with 16Gb of RAM using 32-bit version of Debian GNU/Linux version 6 operational system. The computational times for the simulations are comparable across the three scenarios for a given m, and increases rapidly as m increases.

Table 1. CPU Times (secs) for the Likelihood Computation

Likelihood	Scenario 1	Scenario 2	Scenario 3
Univariate Poisson	8×10^{-3}	8×10^{-3}	8.001×10^{-3}
Bivariate Poisson	8×10^{-3}	1.2×10^{-2}	2.4×10^{-2}
Trivariate Poisson	1.412	18.209	248.788

4. HIERARCHICAL MULTIVARIATE DYNAMIC MODEL (HMDM)

Let $\boldsymbol{Y}_{it} = (Y_{1,it}, \cdots, Y_{m,it})$, for $t = 1, \cdots, T$ denote the m-dimensional time series of counts from location i, where $i = 1, \cdots, N$, and assume that $\boldsymbol{Y}_{it}$ follows a multivariate Poisson distribution (1). The observation equation of the HMDM is

$$
\begin{aligned}
\boldsymbol{Y}_{it}|\boldsymbol{\lambda}_{it} &\sim MP_m(\boldsymbol{y}_{it}|\boldsymbol{\lambda}_{it}) \\
(6) \qquad \log \lambda_{j,it} &= \boldsymbol{D}'_{j,it}\boldsymbol{\delta}_{j,it} + \boldsymbol{S}'_{j,it}\boldsymbol{\eta}_j, \quad j = 1, \cdots, q
\end{aligned}
$$

where log denotes the natural logarithm, $\boldsymbol{D}_{j,it} = (D_{j,it,1}, \cdots, D_{j,it,a_j})'$ is an a_j-dimensional vector of exogenous predictors with location-time varying (dynamic) coefficients $\boldsymbol{\delta}_{j,it} = (\delta_{j,it,1}, \cdots, \delta_{j,it,a_j})'$ and $\boldsymbol{S}_{j,it} = (S_{j,it,1}, \cdots, S_{j,it,b_j})'$ is a b_j-dimensional vector of exogenous predictors with static coefficients $\boldsymbol{\eta}_j = (\eta_{j,1}, \cdots, \eta_{j,it,b_j})'$. We assume that the model either includes $\delta_{j,it,1}$ which represents the location-time varying intercept, or includes $\eta_{j,1}$ which represents the static intercept, i.e., either $D_{j,it,1} = 1$ or $S_{j,it,1} = 1$. A simple formulation of (6) could set $a_j = 1$ for $j = 1, \cdots, q$, set $b_j = b > 1$ for $j = 1, \cdots, m$ and $b_j = 0$ for $j = m+1, \cdots, q$, which implies using only the location-specific and time-dependent intercept to model the Poisson means corresponding to the association portion, and the location-time intercept together with an equal number of static coefficients (corresponding to exogenous predictors) for the main effects portion of the multivariate Poisson specification.

For the reminder of the paper, let $p_d = \sum_{j=1}^q a_j$ and $p_s = \sum_{j=1}^q b_j$. Let $\boldsymbol{\beta}_{it}$ be a p_d-dimensional vector constructed by stacking the a_j coefficients $\boldsymbol{\delta}_{j,it}$ for $j = 1, \cdots, q$. The structural equation of the HMDM relates the location-time varying parameter $\boldsymbol{\beta}_{it}$ to an aggregate (pooled) state parameter $\boldsymbol{\gamma}_t$:

$$
(7) \qquad \boldsymbol{\beta}_{it} = \boldsymbol{\gamma}_t + \boldsymbol{v}_{it}
$$

where the errors $\boldsymbol{v}_{it}$ are assumed to be i.i.d. $N_{p_d}(\boldsymbol{0}, \boldsymbol{V}_i)$. The state (or system) equation of the HMDM is:

$$
(8) \qquad \boldsymbol{\gamma}_t = \boldsymbol{G}\boldsymbol{\gamma}_{t-1} + \boldsymbol{w}_t
$$

where $\boldsymbol{G}$ is a $p_d \times p_d$ state transition matrix and the state errors $\boldsymbol{w}_t$ are assumed to be i.i.d. $N_{p_d}(\boldsymbol{0}, \boldsymbol{W})$.

The HMDM in (6)–(8) simplifies to the Hierarchical DGLM (HDGLM) when $m = 1$, where we replace $MP_m(\boldsymbol{y}_{it}|\boldsymbol{\lambda}_{it})$ by the univariate Poisson pmf.

4.1 Bayesian inference

Let $\boldsymbol{Y}$, $\mathbf{D}$ and $\mathbf{S}$ denote all the responses $\boldsymbol{y}_{it}$, and the dynamic predictors and the static predictors for $t = 1, \cdots, T$ and $i = 1, \cdots, N$. Let $\boldsymbol{\eta}$ and $\boldsymbol{\beta}$ denote all the coefficients $\boldsymbol{\eta}_j$ and $\boldsymbol{\beta}_{it}$ for $j = 1, \cdots, q$, $t = 1, \cdots, T$ and $i = 1, \cdots, N$, and

let $\boldsymbol{\gamma}$ denotes all the coefficients $\boldsymbol{\gamma}_t$ for $t = 1, \cdots, T$. The likelihood function under the model described by (6)–(8) is

$$
\begin{aligned}
L(\boldsymbol{\eta}, \boldsymbol{\beta}, \boldsymbol{\gamma}; \boldsymbol{Y}, \mathbf{D}, \mathbf{S}) &= \prod_{i=1}^N \prod_{t=1}^T MP_m(\boldsymbol{y}_{it}|\boldsymbol{\beta}_{it}, \boldsymbol{\gamma}_t) \\
(9) \quad &\times \; p_{\text{normal}}(\boldsymbol{\eta}) \times p_{\text{normal}}(\boldsymbol{\beta}_{it}|\boldsymbol{\gamma}_t) \times p_{\text{normal}}(\boldsymbol{\gamma}_t|\boldsymbol{\gamma}_{t-1})
\end{aligned}
$$

where we have suppressed the terms $\mathbf{D}$ and $\mathbf{S}$ on the right side for brevity. We assume multivariate Normal priors for the initial state vector and the static coefficients, i.e., $\boldsymbol{\gamma}_0 \sim N_{p_d}(\boldsymbol{m}_0, \boldsymbol{C}_0)$ and $\boldsymbol{\eta} \sim N_{p_s}(\boldsymbol{\mu}_\eta, \boldsymbol{\Sigma}_\eta)$. We assume inverse Wishart priors for the variance terms $\boldsymbol{V}_i$ and $\boldsymbol{W}$, i.e., $\boldsymbol{V}_i \sim IW(n_v, \boldsymbol{S}_v)$, and $\boldsymbol{W} \sim IW(n_w, \boldsymbol{S}_w)$, and we assume a product prior specification. The hyperparameters are selected to correspond to a vague prior specification. Prior elicitation is an important, ongoing problem of considerable interest in Bayesian analysis, and the vast, growing literature includes seminal work on objective priors (Berger, 2006), power priors (Ibrahim and Chen, 2000), expert elicited priors (O'Hagan *et al.*, 2006), etc. In section 6, we discuss prior sensitivity for our analysis.

The joint posterior of the unknown parameters is proportional to the product of the likelihood and the prior:

$$
\pi(\boldsymbol{\beta}_{it}, \boldsymbol{\gamma}_t, \boldsymbol{\eta}, \boldsymbol{V}_i, \boldsymbol{W}|\boldsymbol{Y}, \mathbf{D}, \mathbf{S}) \propto [\prod_{t=1}^T \prod_{i=1}^N MP_m(\boldsymbol{y}_{it}|\boldsymbol{\lambda}_{it})
$$

$$
\times \; |\boldsymbol{V}_i|^{-1/2} \exp\{-\frac{1}{2}(\boldsymbol{\beta}_{it} - \boldsymbol{\gamma}_t)'\boldsymbol{V}_i^{-1}(\boldsymbol{\beta}_{it} - \boldsymbol{\gamma}_t)\}]
$$

$$
\times \; |\boldsymbol{\Sigma}_\eta|^{-1/2} \exp\{-\frac{1}{2}(\boldsymbol{\eta} - \boldsymbol{\mu}_\eta)'\boldsymbol{\Sigma}_\eta^{-1}(\boldsymbol{\eta} - \boldsymbol{\mu}_\eta)\}
$$

$$
\times \; [\prod_{t=1}^T |\boldsymbol{W}|^{-1/2} \exp\{-\frac{1}{2}(\boldsymbol{\gamma}_t - \boldsymbol{G}\boldsymbol{\gamma}_{t-1})'\boldsymbol{W}^{-1}(\boldsymbol{\gamma}_t - \boldsymbol{G}\boldsymbol{\gamma}_{t-1})\}]
$$

$$
\times \; |\boldsymbol{C}_0|^{-1/2} \exp\{-\frac{1}{2}(\boldsymbol{\gamma}_0 - \boldsymbol{m}_0)'\boldsymbol{C}_0^{-1}(\boldsymbol{\gamma}_0 - \boldsymbol{m}_0)\}
$$

$$
\times \; [\prod_{i=1}^N |\boldsymbol{V}_i|^{-n_v/2} \exp\{-\frac{1}{2}tr(\boldsymbol{V}_i^{-1}\boldsymbol{S}_v)\}]
$$

$$
\times \; |\boldsymbol{W}|^{-n_w/2} \exp\{-\frac{1}{2}tr(\boldsymbol{W}^{-1}\boldsymbol{S}_w)\}
$$

(10)

The Gibbs sampler proceeds by sequentially sampling from the complete conditional distributions of the parameters, which are proportional to the joint posterior (10). The complete conditional densities of the unknown parameters are given in the Appendix.

4.2 Details of sampling algorithms

Let $\boldsymbol{\Psi} = (\boldsymbol{V}_1, \cdots, \boldsymbol{V}_N, \boldsymbol{W})'$. Conditional on $\boldsymbol{\beta}_{it}$ and $\boldsymbol{\Psi}$, the structural and state equations of the HMDM have the form of a Gaussian Dynamic HDLM with *observations* $\boldsymbol{\beta}_{it}$ and state $\boldsymbol{\gamma}_t$. Note that $\boldsymbol{\gamma}_t$ is independent of $\boldsymbol{Y}_t$, given $\boldsymbol{\beta}_{it}$. Let $\boldsymbol{F}_t = (\boldsymbol{I}_{p_d}, \cdots, \boldsymbol{I}_{p_d})'$ denote an $Np_d \times p_d$ mapping matrix, and let $\boldsymbol{\beta}_t$ denote the vector obtained by stacking $\boldsymbol{\beta}_{it}$

for $i = 1, \cdots, N$. The complete conditional distribution of $\boldsymbol{\gamma}_t$ may be written as

$$(11) \quad P(\boldsymbol{\gamma}|\boldsymbol{\beta}, \boldsymbol{\Psi}) = P(\boldsymbol{\gamma}|\boldsymbol{\beta}, \boldsymbol{\Psi}) \prod_{t=1}^{T-1} P(\boldsymbol{\gamma}_t|\boldsymbol{\gamma}_{t+1}, \cdots, \boldsymbol{\gamma}_t, \boldsymbol{\beta}, \boldsymbol{\Psi})$$

The structure of the dynamic linear model implies that the second term on the right-hand side of the above equation reduces to $P(\boldsymbol{\gamma}_t|\boldsymbol{\gamma}_{t+1}, \boldsymbol{\beta}_t, \boldsymbol{\Psi})$. To generate a random sample from the complete conditional distribution of $\boldsymbol{\gamma}_t$, for $t = 1, \cdots, T$, using the Forward-Filtering-Backward-Sampling (FFBS) algorithm (Carter and Kohn, 1994, Fruhwirth-Schnatter, 1994), we implement the following steps:

Filter step: For $t = 1, \cdots, T$, we compute the mean $\boldsymbol{m}_t$ and variance matrix $\boldsymbol{C}_t$ of the posterior normal distributions $P(\boldsymbol{\gamma}_t|\boldsymbol{\gamma}_{t+1}, \boldsymbol{\beta}_t, \boldsymbol{\Psi})$, by applying the standard sequential updating results for Gaussian DLMs:

$$\begin{aligned}
\boldsymbol{m}_t &= \boldsymbol{G}\boldsymbol{m}_{t-1} + \boldsymbol{R}_t\boldsymbol{F}_t'\boldsymbol{Q}_t^{-1}(\boldsymbol{\beta}_t - \boldsymbol{F}_t\boldsymbol{G}\boldsymbol{m}_{t-1}) \\
\boldsymbol{C}_t &= \boldsymbol{R}_t - \boldsymbol{R}_t\boldsymbol{F}_t'\boldsymbol{Q}_t^{-1}\boldsymbol{F}_t\boldsymbol{R}_t \\
\boldsymbol{Q}_t &= \boldsymbol{V} + \boldsymbol{F}_t'\boldsymbol{R}_t\boldsymbol{F}_t \\
\boldsymbol{R}_t &= \boldsymbol{W} + \boldsymbol{G}\boldsymbol{C}_{t-1}\boldsymbol{G}',
\end{aligned}$$

where $\boldsymbol{V}$ is an $Np_d \times Np_d$ block diagonal matrix, each block represents the variance covariance matrix $\boldsymbol{V}_i$ for $i = 1, \cdots, N$.

Smooth step: At time $t = T$, we sample the final state vector $\boldsymbol{\gamma}_T$ from the marginal distribution, $P(\boldsymbol{\gamma}_T|\boldsymbol{\beta}_T, \boldsymbol{\Psi})$, which is $N(\boldsymbol{\gamma}_T|\boldsymbol{m}_T, \boldsymbol{C}_T)$. For time periods $t = T-1, \cdots, 0$, we sample from $P(\boldsymbol{\gamma}_t|\boldsymbol{\gamma}_{t+1}, \boldsymbol{\beta}_t, \boldsymbol{\Psi})$ which is $N(\boldsymbol{h}_t, \boldsymbol{H}_t)$ at each time, conditional on the latest value of $\boldsymbol{\gamma}_{t+1}$, where

$$(12) \qquad \boldsymbol{H}_t = (\boldsymbol{C}_t^{-1} + \boldsymbol{G}'\boldsymbol{W}^{-1}\boldsymbol{G})^{-1}$$

$$(13) \qquad \boldsymbol{h}_t = \boldsymbol{H}_t(\boldsymbol{C}_t^{-1}\boldsymbol{m}_t + \boldsymbol{G}'\boldsymbol{W}^{-1}\boldsymbol{\gamma}_{t+1})$$

The results of these steps is a draw $(\boldsymbol{\gamma}_T, \cdots, \boldsymbol{\gamma}_1)$ from its complete conditional distribution.

The complete conditional distributions of the variance terms are inverse Wishart distributions. Let $\widehat{\boldsymbol{\beta}}_{it} = \boldsymbol{\gamma}_t$. For each i, we sample $\boldsymbol{V}_i$ from $IW(n_i, \boldsymbol{S}_i)$, where

$$\begin{aligned}
n_i &= T + n_v \\
\boldsymbol{S}_i &= \boldsymbol{S}_v + \sum_{t=1}^{T} \left[(\boldsymbol{\beta}_{it} - \widehat{\boldsymbol{\beta}}_{it})(\boldsymbol{\beta}_{it} - \widehat{\boldsymbol{\beta}}_{it})'\right]
\end{aligned}$$

We sample $\boldsymbol{W}$ from $IW(n_w, \boldsymbol{S}_w)$ where

$$\begin{aligned}
n_w &= T + n_w \\
\boldsymbol{S}_w &= \boldsymbol{S}_w + \sum_{t=1}^{T} \left[(\boldsymbol{\gamma}_t - \boldsymbol{G}\boldsymbol{\gamma}_{t-1})(\boldsymbol{\gamma}_t - \boldsymbol{G}\boldsymbol{\gamma}_{t-1})'\right]
\end{aligned}$$

The Metropolis-Hastings algorithm is used for sampling $\boldsymbol{\beta}_{it}$ and $\boldsymbol{\eta}$.

The sampling based Bayesian framework for the univariate HDGLM model is similar to that of HMDM, but much simpler, since the conditional sampling distribution of Y_{it} is the univariate Poisson(λ_{it}) distribution. Since $m = q = 1$, this results in lower dimensional vectors and matrices, so that the computations become much faster.

5. MODEL SELECTION AND PREDICTION

We use data on the first T time points from all N locations for model fitting, and then make predictions for the next L times. Predictions for $\boldsymbol{\lambda}_{it}$ are obtained by using the output of the Gibbs sampler to approximate the predictive density $p(\boldsymbol{\lambda}|\boldsymbol{Y}, \boldsymbol{\gamma}, \boldsymbol{\beta}, \boldsymbol{\eta}, \boldsymbol{\Psi})$ using Monte Carlo integration. We also obtain the conditional predictive ordinates (CPO) for each i and t, and aggregate the values to obtain the pseudo-Bayes factor (PsBF) which provides a model selection criterion. The Mean Absolute Deviance (MAD) and the Prediction Mean Absolute Deviance (PMAD) criteria are used for evaluating fits/predictions in the calibration and hold-out data respectively, and are defined as

$$\text{MAD} = \frac{1}{N}\frac{1}{T}\sum_{i=1}^{N}\sum_{t=1}^{T} \text{sign}(y_{it} - \widehat{\lambda}_{it})\sqrt{d_{it}}$$

$$(14) \quad \text{PMAD} = \frac{1}{N}\frac{1}{L}\sum_{i=1}^{N}\sum_{t=T+1}^{T+L} \text{sign}(y_{it} - \widehat{\lambda}_{it})\sqrt{d_{it}}$$

where $\widehat{\lambda}_{it}$ denotes the posterior mean of λ_{it}, and d_{it} denotes an individual deviance contribution and is defined as

$$(15) \quad d_{it} = 2\big(\log(P_{\text{Poisson}}(y_{it}|y_{it})) - \log(P_{\text{Poisson}}(y_{it}|\hat{\lambda}_{it}))\big)$$

where $P_{\text{Poisson}}(y|\lambda)$ denote the univariate Poisson pmf. The MAD and PMAD values reported in sections 6 and 7 are averaged over the Gibbs iterations.

6. SIMULATED DATA RESULTS

Data from the univariate HDGLM with counts Y_{it} ($i = 1, \cdots, 16$ and $t = 1, \cdots, 44$) is simulated according to the following model:

$$\begin{aligned}
Y_{it}|\lambda_{it} &\sim \text{Poisson}(\lambda_{it}) \\
(16) \qquad \log\lambda_{it} &= \delta_{it,1} + \delta_{it,2} + \eta S_{it}
\end{aligned}$$

In (16), S_{it} represents a static predictor, which is simulated from Normal(0,1) and $\eta = 0.5$. Let $\boldsymbol{\beta_{it}} = (\delta_{it,1}, \delta_{it,2})'$ represent a location-time varying coefficient vector. The structural (hierarchical) and state equations are given by (7) and (8), with $\boldsymbol{G} = \text{diag}(1, -1)$. The errors $\boldsymbol{v}_{it}$ in (7) are simulated from $N_2(\boldsymbol{0}, \boldsymbol{V}_i)$ with $\boldsymbol{V}_i = \text{diag}(0.05, 0.01)$ for $i = 1, \cdots, 16$, and $\boldsymbol{w}_t$ from (8) is simulated from $N_2(\boldsymbol{0}, \boldsymbol{W})$ with $\boldsymbol{W} = \text{diag}(10^{-3}, 10^{-4})$. Here, $p_d = 2$ and $p_s = 1$.

Estimation via the Bayesian framework discussed in section 4.1 is carried out on the first $T = 40$ observations, by assuming that the prior distribution of $\boldsymbol{W} = \text{diag}(W_1, W_2)$ is an Inverse-Wishart distribution with $2p_d + 1$ degrees of

freedom and scale matrix $S_w = \mathrm{diag}(10^3, 10^4)$. Similarly the prior distribution of $\boldsymbol{V}_i = \mathrm{diag}(V_{1i}, V_{2i})$ is an Inverse-Wishart distribution with $2p_d + 1$ degrees of freedom and scale matrix $S_v = \mathrm{diag}(25, 100)$. To investigate prior sensitivity, the prior for η_1 is assumed to be $N(0, 10^3)$, $N(0, 10^2)$ or $N(0, 10)$ for simulation scenarios 1, 2 and 3 respectively. In scenario 5, we assume the prior scales $S_w = \mathrm{diag}(10^4, 10^4)$ and $S_v = \mathrm{diag}(10^3, 10^3)$, along with a $N(0, 10^3)$ prior for η. Further, we consider different starting values for the η sampling. Thus scenarios 1 and 4 denote starting from MLE's under a static regression model and 0 respectively, both under a $N(0, 10^3)$ prior. Based on each fitted model, predictions are obtained for the next $L = 4$ times.

Data from a bivariate HMDM with the vector of counts $\boldsymbol{Y}_{it} = (Y_{1,it}, Y_{2,it})'$ ($i = 1, \cdots, 16$ and $t = 1, \cdots, 44$) is simulated according to the following model:

$$
\begin{aligned}
\boldsymbol{Y}_{it}|\lambda_{1,it}, \lambda_{2,it}, \lambda_{3,it} &\sim MP_2(\lambda_{1,it}, \lambda_{2,it}, \lambda_{3,it}) \\
\log \lambda_{1,it} &= \delta_{1,it,1} + \delta_{1,it,2} + \eta_1 S_{it} \\
\log \lambda_{2,it} &= \delta_{2,it,1} + \delta_{2,it,2} + \eta_2 S_{it} \\
(17) \qquad \log \lambda_{3,it} &= \delta_{3,it,1}
\end{aligned}
$$

In (17), S_{it} represents a static predictor, which is simulated from $N(0,1)$ and $\eta_1 = 0.5$ and $\eta_2 = -0.5$. Let $\boldsymbol{\beta}_{it} = (\delta_{1,it,1}, \delta_{2,it,1}, \delta_{3,it,1}, \delta_{1,it,2}, \delta_{2,it,2})'$ represent location-time varying coefficients. The structural and state equations are again given by (7) and (8), with $\boldsymbol{G} = \mathrm{diag}(1,1,1,-1,-1)$. The errors $\boldsymbol{v}_{it}$ are simulated from $N_5(\boldsymbol{0}, \boldsymbol{V}_i)$ with $\boldsymbol{V}_i = \mathrm{diag}(\boldsymbol{V}_{1i}, \boldsymbol{V}_{2i})$, where $\boldsymbol{V}_{1i}$ is a symmetric matrix with diagonal elements 0.05 and off-diagonal elements 0.001, for $i = 1, \cdots, 16$. Likewise, $\boldsymbol{V}_{2i}$ is a symmetric matrix with diagonal elements 0.01 and off-diagonal elements 0.001 for $i = 1, \cdots, 16$. The errors $\boldsymbol{w}_t$ are simulated from $N_5(\boldsymbol{0}, \boldsymbol{W})$ with $\boldsymbol{W} = \mathrm{diag}(10^{-3}, 10^{-3}, 10^{-3}, 10^{-4}, 10^{-4})$. Here, $p_d = 5$ and $p_s = 2$.

The prior distribution of $\boldsymbol{W} = \mathrm{diag}(W_1, \cdots, W_5)$ is an Inverse-Wishart distribution with $2p_d + 1$ degrees of freedom and scale matrix $S_w = \mathrm{diag}(10^3, 10^3, 10^3, 10^4, 10^4)$. The prior distribution of $\boldsymbol{V}_i = \mathrm{diag}(\boldsymbol{V}_{1i}, \boldsymbol{V}_{2i})$ is an Inverse-Wishart distribution with $2p_d + 1$ degrees of freedom and scale matrix $S_v = \mathrm{diag}(25, 25, 25, 100, 100)$. To investigate prior sensitivity, the prior for $\boldsymbol{\eta} = (\eta_1, \eta_2)'$ is assumed to be $N_2(\boldsymbol{0}, \mathrm{diag}(10^3, 10^3))$, $N_2(\boldsymbol{0}, \mathrm{diag}(10^2, 10^2))$ or $N_2(\boldsymbol{0}, \mathrm{diag}(10, 10))$ for simulation scenarios 1, 2 and 3 respectively. The prior scales $S_w = \mathrm{diag}(10^4, 10^4, 10^4, 10^4, 10^4)$ and $S_v = \mathrm{diag}(10^3, 10^3, 10^3, 10^3, 10^3)$ are considered for scenario 5, along with a $N_2(\boldsymbol{0}, \mathrm{diag}(10^3, 10^3))$ prior for $\boldsymbol{\eta}$. We again consider different starting values for the $\boldsymbol{\eta}$ sampling. Thus, scenarios 1, 2, 3 and 5 denote starting from ML estimates from a static regression model, and scenario 4 from 0 corresponding to a $N_2(\boldsymbol{0}, \mathrm{diag}(10^3, 10^3))$ prior.

We ran a total of 40,500 Gibbs iterations; with a burn-in of 500 iterations, and thinning of 40 iterations to obtain a posterior sample of size 1,000. Summaries from the posterior distributions of the static coefficients are shown in Table 2.

Table 2. Posterior Summaries for HMDM–Simulated Data

Scenario	Parameter	True value	Posterior Mean	Posterior Std. Dev.
Univariate		HDGLM		
1	η	0.5	0.5206	0.0217
2	η	0.5	0.5208	0.0220
3	η	0.5	0.5209	0.0216
4	η	0.5	0.5204	0.0215
5	η	0.5	0.5215	0.0216
Bivariate		HMDM		
1	η_1	0.5	0.4371	0.0209
	η_2	-0.5	-0.3610	0.0301
2	η_1	0.5	0.4372	0.0210
	η_2	-0.5	-0.3611	0.0303
3	η_1	0.5	0.4371	0.0211
	η_2	-0.5	-0.3610	0.0303
4	η_1	0.5	0.4370	0.0211
	η_2	-0.5	-0.3608	0.0303
5	η_1	0.5	0.4525	0.0257
	η_2	-0.5	-0.3746	0.0298

The estimated coefficients for $\boldsymbol{\eta}$ were close to the true values of the parameters across all simulation scenarios. The estimates for the other model parameters were also found to be reasonably close to true values, but have been omitted from the table for the brevity. For simulation scenario 1 for the univariate HDGLM, MAD is 0.9080, PMAD is 0.9829 and log PsBF is 2.6046. For the bivariate HMDM, overall MAD and PMAD across both components of the count vector are respectively 0.8879 and 1.0389, while log PsBF is 3.3254. The $\boldsymbol{\eta}$ estimates are also reasonably sensitive to changes in priors for the variance parameters $\mathbf{W}$ and $\mathbf{V}_i$.

7. ANALYSIS OF GASTROPOD COUNTS

For univariate gastropod counts Y_{it}, an HDGLM is

$$
\begin{aligned}
Y_{it} &\sim \mathrm{Poisson}(\lambda_{it}) \\
\log \lambda_{it} &= \alpha_{it} + Season_{it} + \eta_1 LogRep_t + \eta_2 Elevation_i \\
&+ \eta_3 Slope_i + \eta_4 CO_{it} + \eta_5 PAsp_{it} + \eta_6 PAothers_{it} \\
&+ \eta_7 I(Aspect=5)_i + \eta_8 I(Aspect=7)_i + \eta_9 I(Aspect=8)_i \\
&+ \eta_{10} I(Soil\ Type=2)_i + \eta_{11} I(Soil\ Type=3)_i \\
&+ \eta_{12} I(CC=2)_i + \eta_{13} I(CC=3)_i \\
&= \boldsymbol{D}' \boldsymbol{\beta}_{it} + \boldsymbol{S}' \boldsymbol{\eta} = \boldsymbol{I} \boldsymbol{\beta}_{it} + \boldsymbol{S}' \boldsymbol{\eta}
\end{aligned}
$$
(18)

In (18), let $\boldsymbol{\beta}_{it} = (\alpha_{it}, Season_{it})'$ denote the location-time varying coefficients, let $\boldsymbol{\eta}$ denote a 13-dimensional vector of static coefficients and $\boldsymbol{S}$ denote the corresponding vector of exogenous predictors. The structural and state equations are given by (7) and (8), with $\boldsymbol{G} = \mathrm{diag}(1, -1)$. The specifications for Bayesian inference are what are shown for the univariate HDGLM in Section 6.

Table 3. Model Selection Criteria–Gastropod Data

Model	Species	MAD	PMAD	Log PsBF
Univariate	C. caracolla	1.2298	1.9130	3.9575
	G. nigrolineata	1.3363	2.4915	3.9990
Bivariate	Overall	1.2947	2.2087	4.6727
	C. caracolla	1.3424	1.9082	–
	G. nigrolineata	1.2470	2.5093	–

We next fit a bivariate HMDM to gastropod counts $\mathbf{Y}_{it} = (Carcar_{it}, Gaenig_{it})'$, where $Carcar_{it}$ and $Gaenig_{it}$ represent the abundance of C. caracolla and G. nigrolineata respectively at location i and time t. Let $\boldsymbol{\eta}_j$ correspond to a vector of static coefficients for $j = 1, 2$ and $\boldsymbol{S}$ denote the vector of static predictors as defined in (18). As described in Section 4, $p_d = 5$, and $p_s = 26$. The observation equation of the HMDM is

$$
\begin{aligned}
\mathbf{Y}_{it}|\lambda_{1,it}, \lambda_{2,it}, \lambda_{3,it} &\sim MP_2(\lambda_{1,it}, \lambda_{2,it}, \lambda_{3,it}) \\
\log \lambda_{1,it} &= \alpha_{1,it} + Season_{1,it} + \boldsymbol{S}'\boldsymbol{\eta}_1 \\
\log \lambda_{2,it} &= \alpha_{2,it} + Season_{2,it} + \boldsymbol{S}'\boldsymbol{\eta}_2 \\
\log \lambda_{3,it} &= \alpha_{3,it}
\end{aligned}
$$
(19)

The structural equation for $\boldsymbol{\beta}_{it} = (\alpha_{1,it}, \alpha_{2,it}, \alpha_{3,it}, Season_{1,it}, Season_{2,it})'$ is given in (7) and the state equation is given in (8) with $\mathbf{G} = \text{diag}(1, 1, 1, -1, -1)$.

We ran a total of 40,500 Gibbs iterations; with a burn-in of 500 iterations, and thinning of 40 iterations, to obtain a posterior sample of size 1,000. Table 3 shows prediction evaluation criteria for the two gastropod species, which are quite similar for the univariate and bivariate models. As before, MAD corresponds to the calibration data and PMAD corresponds to the holdout data.

The posterior means and standard deviations of the static coefficients for the univariate and bivariate models are given in Table 4. The posterior means for the first three diagonal elements of the matrix $\mathbf{W}$ are estimated to be 0.0011 with posterior standard deviation of about 0.0003 each, while the other two elements are estimated as 0.0001 with posterior standard deviation of around 0.00002 each. The estimated posterior means for the first three diagonal elements of the matrix $\mathbf{V}_i$ are 0.044 with posterior standard deviations around 0.01, while the other diagonal elements are estimated around 0.01 with posterior standard deviation 0.002 each. All the estimated off-diagonal elements of $\mathbf{W}$ and $\mathbf{V}_i$ reveal no significant values for $i = 1, \cdots, 40$. Recall that the first three diagonal elements of $\mathbf{W}$ and $\mathbf{V}_i$ for $i = 1, \cdots, 40$ are associated with the location-specific and time-dependent intercepts, and the last two elements are associated with the location-specific seasonal components (with period equal to two, corresponding to the dry and wet periods).

Figure 1 shows the fitted means for the counts of C. caracolla and G. nigrolineata from Site 14 (top panel) and from Site 22 (bottom panel) based on the bivariate

HMDM model. Clearly, the fitted model adequately tracks the stochastic pattern over time for almost all time points, especially for C. caracolla. The fit for G. nigrolineata is also good, although at a few time points, the model is unable to track the large empirical counts. This suggests that the abundance of G. nigrolineata is affected by environmental characteristics beyond those measured in this study, at least during some time periods. The significance of coefficients from the univariate and bivariate models (Table 4) are generally in accord for C. caracolla. Aspect and soil type have the strongest influence; the logarithm of the number of replicates has a modest influence, and other environmental characteristics (i.e., elevation, slope, canopy openness, canopy cover class, apparency of sierra palm, and plant apparency of species other than sierra palm) have little to no influence on variation in abundance. More specifically, Cristal soils as well as Preito soils support more individuals than do Zarsal soils; these effects are likely mediated by differences in the nutrient characteristics of the soils and the plant assemblages that they harbor. All of the contrasts associated with aspect (over the baseline aspect level 1) are significant. Such effects could arise via two mechanisms: (i) abiotic effects of aspect on gastropods via influences on microclimate, especially diurnal temperature as well as (ii) biotic effects mediated by the well-documented responses of plant assemblages to aspect (i.e., those associated with insolation and variation in temporal patterns of temperature), especially in topographically steep environments, such as the Luquillo Mountains. Finally, the effect of logarithm of the number of replicates is to be expected, as the MNKA, the methodology used to estimate gastropod abundance, can only increase with increasing effort.

For G. nigrolineata, significance based on the univariate framework (see Table 4) are similar, but not identical, to those based on the bivariate model. In general, differences among sites in soil type and canopy cover class have a modest influence on abundance, whereas all other environmental characteristics have little or no significant effect. Abundance of G. nigrolineata on Cristal soils was lower than that on Zarzal soils, whereas abundance of G. nigrolineata on Prieto soils were higher than that on Zarzal soils. Moreover, abundance on sites within cover class 2 (50–80% canopy openness in 1936) was greater than that on cover class 1 (0–49% canopy openness in 1936). The non-significant effect and negative value of the parameter estimate for the influence of the logarithm of the number of replicates on abundance is unexpected. This may be because the suite of environmental characteristics in this study do not include all variables to which G. nigrolineata responds most strongly (i.e., random effects capture variation in abundances associated with unmeasured environmental characteristics), or may arise because of time lags in the response of this species to environmental variability. This is something to investigate in a future study of this species.

In summary, a comparison of static coefficients in Table 4 do not show a marked difference between the univariate and

Table 4. Posterior Summaries for HMDM–Gastropod Data

Gastropod Species	Parameter η	Univariate Model Posterior Mean (Std. Dev.)	Bivariate Model Posterior Mean (Std. Dev.)
C. caracolla	LogRep	0.5492 (0.0668)	0.5650 (0.0705)
	Elevation	0.0011 (0.0007)	0.0001 (0.0005)
	Slope	0.0046 (0.0018)	0.0038 (0.0017)
	CO	0.0227 (0.0082)	0.0247 (0.0083)
	PAsp	-0.0071 (0.0015)	-0.0067 (0.0015)
	PAothers	-0.0017 (0.0009)	-0.0022 (0.0009)
	Aspect=5	1.0305 (0.0736)	1.0146 (0.0844)
	Aspect=7	0.8145 (0.0742)	0.7998 (0.0845)
	Aspect=8	0.4865 (0.0727)	0.4761 (0.0834)
	Soil Type=2	0.6975 (0.0331)	0.6827 (0.0327)
	Soil Type=3	1.8029 (0.0485)	1.7355 (0.0443)
	CC=2	0.2783 (0.0400)	0.2464 (0.0373)
	CC=3	0.0642 (0.0339)	0.0598 (0.0347)
G. nigrolineata	LogRep	-0.0644 (0.0934)	-0.0911 (0.0876)
	Elevation	0.0049 (0.0009)	0.0039 (0.0010)
	Slope	-0.0006 (0.0026)	-0.0033 (0.0025)
	CO	-0.1419 (0.0133)	-0.1328 (0.0128)
	PAsp	0.0194 (0.0018)	0.0193 (0.0019)
	PAothers	-0.0069 (0.0012)	-0.0061 (0.0012)
	Aspect=5	0.0447 (0.0951)	-0.1296 (0.0875)
	Aspect=7	-0.0207 (0.0934)	-0.0297 (0.0809)
	Aspect=8	0.1904 (0.0910)	0.2022 (0.0798)
	Soil Type=2	-0.1716 (0.0498)	-0.1644 (0.0525)
	Soil Type=3	0.2653 (0.0852)	0.1308 (0.0872)
	CC=2	0.3778 (0.0469)	0.2856 (0.0475)
	CC=3	0.1254 (0.0491)	0.1370 (0.0511)

bivariate models for the two species. The top plot in Figure 2 shows that in the bivariate model, the estimated dynamic level coefficient corresponding to the main effect for *C. caracolla* exhibits a slowly increasing trend over time, while for *G. nigrolineata*, the trend is slowly decreasing over time. The estimated dynamic level coefficient corresponding to the second-order covariance effect is almost flat over time, indicating that the dependence between the species does not change over time. The bottom plot shows the decreasing amplitudes of the dynamic seasonal components for both species; the wet season shows an increased level for *C. caracolla* while the dry season shows an increased level for *G. nigrolineata*. The bivariate analysis is useful since it allows us to examine the temporal behavior of the second-order covariance effect, whose behavior would guide the modeler in terms of preferring a univariate or a multivariate model for understanding variation in species abundances. In this application for the two gastropod species, this effect is approximately constant over time, which explains why the results from the univariate and bivariate models are similar.

8. SUMMARY AND DISCUSSION

This article describes a hierarchical dynamic modeling framework for univariate and multivariate time series of counts, and investigates its utility for an ecological data set on gastropod abundances in Puerto Rico. We describe the use of the multivariate Poisson distribution as the sampling distribution in a fully Bayesian framework for inference in the context of multivariate count time series. Our approach is able to account for species-specific variation in gastropod abundances, and documents the importance of particular environmental characteristics in molding the abundance and distribution of organisms, a primary goal in ecology.
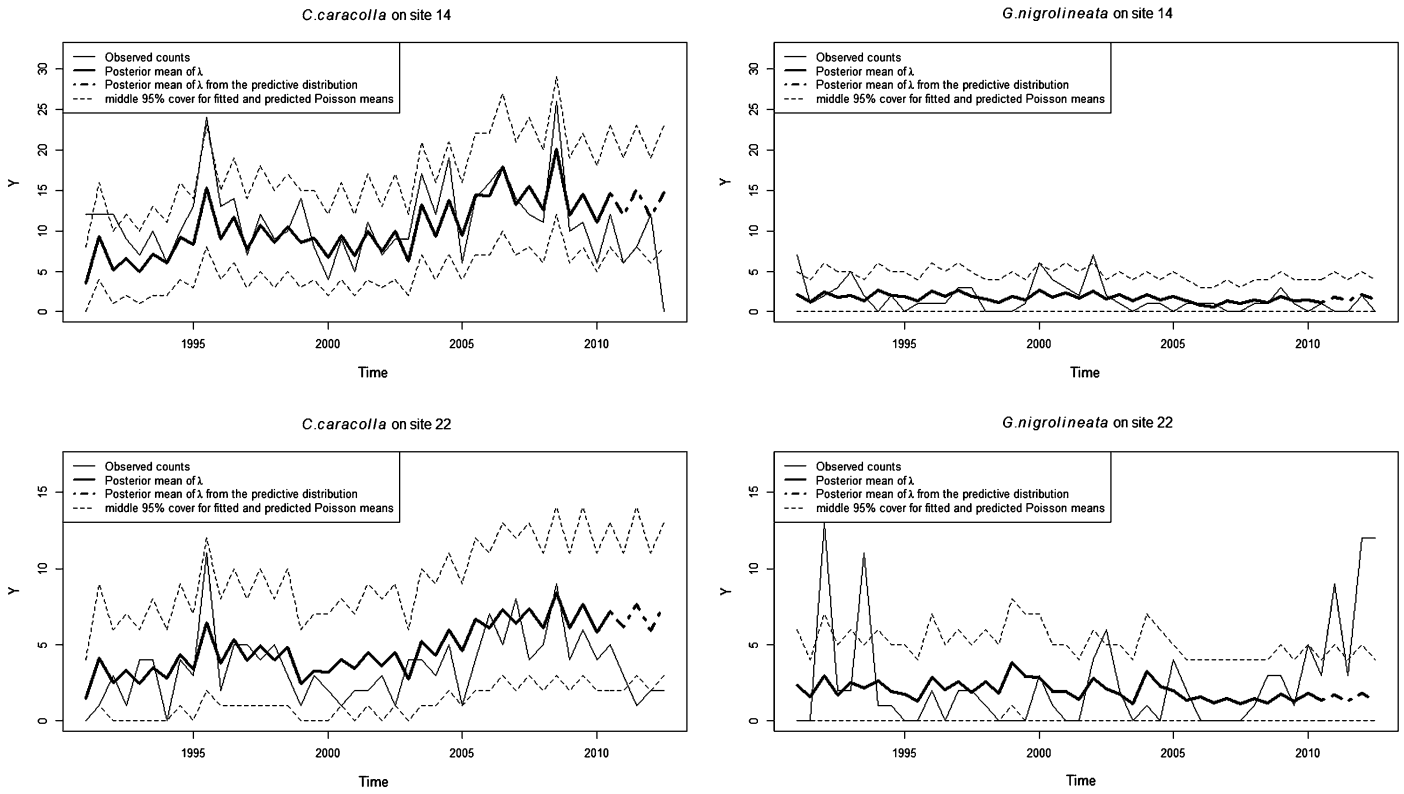
Figure 1. Model fits for counts of C. caracolla and G. nigrolineata on sites 14 and 22.
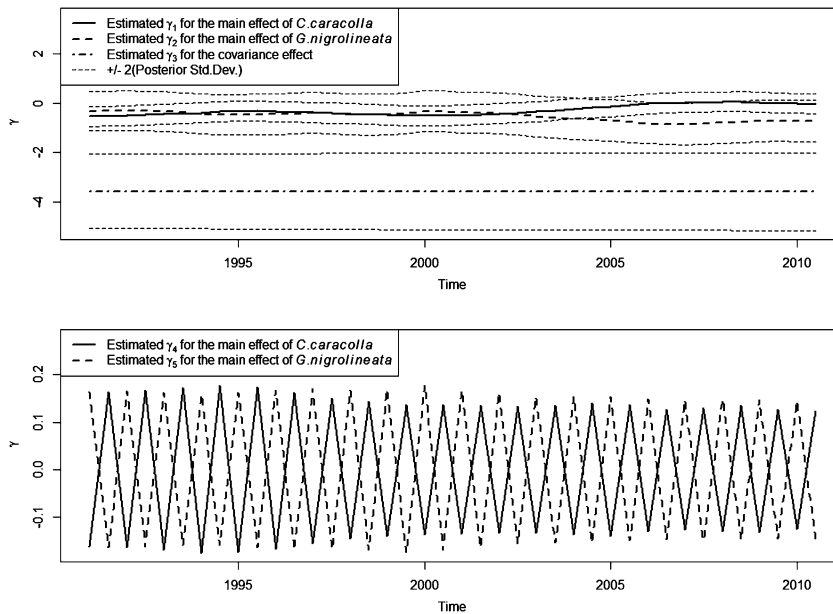
Figure 2. Estimated dynamic components of γ for C. caracolla and G. nigrolineata.

APPENDIX

The complete conditional density of $\boldsymbol{\beta}_{it}$ for $i = 1, \cdots, N$ and $t = 1, \cdots, T$ is

$$\pi(\beta_{it}|\boldsymbol{\gamma}_t, \boldsymbol{\eta}, \boldsymbol{V}_i, \boldsymbol{W}, \boldsymbol{Y}) \propto MP_m(\boldsymbol{y}_{it}|\boldsymbol{\lambda}_{it})$$

$$(20) \qquad \times \quad \exp\{-\frac{1}{2}(\boldsymbol{\beta}_{it} - \boldsymbol{\gamma}_t)'\boldsymbol{V}_i^{-1}(\boldsymbol{\beta}_{it} - \boldsymbol{\gamma}_t)\}$$

The complete conditional density of $\boldsymbol{\gamma}_t$ for $t = 1, \cdots, T$ is

$$\pi(\boldsymbol{\gamma}_t|\boldsymbol{\beta}_{it}, \boldsymbol{\eta}, \boldsymbol{V}_i, \boldsymbol{W}, \boldsymbol{Y}) \propto$$

$$\propto \quad \prod_{i=1}^{N}\exp\{-\frac{1}{2}(\boldsymbol{\beta}_{it} - \boldsymbol{\gamma}_t)'\boldsymbol{V}_i^{-1}(\boldsymbol{\beta}_{it} - \boldsymbol{\gamma}_t)\}$$

$$\times \quad \exp\{-\frac{1}{2}(\boldsymbol{\gamma}_t - \boldsymbol{G}\boldsymbol{\gamma}_{t-1})'\boldsymbol{W}^{-1}(\boldsymbol{\gamma}_t - \boldsymbol{G}\boldsymbol{\gamma}_{t-1})\}$$

$$(21) \quad \times \quad \exp\{-\frac{1}{2}(\boldsymbol{\gamma}_{t+1} - \boldsymbol{G}\boldsymbol{\gamma}_t)'\boldsymbol{W}^{-1}(\boldsymbol{\gamma}_{t+1} - \boldsymbol{G}\boldsymbol{\gamma}_t)\}$$

The complete conditional density of $\boldsymbol{\eta}$ is

$$\pi(\boldsymbol{\eta}|\boldsymbol{\beta}_{it}, \boldsymbol{\gamma}_t, \boldsymbol{V}_i, \boldsymbol{W}, \boldsymbol{Y}) \propto \prod_{t=1}^{T}\prod_{i=1}^{N} MP_m(\boldsymbol{y}_{it}|\boldsymbol{\lambda}_{it})$$

$$(22) \quad \times \quad \exp\{-\frac{1}{2}(\boldsymbol{\eta} - \boldsymbol{\mu}_\eta)'\boldsymbol{\Sigma}_\eta^{-1}(\boldsymbol{\eta} - \boldsymbol{\mu}_\eta)\}$$

The complete conditional density of $\boldsymbol{V}_i$ for $i = 1, \cdots, N$ is

$$\pi(\boldsymbol{V}_i|\boldsymbol{\beta}_{it}, \boldsymbol{\gamma}_t, \boldsymbol{\eta}, \boldsymbol{W}, \boldsymbol{Y}) \propto \prod_{t=1}^{T}|\boldsymbol{V}_i|^{-1/2}$$

$$\times \quad \exp\{-\frac{1}{2}(\boldsymbol{\beta}_{it} - \boldsymbol{\gamma}_t)'\boldsymbol{V}_i^{-1}(\boldsymbol{\beta}_{it} - \boldsymbol{\gamma}_t)\}$$

$$(23) \qquad \times \quad |\boldsymbol{V}_i|^{-n_v/2}\exp\{-\frac{1}{2}tr(\boldsymbol{V}_i^{-1}\boldsymbol{S}_v)\}$$

The complete conditional density of $\boldsymbol{W}$ is

$$\pi(\boldsymbol{W}|\boldsymbol{\beta}_{it}, \boldsymbol{\gamma}_t, \boldsymbol{\eta}, \boldsymbol{V}_i, \boldsymbol{Y}) \propto \prod_{t=1}^{T}|\boldsymbol{W}|^{-1/2}$$

$$\times \quad \exp\{-\frac{1}{2}(\boldsymbol{\gamma}_t - \boldsymbol{G}\boldsymbol{\gamma}_{t-1})'\boldsymbol{W}^{-1}(\boldsymbol{\gamma}_t - \boldsymbol{G}\boldsymbol{\gamma}_{t-1})\}$$

$$\times \quad |\boldsymbol{W}|^{-1/2}\exp\{-\frac{1}{2}(\boldsymbol{\gamma}_0 - \boldsymbol{m}_0)'\boldsymbol{C}_0^{-1}(\boldsymbol{\gamma}_0 - \boldsymbol{m}_0)\}$$

$$(24) \quad \times \quad |\boldsymbol{W}|^{-n_w/2}\exp\{-\frac{1}{2}tr(\boldsymbol{W}^{-1}\boldsymbol{S}_w)\}$$

REFERENCES

AL-OSH, M. A. and ALZAID, A. A. (1987). First-order integer-valued autoregressive (INAR(1)) model. *J. Time Ser. Anal.*, **8**, 261–275. MR0903755

BERGER, J. (2006). The Case for Objective Bayesian Analysis. *Bayesian Analysis*, **1**, 385–402. MR2221271

BLOCH, C. and WILLIG, M. (2006). Context-dependence of long-term responses of terrestrial gastropod populations to large-scale disturbance. *Journal of Tropical Ecology*, **22**, 111–122.

BLOCH, C. and WILLIG, M. (2012). Density compensation suggests interspecific competition is weak among terrestrial snails in tabonuco forest of Puerto Rico. *Caribbean Journal of Science*, **46**, 159–168.

BROWN, S., LUGO, A., SILANDER, S. and LIEGEL, L. (1983). Research history and opportunities in the Luquillo Experimental Forest. *General Technical Report SO–44. United States Department of Agriculture, Forest Service, New Orleans.*

CARLIN, B., POLSON, N. and STOFFER, D. (1992). A Monte Carlo Approach to Nonnormal and Nonlinear State-Space Modeling. *Journal of the American Statistical Association*, **87–418**, 493–500.

CARTER, C. and KOHN, R. (1994). On Gibbs Sampling for State-Space Models. *Biometrika*, **81**, 541–553. MR1311096

CHEN, M.-H., SHAO, Q.-M. and IBRAHIM, J. (2000). *Monte Carlo Methods in Bayesian Computation.* Springer-Verlag: New York. MR1742311

CHIB, S., GREENBERG, E. and WINKELMANN, R. (1998). Posterior simulation and Bayes factors in panel count data models. *Journal of Econometrics*, **86**, 33–54.

CHIB, S. and WINKELMANN, R. (2001). Markov Chain Monte Carlo Analysis of Correlated Count Data. *Journal of Business and Economic Statistics*, **19**, 428–435. MR1947077

COOK, A. (2001). Behavioural ecology: on doing the right thing, in the right place, at the right time. *The biology of terrestrial mollusks. Wallingford: CAB International*, 447–487.

COOK, C. and STUBBENDIECK, J. (1986). Range research: basic problems and techniques. *Society for Range Management, Denver, Colorado.*

DAVIS, R. A., DUNSMUIR, W. T. M. and STREETT, S. B. (2003). Observation-driven models for Poisson counts. *Biometrika*, **90**, 777–790. MR2024757

DURBIN, J. and KOOPMAN, S. J. (2000). Time Series Analysis of Non-Gaussian Observations based on state space models from both classical and Bayesian perspectives. *Journal of the Royal Statistical Society: Series B*, **62**, 3–56. MR1745604

FRUHWIRTH-SCHNATTER, S. (1994). Data Augmentation and Dynamic Linear Models. *Journal of Time Series Analysis*, **24**, 295–320. MR1263889

FRUHWIRTH-SCHNATTER, S. and WAGNER, H. (2006). Auxiliary mixture sampling for parameter-driven models of time series of counts with applications to state space modeling. *Biometrika*, **93**, 827–841. MR2285074

GAMERMAN, D. (1998). Markov Chain Monte Carlo for Dynamic Generalized Linear Models. *Biometrika*, **85**, 215–227. MR1627273

GAMERMAN, D. and MIGON, H. (1993). Dynamic Hierarchical Models. *Journal of the Royal Statistical Society: Series B*, **3**, 629–642. MR1223932

GAMERMAN, D., SANTOS, T. R. and FRANCO, G. C. (2013). A non-Gaussian Family of State-space Models with Exact Marginal Likelihood. *J. Time Ser. Anal.*, **34**, 625–645. MR3127211

HU, S. (2012). Dynamic Modeling of Discrete-valued Time Series with Applications. Ph.D. dissertation, University of Connecticut. MR3121942

IBRAHIM, J. and CHEN, M. H. (2000). Power prior distributions for regression models. *Statistical Science*, **15**, 1–94. MR1842236

JOHNSON, N., KOTZ, S. and BALAKRISHNAN, N. (1997). *Discrete Multivariate Distributions*, Wiley: New York. MR1429617

JORGENSEN, B., LUNDBYE-CHRISTENSEN, S., SONG, P. X-K. and SUN, LI. (1999). A State Space Model for Multivariate Longitudinal Count Data. *Biometrika*, **86**, 169–181. MR1688080

JUNG, R. C. and TREMAYNE, A. R. (2006). Binomial thinning models for integer time series. *Statistical Modelling*, **6**, 21–96. MR2242181

KALMAN, R. (1960). A New Approach to Linear Filtering and Prediction Theory. *Transactions of the ASME. Series D, Journal of Basic Engineering*, **82**, 35–45.

KALMAN, R. and BUCY, R. (1961). New Results in Filtering and Prediction Theory. *Transactions of the ASME. Series D, Journal of Basic Engineering*, **83**, 95–108. MR0234760

KARLIS, D. and MELIGKOTSIDOU, L. (2005). Multivariate Poisson Regression with Covariance Structure. *Statistics and Computing*, **15**, 255–265. MR2205389

KARLIS, D. and MELIGKOTSIDOU, L. (2007). Finite Mixtures of Multivariate Poisson Regression with Application. *Journal of Statistical Planning and Inference*, **137**, 1942–1960. MR2323875

KREBS, C. (1972). Ecology: the experimental analysis of distribution and abundance. *Harper and Row: New York.*

LANDIM, F. and GAMERMAN, D. (2000). Dynamic hierarchical models: an extension to matrix-variate observations. *Computational Statistics and Data Analysis*, **35**, 11–42. MR1802963

LINDLEY, D. and SMITH, A. (1972). Bayes Estimates for the Linear Model (with discussion). *Journal of the Royal Statistical Society: Series B*, **34**, 1–41. MR0415861

LYDEARD, C. et al. (2004). The global decline of nonmarine mollusks. *BioScience*, **54**, 321–330.

MAHAMUNULU, D. (1967). A note on Regression in the Multivariate Poisson Distribution. *Journal of the American Statistical Association*, **62**, 251–258. MR0216613

MCCULLAGH, P. and NELDER, J. (1989). *Generalized Linear Models*, 2nd Edition, Chapman and Hall: London. MR0727836

MCDOWELL, W., SCATENA, F., WAIDE, R., BROKAW, N., CAMILO, G., COVICH, A., CROWL, T., GONZALEZ, G., GREATHOUSE, E., KLAWINSKI, P., LODGE, D., LUGO, A., PRINGLE, C., RICHARDSON, B., RICHARDSON, M., SCHAFER, D., SILVER, W., THOMPSON, J., VOGT, D., VOGT, K., WILLIG, M., WOOLBRIGHT, L., ZOU, X. and ZIMMERMAN, J. (2012). Geographic and ecological setting of the Luquillo Mountains. *A Caribbean Forest Tapestry: The Multidimensional Nature of Disturbance and Response. Oxford University Press, New York, New York*, 72–163.

MCKENZIE, E. (1985). Some simple models for discrete variate time series. *Water Resources Bulletin*, **21**, 645–50.

MCKENZIE, E. (2003). Discrete variate time series. In Handbook of Statistics, **21**, Stochastic models: Modeling and Estimation, eds C. R. Rao and D. N.Shanbhag, Amsterdam: Elsevier Science, 573–606. MR1973555

O'HAGAN, A., BUCK, C. E., DANESHKHAH, A., EISER, R., GARTHWAITE, P., JENKINSON, D., OAKLEY, J. and RAKOW, T. (2006). *Uncertain Judgements: Eliciting Experts' Probabilities.* Wiley: New York.

PEDELI, X. and KARLIS, D. (2011). A bivariate INAR(1) model with application. *Statistical Modelling*, **11**, 325–349. MR2906704

PEDELI, X. and KARLIS, D. (2012). On composite likelihood estimation of a multivariate INAR(1) model. *J. Time Ser. Anal.*, **33**, 903–915. MR3028367

SCHEINER, S. and WILLIG, M. (2011). *Theory of Ecology.* University of Chicago Press: Chicago, Illinois.

SECREST, M. (1995). The impact of Hurricane Hugo on two common tree snails in the Luquillo experimental forest of Puerto Rico: a long-term study. Masters thesis, Texas Tech University, Lubbock, Texas.

SECREST, M., WILLIG, M. and PEPPERS, L. (1996). The legacy of disturbance on habitat associations of terrestrial snails in the Luquillo Experimental Forest, Puerto Rico. *Biotropica*, **28**, 502–514.

SERHIYENKO, V., IVAN, J, RAVISHANKER, N. and ISLAM, S. (2014). Dynamic Compositional Modeling of Pedestrian Crash Counts on Urban Roads in Connecticut. *Accident Analysis & Prevention*, **64**, 78–85.

STEUTEL, F. W. and VAN HARN, K. (1979) Discrete analogues of self–decomposability and stability. *Annals of Probability*, **7**, 893–899. MR0542141

THOMPSON, J., BROKAW, N., ZIMMERMAN,J., WAIDE, R., EVERHAM, E., LODGE, D., TAYLOR, C., GARCIA-MONTIEL, D. and FLUET, M.

(2002). Land use history, environment, and tree composition in a tropical forest. *Ecological Applications*, **12**, 1344–1363.

TSIAMYRTZIS, P. and KARLIS, D. (2004). Strategies for Efficient Computation of Multivariate Poisson Probabilities. *Communications in Statistics–Simulation and Computation*, **33**, 271–292. MR2060113

VENKATESAN, RAJKUMAR, REINARTZ, W. and RAVISHANKER, N. (2012). The Role of Attitudinal Information in CLV-based Customer Management. *MSI Working Paper Series*, 12–107.

WEST, M. and HARRISON, P. (1989). Bayesian Forecasting and Dynamic Models. *Springer, New York.* MR1020301

WILLIG, M., SANDLIN, E. and GANNON, M. (1998a). Structural and taxonomic correlates of habitat selection by a Puerto Rican land snail. *Southwestern Naturalist*, **43**, 70–79.

WILLIG, M, SECREST, M., COX, S., CAMILO, G., CARY, J., ALVAREZ, J. and GANNON, M. (1998b). Long-term monitoring of snails in the Luquillo Experimental Forest of Puerto Rico: Heterogeneity, scale, disturbance, and recovery. *The Parthenon Press, Cranforth, Lancashire, UK*, 293–322.

WILLIG, M., BLOCH, C., BROKAW, N., HIGGINS, C., THOMPSON, J. and ZIMMERMANN, C. (2007). Cross-scale responses of biodiversity to hurricane and anthropogenic disturbance in a tropical forest. *Ecosystems*, **10**, 824–838.

WILLIG, M., PRESLEY, S., BLOCH, C., CASTRO-ARELLANO, I., CISNEROS, L., HIGGINS, C. and KLINGBEIL, B. (2011). Tropical metacommunities and elevational gradients: effects of forest type and other environmental factors. *Oikos*, **120**, 1497–1508.

WILLIG, M., PRESLEY, S., BLOCH, C. and ALVAREZ J. (2012). Population, community, and metacommunity dynamics of terrestrial gastropods in the Luquillo Mountains: a gradient perspective. *Ecological Gradient Analyses in a Tropical Ecosystem. Ecological Bulletins*, **53**.

ZEGER, S. L. (1988). A regression model for time series of counts. *Biometrika*, **75**, 621–629. MR0995107

Nalini Ravishanker
U-4120 University of Connecticut
Department of Statistics
215 Glenbrook Rd.
Storrs, CT 06269
USA
E-mail address: nalini.ravishanker@uconn.edu

Volodymyr Serhiyenko
U-4120 University of Connecticut
Department of Statistics
215 Glenbrook Rd.
Storrs, CT 06269
USA
E-mail address: volodymyr.serhiyenko@uconn.edu

Michael R. Willig
Center for Environmental Sciences and Engineering
University of Connecticut
3107 Horsebarn Hill Road
Building 4 Annex
Storrs, CT 06269
USA
E-mail address: michael.willig@uconn.edu