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, 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.


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 (Ganerman 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 analyzes 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 agile, 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; 18°20 N, 65°49 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 Gaeolus nigrolinnea, 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).

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 type levels are Zarzal (1), Cristal (2), or Prieto (3). 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 $g : N^m \rightarrow N^m, q \geq m$, such that $Y = g(X) = AX$ (Mahamunluu, 1967; Johnson et al., 1997). Here, $X = (X_1, \ldots, X_q)'$ is a vector of unobserved independent Poisson random variables, i.e., $X_r \sim \text{Poisson}(\lambda_r)$ for $r = 1, \ldots, q$, and $A$ is an arbitrary $m \times q$ matrix which determines the properties of the multivariate Poisson distribution. The $m$-dimensional vector $Y = (Y_1, \ldots, Y_m)' = AX$ follows a multivariate Poisson distribution with parameters $\lambda = (\lambda_1, \ldots, \lambda_q)'$ and pmf

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

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

where $g^{-1}(Y)$ denotes the inverse image of $Y \in N^m$ 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 $A = [A_1, A_2]$, where $A_1 = I_m$ captures the main effects; $A_2$ captures the two-way covariance effects; $A_2$ is an $m \times (m(m-1)/2)$ binary matrix; each column of $A_2$ 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 $\lambda$ into two parts, viz., $\lambda^{(1)} = (\lambda_1, \cdots, \lambda_m)'$, which corresponds to the $m$ main effects, and $\lambda^{(2)} = (\lambda_{m+1}, \cdots, \lambda_q)'$ 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 $Y = (Y_1, Y_2)'$ is expressed via $q = 3$ independent Poisson random vari-
ables as:
\[ Y_1 = X_1 + X_3 \]
\[ Y_2 = X_2 + X_3 \]

where \( X_i \sim \text{Poisson}(\lambda_i), \) \( i = 1, 2, 3. \) The joint pmf of \( Y_1 \) and \( Y_2 \) is:
\[
P(Y_1 = y_1, Y_2 = y_2 | X) = \exp\{-\sum_{i=1}^{6} \lambda_i\} \times \sum_{(x_1, x_2, x_3) \in C} \frac{\lambda_1^{y_1} \lambda_2^{y_2} \lambda_3^{y_3}}{y_1! \times y_2! \times y_3!} \cdot \frac{\lambda_4^{x_4} \lambda_5^{x_5} \lambda_6^{x_6}}{(x_4 + x_5 + x_6)!} \cdot \frac{\lambda_1^{y_1} \lambda_2^{y_2} \lambda_3^{y_3}}{(y_1! \times y_2! \times y_3! \times y_4! \times y_5! \times y_6!)}
\]

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

Similarly, when \( m = 3, \) we write \( Y = (Y_1, Y_2, Y_3)' \) as:
\[
Y_1 = X_1 + X_4 + X_5
\]
\[
Y_2 = X_2 + X_4 + X_6
\]
\[
Y_3 = X_3 + X_5 + X_6
\]

where \( X_i \sim \text{Poisson}(\lambda_i) \) for \( i = 1, \ldots, 6. \) The joint probability mass function of \( Y_1, Y_2, \) and \( Y_3 \) is:
\[
P(Y = y_1, y_2, y_3, \lambda) = \exp\{-\sum_{i=1}^{6} \lambda_i\} \times \sum_{(x_1, x_2, x_3, x_4) \in C} \frac{\lambda_1^{y_1} \lambda_2^{y_2} \lambda_3^{y_3} \lambda_4^{x_4} \lambda_5^{x_5} \lambda_6^{x_6}}{(y_1! \times y_2! \times y_3! \times y_4! \times y_5! \times y_6! \times x_4! \times x_5! \times x_6!)}
\]

where the summation is over the set \( C \) such that \( C = \{\{X_1, X_2, X_3\} : (X_1 + X_4 + X_5) \leq y_1 \cap (X_2 + X_4 + X_6) \leq y_2 \cap (X_3 + X_5 + X_6) \leq y_3\} \neq \emptyset\} \]

It is easy to see that the matrix \( A \) has the respective forms shown below for \( m = 2 \) and \( m = 3: \)
\[
\begin{pmatrix}
1 & 0 \\
0 & 1
\end{pmatrix}
\]

\[
\begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 1
\end{pmatrix}
\] and
\[
\begin{pmatrix}
1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1
\end{pmatrix}
\]

The mean vector and variance-covariance matrix of \( Y \) are given by \( E(Y|X) = AX \) and \( \text{Cov}(Y|X) = AXA' \), where \( \Sigma = \text{diag}(\lambda_1, \ldots, \lambda_6). \) As pointed out by Karlis and Meligkotsidou (2005), this covariance cannot accommodate negative associations among the components of \( Y \). When \( m = 1, \) the \( MP_{m}\) 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 \( y_1 \leq y_2, \) 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., \( x_3 = 0, \ldots, \min(y_1, y_2), \) 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 \( y, \) i.e., \( x \in g^{-1}(y). \) The pmf for the bivariate Poisson distribution can be calculated using (3).

When \( m = 3, \) without loss of generality, we assume that \( y_1 \leq y_2 \leq y_3. \) The possible values for \( x_4 \) and \( x_5 \) are in the set \( C_1 = \{0, \ldots, y_1\}, \) and the possible values for \( x_6 \) are in the set \( C_2 = \{0, \ldots, y_2\}. \) We have in total \( K \) different combinations for \( (x_4, x_5, x_6), \) where \( K = (\text{length of set } C_1)^2 \times (\text{length of set } C_2) = (y_1 + 1)^2(y_2 + 1). \) 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 \( Y. \) 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 \( \lambda \)'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 \( Y \) 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, \( \lambda_1 = 2, \lambda_2 = 3, \lambda_3 = 2, \lambda_4 = \lambda_5 = \lambda_6 = 1. \) Under scenario 3, \( \lambda_1 = \lambda_2 = \lambda_3 = 8, \lambda_4 = \lambda_5 = \lambda_6 = 2. \) 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>
<thead>
<tr>
<th>Likelihood</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate Poisson</td>
<td>8 × 10⁻³</td>
<td>8 × 10⁻³</td>
<td>8.001 × 10⁻³</td>
</tr>
<tr>
<td>Bivariate Poisson</td>
<td>8 × 10⁻³</td>
<td>1.2 × 10⁻²</td>
<td>2.4 × 10⁻²</td>
</tr>
<tr>
<td>Trivariate Poisson</td>
<td>1.412</td>
<td>18.209</td>
<td>248.788</td>
</tr>
</tbody>
</table>

562. N. Ravishanker, V. Serhiyenko, and M. R. Willig
4. HIERARCHICAL MULTIVARIATE DYNAMIC MODEL (HMDM)

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

\[
Y_{it} | \lambda_{it} \sim MP_{m}(y_{it} | \lambda_{it})
\]

where the errors \( \varepsilon_{it} \) which represents the static intercept, i.e., either \( \delta_{it} \) for the main effects portion of the multivariate Poisson specification (1), and (7) is

\[
\log \lambda_{j, it} = D_{j, it}' \delta_{j, it} + S_{j, it}' \eta_{j} + v_{j, it}, \quad j = 1, \ldots, q
\]

where \( \log \) denotes the natural logarithm, \( D_{j, it} = (D_{j, it, 1}, \ldots, D_{j, it, a})' \) is an \( a_j \)-dimensional vector of exogenous predictors with location-time varying (dynamic) coefficients \( \delta_{j, it} = (\delta_{j, it, 1}, \ldots, \delta_{j, it, a})' \) and \( S_{j, it} = (S_{j, it, 1}, \ldots, S_{j, it, b})' \) is a \( b_j \)-dimensional vector of exogenous predictors with static coefficients \( \eta_{j} = (\eta_{j, 1}, \ldots, \eta_{j, a})' \).

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 \( j = 1, \ldots, q \), set \( b_j = b = 1 \) for \( j = 1, \ldots, m \), and \( b_j = 0 \) for \( j = m + 1, \ldots, q \), 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 \( \beta_{it} \) be a \( p_d \times p_s \)-dimensional vector constructed by stacking the \( a_j \) coefficients \( \delta_{j, it} \) for \( j = 1, \ldots, q \). The structural equation of the HMDM relates the location-time varying parameter \( \beta_{it} \) to an aggregate (pooled) state parameter \( \gamma_{it} \):

\[
\beta_{it} = \gamma_{it} + v_{it}
\]

where the errors \( v_{it} \) are assumed to be i.i.d. \( N_{p_d}(0, V_s) \). The state (or system) equation of the HMDM is:

\[
\gamma_{it} = G \gamma_{t-1} + w_{it}
\]

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

The HMDM in (6)–(8) simplifies to the Hierarchical DGLM (HDGLM) when \( m = 1 \), where we replace \( MP_{m}(y_{it} | \lambda_{it}) \) by the univariate Poisson pmf.

4.1 Bayesian inference

Let \( Y, D \), and \( S \) denote all the responses \( y_{st} \), the dynamic predictors and the static predictors for \( t = 1, \ldots, T \) and \( i = 1, \ldots, N \). Let \( \eta \) and \( \beta \) denote all the coefficients \( \eta_j \) and \( \beta_{it} \) for \( j = 1, \ldots, q \), \( t = 1, \ldots, T \) and \( i = 1, \ldots, N \), and let \( \gamma \) denotes all the coefficients \( \gamma_t \) for \( t = 1, \ldots, T \). The likelihood function under the model described by (6)–(8) is

\[
L(\eta, \beta, \gamma; Y, D, S) = \prod_{i=1}^{N} \prod_{t=1}^{T} MP_{m}(y_{it} | \beta_{it}, \gamma_t)
\]

\[
\times \text{pnormal}(\eta) \times \text{pnormal}(\beta_{it} | \gamma_t) \times \text{pnormal}(\gamma_t | \gamma_{t-1})
\]

where we have suppressed the terms \( D \) and \( S \) on the right side for brevity. We assume multivariate Normal priors for the initial state vector and the static coefficients, i.e., \( \gamma_0 \sim N_{p_s}(m_0, C_0) \) and \( \eta \sim N_{p_s}(\mu_\eta, \Sigma_\eta) \). We assume inverse Wishart priors for the variance terms \( V_s \) and \( W \), i.e., \( V_s \sim IW(n_s, S_s) \), and \( W \sim IW(n_w, S_w) \), and we assume a product prior specification of the hyperparameters are selected to correspond to a vague prior specification. Prior elicitation is 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:

\[
\pi(\beta_{1t}, \gamma_{1t}, \eta, V_s, W | Y, D, S) \propto \prod_{i=1}^{N} \prod_{t=1}^{T} MP_{m}(y_{it} | \lambda_{it})
\]

\[
\times |V_s|^{-n_s/2} \exp \left\{ -\frac{1}{2} (\beta_{it} - \gamma_t)' V_s^{-1} (\beta_{it} - \gamma_t) \right\}
\]

\[
\times |\Sigma_\eta|^{-n_\eta/2} \exp \left\{ -\frac{1}{2} (\eta - \mu_\eta)' \Sigma_\eta^{-1} (\eta - \mu_\eta) \right\}
\]

\[
\times \left\{ \prod_{t=1}^{T} |W|^{-1/2} \exp \left\{ -\frac{1}{2} \gamma_t - G \gamma_{t-1}' W^{-1} (\gamma_t - G \gamma_{t-1}) \right\} \right\}
\]

\[
\times |C_0|^{-n_0/2} \exp \left\{ -\frac{1}{2} (\gamma_0 - m_0)' C_0^{-1} (\gamma_0 - m_0) \right\}
\]

\[
\times \prod_{i=1}^{N} |V_s|^{-n_s/2} \exp \left\{ -\frac{1}{2} \text{tr}(V_s^{-1} S_s) \right\}
\]

\[
\times |W|^{-n_w/2} \exp \left\{ -\frac{1}{2} \text{tr}(W^{-1} S_w) \right\}
\]

(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 \( \Psi = (V_1, \ldots, V_N, W) \). Conditional on \( \beta_{it} \) and \( \Psi \), the structural and state equations of the HMDM have the form of a Gaussian Dynamic HDLM with observations \( \beta_{it} \) and state \( \gamma_t \). Note that \( \gamma_t \) is independent of \( Y_t \), given \( \beta_{it} \).

Let \( F_t = (I_{p_\theta}, \ldots, I_{p_\theta})' \) denote an \( N_{p_\theta} \times p_d \) mapping matrix, and let \( \beta_{it} \) denote the vector obtained by stacking \( \beta_{it} \).

Hierarchical dynamic models for multivariate times series of counts 563
for $i = 1, \ldots, N$. The complete conditional distribution of $\gamma_t$ may be written as
\begin{equation}
P(\gamma|\beta, \Psi) = P(\gamma|\beta, \Psi) \prod_{t=1}^{T-1} P(\gamma_t|\gamma_{t+1}, \ldots, \gamma_t, \beta, \Psi)
\end{equation}

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

Filter step: For $t = 1, \ldots, T$, we compute the mean $m_t$ and variance matrix $C_t$ of the posterior normal distributions $P(\gamma_t|\gamma_{t+1}, \beta, \Psi)$, by applying the standard sequential updating results for Gaussian DLMs:
\begin{align}
m_t &= Gm_{t-1} + RF_tQ_t^{-1}(\beta_t - F_tGm_{t-1}) \\
C_t &= R_t - R_tF_tQ_t^{-1}F_tR_t \\
Q_t &= V + F_tR_tF_t' \\
R_t &= W + GC_tG'
\end{align}

where $V$ is an $N_{p4} \times N_{p4}$ block diagonal matrix, each block represents the variance covariance matrix $V_i$ for $i = 1, \ldots, N$.

Smooth step: At time $t = T$, we sample the final state vector $\gamma_T$ from the marginal distribution, $P(\gamma_T|\beta, \Psi)$, which is $N(\gamma_T|m_T, C_T)$. For time periods $t = T-1, \ldots, 0$, we sample from $P(\gamma_t|\gamma_{t+1}, \beta, \Psi)$ which is $N(h_t, H_t)$ at each time, conditional on the latest value of $\gamma_{t+1}$, where
\begin{align}
H_t &= (C_t^{-1} + G'W^{-1}G)^{-1} \\
h_t &= H_tC_t^{-1}m_t + G'W^{-1}\gamma_{t+1}
\end{align}

The results of these steps is a draw $(\gamma_T, \ldots, \gamma_1)$ from its complete conditional distribution.

The complete conditional distributions of the variance terms are inverse Wishart distributions. Let $\beta_t = \gamma_t$. For each $i$, we sample $V_i$ from $IW(n_i, S_i)$, where
\begin{align}
n_i &= T + n_v \\
S_i &= S_v + \sum_{t=1}^{T} ((\beta_t - \hat{\beta}_t)(\beta_t - \hat{\beta}_t)')
\end{align}

We sample $W$ from $IW(n_w, S_w)$ where
\begin{align}
n_w &= T + n_w \\
S_w &= S_w + \sum_{t=1}^{T} ((\gamma_t - G\gamma_{t-1})(\gamma_t - G\gamma_{t-1})')
\end{align}

The Metropolis-Hastings algorithm is used for sampling $\beta_t$ and $\eta$.

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($\lambda_{it}$) 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 $\lambda_{it}$ are obtained by using the output of the Gibbs sampler to approximate the predictive density $p(\lambda|Y, \gamma, \beta, \Psi, \eta)$ using Monte Carlo integration. We also obtain the conditional predictive ordinate (CPO) for each $i$ and $t$, and aggregate the values to obtain the pseudo-Bayes factor (PBSF) 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
\begin{align}
MAD &= \frac{1}{N} \frac{1}{T} \sum_{i=1}^{N} \sum_{t=1}^{T} \frac{\text{sign}(y_{it} - \hat{\lambda}_{it})\sqrt{d_{it}}}{d_{it}} \\
PMAD &= \frac{1}{N} \frac{1}{L} \sum_{i=1}^{N} \sum_{t=T+1}^{T+L} \frac{\text{sign}(y_{it} - \hat{\lambda}_{it})\sqrt{d_{it}}}{d_{it}}
\end{align}

where $\hat{\lambda}_{it}$ denotes the posterior mean of $\lambda_{it}$, and $d_{it}$ denotes an individual deviance contribution and is defined as
\begin{equation}
d_{it} = 2(\log(P_{\text{Poisson}}(y_{it}|\lambda_{it})) - \log(P_{\text{Poisson}}(y_{it}|\hat{\lambda}_{it}))
\end{equation}

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 $Y_{it}$ ($i = 1, \ldots, 16$ and $t = 1, \ldots, 44$) is simulated according to the following model:
\begin{align}
Y_{it}|\lambda_{it} &\sim \text{Poisson}(\lambda_{it}) \\
\log \lambda_{it} &= \delta_{it,1} + \delta_{it,2} + \eta S_{it}
\end{align}

In (16), $S_{it}$ represents a static predictor, which is simulated from Normal(0,1) and $\eta = 0.5$. Let $\beta_{it} = (\delta_{it,1}, \delta_{it,2})'$ represent a location-time varying coefficient vector. The structural (hierarchical) and state equations are given by (7) and (8), with $G = \text{diag}(1, -1)$. The errors $\eta_{it}$ in (7) are simulated from $N_2(0, V_i)$ with $V_i = \text{diag}(0.05, 0.01)$ for $i = 1, \ldots, 16$, and $w_t$ from (8) is simulated from $N_2(0, W')$ with $W = \text{diag}(10^{-3}, 10^{-4})$. 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 $W = \text{diag}(W_1, W_2)$ is an Inverse-Wishart distribution with $2p_d + 1$ degrees of
freedom and scale matrix $S_w = \text{diag}(10^3, 10^4)$. Similarly, the prior distribution of $V_i = \text{diag}(V_{1i}, V_{2i})$ is an Inverse-Wishart distribution with $2p_d + 1$ degrees of freedom and scale matrix $S_v = \text{diag}(25, 100)$. To investigate prior sensitivity, the prior for $\eta_i$ 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 $S_w = \text{diag}(10^4, 10^4)$ and $S_v = \text{diag}(10^4, 10^4)$, along with a $N(0, 10^3)$ prior for $\eta$. Further, we consider different starting values for the $\eta$ 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 $Y_{it} = (Y_{1it}, Y_{2it})'$ ($i = 1, \ldots, 16$ and $t = 1, \ldots, 44$) is simulated according to the following model:

$$Y_{it}|\lambda_{i1}, \lambda_{i2}, \lambda_{i3} \sim MP_2(\lambda_{i1}, \lambda_{i2}, \lambda_{i3})$$

$$\log \lambda_{i1} = \delta_{i1} + \eta_1 S_{i1}$$

$$\log \lambda_{i2} = \delta_{i2} + \eta_2 S_{i2}$$

$$\log \lambda_{i3} = \delta_{i3}$$

(17)

In (17), $S_i$ represents a static predictor, which is simulated from $N(0, 1)$ and $\eta_1 = 0.5$ and $\eta_2 = -0.5$. Let $\beta_{it} = (\delta_{i1,1}, \delta_{i2,1}, \delta_{i3,1}, \delta_{i1,2}, \delta_{i2,2})'$ represent location-time varying coefficients. The structural and state equations are again given by (7) and (8), with $G = \text{diag}(1, 1, -1, -1)$. The errors $\epsilon_{it}$ are simulated from $N_5(0, V_i)$ with $V_i = \text{diag}(V_{1i}, V_{2i})$, where $V_{1i}$ is a symmetric matrix with diagonal elements 0.05 and off-diagonal elements 0.001, for $i = 1, \ldots, 25$. Likewise, $V_{2i}$ is a symmetric matrix with diagonal elements 0.01 and off-diagonal elements 0.001 for $i = 1, \ldots, 44$. The errors $\epsilon_{it}$ are simulated from $N_5(0, W)$ with $W = \text{diag}(10^{-3}, 10^{-3}, 10^{-3}, 10^{-5}, 10^{-4})$. Here, $p_d = 5$ and $p_s = 2$.

The prior distribution of $W = \text{diag}(W_1, \ldots, W_5)$ is an Inverse-Wishart distribution with $2p_d + 1$ degrees of freedom and scale matrix $S_w = \text{diag}(10^4, 10^4, 10^4, 10^4)$. The prior distribution of $V_i = \text{diag}(V_{1i}, V_{2i})$ is an Inverse-Wishart distribution with $2p_d + 1$ degrees of freedom and scale matrix $S_v = \text{diag}(25, 25, 25, 100, 100)$. To investigate prior sensitivity, the prior for $\eta = (\eta_1, \eta_2)'$ is assumed to be $N_2(0, \text{diag}(10^3, 10^3))$, $N_2(0, \text{diag}(10^2, 10^2))$ or $N_2(0, \text{diag}(10, 10))$ for simulation scenarios 1, 2 and 3 respectively. The prior scales $S_w = \text{diag}(10^4, 10^4, 10^4, 10^4)$ and $S_v = \text{diag}(10^3, 10^3, 10^3, 10^3)$ are considered for scenario 5, along with a $N_2(0, \text{diag}(10^3, 10^3))$ prior for $\eta$. We again consider different starting values for the $\eta$ 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 $N_2(0, \text{diag}(10^3, 10^3))$ 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

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Parameter</th>
<th>True value</th>
<th>Posterior Mean</th>
<th>Posterior Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate</td>
<td>HDGLM</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$\eta$</td>
<td>0.5</td>
<td>0.5206</td>
<td>0.0217</td>
</tr>
<tr>
<td>2</td>
<td>$\eta$</td>
<td>0.5</td>
<td>0.5208</td>
<td>0.0220</td>
</tr>
<tr>
<td>3</td>
<td>$\eta$</td>
<td>0.5</td>
<td>0.5299</td>
<td>0.0216</td>
</tr>
<tr>
<td>4</td>
<td>$\eta$</td>
<td>0.5</td>
<td>0.5204</td>
<td>0.0215</td>
</tr>
<tr>
<td>5</td>
<td>$\eta$</td>
<td>0.5</td>
<td>0.5215</td>
<td>0.0216</td>
</tr>
<tr>
<td>Bivariate</td>
<td>HDMD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>$\eta_1$</td>
<td>0.5</td>
<td>0.4371</td>
<td>0.0209</td>
</tr>
<tr>
<td>2</td>
<td>$\eta_2$</td>
<td>-0.5</td>
<td>-0.3610</td>
<td>0.0301</td>
</tr>
<tr>
<td>3</td>
<td>$\eta_1$</td>
<td>0.5</td>
<td>0.4372</td>
<td>0.0210</td>
</tr>
<tr>
<td>4</td>
<td>$\eta_2$</td>
<td>-0.5</td>
<td>-0.3610</td>
<td>0.0303</td>
</tr>
<tr>
<td>5</td>
<td>$\eta_1$</td>
<td>0.5</td>
<td>0.4370</td>
<td>0.0211</td>
</tr>
<tr>
<td>6</td>
<td>$\eta_2$</td>
<td>-0.5</td>
<td>-0.3608</td>
<td>0.0303</td>
</tr>
</tbody>
</table>

The estimated coefficients for $\eta$ 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 $\eta$ estimates are also reasonably sensitive to changes in priors for the variance parameters $W$ and $V_i$.

## 7. ANALYSIS OF GASTROPOD COUNTS

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

$$Y_{it} \sim \text{Poisson}(\lambda_{it})$$

$$\log \lambda_{it} = \alpha_{it} + \text{Season}_{it} + \eta_1 \text{LogRep}_{it} + \eta_2 \text{Elevation}_{it} + \eta_3 \text{Slope}_{it} + \eta_4 \text{CO}_{it} + \eta_5 \text{PA}_{it} + \eta_6 \text{PA}_{it}$$

$$+ \eta_7 \text{I(Aspect=5)}_{it} + \eta_8 \text{I(Aspect=7)}_{it} + \eta_9 \text{I(Aspect=8)}_{it}$$

$$+ \eta_{10} \text{I(Soil Type=2)}_{it} + \eta_{11} \text{I(Soil Type=3)}_{it}$$

$$+ \eta_{12} \text{I(CC=2)}_{it} + \eta_{13} \text{I(CC=3)}_{it}$$

$$= D' \beta_{it} + S' \eta = I \beta_{it} + S' \eta$$

(18)

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

*Hierarchical dynamic models for multivariate times series of counts 565*
We next fit a bivariate HMDM to gastropod counts \( Y_{it} = ( \text{Carcar}_{it}, \text{Gaenig}_{it} )' \), where \( \text{Carcar}_{it} \) and \( \text{Gaenig}_{it} \) represent the abundance of \( \text{C. caracolla} \) and \( \text{G. nigrolineata} \) respectively at location \( i \) and time \( t \). Let \( \eta_j \) correspond to a vector of static coefficients for \( j = 1, 2 \) and \( S \) 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

\[
Y_{it} | \lambda_{1, it}, \lambda_{2, it}, \lambda_{3, it} \sim MP_3( \lambda_{1, it}, \lambda_{2, it}, \lambda_{3, it} ) \\
\log \lambda_{1, it} = \alpha_{1, it} + \text{Season}_{1, it} + S' \eta_1 \\
\log \lambda_{2, it} = \alpha_{2, it} + \text{Season}_{2, it} + S' \eta_2 \\
\log \lambda_{3, it} = \alpha_{3, it} 
\]

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

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 \( W \) 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 \( V_i \) 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 \( W \) and \( V_i \) reveal no significant values for \( i = 1, \cdots, 40 \). Recall that the first three diagonal elements of \( W \) and \( V_i \) 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 \( \text{C. caracolla} \) and \( \text{G. nigrolineata} \) 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 \( \text{C. caracolla} \). The fit for \( \text{G. nigrolineata} \) 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 \( \text{G. nigrolineata} \) 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 \( \text{C. caracolla} \). 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, appearance of sierra palm, and plant appearance 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 Zarzal 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 \( \text{G. nigrolineata} \), 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 \( \text{G. nigrolineata} \) on Cristal soils was lower than that on Zarzal soils, whereas abundance of \( \text{G. nigrolineata} \) on Preito 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 \( \text{G. nigrolineata} \) 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>
<thead>
<tr>
<th>Model</th>
<th>Species</th>
<th>MAD</th>
<th>PMAD</th>
<th>Log PsBF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Univariate</td>
<td>C. caracolla</td>
<td>1.2298</td>
<td>1.9130</td>
<td>3.9575</td>
</tr>
<tr>
<td></td>
<td>G. nigrolineata</td>
<td>1.3363</td>
<td>2.4915</td>
<td>3.9990</td>
</tr>
<tr>
<td>Bivariate</td>
<td>Overall</td>
<td>1.2947</td>
<td>2.2087</td>
<td>4.6727</td>
</tr>
<tr>
<td></td>
<td>C. caracolla</td>
<td>1.3424</td>
<td>1.9082</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>G. nigrolineata</td>
<td>1.2470</td>
<td>2.5093</td>
<td>–</td>
</tr>
</tbody>
</table>

566 N. Ravishanker, V. Serhiyenko, and M. R. Willig
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.

**ACKNOWLEDGEMENTS**

We are grateful to a reviewer for very useful suggestions that helped us improve our manuscript. This research was facilitated by grants (DEB-0620910, DEB-0218039, and DEB-0080538) from the National Science Foundation to the Institute of Tropical Ecosystem Studies, University of Puerto Rico, and the International Institute of Tropical Forestry as part of the Long-Term Ecological Research Program in the Luquillo Experimental Forest. Additional support was provided by the USDA Forest Service, the University of Puerto Rico, and the Center for Environmental Sciences and Engineering at the University of Connecticut. The staff of El Verde Field Station provided valu-
able logistical support in Puerto Rico. We also thank the many students and colleagues who assisted with gastropod surveys. The first author is grateful to the University of Connecticut Research Foundation for a Faculty Large Grant that facilitated the early part of this research.
APPENDIX

The complete conditional density of $\beta_{it}$ for $i = 1, \ldots, N$ and $t = 1, \ldots, T$ is

$$
\pi(\beta_{it} | \gamma_i, \eta, V_i, W, Y) \propto MP_m(y_{it} | \lambda_{it}) \exp\left\{ \frac{1}{2} (\beta_{it} - \gamma_i) W^{-1}(\beta_{it} - \gamma_i) \right\}
$$

(20) \times \prod_{t=1}^{T} \exp\left\{ -\frac{1}{2} (\gamma_t - G \gamma_{t-1}) W^{-1}(\gamma_t - G \gamma_{t-1}) \right\}

The complete conditional density of $\gamma_t$ for $t = 1, \ldots, T$ is

$$
\pi(\gamma_t | \beta_{it}, \eta, V_i, W, Y) \propto \prod_{i=1}^{N} \exp\left\{ -\frac{1}{2} (\beta_{it} - \gamma_i) V_i^{-1}(\beta_{it} - \gamma_i) \right\}
$$

(21) \times \exp\left\{ -\frac{1}{2} (\gamma_t - G \gamma_{t-1}) W^{-1}(\gamma_t - G \gamma_{t-1}) \right\}

The complete conditional density of $\eta$ is

$$
\pi(\eta | \beta_{it}, \gamma_i, V_i, W, Y) \propto \prod_{i=1}^{N} \prod_{t=1}^{T} MP_m(y_{it} | \lambda_{it}) \exp\left\{ -\frac{1}{2} (\mu_{it} - \eta) \right\}
$$

(22) \times \exp\left\{ -\frac{1}{2} (\eta - \mu_{it}) \right\}

The complete conditional density of $V_i$ for $i = 1, \ldots, N$ is

$$
\pi(V_i | \beta_{it}, \gamma_t, \eta, W, Y) \propto \prod_{t=1}^{T} |V_i|^{-1/2} \exp\left\{ -\frac{1}{2} (\beta_{it} - \gamma_i) V_i^{-1}(\beta_{it} - \gamma_i) \right\}
$$

(23) \times |V_i|^{-n_{it}/2} \exp\left\{ -\frac{1}{2} tr(V_i^{-1}S_v) \right\}

The complete conditional density of $W$ is

$$
\pi(W | \beta_{it}, \gamma_t, \eta, V_i, Y) \propto \prod_{t=1}^{T} |W|^{-1/2} \exp\left\{ -\frac{1}{2} (\gamma_t - G \gamma_{t-1}) W^{-1}(\gamma_t - G \gamma_{t-1}) \right\}
$$

(24) \times |W|^{-n_{it}/2} \exp\left\{ -\frac{1}{2} tr(W^{-1}S_w) \right\}

Received 18 November 2013

REFERENCES


Hierarchical dynamic models for multivariate times series of counts 569