

## LINKING CHRONIC WASTING DISEASE TO MULE DEER MOVEMENT SCALES: A HIERARCHICAL BAYESIAN APPROACH

MATTHEW L. FARNSWORTH,<sup>1,4</sup> JENNIFER A. HOETING,<sup>2</sup> N. THOMPSON HOBBS,<sup>1</sup> AND MICHAEL W. MILLER<sup>3</sup>

<sup>1</sup>Natural Resources Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80523-1499 USA

<sup>2</sup>Department of Statistics, Colorado State University, Fort Collins, Colorado 80523-1877 USA

<sup>3</sup>Colorado Division of Wildlife, Wildlife Research Center, 317 West Prospect Road, Fort Collins, Colorado 80526-2096 USA

**Abstract.** Observed spatial patterns in natural systems may result from processes acting across multiple spatial and temporal scales. Although spatially explicit data on processes that generate ecological patterns, such as the distribution of disease over a landscape, are frequently unavailable, information about the scales over which processes operate can be used to understand the link between pattern and process. Our goal was to identify scales of mule deer (*Odocoileus hemionus*) movement and mixing that exerted the greatest influence on the spatial pattern of chronic wasting disease (CWD) in northcentral Colorado, USA. We hypothesized that three scales of mixing (individual, winter subpopulation, or summer subpopulation) might control spatial variation in disease prevalence. We developed a fully Bayesian hierarchical model to compare the strength of evidence for each mixing scale. We found strong evidence that the finest mixing scale corresponded best to the spatial distribution of CWD infection. There was also evidence that land ownership and habitat use play a role in exacerbating the disease, along with the known effects of sex and age. Our analysis demonstrates how information on the scales of spatial processes that generate observed patterns can be used to gain insight when process data are sparse or unavailable.

**Key words:** Bayesian analysis; chronic wasting disease; disease ecology; hierarchical models; intrinsic Gaussian conditional autoregressive (ICAR) model; mule deer; *Odocoileus hemionus*; prion disease; spatial models; spatial scale.

### INTRODUCTION

Linking spatial patterns to the processes that generate them offers a fundamental challenge in contemporary ecology (Levin 1992, Sarnelle 1994, Pascual and Levin 1999). This problem is particularly difficult to solve when processes operate over large spatial and temporal scales. When this is the case, it is possible that emergent patterns reflect the outcome of processes operating at more than one nested scale. Hierarchical models provide a natural, unified framework for comparing spatial and temporal processes that operate across a range of scales. Here we use hierarchical modeling to investigate the spatial distribution of an emerging infectious disease of wildlife in North America.

Chronic wasting disease (CWD) (Williams and Young 1980) of North American cervids is the only prion disease known to occur in free-ranging populations (Williams and Miller 2002). The largest known outbreak occurs in a contiguous ~80 000-km<sup>2</sup> area of northeastern Colorado, southeastern Wyoming, and western Nebraska, USA (Williams and Young 1992, Miller et

al. 2000, Williams and Miller 2002). Chronic wasting disease may have been present in free-ranging deer within this area since the 1960s or earlier (Miller et al. 2000). Although the infectious agent causing CWD is contagious in its natural setting (Williams and Young 1992, Miller and Williams 2003), relatively little is known about mechanisms of transmission (Williams and Miller 2002). Possible routes of transmission include animal–animal and animal–environment–animal pathways (Miller and Williams 2003, Miller et al. 2004). Recent discoveries of CWD foci distant to this 80 000-km<sup>2</sup> endemic area motivated our efforts to understand spatial and temporal dynamics. This understanding is needed to inform management strategies for controlling CWD in affected populations and prevent or slow its spread among unaffected populations.

The distribution of CWD in northeastern Colorado appears heterogeneous at both small ( $\leq 50$  km<sup>2</sup>; Wolfe et al. 2002) and large ( $> 38 000$  km<sup>2</sup>; Miller et al. 2000) scales (Fig. 1). Small-scale heterogeneity may result from highly localized contact processes, such as interactions among individuals within matrilineal groups, which are tightly clustered and patchily distributed across their winter home range (Conner and Miller 2004, Miller and Conner 2005). Recent investigations suggest that once the infectious agent is shed into the environment it may persist for several years outside the host (Miller et al. 2004). Consequently, large quantities of agent deposited

Manuscript received 8 June 2005; revised 23 September 2005; accepted 26 September 2005. Corresponding Editor: R. S. Ostfeld.

<sup>4</sup> Present address: USDA-APHIS-WS, National Wildlife Research Center, 4101 LaPorte Avenue, Fort Collins, Colorado 80521 USA. E-mail: matt.farnsworth@aphis.usda.gov

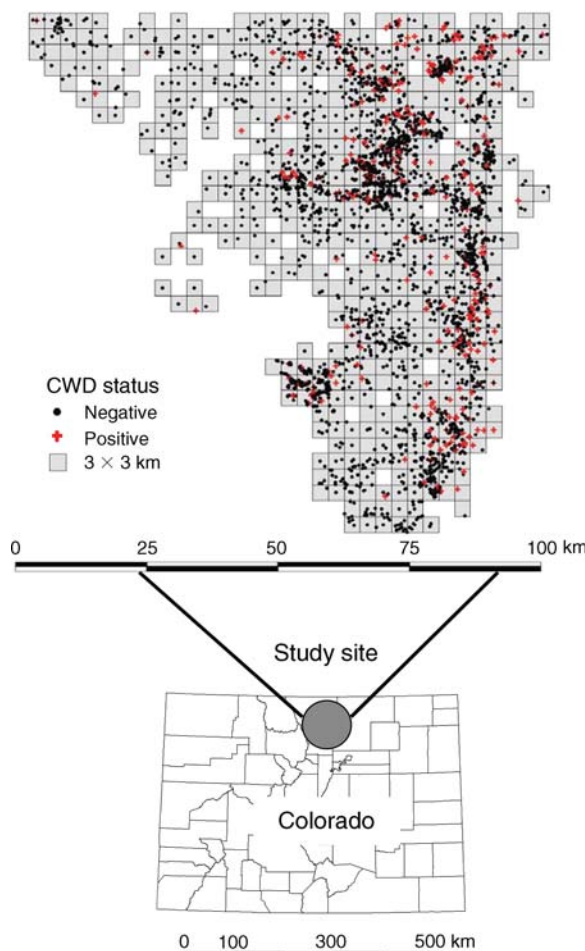


FIG. 1. Study site in north-central Colorado, USA, overlaid with the spatial distribution of chronic wasting disease (CWD) data and the 540 9-km<sup>2</sup> grid cells used in modeling individual-level infection probability at the finest analysis scale.

in a relatively small geographic area could result in “hot spots” of infection scattered across natural landscapes.

The process of mule deer movement in north-central Colorado can be divided into three categories, reflecting the geographic scales of seasonally dependent movement patterns (Conner and Miller 2004). At the largest scale, the summer subpopulation home range, mule deer home range sizes average  $\sim 310$  km<sup>2</sup> and exhibit greater overlap ( $\sim 22\%$ ) than at any other time of year (Conner and Miller 2004). At the winter subpopulation home range scale, deer live in groups with a mean home range size of  $\sim 80$  km<sup>2</sup> and exhibit little overlap ( $< 1\%$ ) among wintering groups (Conner and Miller 2004). Finally, individual mule deer have a characteristic home range size that averages  $\sim 9$  km<sup>2</sup> during the winter when deer are more sedentary (Conner and Miller 2004).

High fidelity to seasonal-use areas and temporally consistent movement patterns of subpopulations (Conner and Miller 2004), and presumably the resulting home range scales, suggest the importance of local,

small-scale contact processes in structuring CWD spatial heterogeneity. At the same time, large-scale movement patterns, such as those that occur when subpopulations expand their home range between winter and summer locations, could result in a greater number of contacts among deer that do not interact during the winter months. If large-scale movements were primarily responsible for the spatial structure of CWD, then the distribution of infected deer should exhibit greater homogeneity at large geographic scales than would be expected if disease transmission occurs predominately at local winter subpopulation or individual movement scales. Heterogeneity in CWD prevalence also may result from different times since disease introduction into the various subpopulations (Miller et al. 2000); unfortunately, time since introduction will remain unknown for most or all infected subpopulations.

Studies examining the manner in which ecological inference changes with analysis scale have been used to understand how avian species richness changes with the scale of observation (Rahbek and Graves 2001), how the risk of disease varies across analysis scales (Jarup et al. 2002, Smith et al. 2003), and how quantitative measures of invasion and extinction of native fish species differ across scales (Olden and Poff 2004). Similarly, to understand the potential relative contribution of different scales of deer movement to the structure of observed spatial heterogeneity of CWD, we specified a set of candidate models reflecting the different scales, combinations of predictor variables, and spatial dependencies affecting the probability of CWD infection in individuals. We then used contemporary model selection techniques to identify those scales and predictors corresponding best to the observed spatial structure of CWD infection across the landscape. In addition to gaining insights into the spatial epidemiology of CWD, we demonstrate how hierarchical modeling can be used to understand the relative contribution of hypothesized generating processes to observed patterns of disease prevalence.

Although some studies of natural populations have examined correlations between the spatial distribution of disease and environmental variables (Van Buskirk and Ostfeld 1998, Giraudoux et al. 2003) and others have proposed biological explanations for the spatial patterns they observe (Jolles et al. 2002, Aylor 2003), few if any have identified the scales over which population-level processes act to shape the spatial distribution of disease in wildlife populations. Hierarchical models provide a natural way to address this problem.

## METHODS

### *Study area*

Our study area included 6500 km<sup>2</sup> in north-central Colorado (Fig. 1) where CWD is endemic in free-ranging cervids (Miller et al. 2000). Elevation ranged from 1500 to 3500 m, rising from east to west. The northeastern quarter of the study area, north of Fort

Collins, consisted of rolling foothills and high prairie where livestock grazing was the main land use. Vegetation was primarily sagebrush-steppe habitat with big sagebrush (*Artemisia tridentata*), antelope bitterbrush (*Purshia tridentata*), mountain mahogany (*Cercocarpus montanus*), and mixed grasses. The southeastern quarter of the study area, from Fort Collins south, consisted of urban centers separated by rural areas with numerous small ranches, agricultural fields, natural areas, and more scattered suburban areas. Vegetation communities in the western half of the study area followed the east-west elevation gradient. Lower elevations were mostly dense mountain mahogany interspersed with grassland openings and small patches of ponderosa pine (*Pinus ponderosa*) that transitioned to mountain shrub habitat with a primarily mixed-conifer overstory at higher elevations. The highest elevations were mainly alpine tundra habitat.

### Data

The Colorado Division of Wildlife (CDOW) provided georeferenced data for 3855 mule deer tested for CWD infection between 1997 and 2003 within the ~6500 km<sup>2</sup> study area. All samples were geo-referenced using either a global positioning system unit or by identifying sample source locations on standardized maps. Sampling methods included deer that were killed by hunters, culled by wildlife managers, or captured and tonsil biopsied; survey and diagnostic methods have been described in detail elsewhere (Miller et al. 2000, Miller and Williams 2002, Wolfe et al. 2002, Hibler et al. 2003). Sampled deer were classified as CWD positive or negative based on immunohistochemistry of retropharyngeal lymph node or tonsil tissue (Miller and Williams 2002). The CDOW established a grid of 9-km<sup>2</sup> cells for identifying where individual deer were sampled (Fig. 1; CDOW, unpublished data). For the current analysis, we partitioned the study area into three grids. Each grid had a resolution representing one of the scales of mule deer movement: 9 km<sup>2</sup>, 81 km<sup>2</sup>, and 324 km<sup>2</sup> overlaid on the map of deer samples. The 9-km<sup>2</sup> grid, being 3 km on a side, was the sampling grid established by the CDOW, with the two coarser scale grids resulting from combining multiple 9-km<sup>2</sup> grid cells. Beginning with the northwestern-most 9-km<sup>2</sup> cell, we used a moving window to combine neighboring 9-km<sup>2</sup> cells until we had achieved a new grid cell with a resolution equal to the next coarsest scale of deer movement. This process was iterated across the entire map at both of the coarser movement scales. For example, combining a 3 × 3 block of 9-km<sup>2</sup> cells resulted in a new cell having an area equal to 81 km<sup>2</sup> (i.e., 9 km on a side). Constructing the grids in this manner resulted in all three grids being aligned with the 9-km<sup>2</sup> sampling grid established by the CDOW. Finally, because CWD prevalence remained relatively constant within the study area between 1996 and 2003 (Miller et al. 2000, Miller and Conner 2005), we aggregated the data across all years.

Data on locations and sex and age class (juvenile or adult), which are known to influence infection probability (Miller and Conner 2005), were recorded for each deer. Adult deer were classified as being two years of age or older. In addition to these individual-level demographic (Demo) covariate effects, three environmental (Env) covariates were calculated using a geographic information system containing data grids representing land ownership and vegetation patterns across the study area at a 90-m resolution. The covariate values were simple measures of proportions and landscape configuration, as described below, calculated for each of the three movement scales used in the model. Each environmental covariate was assumed to exert the same influence on all individuals sampled from the same grid cell. Thus, these covariates were scaled to the map resolution considered in each model.

There is evidence of anthropogenic influence on CWD prevalence in mule deer (Wolfe et al. 2002, Farnsworth et al. 2005); consequently, we considered an environmental covariate, %PRIV, that represented the percentage of private land in a grid cell. Because the configuration of private land may be important, we used a second environmental covariate, DISP, which measured the degree of isolation of private land within a grid cell as an index of the connectivity of private land (McGarigal and Marks 1995) across the entire study area. A final environmental covariate, %HAB, measuring the percentage of low-elevation grassland habitat in each grid cell, was used to represent the amount of wintering habitat to capture the potential influence of deer winter range concentration on CWD prevalence. It is unlikely that any of the values of the covariates changed in a manner that would substantially alter the results during the time period covered by our study.

### Hierarchical model of CWD infection

Two aspects of the data made it difficult to relate movement scales to CWD infection probability. First, we had home range size estimates derived from some deer living within the study area (Conner and Miller 2004). However, we did not have spatially explicit home range data for the deer used in our study. Although data on movement patterns were unavailable across much of the area, we wanted to shed light on potential transmission scales. Second, economic and logistic constraints, in conjunction with the relatively long clinical course and low prevalence of the disease (Miller et al. 2000), made it impossible to follow individual deer as they moved across the landscape prior to being sampled. We addressed these challenges by partitioning the study region into aerial units (i.e., grid cells) reflecting the scales of seasonally dependent mule deer movement patterns. This specification allowed us to compare different movement scales in the face of limited information on the spatial structure of this process across our study area. The hierarchical structure accommodated uncertainties in the point-based CWD

data by treating all individuals sampled from within the same grid cell as having an identical exposure risk to the infectious agent after adjusting for individual-level sex and age effects.

We consider a generalized linear model for disease presence/absence. For each individual deer, we model the probability of being CWD-positive as a function of the covariates and two random effect terms, which account for any unobserved covariates as well as the spatial pattern in the probability of disease presence.

This generalized linear model can be described in three stages: the data model, or likelihood, linking the data to the model parameters; the process model relating the covariates and random effects to the parameters; and the prior distributions for all model parameters (Wikle 2003). Our interest focused on the posterior distribution, the distribution of the process and parameters after being informed by the data. For many ecological problems, the high dimensionality of the model can prohibit the use of standard methods. However, Markov Chain Monte Carlo (MCMC; Geman and Geman 1984, Gelfand and Smith 1990, Gilks et al. 1998) techniques allowed us to estimate the posterior distributions of interest.

*Data model*

The data model relates the known infection status for each deer to the probability of infection. Let  $Y_{ij}$  be the known infection status for deer  $i = 1, \dots, n_j$  in cell  $j = 1, \dots, k$ . We assume that infection status is Bernoulli distributed with parameter  $\pi_{ij}$ :

$$Y_{ij}|\pi_{ij} \sim \text{Bernoulli}(\pi_{ij}) \tag{1}$$

where  $\pi_{ij}$  is the probability of infection for individual  $i$  in cell  $j$ . All observations are assumed to be conditionally independent given this parameter.

*Process model*

The process component of the model relates the probability of infection for each deer,  $\pi_{ij}$ , to the individual and environmental covariates. We include two random effect terms to account for variability that is not accounted for by the covariates. To constrain the Bernoulli-distributed infection probability to the range 0–1, we use a standard logit transform. Thus we model the probability that an individual is infected as

$$\text{logit}(\pi_{ij}) = \mu + \mathbf{x}_{ij}^T \boldsymbol{\beta} + \gamma_j + \delta_j \tag{2}$$

where  $\mu$  is the background infection rate common to all deer,  $\boldsymbol{\beta}$  is an  $m \times 1$  vector of regression coefficients corresponding to the  $\mathbf{x}_{ij}^T$ , the transpose of the  $m \times 1$  vector of individual covariates for the  $i$ th deer and the scale-dependent environmental covariates associated with the  $j$ th grid cell;  $\gamma_j$  is the scale-dependent spatial random effect term for the  $j$ th grid cell; and  $\delta_j$  is the independent random effect term associated with the  $j$ th grid cell. The independent random effects vary with the

scale of analysis, but exhibit no spatial dependency. The random effect terms are described further below.

*Prior and posterior distributions*

Because our analysis is fully Bayesian, we specify prior distributions for all model parameters in the hierarchy. The spatial component, modeled by  $\gamma_j$ , is a key parameter of interest because it models the latent, or unobserved, contact process among mule deer resulting in the local structure of CWD. Recall that  $\gamma_j$  is the extra variation not accounted for by the covariates or unstructured heterogeneity in grid cell  $j$ . We specify the spatially structured variation in infection probability,  $\gamma_j$ , via an intrinsic Gaussian conditional autoregressive (ICAR) model (Besag et al. 1991). For a grid cell  $j$  in our problem, the ICAR model states that  $\gamma_j$  is related to the  $\gamma$  terms for the neighboring grid cells; and, given the  $\gamma$  terms for the neighboring grid cells, each grid cell is independent of all other grid cells outside the local neighborhood. Specifically, let the set of neighbors of cell  $j$  be denoted by  $j_+$ . Then, for each grid cell  $j$ , we assume the following conditional relationship:

$$\gamma_j|\gamma_{j_+} \sim \text{Normal}\left(\frac{1}{n_{j_+}} \sum_{i \in j_+} \gamma_i, \frac{\sigma_\gamma^2}{n_{j_+}}\right) \tag{3}$$

where  $n_{j_+}$  is the number of neighbors of grid cell  $j$  and  $\sigma_\gamma^2$  is the variance for all grid cells. Thus, the conditional mean of  $\gamma_j$  is simply the average value of its neighbors  $\gamma_{j_+}$ , with conditional variance  $\sigma_\gamma^2/n_{j_+}$  inversely proportional to the number of neighbors. In the conditional mean in Eq. 3, the neighboring grid cells are equally weighted so that all neighbors of cell  $j$  influence it equally. Spatial variation in our model is limited to cells sharing a border; however there are no a priori restrictions on specifying the neighborhood structure or cell weights. We use second-order neighborhoods consisting of the eight grid cells surrounding the grid cell in question. Thus, grid cells either sharing a border or immediately diagonal to the focal cell are considered in modeling spatial dependency. We chose a second-order neighborhood instead of something larger because we wish to maintain a sharp distinction between local dependency and global unstructured heterogeneity, which becomes increasingly blurred as the local neighborhood is extended.

The unstructured heterogeneity term,  $\delta_j$ , corresponds to a latent process operating independently in each grid cell at the chosen scale (e.g., home range scale). We let  $\delta_j \sim \text{i.i.d. } \mathcal{N}(0, \sigma_\delta^2 \mathbf{I})$  for  $j = 1, \dots, J$  where  $\mathbf{I}$  is an  $n_j \times n_j$  indicator matrix. This component models the overall, unstructured heterogeneity in the data by assuming no relationship among neighboring grid cells, but with a variance that is common to all grid cells.

The following distributions apply to the remaining model parameters. For the baseline disease risk,  $\mu$  is assumed to follow an improper (flat) prior on the whole real line. This prior distribution, along with a sum-to-

zero constraint placed on the spatial random effects, is necessary to assure identifiability because our model contains ICAR random effects (Besag and Kooperberg 1995). These restrictions result from defining the spatial random effect component conditionally rather than jointly. For the standardized covariates we assume  $\boldsymbol{\beta} \sim \mathcal{N}(0, \sigma_{\boldsymbol{\beta}}^2 \mathbf{I})$ . We specify uniform priors for the variance parameters,  $\sigma_{\boldsymbol{\beta}}^2$ , associated with the fixed effects components because recent investigations suggest that the introduction of prior information using the uniform distribution may be preferable to the more traditionally used gamma distribution (Gelman et al. 2004) when the goal is to provide a conjugate prior that contains little or no information to influence the posterior distribution of model parameters. Based on preliminary modeling using increasingly diffuse hyperparameter distributions we determined that a noninformative hyperparameter distribution for the variance parameters associated with the fixed effects could be specified as  $\sigma_{\boldsymbol{\beta}}^2 \sim \text{Uniform}(0, 100)$ , which was the prior distribution used in all subsequent models. In models containing both spatial and non-spatial random effects, prior distributions for the variances should be specified to allow for equal weighting of prior information, termed a “fair” prior (Carlin and Perez 2000, Banerjee et al. 2004). The idea behind fair-prior specification is that the prior distribution for the spatial random effect should be parameterized so that the spatial and nonspatial random effects are equally likely a priori. Bernardinelli et al. (1995) note that the prior marginal standard deviation of the  $\gamma_j$  parameters is approximately equal to the prior conditional standard deviation,  $\sigma_{\gamma_j}$ , divided by 0.7. Thus a scale that delivers

$$\text{SD}(\delta_j) = \frac{1}{\sigma_{\delta}} \approx \frac{1}{0.7 \bar{m} \sigma_{\gamma}} \approx \text{SD}(\gamma_j) \quad (4)$$

where  $\bar{m}$  is the mean number of neighbors, may offer a reasonably “fair” specification (Banerjee et al. 2004). The fair priors differ for each analysis scale due to variations in  $\bar{m}$ . The gamma-distributed fair priors for  $\sigma_{\gamma}^2$  and  $\sigma_{\delta}^2$ , respectively, are  $G(1.0, 1.0)$ ,  $G(10.37, 3.22)$  for the 9-km<sup>2</sup> scale,  $G(1.0, 1.0)$ ,  $G(9.77, 3.13)$  for the 81-km<sup>2</sup> scale, and  $G(1.0, 1.0)$ ,  $G(7.77, 2.79)$  for the 324-km<sup>2</sup> scale of analysis.

Because the conditional posterior distributions for the random effect parameters are free of the data  $Y$ , they are Bayesianly unidentified in the convolution model (Banerjee et al. 2004); however, this does not preclude learning about the behavior of the random effects parameters (i.e., prior to posterior movement). For example, we consider a quantity,  $\lambda = \sigma_{\gamma}^2 / (\sigma_{\gamma}^2 + \sigma_{\delta}^2)$ , that measures the proportional contribution from the spatial random effect component of variance to the overall variance due to the random effects. This parameter allows us to understand the relative contributions of spatially structured and unstructured variation in models containing both.

Finally, the joint posterior distribution of all model parameters given the field data is fit at the three scales of interest (grid cells of 9 km<sup>2</sup>, 81 km<sup>2</sup>, and 324 km<sup>2</sup>) for various combinations of predictors selected a priori. Thus, our models contain various combinations of demographic (Demo), environmental (Env), spatially structured (Space), and unstructured (Het) variation. The spatial structure induced by the ICAR model is dependent on the underlying grid (Best et al. 2000), and the three grid scales are not functionally related; thus it is not possible to directly scale up or down between the different models (Best et al. 1999, Hjort 1999). However, for the current analysis this is not a concern since our interest lies in comparing the fit of the models across the three scales and not on determining how the spatial structure itself changes with scale.

All models were fit using WinBUGS software (Spiegelhalter et al. 2002b). The MCMC procedure for these models was run for 50 000 iterations after a burn-in period of 500 000 iterations to ensure convergence of all model parameters.

#### Model comparisons

We use a deviance information criteria (DIC), a generalization of the Akaike Information Criteria (AIC), to compare the set of candidate models (Spiegelhalter et al. 2002a). These criteria are based on the deviance,  $D(\theta) = -2\ln L$ , where  $L$  is the likelihood and  $\theta$  is the vector of model parameters, and a penalty for model complexity. For AIC, the penalty is two times the number of parameters in the model. The complexity of a hierarchical model is measured by the effective number of parameters,  $p_D$ , which can be smaller than the total number of parameters. This complexity is defined as  $p_D = D(\bar{\theta}) - D(\hat{\theta})$ , where  $D(\bar{\theta})$  is the expected deviance over the posterior distribution of parameter vector  $\theta$  taken across all MCMC samples, and  $D(\hat{\theta})$  is the deviance evaluated at the posterior mean of the parameter vector. Finally,  $\text{DIC} = \overline{D(\bar{\theta})} + p_D = 2\overline{D(\bar{\theta})} - D(\hat{\theta})$ . Smaller values of DIC indicate a better-fitting model. As with other penalized likelihood criteria, DIC is a method for comparing a collection of alternative models (Carlin and Louis 2000).

Burnham and Anderson (2002) derived Akaike weights that, when normalized, can be interpreted as a set of weights that sum to one and estimate the probability that model  $r$  is the best Kullback-Leibler model for the data at hand, given the set of models considered. This approach provides a method for assessing model selection uncertainty. An analogous approach based on DIC has been suggested to assess model selection uncertainty within a Bayesian modeling context (Spiegelhalter et al. 2002a). We used DIC weights ( $w_{\text{DIC}}$ ) to estimate model selection uncertainty for each model  $r$  in the candidate set, calculated using the following formula for Akaike weights:

TABLE 1. Model selection results to identify the candidate models best explaining observed spatial patterns of chronic wasting disease (CWD) prevalence in mule deer in north-central Colorado, USA.

Model number	Scale (km <sup>2</sup> )	Model	<i>p</i> <sub>D</sub>	DIC	w <sub>DIC</sub>
1	9	Demo + Space	71.1	2182.76	0.694
2	9	Demo + Env + Space + Het	80.7	2186.28	0.119
3	9	Demo + Env + Space	50.2	2186.73	0.095
4	9	Demo + Space + Het	95.5	2187.04	0.082
5	81	Demo + Env + Space	22.4	2192.42	0.006
6	9	Demo + Env + Het	65.0	2195.10	0.001
7	81	Demo + Env + Het	30.5	2195.39	0.001
8	81	Demo + Env + Space + Het	34.6	2195.67	0.001
9	81	Demo + Env	6.0	2197.57	0.000
10	81	Demo + Space + Het	78.7	2244.20	0.000

Notes: We examined three analysis scales (grid cells of 9 km<sup>2</sup>, 81 km<sup>2</sup>, and 324 km<sup>2</sup>), using models that incorporated various combinations of demographic (Demo), environmental (Env), spatial (Space), and unstructured heterogeneous (Het) variation. Demo = AGE + SEX; Env = %HAB + %PRIV + DISP (see *Methods* for details); Space is the spatial random effect; Het is the unstructured variation in the model; *p*<sub>D</sub> is the effective number of parameters; DIC is deviance information criteria; w<sub>DIC</sub> informally quantifies model selection uncertainty.

$$w_{DIC} = \frac{\exp(-\frac{1}{2}\Delta DIC)}{\sum \exp(-\frac{1}{2}\Delta DIC)} \tag{5}$$

where ΔDIC was the difference between the minimum DIC value in the candidate set and model *r*, and the denominator was the sum over all models in the set under consideration. The DIC weights are an informal measure and allow easier comparison between models than the DIC value itself.

RESULTS

Our analyses revealed strong support for local influences on observed spatial patterns of CWD prevalence in mule deer. For clarity, only results for the top 10 out of 22 models fit are shown in Table 1. Based on the w<sub>DIC</sub> shown in Table 1, the individual home range scale of 9 km<sup>2</sup> is the only one that merits consideration as the process scale corresponding to the spatial structure of the CWD data. The combined weights for Models 1–4 (Table 1), w<sub>DIC</sub> = 0.99, indicated nearly exclusive support for models at the individual-home range scale. Within this set, all four models contained a spatial random effect for estimating the probability of CWD infection.

TABLE 2. Univariate parameter estimates from Model 1 and Model 2.

Variable	Model rank	Mean	SD	2.5% CI	97.5% CI
SEX	1	0.72	0.13	0.47	0.97
AGE	1	1.24	0.24	0.79	1.73
%PRIV	2	0.89	0.33	0.22	1.57
%HAB	2	0.79	0.39	0.04	1.51
DISP	2	0.03	0.07	-0.13	0.19

Note: Estimates for individual-level covariates SEX and AGE are from Model 1, the top deviance information criteria (DIC) model (which did not contain environmental covariates), with environmental covariate effects from Model 2, the best model containing these effects. CI, credible intervals.

Table 2 shows the posterior means, standard deviations, and 95% credible intervals (CI), all on the logit scale, for the univariate parameters from Models 1 and 2. The individual-level effects of SEX and AGE show that infection probability is higher in males (odds ratio = 2.05 = exp(0.72), 95% CI = 1.60, 2.64) and in animals at least 2 yr old (odds ratio = 3.46 = exp(1.24), 95% CI = 2.20, 5.63) after adjusting for the effects due to the environmental variables. These results are in agreement with earlier studies (Miller et al. 2000, Wolfe et al. 2002, Miller and Conner 2005). Estimates of environmental covariate effects from Model 2 show that %PRIV (odds ratio = 2.45, 95% CI = 1.25, 4.77) and %HAB (odds ratio = 2.21, 95% CI = 1.04, 4.54) significantly influenced infection probability; however, these credible intervals appear relatively wide, reflecting a high degree of uncertainty in the estimates.

Lambda (λ), which is the ratio of spatial variability to total random effect variation was 0.66 with a 95% credible interval ranging from 0.57 to 0.74 for Model 2. Thus, across the landscape, the spatial random effect accounted for between 57% and 74% of the variability attributed to the random effects in that model. This observation, along with the fact that the best model did not contain an unstructured heterogeneity random effect, strengthens the argument that the contact process resulting in landscape-scaled disease heterogeneity is local in nature, influenced more by small-scale spatial structure than by overall unstructured heterogeneity. These results make sense considering that across our entire study area CWD prevalence was ~9% for all deer sampled, making overall infection probabilities relatively low. However, within a single 9-km<sup>2</sup> aerial unit, prevalence rates were estimated as high as 35%, emphasizing that locally dependent processes scale with the spatial distribution CWD infection.

The weight of evidence for Model 1, the top model that contained only demographic covariates and spatially structured heterogeneity, was relatively high at

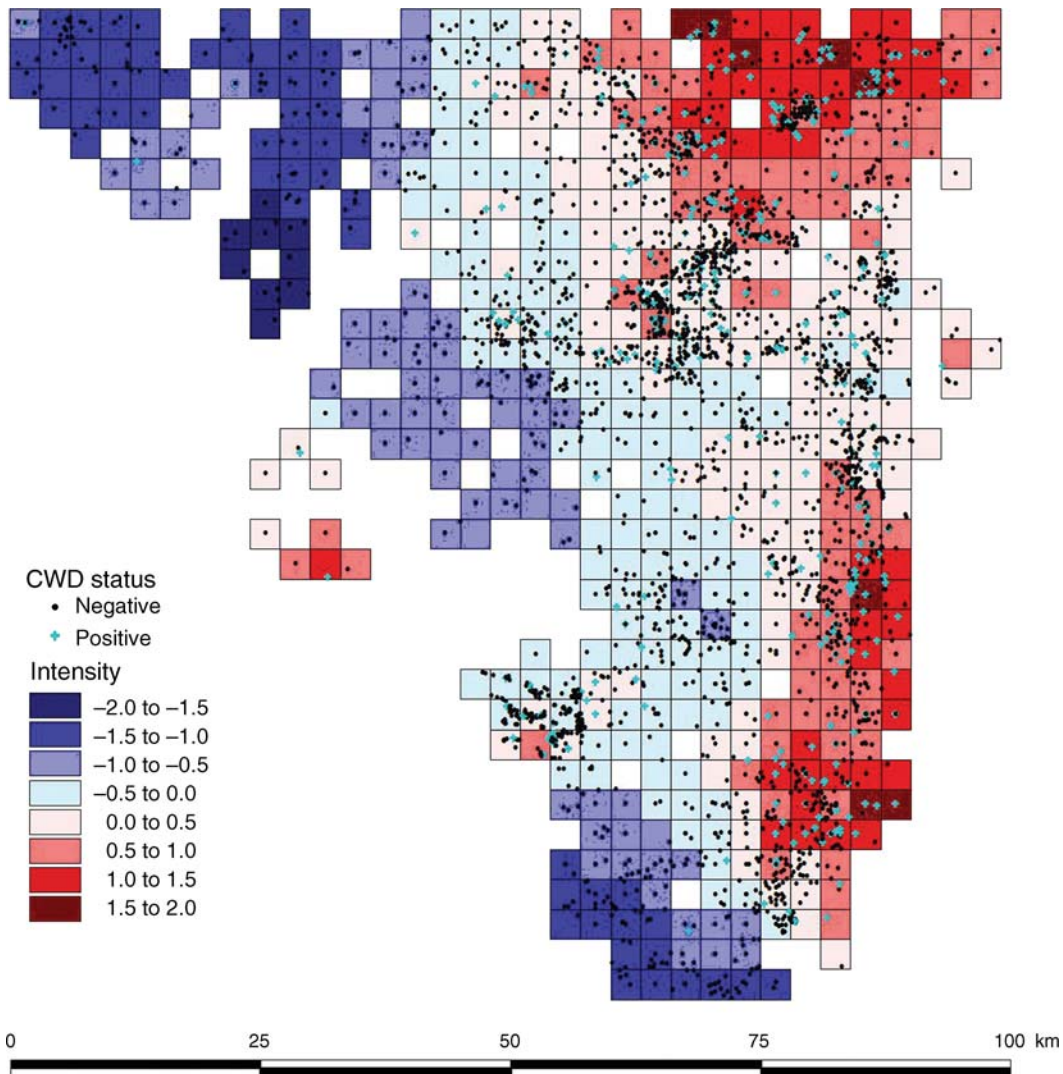


FIG. 2. Posterior estimates of mean spatial random effects for Model 1, the best approximating model of the probability that an individual deer was infected with chronic wasting disease (CWD). The model was fit at a 9-km<sup>2</sup> analysis scale and used the demographic (Demo) and spatial random effects (Space), resulting in a concentrated distribution of infection probabilities.

0.69. Examination of  $p_D$  provides insight into the contribution to model fit made by the two random effect components (Table 1). Recall that Model 2 contained the effects of Model 1 plus the three environmental covariates and unstructured heterogeneity effects. Thus Model 2 has 543 more parameters than Model 1. However,  $p_D$  increased by only 6. This information, combined with a  $w_{DIC}$  of approximately 0.12 for Model 2, suggests that the nonspatial random effects contributed little to the fit of Model 2 and that including these effects, along with the environmental covariates, resulted in a poorer fit to the CWD data than the reduced Model 1.

A visual comparison (Figs. 2 and 3) of the top model containing only demographic covariates and spatial structure (Model 1 in Table 1) with this same model but

with the addition of the environmental covariates (Model 3 in Table 1) shows that Model 1 had a concentrated distribution of posterior spatial random effects across the landscape, while Model 3 had a more diffuse distribution and a lower overall intensity for the spatial random effects. It is not surprising that including environmental covariates in the model has diminished the strength of the local spatial process. This effect can be demonstrated quantitatively by examining the difference in the number of effective parameters,  $p_D$ , between Models 1 and 3, the top models with and without this effect (Table 1). Adding the three environmental covariates to the top model, each contributing a single parameter, reduces  $p_D$  by more than 20 in the resulting third best model. This reduction in  $p_D$  with the addition of three parameters occurred because the environmental

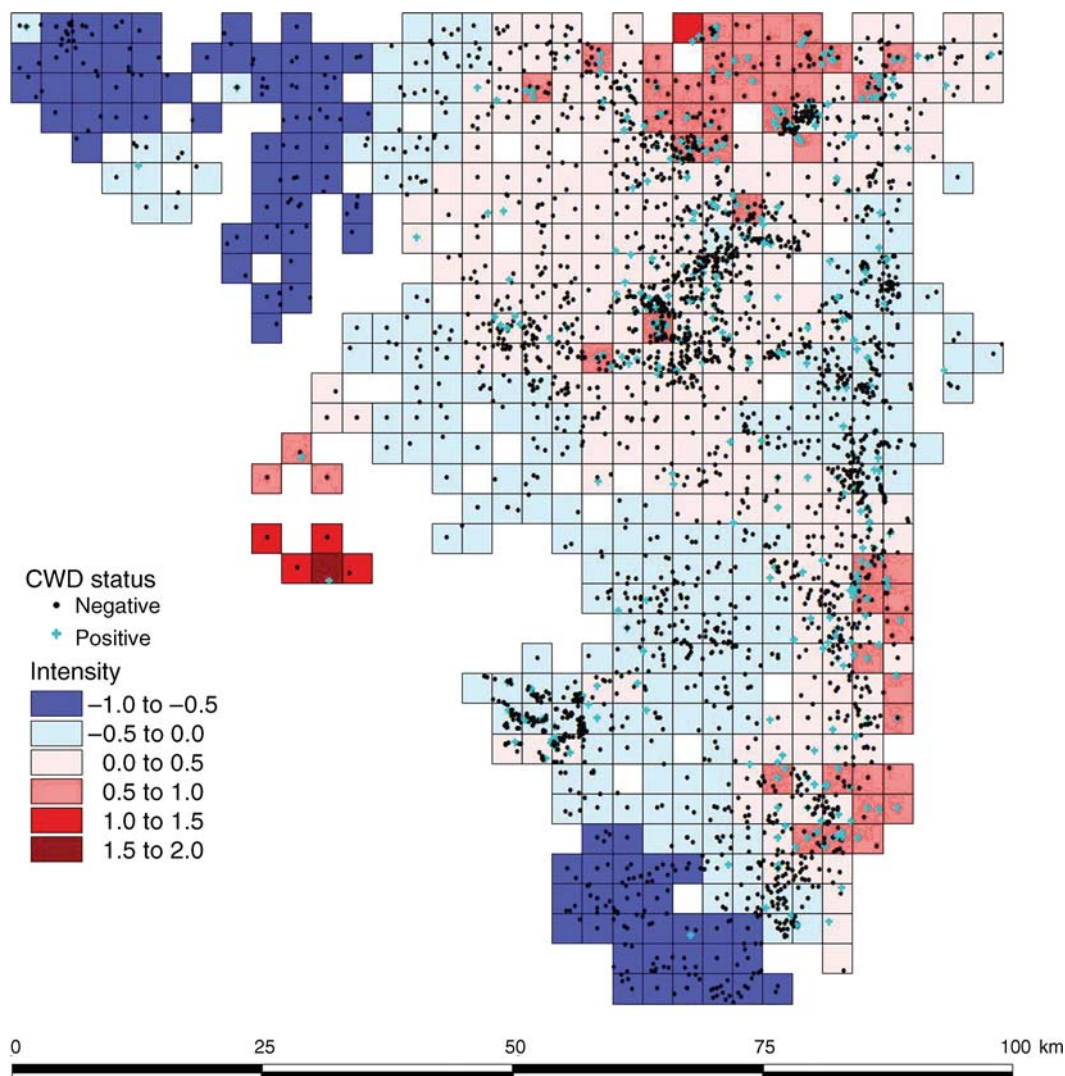


FIG. 3. Posterior estimates of mean spatial random effects for Model 3 of the probability that an individual deer was infected with chronic wasting disease (CWD). The model was fit at a 9-km<sup>2</sup> analysis scale and used the demographic (Demo) and environmental (Env) covariates and spatial (Space) random effects, resulting in a more diffuse distribution of infection probabilities than in Model 1.

covariates “shrunk” the variability in the spatial random effect toward an overall mean effect, thereby reducing the effective number of parameters necessary for modeling the spatial variation in CWD infection probability. Thus, the environmental covariates appear to have compensated for a portion of the residual spatial variation observed when the covariates were absent. To illustrate this, Model 3, which included the landscape covariates, had posterior spatial random effects with 95% credible intervals that overlapped zero for all grid cells, while Model 1 without environmental covariates had spatial random effects that were significantly different from zero for 11% of the grid cells. Still, the environmental covariates produced a poorer fitting model that was less consistent with the observed spatial distribution of CWD, as measured by DIC.

## DISCUSSION

An important first step toward understanding host–pathogen dynamics in spatially structured populations is to understand the scales over which population processes and landscape features shape the host–pathogen relationship. The dynamics of CWD transmission in mule deer are potentially structured by interactions that occur across a range of nested spatial scales in a heterogeneous environment (Miller et al. 2000). Unfortunately, little empirical data exist for relating transmission dynamics, and the resulting spatial distribution of CWD infection, to the scales of seasonal movements that likely help shape observed prevalence patterns. Previous research on the ecology of CWD in mule deer populations identified plausible natural transmission

mechanisms (Miller et al. 2004), as well as apparent influences of demography (Miller and Conner 2005), seasonal movement patterns (Conner and Miller 2004), and land use (Farnsworth et al. 2005) on spatial epidemiology; however, quantitative analyses comparing potential movement scales structuring the spatial distribution of CWD have not been addressed until now.

Our results provide evidence that the spatial structure of CWD results from small-scale, local-contact processes, which likely occur primarily during the winter season when subpopulation home ranges are reduced in size and the potential for infectious contacts among sympatric individuals is possibly increased (Miller and Williams 2003, Conner and Miller 2004, Miller et al. 2004, Miller and Conner 2005). The key to arriving at this conclusion was the use of information about the scales over which the structuring process of movement occurs. Because we specified analysis scales to correspond to the crucial epidemiological process of seasonal movements, we were able to better understand which type of movement pattern (wintering individual, winter subpopulation, or summer subpopulation) appears to be the most plausible process scale underlying observed spatial patterns of CWD prevalence. This result, in turn, provides critical information for further process-based investigations. Our work demonstrates how hypotheses regarding potential generating processes can be evaluated in the face of sparse empirical data. Enumerating the relative contributions from each of these process scales was made possible by the hierarchical modeling framework.

There are biological explanations for the effects of covariates that emerged as important influences on CWD infection probability. Because the clinical course of CWD is protracted, lasting about two years on average (Williams and Miller 2002), we expected an effect of age such that animals older than two years were more likely to test positive for CWD (Miller et al. 2000, Miller and Conner 2005); this highly significant effect is reflected by the models. Private land ownership and developed areas of the landscape often restrict hunter access to deer populations and may encourage their use as refugia (Farnsworth et al. 2005); this restriction possibly leads to populations with larger numbers of males and an older age distribution, both of which are associated with an elevated probability of infection (Miller and Conner 2005). Also, because CWD transmission can occur via exposure to infected animals or environments contaminated by excreta and carcasses from infected animals (Miller et al. 2004), changes in deer distribution or movements affected by land ownership patterns and human alteration of deer habitats could possibly lead to higher local prevalence in these areas (Wolfe et al. 2002, Farnsworth et al. 2005). The sex effect, per se, is more difficult to explain, but could be due to the polygamous mating structure of deer populations combined with effects of predominantly male hunting (Miller and Conner 2005). In this

situation, a sex ratio skewed toward females coupled with the polygamous mating system in which individual males contact groups of females may act to increase the contact rate of males relative to females, thereby increasing their probability of infection.

Our research differs fundamentally from previous pattern-based analyses used in the few wildlife landscape epidemiological investigations that have been undertaken to date. Unlike other studies, we incorporated both host endogenous correlates (sex and age) of the disease and exogenous features of the environment thought a priori to be important predictors of CWD spatial heterogeneity. More importantly, by formulating our approach in terms of a hierarchical model we were able to simultaneously consider the contributions made by these covariates as well as from local spatial structure and overall landscape heterogeneity to the risk of disease occurrence at each of the three movement scales.

Unlike earlier landscape epidemiological investigations in natural systems, we show how an exploration of disease patterns across multiple, process-based, spatial scales can lead to important epidemiological inference regarding plausible biological scales and mechanisms of disease spread. This provides a spatial perspective for further research into the etiology of disease. To fully understand the relationship between host, pathogen, and environment requires in-depth knowledge about host population dynamics and movement patterns, disease etiology, interactions between host and pathogen, and the effects of environmental variation on host and pathogen distributions and dynamics. Achieving this level of understanding requires process-based, biological investigations, informed by statistical analyses that identify relevant scales and correlations among the system's components.

The hierarchical method we have demonstrated provides a powerful approach for a difficult problem in ecology: linking spatial patterns to the scales over which generating processes operate. By maintaining a constant data structure at the lowest (e.g., individual) level in the hierarchy, while varying the scale of the spatial process component of the model, a hierarchical approach allows for direct comparisons of the effect of various process scales on the spatial structure of host–pathogen relationships. In contrast, classical approaches to multi-scaled spatial analysis can suffer from what is known as ecological fallacy, a situation in which inference at one level is based on data collected at a different level (Schwartz 1994, Diez-Roux 1998, Wakefield 2003). Ecological fallacy frequently occurs when the data structure is altered to accommodate multiple scales of analysis, obviating any direct comparison of how different process scales are related to the spatial structure of disease.

Our approach is applicable to many ecological questions for which georeferenced data are available, for example, presence/absence data or counts of

individuals that can be tied to specific locations on the landscape. Further, although our data were in the form of discrete individuals, the hierarchical model structure extends to continuous spatial processes, such as the distribution of nutrients in soils or hydrologic systems, although this extension can introduce additional analytical and computational complexity (Banerjee et al. 2004). Thus, the generality of this approach is applicable across a wide range of ecological questions framed within a spatial context.

Within the Bayesian paradigm there are numerous possible ways to examine the hierarchical model results, and these estimates are accompanied by standard errors that provide for a complete assessment of model uncertainty. For example, in addition to the results considered here, we can construct maps of the mean and standard deviation of cell-level prevalence over the landscape to obtain an aerial estimate of disease prevalence. We can also compare the posterior distribution of disease prevalence over a larger region such as the management units that are used in developing strategies for managing the disease. With additional analysis, we can also compare the spatial cumulative distribution function (Lahiri et al. 1999, Banerjee et al. 2004) for prevalence for two or more regions to determine which areas have the greatest concentration of CWD-infected deer.

Determining the processes that give rise to patterns in nature is difficult in any ecological system. This problem is particularly challenging when we seek to understand processes that act over large geographical or temporal scales. In such cases, data are often limited, and several plausible mechanisms can be identified as potential causes for observed patterns; consequently, it is not always clear how to proceed with testing scientific hypotheses about generating mechanisms. By casting hypotheses regarding scale-dependent processes in terms of models that can be quantitatively compared, hierarchical analyses provide a powerful tool for gaining insight in this context, even when data are limited.

#### ACKNOWLEDGMENTS

Our work was supported by the National Science Foundation–National Institutes of Health Grant DEB-0091961 and National Science Foundation grant DGE-0221595003, CDOW, Federal Aid in Wildlife Restoration Project W-153-R, and the University of Wyoming. We thank K. Burnham for providing helpful comments and suggestions and M. Conner for earlier analyses and database management.

#### LITERATURE CITED

- Aylor, D. E. 2003. Spread of plant disease on a continental scale: role of aerial dispersal of pathogens. *Ecology* **84**:1989–1997.
- Banerjee, S., B. P. Carlin, and A. E. Gelfand. 2004. Hierarchical modeling and analysis for spatial data. Chapman and Hall/CRC, Boca Raton, Florida, USA.
- Bernardinelli, L., D. Clayton, and C. Montomoli. 1995. Bayesian estimates of disease maps—how important are priors? *Statistics in Medicine* **14**:2411–2431.
- Besag, J., and C. Kooperberg. 1995. On conditional and intrinsic autoregressions. *Biometrika* **82**:733–746.
- Besag, J., J. York, and A. Mollie. 1991. Bayesian image-restoration, with 2 applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* **43**:1–20.
- Best, N. G., R. A. Arnold, A. Thomas, L. A. Waller, and E. M. Conlon. 1999. Bayesian models for spatially correlated disease data. Pages 131–156 in A. F. M. Smith, editor. *Bayesian statistics 6*. Oxford University Press, Oxford, UK.
- Best, N. G., K. Ickstadt, and R. L. Wolpert. 2000. Spatial Poisson regression for health and exposure data measured at disparate resolutions. *Journal of the American Statistical Association* **95**:1076–1088.
- Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer, New York, New York, USA.
- Carlin, B. P., and T. A. Louis. 2000. Bayes and empirical Bayes methods for data analysis. Second edition. Chapman and Hall/CRC, Boca Raton, Florida, USA.
- Carlin, B. R., and M.-E. Perez. 2000. Robust Bayesian analysis in medical and epidemiological settings. Pages 351–372 in F. Ruggeri, *Robust Bayesian analysis. Lecture notes in statistics*. Volume 152. Springer-Verlag, New York, New York, USA.
- Conner, M. M., and M. W. Miller. 2004. Movement patterns and spatial epidemiology of a prion disease in mule deer population units. *Ecological Applications* **14**:1870–1881.
- Diez-Roux, A. V. 1998. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *American Journal of Public Health* **88**:216–222.
- Farnsworth, M. L., L. L. Wolfe, N. T. Hobbs, K. P. Burnham, E. S. Williams, D. M. Theobald, M. M. Conner, and M. W. Miller. 2005. Human land use influences chronic wasting disease prevalence in mule deer. *Ecological Applications* **15**:119–126.
- Gelfand, A. E., and A. F. M. Smith. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**:398–409.
- Gelman, A., B. R. Carlin, H. S. Stern, and D. B. Rubin. 2004. *Bayesian data analysis*. Second edition. Chapman and Hall/CRC, Boca Raton, Florida, USA.
- Geman, S., and D. Geman. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**:721–741.
- Gilks, W. R., S. Richardson, and D. J. Spiegelhalter. 1998. *Markov chain Monte Carlo in practice*. Chapman and Hall, Boca Raton, Florida, USA.
- Giraudoux, P., P. S. Craig, P. Delattre, G. Bao, B. Bartholomot, S. Harraga, J. P. Quere, F. Raoul, Y. Wang, D. Shi, and D. A. Vuitton. 2003. Interactions between landscape changes and host communities can regulate *Echinococcus multilocularis* transmission. *Parasitology* **127**:S121–S131.
- Hibler, C. P., et al. 2003. Field validation and assessment of an enzyme-linked immunosorbent assay for detecting chronic wasting disease in mule deer (*Odocoileus hemionus*), white-tailed deer (*Odocoileus virginianus*), and Rocky Mountain elk (*Cervus elaphus nelsoni*). *Journal of Veterinary Diagnostic Investigation* **15**:311–319.
- Hjort, N. L. 1999. Discussion. Pages 150–151 in A. F. M. Smith, editor. *Bayesian statistics 6*. Oxford University Press, Oxford, UK.
- Jarup, L., N. Best, M. B. Toledano, J. Wakefield, and P. Elliott. 2002. Geographical epidemiology of prostate cancer in Great Britain. *International Journal of Cancer* **97**:695–699.
- Jolles, A. E., P. Sullivan, A. P. Alker, and C. D. Harvell. 2002. Disease transmission of aspergillosis in sea fans: inferring process from spatial pattern. *Ecology* **83**:2373–2378.
- Lahiri, S. N., M. S. Kaiser, N. Cressie, and N. J. Hsu. 1999. Prediction of spatial cumulative distribution functions using

- subsampling. *Journal of the American Statistical Association* **94**:86–97.
- Levin, S. A. 1992. The problem of pattern and scale in ecology. *Ecology* **73**:1943–1967.
- McGarigal, K., and B. J. Marks. 1995. FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. USDA Forest Service General Technical Report PNW-351.
- Miller, M. W., and M. M. Conner. 2005. Epidemiology of chronic wasting disease in free ranging mule deer; spatial, temporal, and demographic influences on observed prevalence patterns. *Journal of Wildlife Diseases* **41**:275–290.
- Miller, M. W., and E. S. Williams. 2002. Detection of PrPCWD in mule deer by immunohistochemistry of lymphoid tissues. *Veterinary Record* **151**:610–612.
- Miller, M. W., and E. S. Williams. 2003. Horizontal prion transmission in mule deer. *Nature* **425**:35–36.
- Miller, M. W., E. S. Williams, N. T. Hobbs, and L. L. Wolfe. 2004. Environmental sources of prion transmission in mule deer. *Emerging Infectious Diseases* **10**:1003–1006.
- Miller, M. W., E. S. Williams, C. W. McCarty, T. R. Spraker, T. J. Kreeger, C. T. Larsen, and E. T. Thorne. 2000. Epizootiology of chronic wasting disease in free-ranging cervids in Colorado and Wyoming. *Journal of Wildlife Diseases* **36**:676–690.
- Olden, J. D., and L. N. Poff. 2004. Ecological processes driving biotic homogenization: testing a mechanistic model using fish faunas. *Ecology* **85**:1867–1875.
- Pascual, M., and S. A. Levin. 1999. From individuals to population densities: searching for the intermediate scale of nontrivial determinism. *Ecology* **80**:2225–2236.
- Rahbek, C., and G. R. Graves. 2001. Multiscale assessment of patterns of avian species richness. *Proceedings of the National Academy of Sciences (USA)* **98**:4534–4539.
- Sarnelle, O. 1994. Inferring process from pattern: trophic level abundances and imbedded interactions. *Ecology* **75**:1835–1841.
- Schwartz, S. 1994. The fallacy of the ecological fallacy—the potential misuse of a concept and the consequences. *American Journal of Public Health* **84**:819–824.
- Smith, D. L., L. Ericson, and J. J. Burdon. 2003. Epidemiological patterns at multiple spatial scales: an 11-year study of a *Triphragmium ulmariae*–*Filipendula ulmaria* metapopulation. *Journal of Ecology* **91**:890–903.
- Spiegelhalter, D. J., N. G. Best, B. R. Carlin, and A. van der Linde. 2002a. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B, Statistical Methodology* **64**:583–616.
- Spiegelhalter, D. J., A. Thomas, and N. G. Best. 2002b. WinBUGS. User manual. Version 1.4. MRC Biostatistics Unit, Cambridge, UK.
- Van Buskirk, J., and R. S. Ostfeld. 1998. Habitat heterogeneity, dispersal, and local risk of exposure to Lyme disease. *Ecological Applications* **8**:365–378.
- Wakefield, J. 2003. Sensitivity analyses for ecological regression. *Biometrics* **59**:9–17.
- Wikle, C. K. 2003. Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology* **84**:1382–1394.
- Williams, E. S., and M. W. Miller. 2002. Chronic wasting disease in deer and elk in North America. Scientific and Technical Review Office of International Epizootics **21**:305–316.
- Williams, E. S., and S. Young. 1980. Chronic wasting disease of captive mule deer: a spongiform encephalopathy. *Journal of Wildlife Diseases* **16**:89–98.
- Williams, E. S., and S. Young. 1992. Spongiform encephalopathies of Cervidae. Scientific and Technical Review Office of International Epizootics **11**:551–567.
- Williams, E. S., and S. Young. 1993. Neuropathology of chronic wasting disease of mule deer (*Odocoileus hemionus*) and elk (*Cervus elaphus nelsoni*). *Veterinary Pathology* **30**:36–45.
- Wolfe, L. L., M. M. Conner, T. H. Baker, V. J. Dreitz, K. P. Burnham, E. S. Williams, N. T. Hobbs, and M. W. Miller. 2002. Evaluation of antemortem sampling to estimate chronic wasting disease prevalence in free-ranging mule deer. *Journal of Wildlife Management* **66**:564–573.