

# MARKOV CHAIN MONTE CARLO ESTIMATION OF SPECIES DISTRIBUTIONS: A CASE STUDY OF THE SWIFT FOX IN WESTERN KANSAS

GLEN A. SARGEANT,<sup>1</sup> U.S. Geological Survey, Northern Prairie Wildlife Research Center, 8711 37th Street SE, Jamestown, ND 58401, USA

MARSHA A. SOVADA, U.S. Geological Survey, Northern Prairie Wildlife Research Center, 8711 37th Street SE, Jamestown, ND 58401, USA

CHRISTIANE C. SLIVINSKI,<sup>2</sup> Kansas Department of Wildlife and Parks, 1830 Merchant, Emporia, KS 66801, USA

DOUGLAS H. JOHNSON, U.S. Geological Survey, Northern Prairie Wildlife Research Center, 8711 37th Street SE, Jamestown, ND 58401, USA

**Abstract:** Accurate maps of species distributions are essential tools for wildlife research and conservation. Unfortunately, biologists often are forced to rely on maps derived from observed occurrences recorded opportunistically during observation periods of variable length. Spurious inferences are likely to result because such maps are profoundly affected by the duration and intensity of observation and by methods used to delineate distributions, especially when detection is uncertain. We conducted a systematic survey of swift fox (*Vulpes velox*) distribution in western Kansas, USA, and used Markov chain Monte Carlo (MCMC) image restoration to rectify these problems. During 1997–1999, we searched 355 townships (ca. 93 km<sup>2</sup>) 1–3 times each for an average cost of \$7,315 per year and achieved a detection rate (probability of detecting swift foxes, if present, during a single search) of  $\theta = 0.69$  (95% Bayesian confidence interval [BCI] = [0.60, 0.77]). Our analysis produced an estimate of the underlying distribution, rather than a map of observed occurrences, that reflected the uncertainty associated with estimates of model parameters. To evaluate our results, we analyzed simulated data with similar properties. Results of our simulations suggest negligible bias and good precision when probabilities of detection on  $\geq 1$  survey occasions (cumulative probabilities of detection) exceed 0.65. Although the use of MCMC image restoration has been limited by theoretical and computational complexities, alternatives do not possess the same advantages. Image models accommodate uncertain detection, do not require spatially independent data or a census of map units, and can be used to estimate species distributions directly from observations without relying on habitat covariates or parameters that must be estimated subjectively. These features facilitate economical surveys of large regions, the detection of temporal trends in distribution, and assessments of landscape-level relations between species and habitats. Requirements for the use of MCMC image restoration include study areas that can be partitioned into regular grids of mapping units, spatially contiguous species distributions, reliable methods for identifying target species, and cumulative probabilities of detection  $\geq 0.65$ .

*JOURNAL OF WILDLIFE MANAGEMENT* 69(2):483–497; 2005

**Key words:** Bayesian estimation, carnivore survey, image restoration, Kansas, Markov chain Monte Carlo, species distributions, species-habitat associations, swift fox, *Vulpes velox*, wildlife surveys.

In 1995, the U.S. Fish and Wildlife Service (FWS) issued a long-delayed finding on a petition to list the swift fox (*Vulpes velox*) as an endangered species (60 FR 31663; 16 Jun 1995). According to the FWS, the delay resulted from a general lack of verifiable, quantitative information documenting the current distribution and status of swift foxes. Consequently, an accurate determination of distribution was identified as 1 of the highest-priority information needs for swift fox conservation (U.S. Fish and Wildlife Service 1995).

Unfortunately, swift foxes exemplify the reasons distributions are uncertain for many species of special concern. The potential range of the

swift fox encompasses the shortgrass and mixed-grass prairie regions of 10 states and 3 Canadian provinces (Sovada and Scheick 1999). This broad region encompasses diverse landscapes, vegetation associations, soils, and climates (Allardyce and Sovada 2003, Harrison and Whitaker-Hoagland 2003), and swift foxes are absent from many suitable areas where they formerly occurred (Sovada and Scheick 1999). Habitat requirements of swift foxes are thus sufficiently general, while other factors are sufficiently limiting, to preclude the prediction of swift fox distribution from easily measured surrogates. Instead, the distribution of swift foxes must be estimated directly from observations.

Opportunistic observations often are used to map species distributions; however, such maps are notoriously susceptible to bias resulting from

<sup>1</sup> E-mail: glen\_sargeant@usgs.gov

<sup>2</sup> Present address: 11642 Wall Road, Caledonia, MN 55921, USA.

uneven coverage and variable reporting rates (Bircham and Jordan 1997, Johnson and Sargeant 2002). This may be especially true for swift foxes, which are secretive (Sovada et al. 1998), occur at low densities (Allardyce and Sovada 2003), and may share the canid predilection for long-range dispersal (Schauster et al. 2002a, Sovada et al. 2003). In combination, these features are likely to lead to low reporting rates within core range, yet produce noteworthy, extralimital observations. Distribution maps that are substantially affected by variation in detection rates, reporting rates, or influential observations do not support inferences about temporal trends in distribution. Such trends are of particular interest for swift foxes because relatively rapid changes in distribution and abundance are common among North American canids (e.g., Fuller et al. 1992, Sargeant et al. 1993, Roemer et al. 2001).

Information requirements for swift fox conservation thus create a need for rigorous, repeatable presence/absence surveys; unfortunately, the geographic extent of the problem poses substantial challenges for implementation. Mapping units should be relatively large because temporary absences from habitat patches or even individual home ranges are not germane to the coarse-grained issue of geographic distribution. However, units larger than individual home ranges are difficult to search exhaustively for evidence of a cryptic, secretive species like the swift fox. Consequently, swift foxes are likely to be overlooked in occupied mapping units, perhaps even in the majority of these. The need to accommodate uncertain detection is, therefore, a key consideration for analyses.

When observations of a species are plotted on a map grid, cells in the grid are analogous to the pixels of a digital image. If detection is uncertain and some cells are not searched, as usually is the case for wildlife surveys, this image is imperfect and incomplete. In this sense, maps of observations are analogous to degraded digital images, which commonly are restored via Markov chain Monte Carlo (MCMC) image restoration (Green 1996). The core of image restoration is a Bayesian model for spatial relations between pixels, which can be used to correct erroneous entries and estimate missing values from attributes of neighboring cells.

Parallels between digital images and observations of species were first exploited by Heikkinen and Högmänder (1994) and Högmänder and Møller (1995), who used image restoration to estimate distributions of common toads (*Bufo*

*bufo*) and breeding birds in Finland. However, those applications were published in statistical journals and presupposed a specialized understanding of the notation, terminology, and statistical theory of Bayesian estimation via MCMC simulation. These topics are unfamiliar to most wildlife biologists (Link et al. 2002) and represent a substantial barrier to implementation. As a result, promising results of early applications have inspired minimal subsequent attention despite conceptual advantages over other methods. For example, image models exploit spatial patterns to strengthen estimates and do not require independent data. They can be used to produce estimates for map units that have not been sampled and also accommodate uncertain detection, thereby relieving investigators of the need to census study areas and search map units exhaustively. Habitat covariates may be incorporated but are not required; therefore, image models are useful for mapping distributions of habitat generalists and distributions constrained by other factors to a subset of suitable habitat. Finally, observed occurrences are indices that reflect changes in detection and reporting rates as well as changes in distribution (Anderson 2001). Image models can be used to estimate actual distributions from observations and thus support inferences about changes in distribution, even if detection rates change over time. In short, the use of MCMC image restoration, when coupled with an appropriate sampling design, can resolve many of the most vexing problems that plague surveys of wildlife distributions.

During 1997–1999, we surveyed swift foxes in western Kansas and used Bayesian MCMC image restoration to estimate the species' distribution from observations. In this paper, we (1) describe our survey, (2) derive an image model based on a geometric distribution for search time, (3) fit our model and present the results, (4) use simulation to evaluate the performance of our model, and (5) compare and contrast MCMC image restoration with alternative approaches. Ultimately, we present a rigorous estimate of the distribution of swift foxes in Kansas and provide practicing wildlife biologists with a relatively intuitive introduction to a powerful and practical tool for estimating distributions of rare species. The most difficult details of our analytical methods are not critical to an intuitive understanding, but they will be invaluable for those attempting a similar analysis or seeking an introduction to more specialized statistical literature.

**METHODS**

**Data Collection**

Our study area in western Kansas comprised a nonoverlapping grid of 748 townships (ca. 93 km<sup>2</sup>) that served as mapping units for our survey and analysis. From this grid, we selected a systematic sample of townships (Fig. 1) for 1–3 annual time-constrained track searches. Numbers of searches/township varied because we (1) searched the core of suspected swift fox range during the first year, (2) did not repeat searches in townships after we confirmed the presence of swift foxes, and (3) added townships along the periphery of our study area each year. These steps helped us minimize costs of data collection and distribute them throughout our 3-year study.

For data collection, we contracted with local trappers who were experienced in track identification, knowledgeable about habits of furbearers, and familiar with the search areas. Each observer surveyed 20–80 sample townships/year for 2 hr each or until unambiguous swift fox tracks were found. To detect tracks, observers searched naturally occurring substrates in areas that could be accessed without first securing permission (e.g., secondary and low-maintenance roads, section lines, power line rights-of-way). Because townships could not be searched exhaustively, observers were given detailed maps and told to focus on areas they believed were most likely to be occupied by swift foxes.

Searches were conducted during September–October 1997 and August–September of 1998 and 1999. After precipitation or periods of high winds ( $\geq 24$  km/hr), we suspended searches for  $\geq 24$  hrs to allow tracks to accumulate. Observers recorded the starting time for each search and, if swift fox tracks were found, the time required for detection. To minimize the potential for spurious detections, tracks were photographed, measured, and reviewed by a principal investigator. Only distinct, clearly identifiable tracks were accepted as confirmed occurrences.

**Notation**

We associated 3 random variables with each township, indexed by  $i$ . These included  $x_i$ , the true status of swift foxes (0 = absent, 1 = present);  $v_i$ , the number of years we searched ( $v_i \in \{0, 1, 2, 3\}$ ); and  $y_i$ , the result of our searches (0 = foxes not detected, 1 = foxes detected). Bold-faced type represented vectors (e.g.,  $\mathbf{x} = \{x_1, x_2, \dots, x_{748}\}$ ), and negative subscripts represented the omission of elements from vectors (e.g.,  $\mathbf{x}_{-i} = \{x_1, x_2, \dots, x_{i-1},$

$x_{i+1}, \dots, x_{748}\}$ ). Townships were termed “neighbors” if they were laterally or diagonally adjacent, and the neighbors of township  $i$  (which did not include township  $i$ ) were represented by  $\mathbf{x}'_{-i}$ . Other variables included a coefficient of spatial contagion ( $\beta$ ) and the probability of detecting swift foxes, when they were present, during a single search ( $\theta$ ). We used italic type to represent random variables (e.g.,  $\mathbf{x}, \beta, \theta$ ) and Roman type to represent constants (e.g.,  $\mathbf{x}, \beta, \theta$ ; i.e., when conditioning on specific values of random variables). The ultimate objective of our analysis was to estimate  $\mathbf{x}$ , the true distribution of swift foxes, from the successive states of a Markov random field that were represented by  $\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(J)}$ . Probability densities and distributions play distinctive roles in MCMC simulation. To help distinguish these roles, we used different notation for posterior distributions ( $\pi(\cdot)$ ), likelihoods ( $\ell(\cdot)$ ), prior distributions ( $f(\cdot)$ ), and proposal distributions ( $q(\cdot)$ ). For general references to probabilities of events we used  $\text{Pr}(\cdot)$ .

**Conceptual Model**

Biologists routinely model observations as functions of measured covariates and parameters that are regarded as constants. In contrast, Bayesian models treat both data and parameters as random variables and are envisioned as probability densities for parameters, conditional on data. Image restoration begins with such a conceptual model: in our case, a joint posterior ( $\pi(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y})$ ) for  $\mathbf{x}, \beta$ , and  $\theta$ , conditional on search results that included numbers of searches ( $\mathbf{v}$ ) as well as observations ( $\mathbf{y}$ ).

Our joint posterior was intractable; however, Bayes’ theorem can be used to show (Appendix A) that it was proportional to the product of 4 tractable probability densities:

$$\pi(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y}) \propto \ell(\mathbf{v}, \mathbf{y} | \mathbf{x}, \theta) \times f(\mathbf{x} | \beta) \times f(\theta) \times f(\beta) \tag{1}$$

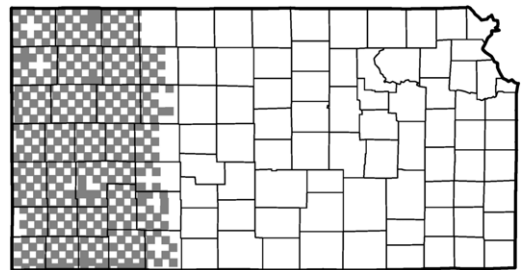


Fig. 1. Study area in western Kansas, USA. Shaded townships were surveyed for evidence of swift foxes during 1997–1999.

or, in the abbreviated terminology customarily employed for Bayesian models,

$$\text{Posterior} \propto \text{Likelihood}(\mathbf{v}, \mathbf{y}) \times \text{Prior}(\mathbf{x}) \times \text{Prior}(\theta) \times \text{Prior}(\beta).$$

This relationship admits a possible solution because MCMC simulation can be used to estimate parameters of distributions known only up to proportionality.

### Model Components

*A Likelihood for Observations.*—Bayesian image models typically incorporate data via likelihoods for observed features of images. In contrast, we used a binomial waiting-time distribution (the geometric distribution) to describe observations and numbers of searches required to detect swift foxes when they were present:

$$\ell(\mathbf{v}, \mathbf{y} \mid x_i = 1, \theta) = [(1 - \theta)^{v_i}]^{(1 - y_i)} \cdot [\theta \cdot (1 - \theta)^{v_i - 1}]^{y_i} \quad (2)$$

Intuitively, our likelihood described the 2 results that were possible for each occupied township: we either failed to detect swift foxes in  $v_i$  visits (i.e.,  $\ell(v_i, y_i = 0 \mid x_i = 1, \theta) = (1 - \theta)^{v_i}$ ) or detected foxes on visit  $v_i$  (i.e.,  $\ell(v_i, y_i = 1 \mid x_i = 1, \theta) = \theta \times (1 - \theta)^{v_i - 1}$ ). False detections were unlikely because we accepted only clearly identifiable tracks as observations. We thus assumed a detection probability of zero ( $\theta = 0.0$ ) where foxes did not occur.

Observations customarily are considered to be conditionally independent, given characteristics of the underlying image to be estimated (see Besag 1986:261, Green 1996:383). Accordingly, we modeled spatial dependence as a feature of the distribution of swift foxes and not as a feature of observations. The joint likelihood for our survey area was thus the product of conditional likelihoods for individual townships, i.e.,

$$\ell(\mathbf{v}, \mathbf{y} \mid \mathbf{x}, \theta) = \prod_i \ell(v_i, y_i \mid x_i, \theta). \quad (3)$$

*A Prior for the Distribution of Swift Foxes.*—In contrast with likelihoods, prior distributions summarize information derived from other sources and do not involve data. In effect, we began our analysis with preconceptions expressed in the form of prior distributions for  $\mathbf{x}$ ,  $\beta$ , and  $\theta$ , and we then used survey data to modify those preconceptions.

Prior to data collection, we did not know which townships would contain swift foxes; hence it might seem that we knew nothing about  $\mathbf{x}$ . However, we did know that species are seldom ran-

domly distributed. Instead, coarse-grained maps of species distributions tend to exhibit spatial dependencies (i.e., occurrences tend to be clustered). In MCMC image restoration, such spatial dependencies are modeled with Markov random fields (Geman and Geman 1984, Besag 1986, Green 1996), which are multidimensional extensions of Markov chains (Geman 1997). Markov chains are sequences of dependent random variables such that the distribution of each variable depends only on the preceding value in the sequence (i.e., if  $\mathbf{x}$  is a sequence,  $\Pr(x_i \mid x_{i-1}, x_{i-2}, \dots, x_1) = \Pr(x_i \mid x_{i-1})$ ; Grimmett and Stirzaker 1994, Link et al. 2002). Markov random fields describe spatial dependencies similarly, in terms of “neighborhoods” surrounding individual cells: (i.e.,  $\Pr(x_i \mid \mathbf{x}_{-i}) = \Pr(x_i \mid \mathbf{x}'_{-i})$ ; Besag 1974).

Our prior for  $\mathbf{x}$  was a variation of the widely used Potts model (Green 1996):

$$f(x_i \mid \mathbf{x}'_{-i}, \beta) \propto \exp \left[ -\beta \cdot w_i \cdot \sum_{x_j \in \mathbf{x}'_{-i}} I(x_i \neq x_j) \right] \quad (4)$$

Intuitively, equation 4 formalized the notion that township  $i$  was likely to resemble its  $J$  neighboring townships with respect to the presence of swift foxes. Components included a coefficient of spatial contagion ( $\beta$ ) that controlled the strength of the resemblance. This indicator function,

$$I(x_i \neq x_j),$$

tallied the number of neighbors that differed from township  $i$  with respect to the presence of swift foxes. Weights ( $w_i = 8/J$ ) scaled the indicator function so it ranged from zero to 8 for townships with  $<8$  neighbors (i.e., where township lines were offset along “correction lines” [Clawson 1968:49], and along the survey area boundary).

*Priors for  $\beta$  and  $\theta$ .*—For  $\beta$  and  $\theta$ , we could specify ranges for plausible solutions but had no reason to prefer specific values within those ranges. We thus used uniform [0,1] priors for both of these variables. The interval [0,1] encompassed admissible values for  $\theta$  (a proportion) and degrees of spatial contagion ranging from independence to strong spatial dependence.

### Estimation

*MCMC Simulation.*—MCMC simulation refers to a group of procedures for generating sequences of dependent random variables with desired limiting distributions (Gilks et al. 1996, Geman 1997, Link et al. 2002). These procedures can be

used to solve intractable statistical problems because any characteristic of a limiting distribution (e.g., mean, median, variance) can be deduced from the distribution of a sufficiently long chain of simulated values. MCMC procedures are especially useful because Bayesian models frequently take the form of a posterior distribution known only up to proportionality. MCMC methods are invaluable in such cases because a function proportional to the target distribution is the only requirement for simulation.

Link et al. (2002) described the Metropolis-Hastings algorithm and a special case, the Gibbs sampler, which are the essential tools of MCMC simulation. In simple cases, an iteration of the Metropolis-Hastings algorithm involves the comparison of the current state of a Markov chain with a proposed new value drawn from a proposal distribution. If the new value is accepted (according to rules defined by the algorithm), it becomes the next state. If it is rejected, the chain remains at the current value for another cycle. Unfortunately, estimating posterior distributions is somewhat more complicated for image models because images comprise large numbers of spatially related elements and are based on prior distributions (e.g., equation 4) specified in terms of individual elements. For these reasons, image models typically are fit with single-component Metropolis-Hastings algorithms (see Gilks et al. 1996:10) that update elements sequentially from full conditional distributions.

*Full Conditional Distributions.*—Map units and model parameters can be updated sequentially because joint posterior distributions (e.g., our  $\pi(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y})$ , equation 1) are uniquely determined by corresponding sets of full conditional distributions (Besag 1974). Full conditional distributions are constructed from posteriors by conditioning, for each variable in turn, on the current states of all other variables to be estimated (Gilks 1996). Conditioning on other variables (Appendix A) and substituting equations 2–4 for corresponding terms led to the full conditionals we used in our analysis:

$$\pi(x_i | \mathbf{x}_{-i}, \mathbf{v}, \beta, \theta) \propto [(1 - \theta)^{v_i}]^{x_i} \cdot \exp[-\beta \cdot w_i \sum_{x_j \in \mathcal{X}_{-i}} I(x_i \neq x_j)]$$

for  $y_i = 0$  (1 otherwise) (5)

$$\pi(\beta | \mathbf{x}) \propto \prod_i \exp[-\beta \cdot w_i \sum_{x_j \in \mathcal{X}_{-i}} I(x_i \neq x_j)]$$

for  $0 \leq \beta \leq 1$  (0 otherwise) (6)

$$\pi(\theta | \mathbf{x}, \mathbf{v}, \mathbf{y}) \propto \prod_i \left\{ [(1 - \theta)^{v_i}]^{(1 - y_i)} \cdot [\theta \cdot (1 - \theta)^{v_i - 1}]^{y_i} \right\}$$

for  $0 \leq \theta \leq 1$  (0 otherwise) (7)

Only 2 states were possible for each township (occupied or unoccupied) and associated probabilities summed to 1. Equation 5 can thus be normalized to produce actual probabilities of occupancy by noting that

$$\Pr(x_i = 1) = \frac{\pi(x_i = 1)}{\pi(x_i = 0) + \pi(x_i = 1)} \quad (8)$$

Normalizing equation 5 is advantageous because townships can be updated by drawing new states directly from binomial distributions when actual probabilities of occupancy are known.

*Implementation.*—We used R statistical software (R Development Core Team 2003), available at <http://www.r-project.org>, to execute our functions and S-Plus (Insightful Corporation 2002) to prepare maps. Our code implemented a single-component Metropolis-Hastings algorithm described by Heikkinen and Högmänder (1994):

- (1) Let the superscript  $j$  indicate the current state of each variable and  $j + 1$  the next state. Assign initial values, denoted  $\mathbf{x}^{(0)}$ ,  $\beta^{(0)}$ , and  $\theta^{(0)}$ .
- (2) Update  $\mathbf{x}^{(j)} \rightarrow \mathbf{x}^{(j+1)}$  by drawing a new state,  $x_i^{(j+1)}$ , for each township in turn. Generate updates by Gibbs sampling using probabilities of occupancy given by equations 5 and 8.
- (3) Draw a candidate value for  $\beta^{(j+1)}$  from a proposal distribution,  $q(\beta)$ . Let  $\beta'$  represent this candidate and let  $\pi(\beta)$  represent the full conditional distribution for  $\beta$  (equation 6). Accept the candidate with probability given by the Metropolis-Hastings algorithm:

$$\Pr(\beta^{(j)} \rightarrow \beta') = \min \left( 1, \frac{\pi(\beta') \cdot q(\beta^{(j)})}{\pi(\beta^{(j)}) \cdot q(\beta')} \right)$$

If the candidate is rejected, let  $\beta^{(j+1)} = \beta^{(j)}$ .

- (4) Use  $q(\theta)$  and  $\pi(\theta)$  from equation 7, and follow the steps in 3 to update  $\theta^{(j)}$ .

(5) Repeat steps 2–4 until the distributions of resulting values approximate the marginal posterior distributions of  $\mathbf{x}$ ,  $\beta$ , and  $\theta$ .

(6) Average the replicates for each township to produce an estimated probability of occupancy,  $\bar{x}_i$ . Note that proposal distributions in steps 3 and 4 serve merely to generate new candidate values for

consideration. Consequently, the distributions used to generate proposals ( $q(\beta')$  and  $q(\theta')$ ) were not critical as long as they spanned distributions they were used to estimate; however, proposal distributions that result in high acceptance rates for proposals and chains with low correlations between successive states were desirable because they led to rapid mixing and efficient estimation. We achieved satisfactory results by drawing proposals from normal distributions with means matching current states and standard deviations of 0.05, which we chose by trial and error.

An intuitive understanding of steps 3 and 4 can be gained by considering the Metropolis-Hastings algorithm to be the product of 2 ratios. In step 3, for example,

$$\frac{\pi(\beta')}{\pi(\beta^{(j)})}$$

favors proposals (i.e., is  $>1$ ) when they are more credible solutions than current values. In contrast,

$$\frac{q(\beta^{(j)})}{q(\beta')}$$

penalizes proposals (i.e., is  $<1$ ) when they are more likely to be proposed than current values. It is the balance of these 2 influences that results in the convergence of a Markov chain to its target distribution. For additional details of Gibbs sampling, the Metropolis-Hastings algorithm, and the method of sampling from full conditional distributions, see Casella and George (1992), Besag et al. (1995), Gilks (1996), Gilks et al. (1996), and Gamerman (1997).

### Model Assessment

We used simulation to assess the potential for confounding of parameter estimates, the consequences of violating assumptions, and the behavior of our estimators over a range of detection rates. To simulate data, we first used estimated probabilities of occupancy to classify each township as occupied ( $\bar{x}_i > 0.5$ ) or unoccupied ( $\bar{x}_i \leq 0.5$ ). The result constituted a target distribution to be estimated from sample data. We then used sequential Bernoulli trials to develop a search and detection history for each of our sample townships (e.g., a township with a history of 000 was searched 3 times without a detection, and a township with a history of 01 was searched twice with a detection on the second occasion). We simulated detection rates in 3 different ways:

fixed detection rates (0.10, 0.20, 0.30, 0.70), detection rates that varied randomly around a fixed mean of 0.70, and detection rates that increased from east to west from 0.56 to 0.84 ( $0.70 \pm 20\%$ ). To simulate random variation around a fixed mean, we drew detection rates from a beta distribution with shape parameters  $a = 4.673$  and  $b = 2$  ( $\theta^- = 0.70$ ,  $\sigma_\theta = 0.17$ ). We also investigated consequences of missing data or sparser sampling schemes by randomly deleting 20% and 40% of sample cells from simulations with detection rates of 0.30 and 0.70.

We analyzed 10 simulated data sets for each setting of  $\theta$ . Initially, each analysis consisted of 6,000 iterations; however, satisfactory convergence required 10,000 iterations for  $\theta = 0.10$ . We discarded the first 1,000 iterations from each set to permit estimates to diverge from initial values and estimated  $\beta$ ,  $\theta$ , and  $\bar{x}$  from the remainder. When computing error rates for maps, we considered estimated probabilities of occurrence  $>0.5$  to suggest occupancy and those  $\leq 0.5$  to suggest absence.

## RESULTS

### Swift Fox Surveys

During 1997–1999, we conducted 619 searches (264 in 1997, 224 in 1998, and 131 in 1999) of 355 townships and detected swift foxes in 173 townships (49%). We detected swift foxes in 110 townships during 1997. In 1998, we dropped the townships with detections in 1997 and 10 others for logistical reasons, repeated searches of 144 townships, searched 80 new townships, and found swift foxes in 56 townships. In 1999, we (1) dropped the 56 townships with detections in 1998 and an additional 50 townships for logistical reasons, (2) repeated searches of 2 townships searched only in 1997, 38 townships searched only in 1998, and 80 searched in both 1997 and 1998, and (3) added 11 townships; and (4) detected foxes in 7 townships.

Detections accumulated at rates that decreased rapidly during the 2-hr search period allotted for each township: 52% occurred during the first 30 min of township searches, 79% occurred during the first 60 min, and 93% occurred in the first 90 min. Most detections occurred the first (75%) or second time (22%) a township was searched. Total costs (paid to contractors who bid on a cost-per-township basis) were \$9,700 (\$36.74/township) in 1997, \$7,600 (\$33.93/township) in 1998, and \$4,644 (\$35.45/township) in 1999.

**MCMC Estimates**

We achieved relatively high acceptance rates for proposals (63% for  $\beta$  and 69% for  $\theta$ ) and modest correlations between successive states of chains (0.85 for  $\beta$  and 0.80 for  $\theta$ ). High acceptance rates and modest correlations between states were desirable because they contributed to rapid convergence and mixing, thereby limiting the number of iterations required for satisfactory parameter estimates. Although results in this section are based on 20,000 iterations preceded by a burn-in of 5,000 iterations, much shorter runs would have sufficed. For example, estimated probabilities of occupancy based on 1,000 iterations preceded by a burn-in period of 100 iterations consistently deviated from results of longer runs by  $\leq 0.02$  for 95–97% of townships, regardless of starting values. Discrepancies between estimates of  $\beta$  and  $\theta$  were consistently  $< 0.005$ .

The distribution of swift foxes was contagious ( $\hat{\beta} = 0.39$ ; 95% BCI = [0.30, 0.48]); consequently, neighbors provided considerable evidence about the presence of swift foxes in townships that were not searched (Fig. 2). However, swift foxes were very likely to be detected when present:  $\hat{\theta} = 0.69$  for townships searched once (95% BCI = [0.60, 0.77]),  $\hat{\theta} = 0.90$  for townships searched twice (95% BCI = [0.84, 0.95]), and  $\hat{\theta} = 0.97$  for townships searched 3 times (95% BCI = [0.94, 0.99]). Search results, therefore, exerted a substantial influence over estimated probabilities of occupancy for townships that were searched (Fig. 2). For example, note that a township searched 3 times without a detection would not have achieved a 50% probability of occupancy, even if every neighboring township were occupied.

Estimated probabilities of occupancy were dominated by high and low values: 68% were  $< 0.10$  or  $> 0.90$ , and 82% were  $< 0.20$  or  $> 0.80$ . Our restoration thus produced a relatively unambiguous estimate of the geographic distribution of swift foxes in western Kansas (Fig. 3). Note that the strong resemblance between the map of our observations and our estimate (Fig. 3) was a fortuitous consequence of a high detection rate and should not be expected in every case.

**Model Assessment Via Simulation**

Analyses of simulated data produced satisfactory estimates of detection rates ranging from 0.10 to 0.70 (Table 1). Although our estimates suggested some positive bias, the apparent bias ( $\approx 2$  percentage points) was modest relative to sampling variation and would not have had practical impli-

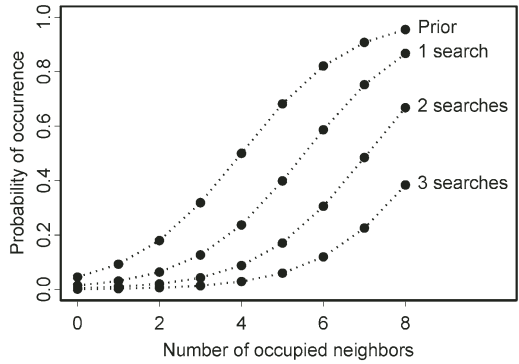


Fig. 2. Prior and full conditional probabilities of occupancy for Kansas, USA, townships without observations of swift foxes, 1997–1999.

cations for the interpretation of survey results. Estimates of mean detection rates were not affected discernibly when we imposed random variation or a spatial trend in detection rates.

In contrast, low detection rates ( $\theta < 0.3$ ) resulted in erratic estimates for  $\beta$  and our target distribution (Fig. 4A) unless we increased the number of searches to compensate. This variability was reflected in (1) large standard deviations for  $\hat{\beta}$  (Table 1), (2) relatively high misclassification rates (Table 1), (3) maps that generally overestimated the extent of our target distribution (e.g., Fig. 4B; Table 1), and (4) relatively ambiguous classifications of map units as occupied or unoccupied (Fig. 5). These features, though undesirable in survey products, were appropriate because they properly conveyed the uncertainty inherent in maps that are not well supported by data. They arose because spatial

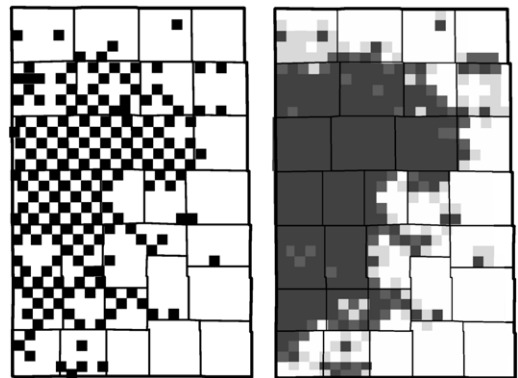


Fig. 3. Observations of swift foxes in Kansas, USA (left), and estimated probabilities of township occupancy (right). At right, unshaded cells indicate strong evidence of absence ( $\text{Pr}[\text{Occupancy}] < 20\%$ ). Shaded cells, from lightest to darkest, represent weak evidence of absence (20–50%), weak evidence of occupancy (50–80%), and strong evidence of occupancy ( $> 80\%$ ).

Table 1. Results of surveys and Markov chain Monte Carlo image restorations for simulated swift fox survey data. Simulations featured 3 sequential searches for each township unless otherwise noted. Per search detection rates were either (1) constant within the range from 0.1 to 0.7, (2) variable but spatially stationary and drawn from a beta distribution ( $\bar{\theta} = 0.70, \sigma_{\theta} = 0.17$ ), or (3) variable, incorporating an east-to-west linear trend ranging from 0.56 to 0.84 ( $\bar{\theta} = 0.70 \pm 20$ ). Misclassification rates are medians; other results represent means and standard deviations (in parentheses) for 10 trials at each parameter setting.

$\theta$	Number of detections	Total searches	$\hat{\theta}$	$\hat{\beta}$	Median proportion of townships misclassified				
					Occupied townships			Unoccupied townships	
					Survey	MCMC estimates		MCMC estimates	
					Sample townships	Sample townships	Other townships	Sample townships	Other townships
0.1	54 (8)	1,006 (9)	0.10 (0.02)	0.49 (0.10)	0.71	0.06	0.01	0.35	0.52
0.1 <sup>a</sup>	122 (8)	2,905 (56)	0.11 (0.02)	0.40 (0.04)	0.35	0.10	0.09	0.04	0.04
0.2	91 (12)	960 (14)	0.20 (0.05)	0.44 (0.07)	0.52	0.11	0.07	0.17	0.24
0.2 <sup>b</sup>	126 (8)	1,470 (30)	0.22 (0.04)	0.40 (0.05)	0.33	0.10	0.10	0.02	0.02
0.3	126 (6)	910 (12)	0.33 (0.05)	0.40 (0.04)	0.32	0.09	0.06	0.02	0.05
0.4	144 (5)	867 (10)	0.45 (0.04)	0.38 (0.02)	0.23	0.11	0.11	<0.01	0.01
0.5	160 (4)	837 (8)	0.52 (0.03)	0.37 (0.02)	0.14	0.09	0.06	<0.01	0.01
0.6	173 (3)	798 (10)	0.62 (0.03)	0.39 (0.02)	0.07	0.06	0.05	0	<0.01
0.7	181 (2)	765 (7)	0.71 (0.03)	0.40 (0.01)	0.02	0.02	0.04	0	<0.01
0.7 <sup>c</sup>	107 (5)	460 (10)	0.71 (0.04)	0.40 (0.01)	0.02	0.02	0.16	0	0.07
beta	175 (4)	771 (6)	0.72 (0.02)	0.38 (0.02)	0.06	0.06	0.09	0	<0.01
[0.56–0.84]	181 (3)	756 (8)	0.75 (0.02)	0.40 (0.01)	0.03	0.03	0.05	0	<0.01

<sup>a</sup> 10 sequential searches of each township.  
<sup>b</sup> 5 sequential searches of each township.  
<sup>c</sup> Data deleted for 40% of sample cells.

arrangements of observations were highly variable and survey results constituted weak evidence of absence when detection rates were low. Estimates of detection rates were not similarly affected because their precision depended on numbers of survey occasions and map units sampled but not on spatial arrangements of observations.

When we increased detection rates to  $\geq 0.30$  or increased numbers of searches to achieve cumulative detection rates (probabilities of detection on  $\geq 1$  occasion) of  $\approx 0.65$ , maps faithfully reflected key features of our target distribution (e.g., Fig. 4C) and misclassification rates were low for both occupied and unoccupied map units (Table 1).

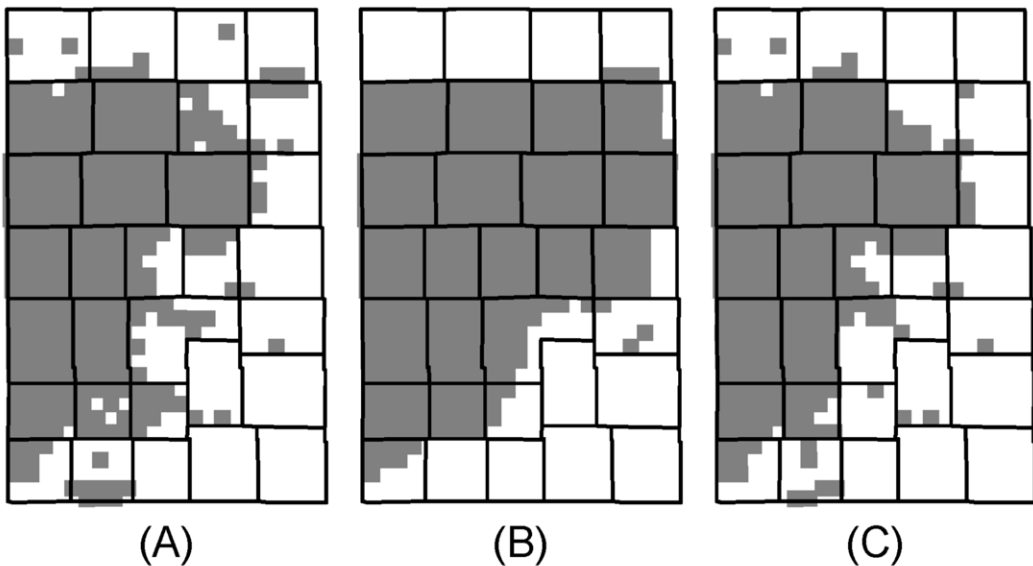


Fig. 4. (A) Target distribution used to simulate survey results for swift foxes in western Kansas, USA, 1997–1999, followed by representative examples of MCMC estimates based on 3 searches per township and per-search detection probabilities of (B)  $\theta = 0.10$ , and (C)  $\theta = 0.30$ .



Imposing spatial variation in detection rates slightly increased misclassification rates for occupied map units (Table 1), but it did not have a meaningful effect on maps or the utility of results.

When we specified high per search detection rates ( $\theta \approx 0.7$ ), simulated survey results and MCMC estimates rarely classified sample units differently; however, we could not have inferred low error rates (Table 1) without estimates of detection rates, and survey results alone would not have supported objective classifications for map units that were not searched. Simulations with less intensive sampling designs than ours (e.g., Fig. 6A) served to illustrate the importance of this latter consideration. We were able to randomly delete  $\geq 40\%$  of sample cells from simulated data, resulting in simulated observations (Fig. 6B) that did not closely resemble our target distribution (Fig. 4A), without compromising estimates (Table 1, Fig. 6C). Despite somewhat higher rates for errors of omission, resulting maps correctly classified 95% of cells (median,  $n = 10$  trials) and faithfully depicted most boundary details.

## DISCUSSION

### Distribution of Swift Foxes in Kansas

Prior to European settlement, swift foxes occupied most of the mixed-grass and short-grass prairie habitat in the western half of Kansas (Zumbaugh and Choate 1985, Sovada and Scheick 1999). Reports of the current distribution agree that swift foxes are much less widespread than they were historically, but the reports are otherwise inconsistent. For example, the FWS (1995) considered the current range of swift foxes to extend from Nebraska to Oklahoma along the Colorado border (Fig. 7A). At about the same time, the Kansas Department of Fish, Wildlife, and Parks (KDFWP; Fox 1993 in U.S. Fish and Wildlife Service 1995) reported opportunistic observations of foxes in Wallace, Logan, Gove, Greeley, Wichita, Hamilton, Kearney, Stanton, Sherman, and Scott counties (Fig. 7B). Our estimate of swift fox distribution (Fig. 7C) is

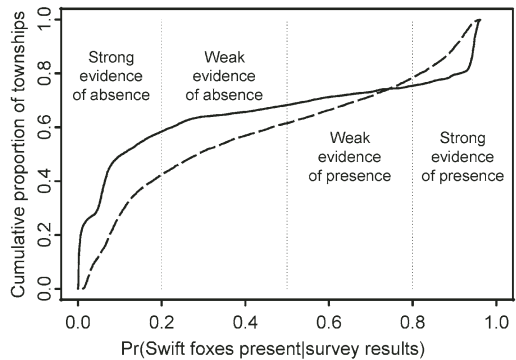


Fig. 5. Effect of detection rate on the strength of evidence for and against the occupancy of townships by swift foxes in western Kansas, USA, 1997–1999. Results derived from simulated survey results based on the target distribution in Fig. 4A, 3 searches per township, and per-search detection probabilities of  $\theta = 0.30$  (dashed line) and  $\theta = 0.70$  (solid line).

similar to, but more extensive than, the distribution attributed to the KDFWP. It does not resemble the distribution reported by the FWS.

Differences among the maps in Fig. 7 exemplify ambiguities that arise from opportunistic sampling and ad hoc methods of boundary determination. Distribution maps are not comparable unless, like Fig. 7C, they depict well-defined biological quantities at a specified scale of spatial and temporal resolution. Our estimates thus represent a baseline for monitoring future changes, but they should not be used to infer past changes in the distribution of swift foxes.

### Model Validation and Evaluation

We could not compare our estimate to the true distribution of swift foxes because we could not

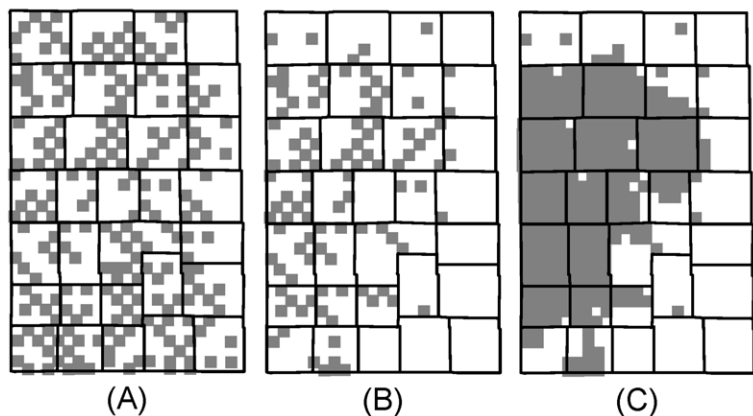


Fig. 6. (A) Simulated sampling grid generated by deleting 40% of sample cells, (B) simulated observations resulting from 3 searches with a detection probability of 0.70 and the target distribution in Fig. 4A, and (C) the resulting MCMC estimate of the distribution in Fig. 4A.

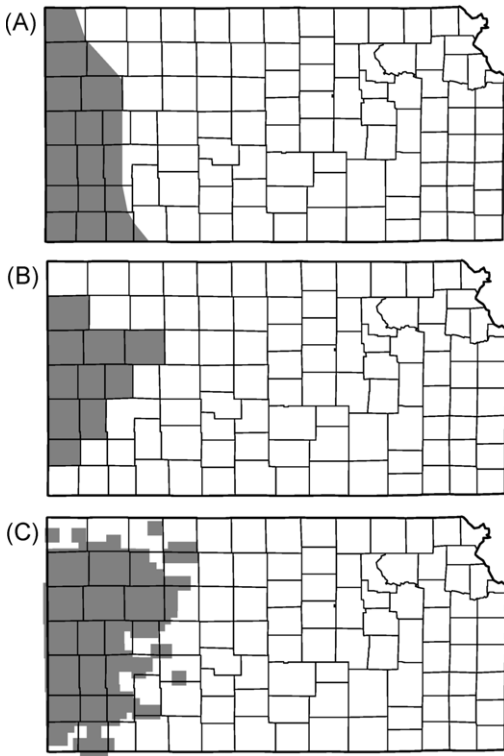


Fig. 7. Range of swift foxes in Kansas (A) redrawn from Fig. 1 of FWS (1995), (B) attributed by FWS (1995) to Lloyd Fox, Kansas Department of Fish, Wildlife and Parks, and (C) estimated via MCMC image restoration from surveys conducted in Kansas, 1997–1999. In (C), shaded cells represent townships with probabilities of occupancy  $>0.50$ .

search townships exhaustively and confirm absences. However, testing models against artificial data with known structure can provide valuable insights about the performance of species distribution models when the “truth” is unknown (Heglund 2002, Stauffer 2002). Insights are possible because binary data are relatively simple manifestations of the complex, interacting influences on species distributions and processes of observation. Data structures that are likely to arise in practice may, therefore, be easy to simulate, even when mechanisms producing them are poorly understood. For swift foxes in Kansas, the most likely complications, regardless of cause, were heterogeneous detection rates and errors of commission. Our methods clearly were robust to these influences (Table 1; subsequent discussion).

### Survey Design

Our time-constrained track searches were inspired by predator surveys conducted by

Sargeant et al. (1993), who searched a sample of quarter-sections in the Prairie Pothole Region of the USA and Canada. Sargeant et al. (1993) were able to search quarter-sections intensively and, therefore, assumed tracks would soon reveal the presence of predators that routinely visited their sample sites. In contrast, the geographic extent and the grain of our study dictated much larger sample units (townships) that we could not search exhaustively; consequently, we expected swift foxes to be overlooked in an unknown proportion of occupied units we surveyed, perhaps even in the majority. Our design objectives were, thus, to facilitate the estimation of detection rates and occupancy probabilities via MCMC image restoration, while simultaneously minimizing survey costs.

Our systematic, checkerboard sample of townships was more economical than a census and represented an efficient allocation of sampling effort for image restoration. Each interior township was assured of  $\geq 4$  neighbors with sample data; hence, our design provided good support for estimates of occupancy probabilities. Search periods of 2 hr proved to be adequate and economical because detections rarely occurred after 90 min. Finally, foxes occurred in most sample townships, and we achieved a relatively high detection rate. Terminating searches of townships when we found evidence of foxes thus resulted in considerable savings and enabled us to expand our study area each year.

We crafted our sampling design to produce satisfactory results under very general conditions. Consequences of using different designs are difficult to generalize because they will depend on detection rates, numbers of surveys conducted, proportions and spatial distributions of cells sampled, and strengths of spatial dependencies in species distributions under study. Investigators can safely assume, however, that sampling intensities greater than ours are desirable when they are feasible. Conversely, sparser designs than ours may merit consideration if detection rates are likely to be high, spatial dependencies are likely to be strong, and sample cells will be well-distributed throughout the area of interest.

### Advantages of MCMC Image Restoration

Our analysis of swift fox distribution encompassed several tasks that have been accomplished by alternative methods. For example, probabilities of occurrence have been estimated with logistic regression (Osborne and Tigar 1992, Buckland and Elston 1993, Robertson et al. 2002), autologistic regression (Augustin et al. 1996, Klute et al. 2002), and models for results of sequential searches (here-

after detection-history models; McArdle 1990, Kéry 2002, MacKenzie et al. 2002). Detection-history models have also been used to estimate detection rates (McArdle 1990, Kéry 2002, MacKenzie et al. 2002). Regression models have been used to predict species distributions from covariates (Osborne and Tigar 1992, Buckland and Elston 1993), and range maps have been prepared by smoothing maps of observations (Johnson and Sargeant 2002). We preferred MCMC image restoration to these alternatives because it combined their strengths and accomplished the same objectives.

Logistic regression, for example, can be used to relate probabilities of occupancy to covariates and then to predict probabilities of occurrence for map units that were not sampled. Unfortunately, actual probabilities of occurrence are confounded with detection rates and the resulting bias cannot be estimated unless detection rates can also be estimated. For this reason, logistic regression models are of doubtful value when detection is uncertain. Predictions based on habitat covariates also require restrictive ecological assumptions (O'Connor 2002*a,b*) that are not met by generalists or by species with distributions restricted by other factors to a subset of potentially suitable habitat. Habitat models are, therefore, of limited utility for constructing maps or for tracking changes in species distributions.

Detection-history models (McArdle 1990, Kéry 2002, MacKenzie et al. 2002) can provide estimates of both detection and occupancy probabilities, but they do not produce estimates for map units that have not been sampled. Like logistic regression, detection-history models are predicated on the independence of sampling units and are thus ill suited for use in the preparation of distribution maps. Because species distributions generally display substantial spatial continuity, the presence of a species in (or absence from) neighboring map units is likely to be an exceptionally useful predictor of occupancy (Heikkinen and Högmander 1994, Högmander and Møller 1995, Augustin et al. 1996). Spatial dependencies thus represent a valuable source of information that should not be ignored when a map of distribution is the ultimate objective of an investigation. Ignoring spatial dependencies can also lead to incorrect inferences (Klute et al. 2002).

Autologistic regression models and spatial smoothing exploit the continuity of species ranges, but the former share other shortcomings of logistic regression. Smoothing can markedly improve the correspondence between maps of observations

and actual species distributions (Johnson and Sargeant 2002), but resulting maps are substantially influenced by the degree of smoothing imposed, which must be estimated from observations or chosen subjectively to produce a visually pleasing result. Estimating smoothing parameters from data is problematic because observations give misleading impressions of spatial continuity when detection rates are low (Heikkinen and Högmander 1994) and the benefits of smoothing are potentially the greatest. Maps developed by smoothing thus give general impressions of species distributions but not in terms of well-defined units (e.g., probabilities of occupancy), and they do not support variance estimates or rigorous assessments of changes in species distributions.

Bayesian image models rectify some deficiencies and combine various advantages of covariate models, capture-history models, and spatial smoothing. Like autologistic regression and spatial smoothing, image models exploit spatial patterns to improve map estimates. Like autologistic regression, they can be used to develop objective estimates of the appropriate degree of smoothing, and like capture-history models, they can accommodate uncertain detection. Desirable features include valid Bayesian confidence intervals (Johnson 1999:770) for model parameters and estimates for map units that are directly interpretable as probabilities of occupancy, conditional on observations.

### Previous Applications

Bayesian image models have a relatively short but distinguished history of use in such diverse fields as statistical physics, medical image processing, epidemiology, archaeology, biogeography, and agricultural field trials (Besag et al. 1995, Green 1996). However, applications generally have been relegated to statistical journals that presuppose a specialized understanding of theory and terminology. We sought to synthesize desirable statistical features of previous applications, identify and dispense with unnecessary complexities, incorporate considerations specific to wildlife surveys, and develop an intuitive, self-contained presentation of key concepts.

Of previous applications, our analysis most closely resembled that of Heikkinen and Högmander (1994) because we incorporated their fully Bayesian approach for the estimation of model parameters. However, they and Högmander and Møller (1995) analyzed atlas data collected for different purposes and did not have accurate measurements of survey effort. Högmander and

Møller (1995) thus highlighted the need for more economical sampling designs that facilitate MCMC image restoration and suggested that sampling effort should be fixed a priori. Critical aspects of model performance also remained to be explored. We extended the work of Heikkinen and Högmander (1994) by (1) developing an economical and effective sampling design, (2) incorporating a waiting-time model that achieved the goal of fixed sampling effort with fewer assumptions and for lower cost, and (3) using simulation to evaluate the performance of our model and assess consequences of assumptions that are not shared by alternative methods of analysis.

### Assumptions

*Accurate Track Identification.*—Track surveys are in widespread use as indicators of swift fox presence (Harrison and Schmitt 2003, Olson et al. 2003, Shaughnessy 2003) and relative abundance (Schauster et al. 2002b, Sargeant et al. 2003). Although accurate track identifications are fundamental to the validity of such methods, users generally have not addressed the potential for, or consequences of, misidentifications.

Our image model was unique among methods used to estimate swift fox distributions because misidentifications of swift fox tracks were accommodated analytically as 1 of several factors contributing to uncertain detection. This feature of our analysis was desirable in its own right but more so because we believe it reduced the potential for errors of commission (i.e., detections of foxes where they do not occur). We believe observers are comparatively unlikely to declare a species to be present when they are uncertain if errors of omission are viewed as acceptable outcomes.

Observers reported several isolated detections of swift fox tracks to the north, east, and south of contiguous swift fox range (Fig. 3). Those observations may have represented errors of commission; however, the few potentially spurious detections in a large region classified predominantly as unoccupied suggests a low error rate. Isolated errors would not have exerted a strong influence on estimated probabilities of occupancy except for townships where they occurred (Fig. 2).

*Homogeneity of Detection Rates.*—Sovada et al. (2003) captured similar numbers of foxes and observed similar home-range sizes in cropland and rangeland, which represented extremes of swift fox habitat that was available within our survey area. In addition, opportunities for detection were not limited by the availability of suitable sites for track

detection, which were superabundant. Finally, we sought to limit the potential for heterogeneity arising from differences in lengths of time required for detection by searching for 2 hr, even though detections rarely occurred after 90 min. Collectively, these considerations justified our use of a single-parameter model for detection probabilities. However, we probably could not entirely eliminate effects of such factors as accessibility, tracking conditions, and species abundance. In principle, likelihoods used to estimate detection rates can reflect the influences of such covariates (Heikkinen and Högmander 1994, MacKenzie et al. 2002). In practice, however, complex spatial models for differences in detection rates will often be difficult to justify, either because sample size requirements will be prohibitive or because likely influences (e.g., abundance) will be difficult or impossible to quantify. Even when sample sizes are sufficient and detection rates are related to covariates that can be measured, conventional methods of model selection will be frustrated by spatial relationships among observations and a joint posterior likelihood known only up to proportionality. Consequently, models for heterogeneity are most likely to be useful when factors affecting detection rates are well understood (e.g., when sample units are relatively homogeneous and can be unambiguously assigned to a few classes that are obviously different).

Given the difficulties of developing realistic models for subtle differences in detection rates, the results of our simulations are reassuring. Neither random variation nor a substantial spatial trend in detection rates had a consequential effect on estimates of mean detection rates or the restoration of a simulated target distribution resembling our results for swift foxes.

*Homogeneity of Spatial Contagion.*—Whereas detection-rate models (McArdle 1990, Kéry 2002, MacKenzie et al. 2002) require independent sample sites, we assumed only that neighboring townships displayed similar degrees of spatial contagion throughout our study area. Given a reason to suspect strong spatial patterns, large study areas can be subdivided and each can be assigned a different parameter, or parameters can be specified as functions of covariates. However, image models usually are based on stationary priors (Green 1996) because such models have been found to perform well in many contexts. Relatively simple, contiguous distributions, like that of swift foxes in western Kansas, do not suggest a need for more complex models.

*Closure.*—MCMC image restoration and detection-history models share a requirement for clo-

sure of map units to changes in occupancy (McArdle 1990, Kéry 2002, MacKenzie et al. 2002). If this assumption is violated, probabilities of detection and probabilities of occupancy are confounded. Due consideration must, therefore, be given to such factors as the longevity, site fidelity, and intra-specific spatial relationships of species being studied, as well as to the potential for abrupt range expansions or contractions. Annual searches of townships may suffice for swift foxes in Kansas but not for shorter-lived or transient species, for species that occur at much lower densities, or where rapid changes in distribution are considered likely. In Minnesota, for example, we are studying species that are more uniformly distributed, and we are concerned about rapid changes in carnivore distribution that may result from recent outbreaks of mange and rabies. As a result, we are searching smaller map units (41 km<sup>2</sup>) for that study and incorporating a sampling schedule that resembles a robust mark-recapture design (Pollock 1982).

## MANAGEMENT IMPLICATIONS

Bayesian image models like ours are worthy of consideration for survey regions that can be readily partitioned into regular grids of appropriately sized map units; when a spatially contagious distribution is likely; when the species of interest can be reliably distinguished from similar species, so that spurious observations are unlikely; and when cumulative detection rates of ≈65% can be achieved. Advantages over alternative methods are likely to be greatest when a census of appropriately sized map units is impractical, detection is uncertain, and distributions do not correspond closely with distributions of covariates that can be measured easily. In such cases, image models can be used to resolve some of the most vexing problems that plague surveys of wildlife distribution. Benefits include substantial cost savings; greater precision than alternatives that do not exploit spatial relations among map units; better support for temporal comparisons when detection rates, or even methods of detection, may change over time; greater objectivity than methods that rely on subjective methods of boundary determination; and relatively weak assumptions. Although image models involve challenging concepts and can be difficult to implement, these benefits justify the necessary effort.

## ACKNOWLEDGMENTS

Funding for this study was provided by the U.S. Geological Survey Northern Prairie Wildlife Re-

search Center and the Kansas Department of Wildlife and Parks. We thank B. C. Lubow, L. D. Mech, M. W. Oehler, M. H. Sherfy, and 2 anonymous referees for constructive criticism of an early draft.

## LITERATURE CITED

- ALLARDYCE, D., AND M. A. SOVADA. 2003. A review of the ecology, distribution, and status of swift foxes in the United States. Pages 3–18 *in* M. A. Sovada and L. Carbyn, editors. The swift fox: ecology and conservation of swift foxes in a changing world. Canadian Plains Research Center, Regina, Saskatchewan, Canada.
- ANDERSON, D. R. 2001. The need to get the basics right in wildlife field studies. *Wildlife Society Bulletin* 29:1 294–1 297.
- AUGUSTIN, N. H., M. A. MUGGLESTON, AND S. T. BUCKLAND. 1996. An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology* 33:339–347.
- BESAG, J. 1974. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society Series B* 36:1 92–236.
- . 1975. Statistical analysis of non-lattice data. *Journal of the Royal Statistical Society Series D* 24:1 79–195.
- . 1986. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society Series B* 48:259–302.
- , P. GREEN, D. HIGDON, AND K. MENSERSEN. 1995. Bayesian computation in stochastic systems. *Statistical Science* 10:3–66.
- BIRCHAM, P. M. M., AND W. J. JORDAN. 1997. A consideration of some of the changes in distribution of “common birds” as revealed by The New Atlas of Breeding Birds in Britain and Ireland. *Ibis* 139:183–186.
- BUCKLAND, S. T., AND D. A. ELSTON. 1993. Empirical models for the spatial distribution of wildlife. *Journal of Applied Ecology* 30:478–495.
- CASELLA, G., AND E. I. GEORGE. 1992. Explaining the Gibbs sampler. *American Statistician* 46:1 67–174.
- CLAWSON, M. 1968. The land system of the United States: an introduction to the history and practice of land use and land tenure. University of Nebraska Press, Lincoln, USA.
- FULLER, T. K., W. E. BERG, G. L. RADDE, M. S. LENARZ, AND G. B. JOSELYN. 1992. A history and current estimate of wolf distribution and numbers in Minnesota. *Wildlife Society Bulletin* 20:42–55.
- GAMERMAN, D. 1997. Markov chain Monte Carlo: stochastic simulation for Bayesian inference. 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. 1996. Full conditional distributions. Pages 75–88 *in* W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. Markov chain Monte Carlo in practice. Chapman and Hall/CRC, Boca Raton, Florida, USA.
- , S. RICHARDSON, AND D. J. SPIEGELHALTER. 1996. Introducing Markov chain Monte Carlo. Pages 1–19 *in* W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. Markov chain Monte Carlo in practice. Chapman and Hall/CRC, Boca Raton, Florida, USA.
- GREEN, P. J. 1996. MCMC in image analysis. Pages 381–399 *in* W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, editors. Markov chain Monte Carlo in practice. Chap-

- man and Hall/CRC, Boca Raton, Florida, USA.
- GRIMMETT, G. R., AND D. R. STIRZAKER. 1994. Probability and random processes. Second edition. Oxford University Press, New York, USA.
- HARRISON, R. L., AND J. WHITAKER-HOAGLAND. 2003. A literature review of swift fox habitat and den-site selection. Pages 79–89 in M. A. Sovada and L. Carbyn, editors. The swift fox: ecology and conservation of swift foxes in a changing world. Canadian Plains Research Center, Regina, Saskatchewan, Canada.
- , AND C. G. SCHMITT. 2003. Current swift fox distribution and habitat selection within areas of historical occurrence in New Mexico. Pages 71–77 in M. A. Sovada and L. Carbyn, editors. The swift fox: ecology and conservation of swift foxes in a changing world. Canadian Plains Research Center, Regina, Saskatchewan, Canada.
- HEGLUND, P. J. 2002. Foundations of species-environment relations. Pages 35–41 in J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.
- HEIKKINEN, J., AND H. HÖGMANDER. 1994. Fully Bayesian approach to image restoration with an application in biogeography. *Journal of the Royal Statistical Society Series C* 43:569–582.
- HÖGMANDER, H., AND J. MØLLER. 1995. Estimating distribution maps from atlas data using methods of statistical image analysis. *Biometrics* 51:393–404.
- INSIGHTFUL CORPORATION. 2002. S-Plus 6.1 for Windows. Insightful Corporation, Seattle, Washington, USA.
- JOHNSON, D. H. 1999. The insignificance of statistical significance testing. *Journal of Wildlife Management* 63:763–772.
- , AND G. A. SARGEANT. 2002. Toward better atlases: improving presence–absence information. Pages 391–398 in J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.
- KÉRY, M. 2002. Inferring the absence of a species—a case study of snakes. *Journal of Wildlife Management* 66:330–338.
- KLUTE, D. S., M. J. LOVALLO, AND W. M. TZILKOWSKI. 2002. Autologistic regression modeling of American woodcock habitat use with spatially dependent data. Pages 335–343 in J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.
- LINK, W. A., E. CAM, J. D. NICHOLS, AND E. G. COOCH. 2002. Of BUGS and birds: Markov chain Monte Carlo for hierarchical modeling in wildlife research. *Journal of Wildlife Management* 66:277–291.
- MACKENZIE, D. I., J. D. NICHOLS, G. B. LACHMAN, S. DROEGE, J. A. ROYLE, AND C. A. LANGTIMM. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology* 83:2248–2255.
- MCARDLE, B. H. 1990. When are rare species not there? *Oikos* 57:276–277.
- O'CONNOR, R. J. 2002a. The conceptual basis of species distribution modeling: time for a paradigm shift? Pages 25–33 in J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.
- . 2002b. Gap conservation and science goals: rethinking the underlying biology. *Gap Analysis Program Bulletin* 11:2–6.
- OLSON, T. L., J. S. DIENI, F. G. LINDZEY, AND S. H. ANDERSON. 2003. Swift fox detection probability using tracking plate transects in southeast Wyoming. Pages 93–98 in M. A. Sovada and L. Carbyn, editors. The swift fox: ecology and conservation of swift foxes in a changing world. Canadian Plains Research Center, Regina, Saskatchewan, Canada.
- OSBORNE, P. E., AND B. J. TIGAR. 1992. Interpreting bird atlas data using logistic models: an example from Lesotho, southern Africa. *Journal of Applied Ecology* 29:55–62.
- POLLOCK, K. H. 1982. A capture–recapture design robust to unequal probability of capture. *Journal of Wildlife Management* 46:757–760.
- R DEVELOPMENT CORE TEAM. 2003. *R*: a language and environment for statistical computing. Foundation for Statistical Computing, Vienna, Austria.
- ROBERTSON, M. J., S. A. TEMPLE, AND J. COLEMAN. 2002. Predicting the distributions of songbirds in forests of central Wisconsin. Pages 399–410 in J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.
- ROEMER, G. W., T. J. COONAN, D. K. GARCELON, J. BASCOMPTE, AND L. LAUGHRIN. 2001. Feral pigs facilitate hyperpredation by golden eagles and indirectly cause the decline of the island fox. *Animal Conservation* 4:307–318.
- SARGEANT, A. B., R. J. GREENWOOD, M. A. SOVADA, AND T. L. SHAFFER. 1993. Distribution and abundance of predators that affect duck production—Prairie Pothole Region. U.S. Fish and Wildlife Resource Publication 194, Washington, D.C., USA.
- SARGEANT, G. A., P. J. WHITE, M. A. SOVADA, AND B. L. CYPHER. 2003. Scent-station survey techniques for swift and kit foxes. Pages 99–105 in M. A. Sovada and L. Carbyn, editors. The swift fox: ecology and conservation of swift foxes in a changing world. Canadian Plains Research Center, Regina, Saskatchewan, Canada.
- SCHAUSTER, E. R., E. M. GESE, AND A. M. KITCHEN. 2002a. Population ecology of swift foxes (*Vulpes velox*) in southeastern Colorado. *Canadian Journal of Zoology* 80:307–319.
- , ———, AND ———. 2002b. An evaluation of survey methods for monitoring swift fox abundance. *Wildlife Society Bulletin* 30:464–477.
- SHAUGHNESSY, M. J., JR. 2003. Swift fox detection methods and distribution in the Oklahoma Panhandle. Pages 61–69 in M. A. Sovada and L. Carbyn, editors. The swift fox: ecology and conservation of swift foxes in a changing world. Canadian Plains Research Center, Regina, Saskatchewan, Canada.
- SOVADA, M. A., C. C. ROY, J. B. BRIGHT, AND J. R. GILLIS. 1998. Causes and rates of mortality of swift foxes in western Kansas. *Journal of Wildlife Management* 62:1300–1306.
- , AND B. K. SCHEICK. 1999. Preliminary report to the swift fox conservation team: historic and recent distribution of swift foxes in North America. Pages 80–147 and appendix in C. G. Schmitt, editor. 1999

annual report of the Swift Fox Conservation Team. New Mexico Department of Game and Fish, Albuquerque, USA.

\_\_\_\_\_, C. C. SLIVINSKI, R. O. WOODWARD, AND M. L. PHILLIPS. 2003. Home range, habitat use, litter size, and pup dispersal of swift foxes in two distinct landscapes of western Kansas. Pages 149–160 in M. A. Sovada and L. Carbyn, editors. The swift fox: ecology and conservation of swift foxes in a changing world. Canadian Plains Research Center, Regina, Saskatchewan, Canada.

STAUFFER, D. F. 2002. Linking populations and habitats: Where have we been? Where are we going? Pages 53–61 in J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Hautler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Washington, D.C., USA.

U.S. FISH AND WILDLIFE SERVICE. 1995. Twelve-month administrative finding on petition to list the swift fox. Region 6 Office, Denver, Colorado, USA.

ZUMBAUGH, D. M., AND J. R. CHOATE. 1985. Historical biogeography of foxes in Kansas. Transactions of the Kansas Academy of Sciences 88:1–13.

Associate Editor: Lubow.

**APPENDIX**

**Derivation of the joint likelihood for  $\mathbf{x}$ ,  $\beta$ , and  $\theta$**

- Step 1: Expand the joint density of  $\mathbf{x}$ ,  $\beta$ ,  $\theta$ , and  $\mathbf{v}$  in 2 ways:  
 (a)  $\Pr(\mathbf{x}, \beta, \theta, \mathbf{v}, \mathbf{y}) = \Pr(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y}) \times \Pr(\mathbf{v}, \mathbf{y})$ , and  
 (b)  $\Pr(\mathbf{x}, \beta, \theta, \mathbf{v}, \mathbf{y}) = \Pr(\mathbf{v}, \mathbf{y} | \mathbf{x}, \beta, \theta) \times \Pr(\mathbf{x} | \beta, \theta) \times \Pr(\theta | \beta) \times \Pr(\beta)$
- Step 2: Equate results (a) and (b) of Step 1:  
 $\Pr(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y}) \times \Pr(\mathbf{v}, \mathbf{y}) = \Pr(\mathbf{v}, \mathbf{y} | \mathbf{x}, \beta, \theta) \times \Pr(\mathbf{x} | \beta, \theta) \times \Pr(\theta | \beta) \times \Pr(\beta)$
- Step 3: Rearrange the results of Step 2:  

$$\Pr(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y}) = \frac{\Pr(\mathbf{v}, \mathbf{y} | \mathbf{x}, \beta, \theta) \times \Pr(\mathbf{x} | \beta, \theta) \times \Pr(\theta | \beta) \times \Pr(\beta)}{\Pr(\mathbf{v}, \mathbf{y})}$$
- Step 4: Spatial patterns in observations arise from patterns in occurrences and thus depend on  $\beta$  only through  $\mathbf{x}$ . Conversely,  $\theta$  affects observations but not actual occurrences. As a result, the likelihood for  $\mathbf{v}$  does not involve  $\beta$ , the prior for  $\mathbf{x}$  does not involve  $\theta$ , and the prior for  $\theta$  does not involve  $\beta$ .  

$$\Pr(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y}) = \frac{\Pr(\mathbf{v}, \mathbf{y} | \mathbf{x}, \theta) \times \Pr(\mathbf{x} | \beta) \times \Pr(\theta) \times \Pr(\beta)}{\Pr(\mathbf{v}, \mathbf{y})}$$
- Step 5:  $\Pr(\mathbf{v}, \mathbf{y})$  does not involve  $\mathbf{x}$ ,  $\beta$ , or  $\theta$ ; therefore, following notation in the text:  
 $\pi(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y}) \propto \ell(\mathbf{v}, \mathbf{y} | \mathbf{x}, \theta) \times f(\mathbf{x} | \beta) \times f(\theta) \times f(\beta)$ .

**Derivation of the full conditional distribution for  $x_j$**

- Step 1: Start with  $\pi(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y}) \propto \ell(\mathbf{v}, \mathbf{y} | \mathbf{x}, \theta) \times f(\mathbf{x} | \beta) \times f(\theta) \times f(\beta)$ , from above.
- Step 2: Condition on current estimates of  $\theta$  and  $\beta$ :  
 $\pi(\mathbf{x} | \mathbf{v}, \mathbf{y}, \beta, \theta) \propto \ell(\mathbf{v}, \mathbf{y} | \mathbf{x}, \theta) \times f(\mathbf{x} | \beta)$ .
- Step 3: Condition on the current estimate of  $\mathbf{x}_{-j}$ . Note also that observations are independent, conditional on the  $\mathbf{x}_p$  and that detections are conclusive.  
 $\pi(x_j | \mathbf{x}_{-j}, \mathbf{v}, \beta, \theta) \propto \ell(v_j | x_j, \theta) \times f(x_j | \mathbf{x}_{-j}, \beta)$  for  $y_j = 0$
- Step 4: Recall that  $\mathbf{x}$  is a Markov random field defined in terms of second-order neighborhoods, hence that  
 $\Pr(x_j | \mathbf{x}_{-j}) = \Pr(x_j | \mathbf{x}_{-j})$ .  
 $\pi(x_j | \mathbf{x}_{-j}, \mathbf{v}, \beta, \theta) \propto \ell(v_j | x_j, \theta) \times f(x_j | \mathbf{x}_{-j}, \beta)$  for  $y_j = 0$

**Derivation of the full conditional distribution for  $\beta$**

- Step 1: Start with  $\pi(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y}) \propto \ell(\mathbf{v}, \mathbf{y} | \mathbf{x}, \theta) \times f(\mathbf{x} | \beta) \times f(\theta) \times f(\beta)$ , from above.
- Step 2: Condition on the current estimates of  $\mathbf{x}$  and  $\theta$ :  
 $\pi(\beta | \mathbf{x}, \theta, \mathbf{v}, \mathbf{y}) \propto \ell(\mathbf{v}, \mathbf{y} | \mathbf{x}, \theta) \times f(\mathbf{x} | \beta) \times f(\beta)$
- Step 3: Eliminate terms that do not involve  $\beta$ :  
 $\pi(\beta | \mathbf{x}) \propto f(\mathbf{x} | \beta) \times f(\beta)$
- Step 4: Substitute a pseudolikelihood approximation for  $f(\mathbf{x} | \beta)$  (Besag 1975, Heikkinen and Högmänder 1994):  
 $\pi(\beta | \mathbf{x}) \propto \Pi f(x_j | \beta) \times f(\beta)$  (approximate).

**Derivation of the full conditional distribution for  $\theta$**

- Step 1: Start with  $\pi(\mathbf{x}, \beta, \theta | \mathbf{v}, \mathbf{y}) \propto \ell(\mathbf{v}, \mathbf{y} | \mathbf{x}, \theta) \times f(\mathbf{x} | \beta) \times f(\theta) \times f(\beta)$ , from above.
- Step 2: Condition on the current estimates of  $\mathbf{x}$  and  $\beta$ :  
 $\pi(\theta | \mathbf{x}, \beta, \mathbf{v}, \mathbf{y}) \propto \ell(\mathbf{v}, \mathbf{y} | \mathbf{x}, \theta) \times f(\mathbf{x} | \beta) \times f(\theta)$
- Step 3: Eliminate terms that do not involve  $\theta$ :  
 $\pi(\theta | \mathbf{x}, \mathbf{v}, \mathbf{y}) \propto \ell(\mathbf{v}, \mathbf{y} | \mathbf{x}, \theta) \times f(\theta)$