

---

## UNMARKED POPULATIONS

Traditional capture-recapture models share the fundamental assumption that each individual in a population can be uniquely identified when captured. Often, this can be accomplished by marking individuals with color bands, ear tags, or some other artificial mark that can be subsequently read in the field. For other species, such as tigers (*Panthera tigris*) or marbled salamanders (*Ambystoma opacum*), individuals can be easily identified using only their natural markings. However, many species do not possess adequate natural markings and are difficult to capture, making it impractical to use standard capture-recapture techniques.

Estimating density when individuals are unmarked can be accomplished using a variety of alternatives to capture-recapture, such as distance sampling (Buckland et al., 2001) and  $N$ -mixture models (Royle, 2004b). These methods, among others, can be very effective when their assumptions are met, but in cases such as when it is not possible to obtain accurate distance data, or when movement complicates the use of fixed-area plots, these methods may not yield unbiased estimates of density (Chandler et al., 2011). In this chapter we highlight the work of Chandler and Royle (In press) who demonstrated that the “individual recognition” assumption of traditional capture-recapture models is not a requirement of spatial capture-recapture models. They showed that, under certain conditions described below, spatially-correlated count data are sufficient for making inference about animal distribution and density even when no individuals are marked. The Chandler and Royle (In press) “spatial count model” (hereafter the SC model) requires neither distance data nor fixed area plots. Instead, the observed data are trap- and occasion-specific counts, which are modeled as a reduced-information summary of the latent encounter histories. Because the model is formulated in terms of the data we wish we had, i.e. the typical encounter history data observed in standard capture-recapture studies of marked animals, the SC model is just a SCR model with a single extension to account for the fact that the encounter history data are unobserved.

The ability to fit SCR models to data from unmarked populations has important implications. For one, it means that SCR models can be applied to data collected using methods like point counts in which observers record simple counts of animals at an array of survey locations. The model can also be fit to camera trapping data collected on

unmarked animals, representing the first formal method for estimating density from such data. So, is this model a free lunch? At face value, it sounds as though it allows for estimation of all the quantities of interest in standard capture-recapture studies, but with very little data. But of course the answer is no – lunch is still not free because with this model come new assumptions, and as was demonstrated by Chandler and Royle (In press), even with “perfect” data, the estimates will typically not be very precise. This should not be surprising given that we are asking so much from simple count data.

The real value of the SC model is two-fold. First, it demonstrates an important theoretical result, namely that spatial correlation in count data carries information about the distribution of individuals. This stands in stark contrast to a prevailing view of spatial correlation as a nuisance to be avoided or modeled out of unsightly residual plots. The second reason why this model is important is that it provides the basis for handling the extremely common phenomenon in which only a subset of individuals in a population can be marked or otherwise distinguishable. Thus, while we do not recommend foregoing the work required to mark animals, this model does provide a method for studying unmarked populations, which *can* yield precise density estimates if some of the individuals are marked, or if prior information about some of the parameters is available. The extension of this model to handle data from marked and unmarked individuals is thoroughly treated in the next chapter. Here, we focus on the case in which all individuals are unmarked. We begin by presenting the basic formulation of the model, and then we proceed by way of a few examples to consider extensions of the model in which ancillary information can be used to increase precision.

## 18.1 EXISTING MODELS FOR INFERENCE ABOUT DENSITY IN UNMARKED POPULATIONS

When capture-recapture methods are not a viable option, ecologists often collect simple count data or even binary detection/non-detection data to estimate parameters such as abundance or occupancy. These data may be analyzed using generalized linear models such as Poisson regression or logistic regression, perhaps with random effects (Zuur et al., 2009). However, these methods will be biased when detection is imperfect, as it usually is. Even when count data or detection/non-detection data are used as an index of abundance or occurrence, standard models may yield unreliable results when covariates affect both the ecological process and the observation process. A classic example is given by Bibby and Buckland (1987) who found that songbird detection probability was negatively related to vegetation height, whereas density was positively associated with vegetation height in restocked conifer plantations. This intuitive phenomenon has been demonstrated repeatedly (e.g. Kéry, 2008; Sillett et al., 2012) and has led to the development of a vast number of models to estimate population size and detection probability when individuals are unmarked. A review of these models is beyond the scope of this chapter, but we mention a few deficiencies of existing methods that warrant the exploration of alternatives for robust inference when standard capture-recapture methods do not apply.

Distance sampling (Buckland et al., 2001), which we briefly introduced in Chapter 1, is perhaps the most widely used method for estimating population density when individuals are unmarked and detection probability is less than one. This class of methods is known to work impeccably when estimating the number of stakes in a field or the number of

15056 duck nests in a wetland. Distance sampling can also work very well in more interesting  
 15057 situations, and it is an extremely powerful method when the assumptions can be met.  
 15058 However, the assumptions that distance data can be recorded without error and that  
 15059 animals are distributed randomly with respect to the transect can be easily violated by  
 15060 common processes such as animal movement and measurement error. Although numerous  
 15061 methods have been proposed to relax some of these assumptions (Royle et al. (2004);  
 15062 Borchers et al. (1998); Johnson (2010); Chandler et al. (2011)), a more important issue  
 15063 is that distance sampling is simply not practical in many settings. For example, many  
 15064 species are so rare and elusive that they can only be reliably surveyed using methods such  
 15065 as camera traps or hair snares.

15066 Other common approaches to estimate density when individuals are unmarked include  
 15067 double-observer sampling, removal sampling, and repeated counts, for which custom mod-  
 15068 els have been developed (Nichols et al., 2000b; Farnsworth et al., 2002; Royle, 2004b,a;  
 15069 Nichols et al., 2009; Fiske and Chandler, 2011). To obtain reliable density estimates using  
 15070 these methods, the area surveyed must be well defined and closed with respect to move-  
 15071 ment and demographic processes. Given a sufficiently short sampling interval, such as a  
 15072 5-min point-count, the closure assumption may be reasonable. However, short sampling  
 15073 intervals limit the number of detections, so observers generally visit each survey location  
 15074 multiple times during a season. But then, animal movement may invalidate the closure  
 15075 assumption, and a model of temporary emigration is required (Kendall et al., 1997; Chan-  
 15076 dler et al., 2011). Furthermore, distance-related heterogeneity in detection probability can  
 15077 introduce bias in these models, although this bias is negligible when the ratio of plot size  
 15078 to the scale parameter of the detection function is low (Efford and Dawson, 2009).

15079 We mention these issues not to suggest that existing models do not have value – indeed  
 15080 we believe that they can be used to obtain reliable density estimates in many situations –  
 15081 rather, our aim is to highlight the need for alternative methods when the assumptions of  
 15082 existing methods cannot be met and when spatially-explicit inference is the objective.

## 18.2 SPATIAL CORRELATION IN COUNT DATA

### 18.2.1 Spatial correlation as information

15083 All of the previous methods require some sort of auxiliary information to model both  
 15084 abundance and detection. For instance, multiple observers or distance data or repeated  
 15085 visits may be required to ensure that model parameters are identifiable (but see (Lele  
 15086 et al., 2012; Sólomos et al., 2012)). The same is true for the SC model, but the auxiliary  
 15087 information comes in the form of spatial correlation, which requires no extra effort to  
 15088 collect (Chandler and Royle, In press).

15090 It is natural to be suspicious of the claim that spatial correlation is a good thing. In  
 15091 fact elaborate methods have been devised to deal with spatial correlation as a nuisance  
 15092 (Lichstein et al., 2002; Dormann et al., 2007), and ecologists have been admonished for  
 15093 failing to obtain “real” replicates uncontaminated by spatial correlation (Hurlbert, 1984).  
 15094 The following heuristic may be helpful for seeing the value of spatial correlation.

15095 Imagine a  $10 \times 10$  grid of camera traps and a single unmarked individual exposed to  
 15096 “capture” whose home range center lies in the center of the trapping grid. If the individual  
 15097 has a small home range size relative to the extent of the trapping grid, we can imagine what  
 15098 the spatial correlation structure of the encounters might look like. If the animal’s home

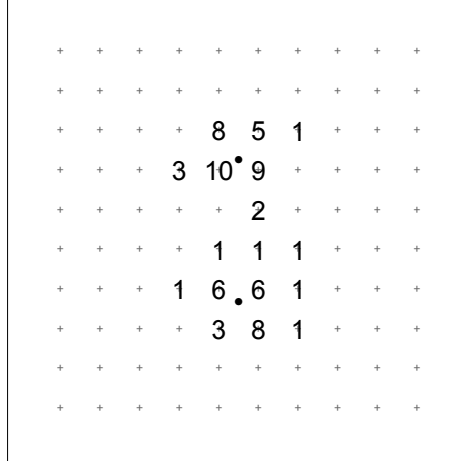
range is symmetric around the activity center then the number of times the individual is detected at each trap (the trap count) is a function of the distance between the home range center and the trap; i.e., traps with the same distance from the activity center will yield counts that are more highly correlated with one another than traps located at different distances from the activity center. Thus, the correlation in counts tells us something about the location of the activity center. It is relatively intuitive that spatial correlation carries information about distribution, but what about density?

Imagine now that there are **two activity centers** located in the trapping grid. Using trap counts alone, is it possible to determine the number and location of these activity centers? The answer is yes, at least under certain circumstances. Figure 18.1 shows the locations of the two hypothetical activity centers, and the total counts made at each trap after 10 survey occasions. Assuming that animals have bivariate normal home ranges, the fact that there are two areas in the map with high counts that dissipate with distance suggests that the most likely number of individuals given these data is 2. Furthermore, the degree to which the counts dissipate from the two areas of highest intensity is information about the home range size parameter  $\sigma$ . These two pieces of information are enough to estimate the number of individuals exposed to sampling – again, given that a bivariate normal home range is a valid assumption. **Of course, the data could just as well have been generated by a single individual whose home range is distinctly bimodal**, and thus *always* the assumptions of our model need to be carefully examined using our biological knowledge of the system. If the assumptions do not hold, it is almost always possible to relax them, for instance by allowing for non-stationary home ranges as we demonstrated in Chapt. 12 and 13.

## 18.2.2 Two types of spatial correlation

The spatial correlation ~~dealt with~~ by the SC model is assumed to arise from animal movement; however, this is just one type of spatial correlation that may exist in ecological count data. Another common type of spatial correlation results from the spatial correlation of environmental covariates. Habitat variables, such as the percent cover of deciduous forest in North America, will often be patchy rather than randomly distributed, and this can result in spatial correlation in abundance, and hence in count data. Often, this type of spatial correlation can be dealt with by simply including the **habitat covariate** in the model. For example, a simple species distribution model with only a few habitat variables can result in a distribution map that reflects the spatial correlation in abundance (Sillett et al., 2012; Royle et al., 2012b). In such a case, there is no need to use spatially-explicit models (Besag and Kooperberg, 1995; Lichstein et al., 2002; Wikle, 2010). The reason is that the relevant assumption of non-spatial models (e.g. GLMs) is that no spatial correlation exists in the *residuals*, and often, any apparent spatial correlation can be accounted for using covariates. This may be obvious, but it is a point that seems to be frequently misunderstood.

When it does become important to account for environmentally-induced spatial correlation is in situations where correlation exists in the residuals *after* accounting for covariate effects. This may be due to unobserved covariates, and Zuur et al. (2009) offer advice on how to check for and deal with such correlation in the context of GLMs and GLMMs. SCR models, including the SC model dealt with in this chapter, explicitly ac-



**Figure 18.1.** Simulated count data at each of 100 camera traps (crosses) after  $K = 10$  sampling occasions. The black dots are the locations of two animal activity centers. The spatial count model estimates both the location and number of activity centers exposed to sampling using such spatially-referenced count data.

count for movement-induced spatial correlation, and they can also be used to account for environmentally-induced spatial correlation by adopting an inhomogeneous point process model for the activity centers. That is, the point process intensity can be modeled as a function of observed covariates, and theoretically, it should be possible to allow for spatially-correlated random effects to deal with unobserved covariates. See Chapt. ?? for details.

## 18.3 SPATIAL COUNT MODEL

### 18.3.1 Data

Whereas traditional SCR models require spatially-referenced encounter histories, this model requires simple count data. Let  $n_{jk}$  be the count data at sampling location  $j$  on occasion  $k$ . The  $J \times K$  matrix of counts will be denoted  $\mathbf{n}$ . A sampling location in this context could be any device capable of recording count data, such as a human observer or a camera trap, and one of the benefits of the SC model is that it can be applied to data collected using many different survey methods. For ease of presentation, we will refer to sampling devices as traps, but remember that a trap is just something capable of recording count data. As in all SCR models, we also require the coordinates of the  $J$  traps, and we denote the location of trap  $j$  by  $\mathbf{x}_j$ . In some instances, additional data might be available such as trap-specific covariates, state-space covariates, information on the identities of a subset of individuals, or perhaps even distance data. We consider some of these model extensions in Sec. 18.8, but for the time being we ignore these possibilities so that we can

15162 focus on the basic model.

### 15163 18.3.2 Model

15164 The state model is exactly the same as the one we have dealt with throughout this book.  
 15165 It is a point process describing the number and distribution of activity centers in the  
 15166 state-space  $\mathcal{S}$ . Although it might be possible to fit inhomogeneous point process models  
 15167 using the methods described in Chapt. 11, given the simplicity of the data, we concentrate  
 15168 on a homogeneous point process  $\{\mathbf{s}_i, \dots, \mathbf{s}_N\} \sim \text{Uniform}(\mathcal{S})$  where  $\mathbf{s}_i$  is the activity center  
 15169 of individual  $i$  in the population of size  $N$ . For the moment, we will assume that  $N$  is  
 15170 known.

15171 The observation model is the same as in other SCR models in the sense that it de-  
 15172 scribes the probability of encountering individual  $i$  at trap  $j$ , conditional on the location  
 15173 of the individual's activity center. The specific encounter process will depend on the sam-  
 15174 pling method, and here we consider the standard camera trapping situation in which an  
 15175 individual can be encountered at multiple traps during a single time period, say one night  
 15176 during a camera-trapping study, and it can be detected multiple times at a single trap  
 15177 during an occasion. This is the Poisson encounter model (a.k.a. the proximity detector  
 15178 case) described in Chapt. 9. Our experience with alternative observation models such as  
 15179 the Bernoulli and multinomial models suggests that the parameters of the model may not  
 15180 be identifiable in these cases, at least when no additional information is available. This is  
 15181 a subject of ongoing research.

15182 As before, we define  $y_{ijk}$  as the encounter data for individual  $i$  at trap  $j$  on occasion  
 15183  $k$ , which we model as:

$$y_{ijk} \sim \text{Poisson}(\lambda_{ij}) \quad (18.3.1)$$

15184 where  $\lambda_{ij}$  is the encounter rate. A common encounter rate model is the Gaussian, or  
 15185 half-normal, model:

$$\lambda_{ij} = \lambda_0 \exp(-\|\mathbf{x}_j - \mathbf{s}_i\|/2\sigma^2)$$

15186 in which  $\lambda_0$  is the baseline encounter rate,  $\|\mathbf{x}_j - \mathbf{s}_i\|$  is the Euclidean distance between the  
 15187 trap and activity center, and  $\sigma$  is the scale parameter determining the degree to which  
 15188 encounter rate decreases with distance. In this context,  $\sigma$  also determines the amount of  
 15189 correlation among the counts because if  $\sigma$  is low relative to the trap spacing, then it is  
 15190 unlikely that an individual will be detected at multiple traps.

15191 When individuals cannot be uniquely identified, the encounter histories cannot be  
 15192 directly observed, which seems like a massively insurmountable problem. The solution of  
 15193 Chandler and Royle (In press) is the same one we routinely apply when we cannot directly  
 15194 observe the process of interest – we regard the encounter histories as latent variables. This  
 15195 leaves the remaining task of specifying the relationship between the count data and the  
 15196 encounter histories, i.e. we need a model of  $[\mathbf{n}|\mathbf{y}]$  where  $\mathbf{y}$  represents the entire collection  
 15197 of encounter histories. In this case, there is only one possibility because, by definition, the  
 15198 count data are simply a reduced-information summary of the latent encounter histories.  
 15199 That is, they are the sample- and trap-specific totals, aggregated over all individuals:

$$n_{jk} = \sum_{i=1}^N y_{ijk}. \quad (18.3.2)$$

15200 So, unlike most model-development problems faced in this book, we don't have to con-  
 15201 sider competing probability models for  $[\mathbf{n}|\mathbf{y}]$ , but instead, we recognize the fact that the  
 15202 relationship between the counts and the latent encounter histories is deterministic. This  
 15203 deterministic constraint poses some computational challenges, which we discuss below.  
 15204 But first we present some alternative formulations of the model.

15205 Recall from Chapt. 2 that the sum of two or more Poisson random variables is also  
 15206 a Poisson random variable. Specifically, if  $x_1 \sim \text{Poisson}(\lambda_1)$  and  $x_2 \sim \text{Poisson}(\lambda_2)$ , then  
 15207  $(x_1 + x_2) \sim \text{Poisson}(\lambda_1 + \lambda_2)$ . Thus, under this Poisson model for the latent encounter  
 15208 histories, the count data can be modeled as Poisson:

$$n_{jk} \sim \text{Poisson}(\Lambda_j) \quad (18.3.3)$$

15209 where

$$\Lambda_j = \lambda_0 \sum_i \exp(-\|\mathbf{x}_j - \mathbf{s}_i\|/2\sigma^2),$$


15210 and because  $\Lambda_j$  does not depend on  $k$ , we can aggregate the replicated counts, defining  
 15211  $n_{j\cdot} = \sum_k n_{jk}$  and then

$$n_{j\cdot} \sim \text{Poisson}(K\Lambda_j).$$

15212 As such,  $K$  and  $\lambda_0$  serve equivalent roles as affecting baseline encounter rate. Formulating  
 15213 the model in terms of the aggregated count data demonstrates that the model can be  
 15214 applied to data from a single sampling occasion ( $J \equiv 1$ ) as has been noted elsewhere for  
 15215 standard SCR models (Efford et al., 2009b). In the context of studying marked popula-  
 15216 tions, the model parameters will only be identifiable in the  $J \equiv 1$  case if an animal can be  
 15217 captured at multiple traps during a single occasion. The SC model essentially requires the  
 15218 same thing, which is to say that it requires correlation in the count data resulting from  
 15219 an individual being captured in multiple, closely-spaced traps.

15220 This formulation of the model in terms of the aggregate count also simplifies computa-  
 15221 tions as the latent encounter histories do not need to be updated in the MCMC estimation  
 15222 scheme; however, retaining them in the formulation of the model is important if some in-  
 15223 dividuals are uniquely marked. This is because uniquely identifiable individuals produce  
 15224 observations of some of the  $y_{ijk}$  variables, which we elaborate on in the subsequent chapter.

## 18.4 HOW MUCH CORRELATION IS ENOUGH?

15225  $\sigma$  shouldn't be too small or too large relative to trap spacing or else the counts will be  
 15226 i.i.d. Poisson random variables. So how much correlation is enough? Phrased differently,  
 15227 what is the ideal ratio of  $\sigma$  to trap spacing to ensure correlation and minimize the variance  
 15228 of the posterior distributions. We see two options for answering this questions, both of  
 15229 which are topics in need of additional research. The first approach is to use the methods  
 15230 described in Chapt. 10, i.e. by either conducting simulation studies with various trap  
 15231 spacing to  $\sigma$  ratios, or to analytically minimize a variance criterion for a given set of  
 15232 sampling conditions and effort. The former approach was used by Chandler and Royle  
 15233 (In press) whose limited simulation study indicated that an ideal ratio is approximately 2  
 15234  This agrees with findings from previous research on the optimal design of SCR studies  
 15235 (Chapt. 10), as it should.

15236 A second approach that may be of use if a data set has already been collected is to  
 15237 use standard techniques from spatial statistics to determine if adequate correlation exists

in the counts. For example, one might compute Ripley's  $K$ -statistic or generate (semi-)variograms (Illian et al., 2008). We have not studied the utility of such approaches, but it seems worthy of investigation.

#### 18.4.1 On $N$ being unknown

Population size,  $N$ , is never known in practice, and thus we need a model for it. For homogeneous point process models,  $N$  is typically modeled as  $N \sim \text{Poisson}(\mu|\mathcal{S}|)$  or  $N \sim \text{Binomial}(M, \psi)$ , the latter of which is equivalent to a discrete uniform prior if  $\psi \sim \text{Unif}(0, 1)$ . In Chapt. 11 and elsewhere, we demonstrated that the choice of prior has very little influence on parameter estimates, and so we favor the binomial prior because of its convenience when using MCMC, i.e. it allows us to fix the dimensions of the parameter space by setting  $M$  to some arbitrarily large integer.

A binomial model is equivalent to a series of  $M$  independent Bernoulli trials, hence we can rewrite  $N \sim \text{Binomial}(M, \psi)$  as  $z_i \sim \text{Bernoulli}(\psi)$  where  $z_i$  is an auxiliary variable indicating if individual  $i$  is a member of the population, i.e.  $N = \sum_{i=1}^M z_i$ . Having expanded the model to include a prior on  $N$ , we can summarize the SC model, with a Gaussian observation model, as follows:

$$\begin{aligned} z_i &\sim \text{Bern}(\psi) \\ y_{ijk} &\sim \text{Poisson}(\lambda_{ijk} z_i) \\ \lambda_{ijk} &= \lambda_0 \exp(-\|\mathbf{x}_j - \mathbf{s}_i\|^2 / (2\sigma^2)) \\ n_{jk} &= \sum_{i=1}^M y_{ijk} \end{aligned}$$

Bayesian analysis can proceed once suitable priors have been put on the hyperparameters  $\psi$ ,  $\sigma$ , and  $\lambda_0$ . Chandler and Royle (In press) provided **R** code for fitting the model using MCMC, and they evaluated the model's performance with uniform priors on the three hyperparameters. They also discussed the possibilities and effects of including prior knowledge about  $\sigma$  into the model. In the next section, we explain how the model can be implemented using **JAGS**, but first we briefly contemplate the viability of classical analysis of this model.

The obvious challenge faced when conducting a classical analysis of this model is that the number of latent variables is huge. In all SCR models, the activity centers are latent, but now, even the encounter histories are latent. Maximizing likelihoods with latent variables (random effects) involves integrating (or summing) over all possible values of the latent variables. For the activity centers, this is typically accomplished by integrating the conditional-on- $\mathbf{s}$  likelihood  $[\mathbf{y}_i | \mathbf{s}_i]$  over the two-dimensional state-space  $\mathcal{S}$  (Chapt. 6). However, with the SC model, we have to sum over all possible encounter histories meeting the constraint of Eq. 18.3.2. The number of possible encounter histories will, in general, be too high to make the likelihood tractable, and thus we do not think that maximum likelihood is a viable option for analyzing this model. However, one might be able to obtain approximate maximum likelihood estimates using simulation-based methods (Lele et al., 2010), which will typically be more computationally challenging than the Bayesian analysis.



## 18.5 SIMULATION EXAMPLE

15269 Simulating data under the SC model proceeds by first simulating standard SCR encounter  
 15270 history data and then collapsing it into count data. The following blocks of **R** code  
 15271 generate data from the model shown in Sec. 18.4.1, with parameters  $\sigma = 5$ ,  $\lambda_0 = 0.4$ , and  
 15272  $N = 50$ . The state-space is a  $[0, 100] \times [0, 100]$  square, and a grid of 100 traps is centered  
 15273 in the middle. These simulated data resemble actual data from camera trap studies in  
 15274 which individuals can be detected multiple times at a trap during a single occasion, and at  
 15275 multiple traps during an occasion. The first block of code generates the trap coordinates  
 15276  $X$  and the  $N = 50$  activity centers:

```
15277 > tr <- seq(15, 85, length=10)
15278 > X <- cbind(rep(tr, each=length(tr)),
15279 +           rep(tr, times=length(tr))) # 100 trap coords
15280 > set.seed(10)
15281 > xlim <- c(0, 100); ylim <- c(0, 100) # S is [0,100]x[0,100] square
15282 > A <- (xlim[2]-xlim[1])*(ylim[2]-ylim[1])/1e4 # area of S
15283 > mu <- 50 # density (animals/unit area)
15284 > (N <- rpois(1, mu*A)) # Generate N=50 as Poisson deviate
15285 [1] 50
15286 > s <- cbind(runif(N, xlim[1], xlim[2]), runif(N, ylim[1], ylim[2]))
```

15287 We could have set  $N = 50$  directly, but instead we treated density as a fixed parameter  
 15288 ( $\mu = 50$ ) and generated  $N$  as a random variable – it just so happens that with the specified  
 15289 random seed,  $N$  equals 50.

15290 Now we can generate the encounter histories under the Poisson observation model.  
 15291 Let's suppose that sampling is conducted over  $K = 5$  nights.

```
15292 > sigma <- 5
15293 > lam0 <- 0.4
15294 > J <- nrow(X)
15295 > K <- 5
15296 > y <- array(NA, c(N, J, K))
15297 > for(j in 1:J) {
15298 +   dist <- sqrt((X[j,1]-s[,1])^2 + (X[j,2] - s[,2])^2)
15299 +   lambda <- lam0*exp(-dist^2/(2*sigma^2))
15300 +   for(k in 1:K) {
15301 +     y[,j,k] <- rpois(N, lambda)
15302 +   }
15303 + }
```

15304 The object  $y$  is the  $N \times J \times K$  array of encounter data, which cannot be directly ob-  
 15305 served if the animals are unmarked. Converting the encounter data to count data can be  
 15306 accomplished using a single **apply** command.

```
15307 > n <- apply(y, c(2,3), sum)
15308 > dimnames(n) <- list(paste("trap", 1:J, sep=""),
15309 +                     paste("night", 1:K, sep=""))
15310 > n[1:4,]
```

```

15311      night1 night2 night3 night4 night5
15312 trap1      1      0      0      0      0
15313 trap2      1      2      2      0      1
15314 trap3      1      0      0      1      0
15315 trap4      0      0      0      0      0

```

15316 This displays the first 4 rows of  $\mathbf{n}$ , the  $J \times K$  matrix of counts. It is worth contemplating  
 15317 how common such count data is in ecology and how many different mechanisms might  
 15318 generate it. Although the list of possibilities is immense, the SC model has advantages over  
 15319 some alternatives in that it includes an explicit model for the distribution of individuals in  
 15320 space *and* it includes a model describing how detections are generated given the distance  
 15321 between traps and individual activity centers. It also provides a foundation for extending  
 15322 the model in many ways as we discuss in Sec. 18.8 and in the next chapter.

15323 The question now is: Is it possible to estimate the parameters? In our simulated  
 15324 dataset we have  $J \times K = 500$  data points, but how many parameters do we need to  
 15325 estimate with this rather small set of data? A frequentist might say that there are only  
 15326 3 parameters:  $\lambda_0$ ,  $\sigma$ , and  $N$  (or density  $\mu$ ) because inference about the latent parameters  
 15327 is carried out using prediction methods after the 3 hyperparameters have been estimated.  
 15328 However, a Bayesian would probably say that each  $\mathbf{s}$  and each element of the latent  
 15329 encounter array  $\mathbf{y}$  is a parameter in need of a posterior. From this perspective there are  
 15330 far more parameters than data points, and thus it would appear as though the situation  
 15331 is dire. Whether or not the parameters are actually estimable is a rather difficult question  
 15332 to answer. One simplistic, but not definitive, approach for addressing the question is  
 15333 to conduct a simulation study and evaluate the frequentist performance of the model by  
 15334 asking how often the data-generating values are included in confidence/credible intervals,  
 15335 and how biased are point estimates. Chandler and Royle (In press) conducted such a  
 15336 simulation study and found that, while the variance of the posterior distributions was  
 15337 high by most standards, the bias of the posterior mode of  $N$  was small and the coverage  
 15338 of the credible intervals was close to nominal. Moreover, they found no evidence that the  
 15339 posterior distributions were dominated by the priors, further supporting the conclusion  
 15340 that spatial correlation in the count data is sufficient for estimating density and encounter  
 15341 probability parameters. However, in such cases where identifiability has not formally  
 15342 demonstrated, it may be wise to compare the results of models fit using both proper and  
 15343 improper priors, as we do below.

15344 At this point in time the SC model can only be fit using one of the **BUGS** en-  
 15345 gines, or using custom software like the **R** code accompanying Chandler and Royle (In  
 15346 press). Although **BUGS** might provide the most flexible option for fitting the model,  
 15347 it is not straight-forward because of the constraints in the model. In **WinBUGS**, the  
 15348  $n_{jk} = \sum_i y_{ijk}$  constraint can be enforced using the so-called “ones-trick”, but we prefer  
 15349 **JAGS** because it has a distribution called **dsum** that was designed for this type of situ-  
 15350 ation in which the observed data are a sum of random variables. Panel 18.1 shows the  
 15351 **JAGS** code, but we abbreviated the arguments to **dsum** because in practice you need to  
 15352 provide all  $M$  of them. The code looks slightly unwieldy if  $M$  is large, but you can easily  
 15353 create it using the **paste** function in **R**. Here is an example, with an unrealistically small  
 15354 value of  $M = 10$ :

```

15355 > paste("y[, 1:10, ", j,k]", sep="", collapse=", ")

```

---

```

15356 [1] "y[1,j,k], y[2,j,k], y[3,j,k], y[4,j,k], y[5,j,k], y[6,j,k],
15357 y[7,j,k], y[8,j,k], y[9,j,k], y[10,j,k]"

model{
  sigma ~ dunif(0, 200) # Tailor this to your state-space
  lam0 ~ dunif(0, 5)    # consider dgamma() as an alternative
  psi ~ dbeta(1,1)
  for(i in 1:M) {
    z[i] ~ dbern(psi)
    s[i,1] ~ dunif(xlim[1], xlim[2])
    s[i,2] ~ dunif(ylim[1], ylim[2])
    for(j in 1:J) { # Number of traps
      distsq[i,j] <- (s[i,1] - X[j,1])^2 + (s[i,2] - X[j,2])^2
      lam[i,j] <- lam0 * exp(-distsq[i,j] / (2*sigma^2))
      for(k in 1:K) { # Number of occasions
        y[i,j,k] ~ dpois(lam[i,j]*z[i])
      }
    }
  }
  for(j in 1:J) {
    for(k in 1:K) {
      n[j,k] ~ dsum(y[1,j,k], y[2,j,k], ..., y[200,j,k]) # Code abbreviated!!
    }
  }
  N <- sum(z[]) # Realized population size
  A <- (xlim[2]-xlim[1])*(ylim[2]-ylim[1]) # Area of state-space
  D <- N / A # Realized density
  ED <- (M*psi)/A # Expected density
}

```

---

Panel 18.1: **JAGS** code to fit the spatial count model. This version includes the latent encounter histories.

The **JAGS** model in Panel 18.1 can be used to fit the version of the model in which the latent encounters are updated at each Monte Carlo iteration. One challenge faced when using this version of the model is that **JAGS** cannot auto-generate initial values that honor the constraints in the model, so it is necessary to provide them. The following code presents one fairly general way of creating acceptable starting values and formatting the data for analysis using the **rjags** package:

```

15364 library(rjags)
15365 dat1 <- list(n=n, X=X, J=J, K=K, M=200, xlim=xlim, ylim=ylim)

```

```

15366 init1 <- function() {
15367   yi <- array(0, c(dat1$M, dat1$J, dat1$K))
15368   for(j in 1:dat1$J) {
15369     for(k in 1:dat1$K) {
15370       yi[sample(1:dat1$M, dat1$n[j,k]),j,k] <- 1
15371     }
15372   }
15373   list(sigma=runif(1, 1, 2), lam0=runif(1),
15374        y=yi, z=rep(1, dat1$M))
15375 }
15376 pars1 <- c("lam0", "sigma", "N", "mu")

```

The code in Panel 18.1 is useful because it shows how closely this model is related to standard SCR models, and it provides the basis for including data on both marked and unmarked individuals, as will be discussed in the next chapter. However, this model runs very slowly, even when using a fast 64-bit machine with chains run in parallel. The code in Panel 18.2 runs much faster because it does not include the latent encounter histories.

An even faster (but perhaps less efficient) alternative is to use the `scrUN` function in `scrbook`. The usage is as follows:

```

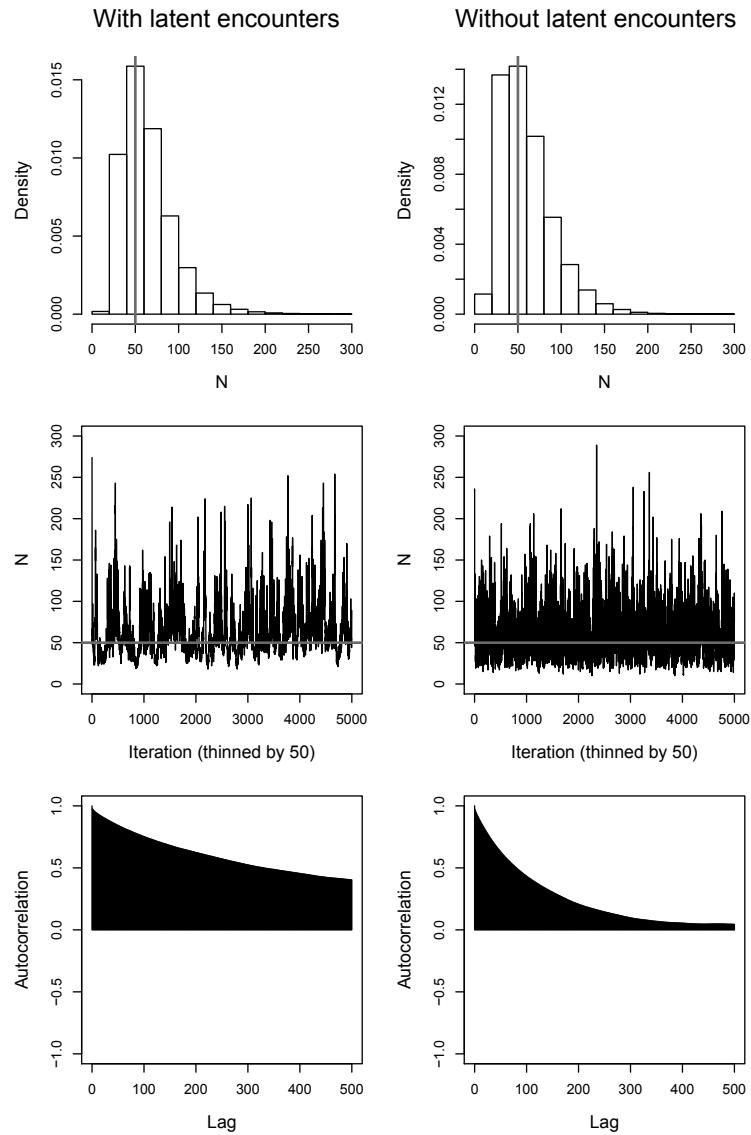
15384 out1 <- scrUN(n=n, X=X, M=300, niter=25000, xlims=xlim, ylims=ylim,
15385              inits=list(lam0=0.3, sigma=rnorm(1, 5, 0.1)), updateY=TRUE,
15386              tune=c(0.004, 0.09, 0.35))

```

where `n` is the matrix of counts, `X` is the trap coordinate matrix, `M` sets the size of the data-augmented latent data, `xlims` and `ylims` define the rectangular state-space, `inits` is a list of starting values, and `updateY` determines if the latent encounter histories are updated as part of the MCMC algorithm. In general, we recommend using the option `updateY=FALSE` because the Markov chains tend to mix better. Even so, it can be important to fiddle with the tuning parameters until the acceptance rates are between 40–60%. Otherwise, the Markov chains will exhibit extremely high autocorrelation. This is one reason to favor **JAGS** over our implementation in `scrbook` since **JAGS** finds suitable tuning parameters automatically during the adaptive phase (when using Metropolis updates).

We fit the model to the simulated data using both formulations – with and without the latent encounter histories – and the results are given in Table 18.1 and Fig. 18.2. Table 18.1 shows summarizes of 10000 posterior draws, and suggests that while the true parameter values are easily covered by the 95% credible intervals, the intervals are rather wide. In many cases, knowing that there are between 21 and 113 individuals in an area will be considered relatively imprecise. This is not just a peculiarity of this particular data set – in general, posterior precision will be low, as noted by Chandler and Royle (In press). Furthermore, as indicated by Fig. 18.2, the autocorrelation of the samples is high, and thus it may take many iterations to achieve convergence. Moreover, the algorithm that includes the latent encounter histories seems to have a hard time exploring the region of the posterior in which  $N$  is low (○). Given these technical difficulties, we recommend using the **JAGS** implementation (based on Panel 18.2), and it is always a good idea to use MCMC diagnostic tools such as those available in the `coda` package.

The take-home message is that, even with simulated data, the precision of the posterior distributions is low and mixing is poor. This should be expected given that we are asking so



**Figure 18.2.** MCMC results for the parameter  $N$  from the two algorithms (with and without the latent encounter histories). The first row contains the histograms of the posterior distributions, the second row contains the history plots, the third row shows the autocorrelation plots.

---

```

model{
sigma ~ dunif(0, 200)
lam0 ~ dunif(0, 5)
psi ~ dbeta(1,1)
for(i in 1:M) {
  z[i] ~ dbern(psi)
  s[i,1] ~ dunif(xlim[1], xlim[2])
  s[i,2] ~ dunif(ylim[1], ylim[2])
  for(j in 1:J) { # Number of traps
    distsq[i,j] <- (s[i,1] - X[j,1])^2 + (s[i,2] - X[j,2])^2
    lam[i,j] <- lam0 * exp(-distsq[i,j] / (2*sigma^2)) * z[i]
  }
}
for(j in 1:J) {
  bigLambda[j] <- sum(lam[,j])
  for(k in 1:K) {
    n[j,k] ~ dpois(bigLambda[j])
  }
}
N <- sum(z[])
A <- (xlim[2]-xlim[1])*(ylim[2]-ylim[1]) * 10000 # Area of state-space (ha)
D <- N / A      # Realized density
ED <- (M*psi)/A # Expected density
}

```

---

Panel 18.2: **JAGS** code to fit the spatial count model. This version does not include the latent encounter histories, and thus runs much faster than the code in Panel 18.1.

15411 much from so little data. In essence we are trying to fit a point process model while being  
 15412 twice removed from the actual point (activity center) locations. These difficulties may  
 15413 warrant the investigation of simpler models at the expense of the mechanistic description  
 15414 of the system. Another option is to figure out ways of improving model precision – options  
 15415 we discuss in Sec. 18.7. Before doing so, we re-analyze the Northern Parula (*Parula*  
 15416 *americana*) data described in Chandler and Royle (In press)

## 18.6 THE MARYLAND NORTHERN PARULA STUDY

15417 The parula data are standard avian point count data, with one exception. Typically,  
 15418 points are spaced by > 200 m when studying passerines in order to maintain statistical  
 15419 independence. In contrast, the parula data were collected at 105 points located on a 50-m  
 15420 grid, which virtually ensures spatial correlation since the parula song can be heard from

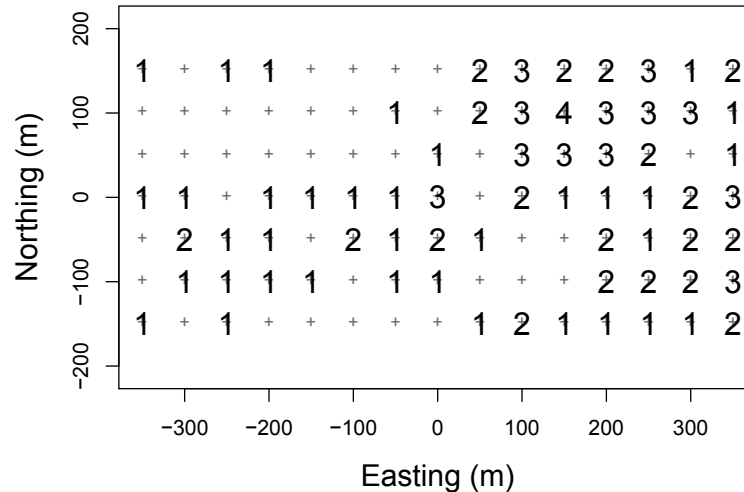
**Table 18.1.** Posterior summaries from the spatial count ("SC") model applied to simulated data using `scrbook` and **JAGS**. 25000 samples were generated, but substantial Monte Carlo error is still evident.

Parameter	Mean	SD	2.5%	50%	97.5%
<b>scrUN(..., updateY=FALSE)</b>					
$\sigma = 5$	4.718	0.922	3.239	4.615	6.833
$\lambda_0 = 0.4$	0.500	0.136	0.268	0.489	0.793
$N = 50$	60.653	31.067	21.000	54.000	137.000
<b>scrUN(..., updateY=TRUE)</b>					
$\sigma$	4.554	0.784	3.216	4.486	6.264
$\lambda_0$	0.489	0.131	0.262	0.479	0.775
$N$	64.772	30.162	26.000	59.000	140.000
<b>JAGS (without latent encounter histories)</b>					
$\sigma$	4.70	0.88	3.24	4.66	6.63
$\lambda_0$	0.52	0.14	0.27	0.52	0.80
$N$	58.55	30.30	20.00	52.00	135.00

distances  $> 50$  m. Each point was surveyed 3 times during June 2006, and Fig. 18.3 depicts the resulting spatially-correlated counts ( $n_j$ ). A total of 226 detections were made with a maximum count of 4 during a single survey. At 38 points, no warblers were detected. All but one of the detections were of singing males, and this one observation was not included in the analysis.

We fit the model using **JAGS** and the code from Panel 18.2, which does not include the latent encounter histories. For comparative purposes, we used proper priors rather than the improper priors used Chandler and Royle (In press), but all other aspects of the analysis were the same, including  $M = 300$  and a state-space created by buffering the grid of point count locations by 250 m. To reduce computation time, we used the `parallel` package and distributed 3 chains to 3 separate cores. The entire example can be reproduced using the code on the help page for `nopa` in our **R** package `scrbook`. The following illustrates the essential elements:

```
library(scrbook)
library(rjags)
dat2 <- list(n = nopa$n, X = nopa$X, M=300, J=nrow(nopa$n), K=ncol(nopa$n),
            xlim=c(-600, 600), ylim=c(-400, 400))
init2 <- function() {
  list(sigma=rnorm(1, 100), lam0=runif(1), z=rep(1, dat2$M))
}
cl2 <- makeCluster(3) # Open 3 parallel R instances
clusterExport(cl2, c("dat2", "init2", "pars1")) # send objects to 3 cores
system.time({
  out2 <- clusterEvalQ(cl2, { # executes the following command on each core
    library(rjags)
    jm <- jags.model("nopa2.jag", dat2, init2, n.chains=1, n.adapt=500)
    jc <- coda.samples(jm, pars1, n.iter=2500)
    return(as.mcmc(jc))
  })
})
```



**Figure 18.3.** Spatially-correlated counts of northern parula. Gray crosses are the locations of the 105 point count stations. Superimposed are the number of detections after 3 survey occasions.

```

15450 } ) # 6242
15451 mc2 <- mcmc.list(out2) # put the 3 chains together
15452 plot(mc2)
15453 summary(mc2)

```

XXXX THE RESULTS OF THE TWO ANALYSES LOOK VERY SIMILAR AS  
 EXPECTED XXXX

Several aspects of this analysis could be improved via model extensions. In particular, we note that a more appropriate observation model would recognize the fact that detection in this case is the result of two processes. Specifically, an ideal encounter probability model would include a process describing the location of the bird (not just its home range center) as well as the probability of detecting it, given it's location during the survey. Essentially, the model we would like to fit could be thought of as a latent distance sampling model allowing for movement. As it turns out, a very rudimentary form of distance data were collected – birds were determined to be either within 150 m or beyond 150 m from the observer. In Sec. 18.8, we propose a model to accommodate these auxiliary data.

## 18.7 IMPROVING PRECISION

Chandler and Royle (In press) recommended two strategies for improving the precision of the posterior distributions obtained under the SC model: (1) mark a subset of individuals or (2) elicit informative priors from the published literature. The first option is the subject of the next chapter. The second option should be readily accomplished in many studies



15469 because extensive information on home range size has been compiled for many species in  
 15470 diverse habitats (*e.g.*, DeGraaf and Yamasaki, 2001). It is easy to ~~embody~~ this information  
 15471 as a prior distribution in Bayesian analyses (Chandler and Royle (In press), Chapt. 5).

15472 In some cases, it may not be possible to mark any individuals, and no prior information  
 15473 may exist about encounter parameters; however, it may be possible to collect axillary data,  
 15474 such as the distance measurements recorded in the parula study. Other sources of auxiliary  
 15475 data could include removal counts ~~of~~ double observer counts, which are routinely collected  
 15476 in wildlife studies. Extending the model to accommodate such data is treated in the next  
 15477 section.

## 18.8 EXTENSIONS OF THE SPATIAL COUNT MODEL

15478 If ancillary data such as distance measurements exist, why bother with the SC model  
 15479 at all? Isn't density estimable using the distance data alone? Yes, in fact it is, and  
 15480 in many situations a simple distance sampling model will be sufficient. However, unlike  
 15481 the situation we described earlier in this chapter where we viewed spatial correlation as  
 15482 a good thing, the model extension we describe now provides a means of dealing with  
 15483 spatial correlation when it is unwanted or perhaps unavoidable. ~~In addition, extensions~~  
 15484 ~~of this model~~ We suspect that the SC model could can be used to make inferences about  
 15485 multiple processes in addition to spatial and temporal variation, such as home range size  
 15486 and movement.

15487 As an example, consider again the northern parula data. As it turns out, observers  
 15488 recorded rudimentary distance sampling data by determining if each detected individual  
 15489 was within or beyond 100 m. Although not ideal, distance data binned into 2 intervals  
 15490 are sufficient for estimating the scale parameter of a distance sampling detection function,  
 15491 and thus we should be able to use that information to increase precision and develop a  
 15492 more realistic encounter model. Doing so requires that we consider not only the activity  
 15493 centers, but also the actual locations of individuals during each survey – much like in  
 15494 search-encounter models (Chapt. 15).

15495 By including both activity centers ( $\mathbf{s}$ ) and actual locations ( $\mathbf{u}$ ) in the model, abundance  
 15496 in any region  $\mathcal{B}$  is given by

$$N(\mathcal{B}) = \sum_i I(\mathbf{u}_i \in \mathcal{B}). \quad (18.8.1)$$

15497 Thus, in the context of distance sampling studies in which the distance data are recorded  
 15498 in discrete intervals, the region  $\mathcal{B}$  would be the area corresponding to a particular distance  
 15499 interval. The probability of detecting the individuals  $N(\mathcal{B})$  would be the average detection  
 15500 probability  $\bar{p}$ , which is computed by integrating a distance-based detection function over  
 15501 the distance interval.

In other contexts, such as when conducting removal surveys, the region  $\mathcal{B}$  could be a  
 fixed-area plot, such as a stream segment. Again, Eq. 18.8.1 could be used to model local  
 abundance ( $n(\mathcal{B})$ ), and detection probability within the region could be modeled condi-  
 tional on  $n(\mathcal{B})$ . For instance, if a stream segment is uniformly surveyed using electrofishing  
 equipment, then a standard non-spatial removal model could be used to estimate detec-  
 tion probability  $p$ , conditional on the spatially-explicit model of abundance. A reasonably

general description of this model is as follows:

$$\begin{aligned} \mathbf{s}_i &\sim \text{Unif}(\mathcal{S}) \\ \mathbf{u}_{ik} &\sim \text{BiNormal}(\mathbf{s}_i, \tau) \\ N(\mathcal{B}_{jk}) &= \sum_{i=1}^M I(\mathbf{u}_{ik} \in \mathcal{B}_{jk}) \\ n_{jkl} &\sim \text{Binomial}(N(\mathcal{B}_{jkl}), p) \end{aligned}$$

where  $\tau$  is the parameter of a bivariate normal distribution (with correlation  $\rho = 0$ ) describing the locations of individuals on occasion  $k$ . The interpretation of the parameter  $p$  will depend upon the survey protocol.

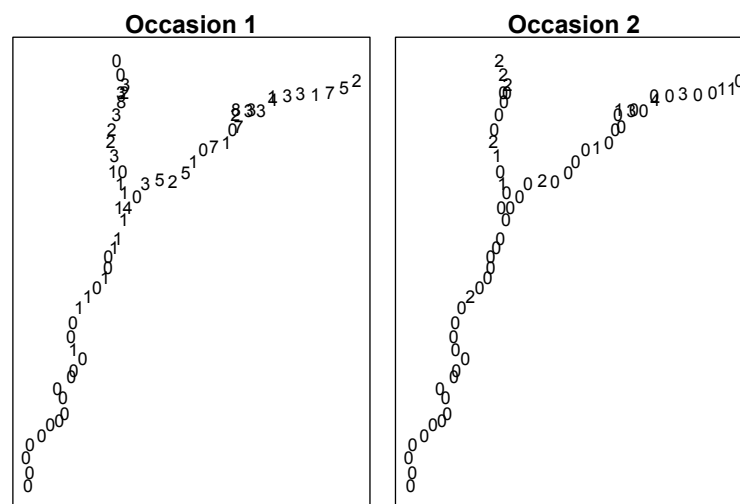
When plots are far enough apart that individuals cannot move between them, the counts will be uncorrelated and the model can be approximated using a non-spatial  $N$ -mixture model allowing for temporary emigration (Chandler et al., 2011). In the next example, we consider data in which the plots are obviously not independent.

## 18.9 THE MARYLAND DUSKY SALAMANDER STUDY

The independence assumption of the Chandler et al. (2011) model will not always hold. A prime example is in studies of aquatic species in stream networks. For example, consider the data depicted in Fig. 18.4. What is this spaghetti soup, you say? These are streams of numbers, i.e. counts of northern dusky salamanders (*Desmognathus fuscus*) in 25-m stretches on a small stream in the Chesapeake and Ohio National Historic Park. The data were collected by E.H.C. Grant and colleagues with the objective of understanding the spatial and temporal dynamics of salamander populations in response to seasonal and annual variations in stream hydrology. In addition, movement processes, including dispersal are studied between years (see Grant et al. (2010) for more details).

To sample the population, the stream networks are divided into 25-m stretches as illustrated in Fig. 18.4. In each stretch, “temporary” removal sampling is used, which involves capturing and removing salamanders on 3 consecutive passes. The salamanders are placed in a bucket of water for the brief 10–20 min duration of sampling, and then they are released at the location of capture. The entire process is repeated 3–4 times per season (May–Aug). In a subset of streams and years, individuals are marked, but in general it is too expensive to mark the entire population, and the data considered here consists entirely of unmarked individuals.

The sampling protocol may be thought of as a “robust design” (Pollock, 1982), with “occasions” (typically 1 day) being the primary period, and secondary samples being the removal passes within the primary periods. An obvious feature of these data is that the neighboring counts are spatially correlated. In this case, we have reason to believe that this correlation is the result of habitat preferences, with individuals actively selecting habitat in the upper reaches of the streams. This could be modeled as a function of a covariate describing the distance from the mouth of the stream. Another obvious feature of this data is that the pattern of spatial correlation remains consistent between occasions, but the overall counts decline markedly over the course of the season. These phenomenon can be explained by the fact that the salamanders have relatively small home ranges, and



**Figure 18.4.** Stream segment counts of northern dusky salamanders in the Chesapeake and Ohio National Historic Park, VA/MD. Each number is the count associated with a 25-m stretch in which 3 removal passes were made on 3 occasions each summer (only 2 occasions are shown here). Notice the consistency of the spatial correlation between occasions and the temporal decline in the counts.

**Table 18.2.** Posterior summaries from removal model of salamander counts allowing for movement and decreasing population size over the course of a breeding season.

Parameter	Mean	SD	2.5%	50%	97.5%
$N_1$	178.393	16.346	151.000	177.000	214.000
$N_2$	62.322	6.884	51.000	62.000	77.000
$N_3$	21.202	3.695	15.000	21.000	29.000
$\phi$	0.348	0.038	0.275	0.348	0.425
$\tau$	27.427	3.200	21.293	27.173	33.706
$p$	0.396	0.053	0.294	0.394	0.502

this results in the consistent pattern of correlation among occasions. Furthermore, as the season progresses, the streams dry out, and many individuals move underground.

Given the importance of movement within home ranges, which determines the correlation among occasions, and movement underground, which results in a decreasing number of individuals being available for sampling, it would be helpful to have a model that describes both processes and allows for evaluation of hypotheses regarding the effects of environmental variables. For example, one might ask how stream flow is related to the probability that an individual remains active. A model describing this process could be used to predict activity levels under future conditions. Although we do not investigate covariate effects in this section, we do present a general model allowing for movement among occasions, and for decreasing availability over the season.

This expanded model is founded on the one described in the previous section, but it also includes a removal model for the observation process, and it includes a basic “open” population model to allow for a decline in abundance over time (Chapt. 16). Actually, the population is not thought to actually decline substantially during the season, but rather, the number of individuals *available* for detection declines because many individuals move underground as the streams dry. Each of these model components is included in the BUGS description of the model presented in Panel 18.3.

We fit this model to the data and obtained the posterior distributions summarized in Table 18.2. The results indicate that the population size available for detection did decrease rapidly during the season, the rate of which is determined by the  $\phi$  parameter. Modeling this parameter as a function of water flow or volume would allow one to predict salamander activity under future environmental conditions. Another result of the analysis is that the movement parameter,  $\tau$ , was relatively low, indicating that adult salamanders rarely move more than 100 m from their home range center during a season. This explains why the distribution of individuals within the stream remains relatively constant over time. Including this parameter in the model also provides a general mechanism for modeling temporal correlation in count data.

## 18.10 SUMMARY AND OUTLOOK

Unlike traditional models of count data used in ecology, the SC model is parameterized in terms of *individuals* – individuals that just so happen not to be directly observed. The reason for accommodating this latent structure is that it provides a more mechanistic description of ecological systems. For example, the model allows us to attach a mechanism

15568 – movement – to the widely observed phenomenon of spatial correlation in count data.  
15569 In addition, by parameterizing the model in terms of individuals, it makes it possible to  
15570 incorporate data from both marked and unmarked individuals, as will be described in the  
15571 next section. This ability to combine different types of data should make it possible to  
15572 design effective monitoring programs when resources are too limited to conduct ~~extensive~~  
15573 spatially extensive capture-recapture studies, as has been done ~~in~~ with ~~non-spatial models~~  
15574 (Conroy et al., 2008).

15575 The SC model is a conceptually simple extension of standard SCR models, but in  
15576 terms of computational requirements and latent structure, it is perhaps at the extreme  
15577 end of what is possible to do with count data. As is always true, the harder we try to  
15578 mirror reality with our models, the harder it becomes to estimate the parameters of the  
15579 system. In this chapter, we tried to emphasize that as conceptually appealing as the SC  
15580 model may be, it is unlikely to produce satisfying results in the absence of additional  
15581 information. However, additional information such as home range size estimates will often  
15582 be available for many species, and if not, we have provided an alternative method of  
15583 accommodating additional data in the form of distance measurements or removal counts.  
15584 This can greatly increase precision in studies designed to make spatially-explicit inferences  
15585 about population processes.