Hierarchical models also propose a consistent probabilistic rationale for prediction. Inferences about the random effect for a new group, Z^{new}, can be derived through the predictive distribution conditional on the observed data ([117]). It is an average of the conditional population distribution $[Z|\theta]$ over the posterior distribution of the parameters θ conditioned on all observed data:

$$[Z^{new}|y_{1:n}] = \int_{\theta} [Z^{new}, \theta|y_{1:n}]\, d\theta$$

$$= \int_{\theta} [Z^{new}|\theta] \times [\theta|y_{1:n}]\, d\theta \qquad (9.5)$$

The difference between the marginal prior predictive $[Z] = \int_{\theta}[Z|\theta] \times [\theta]\, d\theta$ and the posterior predictive $[Z|y_{1:n}] = \int_{\theta}[Z|\theta] \times [\theta|y_{1:n}]\, d\theta$ in Eq. (9.5) reflects the amount of information brought by the data $y_{1:n}$ of all groups $k = 1..n$ to update the prior distribution common to all groups. Additionally in the examples, writing $y_k = (y_k^1, ...y_k^{p_k})$, the number p_k of elementary data may differ between groups.

9.2 Hierarchical exchangeable Binomial model for capture-mark-recapture data

This section is a natural followup to Chapter 4. The main idea is to link the capture-mark-recapture (CMR) models for yearly observations together by a hierarchical structure as the one represented by the DAG of Fig. 1.13. In this case study inspired by the article of Rivot and Prévost ([255]), the years are the statistical units that look alike and that hypothesized resemblance allows for transferring information from a given year to the other years. The only differences with the baseball players' model from Appendix E are that the data structure is a little more sophisticated and that the latent variables are two component vectors instead of simple real numbers.

9.2.1 Data

Relatively long but sparse (small sample size) series of data are quite common when dealing with CMR surveys aimed at estimating the abundance of wild populations over a series of years. For instance on the Oir

FIGURE 9.1: Marking a spawner entering the Scorff River.

River, already presented for the smolt runs in Chapter 4, the rangers from the French National Research Institute for Agronomy (INRA) and from the National Office of Water Management (ONEMA) have collected CMR data about adult salmon that swam back to spawn in the Oir River for the years 1984–2000.

Data are shown in Table 9.1. For each year t from 1984 to 2000, $y_{1,t}$ denotes the number of fish trapped at the Cerisel station (close to the mouth of the river, see Fig. 4.3 of Chapter 4 for more details about the trapping device). $y_{2,t} + y_{3,t}$ individuals from the captured ones are not replaced upstream, either because they died during manipulation or because they are removed for experimental use or for hatchery production. Let $y_{4,t} = y_{1,t} - (y_{2,t} + y_{3,t})$ the number of (tagged) fish released. These spawners are individually marked before they keep on swimming upstream (Fig. 9.1). The recapture sample is gathered during and after spawning (see more details on recapture conditions hereafter). Let us denote as $y_{5,t}$ and $y_{6,t}$ the number of marked and unmarked fish among recaptured fish, respectively.

9.2.2 Observation submodels for the first phase (Cerisel trapping place)

Let denote ν_t the unknown of interest, *i.e.*, the population size of spawners at year t and π_t^1 the unknown trapping efficiency. Assuming all

| Years	y_1	y_2	y_3	y_4	y_5	y_6
1984	167	10	3	154	12	10
1985	264	37	11	216	21	4
1986	130	28	9	93	5	4
1987	16	3	1	12	2	22
1988	226	35	8	183	12	0
1989	235	31	5	199	56	0
1990	15	4	4	7	2	15
1991	44	0	0	44	23	1
1992	31	10	1	20	4	5
1993	100	17	2	81	4	3
1994	32	12	2	18	1	4
1995	109	6	1	102	39	7
1996	70	13	2	55	25	57
1997	56	19	3	34	12	3
1998	34	3	1	30	6	30
1999	154	5	1	148	13	22
2000	53	0	0	53	4	33

TABLE 9.1: Capture-mark-recapture data for spawners by spawning migration year in the Oir River. y_1: Number of fish trapped at the counting fence during the upstream migration time; y_2, y_3: One sea-winter (resp. two sea-winter) fish removed from the population; y_4: Tagged and released fish; y_5, y_6: Number of marked (resp. unmarked) recaptured fish.

of the ν_t spawners are independently and equally catchable in the trap, with a probability π_t^1 considered constant over the migration season, the migration of the ν_t spawners are independent Bernoulli experiments with probability of "success" π_t^1. Accordingly, $y_{1,t}$ is the observed result of a Binomial experiment as given by Eq. (9.6):

$$y_{1,t} \sim Binomial(\nu_t, \pi_t^1) \tag{9.6}$$

9.2.3 Observation submodels for the second phase (recollection during and after spawning)

The fate of a salmon swimming upstream to spawn is described in great detail in Chapter 5. The recapture sample is obtained by three methods: electrofishing on the spawning grounds, collection of dead fish after spawning, and trapping of spent fish at the downstream trap of the Cerisel facility. Due to the available data here, we adopt a simplified
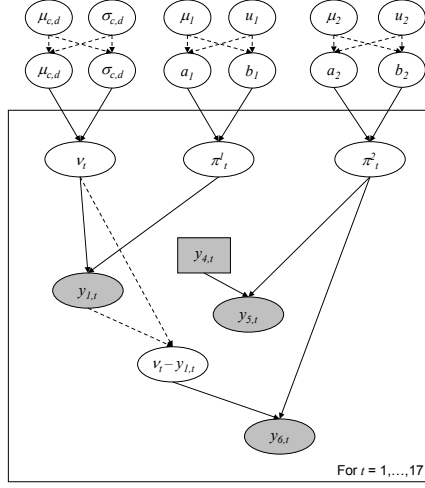
FIGURE 9.2: Directed Acyclic Graph representation of the hierarchical structure for the joint modeling of capture-mark-recapture experiments for the 17 years.

version of the model. As a first approximation to this rather complicated scheme, recapture Binomial experiments with efficiency π_t^2 for marked fish and untagged ones are acceptable, providing one is willing to assume the three following hypotheses: No spawner runs downstream after getting over the trap ($H1$); there is no tag shedding ($H2$); the recapture probability π_t^2 is the same for all the fish whether or not marked ($H3$):

$$\begin{cases} y_{5,t} \sim Binomial(y_{4,t}, \pi_t^2) \\ y_{6,t} \sim Binomial(\nu_t - y_{1,t}, \pi_t^2) \end{cases} \tag{9.7}$$

9.2.4 Latent layers

Writing latent vector $Z_t = (\nu_t, \pi_t^1, \pi_t^2)$ and observation $y_t = (y_{1,t}, y_{5,t}, y_{6,t})$ to cope with the notations of Fig. 1.13, we are now in search of a hierarchical structure to express that, to some extent, years may look like one another. The natural choice for the latent distributions are the Beta distribution (see Eq. (2.6), page 54) for π_t^1 and for π_t^2

$$\begin{cases} \pi_t^1 \sim Beta(a_1, b_1) \\ \pi_t^2 \sim Beta(a_2, b_2) \end{cases} \tag{9.8}$$
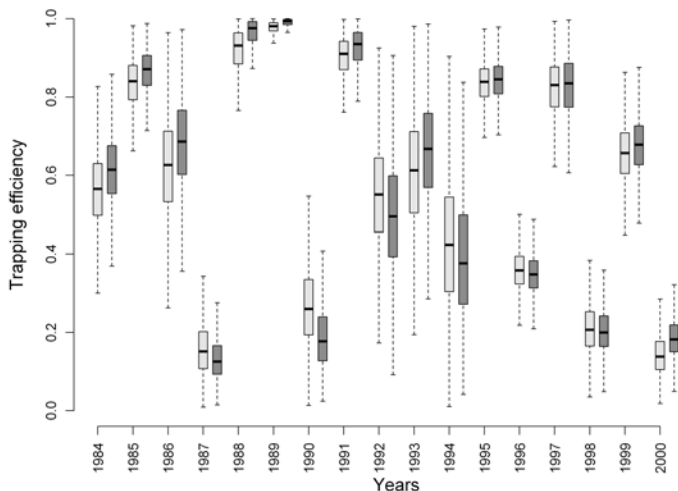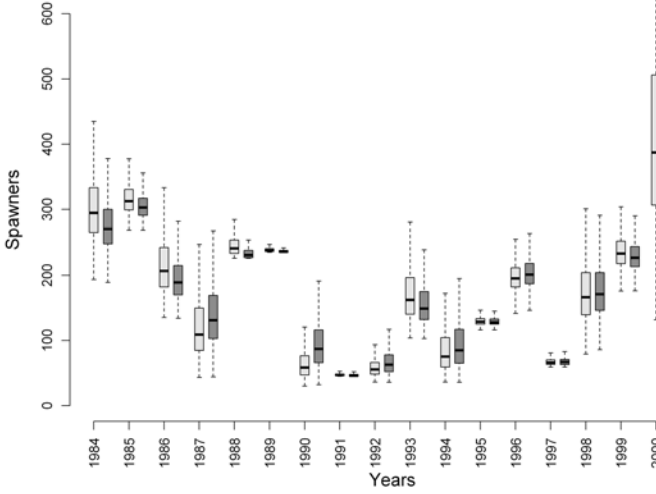
FIGURE 9.3: Box and whisker plots of a size-15,000 MCMC sample from the marginal posterior distribution of the trapping efficiency at the Cerisel trap by migration year obtained under two model configurations, the model in which years are treated independently (light gray) and the hierarchical model (dark gray). The boxes indicate the interquartile range and the median.

and the Negative Binomial distribution (already met at Eq. (4.11)) for ν_t

$$\nu_t \sim NegBinomial(c, d) \qquad (9.9)$$

Figure 9.2 shows a DAG representation of the exchangeable hierarchical model for the joint modeling of the capture-mark-recapture for the 17 years. The higher level parameters $(a_1, a_2, b_1, b_2, c, d)$, denoted θ consistently with notations of Fig. 1.13, are generally assigned a diffuse prior distribution to reflect some ignorance about them. A common practice is to set a prior on some appropriate one-to-one transformed parameters and then to go back to the original parameters via the inverse transformation. Most often, the transformation recovers the mean and variance because of their well-understood meaning. Sometimes, the mean can be assigned a rather informative prior but it is generally not the case of the variance (that describes between years variability), an uppermost unknown quantity.

FIGURE 9.4: Box and whisker plots of a size-15,000 MCMC sample from the marginal posterior distribution of the number of spawners by migration year obtained under two model configurations, the model in which years are treated independently (light gray) and the hierarchical model (dark gray). The boxes indicate the interquartile range and the median.

For the capture and recapture efficiencies, a diffuse (noninformative) prior can be put on the mean $\mu_{a,b} = \frac{a}{a+b}$ and variance $\sigma_{a,b}^2 = \frac{ab}{(a+b)^2(a+b+1)}$ of the Beta distributions. But this transformation introduces some unnecessary difficulty in order to ensure that both parameters a and b are positive, the variance must satisfy the constraint

$$\sigma_{a,b}^2 < \mu_{ab}(1 - \mu_{ab})$$

We used a rather simpler transformation by considering the transformation $\mu_{a,b} = \frac{a}{a+b}$ and $u_{a,b} = (a+b)$ and by drawing $\mu_{a,b}$ in a diffuse prior distribution (we took a $Beta(1.5, 1.5)$) and $log(u_{a,b})$ in a $Uniform(0, 10)$ distribution. Keeping in mind the interpretation of Beta coefficients (a, b) as prior distributions for Binomial trials (see Chapter 2), a and b are to be interpreted as prior number of success and failures, respectively. Then, $u = (a+b)$ is interpreted as a prior sample size that scales the variance of the Beta prior distribution, and a Uniform distribution on the log-scale

is appropriate for a diffuse prior (see also [187] for another utilization of such a parameterization).

For the number of spawners varying between years, the mean and variance of the Negative Binomial distribution are respectively $\mu_{c,d} = \frac{c}{d}$ and $\sigma_{c,d}^2 = \frac{c(d+1)}{d^2}$. We assigned to $\mu_{c,d}$ a bounded Uniform distribution over $]0, \mu_{max}]$. In the case study, μ_{max} is set to 3000 individuals, a reasonable upper bound for a salmon fish population on the Oir River due to bio-ecosystemic constraints. To ensure that (c, d) are positive, we draw $log(\sigma_{c,d}^2)$ in a diffuse Uniform distribution over the bounded range $[log(\mu_{c,d}), log(\sigma_{max}^2)]$ with $\sigma_{max}^2 = 12$ since we do not believe that the standard deviation might exceed 400 fish $(400 \approx \sqrt{\exp(12)})$.

Of course more informative priors should be used when available and robustness to the choice of prior must be investigated as in Rivot and Prévost [255].

9.2.5 Results

To show how a transfer of information between years is organized by the hierarchical model, we compare its results with the model assuming independence between years. For the models with independence, independent prior distributions with known parameters were set on $(\nu_t, \pi_t^1, \pi_t^2)$:

$$\begin{cases} \pi_t^1 \sim Beta(1.5, 1.5) \\ \pi_t^2 \sim Beta(1.5, 1.5) \\ \nu_t \sim Uniform(1, 3000) \end{cases} \qquad (9.10)$$

Inference has been performed via WinBUGS (see the supplementary material available from the book's website *hbm-for-ecology.org*).

Results highlight that hierarchical modeling has no effect on the inferences on the capture efficiencies (Fig. 9.3), but greatly improves posterior inferences for the number of spawners migrating back to the Oir River (Fig. 9.4).

Posterior mean values of the capture probabilities π_t^1 do not seem to shrink much toward their overall grand mean (Fig. 9.3) and the recapture probabilities π_t^2's (not shown) are only slightly subjected to the shrinkage effect. In Fig. 9.3, there remains a lot of between-year variability in the experimental conditions at the Cerisel trapping facility.

Conversely, the hierarchical structure hypothesized on ν_t's strongly reduces the skewness and uncertainty in the estimation of the number of spawners. The grey boxplots of Fig. 9.4 clearly point out that the most precise inferences are obtained under the hierarchical model, especially for the years with sparse CMR data, *i.e.*, low number of marked released
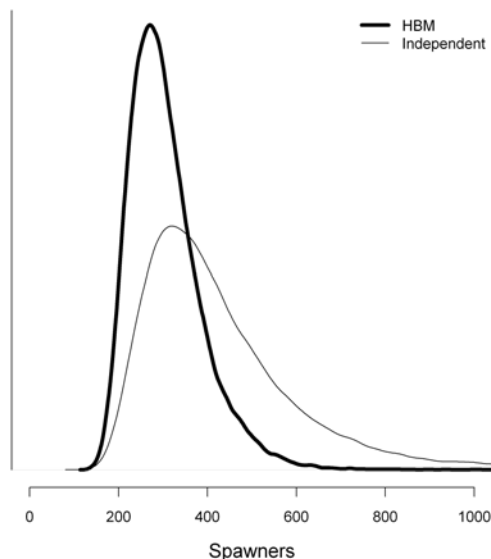
FIGURE 9.5: Marginal posterior distribution of the number of spawners ν_{2000} obtained under two model configurations, the model in which years are treated independently (thin line) and the hierarchical model (bold line).

or, more importantly, low number of recaptures of previously marked fish yield (*e.g.*, years 1987, 1990, 1994 and 2000). For this latter year, the upper bounds of the 95% Bayesian credibility intervals obtained with the model assuming independence between years appears unrealistically high given the size of the Oir River and the available knowledge on the biology and ecology of Atlantic salmon as exemplified for year 2000 in Fig. 9.5.

A straightforward result of the hierarchical model are the posterior predictive distributions of the trapping or recapture efficiencies and of the number of returns , denoted $[\pi^{1,new}|data_{1984:2000}]$ and $[\pi^{2,new}|data_{1984:2000}]$, and $[\nu^{new}|data_{1984:2000}]$, respectively. The posterior predictive of the trapping efficiency is an informative distribution with a mean value 0.124 and 95% of its density in the range $[0.016, 0.567]$ (Fig. 9.6). The posterior predictive of the returns has a mean value around 230 fish and 95% of its density in the range $[40, 610]$ (Fig. 9.6). Thus, the data of all years combined allow discarding *a priori* the pos-
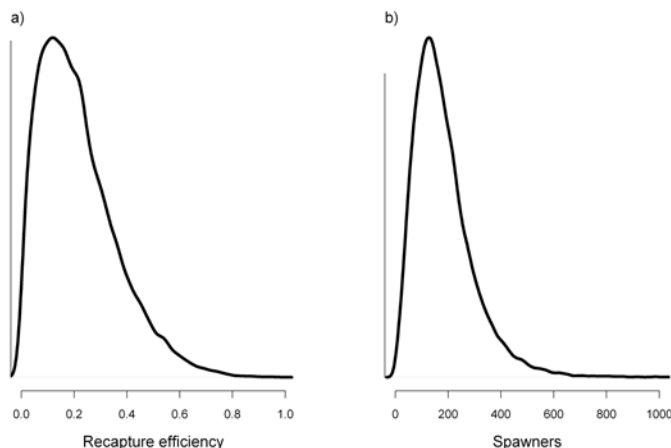
FIGURE 9.6: Posterior predictive distribution of of a) the trapping efficiency $[\pi^{1,new}|data_{1984:2000}]$, and b) adult returns $[\nu^{new}|data_{1984:2000}]$.

sibility of very high trapping efficiency (*i.e.*, greater than 0.5) or high spawner population size (*i.e.*, greater than a thousand) in any additional year.

In addition, the posterior inferences derived under the hierarchical model are rather insensitive to changing priors on $\mu_{a,b}$, whilst the model assuming independence is not, especially for the spawner stock of the sparse data years ν_{1987}, ν_{1990}, and ν_{1994} (results not shown). In-depth sensitivity analyses and another observation model for the recaptures can be found in Rivot and Prévost [255].

9.3 Hierarchical stock-recruitment analysis

This section develops a hierarchical extension of the Ricker Stock Recruitment model introduced in Chapter 7. The data and models are based on a published paper by Prévost et al. [239]; but see also [238].

We show that the hierarchical assemblage of several salmon populations (the biological analogs of baseball players from Appendix E) which we model as exchangeable units appears once again as the work-