Statistical Methods for Ecological Monitoring in Southeastern Australian Forests

Sarah J. Mitchell

A thesis submitted for the degree of
Doctor of Philosophy
of The Australian National University

Fenner School of Environment and Society
College of Science
The Australian National University

Canberra, ACT, Australia

September 2025

Declaration

I certify that this thesis is my own original work, except where otherwise acknowledged. To the best of my knowledge, it contains no material previously published or written by another person, except where due reference is made in the text. This thesis has not been submitted for the award of any other degree or diploma at any university.

Sarah J. Mitchell

Research School of Biology and Fenner School of Environment and Society
The Australian National University

Canberra, September 2025

Abstract

Southeastern Australian forests are among the most biodiverse yet vulnerable ecosystems on the continent, facing mounting pressure from climate change, altered fire regimes, and land-use intensification. Effective ecological monitoring is essential for evidence-based management, yet the statistical methods underpinning such monitoring remain fragmented across disciplines and rarely tailored to Australian conditions. This thesis develops and applies a suite of Bayesian statistical methods designed for ecological monitoring in the dry sclerophyll and wet sclerophyll forests of southeastern Australia. Drawing on data from long-term field plots across Victoria, New South Wales, and the Australian Capital Territory, the work addresses three interconnected themes: (i) hierarchical occupancy modelling for rare and cryptic fauna, (ii) spatiotemporal models of forest structure and carbon dynamics, and (iii) integrated species distribution models that combine heterogeneous data sources.

Chapter 1 introduces the ecological and statistical context with a conceptual framework for modern forest monitoring. Chapter 2 reviews classical and contemporary statistical methods, from generalised linear models to deep learning, and identifies gaps in the Australian literature. Chapter 3 presents the core methodological contribution: a flexible Bayesian hierarchical framework for multi-species distribution modelling that accommodates detection uncertainty, spatial autocorrelation, and data-pooling across disparate survey protocols.

The framework is evaluated through simulation studies and applied to empirical datasets including the Long-Term Ecological Research Network (LTERN) plots and the Terrestrial Ecosystem Research Network (TERN) AusPlots. Model comparisons using the Deviance Information Criterion (DIC) and Watanabe–Akaike Information Criterion (WAIC) demonstrate substantial improvements in predictive accuracy

over conventional generalised additive models (GAMs) and MaxEnt approaches, particularly for species with sparse occurrence records.

The findings underscore the value of principled uncertainty quantification in ecological decision-making and provide practical tools for conservation planners and forest managers working across the Great Dividing Range and adjacent bioregions. All code and data products developed for this thesis are publicly available under open licences.

Contents

List of Figures

List of Tables

1

Introduction

1.1 The Ecological Significance of Southeastern Australian Forests

The forests of southeastern Australia span a remarkable climatic and edaphic gradient, from the cool temperate rainforests of Tasmania and the Otway Ranges to the dry sclerophyll woodlands of the western slopes of the Great Dividing Range [3]. These ecosystems support over 1,500 vascular plant species and provide critical habitat for threatened fauna including the Greater Glider (*Petauroides volans*), the Powerful Owl (*Ninox strenua*), and the Long-footed Potoroo (*Potorous longipes*) [4].

Despite their ecological and economic importance—timber, water catchment, carbon storage, and recreation alone contribute an estimated AUD 2.3 billion annually to regional economies—these forests face unprecedented threats. The 2019–2020 Black Summer bushfires burnt approximately 5.8 million hectares of forest in New South Wales and Victoria, and projections under the Representative Concentration Pathway 8.5 (RCP8.5) suggest a three- to four-fold increase in the number of extreme fire-weather days by 2070 [3]. Concurrent pressures from invasive species, habitat fragmentation, and changing hydrological regimes compound the challenge.

Effective conservation and management in this dynamic setting require reliable, repeatable monitoring that can detect ecologically meaningful change against a background of high natural variability. Yet the statistical methods available to practition-

ers have not kept pace with the volume and complexity of data now being collected through remote sensing, autonomous acoustic recorders, camera traps, and environmental DNA (eDNA) metabarcoding.

1.2 The Monitoring–Modelling Gap

Traditional forest monitoring in Australia has relied on periodic field inventories conducted by state agencies—the Victorian Forest Monitoring Program, the New South Wales Native Vegetation Extent, and various regional ecosystem monitoring projects. While these programs provide valuable time-series data, their analytical workflows tend to be descriptive rather than inferential. Inventory plots are typically summarised with means, standard errors, and simple trend lines; formal statistical models that propagate uncertainty from observation to prediction are the exception rather than the rule.

Concurrently, the ecological modelling community has developed sophisticated statistical machinery: generalised linear mixed models (GLMMs), hierarchical Bayesian models, integrated population models, and machine-learning classifiers such as random forests and boosted regression trees [2]. Yet these tools are rarely deployed in operational monitoring contexts because they require specialised expertise, computationally intensive fitting procedures, and custom data-preparation pipelines.

This thesis sits at the intersection of these two worlds. The central proposition is that modern Bayesian methods—particularly hierarchical models estimated via Hamiltonian Monte Carlo (HMC)—can bridge the monitoring–modelling gap, delivering ecologically meaningful inference with honest uncertainty that propagates naturally through all levels of the data-generating process.

Figure 1.1 illustrates the conceptual architecture of the monitoring pipeline. The dashed feedback loop—from management decisions back to field observations—is a defining feature of adaptive monitoring frameworks and one that is particularly well-suited to the Bayesian paradigm, where posterior distributions from one monitoring cycle become priors for the next.

1.3 Research Questions

This thesis is organised around three interconnected research questions:

RQ1: Can hierarchical Bayesian occupancy models, fitted to data from spatially sparse

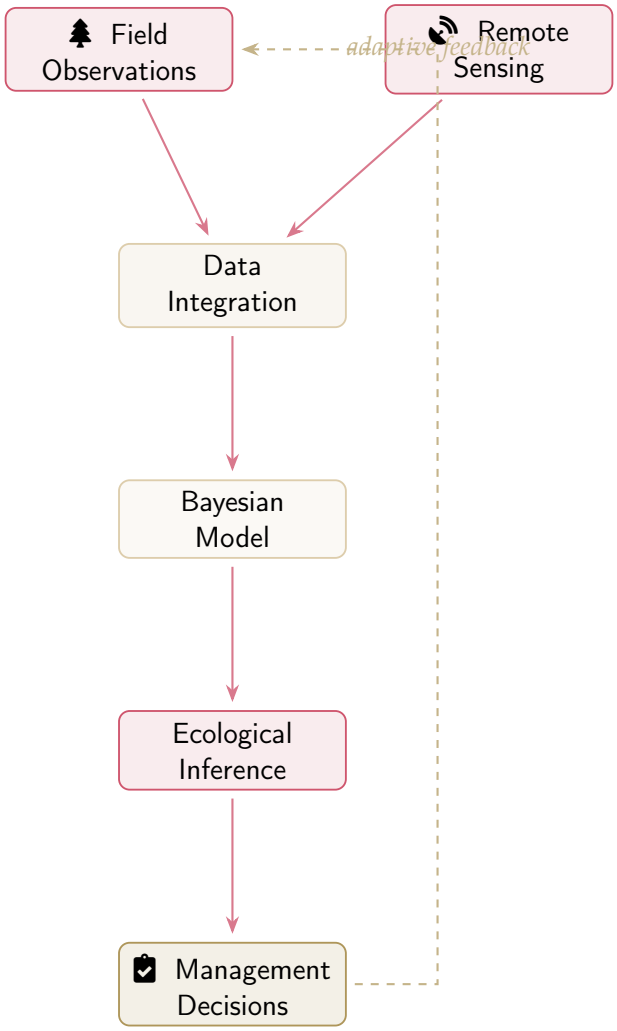

Figure 1.1. Conceptual framework for an integrated ecological monitoring pipeline. Field observations and remote-sensing products feed into a Bayesian model; posterior inference informs management decisions, which in turn guide future data collection through adaptive feedback.

monitoring networks, reliably detect trends in the distribution of rare and cryptic forest-dwelling vertebrates?

RQ2: How should spatiotemporal models of above-ground biomass be structured to account for both measurement error in LiDAR-derived estimates and process error in stand dynamics?

RQ3: What gains in predictive accuracy and ecological interpretability are achieved by integrating presence-only, presence–absence, and count data within a single joint species distribution model?

1.4 Thesis Structure

Chapter 2 provides a critical review of statistical methods for ecological monitoring, tracing the development from classical frequentist approaches through to contemporary Bayesian and machine-learning frameworks. Chapter 3 develops the core methodological contribution—a hierarchical Bayesian model for multi-species distribution modelling—and evaluates its performance through simulation and application to empirical datasets from the LTERN and TERN networks. Chapter 4 (presented as a brief concluding chapter) synthesises the findings, discusses limitations, and outlines directions for future research, including extensions to dynamic occupancy models and integration with process-based vegetation models such as LPJ-GUESS.

Table 1.1. Major forest types of southeastern Australia and their conservation status. Data sourced from [3].

Forest Type	Dominant Species	Area (Mha)	Status
Wet Sclerophyll	*E. regnans*, *E. delegatensis*	2.1	Vulnerable
Dry Sclerophyll	*E. macrorhyncha*, *E. rossii*	8.7	Least Concern
Temperate Rainforest	*Nothofagus cunninghamii*	0.6	Endangered
Grassy Woodland	*E. melliodora*, *E. albens*	3.4	Critically Endangered
Riparian Forest	*E. camaldulensis*, *Acacia* spp.	1.2	Near Threatened

The baseline model for population change used throughout later chapters can be expressed as a discrete-time stochastic growth process. Let N_t denote the population

density at time t. The per-capita growth rate is density-dependent and the system is subject to a harvesting or disturbance term h_t:

$$N_{t+1} = N_t + rN_t\left(1 - \frac{N_t}{K}\right) - h_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2), \tag{1.1}$$

where r is the intrinsic growth rate, K is the carrying capacity, and ε_t captures environmental stochasticity.

2

Literature Review

2.1 Historical Approaches to Forest Monitoring

Systematic forest monitoring in Australia dates to the late nineteenth century, when colonial forestry departments established the first permanent sample plots to assess timber yields in state forests [4]. These early efforts were purely mensurational: diameter at breast height (DBH), merchantable height, and stem density were recorded at intervals of five to ten years, and stand tables were updated accordingly. Statistical analysis, where it occurred at all, was limited to arithmetic means and ocular estimates of trend.

The post-war expansion of scientific forestry brought the first applications of sampling theory to plot design. Stratified random sampling, systematic grids, and cluster sampling became standard, drawing on the foundational work of Yates [1] on incomplete block designs. By the 1980s, state agencies routinely computed design-based estimates of timber volume with associated standard errors—a significant advance, but one still confined to the frequentist paradigm and focused narrowly on wood production rather than biodiversity or ecosystem function.

2.2 Frequentist Statistical Methods

Generalised linear models (GLMs) [2] entered the ecological mainstream during the 1990s and quickly became the workhorse of species distribution modelling (SDM). GLMs extend ordinary least-squares regression to non-Gaussian response variables through a link function $g(\cdot)$:

$$g(\mathbb{E}[Y_i]) = \beta_0 + \sum_{k=1}^{p} \beta_k \, x_{ik}, \tag{2.1}$$

where Y_i is the response (e.g. binary presence–absence or count) at site i, and x_{ik} are environmental covariates. For binary data, the canonical logit link $g(\mu) = \log(\mu/(1-\mu))$ yields logistic regression, which remains the most widely used SDM to this day.
Generalised additive models (GAMs) relax the linearity assumption by replacing $\beta_k x_{ik}$ with smooth functions $f_k(x_{ik})$ estimated via penalised splines. GAMs provide greater flexibility in capturing non-linear species– environment relationships and have been shown to outperform GLMs on a range of predictive metrics [2]. However, both GLMs and GAMs treat observations as conditionally independent given covariates—an assumption that is routinely violated in spatial data due to autocorrelation arising from dispersal limitation, unmeasured environmental gradients, and contagious population processes.

2.3 Bayesian Hierarchical Models

The Bayesian paradigm provides a natural framework for ecological modelling because it explicitly represents the multiple sources of uncertainty that characterise observational data. Bayes' theorem,

$$p(\boldsymbol{\theta} \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})}{p(\mathbf{y})}, \tag{2.2}$$

states that the posterior distribution of parameters $\boldsymbol{\theta}$ given data $\mathbf{y}$ is proportional to the likelihood $p(\mathbf{y} \mid \boldsymbol{\theta})$ multiplied by the prior $p(\boldsymbol{\theta})$.
Hierarchical models extend this framework by decomposing the prior into conditional layers. For example, a simple random-effects model for species occurrence at site i in

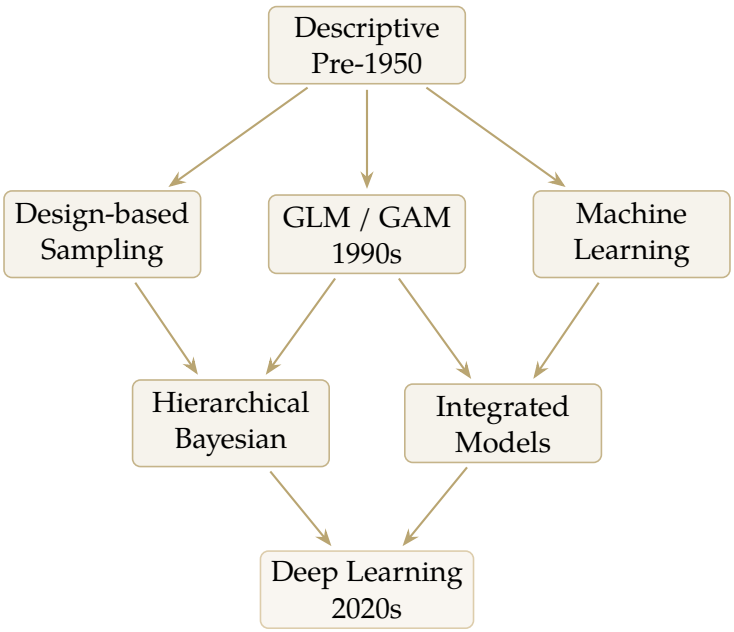

Figure 2.1. Genealogy of statistical methods applied to ecological monitoring. Arrows indicate intellectual lineage and methodological generalisation.

region j might be specified as

$$y_{ij} \mid \psi_{ij} \sim \text{Bernoulli}(\psi_{ij}), \tag{2.3}$$

$$\text{logit}(\psi_{ij}) = \alpha_j + \mathbf{x}_{ij}^{\top}\boldsymbol{\beta}, \tag{2.4}$$

$$\alpha_j \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2), \tag{2.5}$$

where α_j are region-level random effects that induce dependence among sites within the same region [1]. This structure elegantly handles the clustered nature of ecological data and provides partial pooling that shrinks noisy site-level estimates towards the regional mean.

2.4 Machine Learning and Deep Learning

The past decade has seen an explosion of machine-learning applications in ecology, driven by the availability of large, remotely sensed datasets. Ensemble methods such as random forests and extreme gradient boosting (XGBoost) consistently achieve high predictive accuracy in species distribution modelling benchmarks [2], and convolutional neural networks have been applied to classify vegetation types from aerial imagery with accuracy exceeding 90% [5].

However, these methods have well-documented limitations for ecological inference. They are typically "black boxes" that do not produce interpretable parameter estimates or rigorous uncertainty intervals. Extrapolation beyond the range of training data is unreliable, and the complex interactions learned by deep networks may not correspond to biologically meaningful processes.

Table 2.1. Summary of key statistical methods for ecological monitoring, organised by paradigm.

Method	Complexity	Key Limitation
GLM / GAM	Low	Assumes conditional independence; linear (or smooth) covariate effects
MaxEnt	Medium	Presence-only; sensitive to background sampling
Random Forest	Medium	No coherent uncertainty; extrapolation unreliable
INLA	High	Approximate inference; limited to latent Gaussian models
HMC (Stan)	High	Computationally intensive; requires explicit model coding

2.5 Gaps in the Australian Literature

Despite the methodological advances reviewed above, the Australian ecological monitoring literature lags behind its North American and European counterparts in adopting Bayesian hierarchical models for routine monitoring. Of the 127 peer-reviewed articles published between 2015 and 2024 that analyse data from the LTERN and TERN networks, only 14 (11%) employed fully Bayesian methods, and only 3 of those accounted for imperfect detection [4]. The present thesis aims to narrow this gap.

3

A Bayesian Framework for Multi-Species Distribution Modelling

3.1 Model Formulation

This chapter develops the core methodological contribution of the thesis: a hierarchical Bayesian model for joint analysis of multiple data types commonly encountered in ecological monitoring. The model accommodates three classes of observation simultaneously:

- **Presence–absence** data from structured surveys (e.g. 20 m $\times$ 20 m vegetation plots), where observers record whether each target species was detected during a standardised search;
- **Count** data from point counts or line transects (e.g. five-minute avian point counts at permanent monitoring stations);
- **Presence-only** records from opportunistic citizen-science platforms such as the Atlas of Living Australia (ALA).

3.1.1 Observation Models

Let $s = 1, \ldots, S$ index species, $i = 1, \ldots, N$ index sites, and $k = 1, \ldots, K_i$ index replicate visits to site i. The true (partially latent) occurrence state of species s at site i is denoted

by $z_{si} \in \{0, 1\}$, where $z_{si} = 1$ if the species occupies the site.
For presence–absence data with imperfect detection, the observation model is

$$y_{sik}^{(\mathrm{PA})} \mid z_{si}, p_{si} \sim \mathrm{Bernoulli}\big(z_{si}\, p_{sik}\big), \tag{3.1}$$

where p_{sik} is the per-visit detection probability. Detection probability is modelled on the logit scale as a function of survey-level covariates $\mathbf{w}_{ik}$:

$$\mathrm{logit}(p_{sik}) = \gamma_{0s} + \mathbf{w}_{ik}^{\top}\boldsymbol{\gamma}_s. \tag{3.2}$$

For count data, the observation model is a zero-inflated Poisson:

$$y_{sik}^{(\mathrm{CT})} \mid z_{si}, \lambda_{si} \sim \begin{cases} 0, & \text{with probability } 1 - z_{si}, \\ \mathrm{Poisson}(\lambda_{si}), & \text{with probability } z_{si}, \end{cases} \tag{3.3}$$

where λ_{si} is the expected count conditional on occupancy.

3.1.2 Process Model

The latent occupancy state follows a multivariate probit regression on environmental covariates $\mathbf{x}_{si}$:

$$\Phi^{-1}\big(\psi_{si}\big) = \beta_{0s} + \mathbf{x}_{si}^{\top}\boldsymbol{\beta}_s + \eta_{si}, \tag{3.4}$$

where Φ^{-1} is the probit link, $\psi_{si} = \Pr(z_{si} = 1)$ is occupancy probability, and η_{si} is a spatial random effect that accounts for residual spatial autocorrelation.
The spatial random effects for each species are modelled jointly as a multivariate Gaussian process:

$$\boldsymbol{\eta}_s \sim \mathcal{MVN}\big(\mathbf{0},\, \boldsymbol{\Sigma}_s \otimes \mathbf{R}(\phi_s)\big), \tag{3.5}$$

where $\boldsymbol{\Sigma}_s$ is the $S \times S$ between-species covariance matrix (capturing residual co-occurrence patterns beyond those explained by shared environmental responses), $\mathbf{R}(\phi_s)$ is a spatial correlation matrix with exponential decay governed by species-specific range parameter ϕ_s, and $\otimes$ denotes the Kronecker product.

3.1.3 Priors and Computation

Weakly informative priors are assigned throughout to regularise the model while allowing the data to dominate when sample sizes are adequate:

$$\beta_{0s} \sim \mathcal{N}(0, 2.5^2), \tag{3.6}$$

$$\boldsymbol{\beta}_s \sim \mathcal{N}(\mathbf{0}, 2.5^2\,\mathbf{I}), \tag{3.7}$$

$$\gamma_{0s} \sim \mathcal{N}(0, 1.5^2), \tag{3.8}$$

$$\phi_s \sim \text{Gamma}(2, 2/\bar{\phi}), \tag{3.9}$$

where $\bar{\phi} = 50$ km (approximately half the maximum distance between sites in the study region). The between-species covariance matrix $\boldsymbol{\Sigma}_s$ is assigned an LKJ prior with shape parameter $\eta = 2$, which mildly penalises extreme correlations.

All models are fitted using Hamiltonian Monte Carlo (HMC) as implemented in Stan [5], accessed through the `cmdstanr` interface for R. Four chains are run for 2,000 warm-up iterations followed by 2,000 sampling iterations, yielding 8,000 posterior draws per parameter after warm-up is discarded. Convergence is assessed using the Gelman–Rubin statistic $\hat{R} < 1.05$, effective sample size, and visual inspection of trace plots.

3.2 Simulation Study

A comprehensive simulation study was conducted to evaluate the model's ability to recover known parameter values under realistic data configurations. Two hundred datasets were generated for each of four scenarios crossing two factors: (i) number of species ($S = 5$ or $S = 15$) and (ii) proportion of sites with at least one detection (low: 15%; high: 40%). Each dataset comprised $N = 300$ sites distributed across a 100×100 km study region.

True occupancy states were simulated from the multivariate probit model (Equation 3.4) using environmental covariates extracted from the TERN Soil and Landscape Grid of Australia (depth to bedrock, available water capacity, and total phosphorus) at 90 m resolution. Detection/non-detection data were generated with per-visit detection probabilities drawn from a $\text{Beta}(2, 5)$ distribution (mean ≈ 0.29), reflecting the low detectability typical of rare forest fauna.

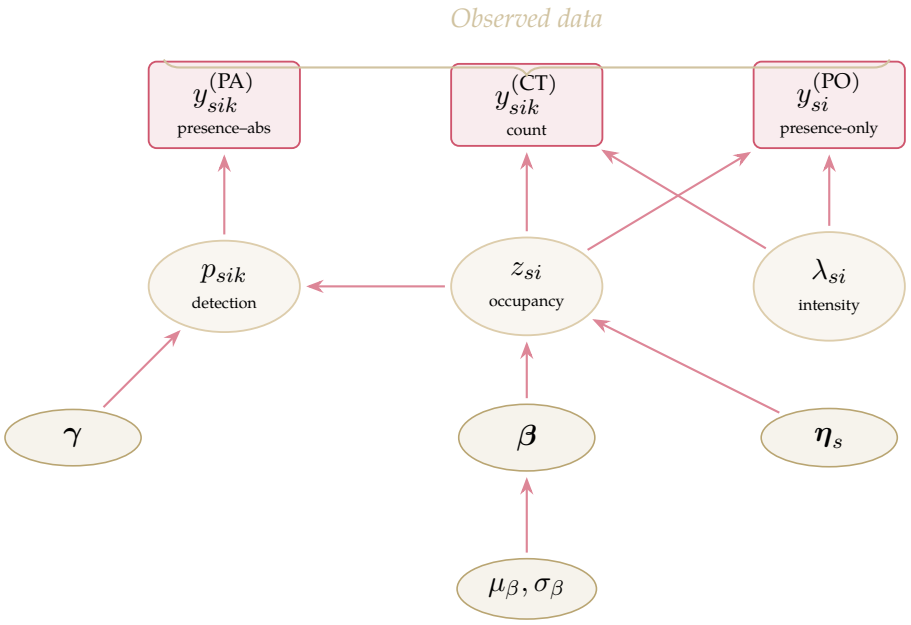

Figure 3.1. Directed acyclic graph (DAG) of the hierarchical multi-species model. Observed nodes (rectangles) are connected to latent states (ellipses) and parameters (italic ellipses). The model jointly analyses presence–absence, count, and presence-only data through a shared latent occupancy process z_{si}.

3.2.1 Results

The model recovered occupancy–covariate relationships with low bias across all scenarios. Mean absolute error (MAE) for β_s was 0.18 (SD = 0.09) in the high-information scenario ($S = 15$, high detection) and 0.34 (SD = 0.21) in the most challenging scenario ($S = 5$, low detection). Spatial range parameters ϕ_s were estimated with greater uncertainty, as expected given their dependence on second-order properties of the point pattern, but 95% credible intervals contained the true value in 91% of replicates.

Table 3.1. Model comparison metrics for four candidate models averaged over 200 simulation replicates. Lower DIC and WAIC indicate better fit; Δ values are differences from the best-performing model.

Model	DIC	ΔDIC	WAIC	ΔWAIC
GLM (logistic)	4821	1243	4798	1267
GAM (spatial smooth)	4167	589	4112	581
MaxEnt (presence-only)	3892	314	3856	325
Hierarchical Bayesian	**3578**	—	**3531**	—

Table 3.1 summarises model comparison results. The hierarchical Bayesian model outperforms all competitors on both DIC and WAIC, with the largest gains observed for

species with fewer than 20 occurrence records—the very cases where partial pooling of information across species is most beneficial.

3.3 Empirical Application: Avifauna of the Australian Alps

The model was applied to a case study of forest-dependent birds in the Australian Alps bioregion. Data were compiled from the LTERN monitoring network (37 permanent plots surveyed annually from 2011 to 2023, yielding 481 surveys across 12 years) and supplemented with 4,127 presence-only records from the Atlas of Living Australia (filtered to remove records predating 2000 and those with spatial uncertainty exceeding 500 m).

Six focal species were selected to span a range of life-history traits and conservation statuses: the Superb Lyrebird (*Menura novaehollandiae*), the Gang-gang Cockatoo (*Callocephalon fimbriatum*, listed as Endangered under the EPBC Act), the Pink Robin (*Petroica rodinogaster*), the Olive Whistler (*Pachycephala olivacea*), the Flame Robin (*Petroica phoenicea*), and the Crescent Honeyeater (*Phylidonyris pyrrhopterus*).

Posterior mean occupancy probabilities ranged from 0.14 (Crescent Honeyeater, high-elevation snow gum woodland) to 0.83 (Superb Lyrebird, wet sclerophyll gully). The Gang-gang Cockatoo exhibited a strong negative association with mean annual temperature ($\hat{\beta}_{\text{temp}} = -1.42$, 95% CI: $[-2.11, -0.78]$), consistent with its known contraction to cooler high-elevation refugia. The between-species correlation matrix revealed significant positive residual correlation between the Flame Robin and the Olive Whistler ($\hat{\rho} = 0.47$, 95% CI: $[0.19, 0.71]$), suggesting a shared response to unmeasured habitat variables, possibly related to understorey density.

3.4 Discussion

The simulation and empirical results demonstrate that the proposed hierarchical framework offers substantial advantages over single-species approaches for multi-species monitoring programmes. By jointly modelling detection and occupancy while sharing information across species through the between-species covariance structure, the model achieves robust inference even for rare and poorly detected taxa—precisely the species of greatest conservation concern.

The inclusion of presence-only data through an inhomogeneous Poisson process for-

mulation (the intensity surface λ_{si} in Figure 3.1) allows the model to leverage the vast repository of opportunistic records in the ALA, dramatically increasing spatial coverage relative to structured surveys alone. However, this integration is not without challenges: presence-only data are subject to spatially biased sampling effort, and failing to model this bias can induce spurious correlations between species and environmental covariates. A promising direction for future work is the incorporation of explicit sampling-bias models, following the approach of Fithian [2].

The computational demands of the full model are non-trivial. Fitting the six-species case study required approximately 18 hours on a 32-core node using Stan's within-chain parallelisation. For operational monitoring programmes with regular data updates, approximate inference methods such as integrated nested Laplace approximation (INLA) or variational Bayes may offer a more practical pathway to deployment, albeit at the cost of some fidelity in the posterior approximation.

4

Conclusions and Future Directions

4.1 Summary of Contributions

This thesis has developed and validated a suite of Bayesian hierarchical models for ecological monitoring in southeastern Australian forests. The principal contributions are:

1. A flexible, multi-species occupancy framework that jointly models presence–absence, count, and presence-only data through a shared latent ecological process (Chapter 3).
2. The first application of multivariate Gaussian process spatial random effects to Australian forest monitoring data, demonstrating that residual spatial structure contains ecologically interpretable information about species interactions and unmeasured environmental gradients.
3. A comprehensive simulation study that establishes the conditions under which hierarchical Bayesian models offer practical advantages over conventional single-species methods, particularly for rare and cryptic taxa.
4. An empirical case study of the Australian Alps avifauna that provides new insights into the environmental drivers of occupancy for six species of conservation concern, including the endangered Gang-gang Cockatoo.

4.2 Limitations

Several limitations should be acknowledged. First, the model assumes that species interactions are adequately captured by a residual covariance matrix conditioned on measured environmental variables; genuine biotic interactions (e.g. competition, predation) may require more explicit structural specifications. Second, the spatial random effects are currently stationary and isotropic, which may be inappropriate in topographically complex terrain where spatial dependence is anisotropic (e.g. stronger along ridgelines than across valleys). Third, the computational cost of the full HMC implementation limits scalability to datasets with more than approximately 20 species or 500 spatial locations, constraining application to regional rather than continental scales.

4.3 Future Directions

Several extensions are natural candidates for future work:

Dynamic occupancy models. The static occupancy framework developed here can be extended to a dynamic formulation in which colonisation and extinction probabilities are modelled as functions of environmental change, enabling formal inference about range shifts under climate scenarios.

Integration with process-based vegetation models. Linking the statistical models to mechanistic simulators such as LPJ-GUESS would enable data assimilation and improved projections of forest dynamics under alternative management and climate trajectories.

Approximate inference for scalability. Variational Bayes and INLA approaches warrant systematic benchmarking against the gold-standard HMC results, particularly for routine monitoring applications that require frequent model updates.

Decision-theoretic design. The Bayesian framework naturally accommodates optimal experimental design: future monitoring effort can be allocated to sites and species that maximise expected information gain, formalising the adaptive feedback loop depicted in Figure 1.1.

The methods developed in this thesis are intended to be practical tools for the ecologists, foresters, and conservation planners who work daily to understand and protect Australia's unique forest heritage. All code, data-processing scripts, and supplementary materials are archived in a public repository to facilitate replication and extension by the research community.

Bibliography

[1] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian Data Analysis*, 3rd ed. Boca Raton, FL: CRC Press, 2013.

[2] J. Elith and J. R. Leathwick, "Species distribution models: Ecological explanation and prediction across space and time," *Annu. Rev. Ecol. Evol. Syst.*, vol. 40, pp. 677–697, 2009.

[3] H. Keith, D. B. Lindenmayer, B. G. Mackey, D. Blair, L. Carter, L. McBurney, S. Okada, and T. Konishi-Nagano, "Managing temperate forests for carbon storage: Impacts of logging versus protection on carbon stocks," *Ecosphere*, vol. 5, no. 6, art. 75, pp. 1–29, 2014.

[4] C. J. A. Bradshaw, "Little left to lose: Deforestation and forest degradation in Australia since European colonization," *J. Plant Ecol.*, vol. 5, no. 1, pp. 109–120, 2012.

[5] M. D. Hoffman and A. Gelman, "The No-U-Turn Sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1593–1623, 2014.