Near-term ecological forecasting: A Bayesian framework for modeling spatio-temporal population dynamics in extreme environments

Hashini Vihanga Goluwa Makkala Gunadasa

Supervisor: Professor Glenda Wardle
Associate Supervisor: Dr Aaron Greenville

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

This research reported in this thesis was supported by the award of an Australian Research Council (ARC) scholarship through the ARC Training Centre in Data Analytics for Resources and Environments (DARE) – Project ID: IC190100031

School of Life and Environmental Sciences
Faculty of Science
The University of Sydney
Australia

28 July 2025

Declaration of Originality

I hereby declare that this thesis is my own work and contains the results of an original investigation, except where sources have been referenced or acknowledged. This research was conducted while I was enrolled as a Doctor of Philosophy candidate in the School of Life and Environmental Sciences at The University of Sydney. This thesis, in whole or in part, has not been previously submitted for any degree at this or any other institution.

Furthermore, I would like to acknowledge that during the preparation of this thesis, AI tools, namely ChatGPT and Writefull, were used for the purposes of text enhancement, including grammar, spelling and review of proper language use. Where any text was modified by generative AI, I then reviewed the resulting content for any errors, inaccuracies or biases, and modified it as required and suitable for the context of the thesis. I take full responsibility for the submitted thesis and ensure the work is my own and has used generative AI within the parameters of use, see University of Sydney generative AI guide for researchers.

Hashini Vihanga Goluwa Makkala Gunadasa
Date : 28 / 02 / 2025

Thesis Abstract

Ecological forecasting plays a crucial role in understanding and predicting ecosystem dynamics, enabling proactive conservation and management decisions. In recent years, near-term ecological forecasting has gained significant attention due to the increasing unpredictability of ecosystems in a changing environment. Forecasting species population abundances is particularly critical, as population dynamics are shaped by complex interactions between biotic (species-species) factors that influence population growth and abiotic (species-environment) factors that impose environmental constraints. These interactions operate across multiple spatial and temporal scales, presenting challenges such as missing data, short time series that may not fully capture species responses to environmental change, and uncertainties from sampling effort and ecological stochasticity; all of which increase the difficulty of accurately forecasting population abundance.

Integrating data science with ecology has become essential to address critical challenges in producing accurate forecasts that support timely and informed decision-making. Advancements in data science offer powerful tools for combining diverse data sources and developing sophisticated modeling frameworks. By using a probabilistic approach to integrate domain knowledge with observational data and state-space models, we can account for uncertainty and interactions across multiple species or locations, thereby improving predictive accuracy. However, it is crucial that these approaches maintain ecological interpretability to ensure their relevance for conservation and management. By combining statistical rigor with ecological principles, data-driven methods can provide actionable insights while preserving the ecological realism necessary for effective decision-making.

As the impacts of climate change intensify, arid ecosystems are emerging at the forefront of ecological uncertainty. These environments are defined by extreme climatic variability, with resource pulses driven by highly unpredictable rainfall events that vary significantly

across both time and space. In addition, the nonlinear interactions between species and environmental factors further complicate ecosystem dynamics, making arid regions particularly vulnerable to climate shifts. Given these complexities, developing robust models to forecast population abundance in arid ecosystems is crucial for guiding conservation strategies and ensuring the resilience of species in these landscapes, particularly as climate change continues to exacerbate environmental instability. The framing of this research project was therefore to examine carefully how we should design, implement and evaluate models when forecasting population abundance.

The first aim of my thesis was to evaluate the robustness and limitations of multivariate auto-regressive state-space (MARSS) models in modeling and forecasting population abundance from both ecological and data science perspectives (Chapter 2). Using a case study that applied MARSS models to infer the abundance of native small mammal populations across nine locations in central Australia over 17–22 years, I examined whether the statistical assumptions of these models align with the ecological processes they aim to represent. The long-term desert mammal live-trapping study conducted by the Desert Ecology Research Group (DERG) used in the case study proved to be a highly valuable resource for ecological forecasting, as it maintained consistent trapping methodologies, broad spatial coverage, and multi-species monitoring. I found that MARSS models assumed a common process across locations when species exhibited similar dynamics across sites, whereas for species with site-specific dynamics, abundance variability was influenced by different processes at each location. Density dependence and two-way species interactions were significant across all species. On the other hand, the white-noise properties of residuals depended on both the length of the time series and species-specific responses to environmental variability and non-normal residuals appeared randomly, with no consistent pattern, which is expected in a highly variable environment. These findings highlight that MARSS models are most robust when their assumptions align with species-specific ecological dynamics, account for spatial variability in biotic and abiotic factors, and incorporate appropriate error structures to reflect environmental uncertainty. However, MARSS models face limitations in forecasting population abundance, particularly in capturing nonlinear interactions, and their computational efficiency declines as the number of species or covariates increases.

The second aim of my thesis was to evaluate how different imputation methods respond to factors such as time series length, missing data patterns, and sample variations, and to determine their impact on the precision of population abundance predictions (Chapter 3). To investigate this, I conducted a simulation study using 17 years of abundance data for a single rodent species at one location, drawn from the long-term DERG dataset analyzed in Chapter 2. Applying three single imputation techniques and a multiple imputation method, I found that in time-dependent population abundance datasets, the choice of imputation method had minimal influence on abundance predictions, regardless of the quantity or temporal placement of missing values. This observation was consistent, even when the compounding effect of missing values was considered in a cumulative approach and the performance of each imputation method was evaluated. As a result, the interpretability and ecological relevance of predicted abundances remained largely unaffected by the selected imputation approach. However, as single imputation methods fail to capture the complexity of ecological time series, multiple imputation methods that account for covariate relationships such as rainfall, more effectively preserve trends and reduce prediction uncertainty.

The third aim of my thesis was to determine the optimal balance between model training length and forecast horizon for forecasting population abundance (Chapter 4). Given the limitations of MARSS models identified in Chapter 2, I employed multivariate generalized additive models (MVGAM) with various trend models to capture spatial and species-level dynamics while incorporating nonlinear environmental and climatic effects specific to each species. For this analysis, I used an extended version of the DERG dataset, including additional data from 2013 to 2022, focusing on two small mammal species with contrasting population dynamics across seventeen sites. By testing two training lengths and three forecast horizons, I found that even models trained on smaller datasets can generate reliable forecasts if they effectively capture varying high-low trends. However, forecast uncertainty increases with longer horizons and for highly dynamic species. By exploring different trend models, I emphasized the importance of balancing computational efficiency with model accuracy. Importantly, I found that incorporating dominant vegetation species as a proxy for broader environmental conditions can significantly reduce the need for exhaustive data collection while maintaining predictive power.

The fourth and final aim of my thesis was to generate near-term forecasts of population abundance (Chapter 5). Building on the analysis in Chapter 4, I used the same extended DERG dataset, species, and MVGAM model to produce real abundance forecasts for two near-term horizons. Since future covariate measurements are unavailable, I incorporated projected rainfall data under different climate change scenarios and assessed the uncertainty in abundance forecasts. I found that highly variable species are particularly sensitive to environmental shifts, with climate change scenarios triggering significant changes in their population dynamics particularly in the short near-term time frame. In contrast, more stable species exhibit muted responses, suggesting greater population resilience across both the horizons. Therefore, evaluating forecast envelopes over near-term time frames is crucial for aligning conservation efforts with the expected temporal scale of ecological change. In addition, I found that the impact of climate change scenarios on forecasts also depends on the availability of observed data, highlighting the need for continued monitoring to improve predictive accuracy. However, throughout the analysis in each chapter, we observed that the significance of environmental and climatic variables affecting the species population abundance varied according to the statistical model used and the range of data used to train the model.

The results of this thesis underscore the importance of integrating data science with ecological principles to enhance the accuracy and interpretability of population abundance forecasts. By critically evaluating modeling approaches, imputation methods, and forecasting strategies, I have demonstrated how data-driven techniques can improve our ability to anticipate population dynamics, particularly in arid ecosystems experiencing rapid environmental change. The findings highlight the need for flexible and ecologically informed models that account for nonlinear interactions, spatial heterogeneity, and uncertainty in environmental drivers. As ecological forecasting continues to evolve, interdisciplinary approaches will be essential for supporting proactive conservation and management efforts, ensuring the resilience of small mammal populations in the face of ongoing climate change. Future research should focus on using more sophisticated statistical models that better capture ecosystem complexities such as lagged dependencies of environmental and climate variables, integrate diverse data sources, such as field-surveyed abundance data with remote sensing

or camera trap data to enhance temporal resolution and incorporate projected vegetation variables alongside climate projections to provide a more comprehensive understanding of habitat dynamics and species responses to environmental change. These advancements will further refine ecological forecasts, supporting more effective conservation strategies in an increasingly uncertain world.

Acknowledgements

The first spark of pursuing a PhD ignited when I was asked to write an article on *'Where would I see myself in 10 years?'* during an English lessons in Grade 10 (2010). Coming from the beautiful yet small island of Sri Lanka, I wouldn't have been able to achieve this dream if not for the unwavering support and guidance of my supervisors, Professor Glenda Wardle and Dr Aaron Greenville. Thank you so much!

Due to travel restrictions of COVID-19, our first meeting took place online via Zoom rather than in person. From that very first day, Professor Glenda made sure that I comfortably learnt about species population dynamics and ecological theory, despite my background being in statistics. My love for ecology today is entirely because of her. Throughout these 3.5 years, her constant encouragement, mentorship and support was always there, even on weekends, holidays, or even field trips with no reception. She always supported me to develop my skills, open my wings to opportunities at the right time, and most of all, always motivated and guided me to be the best version of myself, both professionally and personally. It is difficult to put into words the depth of my gratitude, but I should say that Professor Glenda is one of the best supervisors one could ever ask for.

Dr Aaron's support has been equally invaluable. His guidance in looking at ecology through the lens of statistics was indispensable throughout this journey. Even when Professor Glenda was away at field-work with no reception, Dr Aaron's support was tremendous. Beyond our weekly meetings, I'm so glad to say both of my supervisors were always just an email away, making this a memorable journey.

I would also like to extend my heartfelt gratitude to Professor Robert Kohn at The University of New South Wales, for sharing his statistical expertise during the early stages of my PhD. His advice and knowledge greatly contributed to shaping my statistical knowledge.

A special thanks to the Desert Ecology Research Group and the countless volunteers for their effort in collecting field-data for the past 35 years to get my hands on such a valuable,

real-world dataset. Having been on one of the field trips, I immensely enjoyed the new experience and at the same time truly appreciate the strength of hard work behind this massive, beautifully curated dataset. Special thanks to EcoDynamics Lab led by Dr Aaron for providing the platform to share our project challenges, practice and receive feedback on our posters and presentations.

I have often heard people say being alone in the PhD journey would be a nightmare, but I was fortunate to be a part of a wonderful group at Data Analytics for Resources and Environments (DARE). My sincere thanks to everyone at DARE for making this experience smooth and enjoyable. Special thanks to Dilani Kaveendri for working with me together during the hardest stages of the journey.

I'm very much thankful for all my lecturers at the Department of Statistics at the University of Sri Jayewardenepura, Sri Lanka and all my school teachers for laying the foundation to achieve my goals. A big shout out to Janith Wanniarachchi, my fellow alumnus from the university, for all the support since the first day to overcome all sorts of computational challenges and sharing new techniques and coding knowledge.

To my parents, Mom and Dad, the two pillars of my life. I'm very much grateful for you both for nurturing me to be a strong independent girl to fly away from home and supporting in every step I take to achieve my dreams. Thank you for being open-minded to let me see the world, even as your only child. *'Whatever happens, we are here for you'* was what you both always told me and what led me to fly high with obstacles. I wouldn't be the person whom I'm today, if not for your love, care and all the sacrifices you have done for me.

Not to forget, my partner Dilan for the encouragement and support during all my ups and downs, love, care and always making sure that I had a smile at the end of every call even being 900 miles away.

Have a look at our amazing experience of catching small mammals in the field !!

Author contributions

This thesis is structured as a series of six chapters formatted for publication, an option available to doctoral candidates at The University of Sydney. While these chapters have not yet been published, they are intended for peer review, which may result in some repetition across chapters.

The conceptualization of this study was led by Professor Glenda Wardle and Dr Aaron Greenville, who provided supervision throughout the project, including discussions on study design, data analysis, domain-specific concepts, interpretation of results, and constant constructive feedback to enhance the thesis. In Chapter 2, the statistical methodology was developed with additional guidance from Professor Robert Kohn at The University of New South Wales. The data used in all chapters were provided by the Desert Ecology Research Group (DERG) at The University of Sydney. Extensive data cleaning and pre-processing were required to ensure suitability for analysis. My contributions to each chapter were substantial, and included the execution of conceptualized ideas, data analysis, and drafting and revising all chapters.

As supervisors of this PhD thesis, we concur with the statements above about Ms Gunadasa's contributions.

Professor Glenda Wardle	Dr Aaron Greenville
28 / 02 / 2025	28 / 02 / 2025

Contents

List of Figures

List of Tables

CHAPTER 1

General Introduction

1.1 Introduction

Population dynamics are driven by a complex interaction of biotic and abiotic factors that regulate species abundance and distribution over space and time. Biotic factors, such as density dependence, competition for resources and predation influence the population growth, while abiotic factors such as climate, resource availability and habitat shape populations by imposing environmental constraints (Andrewartha and Birch, 1954; Elith and Leathwick, 2009). These processes function across multiple spatial and temporal scales, resulting in fluctuations that reflect both deterministic and stochastic influences on ecological systems. Despite well-established ecological principles, obtaining accurate estimates of population abundance is often hindered by challenges such as limited data availability, imperfect detection, and the stochastic nature of population dynamics (Tilman et al., 1997). To overcome these obstacles, modeling frameworks serve as powerful tools for integrating diverse data sources, accounting for uncertainty, and disentangling the contributions of biotic and abiotic drivers of population abundance (Ehrlén and Morris, 2015; Iijima, 2020).

Why do ecologists need data scientists? is a question that has gained increasing relevance in recent years due to the challenges of understanding and modeling population dynamics in complex ecosystems (Blair et al., 2019). Long-term ecological data are often sparse, noisy and collected across multiple spatial and temporal scales, making traditional analytical approaches insufficient to capture the underlying patterns and processes (Clark, 2005). Advances in data science, particularly the shift from frequentist to probabilistic approaches and

the ability to integrate domain knowledge with observational data through Bayesian frameworks, provide ecologists with robust tools to analyze heterogeneous datasets, account for uncertainty and improve predictions of species abundance and distribution (Clark and Bjørnstad, 2004; Cressie et al., 2009). Furthermore, computational efficiency is a critical factor in ecological modeling, as timely decision-making depends on models that can estimate population parameters effectively. The integration of data science into population modeling has significantly enhanced computational performance, enabling more efficient and scalable analyses (Levin et al., 1997; Green et al., 2005; Tikhonov et al., 2020; Zurell et al., 2022). As ecological systems become increasingly influenced by extrinsic and intrinsic factors, the fusion between ecology and data science is important to gain a deeper understanding of complex systems and generate accurate forecasts of population dynamics, ultimately informing conservation and management strategies (Green et al., 2005; Miller et al., 2019; Isaac et al., 2020).

1.2 Coherence of ecological and statistical assumptions

Ecologists collect data on the number of individuals in populations over time, and across multiple populations. Unlike data automatically captured by electronic systems (e.g., remote sensing technologies; Rowcliffe and Carbone (2008); Boswell et al. (2017); Lahoz-Monfort and Magrath (2021)), these time series of population abundances are both labour and resource-intensive to collect (Humbert et al., 2009) and are therefore susceptible to various sources of error. Observation error refers to inaccuracies in recorded data that arise during the measurement process, such as field conditions, untrained observers, logistical and financial constraints, misidentifications due to lack of technical expertise, inconsistent sampling effort and human error which occurs because true population sizes are rarely known and are typically estimated rather than directly censused. In contrast, process error captures the intrinsic biological and ecological variability in population dynamics due to factors such as birth, death, dispersal, and environmental fluctuations. Importantly, observation error affects our ability to correctly measure the population, whereas process error reflects real drivers of population fluctuations over time.

State-space models have proven highly effective in addressing these challenges, as they can accommodate missing data and simultaneously incorporate both observation and process errors (De Valpine and Hastings, 2002; Clark and Bjørnstad, 2004; Dennis et al., 2006; Royle and Dorazio, 2008; Auger-Méthé et al., 2021; Lofton et al., 2022). One of the main reasons for the widespread use of state-space models in population dynamics is that the failure to account for observation error will result in biased estimation of abundance and population growth parameters (Hostetler and Chandler, 2015), such as density dependence strength, population trends and extinction risks (Link and Nichols, 1994; Shenk et al., 1998; Dennis et al., 2006; Knape and de Valpine, 2012).

In statistics, state-space models are a class of time series models that rely on assumptions about observation and process errors to ensure statistical validity. In ecology, these models are ideally positioned to describe key dynamics, such as density dependence, predator-prey interactions, and spatial correlations in abundance. For example, De Valpine and Hastings (2002) recommended the numerically integrated state-space (NISS) method, which can handle nonlinear, non-Gaussian dynamics, over least-squares methods for the Beverton-Holt model and, for the Ricker model, unless process or observation error is small and the population dynamics are simple. Building on this, De Valpine (2003) highlighted the value of Monte Carlo state-space likelihood methods to improve inference in population models with complex dynamics or multiple species, demonstrating that they provide more accurate parameter estimates than General Linear Models (GLMs) like ANOVA, particularly in the presence of high process noise, and emphasizing the need for computationally intensive yet ecologically realistic statistical approaches. On the other hand, Humbert et al. (2009) highlighted the superiority of exponential growth state-space models (EGSS) over models that either consider process or observation error particularly when dealing with short time series, missing values, and varying levels of uncertainty. Supporting this, Clark and Bjørnstad (2004) warned that failing to account for missing data, autocorrelated errors, and latent population dynamics can result in incorrect parameter estimation, leading to misrepresentation of population trends and their drivers. Despite these advances, previous research highlights a persistent gap between the ecological assumptions about how populations behave in the real world and the statistical assumptions underpinning the models used to represent those

processes. This lack of coherence refers to instances where the structure, error processes, or dynamics specified in the model do not align with ecological theory or observed system behaviour, potentially leading to biased inference, poor predictive performance, or misinterpretation of ecological processes (Rykiel Jr, 1996). In ecological time series and forecasting applications, ensuring coherence is particularly important because model-based decisions often inform conservation and management actions under uncertainty. When model assumptions diverge from ecological realities, the reliability of forecasts and the trust placed in them can be undermined (Schmolke et al., 2010). This disconnect underscores the importance of further investigation to ensure coherence between statistical and ecological assumptions in population abundance modeling (De Valpine, 2003; Earl, 2019).

1.3 Missing values and uneven sampling intervals

Population abundance time series data often exhibit uneven sampling intervals due to missing observations, which arise from constraints such as budget limitations, logistical challenges, time constraints, and insufficient personnel (Humbert et al., 2009; Bowler et al., 2025). For example, the funding cycles of research projects (typically 3-5 years) often hinder the ability to maintain consistent, long-term monitoring, resulting in sparse datasets (when useful data are rare; mostly zero for true absences and missing values) with only occasional sampling opportunities. Furthermore, uncontrollable factors such as extreme weather events can impede access to study sites, disrupting planned data collection efforts (Ter Braak et al., 1994; Bowler et al., 2025). When analyzing data from multiple sites, missing observations frequently occur at the beginning of a time series if new sites are added mid-project as part of revised research objectives. Sampling priorities and strategies also tend to shift over time to address emerging research questions (e.g., Dickman et al. (2014)). Additionally, the allocation of limited resources across multiple sites reduces the likelihood of sampling every location during each survey, further complicating the consistency and completeness of population datasets.

Missing data in time series can present significant analytical challenges, particularly for models requiring complete data or those sensitive to temporal gaps, such as variants of

Autoregressive integrated moving average (ARIMA) models (Ward et al., 2014). If not handled appropriately, missing observations can introduce biases, reduce the precision of estimates, and compromise the validity of conclusions (Nakagawa and Freckleton, 2008; Van Buuren, 2018). Consequently, a range of statistical techniques have been developed to address missing data by imputing values rather than excluding incomplete observations (Carpenter et al., 2025). These methods include single imputation, multiple imputation, and likelihood-based approaches (Schafer and Graham, 2002; Azur et al., 2011; Van Buuren and Groothuis-Oudshoorn, 2011; Penone et al., 2014). Each approach offers distinct advantages and trade-offs, depending on the characteristics of the dataset and the goals of the analysis.

Despite considerable research comparing imputation techniques, important gaps remain in our understanding of their performance in population abundance time series. One critical question is whether the position of missing values within a time series influences the accuracy and reliability of imputed data. For example, missing values at the beginning, middle, or end of a time series may impact imputation quality differently due to variations in temporal dependencies and trends. Another important consideration is the effect of the number of imputed values on the overall integrity of the time series. Excessive imputation may introduce noise or bias, potentially altering the conclusions drawn from the data. These issues are particularly relevant for species abundance datasets, which often exhibit high variability, small sample sizes, and complex temporal structures (Auger-Méthé et al., 2021; Łopucki et al., 2022). Addressing these gaps requires a thorough evaluation of the interplay between imputation methods, the temporal placement of missing values, and the proportion of missing data. Such efforts are essential to ensure that imputation techniques enhance, rather than compromise, the validity of population abundance models and the ecological insights they provide (Chapon et al., 2023).

1.4 Hindcast, Prediction and Forecast

Accurately modeling population abundance of species over time with the goal of forecasting into the future requires a structured approach to data partitioning, model evaluation, and

uncertainty quantification. A common practice in time-series modeling is to split the available dataset into a training set for model fitting and a testing set for evaluating predictive performance. This ensures that models are assessed on unseen data, preventing overfitting and providing a realistic measure of their generalizability (Dietze, 2017). Once a model is fitted on the training set, the next steps involve a three-step process. First, hindcasting involves making predictions to the training set using the fitted model, which is considered as an internal validation technique to assess whether the model has accurately captured the historical patterns and trends in species abundance. This step is critical, particularly for ecological forecasting, where models must replicate known population dynamics before being trusted for future predictions (Dietze et al., 2018). For this step, various model types have been used in literature ranging from state-space models (Auger-Méthé et al., 2021), generalized additive models (Wood, 2017), and hierarchical Bayesian frameworks (Dorazio and Karanth, 2017). The second step in prediction is to estimate the species abundance in the testing dataset, which consists of observations not used during model fitting. This step allows for an objective evaluation of how well the model generalizes beyond the data it was trained on. Key performance metrics such as the root mean squared error (RMSE) and mean absolute error (MAE) are commonly used to quantify the predictive accuracy of point estimates (Hyndman and Koehler, 2006; Ward et al., 2014). However, with the advancement of statistical techniques, probabilistic modeling approaches have led to the development of scoring rules that evaluate predictions based on their entire distribution rather than single-point estimates. These scoring rules offer flexibility, accommodating both univariate and multivariate models with discrete or continuous response variables. For example, species population abundance, when represented as raw counts, can be evaluated using discrete scoring rules, whereas standardized abundance estimates can be assessed using continuous scoring rules (Gneiting and Raftery, 2007; Simonis et al., 2021). On the other hand, probabilistic modeling techniques (e.g., Basyesian approaches) offer the comfort of quantifying the uncertainty, allowing researchers to assess confidence in abundance predictions, which is crucial in ecological applications where decision-making depends on uncertainty estimates (Clark, 2005). Forecasting, step three of the process, extends beyond within-sample prediction by estimating species abundance into the future, often under changing environmental conditions (the

concept of near-term ecological forecasting; Dietze (2017); Dietze et al. (2018)). Forecasts require not only an accurate abundance model but also projections of covariates, such as temperature, rainfall, and vegetation measurements. This poses additional challenges; 1) future covariates must be obtained from external sources, such as climate model projections Beaumont et al. (2007) and 2) the absence of observed future abundances makes direct evaluation of forecast accuracy impossible, leading to iteratively update the modeling framework as new observational data become available (Dietze, 2017).

Two key challenges in the three-step process are determining the optimal training length on which the models are trained on and forecast horizon — the duration into the future for which predictions remain reliable, mostly tested using the testing set. Previous studies have shown that shorter training lengths may fail to capture long-term dynamics, while overly long training periods might introduce outdated spatio-temporal patterns, reducing forecast accuracy (Ward et al., 2014; Petchey et al., 2015). However, a significant knowledge gap remains in the trade-off between the training length and forecast horizon for population models to acquire accurate predictions. In addition, integrating real climate projections into species abundance models remains an ongoing challenge (Schwartz, 2012). Therefore, robust modeling frameworks are needed to effectively generalize abundance forecasts under a range of plausible future climate scenarios to support decision making processes in conservation and management. What is needed to advance the forecasts of population abundance is a system that experiences rapid environmental change and a suite of species with contrasting life-histories that are sufficiently short-lived to respond with the survey time period. This is what motivated our focus on small mammals in arid Australia.

1.5 Small mammals and arid environments

Small mammals in arid environments often experience dramatic fluctuations in population abundance due to variations in resource availability driven by climatic conditions and interspecies interactions (Kelt, 2011; Meserve et al., 2011; Mason-Romo et al., 2018). However,

unlike arid regions in the Northern Hemisphere, Australian deserts are shaped by unpredictable and irregular extreme rainfall events. These events trigger vegetation pulses that lead to population "'booms", followed by sudden drops ("busts") during extended dry periods (Letnic et al., 2011; Morton et al., 2011; Mason-Romo et al., 2018; O'Connell and Hallett, 2019; Stringer et al., 2024). The Simpson Desert in central Australia, characterized by extreme climatic variability (Van Etten, 2009), is home to small mammal populations exhibiting both highly dynamic and stable patterns, making it an ideal study system for understanding population dynamics and forecasting abundance (Dickman et al., 2010; Letnic et al., 2011; Dickman et al., 2014). Our focus was on the rodents and dasyurids, weighing <500 g, for their contrasting population dynamics across both temporal and spatial scales. Further details of these species are provided in the Thesis Overview below and in Chapter 2.

1.6 Thesis Aims

The specific aims of this thesis and chapters that they are addressed in are as follows:

(1) To investigate the statistical robustness of multivariate state-space models while preserving ecological realism in modeling small mammal population abundance (Chapter 2);
(2) To evaluate the sensitivity of imputation methods to time series length, the quantity and temporal positioning of missing values, and variations across different samples. In addition, to assess the impact of imputation method on the precision and accuracy of population abundance predictions (Chapter 3);
(3) To identify the optimal combination of model training length and forecast horizon for forecasting population abundance using Bayesian multivariate generalized additive models (Chapter 4);
(4) To generate near-term forecasts of population abundance using projected rainfall data under climate change scenarios and quantify the associated forecast uncertainty (Chapter 5).

1.7 Overview of the Thesis

This chapter (Chapter 1) provides a general overview of the key challenges and concepts involved in modeling and forecasting population abundance. This chapter then establishes the general aims of the thesis.

The first aim of my thesis was to validate the coherence between ecological and statistical assumptions in multivariate state-space models (Chapter 2). To achieve this, I used Greenville et al. (2016) as a case study, focusing on native small mammal populations (five species of rodents and dasyurids) of the Simpson Desert (across nine spatially separated sites). I began by examining the characteristics of the study system, including the spatio-temporal dynamics of these populations, and investigating how decades-long (from 1990 to 2012 (23 data points)) field surveys were conducted to collect live capture counts, environmental measurements, and climate data from local weather stations. The chapter details the process of converting manually collected count data into population abundance estimates while addressing challenges posed by varying sampling efforts. I refitted the case study's population state-space abundance model and evaluated the model assumptions to ensure both statistical and ecological validity. Statistically, this includes analyzing residuals; the differences between model predictions and observed counts, to check if they exhibit white noise and follow a normal distribution, ensuring no systematic patterns remain unexplained. Ecologically, the chapter examines how species with different population dynamics balance observation and process errors within the model across different locations within the study site.

The second aim of my thesis was to evaluate the sensitivity of different imputation techniques to various characteristics of missing values in a population abundance time series (Chapter 3). Using the same decades-long dataset from Chapter 2, I focused on the spinifex hopping mouse (*Notomys alexis*) at one of nine study sites. However, since the original dataset contained only 23 data points (1990 – 2012), I conducted a simulation study to ensure a more robust evaluation of imputation methods. Time series of varying lengths (60, 90, 900, and 9000 years) were simulated to reflect the characteristics of the original dataset, with 50 datasets generated for each length to capture the species' diverse "boom-bust" dynamics driven

by rainfall variability. Four imputation methods were examined: the novel multivariate imputation by chained equations (MICE) as a multivariate approach, alongside three single imputation techniques. First, I explored the most suitable simulation length to provide insights into the minimum simulated time series length required for effective modeling. Next, the performance of each imputation method was evaluated across three key characteristics of missing values: (1) the temporal position (beginning, middle, or end of the time series), (2) whether missing values occurred as continuous runs or irregularly and (3) the proportion of missing values. The missing values were imputed cumulatively, taking into account their sequential arrangement (Chapter 3). Additionally, the variation in imputation outcomes across different simulated samples was analyzed. Finally, I assessed the effectiveness of each imputation method in predicting the abundance of *N. alexis* under these varying missing data scenarios, providing a comprehensive evaluation of their suitability for modeling population abundance.

The third aim of my thesis was to identify the optimal combination of model training length and forecast horizon for forecasting population abundance (Chapter 4). This analysis was based on the same study site used in the previous chapters, but now extended to include data from 2013 to 2022. The inclusion of these additional years was crucial, as field surveys have been conducted consistently by the Desert Ecology Research group at the University of Sydney since 1990, providing a long-term dataset that enhances the robustness and relevance of the modeling efforts. First, I performed extensive data cleaning on the new dataset, which included population counts of the species as well as environmental and climate factors. Then, I explored two splits of the training length; 50% and 80% and three forecast horizons; 2 years, 10 years and forecasting for all remaining years. I used this approach to examine the trade-off between using a larger portion of the data for training versus allocating more for testing and to predict both short-term and long-term population trends. For this, I used a novel Bayesian multivariate generalized additive model (MVGAM) (Clark and Wells, 2023; Karunarathna et al., 2024) on two species showing contrasting dynamics; rodent, the sandy inland mouse (*Pseudomys hermannsburgensis*) and small insectivorous dasyurid, the lesser hairy-footed dunnart (*Sminthopsis youngsoni*).

The fourth aim of my thesis was to incorporate insights from the previous chapters and generate near-term forecasts of population abundance (Chapter 5). This chapter extends Chapter 4, using the same dataset, species, and model to obtain abundance forecasts for 4 and 12 years into the future. However, since covariate measurements for future years are unavailable, I used projected rainfall data under different climate change scenarios to assess potential impacts on population abundance and forecast variability. Since observed population counts for future years are unavailable, I evaluated the uncertainty associated with these forecasts to provide insights into making informed decisions on the conservation and management of small mammal populations.

Finally, in Chapter 6, I provide a summary discussion and general conclusions based on my work, placing my findings in the context of the current literature. My key results highlight the importance of aligning statistical model assumptions with ecological relevance to enhance the robustness of population forecasts. I demonstrate that effective imputation techniques can maintain reliability across different patterns of missing data, allowing for accurate long-term predictions even with small datasets and high uncertainty. Furthermore, I show that forecast reliability of population abundance differs across near-term horizons, climate change scenarios, and data availability, underscoring the need for adaptive modeling approaches in ecological forecasting. I conclude the chapter by outlining potential directions for future research.

References

Andrewartha, H. G. and Birch, L. C. (1954). *The distribution and abundance of animals.* Number Edn 1. University of Chicago Press, Chicago, USA.

Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., Leos-Barajas, V., Mills Flemming, J., Nielsen, A., Petris, G., et al. (2021). A guide to state–space modeling of ecological time series. *Ecological Monographs*, 91(4):e01470.

Azur, M. J., Stuart, E. A., Frangakis, C., and Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *International Journal of Methods in Psychiatric Research*, 20(1):40–49.

Beaumont, L. J., Pitman, A., Poulsen, M., and Hughes, L. (2007). Where will species go? incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology*, 13(7):1368–1385.

Blair, G. S., Henrys, P., Leeson, A., Watkins, J., Eastoe, E., Jarvis, S., and Young, P. J. (2019). Data science of the natural environment: a research roadmap. *Frontiers in Environmental Science*, 7:121.

Boswell, A., Petersen, S., Roundy, B., Jensen, R., Summers, D., and Hulet, A. (2017). Rangeland monitoring using remote sensing: comparison of cover estimates from field measurements and image analysis. *AIMS Environmental Science*, 4(1).

Bowler, D. E., Boyd, R. J., Callaghan, C. T., Robinson, R. A., Isaac, N. J., and Pocock, M. J. (2025). Treating gaps and biases in biodiversity data as a missing data problem. *Biological Reviews*, 100(1):50–67.

Carpenter, D., Deyle, E., Park, J., Saberski, E., and Sugihara, G. (2025). Repairing gaps in ecological time series. *Methods in Ecology and Evolution*. doi:10.1111/2041-210X.14491.

Chapon, A., Ouarda, T. B., and Hamdi, Y. (2023). Imputation of missing values in environmental time series by D-vine copulas. *Weather and Climate Extremes*, 41:100591.

Clark, J. S. (2005). Why environmental scientists are becoming bayesians. *Ecology Letters*, 8(1):2–14.

Clark, J. S. and Bjørnstad, O. N. (2004). Population time series: process variability, observation errors, missing values, lags, and hidden states. *Ecology*, 85(11):3140–3150.

Clark, N. J. and Wells, K. (2023). Dynamic generalised additive models (DGAMs) for forecasting discrete ecological time series. *Methods in Ecology and Evolution*, 14(3):771–784.

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., and Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, 19(3):553–570.

De Valpine, P. (2003). Better inferences from population-dynamics experiments using Monte Carlo state-space likelihood methods. *Ecology*, 84(11):3064–3077.

De Valpine, P. and Hastings, A. (2002). Fitting population models incorporating process noise and observation error. *Ecological Monographs*, 72(1):57–76.

Dennis, B., Ponciano, J. M., Lele, S. R., Taper, M. L., and Staples, D. F. (2006). Estimating density dependence, process noise, and observation error. *Ecological Monographs*, 76(3):323–341.

Dickman, C., Wardle, G., Foulkes, J., and Preu, N. d. (2014). Desert complex environments. In Burns, E., Lowe, A., Lindenmayer, D., and Thurgate, N., editors, *Biodiversity and environmental change: monitoring, challenges and direction*, chapter 10, pages 379–438. CSIRO Publishing.

Dickman, C. R., Greenville, A. C., Beh, C.-L., Tamayo, B., and Wardle, G. M. (2010). Social organization and movements of desert rodents during population "booms" and "busts" in central Australia. *Journal of Mammalogy*, 91(4):798–810.

Dietze, M. C. (2017). *Ecological Forecasting*. Princeton University Press.

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., et al. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences*, 115(7):1424–1432.

Dorazio, R. M. and Karanth, K. U. (2017). A hierarchical model for estimating the spatial distribution and abundance of animals detected by continuous-time recorders. *PloS One*, 12(5):e0176966.

Earl, J. E. (2019). Evaluating the assumptions of population projection models used for conservation. *Biological Conservation*, 237:145–154.

Ehrlén, J. and Morris, W. F. (2015). Predicting changes in the distribution and abundance of species under environmental change. *Ecology Letters*, 18(3):303–314.

Elith, J. and Leathwick, J. R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, 40(1):677–697.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378.

Green, J. L., Hastings, A., Arzberger, P., Ayala, F. J., Cottingham, K. L., Cuddington, K., Davis, F., Dunne, J. A., Fortin, M.-J., Gerber, L., et al. (2005). Complexity in ecology and conservation: mathematical, statistical, and computational challenges. *BioScience*, 55(6):501–510.

Greenville, A. C., Wardle, G. M., Nguyen, V., and Dickman, C. R. (2016). Population dynamics of desert mammals: similarities and contrasts within a multispecies assemblage. *Ecosphere*, 7(5):e01343.

Hostetler, J. A. and Chandler, R. B. (2015). Improved state-space models for inference about spatial and temporal variation in abundance from count data. *Ecology*, 96(6):1713–1723.

Humbert, J. Y., Scott Mills, L., Horne, J. S., and Dennis, B. (2009). A better way to estimate population trends. *Oikos*, 118(12):1940–1946.

Hyndman, R. J. and Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688.

Iijima, H. (2020). A review of wildlife abundance estimation models: comparison of models for correct application. *Mammal Study*, 45(3):177–188.

Isaac, N. J., Jarzyna, M. A., Keil, P., Dambly, L. I., Boersch-Supan, P. H., Browning, E., Freeman, S. N., Golding, N., Guillera-Arroita, G., Henrys, P. A., et al. (2020). Data integration for large-scale models of species distributions. *Trends in ecology & evolution*, 35(1):56–67.

Karunarathna, K., Wells, K., and Clark, N. J. (2024). Modelling nonlinear responses of a desert rodent species to environmental change with hierarchical dynamic generalized additive models. *Ecological Modelling*, 490:110648.

Kelt, D. A. (2011). Comparative ecology of desert small mammals: a selective review of the past 30 years. *Journal of Mammalogy*, 92(6):1158–1178.

Knape, J. and de Valpine, P. (2012). Are patterns of density dependence in the global population dynamics database driven by uncertainty about population abundance? *Ecology Letters*, 15(1):17–23.

Lahoz-Monfort, J. J. and Magrath, M. J. (2021). A comprehensive overview of technologies for species and habitat monitoring and conservation. *BioScience*, 71(10):1038–1062.

Letnic, M., Story, P., Story, G., Field, J., Brown, O., and Dickman, C. R. (2011). Resource pulses, switching trophic control, and the dynamics of small mammal assemblages in arid Australia. *Journal of Mammalogy*, 92(6):1210–1222.

Levin, S. A., Grenfell, B., Hastings, A., and Perelson, A. S. (1997). Mathematical and computational challenges in population biology and ecosystems science. *Science*, 275(5298):334–343.

Link, W. A. and Nichols, J. D. (1994). On the importance of sampling variance to investigations of temporal variation in animal population size. *Oikos*, pages 539–544.

Lofton, M. E., Brentrup, J. A., Beck, W. S., Zwart, J. A., Bhattacharya, R., Brighenti, L. S., Burnet, S. H., McCullough, I. M., Steele, B. G., Carey, C. C., et al. (2022). Using near-term forecasts and uncertainty partitioning to inform prediction of oligotrophic lake cyanobacterial density. *Ecological Applications*, page e2590.

Łopucki, R., Kiersztyn, A., Pitucha, G., and Kitowski, I. (2022). Handling missing data in ecological studies: Ignoring gaps in the dataset can distort the inference. *Ecological Modelling*, 468:109964.

Mason-Romo, E. D., Ceballos, G., Lima, M., Martínez-Yrízar, A., Jaramillo, V. J., and Maass, M. (2018). Long-term population dynamics of small mammals in tropical dry forests, effects of unusual climate events, and implications for management and conservation. *Forest Ecology and Management*, 426:123–133.

Meserve, P. L., Kelt, D. A., Previtali, M. A., Milstead, W. B., and Gutiérrez, J. R. (2011). Global climate change and small mammal populations in north-central Chile. *Journal of Mammalogy*, 92(6):1223–1235.

Miller, D. A., Pacifici, K., Sanderlin, J. S., and Reich, B. J. (2019). The recent past and promising future for data integration methods to estimate species' distributions. *Methods in Ecology and Evolution*, 10(1):22–37.

Morton, S., Smith, D. S., Dickman, C. R., Dunkerley, D., Friedel, M., McAllister, R., Reid, J., Roshier, D., Smith, M., Walsh, F., et al. (2011). A fresh framework for the ecology of arid Australia. *Journal of Arid Environments*, 75(4):313–329.

Nakagawa, S. and Freckleton, R. P. (2008). Missing inaction: the dangers of ignoring missing data. *Trends in Ecology & Evolution*, 23(11):592–596.

O'Connell, M. A. and Hallett, J. G. (2019). Community ecology of mammals: deserts, islands, and anthropogenic impacts. *Journal of Mammalogy*, 100(3):1019–1043.

Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E., Graham, C. H., and Costa, G. C. (2014). Imputation of missing data in life-history trait datasets: which approach performs the best? *Methods in Ecology and Evolution*, 5(9):961–970.

Petchey, O. L., Pontarp, M., Massie, T. M., Kéfi, S., Ozgul, A., Weilenmann, M., Palamara, G. M., Altermatt, F., Matthews, B., Levine, J. M., et al. (2015). The ecological forecast horizon, and examples of its uses and determinants. *Ecology Letters*, 18(7):597–611.

Rowcliffe, J. M. and Carbone, C. (2008). Surveys using camera traps: are we looking to a brighter future? *Animal Conservation*, 11(3).

Royle, J. A. and Dorazio, R. M. (2008). *Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations and communities*. Elsevier.

Rykiel Jr, E. J. (1996). Testing ecological models: the meaning of validation. *Ecological modelling*, 90(3):229–244.

Schafer, J. L. and Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, 7(2):147.

Schmolke, A., Thorbek, P., DeAngelis, D. L., and Grimm, V. (2010). Ecological models supporting environmental decision making: a strategy for the future. *Trends in ecology & evolution*, 25(8):479–486.

Schwartz, M. W. (2012). Using niche models with climate projections to inform conservation management decisions. *Biological Conservation*, 155:149–156.

Shenk, T. M., White, G. C., and Burnham, K. P. (1998). Sampling-variance effects on detecting density dependence from temporal trends in natural populations. *Ecological Monographs*, 68(3):445–463.

Simonis, J. L., White, E. P., and Ernest, S. M. (2021). Evaluating probabilistic ecological forecasts. *Ecology*, 102(8):e03431.

Stringer, E. J., Gruber, B., Sarre, S. D., Wardle, G. M., Edwards, S. V., Dickman, C. R., Greenville, A. C., and Duncan, R. P. (2024). Boom-bust population dynamics drive rapid genetic change. *Proceedings of the National Academy of Sciences*, 121(18):e2320590121.

Ter Braak, C., Van Strien, A., Meijer, R., and Verstrael, T. (1994). Analysis of monitoring data with many missing values: which method? *Bird*, 1992:663–673.

Tikhonov, G., Duan, L., Abrego, N., Newell, G., White, M., Dunson, D., and Ovaskainen, O. (2020). Computationally efficient joint species distribution modeling of big spatial data. *Ecology*, 101(2):e02929.

Tilman, D., Lehman, C. L., and Yin, C. (1997). Habitat destruction, dispersal, and deterministic extinction in competitive communities. *The American Naturalist*, 149(3):407–435.

Van Buuren, S. (2018). *Flexible imputation of missing data*. CRC press.

Van Buuren, S. and Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45:1–67.

Van Etten, E. J. (2009). Inter-annual rainfall variability of arid Australia: greater than elsewhere? *Australian Geographer*, 40(1):109–120.

Ward, E. J., Holmes, E. E., Thorson, J. T., and Collen, B. (2014). Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. *Oikos*, 123(6):652–661.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

Zurell, D., König, C., Malchow, A.-K., Kapitza, S., Bocedi, G., Travis, J., and Fandos, G. (2022). Spatially explicit models for decision-making in animal conservation and restoration. *Ecography*, 2022(4).

CHAPTER 2

Ecological relevance and statistical robustness of multivariate state-space models for population abundance modeling

---

2.1 Introduction

Long-term ecological monitoring plays a vital role in providing the foundational data required to understand ecological systems and validate models. Such datasets provide invaluable insights into population dynamics as they allow researchers to capture trends and patterns that emerge over extended temporal scales. They facilitate the detection of trends over an extended period of time, spatial heterogeneity, and interactions between populations and environmental factors. However, obtaining and maintaining long-term datasets is fraught with challenges (Lindenmayer et al., 2012; Kuebbing et al., 2018). Longer (>10 years) and multi-site studies facilitate datasets that often span periods longer than government funding cycles or the tenure of research programs, making sustained support difficult (Lindenmayer and Likens, 2011). As a result, researcher-driven long-term datasets are often built from smaller, short-term studies aggregated over time, rather than being explicitly designed for long-term monitoring, but for an exception of a species abundance dataset collected over 40-years see The Portal Project (White et al., 2019). Alternatives driven by the need to support long-term research funding and infrastructure include nationally-coordinated programs such as the US National Science Foundation Long Term Ecological Reseach (LTER) network which has supported a mix of up to 27 sites since 1980, with 10-year reviews needed (https://lternet.edu/vision-mission/). In Australia, the Terrestrial Ecosystem Research Network (TERN), established in 2012, is an initiative that provides infrastructure for data collection, storage, and sharing across Australian ecosystems. A key component of TERN is its

integration of new and pre-existing long-term terrestrial plot networks, with data spanning from several years to decades across various habitats (Burns et al., 2014). TERN has contributed support to the ongoing DERG datasets since 2012. Other national and international networks that support long-term data collection include the Amazon Forest Inventory Network (RAINFOR; Malhi et al. (2002)), the Landsat archives (Wulder et al., 2012), the National Ecological Observatory Network (NEON; Keller et al. (2008)), and the International Long Term Ecological Research Network (ILTER), which coordinates regional networks across Asia/Pacific, Europe, Africa, North America and Central / South America. These initiatives have the potential to create valuable global datasets for current and future research (Kim, 2006). However, they are relatively recent and have limited temporal coverage, with persistent challenges such as inconsistent methodologies (White et al., 2007), restricted spatial representation, and the absence of comprehensive ecological contexts in publicly accessible data (but see the Global Population Dynamics Database (GPDD), which compiles data on population dynamics across a range of species and ecosystems) (Malhi et al., 2002; Keller et al., 2008; Wulder et al., 2012).

Even with extensive datasets, the observed time series of population abundances cannot be directly equated with true abundances. Observations are affected by two key sources of uncertainty: observation error, which results from measurement inaccuracies and sampling variability, and process error, which arises from the stochasticity inherent in demographic and environmental processes. Historically, models often assumed one of these errors to be constant, a simplification driven by computational limitations (Humbert et al., 2009). The advent of state-space models addressed this issue by allowing researchers to simultaneously account for both types of error, offering a more nuanced and realistic representation of population dynamics (De Valpine, 2003; Buckland et al., 2004; Fukasawa et al., 2013; Iijima et al., 2013; Auger-Méthé et al., 2021; Lofton et al., 2022).

The multivariate autoregressive state-space (MARSS) framework extends state-space models further by incorporating multi-dimensional data, such as time series from multiple sites or species, to improve parameter estimation and provide insights into spatial and temporal dynamics (Holmes et al., 2012; Greenville et al., 2016). This approach is particularly valuable

for identifying spatial heterogeneity and understanding the drivers of sub-population structures, both of which are critical for effective conservation and management. The Bayesian framework of the MARSS model further extends the traditional approach by integrating prior ecological knowledge of the system and providing a more comprehensive quantification of uncertainty. However, the coherence between the statistical assumptions underpinning MARSS models and the ecological realities they aim to represent remains an underexplored aspect in the literature. For instance, how changes in assumptions about observation and process errors such as assuming variable error variances across locations and/or if year-to-year population changes correlate among locations influence model outcomes and whether these models remain statistically sound under varying ecological scenarios are critical questions that require further exploration.

For our study, we used a long-term plot-based monitoring study in the Simpson Desert, central Australia (hereafter referred to as the 'DERG dataset'). This ongoing field-survey focuses on native small mammal populations in a highly variable environment characterized by extremes, making it one of the most challenging datasets to model and forecast (Dickman et al., 2014). Given the increasing variability in global environments, studying ecosystems like the Simpson Desert; prone to environmental conditions such as prolonged drought, floods and wildfires resulting from extreme climate fluctuations, becomes crucial. This dataset exemplifies the typical challenges faced in the collection of ecological data, where limited resources and practical concerns, such as safety and access to study sites, can complicate the process. In addition, the DERG dataset includes time series of varying lengths across different sites due to the gradual expansion of monitoring efforts and the addition of new study locations over time. Despite growing from a combination of short-term targeted projects into a long-term survey, the strength of the DERG dataset lies in its consistent data collection methodology over time and across spatially separated sites (Dickman et al., 2014).

Building upon the work of Greenville et al. (2016), which used the DERG dataset (from 1990 to 2012) and MARSS models in the Bayesian framework, we aim to address the following questions: (1) Does the DERG dataset effectively address the common challenges

faced by other long-term population datasets, and how suitable is it for forecasting population abundance?; (2) How robust are MARSS models in modeling population abundance when assessed from both ecological and data science perspectives?; and (3) What are the limitations of MARSS models in modeling population abundance, and how reliable are they for forecasting?

For question 1, we expect that the DERG dataset has accounted for common challenges in long-term population datasets, such as inconsistencies in data collection, spatial heterogeneity, and missing data. We predict that the dataset's long temporal span, diverse habitat types, and consistent trapping methods across sites will make it well-suited for forecasting population abundance. For question 2, we anticipate that the robustness of MARSS models will be strengthened when the statistical assumptions underlying the models align with the ecological processes they aim to represent. We expect the validity of these assumptions to be influenced by the length of the time series for each species at each site, as well as by the species' responses to varying environmental conditions. However, for question 3, we predict that the MARSS models may face limitations, such as difficulty in capturing nonlinear effects, becoming computationally complex and parameter-heavy as the number of species or covariates increases, and potential biases in forecasts if observational or process errors are misspecified or underestimated.

2.2 Material and methods

2.2.1 Study location

Long-term sampling in the Simpson Desert, Australia, has been conducted by the Desert Ecology Research Group (DERG) at the University of Sydney since 1990 (Wardle and Dickman (2015), Wardle and Dickman (2018a), Wardle and Dickman (2018b), Wardle and Dickman (2018c), unpublished data from 2018). The Simpson Desert covers 170,000 km^2 of mainly dune fields (73%), with the remainder comprising clay pans, rocky outcrops, and gibber plains (Shephard, 1992). The sand dunes, influenced by the prevailing southerly

winds, run in a parallel north-south direction, reaching heights of up to 10 m and spaced 0.6 to 1 km apart (Purdie, 1984). The vegetation of the dune fields is characterized by spinifex grasslands (*Triodia basedowii*) with scattered shrubs (*Acacia*, *Eremophilla*, *Grevillea*), ephmeral forbs and a variety of other perennial and annual grasses; interspersed are wide swales dominated by Georgina Gidgee woodlands (*Acacia georginae*), or mallee eucalypts (Wardle et al., 2015).

The climate of the Simpson Desert exhibits significant variability, alternating between prolonged dry "bust" periods and short yet highly productive "boom" periods driven by summer rainfall (Dickman et al., 2010; Greenville et al., 2013). Daily temperatures usually exceed 40°C in summer and fall below 5°C in winter (Purdie, 1984). Even though highest rainfall occurs in summer, heavy rainfall events may occur on local or regional scales throughout the year (Greenville et al., 2016). Boulia Airport (1886 - 2024), Bedourie (1932 - 2024), and Birdsville Airport (2000 - 2024) are three of the closest long-term weather stations to our study site that have annual average rainfall of 258.5 mm (n = 136 years), 203 mm (n = 79 years) and 157 mm (n = 24 years), respectively (Bureau of Meteorology, 2024).

Vegetation and climatic conditions, which significantly influence species population abundance, are largely shaped by the impacts of wildfires (Letnic, 2004) and cattle grazing (Frank et al., 2013; Tulloch et al., 2023). Since 1990, the region has witnessed two large-scale wildfires (exceeding 1,000 km^2) (Greenville et al., 2009). The mean minimum wildfire return interval is 27 years (updated from Greenville et al. (2009) which reported 26 years). On the other hand, cattle grazing in the eastern parts of the Simpson Desert has been occasional since the mid-twentieth century (Tulloch et al., 2023), and swales have been preferred over dunes due to their relatively higher productivity (Frank et al., 2012). However, changes in land use led to the removal of cattle in 2004 in the Ethabuka Reserve and 2006 in the Pilungah Reserve (Frank et al., 2016). These shifts in land management highlight the dynamic nature of ecosystems and emphasize the importance of long-term studies in capturing the effects of such changes on vegetation and species populations over time.

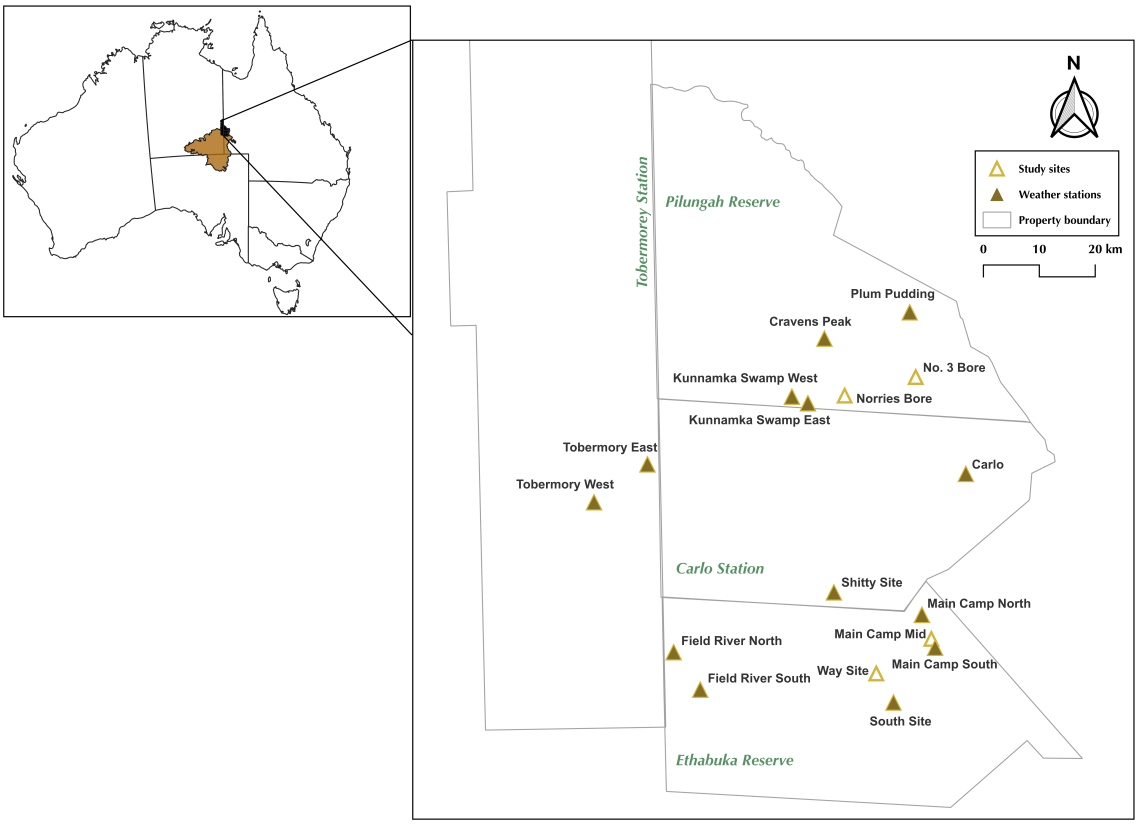

FIGURE 2.1. Seventeen study sites (triangles with outlines) and the approximate weather station (coloured triangles) locations across Pilungah Reserve (previously Cravens Peak Reserve), Tobermorey Station, Carlo Station and Ethabuka Reserve, Simpson Desert, Australia. Inset shows location of study region within the Simpson Desert in Australia.

2.2.2 Small mammals

Live-trapping was carried out at seventeen sites across Carlo Station, Tobermorey Station, Pilungah Reserve (formerly known as Cravens Peak) and Ethabuka Reserve, covering 8000 km^2 of the study location (Fig. 2.1). Small mammals (<500 g) were live-trapped using 36 pitfall traps (16 cm diameter, 60 cm deep) arranged in the form of a grid (plot) comprising of six by six lines of pitfall traps spaced at 20 m apart to cover 1 ha. The top line of traps was placed along a dune crest, while the bottom line was situated 100 m away in the swale, allowing each grid to sample both dune and swale topographies. Each pitfall trap was paired with a 5 m drift fence made of aluminum flywire to improve trap efficiency (Friend et al., 1989). Traps were opened for 2-6 nights on 2-6 times annually from 1990 to 2022 at three sites (Main Camp North, Main Camp Mid and Main Camp South), from 1995 to 2022 at

eleven more sites (Carlo, Cravens Peak, Field River South, Field River North, South Site, Kunnamuka Swamp East, Kunnamuka Swamp West, Tobermorey East, Tobermorey West, Shitty Site and Plum Pudding) and from 2004 to 2022 at No. 3 Bore, Norries and Way Site. Not all sites were sampled each year, thus, when a site was not sampled, it was represented by a missing value in the time series. This included 2020, when sampling was disrupted due to COVID-19 related travel restrictions. Each site in the study contained between 2 and 12 grids, with grid locations randomly selected within the site boundaries, and grid spacing ranging from 0.5 to 2 km.

The number of traps opened (trapping effort) across each sampling trip and site varied due to weather conditions limiting access to certain locations, and some sites were introduced in later years as part of new projects. To account for these differences in the trapping effort, live capture counts (total captures minus recaptures) were standardized to 100 trap nights (100TN). This standardization not only facilitates comparisons with previous studies, which is a key motivation in ecological research, but also simplifies model development by eliminating the need for offset terms. While the standardized captures may result in decimal values, they provide a reliable estimate of total captures, distributing trapping effort equally across sites and sampling trip.

Long-term live-trapping (193 896 trap nights across 34 grids) across 28 - 33 years (151 sampling trips) yielded data on six species of rodents: *Pseudomys hermannsburgensis* (sandy inland mouse; 9521 captures), *Notomys alexis* (spinifex hopping mouse; 6118 captures), *Pseudomys desertor* (desert mouse; 1417 captures), *Rattus villosissimus* (long-haired rat; 1250 captures), *Mus musculus* (house mouse; 1224 captures), and *Leggadina forresti* (short-tailed mouse; five captures), and eight species of dasyurid marsupials: *Dasycercus blythi* (brush-tailed mulgara; 997 captures), *Ningaui ridei* (wongai ningaui; 1122 captures), *Sminthopsis youngsoni* (lesser hairy-footed dunnart; 3106 captures), *Sminthopsis hirtipes* (hairy-footed dunnart; 548 captures), *Sminthopsis macroura* (striped-faced dunnart; 26 captures), *Sminthopsis crassicaudata* (fat-tailed dunnart; 16 captures), *Planigale gilesi* (Giles' planigale; three captures), and *Planigale tenuirostris* (narrow-nosed planigale; 12 captures) (see

Dickman et al. (2014) for more details). Of these thirteen species, only *P. hermannsburgensis* (12 g), *N. alexis* (35 g), *P. desertor* (37 g), *R. villosissimus* (220 g), *M. musculus* (12 - 30 g), the carnivorous *D. blythi* (100 g), the insectivorous *S. youngsoni* (10 g), *S. hirtipes* (15 g) and *N. ridei* (8 g) had sufficient time series data in terms of both length and number of captures across the seventeen sites to support population modeling. The remaining five species had insufficient capture data for modeling, limiting their analysis to observational studies based on available data. Nevertheless, all species reported in sampling highlight the importance of long-term and intensive monitoring in detecting rarer species, as such efforts provide critical insights into population trends and distribution patterns that may otherwise go unnoticed.

2.2.3 Covariate measurements

Small mammal species rely on vegetation cover, seed and flower resources as critical factors that shape their habitat, food availability, and foraging behavior, influencing their distribution and population dynamics (Murray and Dickman, 1994; Murray et al., 1999; Greenville et al., 2016). To assess the cover of the dominant vegetation species, spinifex (*Triodia basedowii*), we visually estimated ground cover as a percentage within 2.5 m radius plots around the same six traps (one trap per row from the 6 x 6 grid) on each small mammal trapping grid in each survey. Furthermore, spinifex seeding was evaluated in each plot using a seed productivity index (ranging from 0 to 5, where 0 represents the absence of seeding and 5 signifies abundant seeding across all plants). We selected spinifex seed for its dominance in the region and its crucial role in the study rodents' diet, making up more than 50% of their diet (Murray and Dickman, 1994; Murray et al., 1999). A flowering index (0–5, where 0 indicates no flowering and 5 signifies abundant flowering across all plants) was also applied to assess the extent of spinifex flowering.

Daily data were collected from automatic weather stations (Environdata, Warwick, Queensland) located at each of the seventeen sites. These weather stations were operational from 1995 to 2022.

2.2.4 Data preparation

The taxonomy of the genus *Dasycercus* was previously unclear until Wolley (2005) clarified that what was once referred to as *D. cristicauda* is now recognized as *D. blythi*. Therefore, for analysis, we adopt the updated naming convention (*D. blythi*) but the DERG dataset retains the original naming (*D. cristicauda*) to maintain consistency across all years and prevent confusion among researchers.

To align our analysis with Greenville et al. (2016), we restricted our dataset to the period from 1990 to 2012, aggregating data at the grid level to the site level. From this, we selected nine sites: Main Camp South (considered as Main Camp), Shitty Site, Field River South, Field River North, South Site, Kunnamuka Swamp East, Carlo, Tobermorey East, and Tobermorey West. Based on capture data up to 2012, only *P. hermannsburgensis*, *N. alexis*, the predatory mulgara *D. blythi*, and the small insectivorous dasyurid *S. youngsoni* and *N. ridei* had sufficient time series data to model species abundance (Greenville et al., 2016).

We incorporated spinifex cover, seed, and rainfall as covariates in the population models. Daily rainfall data from the nearest weather stations at Glenormiston, Sandringham, Boulia, Bedourie and Birdsville were averaged to estimate rainfall data prior to 1995 for Main Camp, where trapping began in 1990. The multivariate autoregressive state-space (MARSS) framework used for modeling population abundance required the time series data to be equally spaced, so we aggregated the data annually. Additionally, capture data were log-transformed (log +1) to align with the model's log-based formulation (Hinrichsen and Holmes, 2009; Ward et al., 2010). All variables were standardized by subtracting the mean and dividing by the standard deviation (z-scored) to facilitate direct comparisons between covariates (Hinrichsen and Holmes, 2009; Ward et al., 2010).

2.2.5 Modeling framework

The MARSS model represents a stochastic exponential growth model (Gompertz population model), where the process component models changes in true abundance on a logarithmic scale, while the observation component, also known as the observation equation, accounts

for measurement error by incorporating additional variation. This framework enables the simultaneous modeling of multiple sites, enhancing parameter estimation and providing insights into density dependence, the spatial structure and environmental factors influencing sub-populations (Holmes et al., 2012). The first-order discrete-time, stochastic Gompertz model for the process component of the MARSS model when $\mathbf{N_t}$ is defined as the raw population abundance at time t is:

$$\mathbf{N_t} = \mathbf{N_{t-1}} \exp\left(\mathbf{u} + \mathbf{b} \ln \mathbf{N_{t-1}} + \mathbf{w_t}\right), \tag{2.1}$$

where $\mathbf{u}$ and $\mathbf{b}$ are constants and $\mathbf{w_t}$ represents the process noise, which follows a multivariate normal distribution with mean 0 and variance-covariance structure $\mathbf{Q}$. The random variables of $\mathbf{w_t}$ are assumed to be uncorrelated.

On the logarithmic scale, the Gompertz process model (Eq. 2.1) is a first-order autoregressive time series model written as:

$$\mathbf{X_t} = \mathbf{u} + \mathbf{B}\mathbf{X}_{t-1} + \mathbf{w}_t; \quad \mathbf{w}_t \sim MVN(0, \mathbf{Q}), \tag{2.2}$$

where $\mathbf{X_t} = \ln \mathbf{N_t}$ and $\mathbf{B} = \mathbf{b} + \mathbf{1}$.

The state vector $\mathbf{X}_t$ represents a $m \times 1$ vector, where m denotes the total number of state values at time t. In this study, it corresponds to all nine sub-populations (the nine selected sites) for each species at time t. Time is considered annual, as observations are aggregated over a year due to unequally spaced data collection. The parameters $\mathbf{B}$, $\mathbf{u}$, and $\mathbf{Q}$ define the state process. The matrix $\mathbf{B}$ allows for interactions between state processes, while $\mathbf{u}$ is the intercept that captures the mean trend. The vector $\mathbf{u}$, which describes the trend of the sub-population, is set to zero due to standardization (Greenville et al., 2016). As exogenous factors (covariates) are incorporated into the process model, eliminating $\mathbf{u}$ shifts the focus to the effects of covariates on sub-population estimates. The diagonal elements of the matrix $\mathbf{B}$ represent the strength of density dependence, where a value less than one indicates density dependence; implying that species abundance is regulated by its own population density.

The off-diagonal elements of $\mathbf{B}$ capture interactions between species populations. Greenville et al. (2016) models only one-way species interactions (to test if *D. blythi* reduces the counts of rodents) but considering a two-way interaction we can also test if the prey numbers (in our case, the rodents) drive the counts of the predator, *D. blythi*. As described in Eq. 2.1, the process errors, $\mathbf{w}_t$ are assumed to be independent and follow a multivariate normal distribution, capturing the variation in population size over time due to environmental and spatial stochasticity. The structure of $\mathbf{Q}$ determines the correlation of process deviations. The diagonal elements of $\mathbf{Q}$ represent process variance, allowing for different variances across sites to account for the spatial differences in process variation. The off-diagonal elements of $\mathbf{Q}$, which describe the correlations between subpopulations, were set to zero (Greenville et al., 2016).

Covariates are incorporated into the process equation (Eq. 2.2) through the elements $\mathbf{C}$, which represent the coefficients of the covariates, and $\mathbf{c}_t$, which represent the covariates over time t. For *P. hermannsburgensis* and *N. alexis*, the model included spinifex cover, the seed productivity index from the previous year (Levine et al., 2024), annual rainfall from the previous year and captures of the predatory *D. blythi*. For the dasyurids *S. youngsoni* and *N. ridei*, the model included spinifex cover, annual rainfall from the previous year and captures of *D. blythi*, which is regarded as both a predator and a competitor. For *D. blythi*, the model incorporated spinifex cover, annual rainfall from the previous year, and total rodent population size (sum of captures of the rodents considered in the study) in the previous year, as rodents are a key food source (see Greenville et al. (2016) for more details). The modified version of Eq. 2.2 with covariates is outlined through Eq. 2.3.

$$\mathbf{X}_t = \mathbf{u} + \mathbf{B}\mathbf{X}_{t-1} + \mathbf{C}\mathbf{c}_t + \mathbf{w}_t; \quad \mathbf{w}_t \sim MVN(0, \mathbf{Q}) \tag{2.3}$$

The multivariate observation component in the MARSS framework is expressed in log space as:

$$\mathbf{Y}_t = \mathbf{a} + \mathbf{Z}\mathbf{X}_t + \mathbf{v}_t; \quad \mathbf{v}_t \sim MVN(0, \mathbf{R}), \tag{2.4}$$

$\mathbf{Y_t}$ represents an $n \times 1$ vector of observations at time t, where in the study, it corresponds to the number of sites at time t, which may differ from the number of sub-populations at that particular time point. $\mathbf{a}$, an $n \times 1$ matrix, represents the mean bias between sites and is set to zero through standardization (Greenville et al., 2016). $\mathbf{Z}$ is an $n \times m$ matrix of zeros and ones, set according to domain knowledge reflecting the synchrony or asynchrony of the sub-populations (spatial population structure). This structure accounts for the possibility that species dynamics may vary or align across locations due to differences in their responses to biological and environmental factors. The observation variability, denoted by $\mathbf{v_t}$, follows a multivariate normal distribution with a mean of zero and a variance-covariance matrix $\mathbf{R}$, which specifies the correlation structure of the observation errors and is assumed to be uncorrelated. The observation error includes sampling error, which can occur due to temporal changes in detectability, or error arising from only a sub-sample of the population being counted. The off-diagonal elements of the variance-covariance matrix $\mathbf{R}$ were set to zero, while the diagonal elements, representing variance across sites, were set to be equal within a species, as the same trapping methods were used throughout the study (Greenville et al., 2016).

According to the synchronicity of the population abundance of the species, two forms of the MARSS model were fitted. A 1-state (synchronous) MARSS model was constructed for the two rodent species and *D. blythi*, while the second model, treating all nine sub-populations as distinct (asynchronous), was fitted for *S. youngsoni* and *N. ridei* (Greenville et al., 2016). The computational specifications were consistent with those used in the case study, implementing a Bayesian approach for inference. For example, prior distributions were assigned as follows: a Uniform distribution for matrix $\mathbf{B}$, Normal distributions for covariates ($\mathbf{C}$) and their coefficients ($\mathbf{c}$), and Gamma distributions for both the process noise covariance ($\mathbf{Q}$) and the observation error covariance ($\mathbf{R}$). Each model was fitted using three Markov chains, with 10,000 MCMC iterations per chain, a thinning interval of 10, and the first 6,000 iterations discarded, resulting in 4,000 retained samples. Although the number of iterations and burn-in period differed from those used in Greenville et al. (2016), we were able to replicate Greenville et al. (2016)'s results.

2.2.6 Examining assumptions

2.2.6.1 Ecological assumptions

Incorporating domain knowledge into modeling often involves setting parameter assumptions to better reflect ecological processes. To assess the validity of these assumptions in the MARSS framework used by Greenville et al. (2016) (hereafter referred to as the 'Original Model'), we fitted the same model without predefined constraints. If the assumptions underlying the Original Model are valid, then both models; constrained and unconstrained, should yield similar results. For example, if relaxing assumptions on parameters in Eq. 2.3 and Eq. 2.4 leads to comparable outcomes, this suggests that the Original Model appropriately captures the underlying ecological dynamics. However, if discrepancies arise, they may indicate that the data suggest a different underlying ecological process. By comparing the two models, we can assess whether the imposed ecological assumptions accurately capture species population dynamics or if a more flexible approach is needed to better represent the system. Importantly, such a distinction is only reliable when the data are relatively complete, unbiased, and precise. If the data are sparse, noisy, or biased, it becomes difficult to discern whether differences between models stem from misspecified assumptions or data limitations. This evaluation helps to ensure that the model provides both statistical robustness and ecological realism.

The MARSS framework incorporates three state process parameters; $\mathbf{B}$, $\mathbf{Q}$, and $\mathbf{R}$, each subject to specific assumptions. Relaxing these assumptions results in seven variants of the Original Model, with each variation reflecting a distinct ecological reasoning (Fig. 2.2).

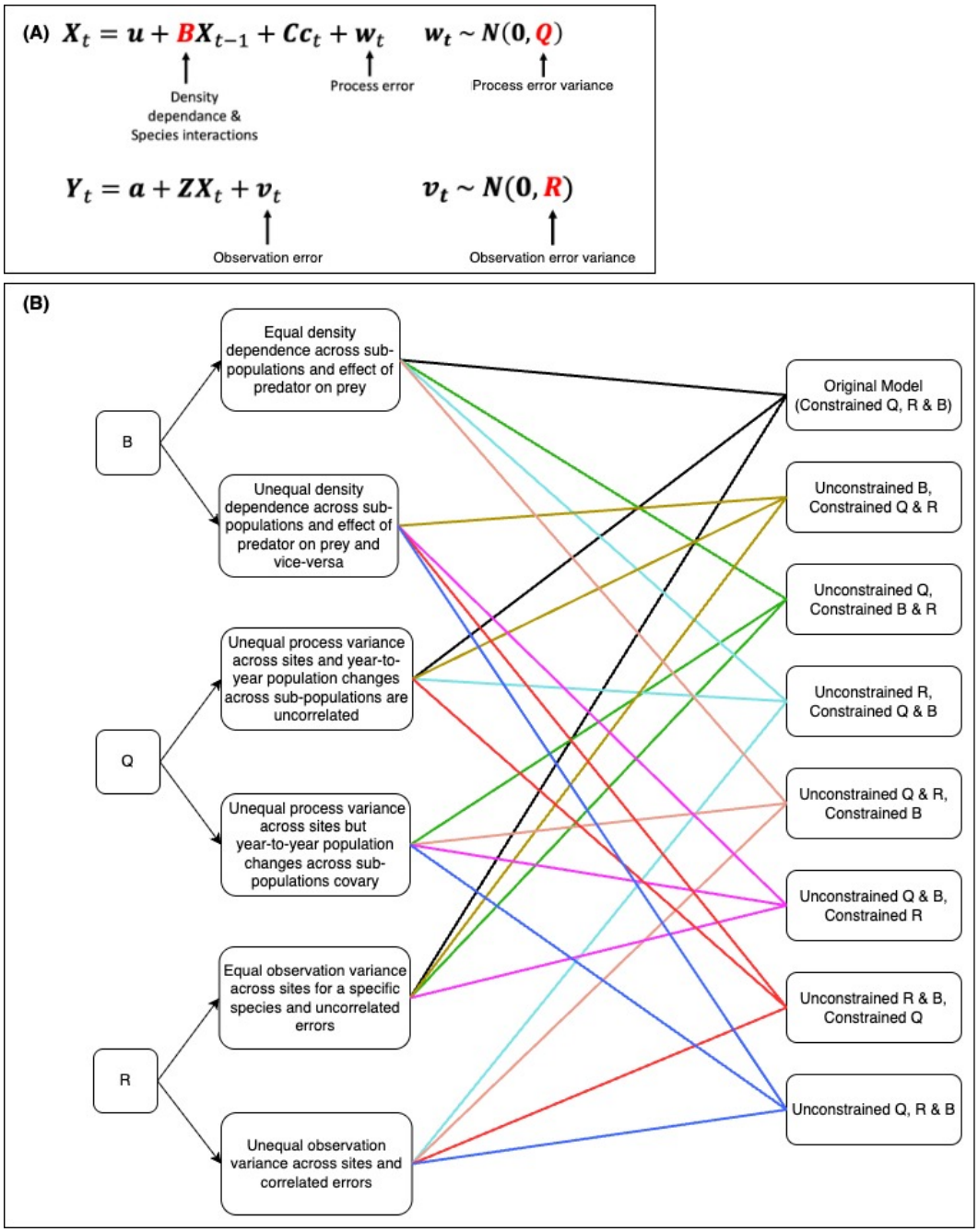

FIGURE 2.2. Variations in ecological assumptions of multivariate autoregressive state-space model (MARSS). (A) $\mathbf{B}$ and $\mathbf{Q}$ are parameters of the process equation while $\mathbf{R}$ is a parameter in the observation equation of the MARSS framework. (B) Combinations of models constructed by applying constraints to the state process parameters. For each $\mathbf{B}$, $\mathbf{Q}$, $\mathbf{R}$ node, the upward arrow indicates the ecological perspective for restricting the parameter, while the downward arrow represents the effect of relaxing the assumption. Lines of the same colour represent the incorporation of constrained and unconstrained matrices. In total, seven alternative model variants (excluding the Original Model) were fitted to evaluate the robustness of the ecological assumptions.

The Akaike Information Criterion (AIC) is widely used to assess the predictive power of statistical models by balancing goodness of fit and model complexity (Akaike, 1998). To address AIC's bias in small sample sizes, the corrected Akaike Information Criterion (AICc) was introduced as an extension (Sugiura, 1978; Hurvich and Tsai, 1989). The AICc is calculated by first computing the standard AIC as:

$$\text{AIC} = 2k - 2\log L,$$

where k is the number of estimated parameters and $\log L$ is the log-likelihood of the model. The AICc then adjusts the AIC for small sample sizes using the following formula:

$$\text{AICc} = \text{AIC} + \left(\frac{2k^2 + 2k}{n - k - 1}\right),$$

where n is the number of observed (non-missing) data points. This correction accounts for potential bias in model selection when the sample size is small relative to the number of parameters. Given the relatively limited sample size in the case study (17 - 22 years of data for each species), AICc provides a more reliable metric for model selection. We use AICc to compare the seven variants (Fig. 2.2) with the Original Model, evaluating how well each model captures variation in the species dynamics. A lower AICc value indicates a model with better predictive performance.

2.2.6.2 Statistical assumptions

The MARSS framework assumes that the residuals from the observation model (known as model residuals and hereafter referred to as 'residuals') should follow an independently distributed white noise pattern (i.e. residuals that have a mean of zero, constant variance, and no autocorrelation) (Holmes et al., 2014). To test this assumption, we used the Ljung-Box test (Ljung and Box, 1978) to asses whether the residuals of the Original Model exhibit a white noise process (Mahan et al., 2015). The null and alternative hypotheses for this test are as follows:

H_0 : The series exhibits a white noise pattern

H_1 : The series does not exhibit a white noise pattern

The residuals should not only follow a white-noise process but also exhibit a normal distribution. To assess normality, we used the Shapiro-Wilk Normality test (Shapiro and Wilk, 1965) for each species at each site and the null and alternative hypotheses are:

$$H_0 : \text{The data are normally distributed}$$

$$H_1 : \text{The data are not normally distributed}$$

To evaluate the independence of residuals, we generated Autocorrelation function (ACF) plots as a graphical representation.

Statistical coherence of the Original Model is crucial to confirm the validity of the ecological assumptions. However, a minor violation of the assumptions regarding the residuals being white noise and normally distributed does not necessarily invalidate the results. In such cases, the model results can still be interpreted with caution, provided the likelihood of a type II error remains low (Shreffler and Huecker, 2023). For this study, we set the threshold for acceptable deviation from the assumptions at 5%.

2.3 Results

2.3.1 Assessing the suitability of the DERG dataset for forecasting population abundance

We first pre-processed the raw data and reproduced the results of Greenville et al. (2016). Through this process, we understood the value of the consistent methodological structure of the DERG dataset, which resulted in population abundance predictions consistent with the case study. The sampling of multiple grids, which effectively produced replicate sampling, helped to mitigate observation error, which is a crucial factor for ensuring data quality. One of the most important features of the DERG dataset is its extensive temporal coverage since 1990, coupled with the high spatial scale of sampling. In addition, not only are species counts

of a variety of species recorded, but other covariate measurements are also consistently collected throughout the years. This consistency ensures that all relevant data is captured within the study, eliminating the need to rely on external covariate data that may have different methodologies or be subject to different biases. This combination of temporal, spatial, and methodological consistency makes the DERG dataset particularly valuable for modeling and forecasting long-term trends in species abundance. As a result, it provides a solid foundation for addressing the objectives outlined in the subsequent chapters of this thesis.

The R code used to implement the pre-processing steps and the MARSS modeling framework was then reorganized and refactored into functions, making the code more generalizable and reproducible for other studies with similar objectives. The code can be provided on request.

2.3.2 Validating ecological assumptions

The best model to predict the population abundance of the species was one of the seven variants of the Original Model, as indicated by the comparison of AICc values (Fig. 2.3; AICc values of all seven variants are shown in Table A1.1). For *P. hermannsburgensis*, *N. alexis* and *D. blythi*), the single population MARSS models assumed different observation errors across sites (Unconstrained $\mathbf{R}$ & $\mathbf{B}$, Constrained $\mathbf{Q}$, but note there is only a single estimate for $\mathbf{Q}$ because this is a single-population MARSS model). In contrast, the nine independent sub-population MARSS models fitted for the two species of small dasyurid marsupials (*S. youngsoni* and *N. ridei*) assumed different process errors across sites (Unconstrained $\mathbf{Q}$ & $\mathbf{B}$, Constrained $\mathbf{R}$). This distinction reflects the species' behaviour: species exhibiting similar dynamics across space assumed the same process across sites, with the variability in abundance primarily arising from sampling. In contrast, species that showed different dynamics across sites were assumed to have different processes at each site, where variability in abundance was less influenced by sampling error (Fig. 2.3 & Table A1.2 & A1.3). In addition, the density dependence was significant in both model structures (Table A1.2 & A1.3) and the nine independent sub-population MARSS models assumed the density dependence to vary

across sites (Table A1.3). However, two-way interactions between species (effect of predator on prey and vice-versa) appeared to be non-significant for all study species, regardless of their population synchrony across space (Table A1.2 & A1.3). The significance of the covariates in the single-population MARSS models for *P. hermannsburgensis* and *N. alexis* were similar to those of the Original Model, with previous year rainfall and spinifex seed being significant. However, in the single-population MARSS model for *D. blythi*, only previous year rainfall was significant, contrasting with the Original Model where in addition, spinifex cover and rodent population effects were significant (Table A1.2). In contrast, none of the covariates in the nine independent sub-population MARSS models fitted for both *S. youngsoni* and *N. ridei* were significant (Table A1.3). Although these observations differ from the Original Model (Greenville et al., 2016), the statistical validity of these models highlights unique model specifications that are more closely aligned with the population dynamics of each species and provide greater ecological relevance.

We observed that the number of parameters to be fitted increased with the relaxing of the assumptions where the model with Unconstrained $\mathbf{B}$, $\mathbf{Q}$, and $\mathbf{R}$ was equipped with the largest number of parameters to be estimated (Fig. 2.2). However, it is important to note that models with a higher number of parameters tend to yield a lower AICc, given their greater capacity to represent the data, which introduces a caveat in the interpretation of the results.

Species	AICc of Original Model	Selected model variant out of the 7 variants of the Original Model	AICc of selected model variant	Best model to predict population abundance (Type 1 or Type 2 or Original Model)
(A)  *Pseudomys hermannsburgensis*	889.716	**Unconstrained R** (assumes unequal observation variance across sites and correlated errors) **and B** (assumes unequal density dependence across sub-populations and effect of predator on prey and vice-versa) but **Constrained Q** (assumes unequal process variance across sites and year-to-year population changes across sub-populations are uncorrelated) - **TYPE 1**	-1870.791	Type 1
(B)  *Notomys alexis*	877.494		-1884.073	Type 1
(C)  *Dasycercus blythi*	924.228		-1894.779	Type 1
(D)  *Sminthopsis youngsoni*	-112.241	**Unconstrained Q** (assumes unequal process variance across sites and year-to-year population changes across sub-populations covary) **and B** with **Constrained R** (assumes equal observation variance across sites for a specific species and uncorrelated errors) - **TYPE 2**	-938.529	Type 2
(E)  *Ningaui ridei*	-84.011		-1004.246	Type 2

FIGURE 2.3. AICc of the Original Model and selected model variant out of the seven variants of the Original Model for the five study species. The model selected from the seven variants is based on the lowest AICc value for each species. Constraints of the Original Model are: **Constrained** Q (assumes unequal process variance across sites and year-to-year population changes across sub-populations are uncorrelated, but note for single-population MARSS models there are only single estimates for Q), **Constrained** B (assumes equal density dependence across sub-populations and effect of predator-on prey) and **Constrained** R (assumes equal observation variance across sites for a specific species and uncorrelated errors). AICc values are compared across each row, with the model yielding the lowest AICc selected as the best model (out of the Original Model and the selected model variant) to predict population abundance for each species.

2.3.3 Testing statistical assumptions

The results from the Ljung-Box test indicate that the assumption of model residuals following a white-noise process is violated for certain species at specific sites. This suggests that the time series of population abundance varies across locations due to differences in how each species responds to local environmental factors. For example, in the 1-state MARSS model fitted for *N. alexis*, the white-noise assumption was violated at Tobermorey East. Similarly, in the nine-independent sub-population MARSS models for *S. youngsoni*, the violation occurred at Field River North, and for *N. ridei*, it was observed at both Kunnamuka Swamp East and Tobermorey East (Table 2.1). At Main Camp; the site with the longest time series

for all species, the model residuals followed a white noise process, with higher p-values from the Original Model. This suggests that longer time series may capture a more comprehensive representation of how species respond to environmental factors, while shorter time series may only reveal partial responses, as indicated by the significant p-values (in bold; Table 2.1) observed for species at other sites. Even at sites with shorter time series, not all species exhibited uncaptured temporal dependence in the residuals (Table 2.1). This discrepancy suggests that, even during boom or bust periods, species may respond differently to extreme environmental conditions across sites, leading the model to fail in capturing the peaks or troughs in the time series and leaving unexplained variation in the residuals. The ACF plots obtained based on the residuals of the observation equation of the Original Model confirmed these observations as the plots indicated correlations within the residuals (see Fig. A1.1). These violations indicate that the MARSS models for these species at the respective sites have not fully captured the time-dependent structure of the data. Consequently, further model refinements, such as relaxing ecological assumptions of state process parameters or incorporating additional covariates, may be required to improve the model fit and better account for these temporal dependencies.

Additionally, the results of the Shapiro-Wilk test for normality of model residuals reveal violations of the normality assumption for several species at various sites. Specifically, for the 1-state MARSS models fitted for *P. hermannsburgensis*, normality was violated at Carlo and Tobermorey West and for *N. alexis*, the normality assumption was violated at Field River North and Tobermorey West. In the nine-independent sub-population MARSS models fitted for *D. blythi*, normality was violated at Main Camp, Kunnamuka Swamp East, Tobermorey East and Tobermorey West, and for *S. youngsoni* showed similar violations at Main Camp (Table 2.1). These violations are likely due to population booms or busts that result in outliers or high-leverage points that extend the tails of the population abundance distribution. Such heavy-tailed values lead to residuals deviating from normality, as extreme fluctuations generate large residuals that the normal distribution fails to accommodate. This suggests that the underlying ecological processes may involve nonlinear dynamics or abrupt population shifts that are not fully captured by the model's normality assumptions. Consequently, these

deviations may affect the validity of statistical inferences relying on normality. It is also important to note that the data were aggregated annually, and the relatively small sample sizes at each site may have contributed to these violations. Normality tests, such as the Shapiro-Wilk test, are sensitive to sample size, and with limited data points, the test becomes more susceptible to Type II errors, incorrectly concluding that residuals follow a normal distribution when in reality they do not. Therefore, caution is needed when interpreting results based on the assumption of normality, as smaller sample sizes increase the likelihood of such errors (Sánchez-Espigares et al., 2018).

TABLE 2.1. The results of the Ljung-Box test and Shapiro-Wilk test for white noise and normality on the residuals of the MARSS observation models, respectively. Single-population MARSS models were fitted for three small mammal species (*Pseudomys hermannsburgensis*, *Notomys alexis*, and *Dasycercus blythi*), while nine independent sub-population MARSS models were fitted for two small dasyurid marsupials (*Sminthopsis youngsoni* and *Ningaui ridei*). P-values in bold indicate cases where the assumptions of white noise or normality were not met.

Site	**Main Camp**	**Carlo**	**Field River South**	**Kunnamuka Swamp East**	**Shitty Site**	**South Site**	**Field River North**	**Tobermorey East**	**Tobermorey West**
White-noise residuals									
P. hermannsburgensis	0.79	0.53	0.89	0.66	0.91	0.87	0.42	0.63	0.42
N. alexis	0.75	0.23	0.37	0.59	0.46	0.81	0.38	**0.03**	0.93
D. blythi	0.40	0.43	0.33	0.96	0.67	0.70	0.19	1.00	1.00
S. youngsoni	0.64	0.89	0.59	0.44	0.86	0.09	**0.03**	0.92	0.81
N. ridei	0.69	0.21	0.42	**0.02**	0.72	0.21	0.60	**< 0.01**	0.76
Normality of residuals									
P. hermannsburgensis	0.58	**0.01**	0.95	1.00	0.28	0.13	0.74	0.10	**< 0.01**
N. alexis	0.43	0.16	0.33	0.44	0.82	0.83	**0.05**	0.72	**< 0.01**
D. blythi	**< 0.01**	0.07	0.99	**< 0.01**	0.16	0.44	0.34	**< 0.01**	**0.01**
S. youngsoni	**0.05**	0.20	0.45	0.48	0.71	0.06	0.08	0.57	0.27
N. ridei	0.46	0.48	0.68	0.92	0.73	0.16	0.06	0.88	0.72

2.4 Discussion

Understanding and forecasting population dynamics requires robust models that accurately capture ecological complexities. However, the effectiveness of these models depends on the availability of long-term high-quality datasets that accurately represent key ecological factors, including environmental drivers, stochastic fluctuations, and spatial variability. The Desert Ecology Research Group (DERG) dataset offers a valuable resource to address these challenges, allowing detailed investigations of species abundance patterns over time. This study evaluated the suitability of the DERG dataset for forecasting, assessed the robustness of MARSS models to capture population dynamics, and identified key limitations that influence their applicability for ecological forecasting. We found that the MARSS framework is not effective for forecasting population abundance due to several key limitations: its inability to capture nonlinear ecological processes, the requirement for equally spaced time intervals, increased computational demands with higher-dimensional models, and challenges in handling zero species counts.

2.4.1 Evaluating the DERG dataset's suitability for forecasting population abundance

Arid systems in Australia are inherently unpredictable (Letnic and Dickman, 2010; Letnic et al., 2011; Morton et al., 2011), due to their high variability, which makes forecasting in these environments particularly challenging. Collected from study sites across the Simpson Desert, the DERG dataset provides a detailed record of species abundance under extreme environmental conditions. Based on a conceptual model of the ecosystem (Dickman et al., 2014), the dataset adheres to sound statistical methodologies for data collection, maintains high-quality records, and includes multiple species to capture the variability in dynamics. In doing so, the DERG dataset fulfills the key criteria for an effective long-term ecological study.

Despite the dataset's strength in its extensive temporal range, covering multiple decades (since 1990), and its sampling conducted at multiple sites; effectively producing replicate

samples, its value is further enhanced by the reduction of observation error. This replication improves statistical modeling and enables robust predictions of species abundance across spatial scales. Unlike many long-term population datasets, which often suffer from inconsistent methodologies and a lack of comprehensive ecological contexts to support ecologically meaningful analysis by data scientists (Likens and Lindenmayer, 2018), the DERG dataset employs standardized trapping methodologies from its inception, ensuring data consistency and providing a thorough description of the ecological context.

Despite these advantages, challenges remain, particularly in relation to missing data, which is an inevitable issue in long-term ecological studies. In the context of the DERG dataset, missing data arises due to two main factors: 1) logistical and financial constraints, which can limit the frequency of data collection efforts, and 2) road closures caused by extreme rainfall, which impede access to study locations and prevent data collection during critical periods. The latter issue can occur even when researchers are already at the study site, as extreme weather conditions may restrict access to certain study locations, leading to missing values for some sites in specific years. However, unlike many other long-term datasets that often lack clarity on the causes of missing data, the DERG dataset provides a clear understanding of the reasons behind these gaps. Therefore, the DERG dataset provides a comprehensive and transparent long-term record of population dynamics, making it a highly valuable resource for accurate ecological forecasting.

Therefore, to achieve the aims of the subsequent chapters in the thesis, which focus on forecasting species population abundance, we will use the DERG dataset as the foundational data source. Following this chapter, Chapter 3 will focus on the same data period analyzed in the case study considered in this chapter (Greenville et al., 2016), while Chapters 4 and 5 will extend the analysis incorporating an additional 11 years of data collected since Greenville et al. (2016) to address their respective objectives.

In considering the potential for forecasting population abundance, we evaluated the suitability of other data sources to enhance model accuracy and predictions. The Portal Project, a long-term ecological study ongoing since 1977 in the Chihuahuan Desert near Portal, Arizona, USA, offers valuable data for this purpose (Karunarathna et al., 2024; Clark et al.,

2025). Although data on rodent and plant species abundance, as well as climatic factors such as air temperature and precipitation, have been collected at a higher frequency than the DERG dataset, inconsistencies in data collection methodologies over time pose challenges for reliable forecasting (Ernest et al., 2018). While White et al. (2019) developed an automated forecasting pipeline for rodent abundances, no attempts have been made yet to generate forecasts projecting into the future, but this may be forthcoming (Karunarathna et al., 2024; Clark et al., 2025). Therefore, considering the high spatial and temporal resolution and diverse range of data collected, the Portal Project offers a promising avenue for future research in forecasting population abundance.

2.4.2 Assessing the robustness of MARSS models for population abundance

MARSS models offer a powerful framework for analyzing multivariate time series data in ecological studies. These models explicitly account for both process and observation errors, making them well-suited for noisy ecological datasets (Holmes et al., 2012). However, their effectiveness depends on the validity of the underlying ecological assumptions and the extent to which they capture species-specific population dynamics.

Expanding the ecological assumptions tested on MARSS models by Greenville et al. (2016), our analysis of the DERG dataset revealed that MARSS models fitted to population abundance data accounted for site-specific errors that varied depending on species dynamics. For species exhibiting synchronous population fluctuations across study sites, such as the two rodent species *Pseudomys hermannsburgensis* and *Notomys alexis*, as well as the predatory mulgara *Dasycercus blythi*, the models assumed different observation errors, likely reflecting variations in detectability rather than differences in true population processes. Conversely, for species exhibiting asynchronous dynamics, such as the small dasyurid marsupials *Sminthopsis youngsoni* and *Ningaui ridei*, the models assumed different process errors. In addition, although previous year rainfall and spinifex seed significantly influenced the abundance of *Pseudomys hermannsburgensis* and *Notomys alexis*, the abundance of *Dasycercus blythi* was only affected by last year's rainfall. Surprisingly, none of the covariates were found to be

significant at any site for *Sminthopsis youngsoni* and *Ningaui ridei*, which contrasts with the findings of Greenville et al. (2016). However, density dependence appeared significant among all species regardless of their different dynamics across space. Therefore, our results suggest that both intrinsic and extrinsic factors, along with sampling and environmental variability, play dominant roles in shaping the population dynamics of species with contrasting dynamics (Lawton, 1988; Kent et al., 2007; Mutshinda et al., 2009).

Another key consideration is the trade-off between ecological interpretability and statistical robustness. MARSS models are effective in capturing species' responses under stable conditions, but extreme fluctuations, such as population booms, may challenge statistical assumptions. A key issue arises when temporal dependencies are not fully accounted for, leading to autocorrelated residuals that indicate unmodeled variability in the system (Holmes et al., 2012). Our analysis suggests that the white-noise properties of residuals depend on both the length of the time series and species-specific responses to environmental variability across spatial scales. For instance, at Main Camp; the site with the longest continuous time series, residuals exhibited white-noise properties for all species, suggesting that the MARSS models effectively captured the underlying population dynamics. However, at other sites with shorter time series, not all species exhibited white-noise residuals, indicating that some ecological processes remained unexplained. Interestingly, some species still showed white-noise residuals despite differences in data length, likely due to asynchronous population booms across sites. In other words, when population surges occur in some locations but not others, spatial variability in species responses may obscure temporal autocorrelation at individual sites (Dennis et al., 2006; Thibault and Brown, 2008). Additionally, if a species, such as *Dasycercus blythi*, is included at a site where it is absent, the model fails to incorporate meaningful explanatory power from this species, leading to unexplained variability. In such cases, the inclusion of irrelevant covariates may introduce additional noise without contributing to predictive accuracy (Mutshinda et al., 2009). This underscores the importance of considering both spatial and temporal heterogeneity when modeling population dynamics using MARSS.

Moreover, extreme population fluctuations may also lead to non-normal residual distributions, particularly in species with highly skewed or zero-inflated abundance data. MARSS models assume normal observation errors, but sudden booms or busts in population abundance can result in heavy-tailed or asymmetric residual distributions. This phenomenon is often observed in field-surveyed ecological data from highly variable environments, where ecological processes are complex and subject to sudden shifts. In our analysis, we observed non-normal residuals only at a few sites and for a limited number of species, with no consistent pattern emerging across the dataset. Therefore, while it is not advisable to completely disregard predictions or inferences, careful interpretation is required when faced with such non-normal residuals (Dennis et al., 2006; Thibault and Brown, 2008).

2.4.3 Identifying limitations of MARSS models and their applicability for forecasting

While MARSS models provide a structured approach to modeling population dynamics, several limitations affect their forecasting utility. First, they require observations at equally spaced time intervals. Aggregating data annually and inserting rows for missing observations can degrade the signal, introducing additional noise through processing, averaging, and gap filling. Therefore, developing methods that explicitly model uneven observation intervals would enhance the model's applicability in real-world ecological settings. Second, MARSS models struggle to capture nonlinear ecological processes, particularly those driven by environmental and climatic factors. This limitation is particularly relevant in systems characterized by abrupt population changes, where traditional linear state-space formulations may fail to capture key transitions (Newman et al., 2014). Third, these models become computationally intensive as the number of species or covariates increases. Parameter estimation in high-dimensional state-space models requires substantial computational resources, making it challenging to scale analyses across large ecological datasets (Dennis and Ponciano, 2014). Moreover, if the observation or process error is misspecified or underestimated, forecasts can become biased, leading to misleading ecological interpretations. In highly variable

environments, meeting model assumptions to ensure statistical robustness can be particularly difficult, further limiting the generalizability of forecasts. Lastly, the common use of the $\log(x + 1)$ transformation to accommodate zero counts introduces another issue. While this transformation avoids the problem of taking the logarithm of zero, it breaks the direct link to the original population model. A more principled alternative would be to model latent abundance on the log scale and use a Poisson model to link to observed counts, which naturally accommodates zeros and avoids the need for the transformation.

To address these challenges, generalized additive models (GAMs) offer a flexible alternative by capturing nonlinear relationships between species population abundance and environmental drivers (Wood, 2017; Pedersen et al., 2019). Unlike traditional MARSS models, which rely on linear state-space formulations, GAMs use smooth functions to model complex ecological processes, allowing for a more biologically realistic representation of population dynamics. By incorporating environmental covariates as smooth terms, GAMs can effectively model species responses to gradual and abrupt ecological changes, making them particularly valuable in systems where nonlinearity plays a crucial role.

Building on this framework, multivariate generalized additive models (MVGAMs) extend GAMs to simultaneously model multiple species or populations while accounting for temporal dependencies (Clark and Wells, 2023; Karunarathna et al., 2024; Clark et al., 2025). MVGAMs introduce latent dynamic processes that capture shared trends across species or locations, offering a structured yet flexible approach to understanding multi-species interactions and environmental influences, making them conceptually similar to MARSS models. However, the flexibility of MVGAMs to integrate hierarchical structures, dynamic trends, delayed environmental effects (Clark and Wells, 2023), and nonlinear covariate relationships makes them a powerful tool for generating ecologically meaningful and statistically robust forecasts with greater computational efficiency. Given these advantages, we will employ MVGAMs in Chapters 4 and 5 of this thesis to forecast the population abundance of small mammal species, particularly in systems influenced by environmental variability and species interactions.

References

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, pages 199–213. Springer.

Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., Leos-Barajas, V., Mills Flemming, J., Nielsen, A., Petris, G., et al. (2021). A guide to state–space modeling of ecological time series. *Ecological Monographs*, 91(4):e01470.

Buckland, S., Newman, K., Thomas, L., and Koesters, N. (2004). State-space models for the dynamics of wild animal populations. *Ecological Modelling*, 171(1-2):157–175.

Bureau of Meteorology (2024). Climate data online. http://www.bom.gov.au/climate/data/. Accessed on 26/12/2024.

Burns, E., Lindenmayer, D., Tennant, P., Dickman, C., Green, P., Hanigan, I., Hoffmann, A., Keith, D., Metcalfe, D., Nolan, K., Russell-Smith, J., Wardle, G., Welsh, A., Williams, R., and Yates, C. (2014). *Making ecological monitoring successful: Insights and lessons from the Long Term Ecological Research Network*.

Clark, N. J., Ernest, S. M., Senyondo, H., Simonis, J., White, E. P., Yenni, G. M., and Karunarathna, K. (2025). Beyond single-species models: leveraging multispecies forecasts to navigate the dynamics of ecological predictability. *PeerJ 13:e18929*.

Clark, N. J. and Wells, K. (2023). Dynamic generalised additive models (DGAMs) for forecasting discrete ecological time series. *Methods in Ecology and Evolution*, 14(3):771–784.

De Valpine, P. (2003). Better inferences from population-dynamics experiments using Monte Carlo state-space likelihood methods. *Ecology*, 84(11):3064–3077.

Dennis, B. and Ponciano, J. M. (2014). Density-dependent state-space model for population-abundance data with unequal time intervals. *Ecology*, 95(8):2069–2076.

Dennis, B., Ponciano, J. M., Lele, S. R., Taper, M. L., and Staples, D. F. (2006). Estimating density dependence, process noise, and observation error. *Ecological Monographs*, 76(3):323–341.

Dickman, C., Wardle, G., Foulkes, J., and Preu, N. d. (2014). Desert complex environments. In Burns, E., Lowe, A., Lindenmayer, D., and Thurgate, N., editors, *Biodiversity and environmental change: monitoring, challenges and direction*, chapter 10, pages 379–438.

CSIRO Publishing.

Dickman, C. R., Greenville, A. C., Beh, C.-L., Tamayo, B., and Wardle, G. M. (2010). Social organization and movements of desert rodents during population "booms" and "busts" in central Australia. *Journal of Mammalogy*, 91(4):798–810.

Ernest, S. M., Yenni, G. M., Allington, G., Bledsoe, E. K., Christensen, E. M., Diaz, R. M., Geluso, K., Goheen, J. R., Guo, Q., Heske, E., et al. (2018). The Portal Project: a long-term study of a Chihuahuan desert ecosystem. *BioRxiv*, page 332783.

Frank, A. S., Dickman, C. R., and Wardle, G. M. (2012). Habitat use and behaviour of cattle in a heterogeneous desert environment in central Australia. *The Rangeland Journal*, 34(3):319–328.

Frank, A. S., Dickman, C. R., Wardle, G. M., and Greenville, A. C. (2013). Interactions of grazing history, cattle removal and time since rain drive divergent short-term responses by desert biota. *PLoS One*, 8(7):e68466.

Frank, A. S., Wardle, G. M., Greenville, A. C., and Dickman, C. R. (2016). Cattle removal in arid Australia benefits kangaroos in high quality habitat but does not affect camels. *The Rangeland Journal*, 38(1):73–84.

Friend, G., Smith, G., Mitchell, D., and Dickman, C. (1989). Influence of pitfall and drift fence design on capture rates of small vertebrates in semi-arid habitats of Western-Australia. *Wildlife Research*, 16(1):1–10.

Fukasawa, K., Hashimoto, T., Tatara, M., and Abe, S. (2013). Reconstruction and prediction of invasive mongoose population dynamics from history of introduction and management: a Bayesian state-space modelling approach. *Journal of Applied Ecology*, 50(2):469–478.

Greenville, A. C., Dickman, C. R., Wardle, G. M., and Letnic, M. (2009). The fire history of an arid grassland: the influence of antecedent rainfall and ENSO. *International Journal of Wildland Fire*, 18(6):631–639.

Greenville, A. C., Wardle, G. M., and Dickman, C. R. (2013). Extreme rainfall events predict irruptions of rat plagues in central Australia. *Austral Ecology*, 38(7):754–764.

Greenville, A. C., Wardle, G. M., Nguyen, V., and Dickman, C. R. (2016). Population dynamics of desert mammals: similarities and contrasts within a multispecies assemblage. *Ecosphere*, 7(5):e01343.

Hinrichsen, R. and Holmes, E. E. (2009). Using multivariate state-space models to study spatial structure and dynamics. *CRC/Chapman Hall*.

Holmes, E. E., Ward, E. J., and Kellie, W. (2012). MARSS: multivariate autoregressive state-space models for analyzing time-series data. *The R Journal*, 4(1):11.

Holmes, E. E., Ward, E. J., and Scheuerell, M. D. (2014). Analysis of multivariate time-series using the MARSS package. *NOAA Fisheries, Northwest Fisheries Science Center*, 2725:98112.

Humbert, J. Y., Scott Mills, L., Horne, J. S., and Dennis, B. (2009). A better way to estimate population trends. *Oikos*, 118(12):1940–1946.

Hurvich, C. M. and Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307.

Iijima, H., Nagaike, T., and Honda, T. (2013). Estimation of deer population dynamics using a bayesian state-space model with multiple abundance indices. *The Journal of Wildlife Management*, 77(5):1038–1047.

Karunarathna, K., Wells, K., and Clark, N. J. (2024). Modelling nonlinear responses of a desert rodent species to environmental change with hierarchical dynamic generalized additive models. *Ecological Modelling*, 490:110648.

Keller, M., Schimel, D. S., Hargrove, W. W., and Hoffman, F. M. (2008). A continental strategy for the National Ecological Observatory Network. *The Ecological Society of America: 282-284*.

Kent, A. D., Yannarell, A. C., Rusak, J. A., Triplett, E. W., and McMahon, K. D. (2007). Synchrony in aquatic microbial community dynamics. *The ISME journal*, 1(1):38–47.

Kim, E.-S. (2006). Development, potentials, and challenges of the International Long-Term Ecological Research (ILTER) Network. *Ecological Research*, 21(6):788–793.

Kuebbing, S. E., Reimer, A. P., Rosenthal, S. A., Feinberg, G., Leiserowitz, A., Lau, J. A., and Bradford, M. A. (2018). Long-term research in ecology and evolution: A survey of challenges and opportunities. *Ecological Monographs*, 88(2):245–258.

Lawton, J. H. (1988). More time means more variation. *Nature*, 334(6183):563–563.

Letnic, M. (2004). Cattle grazing in a hummock grassland regenerating after fire: the short-term effects of cattle exclusion on vegetation in south-western Queensland. *The Rangeland Journal*, 26(1):34–48.

Letnic, M. and Dickman, C. R. (2010). Resource pulses and mammalian dynamics: conceptual models for hummock grasslands and other Australian desert habitats. *Biological Reviews*, 85(3):501–521.

Letnic, M., Story, P., Story, G., Field, J., Brown, O., and Dickman, C. R. (2011). Resource pulses, switching trophic control, and the dynamics of small mammal assemblages in arid Australia. *Journal of Mammalogy*, 92(6):1210–1222.

Levine, J. M., HilleRisLambers, J., Petry, W. K., Usinowicz, J., and Crowther, T. W. (2024). Demographic but not competitive time lags can transiently amplify climate-induced changes in vegetation carbon storage. *Global Change Biology*, 30(8):e17432.

Likens, G. and Lindenmayer, D. (2018). *Effective ecological monitoring*. CSIRO publishing.

Lindenmayer, D. B. and Likens, G. E. (2011). Losing the culture of ecology. *The Bulletin of the Ecological Society of America*, 92(3):245–246.

Lindenmayer, D. B., Likens, G. E., Andersen, A., Bowman, D., Bull, C. M., Burns, E., Dickman, C. R., Hoffmann, A. A., Keith, D. A., Liddell, M. J., et al. (2012). Value of long-term ecological studies. *Austral Ecology*, 37(7):745–757.

Ljung, G. M. and Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303.

Lofton, M. E., Brentrup, J. A., Beck, W. S., Zwart, J. A., Bhattacharya, R., Brighenti, L. S., Burnet, S. H., McCullough, I. M., Steele, B. G., Carey, C. C., et al. (2022). Using near-term forecasts and uncertainty partitioning to inform prediction of oligotrophic lake cyanobacterial density. *Ecological Applications*, page e2590.

Mahan, M. Y., Chorn, C. R., and Georgopoulos, A. P. (2015). White Noise Test: detecting autocorrelation and nonstationarities in long time series after ARIMA modeling. In *SciPy*, pages 97–104.

Malhi, Y., Phillips, O. L., Lloyd, J., Baker, T., Wright, J., Almeida, S., Arroyo, L., Frederiksen, T., Grace, J., Higuchi, N., et al. (2002). An international network to monitor the structure, composition and dynamics of Amazonian forests (RAINFOR). *Journal of*

Vegetation Science, 13(3):439–450.

Morton, S., Smith, D. S., Dickman, C. R., Dunkerley, D., Friedel, M., McAllister, R., Reid, J., Roshier, D., Smith, M., Walsh, F., et al. (2011). A fresh framework for the ecology of arid Australia. *Journal of Arid Environments*, 75(4):313–329.

Murray, B., Dickman, C., Watts, C., and Morton, S. R. (1999). The dietary ecology of Australian rodents. *Wildlife Research*, 26(6):857–858.

Murray, B. R. and Dickman, C. R. (1994). Granivory and microhabitat use in Australian desert rodents: are seeds important? *Oecologia*, 99:216–225.

Mutshinda, C. M., O'Hara, R. B., and Woiwod, I. P. (2009). What drives community dynamics? *Proceedings of the Royal Society B: Biological Sciences*, 276(1669):2923–2929.

Newman, K., Buckland, S., Morgan, B. J., King, R., Borchers, D., Cole, D. J., Besbeas, P., Gimenez, O., and Thomas, L. (2014). Modelling population dynamics. *Methods in Statistical Ecology. New York, NY: Springer New York*.

Pedersen, E. J., Miller, D. L., Simpson, G. L., and Ross, N. (2019). Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ*, 7:e6876.

Purdie, R. (1984). Land systems of the Simpson desert region. *Natural Resources Series*, 2:1–71.

Sánchez-Espigares, J. A., Grima, P., and Marco-Almagro, L. (2018). Visualizing type II error in normality tests. *The American Statistician*, 72(2):158–162.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4):591–611.

Shephard, M. (1992). *The Simpson Desert: natural history and human endeavour*. Royal Geographical Society of Australasia, Adelaide, Australia.

Shreffler, J. and Huecker, M. R. (2023). *Type I and type II errors and statistical power*. StatPearls Publishing, Treasure Island (FL).

Sugiura, N. (1978). Further analysis of the data by akaike's information criterion and the finite corrections: Further analysis of the data by akaike's. *Communications in Statistics - Theory and Methods*, 7(1):13–26.

Thibault, K. M. and Brown, J. H. (2008). Impact of an extreme climatic event on community assembly. *Proceedings of the National Academy of Sciences*, 105(9):3410–3415.

Tulloch, A. I., Healy, A., Silcock, J., Wardle, G. M., Dickman, C. R., Frank, A. S., Aubault, H., Barton, K., and Greenville, A. C. (2023). Long-term livestock exclusion increases plant richness and reproductive capacity in arid woodlands. *Ecological Applications*, 33(8):e2909.

Ward, E. J., Chirakkal, H., González-Suárez, M., Aurioles-Gamboa, D., Holmes, E. E., and Gerber, L. (2010). Inferring spatial structure from time-series data: using multivariate state-space models to detect metapopulation structure of California sea lions in the Gulf of California, Mexico. *Journal of Applied Ecology*, 47(1):47–56.

Wardle, G. and Dickman, C. (2015). Desert Ecology Plot Network: Mammal, Reptile and Vegetation Data Associated with Weather, Simpson Desert, Western Queensland, Australia, 1990–2011. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern.111.57/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018a). Desert Ecology Plot Network: Mammal Abundance Plot-data, Simpson Desert, Western Queensland, 1990+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.196.8/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018b). Desert Ecology Plot Network: Vegetation Plot-data, Simpson Desert, Western Queensland, 1993+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.195.11/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018c). Desert Ecology Plot Network: Weather Data (daily and monthly), Simpson Desert, Western Queensland, 1995+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.194.7/html. Accessed on 16/01/2025.

Wardle, G. M., Greenville, A. C., Frank, A. S., Tischler, M., Emery, N. J., and Dickman, C. R. (2015). Ecosystem risk assessment of Georgina gidgee woodlands in central Australia. *Austral Ecology*, 40(4):444–459.

White, E. P., Yenni, G. M., Taylor, S. D., Christensen, E. M., Bledsoe, E. K., Simonis, J. L., and Ernest, S. K. M. (2019). Developing an automated iterative near-term forecasting system for an ecological study. *Methods in Ecology and Evolution*, 10(3):332–344.

White, P. J., Bruggeman, J. E., and Garrott, R. A. (2007). Irruptive population dynamics in Yellowstone pronghorn. *Ecological Applications*, 17(6):1598–1606.

Wolley, P. (2005). The species of *Dasycercus* Peters, 1875 (Marsupialia: Dasyuridae). *Memoirs of Museum Victoria*, 62(2):213–221.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., and Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, 122:2–10.

CHAPTER 3

Evaluating the sensitivity of imputation methods to characteristics of missing values in population abundance time series

---

3.1 Introduction

Ecologists collect time series data on population abundances across multiple populations to monitor species dynamics and assess ecosystem health. However, obtaining such datasets is both labour-intensive and resource-demanding (Humbert et al., 2009). Despite these efforts, missing data are a persistent challenge in ecological research, often arising from logistical and financial constraints or environmental disturbances that limit access to field sites (Humbert et al., 2009; Penone et al., 2014; Ellington et al., 2015). These gaps can introduce bias, reduce statistical power, and lead to inaccurate ecological inferences, ultimately affecting conservation and management decisions (Onkelinx et al., 2017; Jóhannesson et al., 2019; Łopucki et al., 2022). Furthermore, missing data are not limited to population abundance but also extend to climatic and environmental variables, which are crucial for modeling species dynamics and understanding ecological processes (Noor and Zainudin, 2008; Sa'adi et al., 2023).

Missing data can occur via three different mechanisms: missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). MCAR occurs when the probability of missing data is entirely independent of both observed (i.e., abundance) and unobserved variables, making the missingness purely random. MAR, a more common scenario in ecological studies, occurs when missing values depend on other observed variables, such as the climatic conditions that influence the survey efforts. In contrast, MNAR arises when missing data depend on unobserved factors, making their absence

systematic and potentially introducing bias (Schafer and Graham, 2002). Since ecological datasets are rarely MCAR, the assumption of MAR is often more appropriate (Penone et al., 2014). However, Bowler et al. (2025) argue that gaps in ecological data can be considered MCAR if the factors influencing missingness are entirely independent of those affecting species abundance. On the other hand, as Nakagawa and Freckleton (2008) demonstrated, deleting cases with missing values can lead to biased parameter estimates when data are not MCAR, highlighting the need for robust imputation methods to address missingness effectively.

Addressing missing data is crucial for ensuring the reliability of ecological analyses, particularly in long-term population studies. Listwise deletion, a common approach, excludes observations with missing values, often resulting in substantial data loss and distorted ecological patterns (Little and Rubin, 2019). In time series data, where consistent observation frequencies are essential, this method can create gaps that complicate modeling due to the temporal dependence of successive observations (Graham, 2009). While listwise deletion may be a reasonable approach when aggregating high-resolution data (e.g., hourly to daily or weekly), it is generally unsuitable for lower-resolution datasets, particularly when missing values also affect covariates essential for modeling species abundance. In such cases, imputation methods offer a robust alternative by estimating missing values based on observed data, preserving temporal structure, and enhancing analytical accuracy (Penone et al., 2014; Ellington et al., 2015; Chapon et al., 2023). However, the effectiveness of different imputation techniques varies depending on the characteristics of the dataset and the proportion of missing values, which requires a careful evaluation of their performance in ecological contexts (Łopucki et al., 2022).

Single imputation methods are among the most commonly used techniques for addressing missing data in ecological studies. These methods replace missing values with a single estimated value, derived from observed data. For example, mean imputation involves substituting missing population counts with the average abundance across the time series (Van Buuren, 2018). Alternatively, the moving average method fills the gaps by averaging nearby observations, which can be particularly useful for smoothing short-term fluctuations in abundance

(Humbert et al., 2009). The Kalman filter, a more advanced approach, uses statistical models to predict missing values based on past data and temporal dependencies (Agbailu et al., 2021). While these methods are computationally efficient, they have limitations. They fail to account for the uncertainty in the imputed values, which can lead to underestimation of variability and skewed inferences about population trends (Khan, 2024). Additionally, these methods often overlook complex relationships between missing data and other ecological factors, which is particularly problematic when climatic and environmental variables are also missing (Penone et al., 2014; Onkelinx et al., 2017).

Multiple imputation methods have been developed to address the limitations of single imputation by incorporating uncertainty into missing data estimates. Unlike single imputation, which replaces each missing value with a single estimate, multiple imputation generates several plausible values based on the distribution of observed data, producing multiple complete datasets (Azur et al., 2011). These datasets are then analyzed separately and the results are pooled to obtain more reliable parameter estimates while preserving the variability inherent in the data. In population abundance studies, multiple imputation accounts for the uncertainty in missing data, reducing bias in population trend estimates and enhancing the robustness of statistical analyses (Callaghan et al., 2021). Moreover, this approach can accommodate complex relationships between variables, such as interactions between species and environmental and climatic factors, making it well-suited for ecological time series (Ellington et al., 2015).

Despite the prevalence of missing data in ecological research, imputation methods have been less frequently employed in ecological studies (Łopucki et al., 2022). Previous research on handling missing values has focused primarily on comparing complete case removal with various single and multiple imputation approaches using compiled datasets from unpublished sources and literature reviews covering experimental and observational data (Ellington et al., 2015). In addition, imputation techniques have been evaluated based on the type of missingness using nearly complete trait datasets, such as those from the mammalian order Carnivora (Penone et al., 2014). However, the application of imputation tools in field-collected population abundance time series remains limited compared to their use in social, biological or

medical sciences (Van Buuren, 2018; Khan, 2024). Therefore, further research is needed to evaluate, compare and refine imputation strategies to handle missing data in population ecology, emphasizing the importance of addressing missing data rather than disregarding them (Łopucki et al., 2022).

In this study, we used long-term population abundance data of a native rodent, the spinifex hopping mouse (*Notomys alexis*) from the Desert Ecology Research Group (DERG) dataset (see Chapter 2 for details). To systematically assess imputation performance, we conducted a simulation study with varying lengths of the time series that replicated the characteristics of the population time series and strategically introduced artificial gaps to create missing values (unlike the simulation performed in Agbailu et al. (2021)). We then applied four imputation techniques: three single imputation methods — Mean, Moving average, and Kalman filter, and a multiple imputation approach, Multivariate Imputation by Chained Equations (MICE) (Van Buuren and Groothuis-Oudshoorn, 2011). These methods were chosen to evaluate the differences between imputing missing values using a single estimate versus incorporating uncertainty through multiple imputations.

The aims of this study were to 1) test the sensitivity of each imputation method to factors such as time series length, quantity, and temporal positioning of missing values, and variations across different samples, and 2) assess whether the imputation procedures significantly impacted the precision of population abundance predictions. To achieve this, we first compared the variability of error between the imputed and simulated values in the training set, followed by an evaluation of how well the predicted abundances aligned with the simulated population abundance values in the testing set. We expect that longer time series will result in more accurate imputations, particularly for MICE, which can leverage the full data range, while single imputation methods may struggle to capture temporal dependencies. We anticipate that Mean and Moving average will perform better with fewer missing values (depending on the proportion and length of the time series), with MICE expected to outperform as the proportion of missing data increases, especially in cases of clustered missing values. The Kalman filter, which accounts for temporal dependencies, is anticipated to perform better than Mean and Moving average, though it may fall short in capturing long-term trends

and abrupt changes in population dynamics when missing values are clustered and/or occur at the beginning of the series. However, we don't expect significant variability in the performance of any imputation method based on whether missing values occur at the beginning, middle, or end of the series. Finally, we anticipate that MICE will provide more accurate and precise population abundance predictions than single imputation methods, with the Kalman filter performing better than Mean and Moving average due to its ability to account for temporal dependencies.

3.2 Material and methods

3.2.1 Study location

The long-term population monitoring study conducted by the Desert Ecology Research Group (DERG) at the University of Sydney was carried out in the Simpson Desert, central Australia, since 1990 (Wardle and Dickman (2015), Wardle and Dickman (2018a), Wardle and Dickman (2018b), Wardle and Dickman (2018c), unpublished data from 2018). The Simpson Desert covers 170,000 km^2 of mainly dune fields (73%), with the remainder comprising clay pans, rocky outcrops, and gibber plains (Shephard, 1992). The study region of 8000 km^2, spans the border between Queensland and the Northern Territory (Fig. 2.1) and the vegetation of the dune fields is dominated by spinifex grasslands (*Triodia basedowii*) with scattered shrubs (*Acacia*, *Eremophilla*, *Grevillea*), ephmeral forbs and a variety of other perennial and annual grasses; interspersed are wide swales dominated by Georgina Gidgee woodlands (*Acacia georginae*), or mallee eucalypts (Wardle et al., 2015). The climate is highly variable, with prolonged dry "bust" periods interrupted by productive "boom" periods following summer rainfall (Dickman et al., 2010; Greenville et al., 2013). Temperatures exceed 40°C in summer and drop below 5°C in winter (Purdie, 1984), while rainfall is seasonal, with annual averages decreasing from northwest to southeast: 258.5 mm at Boulia (136 years), 203 mm at Bedourie (79 years), and 157 mm at Birdsville (24 years) (Bureau of Meteorology, 2024). Wildfires (Letnic, 2004) and cattle grazing (Frank et al., 2013; Tulloch et al., 2023) are key factors that directly affect vegetation, with interactive effects for the abundance of fauna species in the region. The mean minimum wildfire return interval is 27 years (Verhoeven et al.,

2020), and cattle have been removed from the study region in 2004 (Ethabuka Reserve) and 2006 (Pilungah Reserve) (Frank et al., 2016) (see Chapter 2 for details).

3.2.2 Study species

Live-trapping was conducted at the Tobermorey East site within Tobermorey Station, in the study region (Fig. 2.1). Monitoring of this site started in 1995, and all grids within the site were selected (see Chapter 2 for details). Out of all the seventeen sites in the study region, Tobermorey East was specifically chosen as a representative example of a site with a truncated start and potential data gaps due to various factors. Although long-term trapping efforts recorded several small mammal species, this study focused on the native rodent, the spinifex hopping mouse (*Notomys alexis*, 35 g). This species was selected due to its high boom-bust dynamics and strong relationship with rainfall (see Chapter 2; Greenville et al. (2016)), as it relies more heavily on seeds than *Pseudomys hermannsburgensis* (Murray et al., 1999; Dickman et al., 2014) and can disappear during prolonged dry periods. These characteristics make it a more extreme test case for imputation methods compared to the dasyurid marsupials (see Chapter 2 for details) in the DERG dataset. The standardized capture data per 100 trap nights (TN) from 1995 to 2012 were aggregated at the site level, averaged annually and log-transformed to align with Greenville et al. (2016) (see Data preparation performed in Chapter 1). Sampling efforts in some years were interrupted due to extreme weather, road closures, and logistical constraints, resulting in missing data (Fig. 3.1(A)). Additionally, since only a few other sites in the DERG dataset were surveyed from 1990 (see Chapter 1), we considered data from pre-1995 as missing for our selected site.

3.2.3 Covariate measurements

An advantage of multiple imputation techniques is that they account for the relationships between population abundance and related environmental or climatic variables, leading to more accurate imputations. In Chapter 2, we found that the population abundance of the spinifex hopping mouse is significantly impacted by the previous year rainfall. Therefore,

we considered annual rainfall totals (calendar year) calculated from daily records obtained from automatic weather stations (Environdata, Warwick, Queensland) operational from 1995 to 2012 (Fig. 3.1(B)). A 1-year lag was applied to account for the time required for small mammals to respond to significant rainfall events (Chapter 2), such as increased breeding, a method previously used to predict mammal captures (Greenville et al., 2013; Levine et al., 2024).

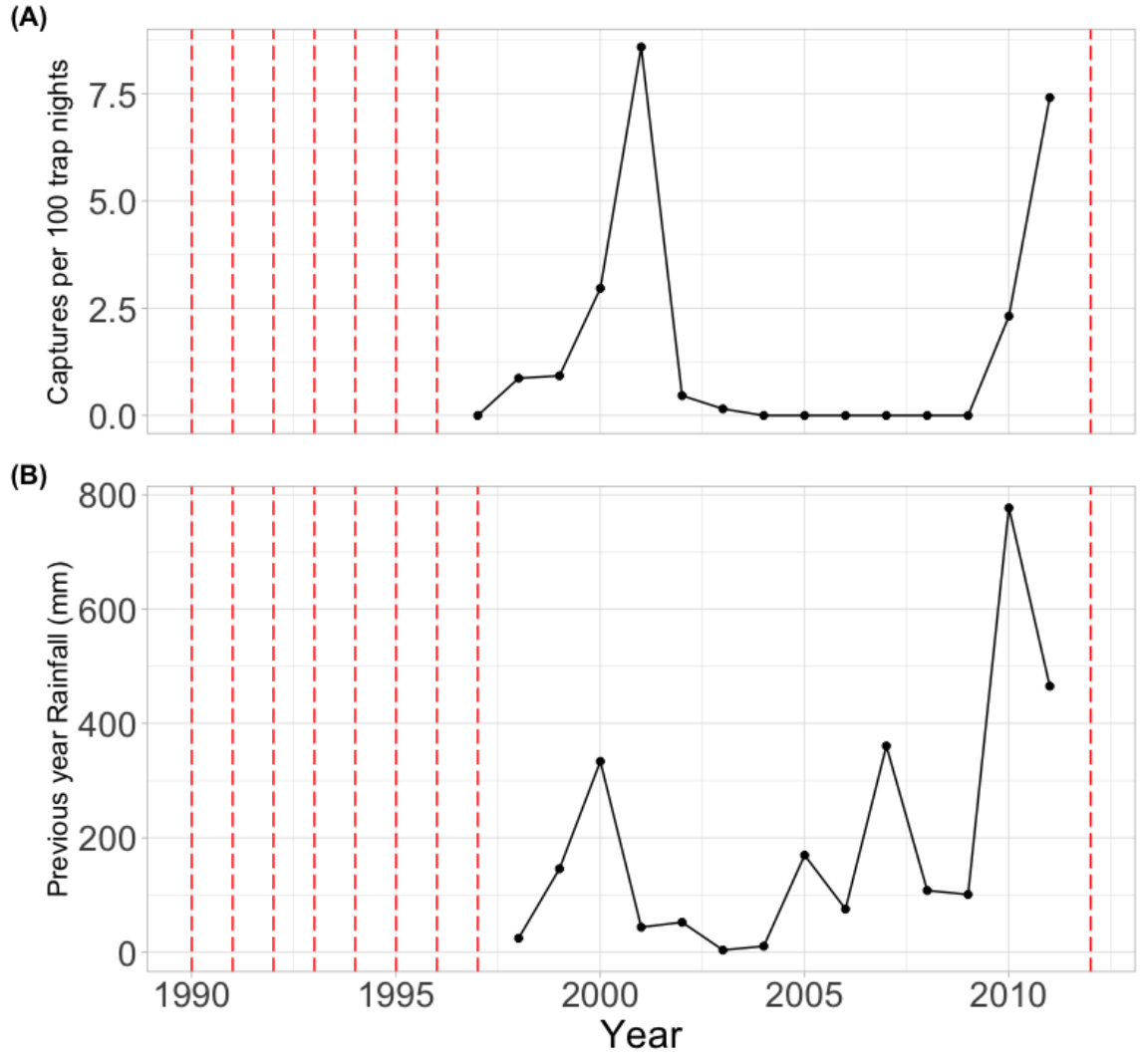

FIGURE 3.1. Time series of (A) captures of *Notomys alexis* standardized per 100 trap nights (TN) and (B) previous year annual rainfall (mm) at Tobermorey East from 1995 - 2012. Red dashed vertical lines indicate missing data: pre-1995, when the site was not established, and post-1995, where data were unavailable.

3.2.4 Data simulation

To evaluate the sensitivity of the selected imputation techniques to different characteristics of the time series and patterns of missing data, we simulated datasets for *N. alexis* based on the DERG dataset and annual rainfall at the study site. The longest continuous non-missing subset from the DERG dataset (1997–2011 for captures and 1998 - 2011 for previous year rainfall; Fig. 3.1) was used to simulate datasets of varying lengths. Simulated time series varied lengths; 60, 90, 900, and 9000 years, with 50 datasets generated for each length, to ensure representativeness of different ecological phenomena that can affect the population abundance of the species at the study site and generate unbiased estimates (Hossie et al., 2021).

The dataset simulations were handled in two different ways (Fig. 3.2). Using the longest non-missing subset, a simple State-Space model (SSM) without any covariates (Auger-Méthé et al., 2021) was fitted for the abundance data while an Autoregressive Integrated Moving Average (ARIMA) (Dhamo and Puka, 2010) model was fitted for the annual rainfall data to simulate the required datasets. The SSM for abundance data was fitted using the *R2jags* package (Inglis et al., 2018) and the ARIMA model for rainfall data was implemented using the *auto.arima()* function from the *forecast* package in R (Dhamo and Puka, 2010). Notably, the manually determined ARIMA model was the same as that of the model configured through the function *auto.arima*.

3.2.5 Missing datasets

We created incomplete datasets from the simulated abundance time series by removing 10 manually selected values from each time series, regardless of its length, to identify a simulation length that was optimal in a pragmatic and realistic sense, one that balances biological plausibility with the need to test model performance under conditions that could feasibly occur in real-world ecological monitoring (Fig. 3.2). These values were chosen to reflect common patterns of missing data, with the same time steps removed across all time series lengths to ensure a consistent basis for comparing imputation errors. Upon the selection of

the optimal simulation length, we further assessed the performance of imputation methods by introducing additional missing values. The additional missing values were introduced only when the selected simulation length was not 60 years (Fig. 3.2). The simulated rainfall datasets remained complete throughout the analysis. All the simulations, imputations and analysis were run in R 4.3.1 (R Core Team, 2024).

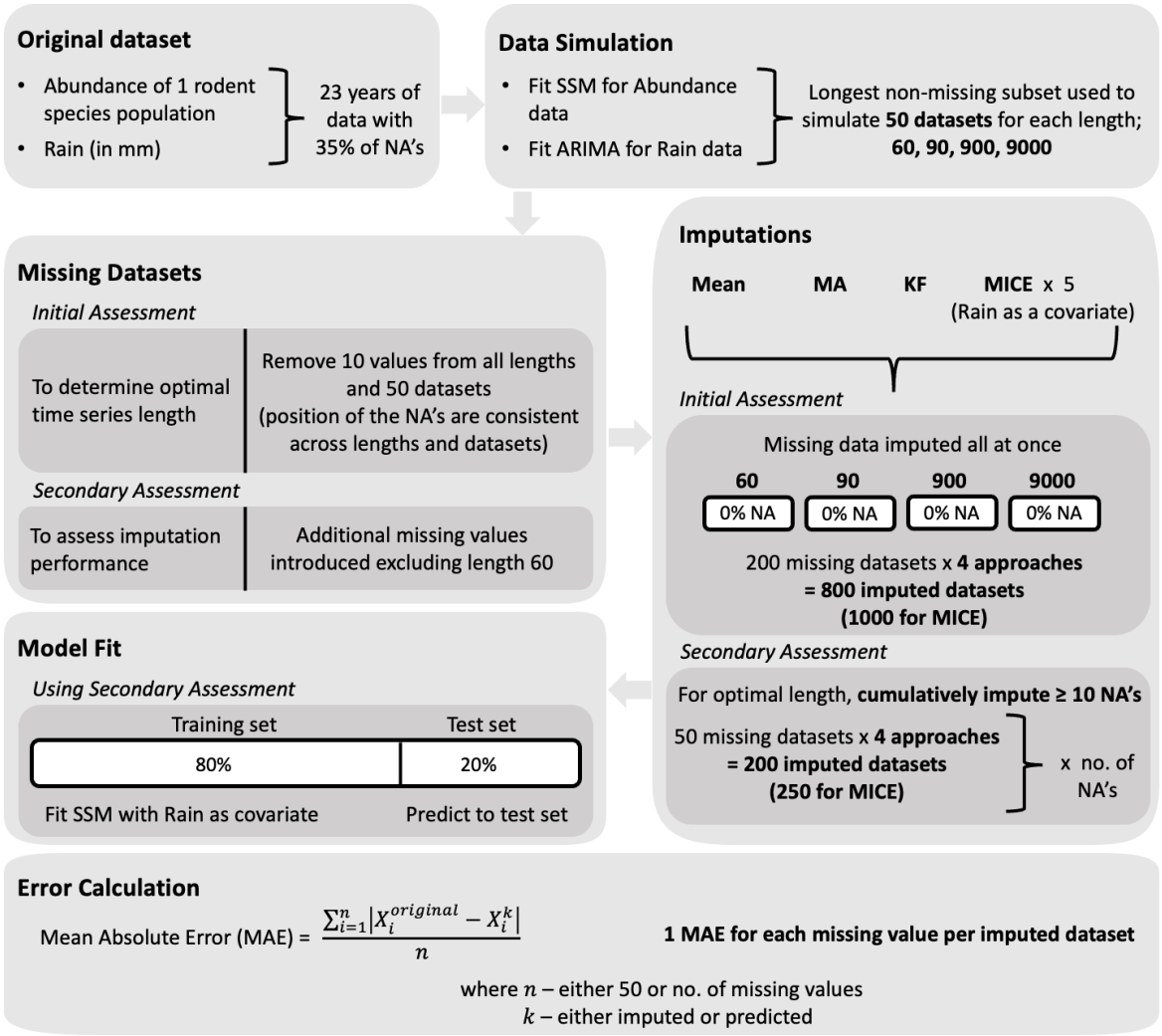

FIGURE 3.2. Overview of the methodology. Using the DERG dataset, we simulated 50 datasets of varying lengths. Missing datasets were generated under two assessments: (1) to determine the optimal simulated time series length and (2) to evaluate imputation performance using the selected optimal length. Missing values were imputed using four techniques in both assessments. The imputed datasets were then partitioned to evaluate the predictive performance of each imputation method. Mean Absolute Errors were calculated by comparing the imputed values with the simulated values.

3.2.6 Missing value imputation

We evaluated four imputation methods encompassing various computational strategies to assess the accuracy in handling missing data under various characteristics of missingness. Single imputation techniques; Mean, Moving average and Kalman filter imputations were performed using the functions *na_mean()*, *na_ma()*, *na_kalman()* respectively from the *ImputeTS* package in R (Moritz and Bartz-Beielstein, 2017). While Mean imputation replaces missing values with the arithmetic average of the available data, Moving average leverages a local window of adjacent data points to impute missing values. Therefore, we used the Exponential Weighted Moving Average Method (EWMA) so that exponentially decreasing weighting factors are used to ensure that recent past observations of abundance exert a more substantial influence on the imputation process compared to earlier observations. We set the integer width of the moving average window to be 4 which means 8 observations (4 left & 4 right). In contrast, the Kalman filter, models the temporal dependencies between data points, taking into account both the available observations and the dynamics of the data-generating process. We applied Kalman smoothing for imputation using a basic trend model with a seasonal component, fitted via maximum likelihood estimation.

In addition to the three single imputation methods, we employed Multivariate Imputation with Chained Equations (MICE) as a multiple imputation technique. The *mice* package in R was used to perform the MICE imputations (Van Buuren and Groothuis-Oudshoorn, 2011). We included simulated rainfall data as a covariate for MICE, as we found it to be highly associated with the probability of missing abundance data in the DERG dataset (Chapter 2; Bouhlila and Sellaouti (2013)). The Predictive Mean Matching (PMM) algorithm was then applied, fitting a linear regression model to the simulated abundance data with missing values. For each missing value, PMM selects an observed value from the donor pool whose predicted value is closest to the predicted value of the missing data (Van Buuren, 2018; Kleinke, 2018; Sa'adi et al., 2023). This method of using a single covariate, followed the guidelines put forward by the authors of MICE (Van Buuren and Groothuis-Oudshoorn, 2011). Five imputed datasets were generated using PMM, and the average of these datasets was taken as the point estimate for the missing values, while the variability across the imputed datasets was used to represent the uncertainty in the imputation process.

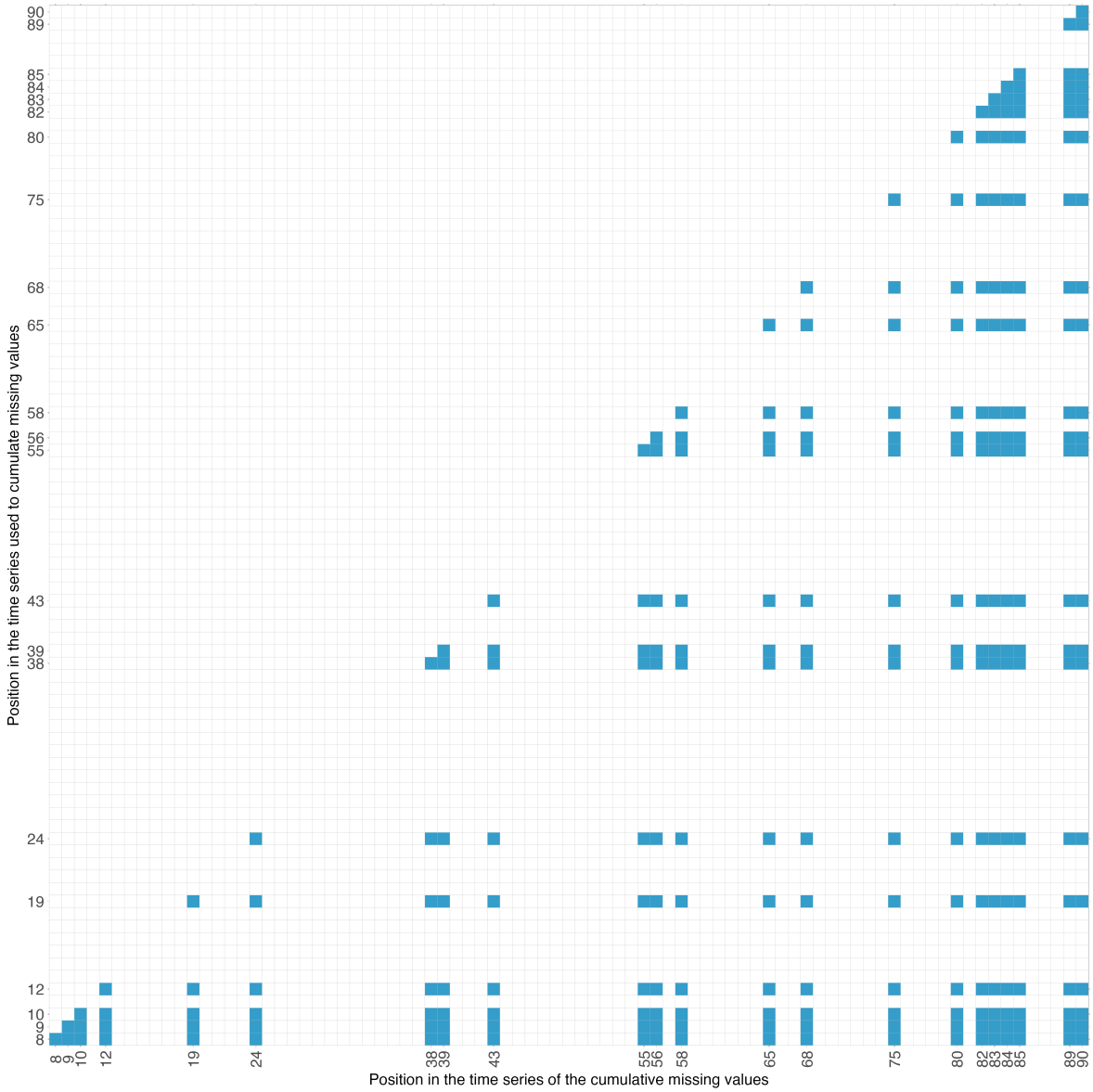

FIGURE 3.3. Performance of Mean, Moving average, Kalman filter, and the Multivariate Imputation by Chained Equations (MICE) imputation methods was evaluated according to the position of the missing value in the dataset and the number of missing values cumulatively imputed at each step. The blue shaded boxes indicate the missing value positions considered at each step. For example, at position 12 on the x-axis, the blue shaded boxes for 8, 9, 10, and 12 on the y-axis shows that these values were treated as missing when imputing abundance at position 12.

The imputation process followed two distinct criteria. First, all 10 missing values were imputed simultaneously for each time series length and across the 50 datasets. The optimal simulation length was then determined by comparing the errors between the imputed and

corresponding simulated values across the four imputation techniques. Then, using the selected optimal time series length, the accuracy of each imputation technique was evaluated based on 1) the impact of the position of a missing value within the time series and 2) the effect of the number of missing values imputed on the overall integrity of the time series. To achieve this, the missing values were imputed cumulatively, considering the sequential arrangement of the missing values, as illustrated in Fig. 3.3. The evaluation was conducted by comparing the errors between the imputed values and the corresponding simulated values, as well as the errors between the predicted values and the simulated values.

3.2.7 Model fit

To evaluate the performance of the four imputation approaches based on the predictions, we first partitioned the simulated datasets into training and testing sets, constituting 80% and 20% proportions, respectively (Fig. 3.2). Thereafter, the missing values in the training set were imputed as illustrated in Fig. 3.3 and a State-Space model (SSM) of the following form was fitted using the *R2jags* package in R (Inglis et al., 2018).

The process component of the SSM is characterized by a first-order autoregressive process given by:

$$X_t = BX_{t-1} + Rr_{t-1} + w_t; \quad w_t \sim N(0, Q) \tag{3.1}$$

Here, X_t signifies the true population abundance of the species at time t, with B and Q representing the process parameters. We incorporated the rainfall of the previous time-step (year) as a covariate denoted by r_{t-1}, and R represents its coefficient. w_t denotes process errors, assumed to be independent and following a normal distribution with a mean of zero and variance Q.

The observation component of the SSM is written as:

$$Y_t = X_t + v_t; \quad v_t \sim N(0, U) \tag{3.2}$$

where Y_t represents the simulated and imputed population abundances at time t. Importantly, Y_t is independent of Y_{t-1}, and all other observations, accounting for the dependence of Y_t on X_t. Observation errors, v_t, are assumed to be uncorrelated and follow a normal distribution with a mean of zero and variance U. Predictions for the testing set were obtained from the fitted SSM.

3.2.8 Imputation evaluation strategy

In assessing the effectiveness of each imputation method, we employed the Mean Absolute Error (MAE), representing the mean of the absolute errors in the imputed values (see equation in Fig. 3.2). Lower MAE values indicate superior imputation performance in approximating the simulated values. Error calculations were undertaken based on three distinct criteria, each tailored to provide insights into different aspects of the imputation process.

First, to determine the optimal simulation length, the MAE for each simulated length were computed by averaging the error between imputed and simulated values across the 10 missing values for each simulated length. Subsequently, focusing only on the optimal simulation length, we then evaluated how the temporal arrangement of missing values influenced imputation accuracy by averaging the error differences solely across the 50 datasets. This criterion allowed us to observe variations in the performance of each imputation method based on the placement of missing values.

Finally, to assess the effectiveness of the imputation approaches on the population abundance predictions for the testing set while considering the temporal placement of the missing values, we averaged the error differences exclusively across the 50 datasets.

3.3 Results

3.3.1 Optimal simulated time series length

Aggregating the errors between the imputed and simulated values across the 10 missing values and 50 datasets revealed that the 60-year population abundance time series had the lowest

MAE (Fig. 3.4). Extending the time series simulation length to 900 and 9000 stabilized the error and revealed that the performance of the imputation methods used in this study are robust across multiple time series lengths. This observation is also evident through the overlapping 95% confidence intervals around the point MAE estimates (Fig. 3.4). In contrast, the time series length of 90 years showed higher imputation errors across all methods in comparison to the other three simulation lengths, with MICE yielding the highest error among the four methods (Fig. 3.4). Given the impracticability of obtaining population abundance data for 900 or 9000 years, we selected the time series length of 90 years as the optimal simulation length for population abundance. This duration balances feasibility and ecological relevance, offering a more practical alternative than shorter time series, which may not capture long-term trends, or excessively long ones, which are less applicable for conservation decision-making.

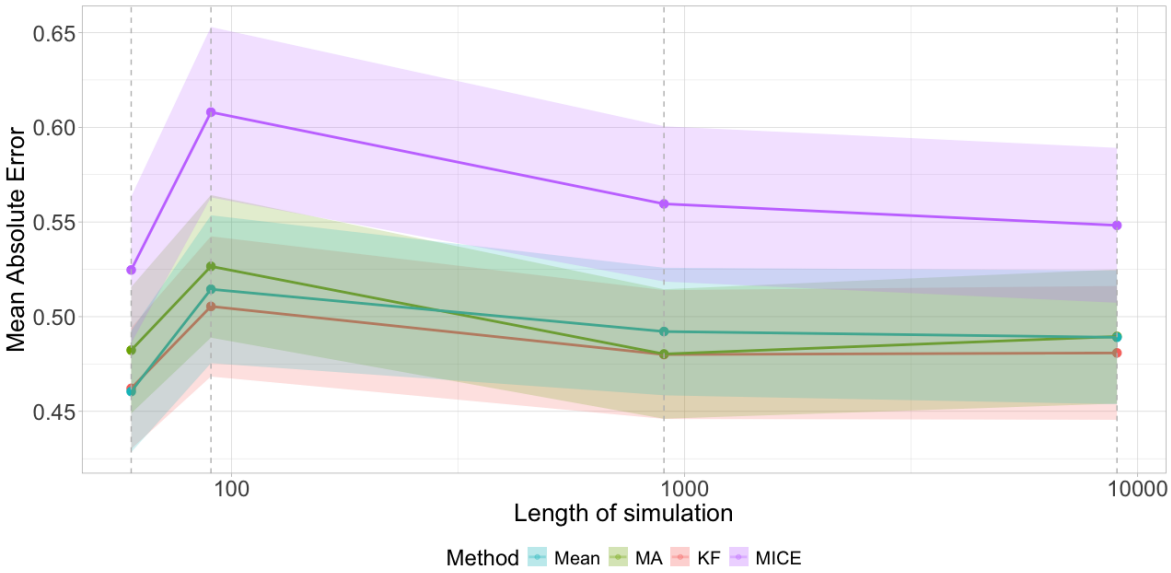

FIGURE 3.4. Mean error (MAE) of imputations averaged across 10 missing population abundances as mentioned in Fig. 3.2 across four different simulation lengths (60, 90, 900, and 9000). Imputation was performed using the Mean, Moving Average (MA), Kalman Filter (KF), and Multivariate Imputation by Chained Equations (MICE) methods. The coloured shaded regions represent the 95% prediction intervals and the grey dashed vertical lines indicate the exact positions of the simulation lengths on the x-axis.

The remaining analyses of this study was based on the simulation length of 90 years. Accordingly, the number of missing values in the dataset was increased to 22, ensuring a more even spread of missing values across the dataset.

3.3.2 Impact of the temporal placement of the missing value on imputation approaches

Overall, the error differences between the imputed population abundance and the simulated values tend to decrease as the number of missing values increases across all the imputation methods considered in this study (Fig. 3.5). Detailed observations reveal that employing MICE to impute missing population abundances yields the lowest median error difference between imputed and simulated values, particularly when up to 2% of missing values are concentrated consecutively within the first quarter of the dataset; however, overall differences among imputation methods were small. As the proportion of missing abundance values exceeds 2% in the last three quarters of the time series, the median MAE tends to increase, even though the variability in the error decreases when using MICE for imputation (Fig. 3.5). Conversely, imputation using the Kalman filter emerges as the superior choice, outperforming other methods in both median MAE and variability in the error, particularly when more than 4% of missing values are present in the latter three quarters of the time series (Fig. 3.5).

It is noteworthy that none of the employed imputation techniques in this study demonstrated inferior performance as the number of missing values increased in the time series. The reduction in the variability of the error differences across the 50 simulated datasets, showed the improvement in imputations, underscoring the dataset's temporal dependence. This suggests that the imputation methods tested in this study were able to effectively capture the underlying patterns and fluctuations in the simulated population dynamics of *N. alexis*, thus improving their ability to learn from the available data, even during sudden booms or busts in population abundance.

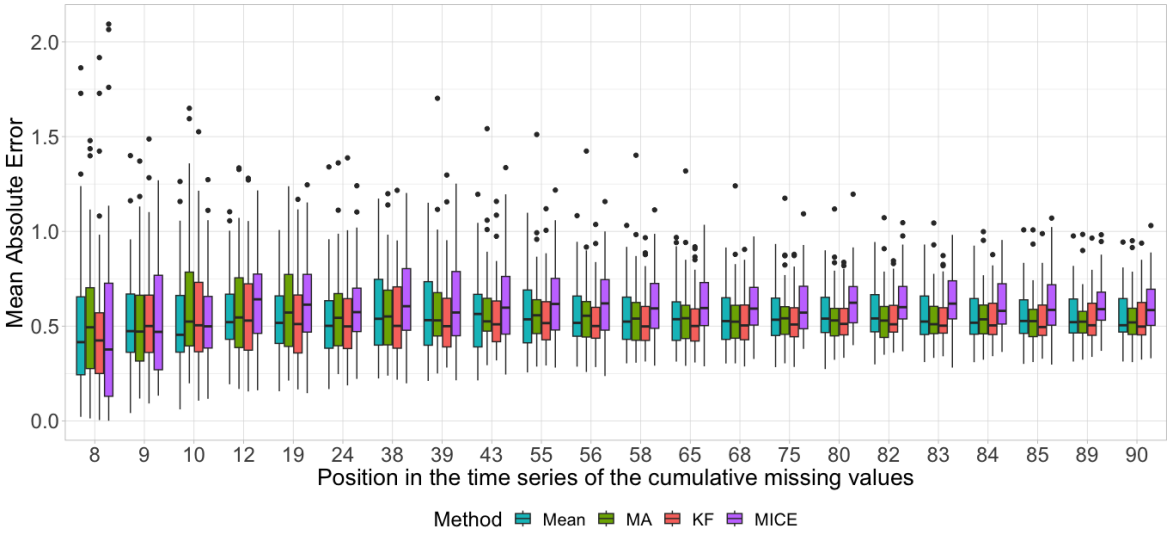

FIGURE 3.5. Mean error (MAE) between imputed and simulated population abundance according to the position of missing values. Imputation was performed using the Mean, Moving Average (MA), Kalman Filter (KF), and Multivariate Imputation by Chained Equations (MICE) methods, as described in Fig. 3.3, on the 90-year simulated time series. Boxplots show the variation in error differences across the 50 simulated datasets.

3.3.3 Effect of the temporal placement of the missing value on predicting population abundance

Error differences in population abundance predictions for the testing set, considering the temporal position of missing values in the training set (72 data points from the simulated 90 year time series) showed no significant improvement when missing values were imputed using any of the techniques applied in this study (Fig. 3.6). Furthermore, regardless of the number of missing values in the training set, none of the imputation methods yielded superior results in predicting population abundance. This is further supported by the stable variability in the error across the 50 datasets, indicating consistent performance across the different imputation techniques (Fig. 3.6). However, our analysis revealed that the median MAE was lower when missing data were imputed using either the Mean or MICE methods, rather than being deleted, when more than 4% of missing values occurred in the last three quarters of the dataset. Nonetheless, the differences among methods were minimal (Fig. 3.6).

As our simulated dataset incorporates the temporal aspect of population abundances, using a single imputation technique such as the Mean would not meet the statistical assumptions (e.g., white noise and normally distributed residuals (see Chapter 2 for more details)) of a time series dataset when predicting population abundance. Therefore, MICE plays a crucial role in imputing missing values for time-dependent datasets, particularly as it also integrates information from a highly dependent covariate.

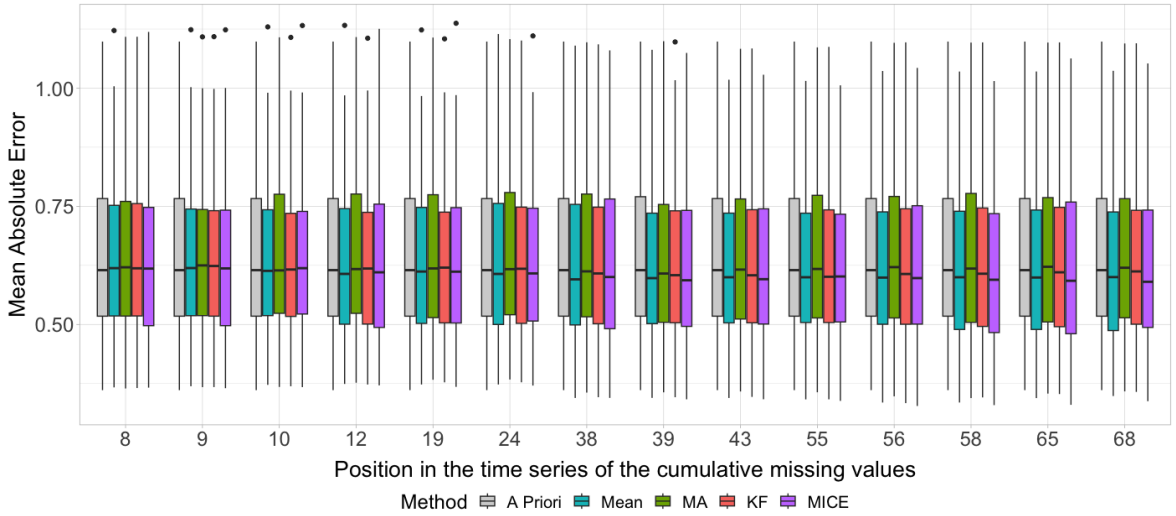

FIGURE 3.6. Mean Absolute Error (MAE) of testing set population abundance predictions when missing values in the training set are imputed cumulatively using the Mean, Moving Average (MA), Kalman Filter (KF), and Multivariate Imputation by Chained Equations (MICE) methods, as described in Fig. 3.3, for the 90-year simulated time series. The MAE of *A Priori* represents the error differences between the testing set and predictions when missing values were excluded from the training set. Boxplots show the variation in error differences across the 50 simulated datasets.

3.4 Discussion

Long-term population abundance time series often contain significant gaps due to logistical and financial constraints and environmental disturbances (Humbert et al., 2009). In datasets collected at low frequencies (e.g., seasonally or yearly), excluding records with missing data is not a practical option (Nakagawa and Freckleton, 2008). Therefore, addressing missing data is essential to ensure the reliability of ecological analyses. Our study focused on understanding the effects of imputation methods on factors such as time series length, the

quantity, the temporal positioning of missing values, and their robustness across different samples to select the most appropriate approach. Overall, our analyses showed that, in time-dependent datasets concerning population abundance, the choice of imputation method has minimal impact on predicting population abundance. This conclusion holds true regardless of the temporal placement or number of missing values in the dataset. Consequently, the interpretability and relevance of domain-specific insights derived from predicted abundances remain largely unaffected by the selected imputation method.

3.4.1 Optimal simulation length

Our results exploring the optimal time series simulation length revealed that a time series of 60 years minimized the imputation error among the four simulated lengths tested in this study (60, 90, 900, and 9000 years). Although extending the simulation length to 900 and 9000 years led to stabilization of error, the impracticality of acquiring such long-term data, led us to select 90 year time series as the optimal simulation length. Importantly, all analyses in this study were rigorously tested on simulated datasets with lengths 900 and 9000 years, yet no notable improvement was observed compared to the performance of the 90 year dataset. Consequently, choosing a 90 year time series provides a balance between methodological rigor and the practical challenges of obtaining extensive long-term data.

3.4.2 Characteristics of missing values affecting imputation approaches

Understanding whether the performance of imputation approaches depends on the temporal placement of missing values (e.g., beginning, middle, or end) and whether they are clustered or randomly distributed is crucial for selecting the most suitable imputation method based on the specific characteristics of missing data in a time series. Our results showed that no imputation approach showed superior performance when comparing error differences between the imputed and simulated population abundance values based on the position of missing data. This observation remained consistent even as the proportion of missing values increased in the simulated time series. The MICE method exhibited the lowest median error

with up to 2% missing values in the initial quarter of the time series; however, the overall differences among imputation methods were small, and MICE's performance declined as the proportion of missing values increased, especially in the latter three quarters. In contrast, the Kalman filter method consistently outperformed others, especially when more than 4% of missing values were present in the latter three quarters of the dataset. Supporting this observation, (Agbailu et al., 2021) found that Kalman filter performs well in both small and medium sample sizes when simulated using a linear trend model.

All imputation methods tested in this study performed similarly, regardless of whether gaps occurred continuously between two observed population abundance values or randomly. In support of this, Carpenter et al. (2025) found that the ability to fill gaps in a time series, generated from a five-species coupled food chain model using the multiview cross mapping (MVCM) approach, was unaffected by the number of continuous gaps between observed values. When applying this to monthly samples of *Planktothrix rubescens* from Lake Zurich, Carpenter et al. (2025) demonstrated that, even with a 6-month gap before and after each observed point, imputation performance was quite similar across multiple linear regression, k-nearest neighbor imputation, and MVCM, the latter effectively capturing the species' abundance dynamics. In contrast, testing the case of randomly occurring missing values, (Łopucki et al., 2022) emphasized that imputation results are highly dependent on the chosen method, with Gradient Boosted Trees model outperforming median imputation in a 7-year monitoring dataset on white stork offspring production. In addition, despite high imputation performance, researchers have found that imputation methods often struggle to capture extreme or high-frequency values (Chapon et al., 2023; Carpenter et al., 2025). For ecological time series datasets, relying on single imputation techniques is not recommended, as the imputed values may not adequately capture the dynamic boom and bust cycles inherent in the dataset (Nakagawa, 2015; Onkelinx et al., 2017; Raja et al., 2020; Carpenter et al., 2025). Similar concerns have been highlighted in other disciplines, including clinical data (Khan, 2024) and climate data (Sa'adi et al., 2023), where imputation methods may not fully preserve underlying patterns and variability.

The choice of imputation method had minimal impact on predicting population abundance, even when the temporal placement of missing values were taken into consideration. Results remained consistent across different imputation techniques and varying proportions of missing data. Although MICE and Mean demonstrated slightly better performance when more than 4% of missing values were present in the last three quarters, the overall influence of the imputation method on predictions was limited.

Our findings have important implications for large-scale ecological databases, such as the Global Population Dynamics Database, which compiles long-term ecological time series data across various species and ecosystems. These large datasets are often subject to missing data due to challenges in data collection, such as sampling errors, logistical limitations, or natural variability (Inchausti and Halley, 2002). By providing a framework for handling missing values in such extensive time series, our study contributes to improving the reliability and completeness of these databases. This is crucial for ongoing ecological monitoring and conservation efforts, where accurate and continuous data on population dynamics are essential for making informed decisions.

Our study demonstrates that imputation emerges as a promising and viable solution to address gaps in population abundance datasets characterized by extensive boom and bust cycles. Despite the flexibility observed in the choice of imputation methods for both imputing missing values for error comparison between the original value and predicting abundances after imputation, the outcomes consistently outperformed scenarios where missing values were simply deleted. Therefore, we encourage ecologists to tackle the challenge of missing values by adopting time-dependent imputation approaches rather than relying on deletion to improve the handling of missing data in time series analyses. Insights from this study will be used in the handling missing values in analyses conducted in Chapters 4 and 5.

References

Agbailu, A. O., Seno, A., and Clement, O. O. (2021). Kalman filter algorithm versus other methods of estimating missing values: time series evidence. *African Journal of Mathematics and Statistics Studies*, 4(2):1–9.

Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., Leos-Barajas, V., Mills Flemming, J., Nielsen, A., Petris, G., et al. (2021). A guide to state–space modeling of ecological time series. *Ecological Monographs*, 91(4):e01470.

Azur, M. J., Stuart, E. A., Frangakis, C., and Leaf, P. J. (2011). Multiple imputation by chained equations: what is it and how does it work? *International Journal of Methods in Psychiatric Research*, 20(1):40–49.

Bouhlila, D. S. and Sellaouti, F. (2013). Multiple imputation using chained equations for missing data in TIMSS: a case study. *Large-scale Assessments in Education*, 1:1–33.

Bowler, D. E., Boyd, R. J., Callaghan, C. T., Robinson, R. A., Isaac, N. J., and Pocock, M. J. (2025). Treating gaps and biases in biodiversity data as a missing data problem. *Biological Reviews*, 100(1):50–67.

Bureau of Meteorology (2024). Climate data online. http://www.bom.gov.au/climate/data/. Accessed on 26/12/2024.

Callaghan, C. T., Nakagawa, S., and Cornwell, W. K. (2021). Global abundance estimates for 9,700 bird species. *Proceedings of the National Academy of Sciences*, 118(21):e2023170118.

Carpenter, D., Deyle, E., Park, J., Saberski, E., and Sugihara, G. (2025). Repairing gaps in ecological time series. *Methods in Ecology and Evolution*. doi:10.1111/2041-210X.14491.

Chapon, A., Ouarda, T. B., and Hamdi, Y. (2023). Imputation of missing values in environmental time series by D-vine copulas. *Weather and Climate Extremes*, 41:100591.

Dhamo, E. and Puka, L. (2010). Using the R-package to forecast time series: ARIMA models and Application. In *International Conference Economic & Social Challenges and Problems*.

Dickman, C., Wardle, G., Foulkes, J., and Preu, N. d. (2014). Desert complex environments. In Burns, E., Lowe, A., Lindenmayer, D., and Thurgate, N., editors, *Biodiversity and*

environmental change: monitoring, challenges and direction, chapter 10, pages 379–438. CSIRO Publishing.

Dickman, C. R., Greenville, A. C., Beh, C.-L., Tamayo, B., and Wardle, G. M. (2010). Social organization and movements of desert rodents during population "booms" and "busts" in central Australia. *Journal of Mammalogy*, 91(4):798–810.

Ellington, E. H., Bastille-Rousseau, G., Austin, C., Landolt, K. N., Pond, B. A., Rees, E. E., Robar, N., and Murray, D. L. (2015). Using multiple imputation to estimate missing data in meta-regression. *Methods in Ecology and Evolution*, 6(2):153–163.

Frank, A. S., Dickman, C. R., Wardle, G. M., and Greenville, A. C. (2013). Interactions of grazing history, cattle removal and time since rain drive divergent short-term responses by desert biota. *PLoS One*, 8(7):e68466.

Frank, A. S., Wardle, G. M., Greenville, A. C., and Dickman, C. R. (2016). Cattle removal in arid Australia benefits kangaroos in high quality habitat but does not affect camels. *The Rangeland Journal*, 38(1):73–84.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual review of psychology*, 60:549–576.

Greenville, A. C., Wardle, G. M., and Dickman, C. R. (2013). Extreme rainfall events predict irruptions of rat plagues in central Australia. *Austral Ecology*, 38(7):754–764.

Greenville, A. C., Wardle, G. M., Nguyen, V., and Dickman, C. R. (2016). Population dynamics of desert mammals: similarities and contrasts within a multispecies assemblage. *Ecosphere*, 7(5):e01343.

Hossie, T. J., Gobin, J., and Murray, D. L. (2021). Confronting missing ecological data in the age of pandemic lockdown. *Frontiers in Ecology and Evolution*, 9:669477.

Humbert, J. Y., Scott Mills, L., Horne, J. S., and Dennis, B. (2009). A better way to estimate population trends. *Oikos*, 118(12):1940–1946.

Inchausti, P. and Halley, J. (2002). The long-term temporal variability and spectral colour of animal populations. *Evolutionary Ecology Research*, 4(7):1033–1048.

Inglis, A., Ahmed, A., Wundervald, B., and Prado, E. (2018). JAGS: Just another Gibbs sampler. *Maynooth:[sn]*.

Jóhannesson, S., Heinonen, J., and Davíðsdóttir, B. (2019). Increasing the accuracy of marine footprint calculations. *Ecological Indicators*, 99:153–158.

Khan, M. A. (2024). A Comparative Study on Imputation Techniques: Introducing a Transformer Model for Robust and Efficient Handling of Missing EEG Amplitude Data. *Bioengineering*, 11(8):740.

Kleinke, K. (2018). Multiple imputation by predictive mean matching when sample size is small. *Methodology*.

Letnic, M. (2004). Cattle grazing in a hummock grassland regenerating after fire: the short-term effects of cattle exclusion on vegetation in south-western Queensland. *The Rangeland Journal*, 26(1):34–48.

Levine, J. M., HilleRisLambers, J., Petry, W. K., Usinowicz, J., and Crowther, T. W. (2024). Demographic but not competitive time lags can transiently amplify climate-induced changes in vegetation carbon storage. *Global Change Biology*, 30(8):e17432.

Little, R. J. and Rubin, D. B. (2019). *Statistical analysis with missing data*, volume 793. John Wiley & Sons.

Łopucki, R., Kiersztyn, A., Pitucha, G., and Kitowski, I. (2022). Handling missing data in ecological studies: Ignoring gaps in the dataset can distort the inference. *Ecological Modelling*, 468:109964.

Moritz, S. and Bartz-Beielstein, T. (2017). imputeTS: Time Series Missing Value Imputation in R. *The R Journal*, 9(1):207–218.

Murray, B., Dickman, C., Watts, C., and Morton, S. R. (1999). The dietary ecology of Australian rodents. *Wildlife Research*, 26(6):857–858.

Nakagawa, S. (2015). Missing data: mechanisms, methods and messages. *Ecological statistics: Contemporary theory and application*, pages 81–105.

Nakagawa, S. and Freckleton, R. P. (2008). Missing inaction: the dangers of ignoring missing data. *Trends in Ecology & Evolution*, 23(11):592–596.

Noor, N. M. and Zainudin, M. L. (2008). A review: Missing values in environmental data sets. In *Proceeding of International Conference on Environment*.

Onkelinx, T., Devos, K., and Quataert, P. (2017). Working with population totals in the presence of missing data comparing imputation methods in terms of bias and precision.

Journal of Ornithology, 158:603–615.

Penone, C., Davidson, A. D., Shoemaker, K. T., Di Marco, M., Rondinini, C., Brooks, T. M., Young, B. E., Graham, C. H., and Costa, G. C. (2014). Imputation of missing data in life-history trait datasets: which approach performs the best? *Methods in Ecology and Evolution*, 5(9):961–970.

Purdie, R. (1984). Land systems of the Simpson desert region. *Natural Resources Series*, 2:1–71.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raja, P., Sasirekha, K., and Thangavel, K. (2020). A novel fuzzy rough clustering parameter-based missing value imputation. *Neural Computing and Applications*, 32(14):10033–10050.

Sa'adi, Z., Yusop, Z., Alias, N. E., Chow, M. F., Muhammad, M. K. I., Ramli, M. W. A., Iqbal, Z., Shiru, M. S., Rohmat, F. I. W., Mohamad, N. A., et al. (2023). Evaluating Imputation Methods for rainfall data under high variability in Johor River Basin, Malaysia. *Applied Computing and Geosciences*, 20:100145.

Schafer, J. L. and Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, 7(2):147.

Shephard, M. (1992). *The Simpson Desert: natural history and human endeavour*. Royal Geographical Society of Australasia, Adelaide, Australia.

Tulloch, A. I., Healy, A., Silcock, J., Wardle, G. M., Dickman, C. R., Frank, A. S., Aubault, H., Barton, K., and Greenville, A. C. (2023). Long-term livestock exclusion increases plant richness and reproductive capacity in arid woodlands. *Ecological Applications*, 33(8):e2909.

Van Buuren, S. (2018). *Flexible imputation of missing data*. CRC press.

Van Buuren, S. and Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45:1–67.

Verhoeven, E. M., Murray, B. R., Dickman, C. R., Wardle, G. M., and Greenville, A. C. (2020). Fire and rain are one: extreme rainfall events predict wildfire extent in an arid grassland. *International Journal of Wildland Fire*, 29(8):702–711.

Wardle, G. and Dickman, C. (2015). Desert Ecology Plot Network: Mammal, Reptile and Vegetation Data Associated with Weather, Simpson Desert, Western Queensland, Australia, 1990–2011. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern.111.57/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018a). Desert Ecology Plot Network: Mammal Abundance Plot-data, Simpson Desert, Western Queensland, 1990+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.196.8/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018b). Desert Ecology Plot Network: Vegetation Plot-data, Simpson Desert, Western Queensland, 1993+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.195.11/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018c). Desert Ecology Plot Network: Weather Data (daily and monthly), Simpson Desert, Western Queensland, 1995+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.194.7/html. Accessed on 16/01/2025.

Wardle, G. M., Greenville, A. C., Frank, A. S., Tischler, M., Emery, N. J., and Dickman, C. R. (2015). Ecosystem risk assessment of Georgina gidgee woodlands in central Australia. *Austral Ecology*, 40(4):444–459.

CHAPTER 4

Identifying optimal forecast horizon and model training length for forecasting population abundance

---

4.1 Introduction

Forecasting population abundance is critical in ecological research and wildlife management, as anticipating future changes in abundance will support decision making for the conservation and management of species (Clark et al., 2001; Dietze et al., 2018). Accurate forecasts can inform proactive measures to mitigate potential declines, protect threatened species, and manage the sustainability of ecosystems. However, ecological systems are inherently complex, characterized by nonlinear interactions, stochastic variability, and external drivers such as climate change and environmental variables (Karunarathna et al., 2024). These complexities pose significant challenges for developing robust forecasting models, emphasizing the need for careful optimization of model configurations, such as forecast horizon and training length (Petchey et al., 2015; Adler et al., 2020).

Ecological forecasting has received increasing attention in recent decades, with the growing recognition of its use in addressing pressing environmental issues (Ward et al., 2014; Lofton et al., 2022). Advances in statistical modeling and computational techniques have facilitated the development of models capable of capturing complex population dynamics (Dennis and Ponciano, 2014; Hobbs et al., 2015; Hooten and Hobbs, 2015). For example, state-space models (Buckland et al., 2004; Hostetler and Chandler, 2015; Auger-Méthé et al., 2021), generalized additive models (Wood, 2017; Pedersen et al., 2019; Clark and Wells, 2023; Karunarathna et al., 2024), and hierarchical Bayesian frameworks (Kéry and Schaub,

2011; Dorazio, 2016; Dorazio and Karanth, 2017) have been extensively employed in ecological studies to incorporate temporal trends, spatial variability, and environmental covariates (Cressie et al., 2009; Dietze, 2017). Multivariate generalized additive models (MVGAMs) stand out for their ability to model nonlinear interactions between species dynamics and abiotic factors while incorporating latent dynamic processes to capture shared trends across species or locations. This flexibility enables MVGAMs to effectively account for dependencies across multiple species and spatiotemporal scales, enhancing their applicability in ecological forecasting (Clark et al., 2025; Karunarathna et al., 2024).

Despite these advancements, the selection of optimal forecast horizon and training length remains under-explored, often resulting in suboptimal predictions that may hinder effective decision-making. The forecast horizon; the length of time into the future for which predictions are made (Coreau et al., 2009), is a critical parameter in ecological modeling. Although longer forecast horizons provide insight into distant future dynamics, they often come at the cost of increased uncertainty and reduced accuracy (Petchey et al., 2015). In contrast, shorter horizons yield more precise predictions, but may limit the scope of management strategies. Striking a balance between these aspects is crucial for ecological forecasting, as inappropriate choices of forecast horizons can compromise the use of predictions. Similarly, the length of historical data used to train models significantly influences their predictive performance. Longer training periods may capture a broader range of ecological variability, but risk overfitting or incorporating changed dynamics, while shorter periods may fail to capture essential long-term trends (Hyndman and Athanasopoulos, 2018).

In the context of population abundance forecasting, optimizing the length of both forecast horizon and training period is particularly important given the dynamic nature of populations and their sensitivity to environmental fluctuations (Dietze, 2017). Population dynamics are often driven by a combination of biotic factors, such as reproductive rates, competition and predation, and abiotic factors, including climate variability, food and shelter availability, and anthropogenic pressures (Dickman et al., 2010; Greenville et al., 2013, 2016). For example, rodent populations in arid regions are known to exhibit boom-and-bust cycles driven by rainfall variability, which requires models that account for these seasonal events while balancing

forecast horizons and training lengths (Lima et al., 2008; Greenville et al., 2016; Clark et al., 2025; Karunarathna et al., 2024). The interplay between these drivers further complicates forecasting efforts, underscoring the importance of model calibration and validation to ensure reliable predictions.

Probabilistic forecasting offers a promising framework for improving the reliability of ecological predictions by explicitly accounting for uncertainty (Dietze, 2017). Unlike point forecasts that provide single-valued predictions, probabilistic forecasts generate distributions of future outcomes, enabling more robust evaluations of uncertainty and model performance. Tools such as scoring rules; widely employed in both univariate and multivariate modeling contexts, enable simultaneous evaluation of forecast accuracy and uncertainty (Gneiting and Raftery, 2007; Keune et al., 2014; Scheuerer and Hamill, 2015; Simonis et al., 2021; Bjerregård et al., 2021; Alexander et al., 2024). These metrics are particularly valuable in ecological forecasting, where uncertainty is driven by intricate interactions among environmental variables, species-specific dynamics, and temporal variability. By leveraging probabilistic approaches, researchers can more effectively assess the trade-offs between forecast horizons and training lengths, optimizing model usage across diverse ecological systems. Adler et al. (2020) highlights that the accuracy of forecasts depends not only on the models' temporal dynamics but also on their consideration of spatial correlations, emphasizing the need for a more integrated approach.

In this study, we used the long-term monitoring data (1990 - 2022) from the Desert Ecology Research Group (DERG dataset; see Chapter 2 for more details) on two species that show contrasting dynamics (i.e., one species that has highly varying population dynamics while the other has stable dynamics). The aims of this study were to 1) identify the most suitable multivariate generalized additive model (MVGAM) structure to capture nonlinear relationships between species dynamics and external drivers (building on the findings from Chapter 2, where we determined that the MARSS framework was not effective for forecasting species abundance), and 2) determine the optimal combination of training length and forecast horizon for each species. To achieve these objectives, we first fitted MVGAMS with three different trend components; Autoregressive (AR), Vector Autoregressive (VAR) and Gaussian

Process (GP). The best-fitting model was selected based on a combination of leave-one-out cross-validation technique, hindcast accuracy, and computational efficiency. We first filtered models based on hindcast performance; specifically, those that closely matched the observed capture patterns and captured a higher proportion of observed counts within the 95% prediction intervals. From this subset, the final model was chosen by prioritizing predictive accuracy alongside reasonable average computation time, ensuring a balance between predictive performance and practical feasibility. Using the selected model, we then explored different combinations of training lengths and forecast horizons, evaluating the predictions of species abundance through a scoring rule. We hypothesize that the MVGAM with the GP trend will offer the best predictive performance due to its ability to capture smooth, nonlinear trends and latent processes driving population fluctuations. We also expect that the MVGAM with the VAR trend will outperform the AR trend model, as it can account for independent trends across different species/locations, whereas the AR model assumes a shared trend across all species/locations. Despite these differences in model structure, we anticipate similar computational efficiency across all three models. Additionally, we hypothesize that longer training lengths will have a significant improvement in the model performance by incorporating a broader range of population variability. However, the specific point at which the training length is split is a critical factor in determining the extent of this improvement. We predict that shorter forecast horizons will yield more precise forecasts, while longer horizons will exhibit increasing uncertainty, with forecast performance declining more rapidly for species with higher population variability. Finally, we expect that the optimal combination of training length and forecast horizon will differ between the two species, reflecting their distinct population dynamics and sensitivities to environmental factors, thus highlighting the importance of species-specific model calibration.

4.2 Material and methods

4.2.1 Study location

The long-term sampling carried out in the Simpson Desert in central Australia, was conducted by the Desert Ecology Research Group (DERG) at the University of Sydney since

1990 (Wardle and Dickman (2015), Wardle and Dickman (2018a), Wardle and Dickman (2018b), Wardle and Dickman (2018c), unpublished data from 2018). The Simpson Desert covers 170,000 km^2 of mainly dune fields (73%), with the remainder comprising clay pans, rocky outcrops, and gibber plains (Shephard, 1992). The study region of 8000 km^2), spans the border between Queensland and the Northern Territory (Fig. 2.1) and the vegetation of the dune fields consists primarily of spinifex grasslands (*Triodia basedowii*)with scattered shrubs (*Acacia*, *Eremophilla*, *Grevillea*), ephmeral forbs and a variety of other perennial and annual grasses; interspersed are wide swales dominated by Georgina Gidgee woodlands (*Acacia georginae*), or mallee eucalypts (Wardle et al., 2015). The region experiences highly variable climatic conditions, with prolonged dry periods, or "busts", elevated by productive "boom" phases following rainfall in summer (Dickman et al., 2010; Greenville et al., 2013). Temperatures can exceed 40°C in summer and drop below 5°C in winter (Purdie, 1984). Rainfall is seasonal, with annual averages decreasing from northwest to southeast: 258.5 mm, 203 mm, and 157 mm at Boulia (136 years), Bedourie (79 years), and Birdsville (24 years) respectively (Bureau of Meteorology, 2024). Vegetation in the region is directly influenced by key ecological factors such as wildfires (Letnic, 2004) and cattle grazing (Frank et al., 2013; Tulloch et al., 2023), with interactive effects shaping the abundance of fauna species in the region. The mean minimum wildfire return interval in the region is 27 years (Verhoeven et al., 2020). Cattle were removed from the study area in 2004 (Ethabuka Reserve) and 2006 (Pilungah Reserve) (Frank et al., 2016) (see Chapter 2 for further details).

4.2.2 Study species

Among the many small mammal species (Dickman et al., 2014) captured during live-trapping, for this study, our primary focus was on two species exhibiting contrasting boom-and-bust dynamics: the rodent *Pseudomys hermannsburgensis* (sandy inland mouse, 12 g) (Fig. A2.3(B)), the small insectivorous dasyurid *Sminthopsis youngsoni* (lesser hairy-footed dunnart, 10 g) (Fig. A2.4(B)), and the carnivorous mulgara *Dasycercus blythi* (brush-tailed mulgara, 100 g), a predator of both the rodent and the small dasyurid, across all the seventeen sites (Fig. 2.1).

Capture data were analyzed at the grid level, as differences in the counts were observed among grids within a site (Fig. A2.3(C) and Fig. A2.4(C)). To ensure data completeness and reduce the bias from having many gaps, we selected grids with missing values (NAs) below the 25^{th} percentile for each site. However, when a site had only two grids, both were included regardless of the missing data criterion. This selection process resulted in a total of 34 grids across all sites (each site having 2 grids; Fig. 4.1).

The standardized capture counts per 100 trap nights (TN) from 1990 - 2022 were averaged biannually according to the mean minimum temperature and mean rainfall across the years 1998 - 2024. This was performed to reduce the loss of information by aggregating annually and to make the most out of the sampling effort across the years, allowing us to better capture ecological processes that might influence population dynamics. Since rainfall in this region primarily occurs during the warmer months and the area lacks distinct four-season cycles, biannual aggregation provided a more ecologically relevant timescale. In addition, a longer timeframe (1998 - 2024) was used to determine the temperature and rainfall thresholds to account for recent climate change, as typically we would use a longer range of years, such as 100 years or the longest available historical data. According to data from two weather stations closest to our study site, Bedourie (1998 - 2004) and Birdsville Airport (2000 - 2024), the mean minimum temperature has remained below 15°C and the mean rainfall has been under 10 mm from May to September and the temperatures have increased during the months from October to April (Fig. 4.2, Fig. A2.1 & Fig. A2.2). Therefore, we considered the bi-annual separation as **'cooler months'** (from May to September) and **'hotter months'** (from October to April) and the species captures were averaged accordingly.

A variable 'time' was created to represent the sampling time points on a continuous scale, ranging from 1 (corresponding to the cooler period of 1990) to 66 (corresponding to the hotter period of 2022), reflecting the biannual separation into cooler and hotter months across the years. Missing records were denoted as NA (not available) in the response variable. Since the non-negative real-valued distribution family used in our modelling framework did not accommodate zero observations (Fig. A2.3(A) & Fig. A2.4(A)), we added 0.01 to the captures (Douma and Weedon, 2019).

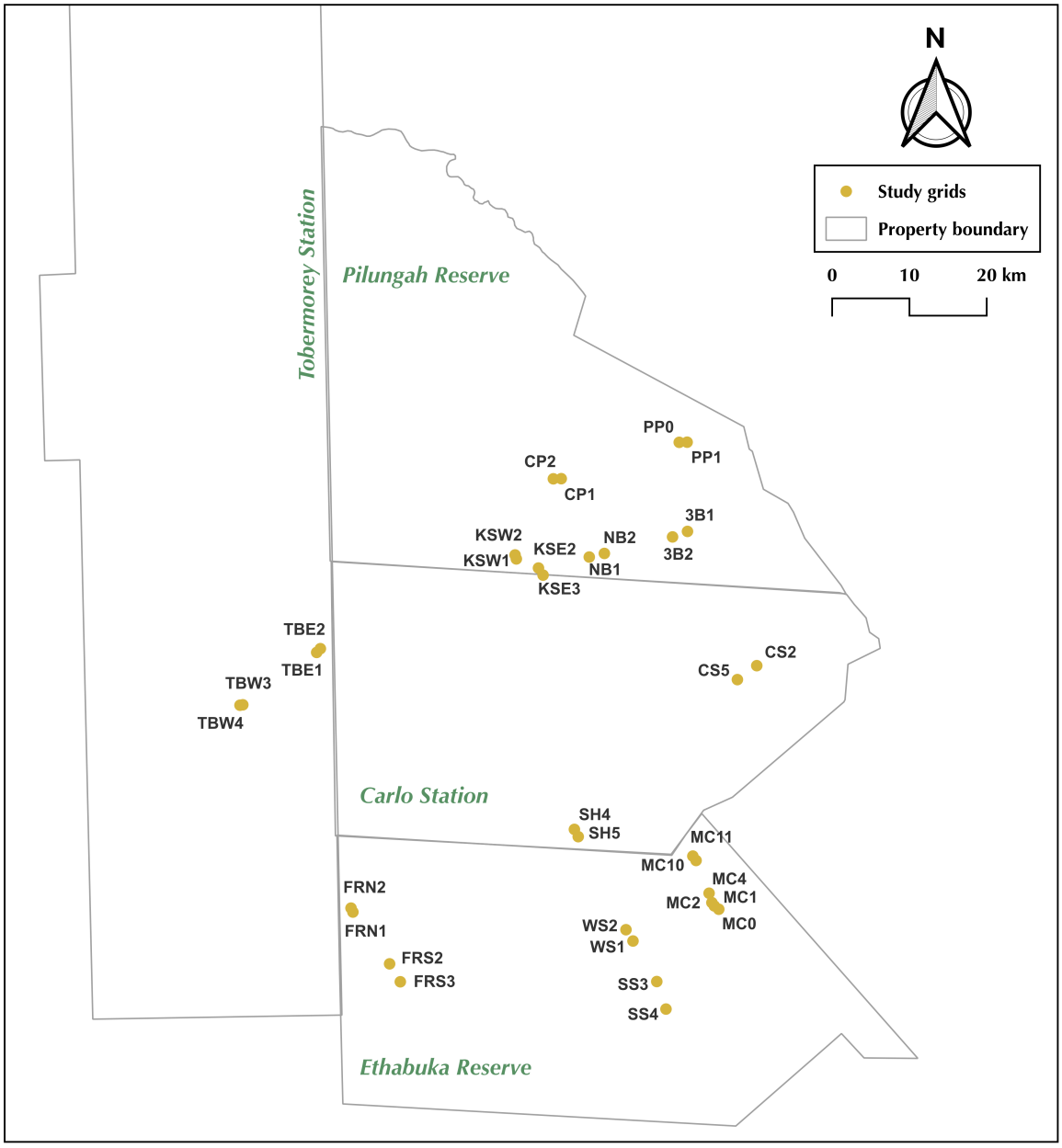

FIGURE 4.1. Thirty-four (34) study grid locations (2 grids per each seventeen sites) across Pilungah Reserve (previously Cravens Peak Reserve), Tobermorey Station, Carlo Station and Ethabuka Reserve, Simpson Desert, Australia.

4.2.3 Covariate measurements

Resource availability and climatic conditions are crucial factors that affect the population abundance of species as they influence the food availability, shelter, weather and possibility of predation (Dickman et al., 2014; Greenville et al., 2016). Therefore, from the DERG dataset, we used cover, seeding, and flowering of the dominant vegetation species, spinifex (*Triodia basedowii*), along with captures of *D. blythi* and the previous year rainfall as predictors in our models (see Chapter 2). Estimates of cover, seed productivity, and flowering were aggregated by grid and trip (herewith 'Avg. Spinifex Cover', 'Spinifex Seeding Score' and 'Spinifex Flowering Score', respectively) over a 33-year period for each of the seventeen sites (34 grids in total).

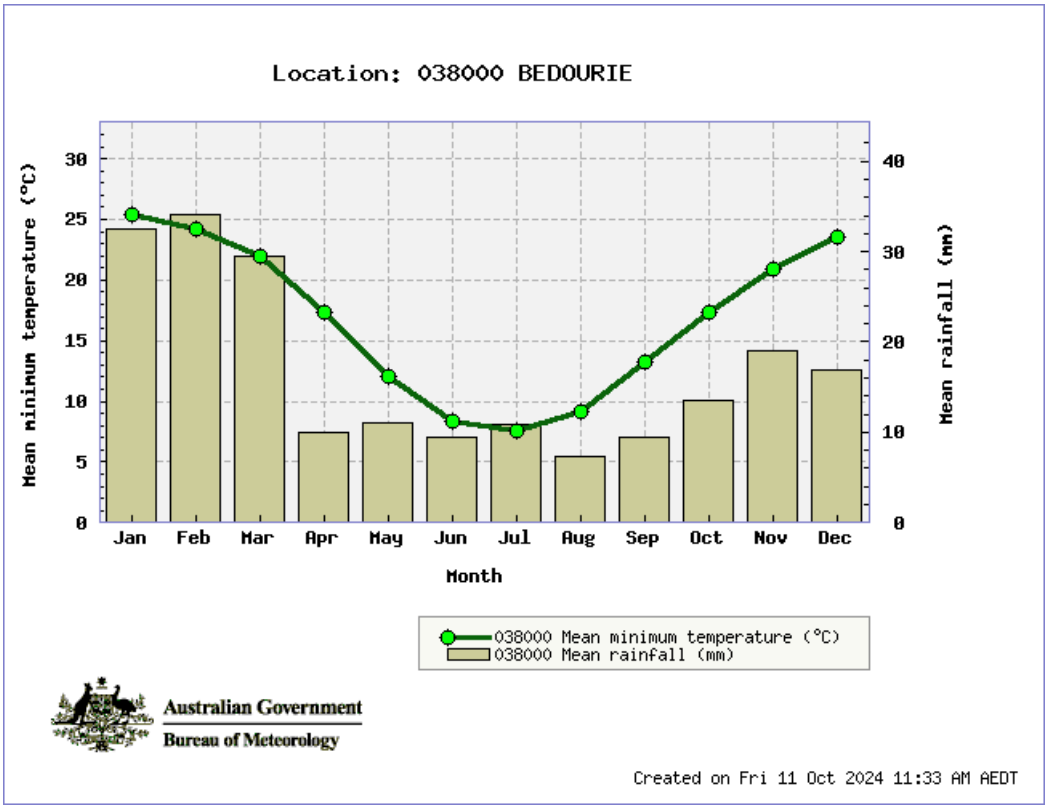

FIGURE 4.2. Mean Minimum Temperature (°C) and Mean Rainfall (mm) across the years 1998 - 2024 at Bedourie weather station.
Source: Bureau of Meteorology, Australian Government

Total rainfall for each sampling month was calculated using daily data from automatic weather stations (Environdata, Warwick, Queensland) located at each of the seventeen sites (see Fig. 2.1 for weather station locations). For consistency, we assumed that all grids within a site experienced the same monthly rainfall. These weather stations were operational from 1995 to 2022. Daily rainfall data prior to 1995 were obtained from the nearest Bureau of Meteorology weather stations to the study region (see Data preparation section of Chapter 2).

All covariate measurements were averaged based on the cooler and hotter months separation described in Section 5.2.1. Avg. Spinifex Cover increased from approximately 10 to 50 by the year 2000, with reduced dispersion (Fig. A2.3(D)). After 2000, however, the dispersion increased, likely due to the heavy rainfall event followed by wildfire (Greenville et al., 2009) that occurred in that year (Fig. A2.3(H)). A nonlinear relationship was also observed between Avg. Spinifex Cover and the previous year's rainfall (Fig. A2.3(N)). The Spinifex Flowering Score and Spinifex Seeding Score did not exhibit clear seasonal patterns. However, both flowering and seeding increased approximately every four years after 2006 (Fig. A2.3(E, F)). Captures of *D. blythi* peaked in 2001 following the heavy rainfall in 2000, but were consistently low in most of the sampled years. Additionally, captures of both *P. hermannsburgensis* and *S. youngsoni* appeared to increase with the previous year rainfall, indicating a nonlinear relationship between the variables (Fig. A2.3(M) & Fig. A2.4(F)). Therefore, we considered interactions between Avg. Spinifex Cover and previous year rainfall when modeling the species abundance.

Exploration of the relationships between captures of *P. hermannsburgensis* and covariate measurements revealed correlations with Avg. Spinifex Cover, Spinifex Seeding Score, and Spinifex Flowering Score; however, these relationships appeared to be nonlinear (Fig. A2.3(I, J, K)). A similar relationship was observed for *S. youngsoni*, but only with Avg. Spinifex Cover (Fig. A2.4(D)), as this species does not rely on Spinifex seeding or flowering (Bennison et al., 2018). Additionally, captures of both *P. hermannsburgensis* and *S. youngsoni* were observed to decrease significantly with increasing *D. blythi* captures, indicating a negative correlation (Fig. A2.3(L)) & (Fig. A2.4(E)). Therefore, nonlinear relationships

between the study species and vegetation variables, previous year rainfall and predator activity were considered in the model.

Since our modelling framework did not allow for missing covariate values, we imputed the missing values using the multiple imputation technique - Multivariate Imputation by Chained Equations as implemented through the *mice* package in R (Van Buuren and Groothuis-Oudshoorn, 2011). We used Predictive Mean Matching (PMM) as the imputation algorithm and averaged the imputations from 5 different imputed datasets for each covariate (see Chapter 3 for more details).

4.2.4 Model training lengths and forecast horizons

To investigate the optimal training length and forecast horizon, we first divided the full dataset of 66 time points, spanning from 1990 to 2022, into training and testing sets using a 70/30 split. This resulted in a training set with 46 time points and a testing set with the remaining 20 time points. To explore the effect of different training lengths, the training set was further divided into two subsets. Training set 1 contained 23 time points, equivalent to 50% of the training set, while Training set 2 included 37 time points, corresponding to 80% of the training set.

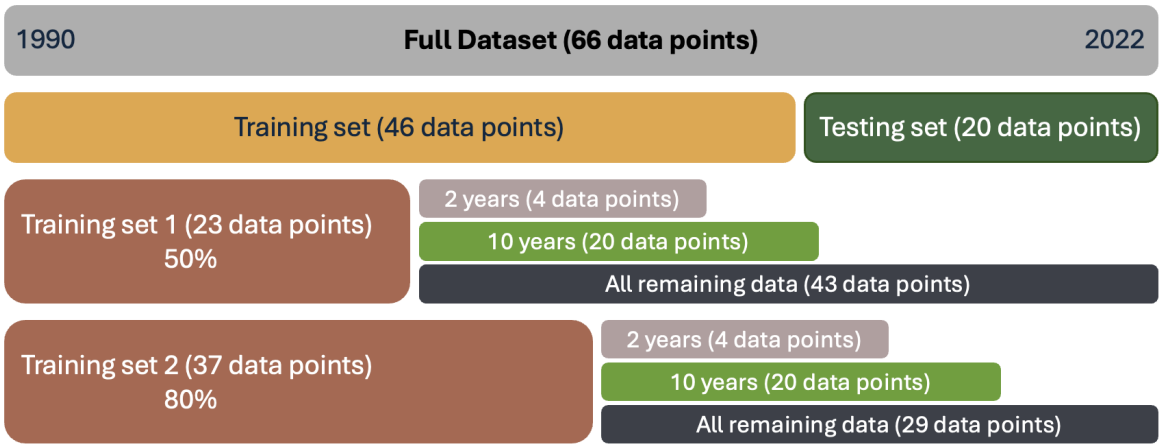

FIGURE 4.3. Full dataset split into two training sets (Training set 1 and Training set 2) and three forecast horizons (2 years, 10 years and all remaining data).

For each training set (i.e. Training set 1 and Training set 2), we examined two forecast horizons: 2 years (4 time points) and 10 years (20 time points), as well as an additional forecast horizon that considered 'all remaining data' until 2022 (43 and 29 time points for Training set 1 and Training set 2 respectively). When comparing the species population abundance forecasts for 2 years and 10 years, we can directly compare them across the two training sets as the forecast horizons are consistent. However, we should be cautious when comparing the forecasts for the 'all remaining data' horizon across the two training sets as the number of time points being forecast differs (Fig. 4.3).

4.2.5 Multivariate generalized additive models (MVGAM)

The multivariate generalized additive framework models the average of captures θ_{it} for each grid i at a given time t based on the observed capture data Y_{it} as:

$$Y_{it} \sim f_{Y_{it}}(\theta_{it}) \tag{4.1}$$
$$\log(\theta_{it}) = S_1(x_1) + S_2(x_2) + \cdots + S_k(x_k)$$

Here, f denotes an exponential probability distribution, while the smooth functions of covariates (X) are represented by $S_i's$. These smooth functions can be multidimensional, such as tensor product smooths (te) or tensor product interactions (ti) involving multiple covariates (Pedersen et al., 2019).

As part of our methodology, we aimed to identify the most suitable model structure to compare performance across different training lengths and forecast horizons. To achieve this, we fitted three Bayesian MVGAMs with the same covariate specification but differing latent trend components (to account for autocorrelation) for *P. hermannsburgensis*. These models were fitted using Training set 1 (Section 4.2.4), and the optimal model structure was selected based on three key criteria: (1) Pareto-smoothed importance sampling leave-one-out cross-validation (PSIS-LOO), which estimates leave-one-out predictive accuracy through the Estimated Log Predictive Density (ELPD) (Vehtari et al., 2017), (2) the computational efficiency, measured by the time required to fit each model, and (3) the hindcast accuracy

evaluated on the training set. The selected model, referred to as the 'final model,' was subsequently used for performance comparisons across different training lengths and forecast horizons.

For *S. youngsoni*, we adopted the same statistical model structure as the final model for *P. hermannsburgensis*. However, the covariate specification differed, as only Avg. Spinifex Cover and previous year rainfall were included (Section 4.2.3). This adjustment reflected differences in ecological drivers between the two species.

To model the grid-level captures, we used a Gamma distribution. The latent trend components evaluated were: an Autoregressive model of order 1 (AR1), a Vector Autoregressive model with uncorrelated errors of order 1 (VAR1), and a Gaussian Process (GP). While AR1 captures short-term dependencies with a simple autoregressive structure, VAR allows for multiple independent trends without capturing cross-correlations (see Appendix 2 for further details), and GP provides a highly flexible, smooth representation of trends by modeling complex, nonlinear dependencies. These models were represented as ***GAM-AR***, ***GAM-VAR***, and ***GAM-GP***, respectively. This approach allowed us to identify the most appropriate model structure while accounting for autocorrelation and species-specific covariate effects, ensuring robust comparisons across training lengths and forecast horizons. The general structure of the three models, ***GAM-AR***, ***GAM-VAR***, and ***GAM-GP***, is the same, with the only difference being how the term Z_{it} is specified. The model specification presented below applies to all three versions, except for the formulation of Z_{it}, which varies across models. These variations are detailed in Appendix 2. Specifications of the final model (***GAM-GP***) are as follows:

$$y_{it} \sim \text{Gamma}(\mu_{it}, \emptyset) \tag{4.2}$$

$$\begin{aligned}\log(\mu_{it}) =& \alpha_{\text{grid}[i]} + S_1(\textit{D. blythi}) + S_2(Spinifex\_Fl\_Score) + \\ & S_3(Spinifex\_Seed\_Score) + S_4(Spinifex\_Avgcover, Rainfall_{(t-2)}) + Z_{it}\end{aligned}$$

where,

$i = 1, 2, \ldots, 34$

$\mu_{it} = E(y_{it})$: is the expected (average) captures per 100TN (TN = Trap Nights) for grid i at time t

$\alpha_{\text{grid}} \sim \text{Normal}(\mu_{\text{grid}}, \sigma^2_{\text{grid}})$: effect of grid i (random effect)

$\mu_{\text{grid}} \sim \text{Normal}(0, 1)$

$\sigma^2_{\text{grid}} \sim \text{Exponential}(0.5)$

$S_1(\textit{D. blythi})$: smooth function represents effects of the predator, *D. blythi*

$\beta_1 \sim \text{MultiNormal}(0, \Sigma^{-1}\lambda)$: coefficients in smooth function $S_1(\textit{D. blythi})$

$S_2(Spinifex\_Fl\_Score)$: smooth function represents effects of Spinifex Flowering Score

$\beta_2 \sim \text{MultiNormal}(0, \Sigma^{-1}\lambda)$: coefficients in smooth function $S_2(Spinifex\_Fl\_Score)$

$S_3(Spinifex\_Seed\_Score)$: smooth function represents effects of Spinifex Seeding Score

$\beta_3 \sim \text{MultiNormal}(0, \Sigma^{-1}\lambda)$: coefficients in smooth function $S_3(Spinifex\_Seed\_Score)$

$S_4(Spinifex\_Avgcover, Rainfall_{(t-2)})$: smooth function represents interaction effects of Avg. Spinifex Cover and previous year Rainfall

$\beta_4 \sim \text{MultiNormal}(0, \Sigma^{-1}\lambda)$: coefficients in smooth function $S_4(Spinifex\_Avgcover, Rainfall_{(t-2)})$

$\Sigma^{-1} \sim$ basis expansion function (supplied by mgcv function)

$\lambda \sim \text{Normal}(30, 25)$

$\frac{1}{\emptyset} \sim \text{Exponential}(5)$

$Z_{it} \sim \text{MultiNormal}(0, K)$: Gaussian Process latent trend component

$$K(t, t') = g^2_{GP} \exp\left(\frac{-|t - t'|^2}{\rho^2_{GP}}\right)$$

$g_{GP} \sim \text{Normal}(0, 0.5)$

$\rho_{GP} \sim \text{InverseGamma}(1.49, 5.67)$

We used smooth terms to model nonlinear relationships between the species capture counts and the covariates. A shrinkage version of the thin plate regression spline (bs = 'ts') was used to model the effect of predator abundance (*D. blythi*) with a single basis function ($k = 1$). Separate smooth terms were specified for Spinifex Flowering Score (k = 2) and Spinifex Seeding Score (k = 2), both using thin plate splines with shrinkage. A tensor product smooth term was used to model the interaction between Avg. Spinifex Cover and previous year rainfall, capturing the combined influence of vegetation cover as an environmental factor and rainfall as a climatic factor. Thin plate splines (bs = 'tp') were employed as the marginal bases, with $k = 4$ basis functions for average cover and $k = 6$ for rainfall. This specification allowed for flexible modeling of the interaction while maintaining interpretability. The marginal smoothing penalties ensured sufficient flexibility without overfitting.

An exponential prior was assigned to the inverse of the dispersion parameter ($\emptyset$) in the Gamma distribution, favoring larger values to accommodate overdispersion. This prior structure allowed the model to capture variance greater than the mean while retaining the flexibility to approximate a standard Gamma distribution when evidence for overdispersion was weak. To account for latent temporal trends in population abundance (X_{it}), a Gaussian Process (GP) with a squared exponential covariance kernel was incorporated.

Models were fitted using the *mvgam()* function from the *mvgam* package in R (Clark and Wells, 2023). To address data sparsity and prevent implausible values, a combination of informative and weakly informative priors was specified. Posterior distributions were estimated using Hamiltonian Monte Carlo (HMC) sampling via the CmdStan interface (Stan Development Team, 2024). The models were run with four parallel Markov chains, each generating 2500 samples, with a burn-in period of 250 and thinning by a factor of 10. Computation was performed using four threads and eight cores for efficiency. Convergence was assessed by examining Rhat values (all of which were < 1.05), effective sample sizes, and HMC divergences, ensuring minimal estimator bias. All analyses were performed in R 4.4.1 (R Core Team, 2024).

4.2.6 Model evaluation strategy - Energy Score

We used the final model to compare performance across two training sets (Training set 1 and Training set 2) and three forecast horizons (see Section 4.2.4). To evaluate out-of-sample forecast accuracy, we employed the Energy Score (ES; Gneiting et al. (2008)), a multivariate proper scoring rule. The Energy Score generalizes the univariate Continuous Ranked Probability Score (CRPS; Matheson and Winkler (1976); Hersbach (2000); Gneiting and Raftery (2007); Alexander et al. (2024)) assessing both the accuracy of marginal distributions (how well each forecast matches the observed counts) and the dependence structure between forecast variables (how well the forecast captures the relationships between different grids).

The Energy Score evaluates probabilistic forecasts by computing the energy distance (typically using Euclidean distance) between the forecast distribution and the observed outcome. Let $\mathbf{y} = (y_1, \ldots, y_n)'$ be an observation of the n-dimensional random vector $\mathbf{Y}$ and let $F_{\mathbf{X}}$ denote the multivariate forecast distribution of $\mathbf{Y}$, where $\mathbf{x_i}$ and $\mathbf{x_j}$ are independent random vectors sampled from $F_{\mathbf{X}}$. By drawing M Monte Carlo realizations from the predictive cumulative density function (CDF) $F_{\mathbf{X}}$, the Energy Score is defined as:

$$\text{ES}(F_{\mathbf{X}}, \mathbf{y}) = \frac{1}{M} \sum_{i=1}^{M} \|\mathbf{x_i} - \mathbf{y}\| - \frac{1}{2M^2} \sum_{i=1}^{M} \sum_{j=1}^{M} \|\mathbf{x_i} - \mathbf{x_j}\|, \tag{4.3}$$

where $\| \cdot \|$ denotes the Euclidean norm. The first term represents the average distance between each forecast sample $\mathbf{x_i}$ and the observed outcome $\mathbf{y}$, while the second term accounts for the pairwise distances between all forecast samples $\mathbf{x_i}$ and $\mathbf{x_j}$, capturing the spread and dependence structure of the forecast distribution. This formulation characterizes probabilistic forecasts with a single scalar value, summarizing both the accuracy of individual forecasts and the relationships between forecast variables. The interpretation of the Energy Score is similar to that of the CRPS: a smaller score indicates a better probabilistic forecast. Specifically, a lower Energy Score reflects 1) improved calibration, meaning the forecast assigns high probability to values close to the observed outcome; 2) sharper predictions, referring to a concentrated forecast distribution that avoids excessive uncertainty; and 3)

a more accurate dependence structure, capturing the correct relationships among multiple forecast variables, the grids.

Pinson and Girard (2012) and Pinson and Tastu (2013) highlight that the Energy Score is sometimes ineffective for evaluating the dependency structure between grids of the forecasted multivariate distribution. Therefore, as an additional approach, we also employed the square root of the Weighted Variogram Score, a scoring rule that penalizes poorly calibrated forecasts and failing to capture inter-species or inter-location correlations, as mentioned in Clark et al. (2025), to assess whether the results would differ significantly from those obtained with the Energy Score.

4.3 Results

4.3.1 Model selection

Training Set 1 (50%) included captures of *P. hermannsburgensis* across 34 grids over 23 time points. All models exhibited comparable hindcasts for captures from 1990 to 2001 (Training Set 1), with the 95% prediction interval successfully capturing nearly all observed counts. For example, results for the Main Camp South grid are shown in Fig. 4.4, and these findings were consistent across all grids. However, the posterior hindcasts tended to be flat and failed to capture the observed boom-bust cycles, underestimating both peaks and troughs. This suggests that while the models were able to account for general trends and uncertainty during the training window, they struggled to represent extreme fluctuations in abundance.

Differences among the models primarily emerged in terms of their generalization to new data within the training window and computational efficiency, as evaluated through ELPD differences and runtime metrics. The ***GAM-VAR*** model, despite having the highest ELPD, indicating better predictive performance, required the most computational time (4.00 hours). In contrast, the ***GAM-GP*** model, though slightly less accurate (ELPD difference of -3.4), was much faster to fit (3.30 minutes), making it the most computationally efficient option.

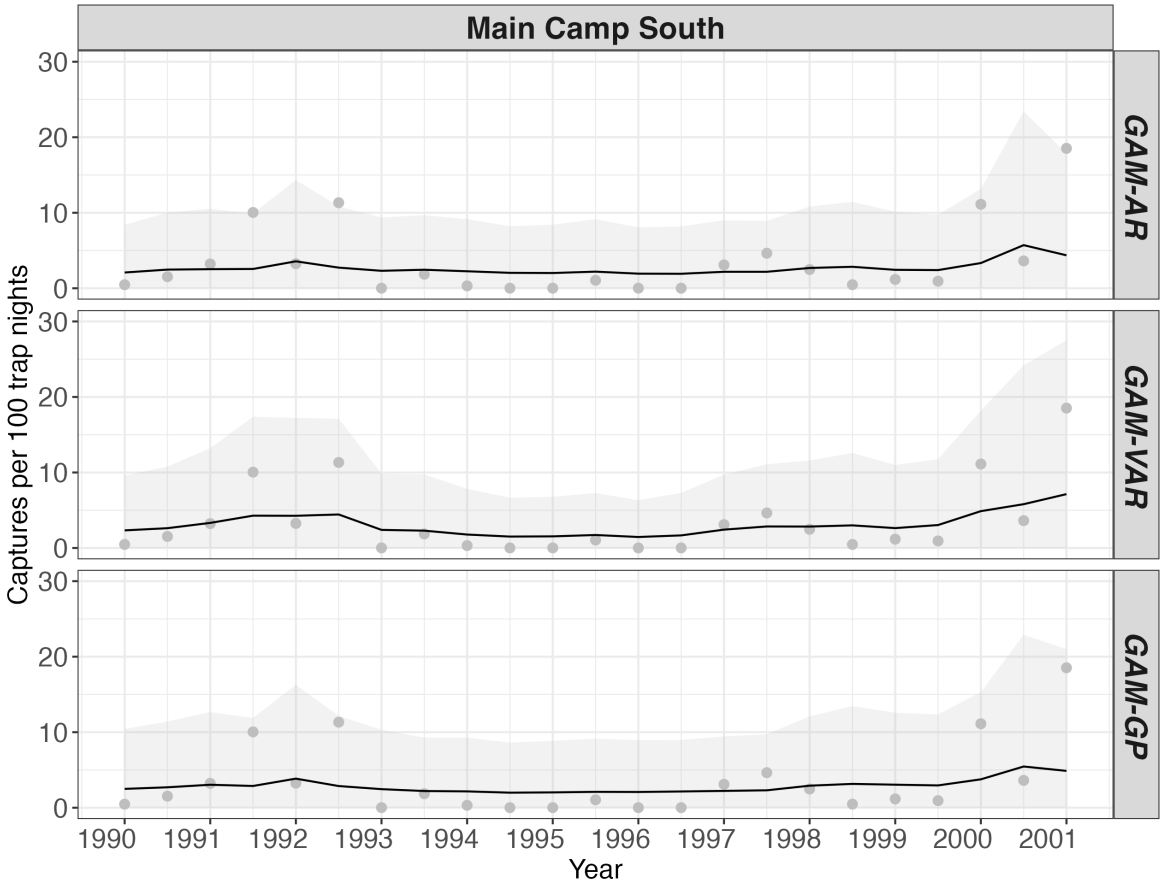

FIGURE 4.4. Hindcasts and uncertainty representation from the three multivariate generalized additive models with three varying latent trend components for *Pseudomys hermannsburgensis* captures at the Main Camp South grid, based on Training set 1 (50%). Dots indicate observed captures, black solid line represents posterior hindcasts, and the grey shaded region depicts the 95% prediction interval.

The ***GAM-AR*** model showed the lowest ELPD and took 37 minutes to fit. Given the trade-off between predictive performance and computational efficiency, the ***GAM-GP*** model was selected as the final model (Table 4.1). Its reasonable predictive performance, combined with its quick fitting time, made it the optimal choice, particularly for scenarios requiring frequent model updates or application to larger datasets.

The significance of covariates in the ***GAM-GP*** model differed from what was observed in the MARSS model representation in Chapter 2. Surprisingly, the significance of the covariates varied between the two training sets. When trained on Training Set 1, no covariates were significant, whereas in Training Set 2, the effect of the predatory mulgara *D. blythi* was found

to be significant (Table A2.1). In contrast, with updated model assumptions for the MARSS model in Chapter 2, we observed that previous year rainfall and spinifex seed were significant predictors of *P. hermannsburgensis* abundance, aligning with findings from Greenville et al. (2016). On the other hand, when the final model was fitted for *S. youngsoni* (***GAM-GP(only cover)***), with revised covariates, none of the predictors was found to be significant (Table A2.1). This result was consistent with the updated MARSS model in Chapter 2 but diverged from the findings of Greenville et al. (2016).

TABLE 4.1. Calculating the Estimated Log Predictive Density (ELPD) using pareto-smoothed importance sampling leave-one-out cross-validation (PSIS-LOO) for the three models fitted on Training set 1 (50%) for *Pseudomys hermannsburgensis*. Higher ELPD values indicate better generalization to new data within the training window. The final column presents the computation times for each model.

Model	Trend Model	ELPD difference	SE of ELPD difference	Model fit time
GAM-VAR	VAR1	0.0	0.0	4.00 hrs
GAM-GP	GP	-3.4	1.1	3.30 mins
GAM-AR	AR1	-6.0	2.1	37.00 mins

4.3.2 Optimal training length and forecast horizon for species with contrasting dynamics

We evaluated the forecast performance across the three forecast horizons, fitting the final model with both Training set 1 and Training set 2 for both species: *P. hermannsburgensis* (Model name: ***GAM-GP***) and *S. youngsoni* (Model name: ***GAM-GP (only cover)***, as only Avg. Spinifex Cover was included as the environmental factor). Despite the contrasting dynamics between the species, both models required no more than 5 minutes for fitting, with minimal differences in time between the two training sets. When forecasting for the 'all remaining data' horizon for captures of *P. hermannsburgensis* using the shorter training set

(Training set 1), approximately 4 minutes were required, which was slightly longer than the time taken to forecast for the other horizons. However, the forecast times remained under 5 minutes for all three forecast horizons (Table 4.2). Overall, despite the contrasting dynamics between the two species, the model fitting times for varying training lengths and forecasting across multiple horizons were efficient. Both models required minimal computational time, even for longer forecast periods. This shows that, regardless of the differing dynamics of the species, the models are computationally efficient and capable of handling both short and long forecast horizons effectively.

TABLE 4.2. Run times for fitting the final models of *Pseudomys hermannsburgensis* (***GAM-GP***) and *Sminthopsis youngsoni* (***GAM-GP(only cover)***) on Training set 1 and Training set 2 and generating forecasts across three forecast horizons (2 years, 10 years and all remaining data).

Model	**Training set**	**Model fit time (mins)**	**Forecast time (mins)**		
			2 years	**10 years**	**All remaining data**
GAM-GP	Set 1 - 50%	3.30	3.26	3.38	4.01
	Set 2 - 80%	5.00	1.56	2.08	2.14
GAM-GP(only cover)	Set 1 - 50%	3.50	1.25	1.35	1.49
	Set 2 - 80%	3.50	2.23	2.32	2.34

For *P. hermannsburgensis*, the Energy Scores (ES) for the final model fitted with Training set 1 (50%) and Training set 2 (80%) show relatively higher mean values compared to those of *S. youngsoni*, particularly for the 2 year and 10 year forecast horizons (Fig. 4.5). This trend aligns with the species' ecological characteristics, as *P. hermannsburgensis* exhibits population dynamics that are more sensitive to environmental variability and seasonal changes, characterized by boom-bust dynamics. As anticipated, the mean ES decreases (when comparing only between the 2 year and 10 year forecast horizons) when the model is trained with a larger dataset, as it benefits from more data, improving its ability to forecast with greater accuracy. However, the overlap in confidence intervals between the 50% and 80% training

sets indicates that the forecast performance does not differ significantly between these training lengths, and this finding is consistent across both the 2 year and 10 year horizons. For the 'all remaining data' horizon, the ES exhibits greater variability, as reflected by wider confidence intervals. This increased variability likely arises from the challenges of forecasting over an extended period, combined with the species' inherently dynamic and fluctuating population trends. Even so, training the model with 50% of the data and forecasting for 43 time points, compared to training with 80% of the data and forecasting for 29 time points, does not result in a significant difference in forecast performance (Fig. 4.5). We also tested the ***GAM-GP*** model by replacing the 'Spinifex' variables with the 'Overall Vegetation' variables, which incorporate all vegetation species surrounding the grids ('Avg. Overall Cover', 'Overall Seeding Score' and 'Overall Flowering Score'). The results showed the same trends (Fig. A2.5), confirming that the combined effect of all vegetation species is equivalent to that of the dominant vegetation species. Therefore, the environmental variables of the dominant vegetation species can serve as a reliable proxy for the overall vegetation species.

In contrast, the ES for *S. youngsoni* are generally lower and exhibit less variability across the 2 year and 10 year forecast horizons, indicating more stable model performance (Fig. 4.5). This stability aligns with the ecological characteristics of *S. youngsoni*, whose population dynamics are relatively consistent and primarily influenced by Avg. Spinifex Cover and the previous year's rainfall - the key covariates included in the model. Similar to *P. hermannsburgensis*, minimal differences in ES are observed between Training set 1 (50%) and Training set 2 (80%), as indicated by overlapping confidence intervals for the 2 year and 10 year horizons. While the 'all remaining data' horizon shows slightly wider confidence intervals, the overall stability in ES supports the conclusion that *S. youngsoni*'s ecological trends remain consistent, even over extended forecast periods (Fig. 4.5).

Overall, these observations indicate that both the ***GAM-GP*** model for *P. hermannsburgensis* and the ***GAM-GP (only cover)*** model for *S. youngsoni* show robust performance across different training splits and forecast horizons (Fig. 4.5). The Weighted Variogram Scores (Fig. A2.6) reveal similar trends, although the values are larger due to the inherent nature of this

scoring metric. This supports the conclusion that the Energy Score adequately captures the dependency structure between the grids.

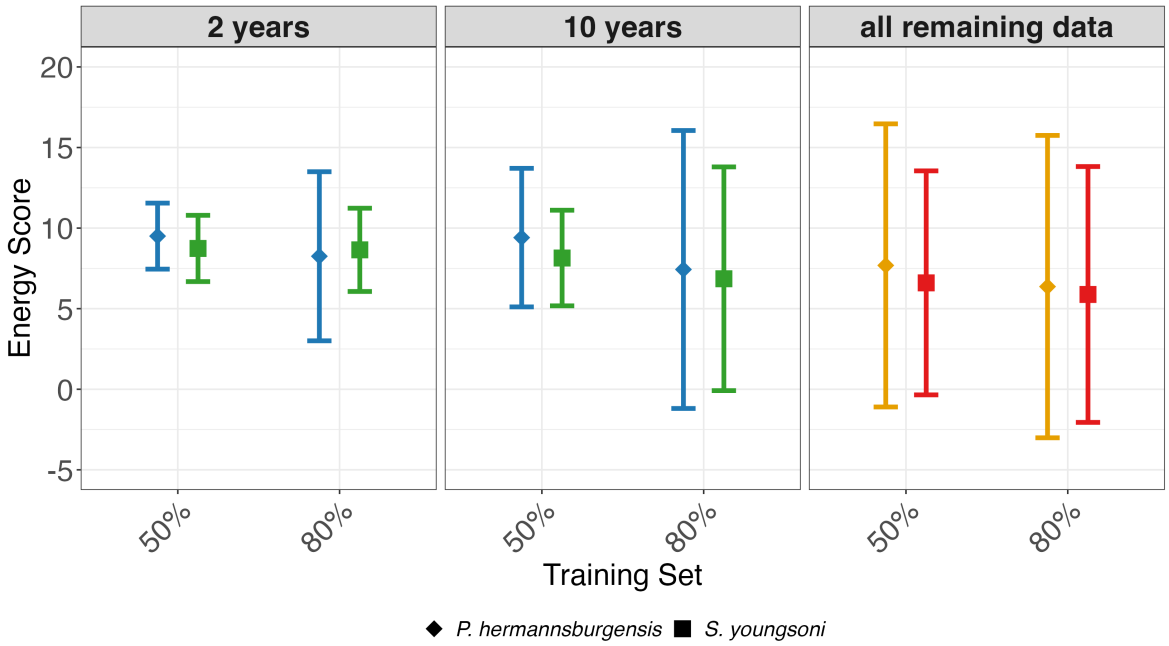

FIGURE 4.5. Energy Scores for the final model of *Pseudomys hermannsburgensis* (***GAM-GP***) and *Sminthopsis youngsoni* (***GAM-GP(only cover)***) fitted for Training set 1 (50%) and Training set 2 (80%) and evaluated for three forecast horizons (2 years, 10 years and all remaining data). Diamond and Square points represent mean scores across the forecasted seasons within each forecast horizon, with error bars indicating the 95% confidence interval. The colour differences within the 2 year and 10 year forecast horizons (blue and green) reflect variations in covariate specifications between *Pseudomys hermannsburgensis* and *Sminthopsis youngsoni*. The 'all remaining data' horizon is represented in distinct colours (yellow & red) to indicate differences in the number of time points, making direct comparisons with the other horizons inappropriate.

Looking at the Energy Scores across all time points in the 'all remaining data' forecast horizon for both *P. hermannsburgensis* (***GAM-GP***) and *S. youngsoni* (***GAM-GP(only cover)***) models fitted for Training set 1 and Training set 2 shows that both the models effectively identify extreme conditions and return to original process (Fig. 4.6). High Energy Scores correspond to identifying boom periods, while a sudden drop in scores at subsequent time points indicates a bust. Consistent with the results above, the performance of each model clearly reflects the contrasting species dynamics, as shown by the variation in Energy Scores across all time points within the forecast horizon.

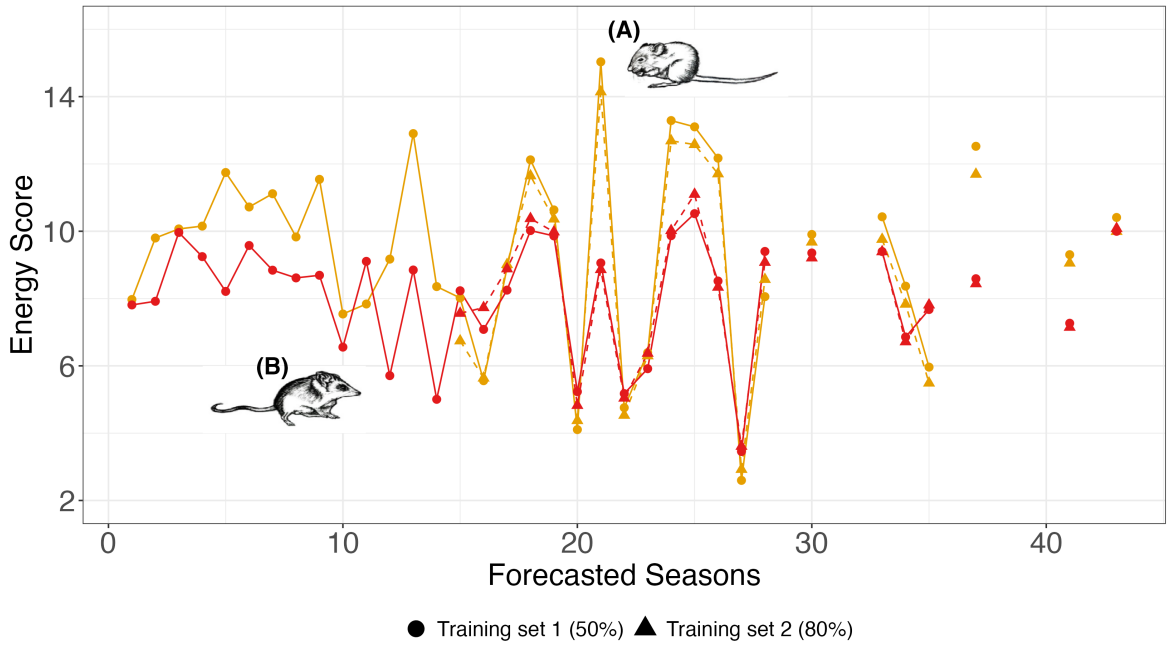

FIGURE 4.6. Energy Scores for the final model of (A) *Pseudomys hermannsburgensis* (***GAM-GP***) (yellow) and (B) *Sminthopsis youngsoni* (***GAM-GP(only cover)***) (red) fitted for Training set 1 (solid line) and Training set 2 (dashed line) and evaluated for each forecasted season within the 'all remaining data' horizon. Gaps were observed during cooler months of 2015, 2016, 2020 & during hotter months of 2017, 2019, 2020, 2021, 2022 due to heavy rainfall causing road closures and COVID-19 impeding travel to the study location.

4.4 Discussion

Our study focused on determining the optimal training length and forecast horizon to improve the use of ecological forecasts for species abundance in dynamic environments. By evaluating how far into the future useful predictions can be made and how ecological features like population variability influence forecast accuracy, we aimed to advance the understanding of the ecological forecast horizon. To achieve this, we focused on two ecologically distinct species, a rodent - *P. hermannsburgensis* and a small dasyurid - *S. youngsoni*, which provided contrasting population dynamics to evaluate model performance under varying conditions.

4.4.1 Model performance and selection

We first identified the best-fitting model structure to compare performance across different training lengths and forecast horizons. Our model incorporated key predictors, including cover, seeding, and flowering of the dominant vegetation species in the study region, spinifex (*Triodia basedowii*), as well as rainfall from the previous year. Rainfall is well-established as a critical driver of desert rodent community dynamics, primarily through its role in stimulating the growth of essential food resources (Ernest et al., 2000; Thibault et al., 2010; Bennison et al., 2018). In a study conducted southwest of the Northern Territory, Australia, Bennison et al. (2018) observed that rodents showed a positive response to previous six months cumulative rainfall, while dasyurids exhibited a similar response within a three-month period. Similarly, Greenville et al. (2016), working in the Simpson Desert (the same study area as the DERG dataset), demonstrated that both *P. hermannsburgensis* and *S. youngsoni* were strongly associated with rainfall from the previous year, highlighting the importance of temporal rainfall patterns in shaping species-specific population dynamics. Rainfall-driven pulses not only enhance food availability, such as seeds and flowers for rodents (Bennison et al., 2018), but also increase vegetation cover, particularly spinifex, which provides essential shelter (Greenville et al., 2016). However, prolonged droughts, fires, and predation by larger dasyurids, such as *D. blythi*, amplify population declines through direct impacts on prey species, including small rodents and dasyurids (Spencer et al., 2014). These findings underscore the complex interplay between abiotic and biotic factors in influencing desert mammal communities. Furthermore, our results confirm the nonlinear relationships between these ecological drivers and species abundance (Lima et al., 2008; Thibault and Brown, 2008; Karunarathna et al., 2024).

We evaluated three MVGAMs with different trend components; ***GAM-AR***, ***GAM-VAR***, and ***GAM-GP***, using 50% of the data (Training set 1) for *P. hermannsburgensis*. While all models generated hindcasts that captured general temporal and spatial variability of *P. hermannsburgensis* within the training set, our analysis highlighted trade-offs between generalization to new data and computational efficiency. Although the ***GAM-VAR*** model exhibited the highest predictive performance, as indicated by its ELPD, the substantial runtime (4.00

hours) made it impractical for applications requiring frequent model updates or scalability to larger datasets. Conversely, the ***GAM-GP*** model, with only a modest reduction in predictive performance (ELPD difference of -3.4), emerged as the most computationally efficient option, requiring just 3.30 minutes to fit. This efficiency is particularly relevant for ecological forecasting, where quick model updates are often required to incorporate new data or adapt to changing environmental conditions (Dietze et al., 2018). The ***GAM-AR*** model, while faster than ***GAM-VAR*** (37 minutes), demonstrated the lowest predictive accuracy and was thus less suitable for the study objectives. These findings align with Clark et al. (2025), who reported that the ***VAR*** model outperformed in most validation procedures but fell slightly behind the ***AR*** model when the training set excluded a major rodent abundance shift caused by drought in multi-species analyses. These findings align with prior studies that emphasize the value of computationally efficient models in ecological research, particularly in time-sensitive conservation efforts or large-scale environmental datasets (Wood, 2017; Simpson et al., 2017; Pedersen et al., 2019). The selection of the ***GAM-GP*** model reflected a pragmatic approach that balances accuracy and efficiency (Karunarathna et al., 2024). This is consistent with previous research highlighting the flexibility and computational advantages of Gaussian Process-based models for ecological forecasting (Vanhatalo et al., 2012; Sigourney et al., 2012; Golding and Purse, 2016).

The significance of the covariates in the ***GAM-GP*** model for *P. hermannsburgensis* varied depending on the training set used. We observed that no covariates were significant when the model was trained on Training Set 1 while only the effect of the predatory mulgara *D. blythi* was significant when Training Set 2 was used. This significance of the predictors that affect the abundance of *P. hermannsburgensis* contrasted with what we found in Chapter 2, where both previous year rainfall and spinifex seeding were significant. In addition, the ***GAM-GP(only cover)*** model fitted for *S. youngsoni* showed no significant covariates that aligned with the updated model assumptions of the MARSS model in Chapter 2 but slightly contrasting from Greenville et al. (2016). These findings underscore that the significance of covariates influencing small mammal abundance is highly dependent on the model, regardless of the underlying species' dynamics. Moreover, the range and variability of the training

data strongly influence covariate significance, particularly in species with high boom-bust dynamics compared to those with more stable dynamics.

4.4.2 Forecast performance across training lengths and forecast horizons

Understanding the optimal amount of data required for model training and the forecast horizon that can provide reliable predictions is a critical challenge for ecologists, especially in studies with sparse data.

The comparison of training lengths and forecast horizons offers valuable insights into the models' flexibility and reliability across different data availability scenarios. Despite the inherent differences in the population dynamics of *P. hermannsburgensis* and *S. youngsoni*, the ***GAM-GP*** model consistently demonstrated both computational efficiency and robust performance. Notably, the model fit times for 50% and 80% training sets, along with the forecast times for 2 year, 10 year, and all remaining data horizons, were all under 5 minutes. This efficiency of the ***GAM-GP*** model enables the rapid execution of multiple comparisons, making it possible to explore a wide range of management scenarios, such as those aimed at increasing plant cover (e.g., through wildfire reduction), or assessing the potential impacts under future climate scenarios. By allowing the model to process different time series lengths and scenarios, more extensive analyses can be conducted within a very limited time, providing a broader understanding of how environmental and management changes may influence species dynamics. This capacity for running numerous simulations supports the development of timely, informed, data-driven decision-making processes, which will be further explored in Chapter 5 with actual future forecasts.

The results showed that the model performed well across all training splits and forecast horizons tested in the study. For *P. hermannsburgensis*, the higher and more variable Energy Scores (ES) observed across the 2 year and 10 year forecast horizons reflect the species' boom-bust dynamics, which are driven by environmental variability and seasonal fluctuations. These findings are consistent with previous studies on arid-zone rodents, which emphasize their sensitivity to rainfall and other environmental drivers (Dickman et al., 2014;

Greenville et al., 2016). The decreasing mean ES with larger training datasets suggests that more data enhances the model's ability to capture complex population dynamics. However, the overlapping confidence intervals indicate no significant difference in forecast performance across different training splits and forecast horizons, pointing to the stability of the model's predictions. However, Simonis et al. (2021) highlights that model performance can vary between short-term and long-term forecast horizons depending on the specific model applied, particularly when analyzing long-term rodent population time series. Therefore, while the model generates robust forecasts regardless of training length or forecast horizon, it is crucial to carefully consider where the training and testing sets are split, ensuring that the training set adequately captures complex dynamics and significant shifts in population trends.

In contrast to *P. hermannsburgensis*, the relatively stable ES observed for *S. youngsoni* across the 2 year and 10 year horizons is consistent with the species' more stable population dynamics, which are primarily influenced by Avg. Spinifex Cover and the previous year's rainfall. The reduced variability in ES across forecast horizons highlights the model's robustness in predicting consistent ecological trends. These findings support earlier work by Letnic et al. (2013), which documented the resilience of *S. youngsoni* populations to environmental fluctuations, such as those driven by droughts and changes in vegetation cover.

The ***GAM-GP*** model's ability to generalize across different forecast horizons is particularly noteworthy. For both species, minimal differences in ES were observed between the 2 year and 10 year horizons, despite the challenges posed by extended forecasting periods. Interestingly, when forecasting for the full testing set (the 'all remaining data' horizon), both species exhibited broader confidence intervals, regardless of the training set length. This outcome reflects the inherent uncertainty associated with longer-term predictions, especially for species with dynamic population trends, such as *P. hermannsburgensis*. This uncertainty quantification aligns with findings from Petchey et al. (2015) and Adler et al. (2020), who emphasized the importance of accounting for forecast uncertainty, particularly over extended periods.

These findings carry practical implications for ecological forecasting, suggesting that models trained on smaller datasets can still provide reliable forecasts, which can reduce data collection costs and facilitate more frequent model updates. Additionally, the ***GAM-GP*** model's ability to handle extended forecast horizons underscores its potential for long-term ecological predictions, even in the face of species-specific challenges, such as dynamic population fluctuations or more stable trends. This capacity makes ***GAM-GP*** a valuable tool for researchers working in areas with limited data or where forecasting future ecological states is critical for management and conservation efforts.

4.4.3 Ecological implications

The substitution of "Spinifex" variables with "Overall Vegetation" variables yielded consistent results, confirming that dominant vegetation species can serve as reliable proxies for overall vegetation dynamics. This finding simplifies the selection of environmental covariates for ecological models, reducing the need for comprehensive vegetation surveys and enabling more streamlined data collection processes. The equivalence of results between dominant and overall vegetation variables aligns with ecological theories emphasizing the disproportionate influence of dominant species on ecosystem structure and function (Grime, 1998; Smith et al., 2009). Furthermore, this supports the notion that dominant species often drive community-level responses to environmental changes (HilleRisLambers et al., 2012).

Additionally, the application of alternative scoring metrics, such as Weighted Variogram Scores (WVS), supported the trends observed with Energy Scores (ES), though with higher absolute values. This finding underscores the importance of employing multiple evaluation metrics to comprehensively assess model performance, particularly in spatially structured data. WVS provides valuable insights into the dependency structure between spatial grids, complementing the broader evaluation provided by ES (Gneiting and Raftery, 2007; Simonis et al., 2021; Alexander et al., 2024). Given the robustness of the ***GAM-GP*** model in producing reliable forecasts for species with contrasting dynamics, in Chapter 5, we will generate real forecasts into the future with the support of projected covariate data.

References

Adler, P. B., White, E. P., and Cortez, M. H. (2020). Matching the forecast horizon with the relevant spatial and temporal processes and data sources. *Ecography*, 43(11):1729–1739.

Alexander, C., Coulon, M., Han, Y., and Meng, X. (2024). Evaluating the discrimination ability of proper multi-variate scoring rules. *Annals of Operations Research*, 334(1):857–883.

Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., Leos-Barajas, V., Mills Flemming, J., Nielsen, A., Petris, G., et al. (2021). A guide to state–space modeling of ecological time series. *Ecological Monographs*, 91(4):e01470.

Bennison, K., Godfree, R., and Dickman, C. R. (2018). Synchronous boom–bust cycles in central Australian rodents and marsupials in response to rainfall and fire. *Journal of Mammalogy*, 99(5):1137–1148.

Bjerregård, M. B., Møller, J. K., and Madsen, H. (2021). An introduction to multivariate probabilistic forecast evaluation. *Energy and AI*, 4:100058.

Buckland, S., Newman, K., Thomas, L., and Koesters, N. (2004). State-space models for the dynamics of wild animal populations. *Ecological Modelling*, 171(1-2):157–175.

Bureau of Meteorology (2024). Climate data online. http://www.bom.gov.au/climate/data/. Accessed on 26/12/2024.

Clark, J. S., Carpenter, S. R., Barber, M., Collins, S., Dobson, A., Foley, J. A., Lodge, D. M., Pascual, M., Pielke Jr, R., Pizer, W., et al. (2001). Ecological forecasts: an emerging imperative. *Science*, 293(5530):657–660.

Clark, N. J., Ernest, S. M., Senyondo, H., Simonis, J., White, E. P., Yenni, G. M., and Karunarathna, K. (2025). Beyond single-species models: leveraging multispecies forecasts to navigate the dynamics of ecological predictability. *PeerJ 13:e18929*.

Clark, N. J. and Wells, K. (2023). Dynamic generalised additive models (DGAMs) for forecasting discrete ecological time series. *Methods in Ecology and Evolution*, 14(3):771–784.

Coreau, A., Pinay, G., Thompson, J. D., Cheptou, P.-O., and Mermet, L. (2009). The rise of research on futures in ecology: rebalancing scenarios and predictions. *Ecology Letters*, 12(12):1277–1286.

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., and Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, 19(3):553–570.

Dennis, B. and Ponciano, J. M. (2014). Density-dependent state-space model for population-abundance data with unequal time intervals. *Ecology*, 95(8):2069–2076.

Dickman, C., Wardle, G., Foulkes, J., and Preu, N. d. (2014). Desert complex environments. In Burns, E., Lowe, A., Lindenmayer, D., and Thurgate, N., editors, *Biodiversity and environmental change: monitoring, challenges and direction*, chapter 10, pages 379–438. CSIRO Publishing.

Dickman, C. R., Greenville, A. C., Beh, C.-L., Tamayo, B., and Wardle, G. M. (2010). Social organization and movements of desert rodents during population "booms" and "busts" in central Australia. *Journal of Mammalogy*, 91(4):798–810.

Dietze, M. C. (2017). *Ecological Forecasting*. Princeton University Press.

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., et al. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences*, 115(7):1424–1432.

Dorazio, R. M. (2016). Bayesian data analysis in population ecology: motivations, methods, and benefits. *Population Ecology*, 58(1):31–44.

Dorazio, R. M. and Karanth, K. U. (2017). A hierarchical model for estimating the spatial distribution and abundance of animals detected by continuous-time recorders. *PloS One*, 12(5):e0176966.

Douma, J. C. and Weedon, J. T. (2019). Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. *Methods in Ecology and Evolution*, 10(9):1412–1430.

Ernest, S. M., Brown, J. H., and Parmenter, R. R. (2000). Rodents, plants, and precipitation: spatial and temporal dynamics of consumers and resources. *Oikos*, 88(3):470–482.

Frank, A. S., Dickman, C. R., Wardle, G. M., and Greenville, A. C. (2013). Interactions of grazing history, cattle removal and time since rain drive divergent short-term responses by desert biota. *PLoS One*, 8(7):e68466.

Frank, A. S., Wardle, G. M., Greenville, A. C., and Dickman, C. R. (2016). Cattle removal in arid Australia benefits kangaroos in high quality habitat but does not affect camels. *The Rangeland Journal*, 38(1):73–84.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378.

Gneiting, T., Stanberry, L. I., Grimit, E. P., Held, L., and Johnson, N. A. (2008). Assessing probabilistic forecasts of multivariate quantities, with an application to ensemble predictions of surface winds. *TEST*, 17:211–235.

Golding, N. and Purse, B. V. (2016). Fast and flexible Bayesian species distribution modelling using Gaussian processes. *Methods in Ecology and Evolution*, 7(5):598–608.

Greenville, A. C., Dickman, C. R., Wardle, G. M., and Letnic, M. (2009). The fire history of an arid grassland: the influence of antecedent rainfall and ENSO. *International Journal of Wildland Fire*, 18(6):631–639.

Greenville, A. C., Wardle, G. M., and Dickman, C. R. (2013). Extreme rainfall events predict irruptions of rat plagues in central Australia. *Austral Ecology*, 38(7):754–764.

Greenville, A. C., Wardle, G. M., Nguyen, V., and Dickman, C. R. (2016). Population dynamics of desert mammals: similarities and contrasts within a multispecies assemblage. *Ecosphere*, 7(5):e01343.

Grime, J. (1998). Benefits of plant diversity to ecosystems: immediate, filter and founder effects. *Journal of Ecology*, 86(6):902–910.

Hersbach, H. (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15(5):559–570.

HilleRisLambers, J., Adler, P. B., Harpole, W. S., Levine, J. M., and Mayfield, M. M. (2012). Rethinking community assembly through the lens of coexistence theory. *Annual Review of Ecology, Evolution, and Systematics*, 43(1):227–248.

Hobbs, N. T., Geremia, C., Treanor, J., Wallen, R., White, P., Hooten, M. B., and Rhyan, J. C. (2015). State-space modeling to support management of brucellosis in the Yellowstone bison population. *Ecological Monographs*, 85(4):525–556.

Hooten, M. B. and Hobbs, N. T. (2015). A guide to bayesian model selection for ecologists. *Ecological Monographs*, 85(1):3–28.

Hostetler, J. A. and Chandler, R. B. (2015). Improved state-space models for inference about spatial and temporal variation in abundance from count data. *Ecology*, 96(6):1713–1723.

Hyndman, R. J. and Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts Melbourne.

Karunarathna, K., Wells, K., and Clark, N. J. (2024). Modelling nonlinear responses of a desert rodent species to environmental change with hierarchical dynamic generalized additive models. *Ecological Modelling*, 490:110648.

Kéry, M. and Schaub, M. (2011). *Bayesian population analysis using WinBUGS: a hierarchical perspective*. Academic Press.

Keune, J., Ohlwein, C., and Hense, A. (2014). Multivariate probabilistic analysis and predictability of medium-range ensemble weather forecasts. *Monthly Weather Review*, 142(11):4074–4090.

Letnic, M. (2004). Cattle grazing in a hummock grassland regenerating after fire: the short-term effects of cattle exclusion on vegetation in south-western Queensland. *The Rangeland Journal*, 26(1):34–48.

Letnic, M., Tischler, M., and Gordon, C. (2013). Desert small mammal responses to wildfire and predation in the aftermath of a La Niña driven resource pulse. *Austral Ecology*, 38(7):841–849.

Lima, M., Ernest, S. M., Brown, J. H., Belgrano, A., and Stenseth, N. C. (2008). Chihuahuan desert kangaroo rats: nonlinear effects of population dynamics, competition, and rainfall. *Ecology*, 89(9):2594–2603.

Lofton, M. E., Brentrup, J. A., Beck, W. S., Zwart, J. A., Bhattacharya, R., Brighenti, L. S., Burnet, S. H., McCullough, I. M., Steele, B. G., Carey, C. C., et al. (2022). Using near-term forecasts and uncertainty partitioning to inform prediction of oligotrophic lake cyanobacterial density. *Ecological Applications*, page e2590.

Matheson, J. E. and Winkler, R. L. (1976). Scoring rules for continuous probability distributions. *Management science*, 22(10):1087–1096.

Pedersen, E. J., Miller, D. L., Simpson, G. L., and Ross, N. (2019). Hierarchical generalized additive models in ecology: an introduction with mgcv. *PeerJ*, 7:e6876.

Petchey, O. L., Pontarp, M., Massie, T. M., Kéfi, S., Ozgul, A., Weilenmann, M., Palamara, G. M., Altermatt, F., Matthews, B., Levine, J. M., et al. (2015). The ecological forecast horizon, and examples of its uses and determinants. *Ecology Letters*, 18(7):597–611.

Pinson, P. and Girard, R. (2012). Evaluating the quality of scenarios of short-term wind power generation. *Applied Energy*, 96:12–20.

Pinson, P. and Tastu, J. (2013). Discrimination ability of the Energy score. *DTU Compute Technical Report-2013*.

Purdie, R. (1984). Land systems of the Simpson desert region. *Natural Resources Series*, 2:1–71.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Scheuerer, M. and Hamill, T. M. (2015). Variogram-based proper scoring rules for probabilistic forecasts of multivariate quantities. *Monthly Weather Review*, 143(4):1321–1334.

Shephard, M. (1992). *The Simpson Desert: natural history and human endeavour*. Royal Geographical Society of Australasia, Adelaide, Australia.

Sigourney, D. B., Munch, S. B., and Letcher, B. H. (2012). Combining a bayesian nonparametric method with a hierarchical framework to estimate individual and temporal variation in growth. *Ecological Modelling*, 247:125–134.

Simonis, J. L., White, E. P., and Ernest, S. M. (2021). Evaluating probabilistic ecological forecasts. *Ecology*, 102(8):e03431.

Simpson, D., Rue, H., Riebler, A., Martins, T. G., and Sørbye, S. H. (2017). Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, 32(1):1–28.

Smith, M. D., Knapp, A. K., and Collins, S. L. (2009). A framework for assessing ecosystem dynamics in response to chronic resource alterations induced by global change. *Ecology*, 90(12):3279–3289.

Spencer, E. E., Crowther, M. S., and Dickman, C. R. (2014). Diet and prey selectivity of three species of sympatric mammalian predators in central Australia. *Journal of Mammalogy*, 95(6):1278–1288.

Stan Development Team (2024). *Stan Modeling Language Users Guide and Reference Manual*. Version 2.36.

Thibault, K. M. and Brown, J. H. (2008). Impact of an extreme climatic event on community assembly. *Proceedings of the National Academy of Sciences*, 105(9):3410–3415.

Thibault, K. M., Ernest, S. M., White, E. P., Brown, J. H., and Goheen, J. R. (2010). Long-term insights into the influence of precipitation on community dynamics in desert rodents. *Journal of Mammalogy*, 91(4):787–797.

Tulloch, A. I., Healy, A., Silcock, J., Wardle, G. M., Dickman, C. R., Frank, A. S., Aubault, H., Barton, K., and Greenville, A. C. (2023). Long-term livestock exclusion increases plant richness and reproductive capacity in arid woodlands. *Ecological Applications*, 33(8):e2909.

Van Buuren, S. and Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45:1–67.

Vanhatalo, J., Veneranta, L., and Hudd, R. (2012). Species distribution modeling with Gaussian processes: A case study with the youngest stages of sea spawning whitefish (Coregonus lavaretus L. sl) larvae. *Ecological Modelling*, 228:49–58.

Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27:1413–1432.

Verhoeven, E. M., Murray, B. R., Dickman, C. R., Wardle, G. M., and Greenville, A. C. (2020). Fire and rain are one: extreme rainfall events predict wildfire extent in an arid grassland. *International Journal of Wildland Fire*, 29(8):702–711.

Ward, E. J., Holmes, E. E., Thorson, J. T., and Collen, B. (2014). Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. *Oikos*, 123(6):652–661.

Wardle, G. and Dickman, C. (2015). Desert Ecology Plot Network: Mammal, Reptile and Vegetation Data Associated with Weather, Simpson Desert, Western Queensland, Australia, 1990–2011. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern.111.57/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018a). Desert Ecology Plot Network: Mammal Abundance Plot-data, Simpson Desert, Western Queensland, 1990+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.196.8/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018b). Desert Ecology Plot Network: Vegetation Plot-data, Simpson Desert, Western Queensland, 1993+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.195.11/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018c). Desert Ecology Plot Network: Weather Data (daily and monthly), Simpson Desert, Western Queensland, 1995+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.194.7/html. Accessed on 16/01/2025.

Wardle, G. M., Greenville, A. C., Frank, A. S., Tischler, M., Emery, N. J., and Dickman, C. R. (2015). Ecosystem risk assessment of Georgina gidgee woodlands in central Australia. *Austral Ecology*, 40(4):444–459.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

CHAPTER 5

Near-term forecasts of population abundances using projected rainfall data under climate change scenarios

---

5.1 Introduction

Ecologists face the critical challenge of providing reliable scientific forecasts of future ecological outcomes. This is more challenging due to the stochasticity of ecosystems, the complex interactions of biotic and abiotic factors, and the uncertainty associated with environmental change. Addressing this global challenge requires anticipating the impacts of environmental change on ecosystem function and biodiversity (Carpenter, 2002; Lewis et al., 2023). In an era of rapid climatic and environmental shifts, the need for reliable ecological forecasts has never been more critical. Accurate forecasts are fundamental to informing conservation management, mitigating biodiversity loss, and ensuring the resilience of ecosystems (Clark et al., 2001). As we enter a phase where management is moving toward climate adaptation, given the challenges and delays in mitigating climate change, the need for near-term ecological forecasts is paramount (Stein et al., 2013). However, achieving this level of predictive capability presents both theoretical and methodological challenges, particularly when extending forecasts beyond the scope of observed data.

Traditionally, ecological forecasting has relied on statistical models evaluated on historical data to predict future conditions. These models often employ a data-partitioning strategy, where the dataset is divided into training and testing subsets. The training set is used to fit the model, while the testing set evaluates its predictive performance (Dietze, 2017). Taking a different approach of predicting change in species distribution by comparing two distinct time periods, Piirainen et al. (2023) highlight the importance of capturing sufficient variation

in the training data which otherwise would challenge the reliability of predicted spatial patterns. While this approach provides important insights into model accuracy and robustness, it remains constrained by the conditions of the observed past. Dietze (2017) highlighted that a truly successful ecological forecast must not only be able to explain past observations but also provide forecasts that extend beyond these observations, offering stronger evidence for hypotheses and greater confidence in future forecasts by reducing the risk of overfitting. This shift from within-sample testing to out-of-sample forecasting presents a critical opportunity to improve the relevance and application of ecological models.

One of the key limitations in ecological forecasting is the challenge of validating forecasts (Beaumont et al., 2007). Without future observations, directly assessing forecast accuracy using formal measures is impossible. This absence of ground truth data raises significant concerns regarding the reliability of models and their forecasts. To address this, uncertainty quantification becomes a cornerstone of ecological forecasting (Thuiller et al., 2004; Gauthier et al., 2016). Accurately understanding and communicating the uncertainty associated with forecasts is critical, especially when those forecasts inform conservation strategies and management decisions that will impact the state of ecosystems for years to come (Valle et al., 2009; Nichols et al., 2011; Mouquet et al., 2015). Methods such as probabilistic forecasting and Bayesian approaches can help to quantify this uncertainty, providing a range of potential outcomes rather than a single deterministic forecast value. For example, Finley et al. (2012) shows how Bayesian methods can incorporate prior knowledge and account for uncertainty in model parameters, thereby offering more robust and transparent forecasts. Moreover, Dietze et al. (2018) advocate for iterative forecasting frameworks that refine forecasts over time as new data become available, thereby reducing uncertainty and improving forecast reliability. On the other hand, matrix population models are commonly used to project population abundance by considering various life stages (e.g., age, size, location) of the species. However, because these models require the collection of additional stage-specific data, which can be challenging, uncertainty tends to propagate through the projections (Fujiwara and Caswell, 2002; Caswell, 2000). Therefore, approaches that quantify uncertainty are essential not only for validating forecasts but also for guiding decision-making under conditions of uncertainty.

Forecasting species population dynamics presents unique challenges due to the influence of external environmental factors such as rainfall, temperature, land-use changes and broader climate cycles such as El Niño-Southern Oscillation (ENSO) (Letnic and Dickman, 2006; Greenville et al., 2009). Rainfall is widely recognized as a critical driver of the dynamics of the desert rodent community, mainly through its role in stimulating the growth of essential food resources (Ernest et al., 2000; Thibault et al., 2010; Bennison et al., 2018) and researchers have found that the temporal effects of rainfall vary between rodents and dasyurids in arid environments Bennison et al. (2018); Greenville et al. (2016). Apart from enhancing food availability (Bennison et al., 2018), rainfall also increases vegetation cover, which provides essential shelter for teh small mammal species (Greenville et al., 2016). Therefore, understanding how these variables interact under different future scenarios is essential to accurately forecast species abundance, identifying potential population declines or risks of extinction, and informing the development of effective conservation strategies. This highlights the critical need to understand the current state of climate change in order to facilitate forecasting population abundance under future climate change scenarios. Beaumont et al. (2007) used projected temperature and rainfall data from climate models available in the Programme for Climate Model Diagnosis and Intercomparison (PCMDI) data archive to project bioclimatic envelopes for nine Australian butterfly species. Their findings revealed greater consistency within individual climate models compared to projections across different models, highlighting the importance of model selection. Similarly, studies using bioclimatic models have shown that forecasts derived from different climate models can yield significant variability in projected species distributions (Thuiller, 2004; Beaumont et al., 2005; Saxon et al., 2005). On the other hand, Tsiftsis et al. (2024) claims that species distribution projections should include all potential environmental factors, not just climatic variables, as neglecting them would question the reliability of the projections. However, there is a contradiction between the more common ecological forecasts that are based on long-term climate change projections, and the shorter time frames required for environmental decision-making (Dietze et al., 2018). Decision-makers typically need near-term forecasts (ranging from daily to decadal) modeled on current data and projections that evaluate alternative management strategies. This underscores the need for downscaled, high-frequency climate projections (e.g. daily,

monthly, seasonal, yearly) and more timely, data-driven models to bridge the gap between research and practical, decision-relevant forecasting.

In Chapter 4, we used historical data to fit multivariate generalized additive models (MVGAMs) and predict population dynamics based on observed climate and vegetation conditions. While this approach yielded valuable insights into the relationships between population dynamics and past environmental variables under different model specifications, it was constrained by its inability to forecast future population trends beyond the scope of observed data. Therefore, this study advances the work in Chapter 4 by incorporating projected climate data into the forecasting process.

Using long-term population abundance data from the DERG dataset (1990 – 2022), this study aims to forecast the abundance of two species with contrasting population dynamics; the rodent *Pseudomys hermannsburgensis* and the small insectivorous dasyurid *Sminthopsis youngsoni*, over two near-term horizons; 4 years, chosen to be relevant to on-ground management actions and 12 years to reflect a period of years within another ENSO cycle is likely to have happened. We generate forecasts by integrating rainfall projections from different climate change scenarios, known as Shared Socioeconomic Pathways (SSPs) (Meinshausen et al., 2020), which represent a range of future greenhouse gas emission trajectories to act as a future resource state of environment. We predict that the forecasts for the rodent, sandy inland mouse (*P. hermannsburgensis*), a species with highly varying dynamics, will be different across the two near-term horizons, with the reliability of forecasted abundance decreasing over the longer horizon (Ward et al., 2014; Petchey et al., 2015; Lewis et al., 2022). However, we expect to see considerable fluctuations in population abundance under more extreme climate change scenarios. These fluctuations will reflect the species' sensitivity to environmental changes, potentially leading to vastly different management and conservation decisions. However, for the lesser hairy-footed dunnart, small insectivorous dasyurid (*S. youngsoni*), the species with more stable dynamics, we anticipate relatively consistent population trends with minimal variations in response to climate change scenarios. This contrast in dynamics will highlight the varying resilience of species to climate-induced changes and emphasize the importance of timely decision-making. In addition, our approach

demonstrates that even in locations where abundance data are sparse or largely missing, the MVGAM method allows for reliable forecasting. Overall, this novel approach represents a significant advancement in ecological forecasting, providing a framework for anticipating future population trends under diverse environmental scenarios and offering actionable insights for conservation and management.

5.2 Material and methods

5.2.1 Study location and study species

The Desert Ecology Research Group (DERG) at the University of Sydney has been conducting long-term ecological monitoring in the Simpson Desert, Australia since 1990 (Wardle and Dickman (2015), Wardle and Dickman (2018a), Wardle and Dickman (2018b), Wardle and Dickman (2018c), unpublished data from 2018). The Simpson Desert covers 170,000 km^2 of mainly dune fields (73%), with the remainder comprising clay pans, rocky outcrops, and gibber plains (Shephard, 1992). The study region of 8000 km^2, spans the border between Queensland and the Northern Territory (Fig. 2.1) and the vegetation of the dune fields is characterized by spinifex hummock grasslands (*Triodia basedowii*) with scattered shrubs (*Acacia*, *Eremophilla*, *Grevillea*), ephmeral forbs and a variety of other perennial and annual grasses; interspersed are wide swales dominated by Georgina Gidgee woodlands (*Acacia georginae*), or mallee eucalypts (Wardle et al., 2015). Climatic conditions are highly variable, alternating between extended dry periods ("busts") and productive "boom" phases following summer rainfall (Dickman et al., 2010; Greenville et al., 2013). Temperatures range from over 40°C in summer to below 5°C in winter (Purdie, 1984), while annual rainfall varies spatially, averaging 258.5 mm in Boulia (136 years), 203 mm in Bedourie (79 years), and 157 mm in Birdsville (24 years) (Bureau of Meteorology, 2024) (see Chapter 2 for further details).

Building on the work presented in Chapter 4, our primary focus remains on the same two species, which exhibit contrasting boom-and-bust dynamics: the rodent *Pseudomys hermannsburgensis* (sandy inland mouse, 12 g) and the small insectivorous dasyurid *Sminthopsis youngsoni* (lesser hairy-footed dunnart, 10 g). Capture data were analyzed at the grid

level due to significant variations in abundances among grids within the same site. Live capture counts of *P. hermannsburgensis* and *S. youngsoni* from 1990 to 2022 were standardized per 100 trap nights across the thirty four grids (Fig. 4.1), averaged biannually (to cooler and hotter months separation), and assigned a continuous time variable ('time') ranging from 1 to 66, while missing records were denoted as NA (see Chapter 4 for more details).

5.2.2 Historical and projected rainfall data

We gathered high-resolution monthly rainfall projections for Australia through the dynamical downscaling of global climate models (GCMs) from the latest phase of the Coupled Model Intercomparison Project (CMIP6). These projections were produced by the Queensland Future Climate Science Program (Syktus et al., 2020; Chapman et al., 2023). Eleven CMIP6 GCMs were selected to generate downscaled climate change simulations, with three of these models run in multiple configurations, resulting in an ensemble of 15 regional climate models (RCMs). The Conformal Cubic Atmospheric Model (CCAM; (McGregor, 2005; McGregor and Dix, 2008)) was used to dynamically downscale the GCMs to a 10 km spatial resolution over Australia in both atmospheric-only and coupled atmosphere-ocean configurations (see Chapman et al. (2024) for further details). This process produced 60 downscaled simulations for historical (1960 – 2014) and future projections (2015 - 2100) based on three Shared Socioeconomic Pathways (SSPs). The SSPs represent possible future socioeconomic pathways and their associated greenhouse gas emissions (O'Neill et al., 2016; Meinshausen et al., 2020). The three climate change scenarios (SSPs) used in obtaining projections are (Meinshausen et al., 2020): 1) SSP126: the 'Sustainability' scenario which assumes a gradual reduction in global emissions, with net zero achieved only after 2050, leading to a stabilization of warming at approximately 1.8°C by 2100, 2) SSP245: the 'Middle-of-the-road' scenario projects emissions remaining steady at current levels before declining in the middle of the century. However, net zero is not achieved by 2100, resulting in a temperature rise of about 2.7°C by the end of the century and 3) SSP370: A medium-high emissions scenario corresponding to the 'Regional rivalry' socioeconomic family where both

emissions and temperatures continue to rise, leading to an estimated warming of approximately 3.6°C by 2100. The added value of high-resolution (10 km) projections for assessing the impacts of climate change is particularly evident in areas with complex terrain, coastal regions, and for climate extremes (Rummukainen, 2016; Giorgi, 2019). These factors make the dataset particularly well-suited for our study site, which features diverse and complex environmental characteristics.

The selection of regional climate model projections for our study followed a three-step process. First, we focused on a subset of models (Table 5.1) developed by local institutions (i.e., CSIRO and the Bureau of Meteorology) to ensure representation of local climatic behavior (see Trancoso et al. (2024) for descriptions of the contributing institutions and the origins of the models).

TABLE 5.1. Subset of the 15 regional climate models (RCMs) developed by CSIRO and the Bureau of Meteorology.

CMIP6 model name	**Model name**	**Ensemble member**
ACCESS-ESM1.5	Australian Community Climate and Earth System Simulator, version 1.5, CCAM atmospheric model version	r6i1p1f1
ACCESS-ESM1.5_oc	Australian Community Climate and Earth System Simulator, version 1.5, CCAM coupled ocean model version	r20i1p1f1 & r40i1p1f1
ACCESS_CM2_oc	Australian Community Climate and Earth System Simulator, version 2, CCAM coupled ocean version	r2i1p1f1

We then subset the models that were run in a coupled ocean-atmosphere configuration, as this setup accounts for critical interactions between the atmosphere and oceans, which are essential for capturing large-scale climate systems (e.g. ENSO, Indian Ocean Dipole; Ashok et al. (2003); Cai et al. (2009)) and our study site is strongly influenced by the ENSO cycle (Letnic and Dickman, 2006; Greenville et al., 2009). This resulted in the selection of two models: ACCESS-ESM1.5_oc and ACCESS-CM2_oc. As the final step, we selected the ACCESS-CM2_oc model for our study due to its focus on physical climate modeling and its newer, more advanced atmospheric, land, and sea-ice components. ACCESS-CM2_oc uses the UK Met Office Unified Model (UM) version 10.6, which includes an updated atmospheric configuration and parameterizations, compared to ACCESS-ESM1.5_oc, which is based on

the older UM version 7.3 (ARC Centre of Excellence for Climate Extremes, 2022). Additionally, ACCESS-CM2_oc exhibits a larger equilibrium climate sensitivity (ECS) of 4.7°C compared to ACCESS-ESM1.5_oc's ECS of 3.9°C, making it more sensitive to greenhouse gas forcing. Given its focus on physical climate modeling, updated components, and greater climate sensitivity as opposed to ACCESS-ESM1.5_oc, which focuses on simulating a fully interactive carbon cycle, ACCESS-CM2_oc was considered more suitable for capturing the complex atmospheric and oceanic interactions critical to our study site.

Although we had access to monthly rainfall data from automatic weather stations located at each of our seventeen sites from 1990 to 2022 (Chapter 4; Fig. 2.1), we note that these data were not on the same scale as the historical simulations from the ACCESS-CM2_oc RCM (see Fig. A3.1) for detailed information). Since, the rainfall projections from ACCESS-CM2_oc RCM were used in our model to forecast population abundances, we avoided using rainfall data on a different measuring scale to minimize confounding the comparisons of our forecasts. Therefore, to ensure consistency and comparability, for our study, we used both historical (1990 – 2022) and 12 year projected (2023 – 2035) rainfall data based on the three climate change scenarios from the ACCESS-CM2_oc RCM (Fig. A3.2).

The historical simulations and projections for rainfall were provided as 'Precipitation' with units kgm^2s^{-1} (Copernicus Climate Change Service, Climate Data Store, 2021), which represents the amount of liquid and frozen water (comprising rain and snow) per unit area and time. However, since our study site experiences precipitation only in the form of rain, we converted the units to millimeters (mm), the standard unit for measuring rainfall, by multiplying by the number of seconds in a month. Thereafter, the monthly data were aggregated based on the separation of cooler and hotter months, as outlined in Chapter 4.

5.2.3 Data processing

A two-step process was used to identify the closest rainfall location to our study sites. First, the longitudes and latitudes of the bi-annually aggregated historical and projected rainfall locations were subsetted to match the spatial extent of our study region. Then, using the

st_distance() function from the *sf* package in R (Pebesma, 2018), we computed the distances between the rainfall data locations and the study sites as great-circle distances (geodesic) (Bullock, 2007), accounting for the Earth's curvature (Fig. 5.1).

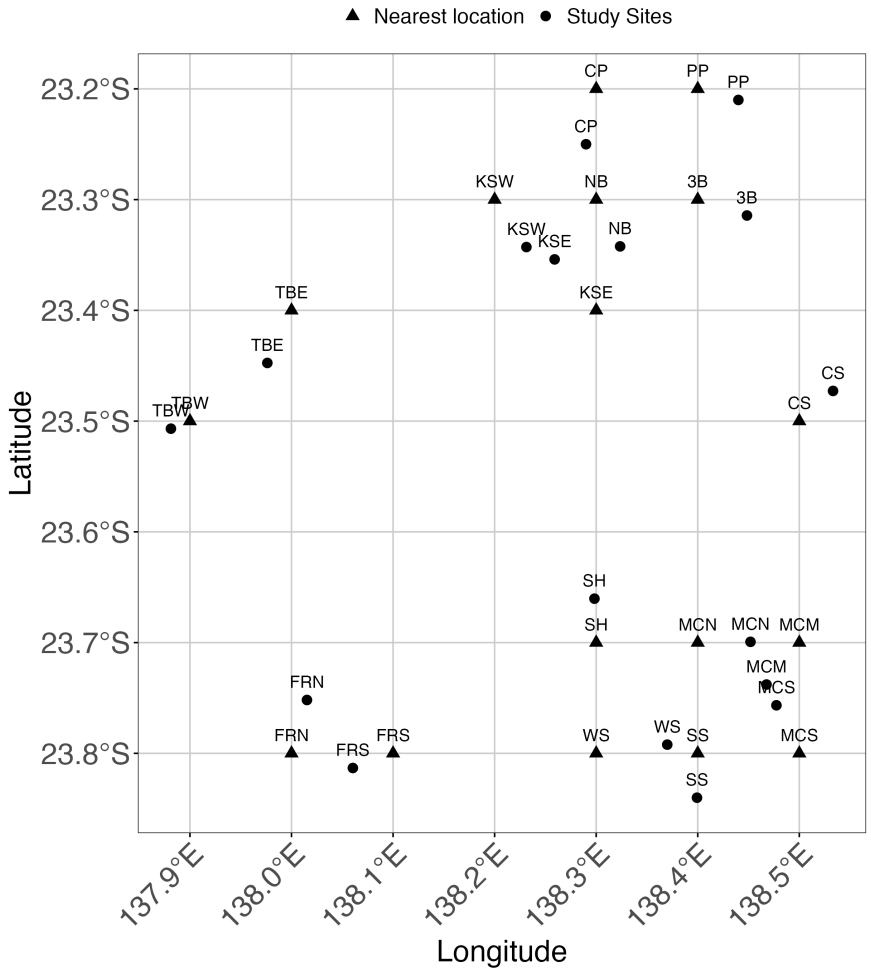

FIGURE 5.1. Nearest locations from a 10 km resolution dataset to each of the seventeen study sites, computed as great-circle distances (geodesic). The circles represent the exact locations of the study sites, while the triangles indicate the nearest location from the down-scaled dataset to each study site. Site abbreviations in alphabetical order: CS - Carlo, CP - Cravens Peak, FRN - Field River North, FRS - Field River South, KSE - Kunnamuka Swamp East, KSW - Kunnamuka Swamp West, MCN - Main Camp North, MCM - Main Camp Mid, MCS - Main Camp South, NB - Norries, 3B - No. 3 Bore, PP - Plum Pudding, SH - Shitty Site, SS - South Site, TBW - Tobermorey West, TBE - Tobermorey East, WS - Way Site.

5.2.4 Model expansion

Using the same generalized additive modeling framework outlined in Chapter 4; (Clark and Wells, 2023; Karunarathna et al., 2024; Clark et al., 2025), we applied the specifications and model representations of the 'final model' (***GAM-GP***) derived from the model selection procedure in that chapter. However, for this study, we lack projected future data for the vegetation variables (*Avg. Spinifex Cover*, *Spinifex Flowering Score*, and *Spinifex Seeding Score*) and the abundances of the predatory mulgara, *D. blythi*. Furthermore, in Chapter 2, we found that rainfall plays a major role in shaping the population abundance of the species (Greenville et al., 2016) even though the interaction between spinifex cover and rainfall was found insignificant in Chapter 4. Therefore, the analysis focused solely on the effect of rainfall in modeling and forecasting the population abundance of the two focal species, *P. hermannsburgensis* and *S. youngsoni*. Model specifications used in this study are as follows:

$$y_{it} \sim \text{Gamma}(\mu_{it}, \emptyset) \tag{5.1}$$

$$\log(\mu_{it}) = \alpha_{\text{grid}[i]} + S_1(Rainfall_{(t-2)}) + Z_{it}$$

where,

$i = 1, 2, \ldots, 34$

$\mu_{it} = E(y_{it})$: is the expected (average) captures per 100TN (TN = Trap Nights) for grid i at time t

$\alpha_{\text{grid}} \sim \text{Normal}(\mu_{\text{grid}}, \sigma^2_{\text{grid}})$: effect of grid i (random effect)

$\mu_{\text{grid}} \sim \text{Normal}(0, 1)$

$\sigma^2_{\text{grid}} \sim \text{Exponential}(0.5)$

$S_1(Rainfall_{(t-2)})$: smooth function represents effects of previous year Rainfall

$\beta_1 \sim \text{MultiNormal}(0, \Sigma^{-1}\lambda)$: coefficients in smooth function $S_1(Rainfall_{(t-2)})$

$\Sigma^{-1} \sim$ basis expansion function (supplied by mgcv function)

$\lambda \sim \text{Normal}(30, 25)$

$\frac{1}{\emptyset} \sim \text{Exponential}(5)$

$Z_{it} \sim \text{MultiNormal}(0, K)$: Gaussian Process latent trend component

$K(t, t') = g_{GP}^2 \exp\left(\frac{-|t - t'|^2}{\rho_{GP}^2}\right)$

$g_{GP} \sim \text{Normal}(0, 0.5)$

$\rho_{GP} \sim \text{InverseGamma}(1.49, 5.67)$

Smooth terms were used to model the nonlinear relationships between the captures and the previous year rainfall. To model the effect of previous year rainfall, we used a penalized cubic regression spline; specifically, a cyclic cubic regression spline with matching ends up to the second derivative (bs = 'cc'), using k = 17 basis functions. All other specifications remain consistent with the final model outlined in Chapter 4.

Modeling was performed using the *mvgam()* function from the *mvgam* package in R (Clark and Wells, 2023). The model was trained on data from 1990 to 2022 (all of the field-surveyed data up to date) and generated forecasts for the period 2023 to 2035, providing insights into how the species abundance might fluctuate over the next 12 years before they are observed through sampling.

5.3 Results

The final model (***GAM-GP***) from Chapter 4 was implemented as a multivariate generalized additive model (MVGAM) using data from all thirty-four grids on the rodent *P. hermannsburgensis* and the small insectivorous dasyurid *S. youngsoni*. We used three rainfall change scenarios to forecast the abundance of the species beyond 2022 and we compared results across three representative grids: Main Camp South, which had the highest number of observed captures (325 standardized counts); South Site, which had an average number (132 standardized counts); and No. 3 Bore (32 standardized counts), which had the lowest number

of observed counts. Our findings indicate that both hindcast and forecast trends in species abundance were influenced by the rainfall change scenario and the amount of data available at each location, regardless of species-specific dynamics. However, the model results showed that rainfall significantly influenced the forecasts for *P. hermannsburgensis*, a species with high boom-bust dynamics, while for the more stable *S. youngsoni*, rainfall alone was not a significant predictor, suggesting that site-specific covariates play a greater role in shaping its population dynamics (Table 5.2).

TABLE 5.2. The results of the final model (***GAM-GP***) for the rodent *Pseudomys hermannsburgensis* & the small insectivorous dasyurid *Sminthopsis youngsoni* using three rainfall change scenarios; SSP126 - Sustainability scenario, SSP245 - Middle-of-the-road scenario, SSP370 - Regional Rivalry scenario. Approximate significance of generalized additive smooth covariates (p-values < 0.05) are shown in bold.

Rainfall change scenarios	**p-value**
P. hermannsburgensis	
SSP126 (1.8°C increase by 2100)	**< 0.002**
SSP245 (2.7°C increase by 2100)	**< 0.002**
SSP370 (3.6°C increase by 2100)	**< 0.002**
S. youngsoni	
SSP126 (1.8°C increase by 2100)	0.70
SSP245 (2.7°C increase by 2100)	0.70
SSP370 (3.6°C increase by 2100)	0.70

5.3.1 Population abundance forecasts of a species with high boom-bust dynamics

Trends in the hindcasts of population abundance for *P. hermannsburgensis* aligned reasonably well with observations with consistent patterns across all the three exemplar grids with high (Main Camp South) medium (South Site) and low (No. 3 Bore) counts. Although the model appeared relatively flat compared to the observed captures in the training set, it successfully identified the years in which capture peaks occurred during boom periods at the study site. However, it was not effective in capturing the magnitude of those peak counts.

Notably, the 95% prediction intervals for the hindcasts successfully encompassed all observed counts across the three grids. However, the width of these intervals indicates a high level of uncertainty in the model's hindcasting phase, rather than strong predictive performance (Fig. 5.2).

The trends for the 4-year forecast horizon (2022 – 2026) observed under the three climate change scenarios vary across the three grids (Fig. 5.2). Under the Sustainability pathway that leads to 1.8°C increase by 2100 (SSP126), the model forecasts that the trend observed in the past will largely continue, with abundances remaining stable and reflecting historical patterns at Main Camp South and No. 3 Bore (even with limited number of observed captures), although at lower levels in the latter, while South Site is expected to sustain moderately high and steady abundances. This aligns with the best-case scenario in which climate change mitigation efforts lead to environmental preservation and careful resource management, helping to maintain stable ecosystems. In contrast, under the assumption of continued historical trends with a 2.7°C increase by 2100 (SSP245 - Middle-of-the-Road pathway), forecasts for Main Camp South indicate fluctuations with noticeable peaks and troughs, followed by a plateau, while at South Site, forecasts are characterized by recurring peaks and troughs with gradual stabilization. At No. 3 Bore, the forecast reflects a sharper decline in species abundance around the hotter months of 2024, aligning with the pathway's moderate environmental pressures and resource use. Under SSP370 (Regional Rivalry pathway) species abundance is projected to decline at both Main Camp South and South Site by 2026, reflecting the pathway's poor environmental management and increasing pressures on resources. However, No. 3 Bore is expected to reach a plateau, indicating some level of stability despite the adverse conditions.

Across all three grids, the 95% prediction intervals for the forecasts show that the model accounts for potential extreme peaks in captures during boom periods, reflecting the inherent variability and uncertainty in species abundance under different scenarios. Compared to the Sustainability scenario (SSP126), slightly lower uncertainty is observed under the Regional Rivalry scenario (SSP370), suggesting greater precision in abundance forecasts with

increased emissions and temperature levels. This highlights the model's capacity to capture the dynamic nature of population trends, particularly in response to a rapidly changing climate (Fig. 5.2).

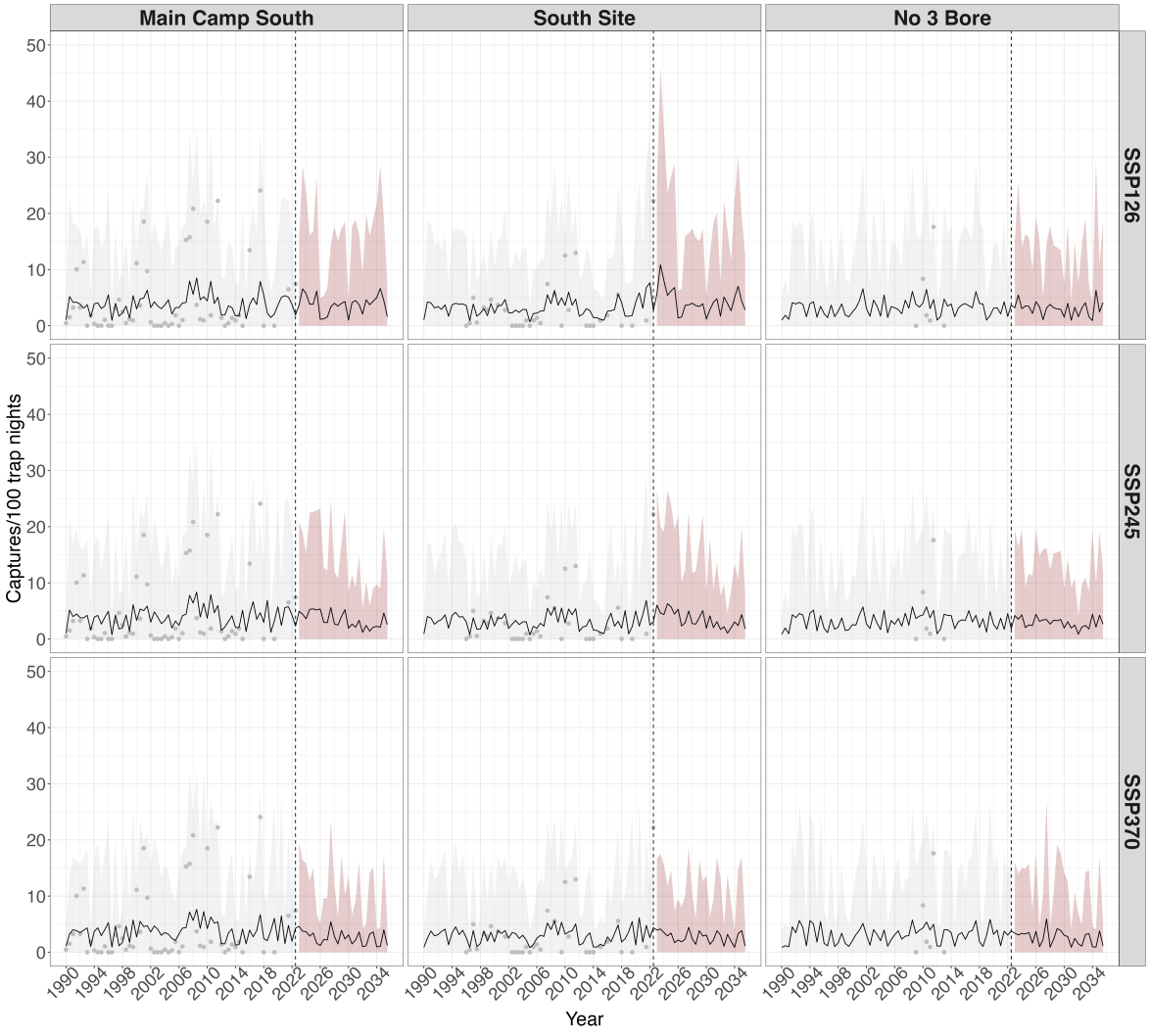

FIGURE 5.2. Hindcasts, forecasts and uncertainty representation for *Pseudomys hermannsburgensis captures at the Main Camp South, South Site and No. 3 Bore grids, based on the model presented in Section 5.2.4*. Each row corresponds to forecasts under a different climate change scenario: SSP126 (1.8°C increase by 2100), SSP245 (2.7°C increase by 2100), and SSP370 (3.6°C increase by 2100). Dots represent observed captures, while the black solid line shows posterior hindcasts (left of 2022) and forecasts (right of 2022). The grey and pink shaded regions illustrate the 95% prediction intervals.

However, when extending the forecast horizon to 12 years (from 2022 to 2035), the observed trend diverges compared to the initial 4-year forecast period. Despite variations in projected

climate change scenarios, the forecasted abundances do not show a drastic divergence across the three grids, confirming the synchronicity of the species in the long run. On the other hand, none of climate change scenarios led to local extinction (zero values) for the species. The main differences are in the magnitude of fluctuations in abundances, which appear to be more strongly influenced by the availability of observed data than by the climate change scenarios. Notably, the uncertainty representation of the 12 year forecasts remains consistent with that of the 4-year horizon, indicating that the model's uncertainty estimate remains stable despite the longer forecast period (Fig. 5.2).

5.3.2 Forecasting abundance of a species with stable dynamics

For *S. youngsoni*, hindcasts at Main Camp South showed occasional peaks in abundance between 1996 and 2012, mainly due to the relatively high abundances observed during this period, though still lower than those of *P. hermannsburgensis*. In contrast, hindcasts for South Site and No. 3 Bore were more stable, reflecting the less dynamic nature of *S. youngsoni*, when compared to *P. hermannsburgensis* (Fig. 5.3).

The 95% prediction interval during the hindcast phase effectively captured all observed peak abundances. This was especially evident at Main Camp South, where the prediction interval showed substantial variation. The fluctuating nature of the prediction interval at this site reflects the inherent uncertainty and variability in population trends, particularly during periods of higher-than-average abundances that contribute to the observed occasional peaks (Fig. 5.3).

When examining forecasted abundances across both 4-year and 12 year forecast horizons, as expected, the forecasts showed less dynamism compared to those for *P. hermannsburgensis*. However, the forecasts remained relatively stable in line with observed counts from previous years. The impact of rainfall under different climate change scenarios, appeared more muted for *S. youngsoni* than for *P. hermannsburgensis*. In instances where fewer captures for *S. youngsoni* were available (e.g. No. 3 Bore), the model's predictive power for varying climate change scenarios diminished, leading to a reduced responsiveness to changes represented by

the scenarios. However, the uncertainty remained consistent with that observed during the hindcast phase, highlighting the expectation of fluctuating, but with less pronounced boom and bust dynamics in the captures (Fig. 5.3).

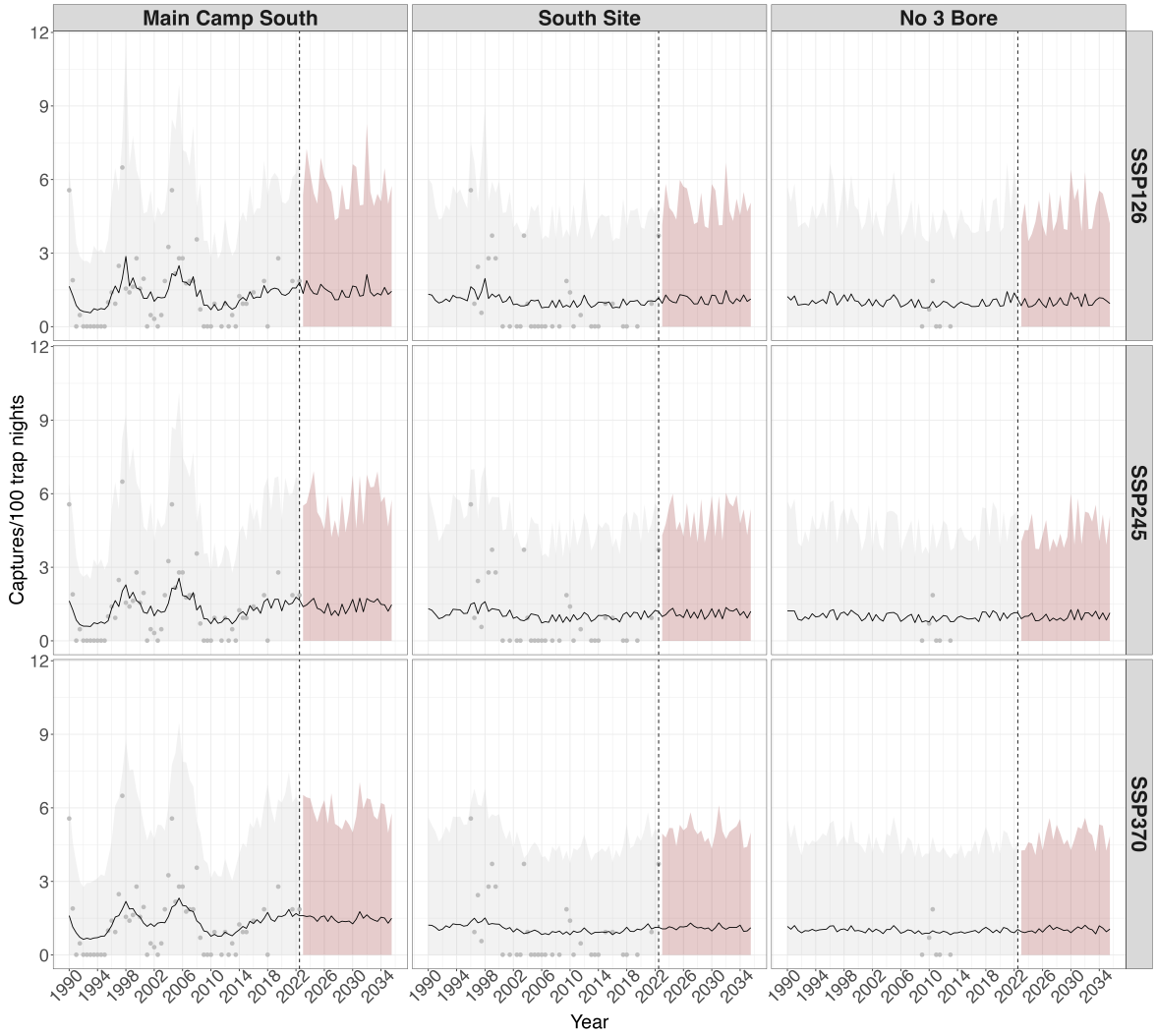

FIGURE 5.3. Hindcasts, forecasts and uncertainty representation for *Sminthopsis youngsoni abundances at the Main Camp South, South Site and No. 3 Bore grids, based on the model presented in Section 5.2.4*. Each row corresponds to forecasts under a different climate change scenario: SSP126 (1.8°C increase by 2100), SSP245 (2.7°C increase by 2100), and SSP370 (3.6°C increase by 2100). Dots represent observed counts, while the black solid line shows posterior hindcasts (left of 2022) and forecasts (right of 2022). The grey and pink shaded regions illustrate the 95% prediction intervals.

5.4 Discussion

Forecasting species dynamics beyond observed data represents a critical step in advancing ecological forecasting. By projecting possible future conditions that influence species population abundance and forecasting for near-term horizons, we can develop adaptive management strategies to mitigate imbalances in the system before they occur. As new data become available, this process becomes iterative, allowing for model updates and continuous improvement of forecasts under anticipated future conditions (Dietze et al., 2018).

Our study used two species with contrasting dynamics; *P. hermannsburgensis* and *S. youngsoni* to highlight the potential of multivariate generalized additive models (MVGAMs) to capture species-specific dynamics and evaluate their responses under varying climate change scenarios. These models incorporate nonlinear interactions between species dynamics and abiotic factors, along with latent dynamic processes that capture shared trends across species or locations, providing a robust framework for forecasting population trends in highly variable environments. We used projected rainfall under three different climate scenarios as the only covariate to forecast abundances 12 years ahead of 2022.

Our results revealed that the species population abundance forecasts are influenced by the change in projected rainfall under the three climate change scenarios and vary based on species-specific dynamics and the availability of observed data. Under the sustainability-focused scenario (SSP126) expecting a 1.8°C increase by 2100, the forecasts for *P. hermannsburgensis* suggested relative stability in population trends across all monitored sites, with the model capturing patterns reflective of the best case scenario if the world acts on climate change mitigation. For *S. youngsoni*, the model forecasts limited fluctuations in abundance across all three grids, reflecting the species' inherently stable population dynamics. While the forecast uncertainty remains consistent across both the near-term horizons, the longer-term projections suggest that the species abundance of *S. youngsoni* is less sensitive to climate variability compared to *P. hermannsburgensis* (Dickman et al., 2001; Letnic and Dickman, 2010; Greenville et al., 2016; Bennison et al., 2013, 2018). However, we found that the availability of observed data influenced the dynamics forecasted in both the species

more than the impact from climate change (Ward et al., 2014; Lewis et al., 2022). Therefore, maintaining sustainable environmental policies that lead to less greenhouse gas emissions and cause less drastic changes in the climate could help preserve species populations.

If the management actions for current greenhouse gas emissions remain the same but leading to 2.7°C increase by 2100, under the Middle-of-the-Road pathway (SSP245), looking on a shorter forecast horizon, *P. hermannsburgensis* forecasts revealed moderate fluctuations with peaks and troughs occurring at regular intervals at all sites but experienced a sharper decline in species abundance when the observed data are limited (e.g., No. 3 Bore). However, the longer forecast horizon (12 year) suggests that fluctuations continued, but do not deviate significantly from historical trends. *Sminthopsis youngsoni* forecasts indicate stable population levels at all sites but with the potential of slight declines over time at sites with limited data. These trends highlight the species' resilience but also suggest that increased environmental stressors may lead to gradual declines (Graham et al., 2019). Therefore, management approaches should focus on monitoring population dynamics closely, especially at sites with limited data, and implementing localized conservation efforts to mitigate anticipated population declines (Pacifici et al., 2015).

Finally, under the Regional Rivalry pathway (SSP370) which leads to a 3.6°C in 2100, *P. hermannsburgensis* populations are forecasted to plateau when the observed data are limited as opposed to significant declines at other sites when looked at the shorter horizon, reflecting poor environmental management and increasing resource pressures. In contrast, *S. youngsoni* forecasts indicate limited population decline, with stable trends continuing under both forecast horizons. The model's forecasts suggest that while *S. youngsoni* is more resistant to climate variability, prolonged poor environmental policies could still lead to long-term population reductions. Conservation strategies under high-emission scenarios should prioritize habitat protection, targeted species interventions, and the establishment of resilient ecosystems to counteract the negative impacts of climate change on more dynamic species.

Overall, we observed that both the species; *P. hermannsburgensis* and *S. youngsoni* appeared realtively resilient to climate change regardless of their species dynamics. However, increased environmental variability, particularly in food availability and species interactions

such as competition and predation, could lead to population declines, especially for species with pronounced boom-bust cycles. While these fluctuations may reduce abundance, they are unlikely to result in stochastic extinction (Letnic and Dickman, 2006; Greenville et al., 2018; Letnic et al., 2011; Greenville et al., 2017). To enhance forecast reliability, particularly for species or locations with more stable population trends, it is essential to obtain sufficient, high-quality observational data and incorporate additional ecological drivers influencing species abundance (Clark et al., 2025). Identifying potential population declines in advance allows targeted interventions such as habitat restoration, resource allocation, or predator control (White et al., 2019; Simonis et al., 2022). On the other hand, the appropriate management strategies may vary depending on the forecast horizon. Short-term forecasts (e.g. 4-years) allow for immediate on-ground actions aimed at mitigating rapid declines or stabilizing populations during critical periods. Long-term forecasts (e.g. 12 years) provide a broader perspective on potential ecological trajectories such as the effect of ENSO cycles, aiding in the formulation of policies for habitat connectivity, long-term monitoring programs, and climate adaptation strategies. This distinction underscores the importance of aligning conservation efforts with the temporal scale of anticipated changes (Pierson et al., 2015; Adler et al., 2020; Tulloch et al., 2020).

5.4.1 Limitations and future directions

One key limitation of the current study is the exclusion of vegetation dynamics, which play a critical role in shaping the abundance and distribution of desert small mammals (Chapter 2, Greenville et al. (2016)). Vegetation provides essential resources, such as food and shelter, which are crucial for the survival and reproductive success of these species (Murray et al., 1999; Spencer et al., 2014; Greenville et al., 2016). Future models would benefit from the incorporation of vegetation dynamics, which could significantly improve our understanding of habitat-mediated population trends. Specifically, projections of vegetation data, such as the Normalized Difference Vegetation Index (NDVI), could offer valuable insights into the relationship between vegetation cover and species abundance. NDVI, which serves as a

proxy for vegetation productivity, could enhance forecasts, especially for species like *P. hermannsburgensis*, which exhibit highly dynamic population fluctuations (Clark et al., 2025; Karunarathna et al., 2024). Alternatively, overall vegetation cover could be replaced by dominant vegetation species data (Chapter 4). By integrating vegetation metrics into predictive models we could improve the accuracy and robustness of population forecasts.

Another challenge faced when generating forecasts is the lack of observed data to evaluate the performance of the forecasts using formal metrics (e.g. Energy Score as outlined in Chapter 4). As a result, the evaluation of model performance is limited to the representation of uncertainty. This reliance on uncertainty estimation underscores the importance of refining the fitted model iteratively as new observed data become available; opening the concept of iterative near-term forecasting (Dietze et al., 2018). Moreover, the incorporation of uncertainty into management decision-making is essential, particularly in the absence of ground-truth data, as it allows for more cautious and informed strategies in species conservation and habitat management (Cressie et al., 2009; Lofton et al., 2022).

By addressing the limitations of the current approach and incorporating additional projected ecological variables, future research can significantly improve our capacity to forecast and respond to population trends in a rapidly changing environment. These advancements will play a crucial role in enhancing our understanding of species-environment interactions and refining management strategies. Ultimately, they will be vital for ensuring the long-term sustainability of biodiversity and the essential ecosystem services in arid environments.

References

Adler, P. B., White, E. P., and Cortez, M. H. (2020). Matching the forecast horizon with the relevant spatial and temporal processes and data sources. *Ecography*, 43(11):1729–1739.

ARC Centre of Excellence for Climate Extremes (2022). Climate-cms. http://climate-cms.wikis.unsw.edu.au/Home. [Accessed on 13/01/2025].

Ashok, K., Guan, Z., and Yamagata, T. (2003). Influence of the Indian Ocean Dipole on the Australian winter rainfall. *Geophysical Research Letters*, 30(15).

Beaumont, L. J., Hughes, L., and Poulsen, M. (2005). Predicting species distributions: use of climatic parameters in bioclim and its impact on predictions of species' current and future distributions. *Ecological Modelling*, 186(2):251–270.

Beaumont, L. J., Pitman, A., Poulsen, M., and Hughes, L. (2007). Where will species go? incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology*, 13(7):1368–1385.

Bennison, K., Dickman, C. R., and Godfree, R. (2013). Habitat use and ecological observations of the Ooldea dunnart *(Sminthopsis ooldea)* at Uluru–Kata Tjuta National Park, Northern Territory. *Australian Mammalogy*, 35(2):175–179.

Bennison, K., Godfree, R., and Dickman, C. R. (2018). Synchronous boom–bust cycles in central Australian rodents and marsupials in response to rainfall and fire. *Journal of Mammalogy*, 99(5):1137–1148.

Bullock, R. (2007). Great circle distances and bearings between two locations. *MDT, June*, 5:1–3.

Bureau of Meteorology (2024). Climate data online. http://www.bom.gov.au/climate/data/. Accessed on 26/12/2024.

Cai, W., Cowan, T., and Sullivan, A. (2009). Recent unprecedented skewness towards positive Indian Ocean Dipole occurrences and its impact on Australian rainfall. *Geophysical Research Letters*, 36(11).

Carpenter, S. R. (2002). Ecological futures: building an ecology of the long now. *Ecology*, 83(8):2069–2083.

Caswell, H. (2000). *Matrix population models*, volume 1. Sinauer Sunderland, MA.

Chapman, S., Syktus, J., Trancoso, R., Thatcher, M., Toombs, N., Wong, K. K. H., and Takbash, A. (2023). Evaluation of Dynamically Downscaled CMIP6-CCAM Models Over Australia. *Earth's Future*, 11(11):e2023EF003548.

Chapman, S., Syktus, J., Trancoso, R., Toombs, N., and Eccles, R. (2024). Projected changes in mean climate and extremes from downscaled high-resolution CMIP6 simulations in Australia. *Weather and Climate Extremes*, 46:100733.

Clark, J. S., Carpenter, S. R., Barber, M., Collins, S., Dobson, A., Foley, J. A., Lodge, D. M., Pascual, M., Pielke Jr, R., Pizer, W., et al. (2001). Ecological forecasts: an emerging imperative. *Science*, 293(5530):657–660.

Clark, N. J., Ernest, S. M., Senyondo, H., Simonis, J., White, E. P., Yenni, G. M., and Karunarathna, K. (2025). Beyond single-species models: leveraging multispecies forecasts to navigate the dynamics of ecological predictability. *PeerJ 13:e18929*.

Clark, N. J. and Wells, K. (2023). Dynamic generalised additive models (DGAMs) for forecasting discrete ecological time series. *Methods in Ecology and Evolution*, 14(3):771–784.

Copernicus Climate Change Service, Climate Data Store (2021). CMIP6 climate projections. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.c866074c. Accessed on 25/11/2024.

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., and Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, 19(3):553–570.

Dickman, C. R., Greenville, A. C., Beh, C.-L., Tamayo, B., and Wardle, G. M. (2010). Social organization and movements of desert rodents during population "booms" and "busts" in central Australia. *Journal of Mammalogy*, 91(4):798–810.

Dickman, C. R., Haythornthwaite, A. S., McNaught, G. H., Mahon, P. S., Tamayo, B., and Letnic, M. (2001). Population dynamics of three species of dasyurid marsupials in arid central Australia: a 10-year study. *Wildlife Research*, 28(5):493–506.

Dietze, M. C. (2017). *Ecological Forecasting*. Princeton University Press.

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., et al. (2018). Iterative

near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences*, 115(7):1424–1432.

Ernest, S. M., Brown, J. H., and Parmenter, R. R. (2000). Rodents, plants, and precipitation: spatial and temporal dynamics of consumers and resources. *Oikos*, 88(3):470–482.

Finley, A. O., Banerjee, S., and Gelfand, A. E. (2012). Bayesian dynamic modeling for large space-time datasets using Gaussian predictive processes. *Journal of Geographical Systems*, 14:29–47.

Fujiwara, M. and Caswell, H. (2002). Estimating population projection matrices from multi-stage mark–recapture data. *Ecology*, 83(12):3257–3265.

Gauthier, G., Péron, G., Lebreton, J.-D., Grenier, P., and van Oudenhove, L. (2016). Partitioning prediction uncertainty in climate-dependent population models. *Proceedings of the Royal Society B: Biological Sciences*, 283(1845):20162353.

Giorgi, F. (2019). Thirty years of regional climate modeling: where are we and where are we going next? *Journal of Geophysical Research: Atmospheres*, 124(11):5696–5723.

Graham, E. M., Reside, A. E., Atkinson, I., Baird, D., Hodgson, L., James, C. S., and VanDerWal, J. J. (2019). Climate change and biodiversity in Australia: a systematic modelling approach to nationwide species distributions. *Australasian Journal of Environmental Management*, 26(2):112–123.

Greenville, A. C., Burns, E., Dickman, C. R., Keith, D. A., Lindenmayer, D. B., Morgan, J. W., Heinze, D., Mansergh, I., Gillespie, G. R., Einoder, L., et al. (2018). Biodiversity responds to increasing climatic extremes in a biome-specific manner. *Science of the Total Environment*, 634:382–393.

Greenville, A. C., Dickman, C. R., Wardle, G. M., and Letnic, M. (2009). The fire history of an arid grassland: the influence of antecedent rainfall and ENSO. *International Journal of Wildland Fire*, 18(6):631–639.

Greenville, A. C., Wardle, G. M., and Dickman, C. R. (2013). Extreme rainfall events predict irruptions of rat plagues in central Australia. *Austral Ecology*, 38(7):754–764.

Greenville, A. C., Wardle, G. M., and Dickman, C. R. (2017). Desert mammal populations are limited by introduced predators rather than future climate change. *Royal Society Open Science*, 4(11):170384.

Greenville, A. C., Wardle, G. M., Nguyen, V., and Dickman, C. R. (2016). Population dynamics of desert mammals: similarities and contrasts within a multispecies assemblage. *Ecosphere*, 7(5):e01343.

Karunarathna, K., Wells, K., and Clark, N. J. (2024). Modelling nonlinear responses of a desert rodent species to environmental change with hierarchical dynamic generalized additive models. *Ecological Modelling*, 490:110648.

Letnic, M. and Dickman, C. R. (2006). Boom means bust: interactions between the El Niño/Southern Oscillation (ENSO), rainfall and the processes threatening mammal species in arid Australia. *Biodiversity & Conservation*, 15:3847–3880.

Letnic, M. and Dickman, C. R. (2010). Resource pulses and mammalian dynamics: conceptual models for hummock grasslands and other Australian desert habitats. *Biological Reviews*, 85(3):501–521.

Letnic, M., Story, P., Story, G., Field, J., Brown, O., and Dickman, C. R. (2011). Resource pulses, switching trophic control, and the dynamics of small mammal assemblages in arid Australia. *Journal of Mammalogy*, 92(6):1210–1222.

Lewis, A. S., Rollinson, C. R., Allyn, A. J., Ashander, J., Brodie, S., Brookson, C. B., Collins, E., Dietze, M. C., Gallinat, A. S., Juvigny-Khenafou, N., et al. (2023). The power of forecasts to advance ecological theory. *Methods in Ecology and Evolution*, 14(3):746–756.

Lewis, A. S., Woelmer, W. M., Wander, H. L., Howard, D. W., Smith, J. W., McClure, R. P., Lofton, M. E., Hammond, N. W., Corrigan, R. S., Thomas, R. Q., et al. (2022). Increased adoption of best practices in ecological forecasting enables comparisons of forecastability. *Ecological Applications*, 32(2):e2500.

Lofton, M. E., Brentrup, J. A., Beck, W. S., Zwart, J. A., Bhattacharya, R., Brighenti, L. S., Burnet, S. H., McCullough, I. M., Steele, B. G., Carey, C. C., et al. (2022). Using near-term forecasts and uncertainty partitioning to inform prediction of oligotrophic lake cyanobacterial density. *Ecological Applications*, page e2590.

McGregor, J. L. (2005). *C-CAM: Geometric aspects and dynamical formulation*, volume 70. CSIRO Atmospheric Research Dickson ACT, Australia.

McGregor, J. L. and Dix, M. R. (2008). An Updated Description of the Conformal-Cubic Atmospheric Model. In *High Resolution Numerical Modelling of the Atmosphere and Ocean*, pages 51–75. Springer.

Meinshausen, M., Nicholls, Z. R., Lewis, J., Gidden, M. J., Vogel, E., Freund, M., Beyerle, U., Gessner, C., Nauels, A., Bauer, N., et al. (2020). The shared socio-economic pathway (SSP) greenhouse gas concentrations and their extensions to 2500. *Geoscientific Model Development*, 13(8):3571–3605.

Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., Faure, D., Garnier, E., Gimenez, O., Huneman, P., et al. (2015). Predictive ecology in a changing world. *Journal of Applied Ecology*, 52(5):1293–1310.

Murray, B., Dickman, C., Watts, C., and Morton, S. R. (1999). The dietary ecology of Australian rodents. *Wildlife Research*, 26(6):857–858.

Nichols, J. D., Koneff, M. D., Heglund, P. J., Knutson, M. G., Seamans, M. E., Lyons, J. E., Morton, J. M., Jones, M. T., Boomer, G. S., and Williams, B. K. (2011). Climate change, uncertainty, and natural resource management. *The Journal of Wildlife Management*, 75(1):6–18.

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., Knutti, R., Kriegler, E., Lamarque, J.-F., Lowe, J., et al. (2016). The scenario model intercomparison project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, 9(9):3461–3482.

Pacifici, M., Foden, W. B., Visconti, P., Watson, J. E., Butchart, S. H., Kovacs, K. M., Scheffers, B. R., Hole, D. G., Martin, T. G., Akçakaya, H. R., et al. (2015). Assessing species vulnerability to climate change. *Nature Climate Change*, 5(3):215–224.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1):439–446.

Petchey, O. L., Pontarp, M., Massie, T. M., Kéfi, S., Ozgul, A., Weilenmann, M., Palamara, G. M., Altermatt, F., Matthews, B., Levine, J. M., et al. (2015). The ecological forecast horizon, and examples of its uses and determinants. *Ecology Letters*, 18(7):597–611.

Pierson, J. C., Beissinger, S. R., Bragg, J. G., Coates, D. J., Oostermeijer, J. G. B., Sunnucks, P., Schumaker, N. H., Trotter, M. V., and Young, A. G. (2015). Incorporating evolutionary

processes into population viability models. *Conservation Biology*, 29(3):755–764.

Piirainen, S., Lehikoinen, A., Husby, M., Kålås, J. A., Lindström, Å., and Ovaskainen, O. (2023). Species distributions models may predict accurately future distributions but poorly how distributions change: A critical perspective on model validation. *Diversity and Distributions*, 29(5):654–665.

Purdie, R. (1984). Land systems of the Simpson desert region. *Natural Resources Series*, 2:1–71.

Rummukainen, M. (2016). Added value in regional climate modeling. *Wiley Interdisciplinary Reviews: Climate Change*, 7(1):145–159.

Saxon, E., Baker, B., Hargrove, W., Hoffman, F., and Zganjar, C. (2005). Mapping environments at risk under different global climate change scenarios. *Ecology Letters*, 8(1):53–60.

Shephard, M. (1992). *The Simpson Desert: natural history and human endeavour*. Royal Geographical Society of Australasia, Adelaide, Australia.

Simonis, J. L., Yenni, G. M., Bledsoe, E. K., Christensen, E. M., Senyondo, H., Taylor, S. D., Ye, H., White, E. P., and Ernest, S. M. (2022). portalcasting: Supporting automated forecasting of rodent populations. *Journal of Open Source Software*, 7(72):3220.

Spencer, E. E., Crowther, M. S., and Dickman, C. R. (2014). Diet and prey selectivity of three species of sympatric mammalian predators in central Australia. *Journal of Mammalogy*, 95(6):1278–1288.

Stein, B. A., Staudt, A., Cross, M. S., Dubois, N. S., Enquist, C., Griffis, R., Hansen, L. J., Hellmann, J. J., Lawler, J. J., Nelson, E. J., et al. (2013). Preparing for and managing change: climate adaptation for biodiversity and ecosystems. *Frontiers in Ecology and the Environment*, 11(9):502–510.

Syktus, J., Trancoso, R., Ahrens, D., Toombs, N., and Wong, K. (2020). Queensland future climate dashboard: downscaled CMIP5 climate projections for Queensland. *Available from: https://www.longpaddock.qld.gov.au/qld-future-climate/*.

Thibault, K. M., Ernest, S. M., White, E. P., Brown, J. H., and Goheen, J. R. (2010). Long-term insights into the influence of precipitation on community dynamics in desert rodents. *Journal of Mammalogy*, 91(4):787–797.

Thuiller, W. (2004). Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, 10(12):2020–2027.

Thuiller, W., Araújo, M. B., Pearson, R. G., Whittaker, R. J., Brotons, L., and Lavorel, S. (2004). Uncertainty in predictions of extinction risk. *Nature*, 430(6995):34–34.

Trancoso, R., McGloin, R., Putland, D., Syktus, J., Ahrens, D., Chapman, S., Owens, D., Toombs, N., Eccles, R., and Zhang, H. (2024). Queensland Future Climate Services - downscaled CMIP6 climate projections for Queensland's regions and locations. Accessed from https://www.longpaddock.qld.gov.au/qld-future-climate/data-info/tern-cmip6/.

Tsiftsis, S., Štípková, Z., Rejmánek, M., and Kindlmann, P. (2024). Predictions of species distributions based only on models estimating future climate change are not reliable. *Scientific Reports*, 14(1):25778.

Tulloch, A. I., Hagger, V., and Greenville, A. C. (2020). Ecological forecasts to inform near-term management of threats to biodiversity. *Global Change Biology*, 26(10):5816–5828.

Valle, D., Staudhammer, C. L., Cropper Jr, W. P., and Gardingen, P. R. (2009). The importance of multimodel projections to assess uncertainty in projections from simulation models. *Ecological Applications*, 19(7):1680–1692.

Ward, E. J., Holmes, E. E., Thorson, J. T., and Collen, B. (2014). Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. *Oikos*, 123(6):652–661.

Wardle, G. and Dickman, C. (2015). Desert Ecology Plot Network: Mammal, Reptile and Vegetation Data Associated with Weather, Simpson Desert, Western Queensland, Australia, 1990–2011. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern.111.57/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018a). Desert Ecology Plot Network: Mammal Abundance Plot-data, Simpson Desert, Western Queensland, 1990+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.196.8/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018b). Desert Ecology Plot Network: Vegetation Plot-data, Simpson Desert, Western Queensland, 1993+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.195.11/html. Accessed on 16/01/2025.

Wardle, G. and Dickman, C. (2018c). Desert Ecology Plot Network: Weather Data (daily and monthly), Simpson Desert, Western Queensland, 1995+. Long Term Ecological Research Network. http://www.ltern.org.au/knb/metacat/ltern6.194.7/html. Accessed on 16/01/2025.

Wardle, G. M., Greenville, A. C., Frank, A. S., Tischler, M., Emery, N. J., and Dickman, C. R. (2015). Ecosystem risk assessment of Georgina gidgee woodlands in central Australia. *Austral Ecology*, 40(4):444–459.

White, E. P., Yenni, G. M., Taylor, S. D., Christensen, E. M., Bledsoe, E. K., Simonis, J. L., and Ernest, S. K. M. (2019). Developing an automated iterative near-term forecasting system for an ecological study. *Methods in Ecology and Evolution*, 10(3):332–344.

CHAPTER 6

General Discussion and Conclusion

6.1 Overview

"Ecological forecasting is a key component of decision support, as all decisions are fundamentally about the future and are sensitive to forecast uncertainties. However, decisions are ultimately about human values in how we balance trade-offs among competing objectives." Dietze (2017)

Ecological forecasting is indispensable for adaptive management (Berkes et al., 2000), offering a foundation for decision-making in the face of uncertainty and environmental change. However, ecosystems are inherently complex and variable, with uncertainties arising from gaps in data, incomplete understanding of ecological processes, and unpredictable environmental changes. To address these challenges, ecological forecasts must be supported by long-term datasets with consistent methodologies, covering multiple locations and species to capture spatiotemporal dynamics accurately (Lindenmayer et al., 2012; Kuebbing et al., 2018).

Key considerations in ecological forecasting include aligning ecological assumptions with statistical models, effectively managing missing data, and selecting models that balance the need for complexity to capture ecological dynamics with the interpretability required for practical application (Lewis et al., 2023). Furthermore, accurate hindcasts and reliable predictions require models to perform across varying forecast horizons while incorporating projected environmental data to provide actionable, real-world forecasts (Dietze, 2017; Adler et al., 2020).

Recent advancements in data science, particularly in statistical modeling and Bayesian approaches, have significantly improved ecological forecasting. These techniques allow for the assimilation of diverse data streams, quantification of forecast uncertainty, and the production of more robust predictions (Dietze, 2017; Lewis et al., 2022). By leveraging these tools, ecological forecasters can provide more informed insights to guide management decisions and prepare for future environmental scenarios.

This thesis explores these themes through the lens of forecasting desert rodent populations in Australian ecosystems. We emphasize the importance of robust datasets, thoughtful model selection, and seamless integration of ecological and statistical frameworks. To bring this together, Table 6.1 summarizes the key areas addressed in each thesis chapter and the main findings. The discussion evaluates the key findings, extends the relevance of these advances through examples of practical application of knowledge and approaches used in the thesis to industry, and examines the limitations, and potential future research directions. By addressing these challenges, this work advances ecological forecasting as a critical tool for managing biodiversity and ecosystem function in an uncertain and changing world.

6.2 Key findings

"Data science is situated in the problem domain and that we have a data science of the natural environment and not data science for the natural environment." Blair et al. (2019)

A key message of this thesis is the importance of ensuring coherance between statistical modeling and ecological complexities. One common challenge in bridging data science and ecology is the lack of understanding of the system being modeled, including species dynamics, environmental influences, data collection methodology and the challenges involved in obtaining data (Blair et al., 2019). In fields like eCommerce, supply chain management, and healthcare, the data collection process may be less critical because the systems are well-defined with predictable variables (Dhar, 2013; Provost and Fawcett, 2013). These domains operate within structured frameworks where key variables are systematically recorded, often through automated processes such as digital transactions, GPS tracking, and electronic health

TABLE 6.1. Summary of the key areas addressed in this thesis and the key findings.

Chapter	Themes addressed	Methods used	Key outcomes
2	Ecological and statistical assumptions	Multivariate state-space modeling	Align statistical model assumptions with ecological relevance to ensure robustness
3	Missing values	Single and multiple imputation techniques	Robustness of imputation techniques across different missing value characteristics
4	Model training length and forecast horizon	Multivariate generalized additive modeling	Reliable long-term forecasts with small datasets despite high uncertainty
5	Near-term forecasts	Multivariate generalized additive modeling with climate change projections	Contrasting forecasts for different near-term horizons, climate change scenarios and data availability

records unlike fields that require extensive human-labour and are prone to inherent complexities. Ecology falls into the latter category, where the data collection process is fundamental to understanding the system due to the complexity and unpredictability of ecological systems, especially in highly variable environments driven by extreme climatic conditions. For example, counting small nocturnal mammals, such as rodents or bats, involves techniques like live trapping, camera traps, or acoustic monitoring (Greenville et al., 2016; Dorazio and Karanth, 2017; Milchram et al., 2020; Bruce et al., 2025). These methods are essential because these animals are hard to detect and are active only at night, making direct observation challenging. In contrast, counting large herbivores like deer, sheep, or cows typically involves methods such as direct observation, aerial surveys, or GPS tracking, as these animals are larger, easier to detect, and often active during the day (Hebblewhite and Haydon, 2010; Terletzky and Ramsey, 2016). However, each method has its own advantages and limitations, and the choice of technique depends on factors like the species' size, behavior, habitat, and the specific research question. Therefore, long-term datasets that maintain consistent

methodologies, spatial coverage, and multi-species monitoring are invaluable, as they allow for a deeper understanding of ecological processes and trends, such as the Desert Ecology Research Group (DERG) dataset used in this study (Chapter 2). Such datasets help capture complex non-linear interactions between species and environmental factors (Lindenmayer et al., 2012; Blair et al., 2019; Kuebbing et al., 2018). Therefore, careful planning of resources, methodology, and system-specific challenges is essential before starting long-term monitoring or integrating datasets. Without this, predictions can be unreliable, which undermines the effectiveness of the model for forecasting and decision making (Chapter 2).

Modeling species population abundance allows for a deeper understanding of data derived from various observation and monitoring techniques, facilitating predictions of future trends and the exploration of "what-if" scenarios. However, for the model to be robust, the statistical assumptions must align with ecological relevance (Chapter 2). This highlights the value of integrating the perspectives of both ecologists and data scientists, ensuring a more comprehensive understanding of the system being studied (Blair et al., 2019). Using a case study that modeled species population abundance from the DERG dataset with a multivariate autoregressive state-space model (MARSS), we found that field-collected population data often exhibit high variability due to observation and process errors (Thomas, 1996; Humbert et al., 2009; Amano et al., 2012), as well as deviations from normal distributions (Chapter 2). As a result, it is unlikely that all statistical assumptions will be fully satisfied (Eastoe and Tawn, 2009). However, this does not justify discarding the model results. Instead, it is crucial to interpret the findings cautiously, acknowledging their limitations while still drawing meaningful insights. By adopting this careful approach, we can ensure that predictions and analyses remain reliable, even in the face of the inherent complexities and uncertainties in ecological data (Chapter 2). When validating ecological assumptions, we found that single-population MARSS models, fitted for species exhibiting similar dynamics across sites, assumed a common process across locations, with variability in abundance primarily driven by sampling error. In contrast, multi-population MARSS models, applied to species displaying distinct site-specific dynamics, assumed different processes at each site, where variability in abundance was less influenced by sampling error. Regardless of response patterns, the density dependence appeared to be significant for all study species. However, for rodent

species showing synchronous behaviour across locations, previous year rainfall and spinifex seed significantly influenced their abundances aligning with the findings of Greenville et al. (2016). Similarly, the predatory mulgara, which also displayed synchronous dynamics, was influenced only by previous year rainfall. However, this contrasts with Greenville et al. (2016), which found additional influences from spinifex cover and rodent interactions. On the other hand, for species exhibiting asynchronous response patterns across locations, none of the environmental or climatic variables were significant, which also contrasts with the findings of Greenville et al. (2016) (Chapter 2). Therefore the statistical validity of these models highlight the importance of tailored model specifications that better capture species-specific population dynamics, enhancing their ecological relevance.

In data science, datasets are often characterized by volume (the amount of data), velocity (the speed at which data is generated and processed), variety (the different types or formats of data), and veracity (the accuracy or reliability of the data) (Jagadish et al., 2014). While fields such as finance, healthcare, and logistics prioritize volume and velocity due to the availability of high-frequency data, in ecology, variety and veracity take precedence (Nakagawa and Freckleton, 2008; Blair et al., 2019). Missing values are a critical issue across all domains, but the approach to handling them differs significantly. In fields with large, dense datasets, removing records with missing values typically has minimal impact on modeling outcomes, as sufficient data remain to capture trends. In contrast, field-surveyed ecological datasets are typically sparse, with fewer records per year due to financial and logistical constraints of sampling, even when collected over decades. Unlike other data collection procedures such as remote sensing or camera traps, we found that deleting missing data in field-collected datasets risks reducing the temporal resolution and reliability of the dataset, making it harder to model population abundances accurately (Chapter 3; Łopucki et al. (2022)). Moreover, substituting missing values with zero is not appropriate for ecological data, as the absence of sampling does not imply species absence; such an approach distorts time series trends and misrepresents population dynamics (De Souto et al., 2015; Zhang, 2016; Khan, 2024). Our analysis highlights the robustness of imputation methods (Hossie et al., 2021; Chapon et al., 2023; Sa'adi et al., 2023; Khan, 2024; Carpenter et al., 2025) in addressing missing values when modeling population dynamics. By employing a cumulative approach to test

different characteristics of missing data (e.g., temporal placement, quantity), we found that the accuracy of population abundance predictions was largely unaffected, underscoring the potential of imputation to preserve trends (Chapter 3; Niako et al. (2024)). However, single imputation methods often fail to adequately account for the variability and complexity inherent in ecological time series (Zhang, 2016; Raja et al., 2020; Khan, 2024). Instead, multiple imputation methods that incorporate relationships between highly dependent covariates, effectively capture trends, and reduce uncertainty in predictions are beneficial for any field requiring time series modeling and forecasting.

Ecological forecasting is a valuable tool for predicting population trends and understanding the effects of environmental changes, but its effectiveness depends on carefully selecting models, covariates, and evaluation methods. In this study, we applied the novel statistical technique of multivariate generalized additive models (MVGAM) to model population abundance, capturing spatial and species-level dynamics while incorporating nonlinear environmental and climatic effects (Clark and Wells, 2023). However, we found that using a Gaussian process to capture the latent trend is well-suited for conservation decisions in data-limited regions, as Gaussian processes provide reliable forecasts without excessive computational demands (Chapter 4). Although more complex trend models, like vector autoregressive (VAR) models (Stock and Watson, 2001; Lui et al., 2007; Hampton et al., 2013), may improve accuracy, their computational intensity can make them impractical when timely decisions are required. This highlights the importance of balancing computational efficiency with model accuracy; an essential trade-off in applied data science for ecological forecasting. We also understood that it is important to include species-specific covariates to the model. However, the significance of the covariates in the models varied depending on the proportion of data used for training and differed from our findings in Chapter 2, resulting in rainfall not being identified as a significant driver of population abundance. This suggests that the influence of covariates on small mammal abundance is highly dependent on both the model and the range of data used for training, regardless of the underlying species dynamics. In addition, we observed that using dominant vegetation species, such as *Triodia basedowii* (Spinifex), as a proxy for broader environmental conditions can reduce the need for exhaustive data collection. This is a key part of data preparation in data science, where using

domain-specific proxies can simplify modeling efforts without sacrificing predictive power (Carroll and Kenkel, 2019; Warner and Sloan, 2021; Bastani, 2021). By integrating data science into ecology, we used a Bayesian framework to generate probabilistic predictions of species population abundance. Models trained on smaller datasets can still produce reliable forecasts if they effectively capture "boom-bust" trends; however, uncertainty increases with longer horizons and for highly dynamic species like *Pseudomys hermannsburgensis*. We emphasize the importance of transparently communicating these uncertainties to support informed decision-making (Chapter 4). Finally, we found that understanding species-specific dynamics and using multivariate evaluation metrics, such as those assessing spatio-temporal relationships (Simonis et al., 2021), are critical for producing accurate and practical forecasts (Chapter 4).

When making near-term ecological forecasts of population abundance, it is crucial to consider the effects of different climate change scenarios on the future dynamics of species and ecosystems (Beaumont et al., 2007). High-resolution projections under varying greenhouse gas emission scenarios, are particularly important in regions with highly variable climates, where extreme conditions can drive significant ecological changes. Projections that consider all relevant climate effects; including oceanic and environmental changes within the study area, are necessary to ensure the reliability and accuracy of forecasts. MVGAMs offer a robust approach as they are especially effective when projecting into the future, as they can incorporate multi-species or multi-site specifications, even when data are sparse for specific species or sites (Chapter 5). Using MVGAMs with projected rainfall data, we found that, before making decisions, it is essential to evaluate envelopes of near-term forecasts (e.g. 4-12 years) because a decline of population abundance observed in a 4-year period would not be the same extended to 12 years (Chapter 5). We observed that for species with high variability, such as *Pseudonomys hermannsburgensis*, climate change scenarios can trigger significant changes in population dynamics, underscoring their sensitivity to environmental shifts. In contrast, more stable species, such as *Sminthopsis youngsoni*, show muted responses to climate change, suggesting a higher level of population resilience (Chapter 5; Dickman et al. (2001); Letnic and Dickman (2010); Greenville et al. (2016); Bennison et al. (2013, 2018)). However, the effect of climate change scenarios on forecasts also varies on

the availability of observed data (Ward et al., 2014; Lewis et al., 2022). On the other hand, the significance of the impact of projected rainfall on the species abundance varied according to the climate change scenario. The absence of future observed data for formal forecast evaluation limits our ability to validate these forecasts fully, making it necessary to rely on measures of uncertainty (Nichols et al., 2011). Incorporating these uncertainties into management strategies is essential, as it allows for more cautious and adaptive decision-making, especially when ground-truth data are lacking. Acknowledging and addressing uncertainty can prevent overconfidence in forecasts and guide more flexible and responsive management approaches (Dietze, 2017). We found that *P. hermannsburgensis* and *S. youngsoni* appeared relatively resilient to climate change, but increased environmental variability, particularly in food availability and species interactions, could drive population declines, however, without causing stochastic extinction (Letnic and Dickman, 2006; Greenville et al., 2018; Letnic et al., 2011; Greenville et al., 2017). Key findings from this study pave the way for practical applications across various industries by providing data-driven insights, refining methodologies, or introducing novel approaches. These findings can inform policy decisions, enhance predictive models, and improve resource management, making them valuable beyond academic research.

6.2.1 Practical application of skills and methods

The key findings of this thesis provide valuable guidance for ecological data collection programs, such as the National Ecological Observatory Network (NEON), and offer insights that can enhance ecological forecasting efforts, including initiatives like the NEON Ecological Forecasting Challenge (Thomas et al., 2023). Our results underscore the importance of integrating both domain expertise and data science knowledge to improve the understanding of data structures and experimental design in ecological forecasting. The use of imputation methods, robust statistical techniques, uncertainty quantification, and projected environmental variables under different climate change scenarios highlights the need for tailored, project-level strategies. We also stress the limitations of simple forecasting models, like ARIMA, in capturing non-continuous or seasonal ecological data, urging researchers to explore

hierarchical Bayesian models and multivariate time-series approaches. It is also important to account for forecast uncertainty and model evaluation, focusing on the integration of climate covariates to improve model accuracy and inform decision-making for ecosystem monitoring.

The practical application of the insights, feedback, and recommendations from this thesis highlights the demand across various industries for advanced analytical tools. One notable example of applying the skills and knowledge gained during my PhD is the internship I undertook at the Department of Climate Change, Energy, the Environment, and Water (DCCEEW). This experience allowed me to work closely with stakeholders who highlighted the pressing need for data-driven solutions to address key environmental challenges. A major focus was understanding biodiversity patterns and environmental drivers to guide the restoration and conservation of New South Wales' (NSW) natural capital. Research on vascular plant diversity within the Australian Regional Climate Modeling (NARCliM) extent has been limited by inconsistent, plot-based survey data. To overcome this challenge, we compiled and harmonized vegetation survey data from NSW and surrounding states, focusing on species richness, which included both native and total (native and introduced) species. This comprehensive dataset provided the foundation for predictive modeling, where we applied a Bayesian Generalized Additive Model (GAM) to predict species richness across the region. To face the challenge of different methodologies used to collect data from different data sources, we used a 400 m^2 (Mokany et al., 2022) plot area standardization approach to obtain consistent results. The model's performance was tested through the creation of training and testing sets, with accuracy assessed by comparing observed and predicted values. The results demonstrated the model's robustness, delivering strong predictions for both original and standardized plot scenarios. This work identified critical environmental drivers of species richness, such as temperature, precipitation, radiation, elevation, and topsoil composition. Insights drawn from the project are highly relevant to industry needs to create fine-resolution spatial predictions of plant diversity. By developing and applying these models, the project highlights the growing demand for advanced data science techniques in biodiversity management, showcasing their potential to predict and monitor ecological changes across multiple scales.

6.3 Limitations and future research

This thesis emphasizes the importance of long-term datasets with consistent methodologies and broad spatial and temporal coverage in ecological studies to deepen our understanding of species dynamics and their environmental drivers. The Desert Ecology Research Group (DERG) dataset exemplifies such a dataset, providing decades of population abundance data collected consistently across multiple sites (Chapter 2). However, like many field-surveyed monitoring projects, the DERG dataset is constrained by logistical and financial challenges, particularly the difficulty of sampling during the same time period each year, resulting in missing data. Traditional time series models as discussed in Chapter 1; both frequentist and probabilistic, require regular intervals between observations, often necessitating aggregation of data into seasonal or annual summaries, which can obscure fine-scale temporal patterns and interactions (Chapter 2, Chapter 4; Hostetler and Chandler (2015)). Future researchers can address these limitations in several ways. First, remote sensing or camera trap data could be integrated with field-surveyed datasets to provide high-frequency observations. However, uncertainties arising from different data collection methods must be explicitly modeled to ensure reliability (Lofton et al., 2022). Second, in addition to combining field surveys and digital data, researchers may consider using only equipment-driven data such as remote sensing, camera trap or radio tracking data, provided the methodologies are consistent and well-documented, and covariate data are also collected directly within the study site rather than relying on external sources (Lahoz-Monfort and Magrath, 2021). Third, the development of continuous-time models could alleviate the reliance on discrete sampling intervals. However, current continuous-time (e.g. Dennis and Ponciano (2014)) models are limited in capturing nonlinear interactions between covariates and population dynamics, highlighting the need for future advancements in this area. In addition, our study also emphasizes the value of interdisciplinary collaboration between ecologists and data scientists, as there is a need to establish more systematic approaches for integrating domain expertise into model design and validation.

This thesis also found that imputing missing values is an effective approach to addressing data gaps, offering a better alternative to disregarding missing records entirely. However,

these findings were based on a limited set of missing value characteristics observed within the DERG dataset and the specific type of missing data it included (Chapter 3). Furthermore, we employed a cumulative approach where as missing values tend to occur in the dataset with time, the compounding effects from previous missing values are considered. One of our aims with this approach was to evaluate how the performance of imputation techniques changed as the proportion of missing values in the dataset increased over time. As such, the generalizability of these results may be restricted to similar contexts and methodology. An intriguing avenue for future research would involve comparing the performance of different imputation techniques across a wider range of missing value characteristics and types, both in highly variable environments like those studied in the DERG dataset and in simpler, more stable ecological systems. This comparative approach would help evaluate the robustness of imputation methods under diverse conditions and inform best practices for addressing missing data in ecological research (Van Buuren, 2018).

Multivariate state-space models offer several advantages over univariate models in capturing the spatial complexities of population dynamics. However, they also come with certain limitations. A key advantage of multivariate generalized additive models (MVGAM) as opposed to the multivariate autoregressive state-space (MARSS) framework is the ability to incorporate nonlinear interactions between covariates and the response variable, allowing for a more nuanced representation of ecological dynamics (Chapter 4, Chapter 5; Greenville et al. (2016); Wood (2017); Clark and Wells (2023)). Additionally, MVGAMs can accommodate fine-resolution spatial data, such as grid-level observations, without significantly complicating parameter estimation or reducing computational efficiency. This flexibility is particularly valuable for ecological studies where spatial heterogeneity plays a critical role (Karunarathna et al., 2024). Unlike the MARSS framework, which typically employs simpler trend models, MVGAMs provide a wide array of latent trend models to capture the underlying temporal dynamics. This diversity allows researchers to balance computational efficiency and predictive accuracy when modeling population abundance, which is crucial in long-term ecological monitoring (Chapter 4). However, a limitation of MVGAMs arises when standardized counts are used to account for variable sampling efforts, resulting in decimal values. While the *mvgam* package supports the Gamma distribution for modeling such standardized counts, it

does not accommodate zeros in the time series. This can be problematic for ecological count data, where zeros are common due to the absence of individuals during certain sampling events (e.g., threatened species). In such a case, one approach is to add a small delta value to shift zero counts slightly above zero (Douma and Weedon, 2019) or raw counts can be used with an offset for survey effort, modeling the data with a Poisson or Negative Binomial distribution. Future developments in MVGAMs could help overcome this limitation, improving their applicability to ecological studies with zero-inflated datasets.

The use of projected covariate data in forecasting population abundance for near-term horizons demonstrates the potential to generalize our findings to broader ecological studies. Incorporating climate projections under various greenhouse gas emission scenarios, such as Shared Socioeconomic Pathways, further strengthens the validity of our study by accounting for potential future states shaped by human activities (Chapter 5; Trancoso et al. (2024)). These projections provide a more dynamic understanding of species population trends in response to climate change and other anthropogenic factors, offering valuable insights for more informed conservation strategies (Beaumont et al., 2007). However, a key gap in the literature and available data products is the limited availability of long-term projections for other critical environmental covariates, such as vegetation cover. While climate models have made considerable progress in generating future projections, long-term datasets for covariates like vegetation coverage remain scarce. This presents a significant avenue for future research, as improved models and datasets are necessary to predict how ecological systems will evolve over the coming decades, particularly under different levels of human intervention. Enhanced projections of environmental covariates would provide deeper insights into ecosystem dynamics, helping to better predict the impacts of climate change on species populations and inform effective conservation planning (Chapter 5).

6.4 Concluding remarks

Understanding the distribution and abundance of species has highlighted the need to bridge the gap between domain knowledge and data science (Chapter 1). The importance of using

long-term data, collected consistently with robust methodologies that cover high spatial and temporal scales, cannot be overstated. Such data are essential to inform researchers and organizations about effective resource management and planning, particularly before embarking on long-term monitoring projects. This thesis highlights the critical role of integrating these data with advanced statistical modeling techniques while ensuring alignment between statistical methods and ecological assumptions. Given the inevitability of missing data in ecological datasets, the application of imputation techniques; tailored to the nature of missingness, is vital for maintaining the integrity of analyses. Together, these approaches lay the groundwork for near-term ecological forecasting by leveraging multivariate statistical modeling frameworks and projected covariate data. This concept opens the avenue for even using short-term ecological datasets, provided they address limitations in long-term data and contain minimal observation noise and missing values. Moreover, while data science techniques have advanced to machine learning algorithms, the need to understand underlying ecological processes, where each data point tells a unique story, makes black-box approaches less suitable for this context. Therefore, throughout this thesis, I demonstrated that integrating solid, robust data science methods with domain-specific knowledge has immense potential to tackle ecological challenges, providing valuable insights for future biodiversity management and policy development.

References

Adler, P. B., White, E. P., and Cortez, M. H. (2020). Matching the forecast horizon with the relevant spatial and temporal processes and data sources. *Ecography*, 43(11):1729–1739.

Amano, T., Okamura, H., Carrizo, S. F., and Sutherland, W. J. (2012). Hierarchical models for smoothed population indices: the importance of considering variations in trends of count data among sites. *Ecological Indicators*, 13(1):243–252.

Bastani, H. (2021). Predicting with proxies: Transfer learning in high dimension. *Management Science*, 67(5):2964–2984.

Beaumont, L. J., Pitman, A., Poulsen, M., and Hughes, L. (2007). Where will species go? incorporating new advances in climate modelling into projections of species distributions. *Global Change Biology*, 13(7):1368–1385.

Bennison, K., Dickman, C. R., and Godfree, R. (2013). Habitat use and ecological observations of the Ooldea dunnart *(Sminthopsis ooldea)* at Uluru–Kata Tjuta National Park, Northern Territory. *Australian Mammalogy*, 35(2):175–179.

Bennison, K., Godfree, R., and Dickman, C. R. (2018). Synchronous boom–bust cycles in central Australian rodents and marsupials in response to rainfall and fire. *Journal of Mammalogy*, 99(5):1137–1148.

Berkes, F., Colding, J., and Folke, C. (2000). Rediscovery of traditional ecological knowledge as adaptive management. *Ecological Applications*, 10(5):1251–1262.

Blair, G. S., Henrys, P., Leeson, A., Watkins, J., Eastoe, E., Jarvis, S., and Young, P. J. (2019). Data science of the natural environment: a research roadmap. *Frontiers in Environmental Science*, 7:121.

Bruce, T., Amir, Z., Allen, B. L., Alting, B. F., Amos, M., Augusteyn, J., Ballard, G.-A., Behrendorff, L. M., Bell, K., Bengsen, A. J., et al. (2025). Large-scale and long-term wildlife research and monitoring using camera traps: a continental synthesis. *Biological Reviews*.

Carpenter, D., Deyle, E., Park, J., Saberski, E., and Sugihara, G. (2025). Repairing gaps in ecological time series. *Methods in Ecology and Evolution*. doi:10.1111/2041-210X.14491.

Carroll, R. J. and Kenkel, B. (2019). Prediction, proxies, and power. *American Journal of Political Science*, 63(3):577–593.

Chapon, A., Ouarda, T. B., and Hamdi, Y. (2023). Imputation of missing values in environmental time series by D-vine copulas. *Weather and Climate Extremes*, 41:100591.

Clark, N. J. and Wells, K. (2023). Dynamic generalised additive models (DGAMs) for forecasting discrete ecological time series. *Methods in Ecology and Evolution*, 14(3):771–784.

De Souto, M. C., Jaskowiak, P. A., and Costa, I. G. (2015). Impact of missing data imputation methods on gene expression clustering and classification. *BMC Bioinformatics*, 16:1–9.

Dennis, B. and Ponciano, J. M. (2014). Density-dependent state-space model for population-abundance data with unequal time intervals. *Ecology*, 95(8):2069–2076.

Dhar, V. (2013). Data science and prediction. *Communications of the ACM*, 56(12):64–73.

Dickman, C. R., Haythornthwaite, A. S., McNaught, G. H., Mahon, P. S., Tamayo, B., and Letnic, M. (2001). Population dynamics of three species of dasyurid marsupials in arid central Australia: a 10-year study. *Wildlife Research*, 28(5):493–506.

Dietze, M. C. (2017). *Ecological Forecasting*. Princeton University Press.

Dorazio, R. M. and Karanth, K. U. (2017). A hierarchical model for estimating the spatial distribution and abundance of animals detected by continuous-time recorders. *PloS One*, 12(5):e0176966.

Douma, J. C. and Weedon, J. T. (2019). Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. *Methods in Ecology and Evolution*, 10(9):1412–1430.

Eastoe, E. F. and Tawn, J. A. (2009). Modelling non-stationary extremes with application to surface level ozone. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 58(1):25–45.

Greenville, A. C., Burns, E., Dickman, C. R., Keith, D. A., Lindenmayer, D. B., Morgan, J. W., Heinze, D., Mansergh, I., Gillespie, G. R., Einoder, L., et al. (2018). Biodiversity responds to increasing climatic extremes in a biome-specific manner. *Science of the Total Environment*, 634:382–393.

Greenville, A. C., Wardle, G. M., and Dickman, C. R. (2017). Desert mammal populations are limited by introduced predators rather than future climate change. *Royal Society Open Science*, 4(11):170384.

Greenville, A. C., Wardle, G. M., Nguyen, V., and Dickman, C. R. (2016). Population dynamics of desert mammals: similarities and contrasts within a multispecies assemblage. *Ecosphere*, 7(5):e01343.

Hampton, S. E., Holmes, E. E., Scheef, L. P., Scheuerell, M. D., Katz, S. L., Pendleton, D. E., and Ward, E. J. (2013). Quantifying effects of abiotic and biotic drivers on community dynamics with multivariate autoregressive (MAR) models. *Ecology*, 94(12):2663–2669.

Hebblewhite, M. and Haydon, D. T. (2010). Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550):2303–2312.

Hossie, T. J., Gobin, J., and Murray, D. L. (2021). Confronting missing ecological data in the age of pandemic lockdown. *Frontiers in Ecology and Evolution*, 9:669477.

Hostetler, J. A. and Chandler, R. B. (2015). Improved state-space models for inference about spatial and temporal variation in abundance from count data. *Ecology*, 96(6):1713–1723.

Humbert, J. Y., Scott Mills, L., Horne, J. S., and Dennis, B. (2009). A better way to estimate population trends. *Oikos*, 118(12):1940–1946.

Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., and Shahabi, C. (2014). Big data and its technical challenges. *Communications of the ACM*, 57(7):86–94.

Karunarathna, K., Wells, K., and Clark, N. J. (2024). Modelling nonlinear responses of a desert rodent species to environmental change with hierarchical dynamic generalized additive models. *Ecological Modelling*, 490:110648.

Khan, M. A. (2024). A Comparative Study on Imputation Techniques: Introducing a Transformer Model for Robust and Efficient Handling of Missing EEG Amplitude Data. *Bioengineering*, 11(8):740.

Kuebbing, S. E., Reimer, A. P., Rosenthal, S. A., Feinberg, G., Leiserowitz, A., Lau, J. A., and Bradford, M. A. (2018). Long-term research in ecology and evolution: A survey of challenges and opportunities. *Ecological Monographs*, 88(2):245–258.

Lahoz-Monfort, J. J. and Magrath, M. J. (2021). A comprehensive overview of technologies for species and habitat monitoring and conservation. *BioScience*, 71(10):1038–1062.

Letnic, M. and Dickman, C. R. (2006). Boom means bust: interactions between the El Niño/Southern Oscillation (ENSO), rainfall and the processes threatening mammal species in arid Australia. *Biodiversity & Conservation*, 15:3847–3880.

Letnic, M. and Dickman, C. R. (2010). Resource pulses and mammalian dynamics: conceptual models for hummock grasslands and other Australian desert habitats. *Biological Reviews*, 85(3):501–521.

Letnic, M., Story, P., Story, G., Field, J., Brown, O., and Dickman, C. R. (2011). Resource pulses, switching trophic control, and the dynamics of small mammal assemblages in arid Australia. *Journal of Mammalogy*, 92(6):1210–1222.

Lewis, A. S., Rollinson, C. R., Allyn, A. J., Ashander, J., Brodie, S., Brookson, C. B., Collins, E., Dietze, M. C., Gallinat, A. S., Juvigny-Khenafou, N., et al. (2023). The power of forecasts to advance ecological theory. *Methods in Ecology and Evolution*, 14(3):746–756.

Lewis, A. S., Woelmer, W. M., Wander, H. L., Howard, D. W., Smith, J. W., McClure, R. P., Lofton, M. E., Hammond, N. W., Corrigan, R. S., Thomas, R. Q., et al. (2022). Increased adoption of best practices in ecological forecasting enables comparisons of forecastability. *Ecological Applications*, 32(2):e2500.

Lindenmayer, D. B., Likens, G. E., Andersen, A., Bowman, D., Bull, C. M., Burns, E., Dickman, C. R., Hoffmann, A. A., Keith, D. A., Liddell, M. J., et al. (2012). Value of long-term ecological studies. *Austral Ecology*, 37(7):745–757.

Lofton, M. E., Brentrup, J. A., Beck, W. S., Zwart, J. A., Bhattacharya, R., Brighenti, L. S., Burnet, S. H., McCullough, I. M., Steele, B. G., Carey, C. C., et al. (2022). Using near-term forecasts and uncertainty partitioning to inform prediction of oligotrophic lake cyanobacterial density. *Ecological Applications*, page e2590.

Łopucki, R., Kiersztyn, A., Pitucha, G., and Kitowski, I. (2022). Handling missing data in ecological studies: Ignoring gaps in the dataset can distort the inference. *Ecological Modelling*, 468:109964.

Lui, G. C., Li, W. K., Leung, K. M., Lee, J. H., and Jayawardena, A. W. (2007). Modelling algal blooms using vector autoregressive model with exogenous variables and long memory filter. *Ecological Modelling*, 200(1-2):130–138.

Milchram, M., Suarez-Rubio, M., Schröder, A., and Bruckner, A. (2020). Estimating population density of insectivorous bats based on stationary acoustic detectors: A case study. *Ecology and Evolution*, 10(3):1135–1144.

Mokany, K., McCarthy, J. K., Falster, D. S., Gallagher, R. V., Harwood, T. D., Kooyman, R., and Westoby, M. (2022). Patterns and drivers of plant diversity across Australia. *Ecography*, 2022(11):e06426.

Nakagawa, S. and Freckleton, R. P. (2008). Missing inaction: the dangers of ignoring missing data. *Trends in Ecology & Evolution*, 23(11):592–596.

Niako, N., Melgarejo, J. D., Maestre, G. E., and Vatcheva, K. P. (2024). Effects of missing data imputation methods on univariate blood pressure time series data analysis and forecasting with ARIMA and LSTM. *BMC Medical Research Methodology*, 24(1):320.

Nichols, J. D., Koneff, M. D., Heglund, P. J., Knutson, M. G., Seamans, M. E., Lyons, J. E., Morton, J. M., Jones, M. T., Boomer, G. S., and Williams, B. K. (2011). Climate change, uncertainty, and natural resource management. *The Journal of Wildlife Management*, 75(1):6–18.

Provost, F. and Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big data*, 1(1):51–59.

Raja, P., Sasirekha, K., and Thangavel, K. (2020). A novel fuzzy rough clustering parameter-based missing value imputation. *Neural Computing and Applications*, 32(14):10033–10050.

Sa'adi, Z., Yusop, Z., Alias, N. E., Chow, M. F., Muhammad, M. K. I., Ramli, M. W. A., Iqbal, Z., Shiru, M. S., Rohmat, F. I. W., Mohamad, N. A., et al. (2023). Evaluating Imputation Methods for rainfall data under high variability in Johor River Basin, Malaysia. *Applied Computing and Geosciences*, 20:100145.

Simonis, J. L., White, E. P., and Ernest, S. M. (2021). Evaluating probabilistic ecological forecasts. *Ecology*, 102(8):e03431.

Stock, J. H. and Watson, M. W. (2001). Vector autoregressions. *Journal of Economic perspectives*, 15(4):101–115.

Terletzky, P. A. and Ramsey, R. D. (2016). Comparison of three techniques to identify and count individual animals in aerial imagery. *Journal of Signal and Information Processing*, 7(3):123–135.

Thomas, L. (1996). Monitoring long-term population change: why are there so many analysis methods? *Ecology*, 77(1):49–58.

Thomas, R. Q., Boettiger, C., Carey, C. C., Dietze, M. C., Johnson, L. R., Kenney, M. A., McLachlan, J. S., Peters, J. A., Sokol, E. R., Weltzin, J. F., et al. (2023). The NEON ecological forecasting challenge. *Frontiers in Ecology and the Environment*, 21(3):112–113.

Trancoso, R., McGloin, R., Putland, D., Syktus, J., Ahrens, D., Chapman, S., Owens, D., Toombs, N., Eccles, R., and Zhang, H. (2024). Queensland Future Climate Services - downscaled CMIP6 climate projections for Queensland's regions and locations. Accessed from https://www.longpaddock.qld.gov.au/qld-future-climate/data-info/tern-cmip6/.

Van Buuren, S. (2018). *Flexible imputation of missing data*. CRC press.

Ward, E. J., Holmes, E. E., Thorson, J. T., and Collen, B. (2014). Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. *Oikos*, 123(6):652–661.

Warner, R. and Sloan, R. H. (2021). Making artificial intelligence transparent: Fairness and the problem of proxy variables. *Criminal Justice Ethics*, 40(1):23–39.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*. Chapman and Hall/CRC.

Zhang, Z. (2016). Missing data imputation: focusing on single imputation. *Annals of Translational Medicine*, 4(1):9.

Appendices

1 Appendix 1 (Chapter 2)

TABLE A1.1. Corrected Akaike Information Criterion (AICc) values of the seven variants of the Original Model for the five study species (Fig. 2.2). Single-population MARSS models were fitted for three small mammal species (*Pseudomys hermannsburgensis*, *Notomys alexis*, and *Dasycercus blythi*) and nine independent sub-population MARSS models were fitted for two small dasyurid marsupials (*Sminthopsis youngsoni* and *Ningaui ridei*).

Species	Unconstrained **Q** Constrained **B, R**	Unconstrained **R** Constrained **B, Q**	Unconstrained **B** Constrained **Q, R**	Unconstrained **Q, R** Constrained **B**	Unconstrained **Q, B** Constrained **R**	Unconstrained **R, B** Constrained **Q**	Unconstrained **B, Q, R**
Pseudomys hermannsburgensis	901.44	-1866.02	889.25	-1812.50	902.63	**-1870.79**	-1817.93
Notomys alexis	885.76	-1877.40	877.90	-1850.14	886.43	**-1884.07**	-1833.81
Dasycercus cristicauda	931.75	-1880.23	923.58	-1817.81	932.99	**-1894.78**	-1846.69
Sminthopsis youngsoni	103.77	-570.14	-934.04	-222.41	**-938.53**	-594.77	-537.40
Ningaui ridei	104.62	-574.28	-992.75	-233.32	**-1004.25**	-594.04	-537.31

TABLE A1.2. The results of the selected model variant of single-population MARSS models fitted for three small mammal species (*Pseudomys hermannsburgensis*, *Notomys alexis*, and *Dasycercus blythi*) - Unconstrained $\mathbf{R}$, $\mathbf{B}$ and Constrained $\mathbf{Q}$, but note there is only a single estimate for $\mathbf{Q}$ because this is a single-population MARSS model. a) Estimates of covariates, density dependence, and process error. Density dependence is indicated when $B_{1,1} < 1$. Covariates that are significant in the selected model (95% credible interval excludes zero) are shown in bold. Presence of density dependence and covariate significance in the Original Model (Constrained $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{B}$) are marked with a plus and an asterisk, respectively. b) Population level observation error. Off-diagonal values representing correlated errors are omitted.

(A)

Model covariates	**Estimate**	**Lower CI**	**Upper CI**
P. hermannsburgensis			
Density dependence ($B_{1,1}$)	0.27^{+}	0.02	0.64
Prey ← Mulgara ($B_{1,2}$)	-0.27	-0.67	0.10
Prey → Mulgara ($B_{2,1}$)	0.03	-0.38	0.43
Previous year rainfall	**0.54***	0.28	0.79
Spinifex cover	-0.26	-0.60	0.08
Spinifex seed	**0.36***	0.03	0.66
Process error (q^2)	0.55	0.37	0.83
N. alexis			
Density dependence ($B_{1,1}$)	0.26^{+}	0.01	0.65
Prey ← Mulgara ($B_{1,2}$)	0.10	-0.36	0.48
Prey → Mulgara ($B_{2,1}$)	-0.17	-0.64	0.29
Previous year rainfall	**0.56***	0.34	0.77
Spinifex cover	-0.20	-0.46	0.09
Spinifex seed	**0.41***	0.11	0.68
Process error (q^2)	0.42	0.29	0.63
D. blythi			
Density dependence ($B_{1,1}$)	0.39^{+}	0.03	0.86
Prey ← Rodent ($B_{1,2}$)	-0.06*	-0.51	0.40
Prey → Rodent ($B_{2,1}$)	0.07	-0.38	0.48
Previous year rainfall	**0.32***	0.05	0.62
Spinifex cover	0.23*	-0.08	0.55
Process error (q^2)	0.37	0.08	0.75

(B)

Observation error (r^2)	**Main Camp**	**Carlo**	**Field River South**	**Kunnamuka Swamp East**	**Shitty Site**	**South Site**	**Field River North**	**Tobermorey East**	**Tobermorey West**
P. hermannsburgensis	38.17	39.66	39.23	35.47	35.83	36.94	35.79	38.96	29.50
	(16.81–66.24)	(20.83–64.62)	(19.42–65.96)	(16.55–61.50)	(17.70–59.09)	(16.18–63.42)	(16.75–63.66)	(18.58–66.36)	(10.36–58.44)
N. alexis	37.04	35.21	37.22	39.33	38.60	32.57	38.56	37.97	27.55
	(18.07–63.49)	(17.72–59.88)	(18.99–64.09)	(18.92–67.29)	(19.87–66.76)	(16.05–56.85)	(19.42–66.34)	(17.08–65.03)	(13.10–51.10)
D. blythi	32.86	37.34	39.22	35.93	35.49	38.14	38.74	37.20	27.22
	(15.14–57.32)	(18.70–64.69)	(19.53–65.18)	(16.65–61.05)	(16.41–60.51)	(19.01–65.17)	(19.63–63.22)	(18.45–61.28)	(11.54–49.63)

TABLE A1.3. The results of the selected model variant of nine independent sub-population MARSS models fitted for two small dasyurid marsupials (*Sminthopsis youngsoni* and *Ningaui ridei*) - Unconstrained $\mathbf{Q}$, $\mathbf{B}$ and Constrained $\mathbf{R}$. Density dependence is indicated when $B_{1,1} < 1$. Significant covariates in the selected model (95% credible interval excludes zero) are shown in bold, while those that were significant in the Original Model (Constrained $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{B}$) are marked with asterisks. Off-diagonal process error values representing correlated errors are omitted.

Model covariates	**Rainfall**	**Prey ← Mulgara** ($B_{i,i+1}$)	**Prey → Mulgara** ($B_{i+1,i}$)	**Spinifex cover**	**Density dependence** (B_{ii})	**Process error** (q^2)	**Observation error** (r^2)
S. youngsoni							
Main Camp	-0.04	0.07	-0.13	-0.01	0.43	34.39	0.37
	(-0.78–0.61)	(-2.58–2.6)	(-2.72–2.19)	(-0.03–0.02)	(-2.34–3.13)	(14.57–61.07)	(0.20–0.55)
Carlo	-0.09	-0.20	0.31	< -0.01	-0.20	34.69	
	(-1.05–0.78)	(-3.26–2.78)	(-2.24–2.82)	(-0.03–0.03)	(-2.80–2.49)	(15.93–61.07)	
Field River South	-0.07	0.35	-0.25	< -0.01	0.52	38.58	
	(-0.55–0.37)	(-2.67–3.35)	(-2.88–2.24)	(-0.05–0.05)	(-2.29–3.06)	(20.47–62.83)	
Kunnamuka Swamp	0.09	-0.03	0.07	< -0.01	-0.25	40.82	
East	(-0.39–0.60)	(-3.48–3.28)	(-2.87–3.04)	(-0.04–0.02)	(-3.33–2.84)	(19.29–67.59)	
Shitty Site	0.16	-0.01	0.68	< 0.01	-0.05	33.79	
	(-0.59–1.12)	(-3.29–3.25)	(-2.22–3.2)	(-0.04–0.05)	(-2.86–2.71)	(16.25–61.60)	
South Site	-0.2	-0.64	-0.26	0.02	-0.07	38.92	
	(-0.89–0.44)	(-4.12–2.92)	(-2.22–1.54)	(-0.04–0.09)	(-2.4–2.33)	(19.15–66.06)	
Field River North	0.13**	-0.13	0.20	< -0.01	0.47	36.20	
	(-0.55–0.85)	(-3.33–2.84)	(-2.94–3.30)	(-0.05–0.04)	(-2.65–3.28)	(17.19–61.77)	
Tobermorey East	-0.02	0.38	-0.27	< 0.01	0.11	38.00	
	(-0.75–0.54)	(-2.06–2.82)	(-3.06–2.53)	(-0.03–0.05)	(-2.49–2.65)	(17.76–66.28)	
Tobermorey West	-0.09	0.05	-0.18	< 0.01**	0.38	37.09	
	(-1.11–0.72)	(-2.71–3.09)	(-3.09–2.81)	(-0.03–0.05)	(-2.51–3.65)	(18.28–63.83)	
N. ridei							
Main Camp	-0.04**	0.36	-0.01	< -0.01**	0.58	33.76	0.36
	(-0.71–0.62)	(-1.64–2.43)	(-2.88–2.86)	(-0.03–0.02)	(-2.06–3.24)	(15.38–59.23)	(0.17–0.53)
Carlo	-0.01	0.01	-0.31	< -0.01	-0.10	37.10	
	(-0.84–0.84)	(-2.80–2.76)	(-3.41–2.83)	(-0.03–0.02)	(-2.86–2.81)	(18.36–63.41)	
Field River South	0.25**	0.04	-0.09	< -0.01	-0.10	37.94	
	(-0.22–0.68)	(-2.93–2.96)	(-2.50–2.47)	(-0.05–0.04)	(-2.47–2.46)	(18.02–64.29)	
Kunnamuka Swamp	-0.05	1.18	-0.45	< 0.01	-0.13	35.45	
East	(-0.83–0.76)	(-2.45–4.56)	(-2.51–1.79)	(-0.04–0.05)	(-2.68–2.33)	(18.42-60.48)	
Shitty Site	0.02**	0.10	0.41	< 0.01	0.44	37.22	
	(-0.80–0.80)	(-3.19–3.52)	(-2.80–3.27)	(-0.03–0.04)	(-2.56–3.35)	(16.69–66.38)	
South Site	0.17	0.28	0.18	-0.02	0.15	34.97	
	(-0.43–0.74)	(-3.09–3.83)	(-2.28–2.67)	(-0.07–0.02)	(-2.45–2.66)	(16.70–59.53)	
Field River North	0.03	-0.09	0.36	< 0.01	0.34	38.07	
	(-0.70–0.67)	(-3.43–3.33)	(-2.62–3.21)	(-0.03–0.04)	(-2.50–3.18)	(17.75–66.01)	
Tobermorey East	-0.13	0.76	0.12	< -0.01	-0.38	36.65	
	(-1.11–0.64)	(-1.95–3.41)	(-2.56–2.89)	(-0.05–0.03)	(-3.03–2.69)	(16.71–65.24)	
Tobermorey West	-0.32	0.09	-0.18	< -0.01	0.12	35.55	
	(-1.24–0.47)	(-2.93–2.81)	(-2.44–2.06)	(-0.05–0.05)	(-2.24–2.43)	(15.18–61.34)	

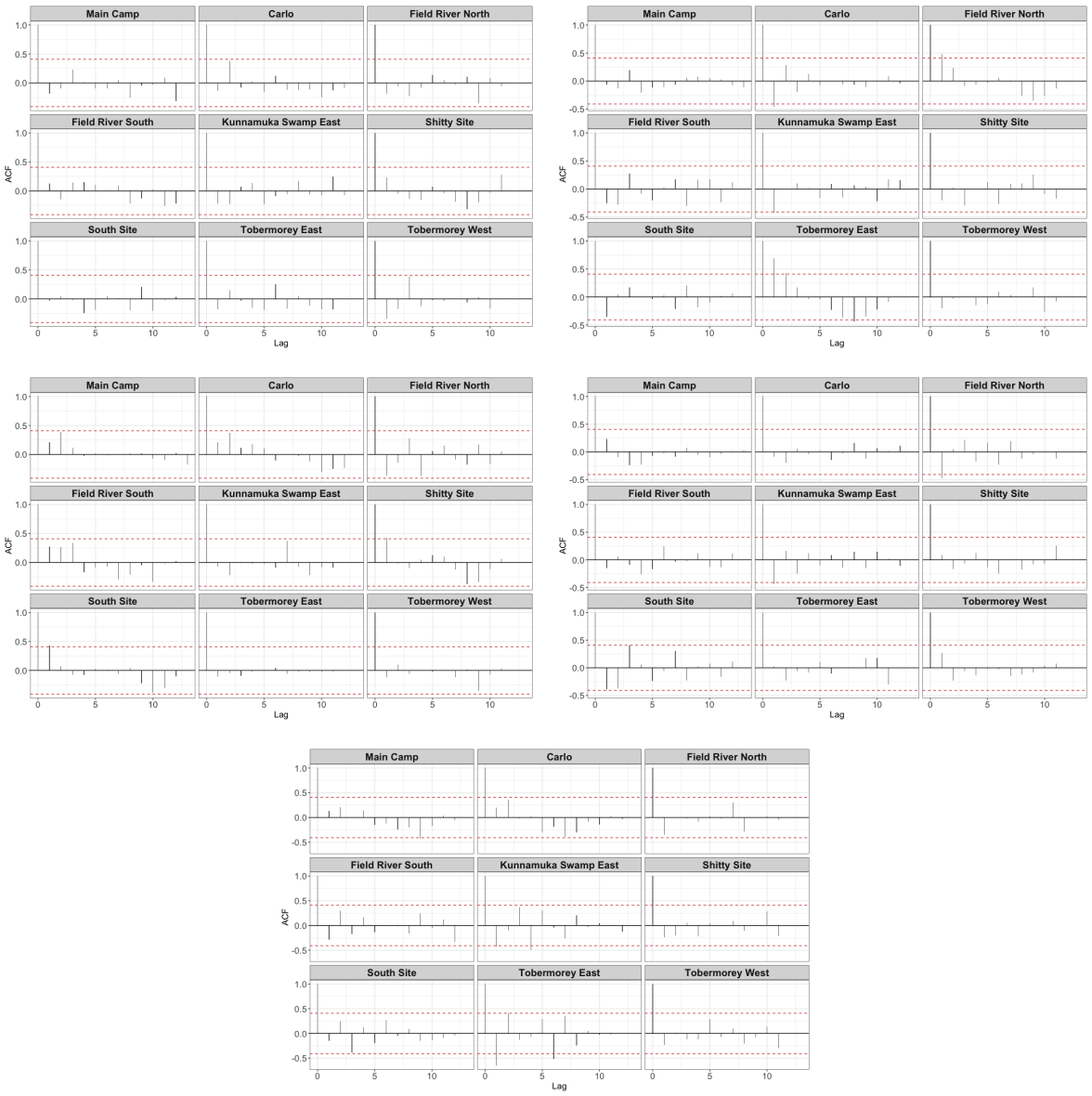

FIGURE A1.1. Autocorrelation function (ACF) plots of residuals from the Original Model for the five study species; (A) *Pseudomys hermannsburgensis*, (B) *Notomys alexis*, (C) *Dasycercus blythi*, (D) *Sminthopsis youngsoni*, (E) *Ningaui ridei* across nine sites (Main Camp, Carlo, Field River North, Field River South, Kunnamuka Swamp East, Shitty Site, South Site, Tobermorey East, Tobermorey West). The horizontal red dashed lines or bands indicate the 95% confidence interval.

2 Appendix 2 (Chapter 4)

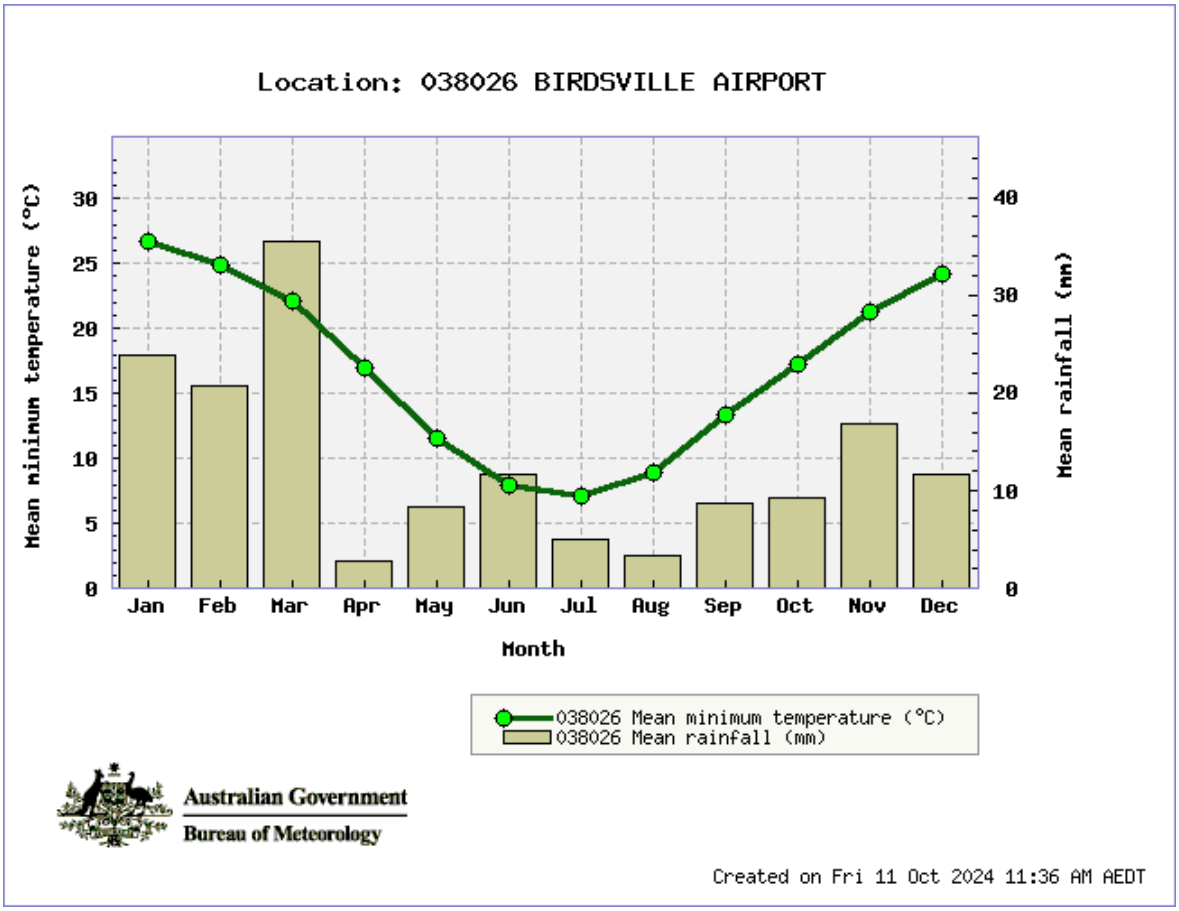

FIGURE A2.1. Mean Minimum Temperature (°C) and Mean Rainfall (mm) across the years 2000 - 2024 at Birdsville Airport weather station.
Source: Bureau of Meteorology, Australian Government

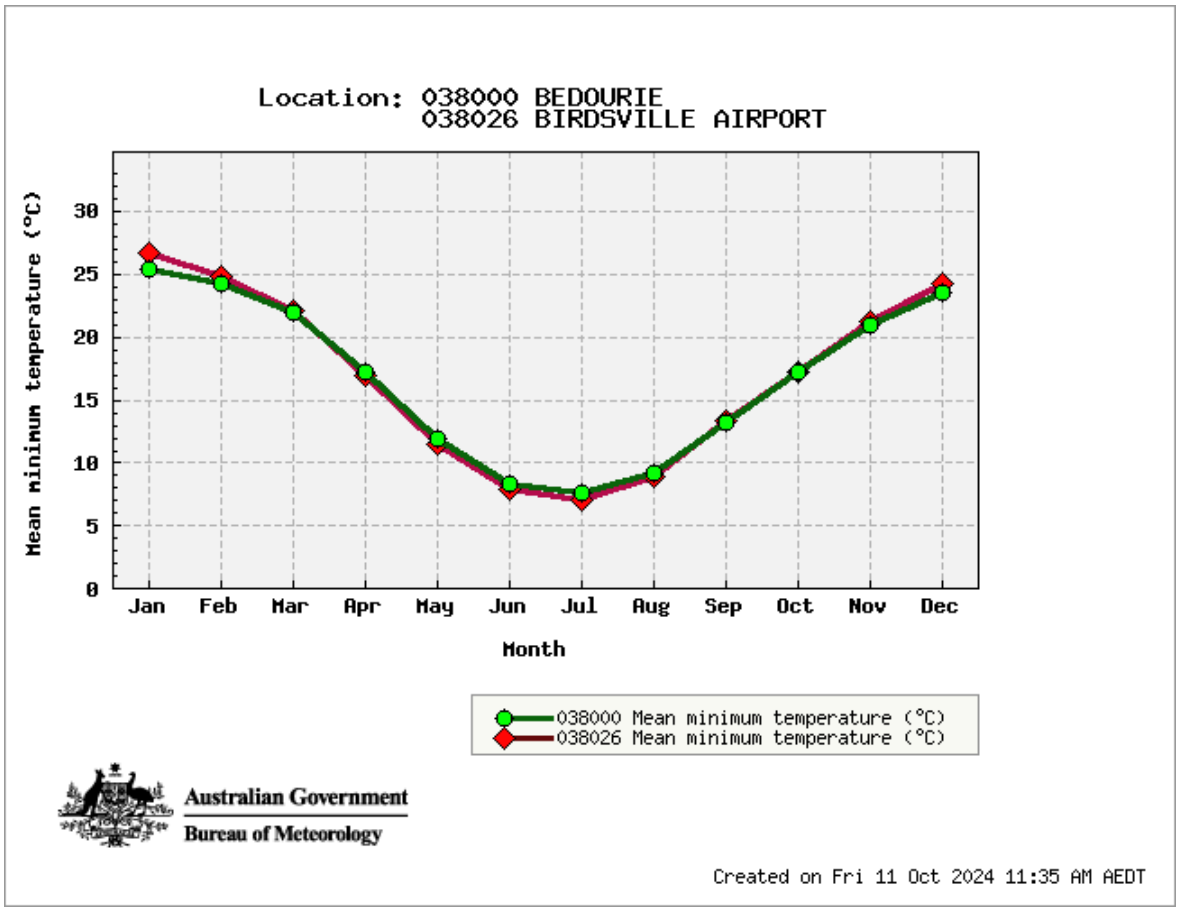

FIGURE A2.2. Mean Minimum Temperatures (°C) of stations Bedourie (1998 - 2024) and Birdsville Airport (2000 - 2024).
Source: Bureau of Meteorology, Australian Government

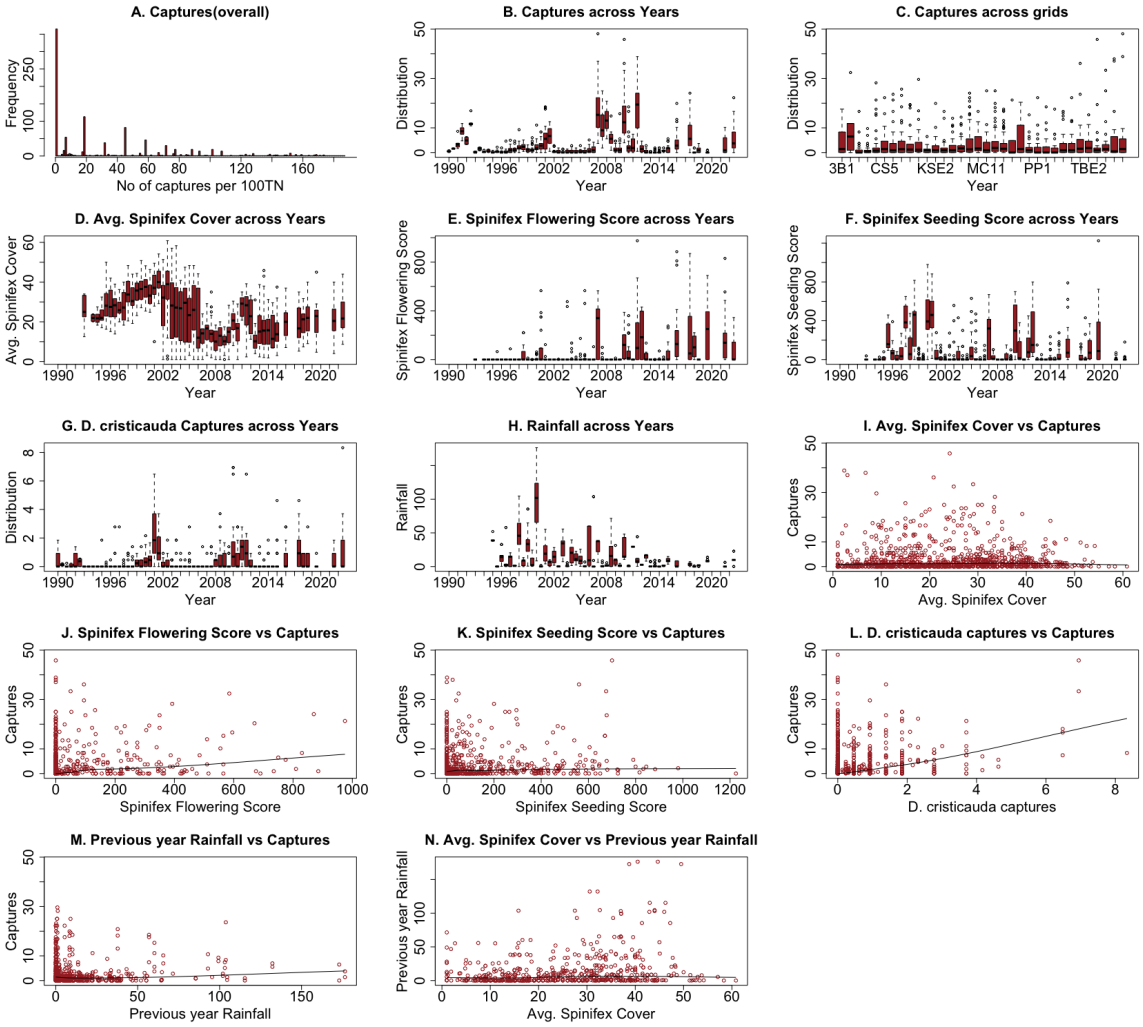

FIGURE A2.3. Distribution of capture rates for *Pseudomys hermannsburgensis* over years and grids. (A) Frequencies of each captures per 100TN. (B) Overall distribution pattern of capture rates across years. (C) Overall distribution pattern of capture rates across grids. (D) Distribution pattern of Avg. Spinifex Cover across years. (E) Distribution pattern of Spinifex Flowering Score across years. (F) Distribution pattern of Spinifex Seeding Score across years. (G) Distribution pattern of *Dasycercus blythi* capture rates across years. (H) Distribution pattern of Total Rainfall across years. (I) Scatterplot of captures vs Avg. Spinifex Cover. (J) Scatterplot of captures vs Spinifex Flowering Score. (K) Scatterplot of captures vs Spinifex Seeding Score. (L) Scatterplot of captures vs *Dasycercus blythi* captures. (M) Scatterplot of Previous year rainfall vs captures. (N) Scatterplot of Avg. Spinifex Cover vs Previous year rainfall.

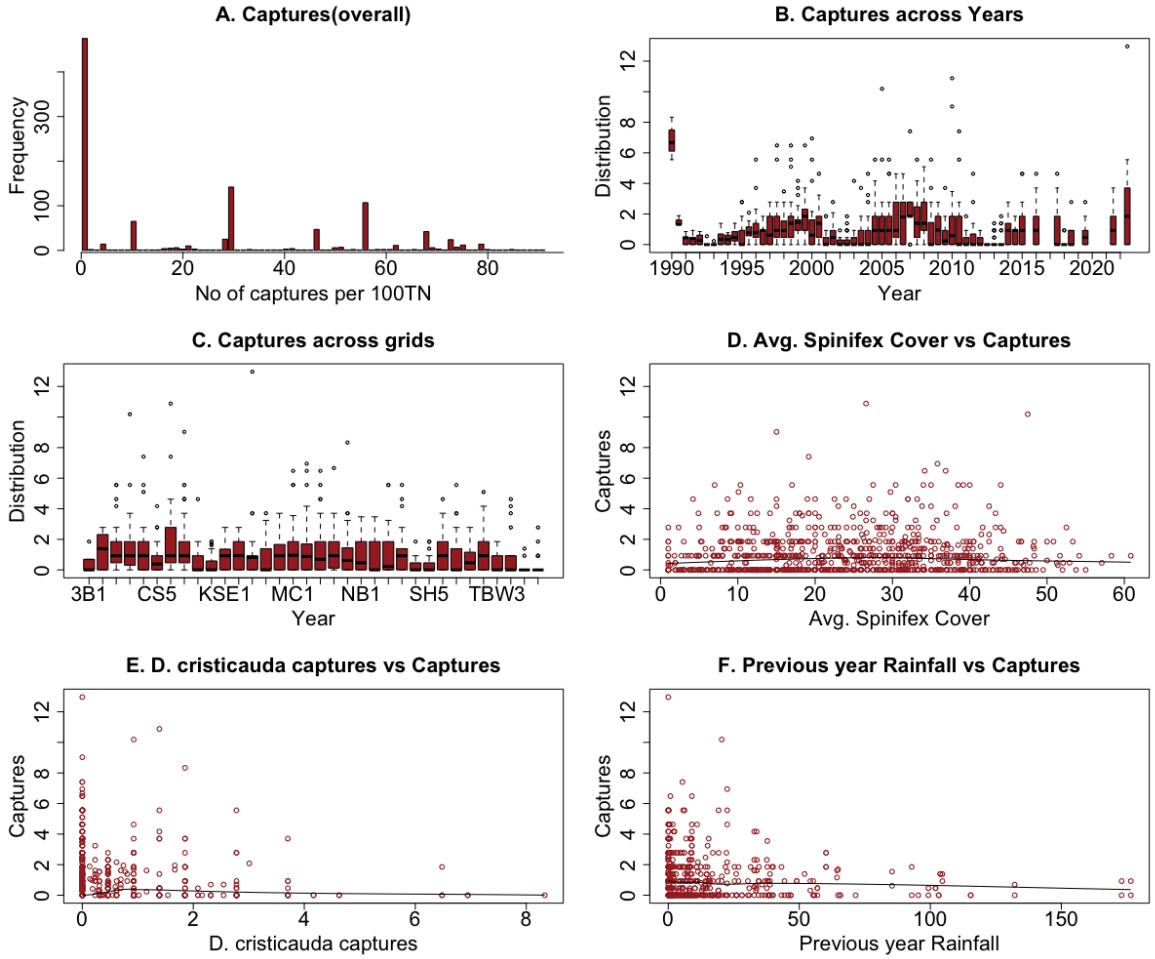

FIGURE A2.4. Distribution of capture rates for *Sminthopsis youngsoni* over years and grids. (A) Frequencies of each captures per 100TN. (B) Overall distribution pattern of capture rates across years. (C) Overall distribution pattern of capture rates across grids. (D) Scatterplot of captures vs Avg. Spinifex Cover. (E) Scatterplot of captures vs *Dasycercus blythi* captures. (F) Scatterplot of Previous year rainfall vs captures.

Model Specifications : Details of *GAM-AR & GAM-VAR*

Model for ***GAM-AR & GAM-VAR*** is defined as follows (similar to ***GAM-GP***):

$$y_{it} \sim \text{Gamma}(\mu_{it}, \emptyset)$$

$$\begin{aligned}\log(\mu_{it}) =& \alpha_{\text{grid}[i]} + S_1(\textbf{\textit{D. blythi}}) + S_2(Spinifex\_Fl\_Score) + \\ & S_3(Spinifex\_Seed\_Score) + S_4(Spinifex\_Avgcover, Rainfall_{(t-2)}) + Z_{it}\end{aligned}$$

While all the parameter specifications are as per specified in Section 4.2.5 of Chapter 4, the latent trend component (Z_{it}) differs as follows:

- ***GAM-AR*** sets a single latent time series shared across grids:

  $$Z_{i,t} \sim \text{Normal}(\text{ar1}_i \cdot Z_{i,t-1}, \sigma_i^2)$$

  where,

  $i = 1, 2, \ldots, 34$

  $ar1_i \sim \text{Normal}(0.5, 0.25)$

  $\sigma_i^2 \sim \text{Exponential}(5)$

- ***GAM-VAR*** which assumes uncorrelated error sets independent trends for each grids as a vector of multiple latent processes:

  $$\mathbf{Z_t} \sim \text{Normal}(A \cdot \mathbf{Z_{t-1}}, \mathbf{\Sigma})$$

  where,

  $A = N_i$ x N_i matrix of autoregressive coefficients where $i = 1, 2, \ldots, 34$

  capturing within grid dependance through diagonals

  $\Sigma \sim \text{Exponential}(5)$

TABLE A2.1. The results of the final model (***GAM-GP***) for the rodent *Pseudomys hermannsburgensis* & the small insectivorous dasyurid *Sminthopsis youngsoni*. Approximate significance of generalized additive smooth covariates (p-values < 0.05) are shown in bold.

Model covariates	**p-value**	
	Training set 1	**Training set 2**
P. hermannsburgensis		
Spinifex Flowering Score	0.61	0.58
Spinifex Seeding Score	0.33	0.53
Mulgara	0.10	**0.01**
Avg. Spinifex Cover $*$ previous year rainfall	1.00	0.99
S. youngsoni		
Mulgara	0.47	0.41
Avg. Spinifex Cover $*$ previous year rainfall	1.00	1.00

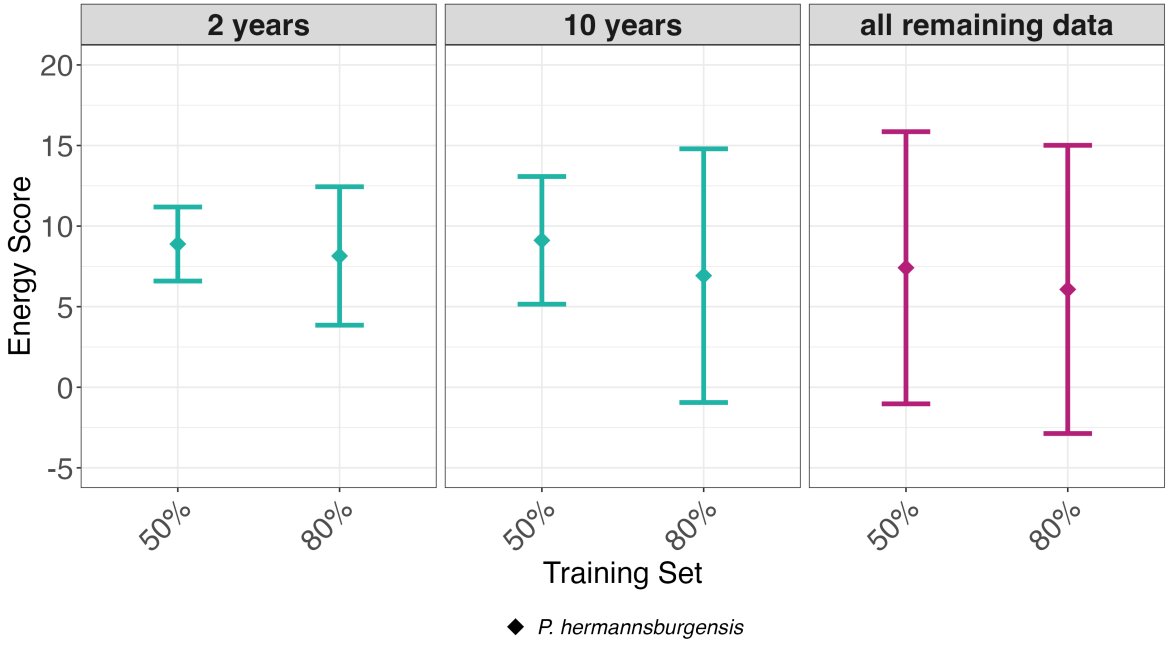

FIGURE A2.5. Energy Scores for the final model of *Pseudomys hermannsburgensis* (***GAM-GP***) with **'Spinifex' variables replaced with 'Overall vegetation' variables** fitted for Training set 1 (50%) and Training set 2 (80%) and evaluated for three forecast horizons (2 years, 10 years and all remaining data). Diamond points represent mean scores across the forecasted seasons within each forecast horizon, with error bars indicating the 95% confidence interval. The 'all remaining data' horizon is represented in distinct colours to indicate differences in the number of data points, making direct comparisons with the other horizons inappropriate.

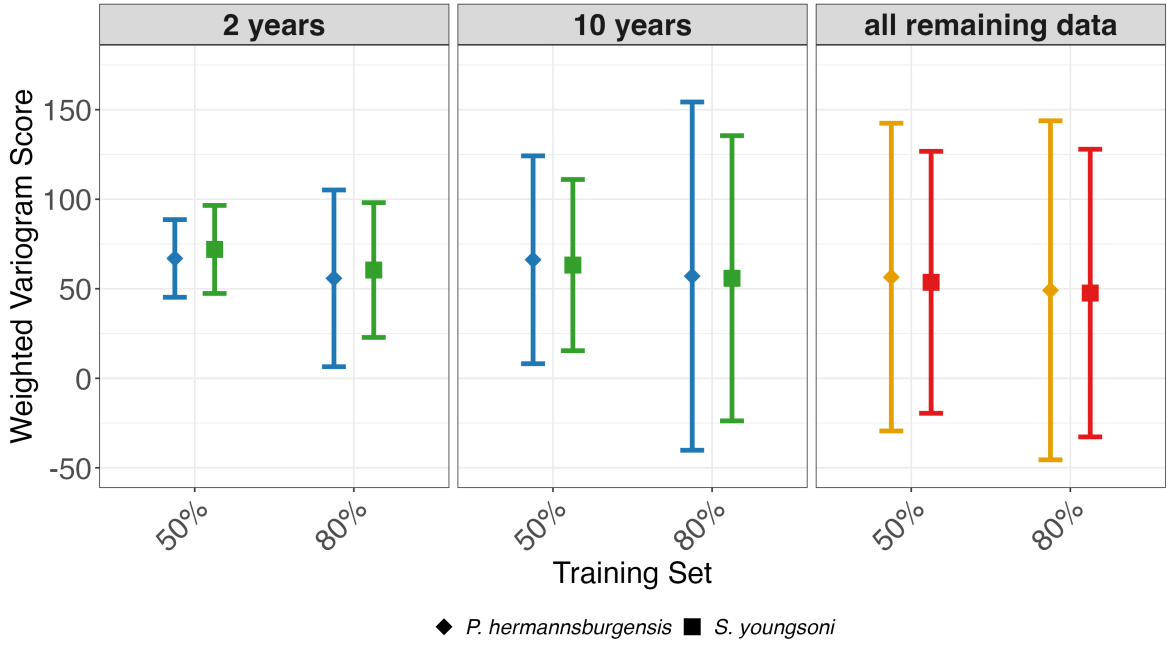

FIGURE A2.6. Square root of the Weighted Variogram Scores for the final model of *Pseudomys hermannsburgensis* (***GAM-GP***) and *Sminthopsis youngsoni* (***GAM-GP(only cover)***) fitted for Training set 1 (50%) and Training set 2 (80%) and evaluated for three forecast horizons (2 years, 10 years and all remaining data). Diamond and Square points represent mean scores across the forecasted seasons within each forecast horizon, with error bars indicating the 95% confidence interval. The colour differences within the 2 year and 10 year forecast horizons (blue and green) reflect variations in covariate specifications between *Pseudomys hermannsburgensis* and *Sminthopsis youngsoni*. The 'all remaining data' horizon is represented in distinct colours to indicate differences in the number of data points, making direct comparisons with the other horizons inappropriate.

3 Appendix 3 (Chapter 5)

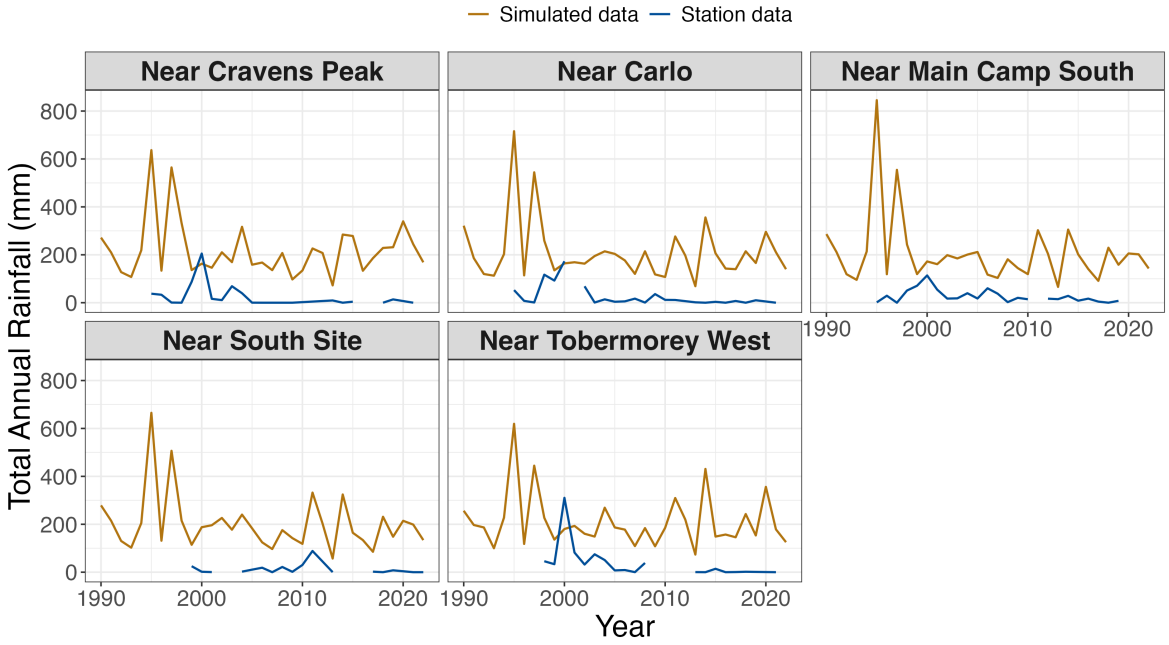

FIGURE A3.1. Comparison of total annual rainfall between weather station data (blue line) and ACCESS-CM2_oc simulated rainfall data (brown line) across five sites, representing all spatial directions within the study area. Rainfall was analyzed at the site level, assuming that both grids within a site experienced the same rainfall (Chapter 4). The selection of five representative sites aimed to determine whether differences in rainfall scale were uniform across the study area or specific to certain sites. Monthly rainfall data from weather stations were unavailable for all 12 months as sampling was hindered due to constraints such as adverse climatic conditions (e.g. road closures), hence, total annual rainfall was used to ensure consistency in the comparison. Simulated rainfall data, derived at a 10 km resolution, were matched to the nearest longitude and latitude points corresponding to the selected sites, which is why the dataset is labeled 'near'.

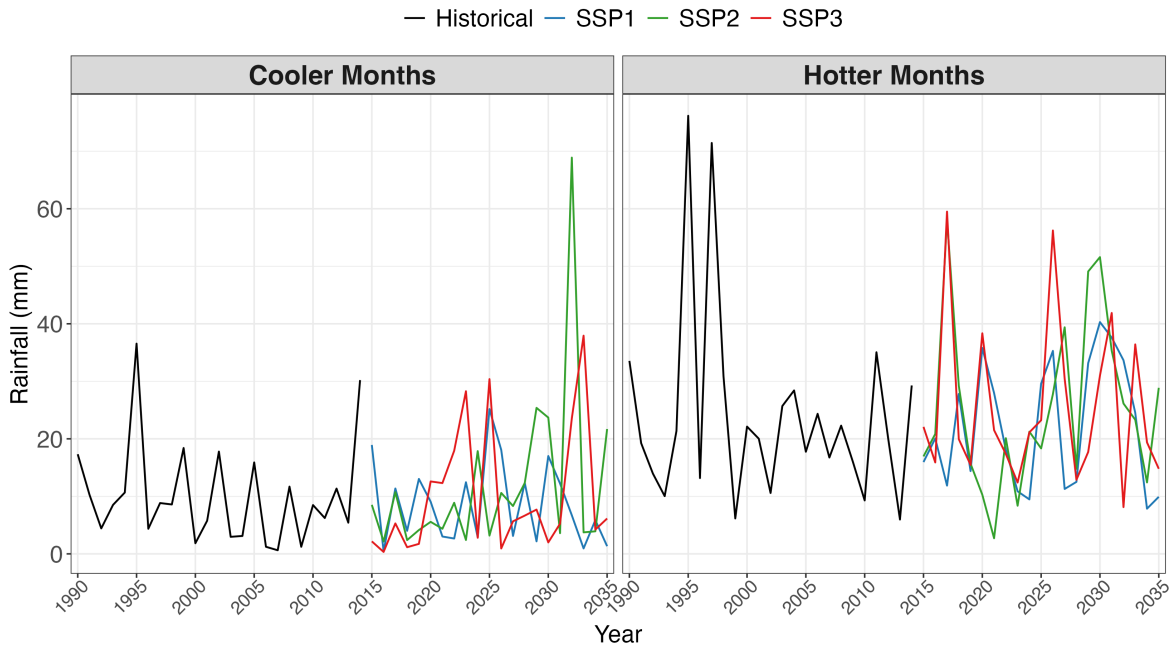

FIGURE A3.2. Historical (1990 – 2014) and 12 year projected rainfall data based on the three Shared Socioeconomic Pathways (SSPs) from the ACCESS-CM2_oc Regional Climate Model (RCM). The black line represents the historical simulated rainfall data, while the blue, green, and red lines correspond to the projected rainfall under SSP1, SSP2, and SSP3 scenarios, respectively. The data is separated for cooler and hotter months. The nearest location to the Carlo site has been selected for reference.