26th International Congress on Modelling and Simulation, Adelaide, SA, Australia, 30 Nov–4 Dec 2025
mssanz.org.au/modsim2025

From data gaps to climate projections: A framework for forecasting species population abundance

V. Gunadasa [a], **A. Greenville** [a] **and G. M. Wardle** [a]

[a] *School of Life and Environmental Sciences, The University of Sydney, 2006, Australia*
Email: vihanga.gunadasa@sydney.edu.au

Abstract: Ecological forecasting has advanced rapidly in recent years, yet meeting the predictive needs of decision-makers requires a major expansion in operationalized forecasts. Forecasting species count data is especially challenging due to ecosystem complexity, limited data availability, and the influence of unpredictable environmental drivers.

Drawing on insights from long-term small mammal monitoring in arid Australia, we propose a five-step framework to improve forecast reliability: (1) align ecological assumptions with statistical models, (2) handle missing data appropriately, (3) select models balancing complexity and interpretability, (4) evaluate performance across training lengths and forecast horizons, and (5) incorporate projected environmental covariates. Using over two decades of population data from spatially separated sites on species with contrasting dynamics, we found that tailoring statistical models to species-specific dynamics, rather than applying a one-size-fits-all approach, markedly improved both interpretability and reliability. We show that common practices such as zero-filling missing values distort population trends, whereas multivariate imputation by chained equations (MICE) preserved ecological signals and enhanced forecast performance across a range of data-gap patterns. Multivariate generalized additive models (MVGAMs) effectively captured nonlinear species–environment relationships, site-level effects, and temporal dynamics, while comparisons of Gaussian process and vector autoregressive (VAR) latent trends highlight trade-offs between accuracy and computational feasibility in data-limited settings. Bayesian approaches provide a powerful framework for probabilistic forecasting, revealing that even relatively short time series can capture boom–bust population dynamics characteristic of arid ecosystems. However, forecast uncertainty increased with horizon length and species variability, underscoring the importance of identifying the appropriate training data requirements for specified forecast horizons. Beyond the data science refinements, we show that the ecological relevance of covariates can shift with horizon length, as for example, proxies such as dominant vegetation cover improved model stability.

We applied these findings to forecast species abundance using rainfall projections under three greenhouse gas emissions management pathways: sustainability-focused with reduced global emissions, middle-of-the-road with emissions remaining steady, and regional rivalry with rising emissions and temperatures, generating forecasts four and twelve years beyond 2022. Results show that rainfall strongly influences abundance, but species-specific dynamics and data availability are equally critical. Under sustainable climate mitigation, the focal species exhibiting variable dynamics, Pseudomys hermannsburgensis, was projected to maintain relatively stable trends, while the more stable species, Sminthopsis youngsoni, exhibited only minor fluctuations, with longer-term projections indicating lower sensitivity to climate variability.

With current management efforts, the species with boom-bust dynamics, P. hermannsburgensis, showed moderate fluctuations and sharper declines where data were limited, though long-term projections remain close to historical patterns. The stable species, S. youngsoni, remained steady highlighting resilience, with slight declines at data-limited sites. Under high emissions and temperature, the species with varying dynamics plateaued or declined more sharply depending on data availability, while the stable species maintained relatively consistent trends but showed gradual reductions over time. Across scenarios, species with inherently stable dynamics were more resistant to climate variability, while those with variable dynamics were more sensitive to poor environmental management. Forecasts for both focal species over the projected intervals did not suggest stochastic extinction but forecast uncertainty did increase with horizon length suggesting caution for longer-term decision frameworks.

Our results emphasize that reliable ecological forecasting depends not only on climate trajectories but also on model design, data quality, and the temporal scale of management. By aligning models with ecological processes, incorporating robust handling of missing data, and integrating projected environmental drivers, we provide a framework for delivering forecasts that can more effectively guide proactive, evidence-based conservation in variable environments.

Keywords: *Ecological forecasting, spatio-temporal data, multivariate imputation, MVGAM, climate projections*

