

But I can't preregister my research: Improving the reproducibility and transparency of ecology and conservation with adaptive preregistration for model-based research

Elliot Gould Chris Jones Jian Yen Hannah Fraser
Henry Wootton Lyndsey Vivian Megan Good
David Duncan Libby Rumpff Fiona Fidler

1. Preregistration is an open-science practice adopted in scientific disciplines that have begun to confront the 'reproducibility crisis', which aims to improve research transparency and mitigate questionable research practices within a study. There has been little uptake of preregistration in ecology and conservation, and many modellers and researchers engaged in non-hypothesis testing research both within and outside of ecology have eschewed pre-registration on the grounds that existing templates are irrelevant because they are focused on null-hypothesis significance testing.
2. We advocate that preregistration can and should be used in model-based research in ecology and conservation, but that preregistration should reflect the norms and practice of the research context in which it is applied, in order to adequately restrict researcher degrees of freedom that are unique to a domain-specific methodology. We took a user-centred design approach to the task of translating preregistration into model-based research contexts for ecology and conservation.
3. We identified a generalised 'modelling workflow' that captures critical analysis decisions in the process of the model development cycle. The workflow informed the content and structure of a preregistration template to be used in model-based research. The iterative nature of model development was identified as a key barrier to implementing preregistration for modelling research because it conflicts with the current format and linear process of preregistration. To better align the internal logic of preregistration with the iterative and non-linear process of ecological modelling, we propose a methodology for implementing 'adaptive preregistration' leveraging git and GitHub. We

tested and evaluated the preregistration template and methodology using a case- study of environmental flows modelling for decision-support in Victoria, Australia.

4. This research provides a template and methodology for implementing adaptive preregistration of ecological models in either pure or applied settings and will facilitate the wider uptake of preregistration in ecology. Although we focus on ecology and conservation in this paper, the template is extensible to alternative fields and methods, and the proposed methodology for implementing adaptive preregistration can be applied in model- based re- search within science more broadly. Modellers need no longer cry “but, I can’t preregister my research.”

Introduction

Science’s replication crisis, and the subsequent recognition of problems with reliability of scientific research findings, has triggered a ‘credibility revolution’ spawning a suite of proposed methodological and policy reforms, including *preregistration* (parker2019?). Preregistration is an open-science practice that aims to distinguish between planned and ad hoc analyses (Parker, Fraser, and Nakagawa 2019) and aims to combat the risk of questionable research practices by requiring researchers to register their methods and analysis plan in a secure and publicly accessible platform prior to collecting and/or analysing data, which cannot be altered after submission (#todo #citation). The purpose of preregistration is to prevent biased research findings by guarding against questionable research practices (such as selective reporting, p-hacking, HARKing, among others #todo #citation) by distinguishing what is a genuine *a priori* planned analysis, and what is not, thereby preventing the opportunistic misuse of researcher degrees of freedom, or those many analytic decisions researchers make during the planning, data collection, data analysis and reporting of results in scientific studies (See Box 1 for background). QRPs have been labelled as one of the primary cause of the replication crisis (Shrout and Rodgers, 2018 in (Fife and Rodgers 2022)) and serve to XYZ bias the overall literature, etc (#todo finish this point).

Despite enthusiastic uptake in some scientific disciplines, such as psychology and medicine (#todo #citation), the use of registered reports and preregistration in ecology and conservation more broadly is new and limited (Desbureaux 2021), particularly for model-based research (#todo Joel Pick preprint?). **Ecologists have brushed preregistration off on the basis that existing templates are irrelevant since they are focussed on null-hypothesis significance testing**, whereas in applied ecology and conservation, we don’t usually explicitly test hypotheses (Betts et al. -3-25 2021), but instead use explanatory models, or increasingly predictive models that generate anticipatory predictions, to inform management decisions (Wood et al. 2020). Model-fitting and evaluation, rather than falsification of hypotheses about natural phenomena, are increasingly prevalent approaches to scientific inquiry in ecology (Connolly et al. 2017)

Preregistration in model-based research has been eschewed on several grounds by modellers and researchers engaged in model-based research both within and outside of ecology. Much of the literature advocating for the use of preregistration assumes a hypothetico-deductive model of scientific inquiry, and the templates used to write preregistrations have been accused of being focussed on null-hypothesis significance testing (e.g. McDermott et al. 2021; Prosperi et al. 2019). Approaches and advocacy for preregistration of MBR across all scientific disciplines are limited, and focussed have focused on the phases of the modelling process and modelling approaches that are more confirmatory (#todo see Ioannidis (2022)). For example, Crüwell and Evans (2019) present a template for preregistration of MBR where the model construction and development phase is explicitly and intentionally excluded from the preregistration process, while much of the discussion around preregistration for models has centred on machine learning (Hildebrandt 2018) or predictive modelling (Hofman et al. 2023) (#todo see [preregistration for modelling] for more refs).

A second argument against preregistration for model-based research is based on the mischaracterisation of the model development process as a intrinsically exploratory process (McDermott et al. 2021) , and since only the preregistration of confirmatory research, such as hypothesis testing, is considered necessary (Prosperi et al. 2019), there is no need to pre-register model-based research (#todo #citation). The description of modelling as a wholly exploratory process is a mischaracterisation based on a misunderstanding of sensemaking and reasoning within the modelling process stemming from a lack of clear formal definitions distinguishing confirmatory research from exploratory research (Flournoy et al. 2020; Devezer et al. 2020; Alspaugh et al. 2019). Although modelling includes some tasks and activities that are exploratory in nature, it can be distinguished from wholly *exploratory analysis* in that tends be goal-driven, directed analysis that is less open-ended (Alspaugh et al. 2019). For model-based research, the exploratory/confirmatory divide is an idealised and simplifying dichotomy that, in practice, is more complicated and nuanced and can be considered as a spectrum on which exploratory analysis sits at the analysis is open-ended, high-level and opportunistic end, to goal-driven, low-level and precise analyses at the other (Alspaugh et al. 2019). In practice the divide is often blurred, and researchers will move back and forth between different types of reasoning and inference throughout the modelling process, switching from activity at one end of the spectrum to the other and back again (Alspaugh et al. 2019), with deductions leading to abductions as the analysis process proceeds (Alspaugh et al. 2019; Prosperi et al. 2019), or even occurring together in a hybridized manner (Prosperi et al. 2019).

When critics of preregistration for MBR describe modelling as intrinsically ‘exploratory’, what is seemingly being referred to is the iterative, adaptive and non-linear nature of model-development. In practice, modelling proceeds in an adaptive manner by which models are iteratively developed and refined (Augusiak, Van den Brink, and Grimm 2014; **Babel2019?**; Bennett et al. 2013; Bodner, Fortin, and Molnár 2020). New analyses are often chosen based on the basis of data exploration and previous analyses of the same data (Dwork et al. 2015), For example, modellers may perform exploratory analyses to validate distributional assumptions are met, such as no over-dispersion or zero-inflation in count-data to ensure valid inference when using particular models (Campbell 2021) (#todo also see Alspaugh et al. (2019) for list

of examples). Such practices violate a key tenet of preregistration, that is ‘data-independent’ analytic decision-making (#todo #citation). Consequently, the iterative nature of model development, wherein models are continually refined through analysing and reanalysing a dataset and the model, is considered to be in conflict with the practice of preregistration; by requiring researchers to commit to a single static plan that prevents changes or deviations from the original plan it is seen to be forcing adaptive analysis to be *non-adaptive* through ‘defining the entire analysis protocol ahead of time’ (Dwork et al. 2015; Benning et al. 2019; MacEachern and Van Zandt 2019; McDermott et al. 2021). In short, we can’t write off preregistration for modelling on the basis that it is exploratory, because in practice it employs a range of modes of reasoning from induction, deduction, to abduction, with analysts switching between those types of reasoning or even hybridising them.

What *is* important to consider when considering the application of preregistration within MBR is whether the reliability of MBR may be potentially threatened by the potential for Questionable Research Practices (Box 1). Evidence from ecology more broadly has revealed that the rate of QRPs is similar to those in other disciplines (Fraser et al. 2018) and while there has been little study of QRPs in MBR (except E. Gould (In prep.) #todo preprint doc & cross-ref here), preliminary evidence reveals that modellers do engage in QRPs such as selective reporting, for example, they may omit explored analysis paths from their reported analyses if they did not corroborate their conclusions (Liu, Althoff, and Heer 2020). Furthermore, QRPs are likely to arise in conditions where there is incomplete and opaque reporting practices (Fidler et al. /04/15 2018), and poor model-reporting practices and an overall lack of transparency about the modelling process are endemic in MBR in ecology. Concerns about the state of model reporting in ecological modelling have been raised repeatedly, arguing that current practices are neither transparent or comprehensive (Schmolke et al. /08/01 2010; de Vos, Janssen, and Van Bussel Proceedings of the 2011). Despite a number of attempts to promote thorough and transparent reporting, (e.g. Jakeman, Letcher, and Norton (/05/01 2006); Alexandrov et al. (/03/01 2011); Bodner, Fortin, and Molnár (2020)) as a critical component of *Good Modelling Practice*, a culture of adequate research transparency in modelling is yet to be established and remains limited (de Vos, Janssen, and Van Bussel Proceedings of the 2011; Grimm et al. 2014). The lack of transparency in describing ecological models and reporting the modelling process and results provides significant opportunity for undisclosed researcher degrees of freedom, and hence, the potential for Questionable Research Practices.

Although the purpose of preregistration is predominantly conceptualised in terms of distinguishing exploratory from confirmatory research (e.g. Nosek et al. (2018), Munafò et al. (2017) and #todo insert other refs), we adopt a more general stance towards the purpose of preregistration, i.e. that preregistration may mitigate against questionable research practices by reducing the potential for cognitive biases to cause opportunistic misuse of researcher degrees of freedom. In this paper, we advocate that preregistration can and should be used in model-based research in ecology and conservation, but its particular form should reflect the norms and practices of the research context in which it is applied, in order to adequately restrict researcher degrees of freedom that may uniquely emerge within this domain-specific methodology.

The goal of this study is to rethink preregistration to better suit the domain-specific challenges and methodological particularities of model-based research in ecology and conservation. Preregistration cannot simply be ported over to other disciplines and adopted as-is. Different research methodologies, and their domain-specific application across different domains and fields of study follow different analytical choices and are therefore susceptible to different QRPs. Consequently, domain-specific templates are necessary for guiding preregistration in such a way that restricts researcher degrees of freedom as they uniquely emerge within a domain or field-specific application of a particular research methodology. If we wish to adopt preregistration for ecological modelling, both the content required to be preregistered, as well as the procedure for creating a preregistration must be adapted to suit model-based research as it is undertaken in ecology.

We first aimed to build a preregistration template for ecological modelling. We identified a generalised ‘modelling workflow’ that captured critical analysis decisions in the process of the model development cycle. The workflow informed the content and structure of a preregistration template to be used in model-based research. The iterative nature of model development was identified as a key barrier to implementing preregistration for modelling research because it conflicts with the current format and linear process of preregistration. To better align the internal logic of preregistration with the iterative and non-linear process of ecological modelling, we devised a methodology for implementing an expanded view of preregistration that we term *Adaptive Preregistration* (sensu Srivastava 2018). Adaptive preregistration allows for the iterative cycle of model development to be reflected in the preregistration and better accounts for the inherently adaptive nature of model development than current approaches to preregistration in which researchers are committed to a static, unalterable and deterministic view of preregistration. Finally, to assess the feasibility and practicality of preregistration for MBR, we tested and evaluated the preregistration template and methodology using a case-study of environmental flows modelling for decision-support in Victoria, Australia.

Box 1: How Researcher Degrees of Freedom become Questionable Research Practices in MBR

Modellers are faced with numerous ‘researcher degrees of freedom’ (Wicherts et al. 2016) or subjective decisions during the course of modelling that may substantially impact the chosen analysis strategy, and therefore the research findings and conclusions drawn from that analysis (Hoffmann et al. 2021; Simonsohn, Simmons, and Nelson 2020; Desbureaux 2021; **gouldSameDataDifferent2023?**). The set of all plausible or reasonable potential decisions that a researcher could take from the beginning to end of analysis has been described as the ‘garden of forking paths’ (Gelman and Loken 2013). ‘Many analyst’-style studies, whereby

multiple separate teams of researchers independently analyse the same dataset and same research question, demonstrate that different pathways in this ‘garden of forking paths’ can lead to potentially conflicting conclusions (Young and Stewart 2021; Liu, Althoff, and Heer 2020; **gouldSameDataDifferent2023?**).

Researcher degrees of freedom arise because there is inherent analytic uncertainty in model-based research whether due to incomplete theoretical understanding of underlying ecological processes, or because modelling is an act of simplification and abstraction; there is always some degree of choice about which ecological processes and phenomena should be included in the model, how they should be represented (**Babel2019?**; Getz et al. 2017). Further analytic uncertainty may arise due to uncertainty in data processing or choice about which analysis method is best (Hoffmann et al. 2021).

- *Analytic Uncertainty* (Walker et al. 2003) talks about uncertainty specifically in terms of *model based* decision-making. #todo

When researchers are faced with ambiguous decisions where there is no single decision or strategy acceptable according to scientific standards and consensus, we are likely to select the decision that results in the most ‘favourable’ finding with “convincing self-justification” (Simmons, Nelson, and Simonsohn 2017; Hoffmann et al. 2021). Researchers routinely execute and explore alternative analysis pathways in service of identifying the most favourable (i.e. publishable) result, selectively reporting only one or several preferred results (Liu, Althoff, and Heer 2020; Simmons, Nelson, and Simonsohn 2017). Thus researcher degrees of freedom precipitate ‘questionable’ research practices, or QRPs, when they remain both undisclosed and exploited opportunistically (Bakker et al. 2018). Such ‘questionable’ research practices have been demonstrated to be surprisingly common upon admission among ecologists, with comparable rates to other scientific disciplines (Fraser et al. 2018). Questionable research practices are categorically distinct from fraudulent practices, arising not from consciously nefarious ‘fishing expedition[s]’ (Gelman and Loken 2013), but rather because researchers are motivated actors working within a ‘publish or perish’ culture who are susceptible to cognitive biases and are willing to change their modelling assumptions when the results conflict with desired outcomes, for example, sincerely and self-convincingly believing that the more favourable model is superior (Young and Stewart 2021).

QRPs may increase the probability of false positive results, may artificially inflate effect size estimates, and ultimately reduce the replicability of published findings, while encouraging overconfidence in the precision of results and exacerbating the risk of accepting and propagating false facts (Wicherts et al. 2016; Hoffmann et al. 2021; Roettger 2019; Nissen et al. 2016). Good intentions, and simply being ‘aware’ of the potential for researcher degrees of freedom to lead to questionable research practices cannot sufficiently mitigate their risk (Nosek et al. 2018).

Methods

We approached the task of translating preregistration into model-based research contexts for ecology and conservation with a user-centred design (UCD) framework. UCD is a philosophy for designing processes, services or products that account for user requirements at all stages of the design process so that the resource generated by this approach is effective and usable for specific people in specific contexts (Pavelin et al. 2012). Because it is difficult to predict how users will actually engage with a resource (Pavelin et al. 2012) and a designer’s intention for preregistration may conflict with the user’s purpose and undermine the overarching goal of preregistration (Pu et al. 2019), we heed the advice of Pu et al. (2019), who recommend treating preregistration as a UCD problem so that preregistration can effectively meet its intended objectives, despite UCD being typically used within software engineering domains, rather than academic domains.

We borrow from the methodology proposed by Pavelin et al. (2012) to design, test and refine a preregistration template suitable for ecological modelling, and a methodology for preregistering model-based research (Figure 1). The methods description is structured following the key steps illustrated in Figure 1.

Define Vision & Objectives

Define the goals of the research under development and identify the audience

The first step in a UCD approach is to articulate the goals, objectives and target audience of the resource, consequently we define the intended purpose of the preregistration template and methodology, the community of users, as well as their research contexts and use-cases. Finally, we explicate the aspirations and desirable attributes of the template and methodology. Much like in ‘Good Modelling Practice’, articulating the overarching purpose, intended audience, as well as the objectives and aspirations will guide the development and evaluation of preregistration for model-based research in ecology and conservation, with important implications for the design of the template – from its structure, content, user-interface and broader workflow in which it is designed to be used.

Articulating Intended purpose of preregistration template & methodology

We define the purposes of the template as:

1. *To delimit researcher degrees of freedom* > when conducting model-based research in applied ecology and conservation so as to mitigate the risk of questionable research practices in order to improve the reliability and credibility of this research.
2. *Increase research transparency* > of model-based research in ecology and conservation. *Research transparency* refers to both production transparency — which includes research artefacts

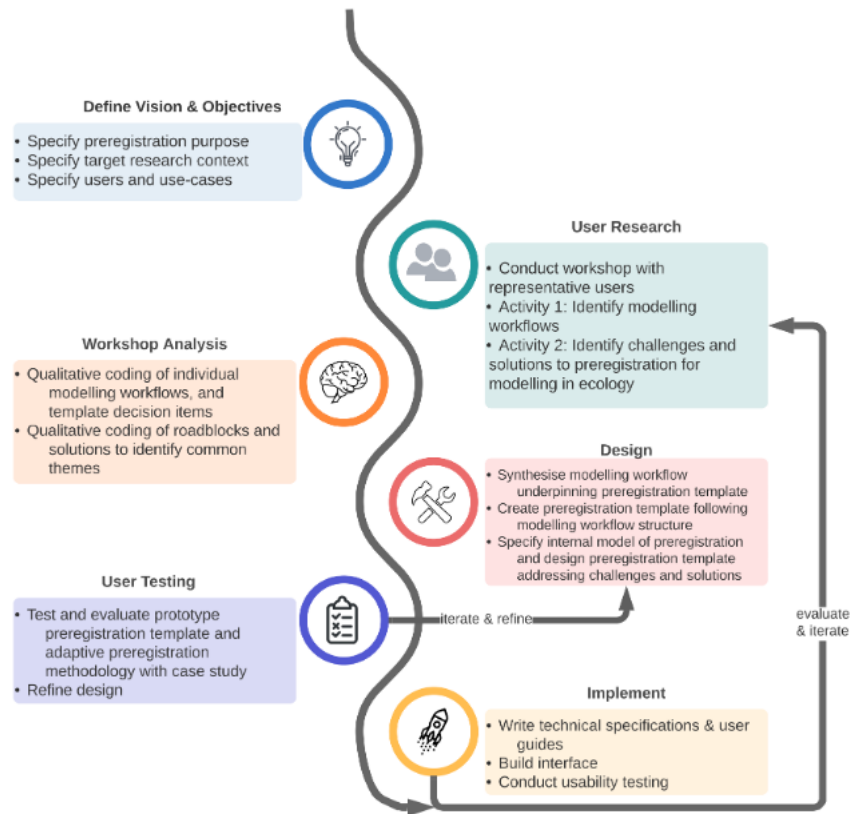


Figure 1: Insert caption. Number each step so i can refer to it in the headings and text below. Make it clear in the figure that we're not doing the build interface and conduct usability testing component of implement, nor the evaluate and iterate. we throw that out to the community of practice to do that part.

like open-access data and materials, or data collection procedures — and analytic transparency — which describes a complete account of how the analytic conclusions are drawn from the data (Lupia and Elman 2014).

Target user community and their research-contexts

We aim for the template and preregistration process developed in this study to be used for model-based research, which we define as any research, empirical or applied, that uses quantitative modelling to answer its research question (Brudvig 2017), where *quantitative modelling* may be defined as any modelling that “describes and/or forecast(s) the behaviour of a system using mathematical and statistical concepts, and whose parameters and their relationships are expressed as quantities” (García-Díaz et al. 2019).

We define model-based research loosely in order to capture the breadth of different modelling goals (e.g. exploration, inference and prediction; Tredennick et al. (2021)) and type of research contributions (e.g. development of a new method, etc #todo see Moallemi, Elsayah, and Ryan (2019) and across both pure and applied ecological modelling.

Specifying Preregistration use-cases

Pu et al. (2019) describe two key use-cases of preregistration, depending on the role assumed in interacting with the preregistration. Users have different requirements of preregistration depending on the role that they assume, that is, study authors and modellers have a different use-case than reviewers, editors and readers, however this latter use-case is often not explicitly designed and accounted for in current platforms/infrastructure/processes (Pu et al. 2019). Reviewers and editors are usually tasked with the process of ‘checking’ the preregistration, comparing the final manuscript to its preregistration and verifying that the preregistered analyses are conducted and reported as specified (Pu et al. 2019). However existing formats are designed primarily for the use-case of the author - they are text-based static documents that are ill-suited to checking the reported study against the preregistration.

Although preregistration authorship is the primary target use-case of the preregistration template and methodology developed in this study, we aim to consider and incorporate features into the preregistration process, such as template formatting, that will facilitate preregistration checking to also address non-authorship use-cases.

Articulating desirable attributes of the preregistration template and adaptive preregistration methodology

The preregistration template should capture critical decision-points that may influence the overall model and analysis as well as any recommendations or decisions drawn from analysis

of the model, thus delimiting researcher degrees of freedom at these junctures in the modelling decision process. The structure and content of the template is necessarily opinionated, encouraging Good Modelling Practices, while reflecting community norms and practices in model-development while still meeting the objective of mitigating the risk of Questionable Research Practices by delimiting researcher degrees of freedom. The template needs to be generalisable enough to apply to a range of common modelling purposes, methods and research contributions, and flexible enough that if a particular decision is not applicable for the planned study, that those items may be removed or altered, and other items may be added onto the template.

Secondary to the aim of mitigating QRPs, is that the the template should also enhance the research transparency of the preregistered study, ensuring that analysis choices are documented with sufficient transparency to facilitate the reader of the preregistered completed study in making informed assessments about the credibility and reliability of the model and any associated analysis and recommended interventions. Borrowing from Bodner, Fortin, and Molnár (2020)’s framework to increase the accuracy, reliability and transparency of predictive models, in order to increase transparency of model-reporting, the template should explicitly outline critical model choices and their assumptions, steps in the modelling process, and explicit expectations about the outputs and analysis of the model.

Adaptive Preregistration Methodology

The Adaptive Preregistration Methodology aims to capture the entire modelling process, including model development and construction, rather than beginning at the point of analysing a pre-existing and parameterised model for deductive inference / prediction. Critically, the methodology must not preclude the iterative and adaptive nature of model construction, while simultaneously restricting the types of data-dependent analytic decision-making that may compromise the reliability and validity of the model and its contingent analysis and conclusions.

User Research & Analysis

Characterise users and their needs

On the 4th and 5th of March 2020, we facilitated a series of structured activities at the Arthur Rylah Institute for Environmental Research which aimed to inform the development of two preregistration templates, one for field-work and sampling design, the other for modelling and data-analysis. We split the activities into two groups, to better focus discussion for each of these templates.

- ☐ Note: might need to reframe this and omit the field study stuff, but should probably state that we did this but decided not to pursue?

The objectives of the collaborative workshop activities were to aid in: 1. Understanding how modelling is done in practice to inform and tailor the preregistration templates to adequately

capture key decision-points in the scientific modelling process; 2. Understanding how science in practice – whether due to logistical, technical, feasibility or normative constraints – differs from some idealised practice, in order to: 3. Identifying challenges and roadblocks to the use and uptake of preregistration in among ecological modellers, both in empirical and applied research contexts; and to 4. Using the above knowledge to inform the preregistration template and process design so they meet their intended purpose, and are adopted by our target user community for their particular use-cases.

Workshop 1: Identifying Ecological Modelling Workflows, Norms, and Practice

The objective of this workshop was to identify modelling workflows that capture critical decision-steps in the model development cycle that should be included in a preregistration template. By “critical”, we mean that without these decision-steps being included, the template would not adequately constrain researcher degrees of freedom nor b) transparently document the analytic decision-process. The activity consisted of two-parts, first, we individually reflected on our scientific process for a recent or memorable research project, listing the steps we undertook from the beginning to end. In the second part of this exercise, we came together as a group to collate each of the steps in our personal workflows onto pre-prepared workflow template (insert specific OSF link to the document, Supp Mat 1). The template consisted of several key ‘phases’ of the modelling process, with commonly implemented decision-steps grouped under each phase. The development of this modelling workflow sheet was derived from previous work by Gould et al. *in prep* (insert link to paper preprint) and is outlined in Appendix 1 (#todo insert). We sorted and categorised each decision-step in our personal workflows as belonging to a particular workflow phase. Next we grouped similar decision-steps that could be included as a single item on a preregistration template. At the end of this process we reviewed the suggested phases of the workflow to assess whether phases needed further discretisation, aggregation, removal, or amendment.

Workshop 2: Understanding ecology- and modelling-specific challenges to adopting preregistration in research practice

This second activity aimed to identify critical challenges and roadblocks to undertaking pre-registration in ecology and conservation in general, and for model-based research in particular. We facilitated a discussion structured around a series of pre-prepared questions and talking points informed by recent debates about preregistration in the literature (#todo Supp Mat 1). These talking points were chosen with the aim of identifying problems in the content and or structuring of the phases and decision steps that were collectively identified in Workshop 1, as well as any conceptual or philosophical issues with preregistration in ecology, conservation, and in particular modelling, that might impact on the template structure, content and procedural implementation of preregistration. Discussion points included the following: - What should a

research workflow look like for a modeller in an applied ecological setting who wishes to pre-register their work? - What barriers are there for preregistering model-based research? - How do we accommodate the iterative cycle of model development into the preregistration process? - At what stage in a modelling or structured decision-making process should preregistration begin? - should we and how can we change the medium of the template and/or archiving platform to accomodate any roadblocks? - How could we change the medium of both the template and the archiving platform to accommodate these issues? - What procedural and or technical solutions could we implement to address these roadblocks to preregistration? - What decision-tools could be used to aid in completing a preregistration? - What sort of parsimony should we aim for in terms of the resolution of the templates and therefore their applicability across different methodologies?

Analysis of Workshop Outputs & Design

Analyse outputs of user-research workshops to inform the design of resource mock-ups

1. Identify Scientific Workflows for ecological modelling, both real/in-practice and idealised.

The goal of this analysis task was to describe a generalised workflow of the modelling cycle that will underpin the structure and content of the preregistration template. We used a combination of inductive and deductive coding (`#todo` is this what it is called and `#citation`) using the qualitative coding software ATLAS.ti (2019) to analyse the workshop outputs and generate a model of scientific workflow for model-based research that would constitute the presumed scientific workflow that would be used to guide the structure of the preregistration template. We collated and coded the workshop outputs using. All workshop outputs can be found here (`#todo` Link to Workshop Outputs OSF repository), and the coded analysis can be found here (`#todo` link to analysis files / screen caps).

A preliminary idealised scientific workflow for ecological modelling research was synthesised after reviewing the literature on ‘good modelling practice’ and structured decision-making literature in ecological management. Steps in the workflow were coded and broken down into ‘phases’ consisting of ‘steps’ and ‘sub-steps’, in line with the ontology we specified earlier for translating a modelling workflow into a preregistration template (section link), (insert link to appendices or just present the final?). Each decision-point in the personal modelling workflows generated from activity one of the collaborative workshop were coded as belonging to one of these phases, steps and sub-steps. We revised the coding structure (inductively is this the word?), according to patterns and themes identified across multiple personal workflows. This workflow is used to provide structure and content to the preregistration template (`#todo` insert link to this workflow).

2. Translating Workflows into Preregistration Templates

We translated the model-based workflow identified in part one of workshop analysis into a text-based preregistration template. Reflecting on existing templates ([#todo link to OSF repository of templates page](#), and others, namely secondary data analysis, see Olmo’s links), we identified common elements / units of preregistration templates that we mapped onto the generalised modelling workflow identified in part 1 of workshop outputs analysis. Following the features of a good guideline identified by Moallemi et al. (2019), we provide examples of choices or illustrations to clarify each step of the preregistration, especially for phases or steps in the the modelling process that are semantically, epistemically and/or philosophically contentious or ambiguous (link to paper describing contentious issues in modelling history)_. We synthesised the following ontology with which to build our preregistration template:

1. Decision Phase - each has a “title” and “description” or “definition”, steps are grouped under phases. (Note, that these are being scrapped from the template and are turned into tool-tips for formatting)
2. Decision Step - each has a “title” and “description” or “definition” choices are grouped into ‘steps’, which corresponds to an activity or set of related activities in a scientific workflow. (again, these are being scrapped and turned into tool-tips. CHECK THIS, it’s changed)
3. Choices —> each ‘choice’ corresponds to a uniquely numbered item on the PRT requiring a response from user. The style should be in the form of a directive or a question. E.g. “explain how you will do xyzzy”. Or “What performance measure will you use for assessing ‘goodness of fit’?”

3. Identifying challenges and solutions for implementing preregistration in model-based research

We analysed the outputs of Activity 2 in the collaborative workshop identifying barriers and solutions to undertaking preregistration in model-based research and to broader adoption of preregistration in the modelling community of practice. The analysis of Activity 2 in the collaborative workshop was used to inform the design of a suitable preregistration procedure for preregistering MBR in the domain of ecology.

User Testing

Build mock-ups of the product and test with users

While user-testing is an essential step in the UCD framework, we also concur with Devezer et al. (2020) that open-science reforms should be evaluated against their stated objectives. We therefore preregistered a real-world model-based research problem as a case-study for evaluating and testing the preregistration templates and proposed adaptive preregistration methodology.

Case Study

We preregistered a real-world study of vegetation responses to environmental flows in Victoria, Australia. This study is one component of a larger set of studies within the Stage 6

Victorian Environmental Flows Monitoring and Assessment Program (VEFMAP), managed by Department of Energy, Environment and Climate Action (DEECA) and delivered through the Arthur Rylah Institute for Environmental Research. The case-study analysis aimed to evaluate the effectiveness of environmental-flows on riparian vegetation with the following objectives: - Quantify short and medium-term vegetation responses to individual flow events and regimes. - Describe detailed vegetation trends associated with season, flow, rainfall and other factors such as livestock grazing and exotic vegetation - Quantify the effect of various factors on vegetation in order to guide future management expectations and planning.

Details of the study background, methods and results are described in Appendix A4.

Evaluation Criteria & Rationale

- *Evaluation Criteria and Rationale (should this be moved to earlier aims section? Seems like there's a little bit of overlap between the two...*

The content of the template should be ‘parsimonious’ in that the included items within the template should adequately a) constrain researcher degrees of freedom and by being specific, precise and exhaustive (Wicherts et al. 2016) b) facilitate transparent documentation of the modelling process and analytic decisions during model development and analysis, while c) not being too detailed to the point that the preregistration is too difficult and time-consuming to complete.

Many modellers (e.g. MacEachern and Van Zandt 2019) #todo #citation are concerned about preregistration making modelling a more difficult task than it already is, and hampering the ability to meet project deadlines, especially in applied settings in which models directly underpin decisions. Modellers already struggle with sufficiently documenting models during publication and reporting , and there is a resource cost to increasing transparency of model-reporting (de Vos, Janssen, and Van Bussel Proceedings of the 2011). Consequently we will also assess the preregistration template and process on its perceived feasibility and practicality.

Evaluation Method

Rather than attempting to reconstruct the process of creating the preregistration through using interview questions alone and digging back through the version history of a completed preregistration (e.g. Pu et al. 2019), we aimed to capture the analysts ‘thinking in the moment’ by utilising the version-control and collaborative project management features of Git and GitHub (Ben, Smith, and Bailey 2018; Blischak, Davenport, and Wilson 2016) as a tool for documenting both the analytic decisions of the researcher, and for using the preregistration process itself to ‘live-develop’ the template template. For example, if it turned out that we had missed an important item, or perhaps the phrasing of an item or the order and structure of the template needed to change whilst in the process of completing the preregistration, the

suggested change, and its justification would be both recorded in GitHub and linked to the actual preregistration.

Finally, we supplemented this ‘in situ’ testing and evaluation process with follow-up semi-structured interviews on 9th May 2024 in order to capture more detailed reflections of the analysts. Lead analysts of the case study were often under strict time constraints for meeting deliverables on the project, and the process of preregistration was a new and unfamiliar task, so capturing reflections on the broader preregistration process during preregistration and analysis of the case-study was not always possible.

Implementation

Implement the product, write technical implementations and build the product

We wrote the initial template as a text-based document in Rmarkdown. The initial template used to preregister the case-study underwent a series of changes throughout the preregistration process, including changes to the organisation, content, help-text and visual appearance. The final preregistration template is made available as a quarto markdown extension (#citation & link to doc) that can be installed and used to populate a new preregistration document for a new modelling study, and as a form that can be used with the **preregr** R package (Peters 2023).

To aid the case-study analysts in using the adaptive preregistration methodology to preregister their study, we wrote a detailed user-guide outlining the key principles of adaptive preregistration and a protocol for implementing adaptive preregistration using git and GitHub. The guide is presented as a publicly available website and also includes the template.

Results

User-research and analysis

The outputs and analysis for both Workshop 1 and 2 is archived at (insert OSF link). The modelling workflow generated from the analysis of Workshop 1 that informed the development of the initial preregistration template is archived at the same link. During user-testing with our case-study, the preregistration template structure, and therefore its underlying modelling workflow on which it is based underwent, a series of changes. Consequently we present the revised workflow alongside the refined preregistration template in section X below.

Discussions from Workshop Activity 2 focussing on challenges to preregistering MBR revealed that the iterative nature of model development is a significant barrier to adopting preregistration for model-based research. Modellers felt that adopting preregistration in their workflows

was impossible because preregistration was understood to be a process that precludes iterative model development because its inner logic presumes a linear research workflow premised on hypothesis-testing research, and because it requires researchers to confirm an unalterable version of the analysis plan prior to collecting or analysing the data (Pu et al. 2019; Mertens and Krypotos 2019).

The key tenets of preregistration are based on NHST-focused accounts of questionable research practices that prescribe the principle of ‘decision-independence’ where analysis decisions are made independently of and *a priori* to analysing the data (Srivastava 2018; Gelman and Loken 2013). However, analysts reflected on their personal modelling workflows and described situations in the modelling process that may violate the principle of ‘decision-independence’ that could not be avoided and were considered to be an intrinsic feature of the model-development process. During modelling a researcher will almost always engage in state-dependent or data-dependent decision-making – that is, decisions about whether to alter the model in a subsequent round of model development depend on the results of model testing, evaluation and analysis at the end of the previous iteration of model development.

Very often in modelling, there are decisions that are necessarily dependent on the outcome of previous analytic decisions in the modelling workflow (Liu, Althoff, and Heer 2020). Firstly, preliminary or investigatory analyses might need to be conducted before being able to specify future decision-steps in the analysis plan. For example, modellers might need to check the distribution of particular variables in order to determine how to specify the model, perform assumption checks, or check for collinearity or spatial autocorrelation. Secondly, data-dependent decisions from model checking may justifiably result in changes to either the analysis or to interpretations of the model (Srivastava 2018). For instance, inspection of the residuals and other model checking tests may force the modeller to return to earlier decision-points and change the planned analysis or model. Other times, the planned analysis and specified model may simply not converge, or the model is saturated and runs out of variation to apportion such that the model and or the model fitting algorithm must be respecified and re-implemented. Other times, the modelling process itself may generate knowledge and learning about the system, and the model structure changes throughout the process of model development.

Data-dependent decision-making is not necessarily intrinsically *questionable*, but it does have the *potential* to reduce the reliability of the model and its findings if they are exploited in service of artificially inflating the accuracy or precision of the model, its predictions or evaluation tests to the effect that the model is perceived to be more credible than it would be if that practice did not occur – essentially data-dependent decisions that lead to consumers of the model placing unsubstantiated belief in the reliability, validity and utility of the model and its outputs are questionable (E. Gould 2019) and can be mitigated through the practice of preregistration.

The concept of preregistering all modelling choices *a priori* was seemingly overwhelming for analysts. They described situations where decisions are too difficult to anticipate, or cannot be made in advance, particularly for decisions occurring at the later phases of model construction and development process (Figure 2, TBC), where downstream decisions depend on the outcomes of earlier decisions and outcomes of modelling analyses. Some decisions might not be

able to be specified on the first attempt at writing the preregistration until the model is fully or at least close to fully specified, for example, specifying precisely what and how sensitivity analyses or uncertainty analyses will be conducted.

- Finally, sometimes there are some decisions that are too difficult to anticipate, or simply cannot be made in advance (Srivastava 2018).
- I think Pu or Pavelin describe the same thing... they coin a name for these types of decisions too but I can't find it.

Design

Adaptive Preregistration: A methodology for preregistering model-based research

In order to address the disjuncture between the iterative nature of model development and the single-use, deterministic preregistration procedure, we propose an expanded view of preregistration, *adaptive reregistration*, which allows for some degree of flexibility and adaptivity in the analysis workflow, while still mitigating against unconstrained and undisclosed researcher degrees of freedom. Adaptive preregistration breaks with the current preregistration workflow, wherein the author writes a single deterministic preregistration containing a single deterministic choice for every decision-step and the research shifts from ideation, to preregistration to execution of the plan (fig or #citation to ref?). Instead, in adaptive preregistration, as the modeller proceeds through the model development process, they will transition from ideation, preregistration, execution, and back to preregistration and execution again. The modeller may repeatedly cycle through the preregistration process rather than proceeding in a linear fashion. Two features facilitate this process: 1. An *adaptive preregistration* contains 'plans to deploy flexible strategies' (Srivastava 2018) whereby the preregistration author can supply a heuristic consisting of different analysis or modelling strategies whose execution depends on the outcome of the previous decision-point or analysis, for example, decision-trees for certain points in the analysis or items in the preregistration template could be preregistered prior to analysing (Baldwin et al. 2022). 2. The preregistration may be iterative and consist of interim preregistrations that mark the transition between phases of modelling and analysis as different parts of the data are observed (Srivastava 2018).

- Other refs to chase up... Benning, Schuwirth, Latihan, hamalainen, hofman, reusable holdout paper, Dwork?

Preregistering Flexibility

When preliminary analyses, model checking or other data-dependent analytic decisions are to be preregistered, the modeller can specify a 'flexible strategy' within their preregistration. This is particularly relevant for when it is uncertain as to whether the planned analyses will be appropriate for the data (Baldwin et al. 2022), which is a commonly encountered issue

in ecological modelling, because often-times the model we might ideally want to fit that best represents our conceptual model of the target phenomena cannot be fitted due to a paucity of data (#todo citation for this? was a motivating factor for the VEFMAP project). Like existing approaches to preregistration, the registered heuristic is still a deterministic one, but there there may be multiple potential choices rather than one, and their realisation depends on the outcome of an earlier decision step or analysis.

To preregister flexibility, there are some critical requirements: 1. State what quantity or outcome needs to be known in order to move forward with the modelling and analysis, and 2. explain what test or analysis will be performed to obtain this quantity or outcome, and which parts of the data will be used in this analysis, and 3. describe how the results will be interpreted, linking each potential decision to its outcome trigger, where possible.

- ☐ Might want to give an example of this...? Or could give example in reporting the case study and cross-ref that? Alternatively, could put an example in the guide, and send readers there? I think second option is best.
- ☐ SEE ALSO OTHER refs in this vault, the idea of registering flexibility e.g. using decision-trees is not new.. which paper was it?- e.g. BENNING?

Interim Preregistrations

Just as model development proceeds in an iterative process of preregistration, so too does the process of adaptive preregistration. As the modeller proceeds, they will implement parts of the analysis plan that were specified in the previous preregistration step, letting the outcomes of that preregistered analysis inform either the specification or the realisation of the pre-registered specification in the next step of the analysis workflow (?@fig-adaptive-preregistration).

The initial preregistration should have as many items as possible, if not all items, in the preregistration template answered. If an item cannot be answered with a single decision because that decision is data-dependent, then the analyst should register flexibility for that item following the aforementioned requirements. Depending on observed outcomes of the registered decision-trees or flexible heuristics, interim preregistrations are created at multiple points in the model development process, each time there is an addition or amendment to the plan.

Integrating additional data handling and modelling practices

We propose additional data handling and analysis techniques be embedded into the both the modelling and adaptive preregistration processes to help avoid the pitfalls of data-dependent decisions, such as partitioning the data to create a holdout dataset or using blinded datasets for data exploration.

Partitioning data into train/test splits is a modelling technique that has been suggested for incorporation into preregistration for predictive modelling (Hofman et al. 2023) or other

non-hypothesis testing research (Baldwin et al. 2022; Dwork et al. 2015). We suggest that activities considered as typically ‘exploratory’ partition the data into ‘training’ and ‘holdout’ datasets where the training dataset can be used to inform potential models or analyses that will be preregistered in later versions. The holdout set may then be used for components of the modelling process that align more closely with confirmatory modes of inquiry, such as model evaluation or testing, forecasting, hind-casting and other predictive methods (Lewis et al. 2023; Pu and Kay 2018; Hofman et al. 2023) #todo add pred modelling citations, e.g. Houlahan et al. (2016) etc. see memo for Lewis et al.). However, partitioning ecological data is not always feasible due to small datasets, studies that are limited in duration or spatial extent compared to the target phenomena, or data patchiness due to the logistical and financial challenges to data acquisition in ecology (Wood et al. 2020).

In these cases where preregistration of flexibility and data partitioning is not feasible or where the analyst may need to check the data in order to pre-specify a decision, e.g. checking for data distributions or missingness, blinded datasets, such as those generated by data scrambling (Baldwin et al. 2022). Data blinding can prevent data-peeking that may influence downstream modelling decisions while aiding the modeller in making informed analytic decisions that are appropriate for the dataset on hand.

Finally, oftentimes interpretation of the model and its outputs are based on subjective assessments, such as visually interpreting plots of model fit quality for indications of issues such as heteroscedasticity, interpreting the results of model performance where there is no clear decision threshold, or interpreting model performance across a suite of different metrics and indicators. We encourage analysts to follow Baldwin et al.’s (2022) suggestion whereby both the research questions and the conditions for interpretation are preregistered. We have embedded this approach to preregistering MBR into our template as well, by including prompts within help and explanatory text that direct the researcher to ensure that the con

Implementing Adaptive Preregistration with GitHub

□ #todo insert summarised version of guide

Draft Template & Modelling Workflow

We developed a draft template for preregistering model-based research in ecology and conservation that was tested and further refined using our case-study of adaptive preregistration (#todo Appendix A2). We report the final preregistration template refined through the case-study, as well as the refined description of the modelling workflow underlying the preregistration template.

□ How do I present this in the manuscript?? Do I just refer to this in the supp mat? Feels important to present but this manuscript is already LONG!

- Also need to either make fig or table based on preregistration template for the underlying model development workflow underpinning it.

User Testing

Model development and adaptive preregistration processes occurred in tandem, and were implemented using git and GitHub, Leveraging GitHub’s version control features. This approach is summarised in the guide https://egouldo.github.io/EcoConsPreReg/adaptive_preregistration_implementation.html, (Elliot Gould et al. 2024). The GitHub repository is located at https://github.com/egouldo/VEFMAP_VEG_Stage6, but is currently embargoed as of today (30 April 2024).

- #todo consider whether we explain the GitHub procedure here... in the ‘Design’ section, in an appendix section, OR leave out and just reference the website and zenodo link. I DO think however, that it has some key features that map onto the adaptive preregistration process that I want to highlight. So perhaps we can present in the ‘Design’ section.

Final Template Version & Changes

The template used in the VEFMAP study underwent a series of changes to both its content and visual formatting during the analysis to better suit the needs of the modellers. The preregistration template was written as an *Rmarkdown* document saved within the project GitHub repository (INSERT LINK TO DOCUMENT #todo). Using knitr (CITE) and pandoc (CITE), this document is converted to a markdown document using RStudio’s `github_document` format, which is automatically rendered into HTML for viewing in the web browser by GitHub (https://rmarkdown.rstudio.com/github_document_format.html).

Improvements to template formatting include the use of dropdown boxes to store template text. Template text was divided into a) text that describes the rationale for the inclusion of a particular preregistration section or item within the template, as well as any explanation of key terms and definitions of some modelling practice / process terms, and b) text that contains specific questions with check-boxes for researchers to address in their preregistration. Each text-type is given a different icon to help differentiate between them. The use of drop-down text-boxes with icons better differentiates template prompts and text from authors’ responses to preregistration items and the VEFMAP modellers noted that these changes greatly improved readability of the preregistration document https://github.com/egouldo/VEFMAP_VEG_Stage6/pull/75#pullrequestreview-1959017213.

In terms of substantive changes to the structure and content of the preregistration template, significant changes were made to the structure of the document (reorganized different section), and to descriptions of item rationale and explanations, as well as the preregistration items themselves. Some items were omitted, some were added, most were reworded for improved expression, and some were re-structured into different sections.

- Collate comments, pull requests and reviews and issues describing problems with the changes see: https://github.com/egouldo/VEFMAP_VEG_Stage6/issues/29 #chrisjones and also other comments from #jdyen

Modellers think about the modelling process differently, and use different terms for the same or similar practices and modelling concepts, depending on their particular modelling background, and at times the language used in the original version of the template was confusing and unclear to the VEFMAP modellers.

- #todo consider as #discussion point in this manuscript #writing

Adaptive Preregistration of the Case Study

We tested and refined both the preregistration template and adaptive preregistration protocol as outlined in Version X of the guide using the VEFMAP modelling study as a case study. We did not preregister this study on an existing official preregistration platform that meets the three core requirements of preregistration outlined by Haroz (2022) because existing platforms are unable to accommodate adaptive preregistration where there are multiple and interim versions of the preregistration. Consequently tracked changes to the preregistration using version control with git and GitHub, and used GitHub's tag and release features to demarcate major changes and interim versions of the preregistration.

The preregistration process was marked by three main phases of analysis and modelling, each marked with an interim preregistration:

1. V1 - Data analysis and cleaning, specification of problem context and background preregistered (insert link to supp mat)
2. V2 - Pilot Analysis preregistered (insert link to supp mat)
3. V3 - Main analysis including decision tree for informing model selection based on results of pilot analysis (insert link to supp mat)

V1: Preregistration of problem context, data analysis and cleaning

Prior to preregistration and analysis, data for the VEFMAP study had already been collected, as this study is a small component of a broader project that has been running for a number of years. The preregistration was iteratively completed and updated prior to beginning any analysis or modelling with the VEFMAP dataset, however, data cleaning and wrangling was undertaken in tandem with completion of the initial preregistration document by different researchers on different git branches due to project time-constraints.

V2: Preregistration of Pilot Analysis

There was a great deal of uncertainty about whether the dataset would support fitting the desired candidate models due to temporal and spatial patchiness and density of the data. The

dataset consisted of over 2 million observations across different waterbodies which were missing data during some years, while other sites had been surveyed more times than others (insert table from meeting with DEECA into supp. mat and cross-ref [[2024-03-27_Meeting_Notes]]). Some preliminary analysis and exploration of the dataset was needed in order to make appropriate decisions about how best to operationalise the models given the patterns in the underlying dataset and the spatio-temporal variation in the density / patchiness of the dataset. For example, simple tests of data distributions within hierarchies needed to be implemented to check for data spread and presence of zero-inflation, and checks for collinearity among candidate predictor variables were needed to reduce overfitting and to ensure the relevant interactions and captured in the models. Consequently, we decided to implement the use of a holdout dataset by partitioning the data into a smaller ‘pilot analysis’ dataset and ‘main analysis’ dataset. The pilot analysis dataset consisted of data from only a single site that had good temporal data coverage, while the dataset for the main analysis consisted of the entire dataset, including the pilot dataset.

The pilot analysis was exploratory in nature and aimed to resolve critical uncertainties about how best to operationalise the models given data constraints in order to identify a candidate set of models that could feasibly be fitted using the full dataset. In addition to the preliminary analyses described above, two initial models for the two response variables, cover and richness, were preregistered. The same model checking procedure was preregistered for both the pilot and main analyses. The pilot analysis was preregistered in section “2.5.3 Describe any data exploration or preliminary data analyses” of the template.

Pilot Analysis Outcomes

The pilot analysis resolved some uncertainties, while some further analytic uncertainty about the best model operationalisation remained. Firstly, for models of both response variables, the data were unable to support a single model containing separate terms representing both flow regimes and flow events. Consequently two separate models, one with each term operationalising flow, were fitted to the pilot dataset instead. Secondly, separate models for each plant functional group were desired to help understand how each group responds to environmental flows. However, there was not enough data for each functional groups split across both native and introduced species to get these models to converge so plant functional group was included as a random effect instead and fitted to the pilot data.

Model checking revealed both high levels of over-dispersion and over-inflation. To account for different proportions of zeros among plant functional groupings, the pilot analysis allowed for zero-inflation to differ among plant functional groups, however did not account for over-dispersion. During the pilot analysis, models were refitted using negative-binomial distributions but failed to converge.

- #todo Report models here!
- #todo Report Decision tree here!

- See notes from the report back meeting here: [[2024-03-25_Meeting_Notes]] and [[2024-03-27_Meeting_Notes]]

V3: Preregistration of Main Analysis and Decision Tree

Although the full models with complex three-way interactions failed to converge on the smaller pilot dataset, it was hoped that they may still converge on the full dataset in the main analysis. Consequently, the full models were also specified in the main analysis, but with the two simplified ‘flow events’ and ‘flow regime’ models preregistered also in the event that the full models were unable to converge and/or provide reliable parameter estimates.

Given that zero-inflated Poisson models both converged, and reliably captured the proportion of zeros within the pilot dataset, this model family was preregistered for all models fitted to the full dataset, including simplified models. However, modellers were concerned that there may be unexpected issues with these approaches due to actual data distributions in the full dataset not matching expectations informed by the pilot analysis, such as unaccounted over-dispersion of the data. Consequently a second flexible strategy was preregistered to account for the case where over-dispersion was identified.

First, the two key vegetation response variables were to be modelled as zero-inflated Poisson distributions with a log-link function. However, in the case that model-checking identified over-dispersion in any of the models, the models were to be refitted using a negative binomial distribution. However, a rule was preregistered that the models re-fitted using the negative binomial distribution would only be accepted over the Poisson models if they *both* improved over-dispersion *and* did not decline in model performance, particularly in terms of model fit, zero-inflation, and posterior predictive checks.

The final decision on which model family should be selected based on model checking in the instance that overdispersion is identified is a fuzzy and subjective one in which explicit weightings between the different model performance criteria cannot be expressed or preregistered. Instead, modellers attempted to provide an explanation and description of the conditions for interpretation informing decisions about the model family:

While we are unable to *a priori* precisely weight these criteria in determining the final distribution, the final decision will be guided by the model’s overall ability to capture key associations reliably. This is especially likely to occur when there is no ‘perfect model’ and there is no dominant alternative choice of model distribution.

Analysis following final preregistration

TBD, briefly summarise what outcomes from the preregistered decision heuristics were realised in the final analysis.

Implementation: User Guides and Documentation

We have developed a guide for modellers on how to implement adaptive preregistration for model-based research packaged as a website at <https://egouldo.github.io/EcoConsPreReg/> (Elliot Gould et al. 2024). The guide first provides background material on the risk of researcher degrees of freedom to transparency and reliability of model-based research in order to motivate the use of preregistration in model-based research. The core materials include an overview of adaptive preregistration, as well as a practical step-by-step guide to implementing adaptive preregistration using [git](#) and [GitHub](#), two commonly used tools by modellers for version-controlling modelling code during model development (Braga et al. 2023). The preregistration template for model-based research in ecology and conservation is also hosted on this website. Components of the guide to adaptive preregistration have been repackaged and incorporated into this manuscript.

Continuous Improvement and Feedback

One of the key features of user-centred design is the principle of ‘continuous feedback’. We therefore encourage modellers to preregister their research using our template and adaptive preregistration protocol outlined in this manuscript, and to provide feedback and suggestions for improvement. Suggested changes may be submitted in the form of a pull-request to the repository ([#todo insert link to how to make a pull request to the new template repository](#)), or by posting a new issue ([#todo link to creating a new issue on the repo and setup template](#)). The repository and issues are publicly available, so any future discussion will be transparent and open to the ecological modelling community to participate in.

Discussion

- intro para summarising problem / objective:
 - Preregistration of model-based research has been rejected both within ecology and in other disciplines.
 - * But the problems that preregistration addresses, i.e. RDFs / QRPs, are likely to exist for MBR too.
 - There is a lack of discipline-specific preregistration templates and this impedes the uptake of open-science tools, such as preregistration, within ecology.
 - This study aimed to expand preregistration for use within ecological modelling and modelling more broadly by developing a preregistration template for ecological modelling research, and a methodology for preregistering model-based research more generally.

We took a user-centred design approach to the task of developing a preregistration template and methodology that accounts for the domain and methodology-specific nuances of ecological modelling. We did this by: 1. Identifying idealised norms and norms of practices to develop

‘idealised’ / ‘generalised’ research workflows that describe sets of commonly undertaken tasks (and some sort of order/sequence to reflect the workflow, both idealised and as practically conducted), within an ecological modelling study. 2. We then transformed these workflows into a ‘preregistration template’, by distilling a set of decision-items that, should a) adequately constrain researcher degrees of freedom to mitigate the risk of Questionable Research Practices in an ecological modelling exercise and b) adequately facilitates research/analytic transparency. We transformed the decision-steps (the highest resolution) into template items posed as directives or questions for preregistration users to complete, in accordance with the particularities of the study. 3. We qualitatively evaluated the template and the adaptive preregistration methodology for its feasibility of use in MBR using a case-study of environmental flows modelling in Victoria, Australia... The VEFMAP modelling study faced particular challenges to preregistration in terms of data complexity and patchiness that made this study a good candidate for trialling adaptive preregistration...

Preregistration Template

Blah what to say?

Adaptive Preregistration Methodology

User research revealed critical features of the model development process, namely the inherent iterative nature of model development, and requirements for model checking that forced us to reconsider the procedure for completing a preregistration suitable to ecological modelling. Consequently, we took an ‘expanded view of preregistration’, “Adaptive Preregistration”. Key features: - plans for preregistering and deploying flexible analysis strategies: - Iterative, cyclical preregistration rather than linear process that maps onto the iterative cycles of model development generating - interim preregistrations marking key points in the analysis

We also propose combining adaptive preregistration with previously proposed technical strategies for reducing the risk of overfitting and bias in MBR. E.g. using data-partitioning with a exploratory/training vs. test/holdout dataset / data scrambling for preliminary or exploratory analyses. We depart from existing research, e.g. Baldwin, by proposing that these preliminary analyses also be pre-registered, even if preregistration is not yet complete. The results of the preliminary or exploratory analyses may be used to inform analytic decision-making that can then be preregistered.

Presently there are no existing preregistration repository services that meet the critical requirements of a preregistration platform outlined by Haroz (2022) that can be used for adaptive preregistration. We proposed and refined an implementation of the Adaptive Preregistration methodology that leverages features of git and GitHub for transparently tracking the evolution of the preregistration document as the modelling progresses.

Reflections (heading TBD, can we make a water pun given case study and concluding sentence or is that lame)

We approached the application and evaluation of both the template and adaptive preregistration and sought to evaluate the feasibility of preregistration for MBR. We genuinely did not know what to expect when we began preregistering the case study, but approached this goal with a ‘proof is in the pudding mentality’ to ask what are the limits of preregistration.... Here we reflect on the parts that worked, and those that did not.

Time & Timing

- “A natural consequence of increased transparency, is that preregistrations become longer as more possible decisions are made transparent: *increased transparency implies increased length*. This trade-off was something we considered important during the design phase, and deciding what was in and out of scope for the preregistration template was particularly challenging.
- However, as Pu et al. note, some participants who were both authors and reviewers didn’t think that the time-costs of preregistration in research were significantly greater than not preregistering research. Instead they thought of the preregistration process invoking a ‘time shift’, that brought ‘decision points’ before data collection instead of during analysis.” P.19. In addition, “some reviewers thought that preregistration saves them the time they might have spent wondering about undisclosed flexibilities when working on an un-preregistered paper.”
- Even in adaptive preregistration where the process is iterative, modellers reflected on the front-heavy reserach timeline forced by the process of adaptive preregistration. Shifting decision-making to the start of the project timeline delayed the start of the analysis, but modellers felt that this resulted in better decision-making because the decision-making was more carefully reasoned. Additionally, the case study researchers felt that the level of detail reported in the preregistration aided in efficiently writing the internal report.
 - This relates to what Pu et al. found where participants ‘felt their research workflow improved by having preregistration as a “forcing function” [...] For example, writing down study plans ahead of time might prompt them to think more thoroughly and to anticipate flexibilities / researchers’ degree of freedmen advance.’ (Page 13). For reviewers ‘participant reports that preregistration makes the review process “easier”’.
- Mention difficulties I had with the internal deadlines of DEECA, and that we could not preregister everything we wanted to (i.e the analysis of the model and its outputs used for presentation in the report, which were extremely vaguely specified, aside from model checking). Preregistration could be challenging for modelling in such settings where there extremely tight timelines. BUT, as we are doing for the stage 2 report preregistration of the second study now, could be used during early consultation / scoping. So the

document could begin to be filled out much earlier than we did here (data already collected).

Resolution of the template

- Competing objectives? Although the purposes of preregistration are not mutually exclusive - as we have decided the purpose of preregistration we have defined here is to both delimit/state flexibility, *and* be flexible. However, “the preference of one purpose over another can have tangible effects on the design of preregistration” p15.. Pu.
- Compared to Hofman (2023), a template for preregistering predictive modelling research, our template was much more detailed and contained many more items to be preregistered.
- Get Jian / Chris’ thoughts on template. I suspect it’s a bit long / overwhelming. But potentially that’s because of the visual presentation / limitations of a static text based medium for the document.
- However, the benefit of a more exhaustive, specific and detailed preregistration is increased transparency (link back to intro on transparency in ecological modelling). By adopting preregistration for MBR we can also improve the transparency of ecological modelling, which is done terribly!!!
- cost to completeness, and specificity / precision is that it is time-consuming. But aside from increasing transparency, it is also critical for appropriately restraining unmitigated and undisclosed RDFs by preventing loose interpretation / operationalisation of the preregistration in instances where the preregistration is vague.
 - We can see this in the disjuncture between all of the modelling outputs conducted and reported for the internal report, but that weren’t fully specified in the preregistration due to time-constraints on the modellers.

Restricting Researcher Degrees of Freedom, not creativity!

Counter arguments to both preregistration and preregistration for modelling express fear that preregistration will restrict creativity and lead to ‘punch card style analysis’ (MacEachern and Van Zandt 2019). However, modellers in our preregistration of the case study found the opposite, that it prompted critical thinking about analytic choices during modelling, that might have only ever been tacit and never explicitly recorded (cite paper saying many modelling decisions never thought of as decisions and therefore never reported).

Combining the use of data partitioning, and splitting the analysis into a ‘pilot’ and ‘main’ analysis aided the modellers in understanding what models could feasibly be preregistered given constraints on the data, while also helping inform the preregistration of flexible analyses. The addition of the pilot study components to the template, and into the modelling process was good. And I think we should recommend doing this where possible. Link to that arxiv paper on preregistering machine learning studies. Might seem less natural for ecological modelling studies since we are data poor! BUT, in machine learning and other computer-science focussed

modelling contexts, this type of practice is the norm. One issue to flag is the potential for [[information leakage]], however. Because part of the data *is* used again in the full dataset. However, given that this study is situated within an ongoing program of monitoring and decision-making, I think it's ok. The modelling can be built into an adaptive monitoring / modelling context where future modelling can be used to test / evaluate and then update the model.

- How do we think the process actually went in terms of restricting researcher degrees of freedom / improving reliability / reproducibility? **What other reflections are there?**
- Tricky to anticipate?? But can certainly say that all of the exploratory analysis undertaken on the pilot dataset would have surely been applied to the full dataset had we not partitioned the dataset and the preregistration process.

Future Implementation & Testing

We did not undertake the ‘usability testing’ component of User-Centered design set out in fig 1. We offer this template and adaptive preregistration methodology to the ecological modelling community, and modelling communities across disciplines, to use, refine and adapt in their modelling practices moving forward. We hope insights gained through trial and error of its application will be used to further refine and improve the preregistration template and adaptive preregistration methodology.

Advocacy and uptake of preregistration is based on the working hypothesis that it is able to mitigate against QRPs by delimiting opportunistic exploitation of researcher degrees of freedom while improving transparency. While initial evidence suggests that preregistration may improve the reliability and reproducibility of research (cite study looking at replication rates of preregistered vs not), empirical testing of this working hypothesis is necessary for evaluating whether this new practice meets its intended purpose (Devezer et al. 2020) (and others). Future research should evaluate the effectiveness of adaptive preregistration for improving the reliability and transparency of MBR.

Hosting the preregistration template on GitHub, has the added benefit of providing a mechanism for dealing with *uncertain norms* in modelling practice (Pu et al. 2019), and for following modelling methods and decision contexts that might not fit the presumed modelling development cycle underpinning the preregistration template. Anyone is free to *fork* our GitHub repository, meaning that the entire contents of the repository can be copied to a users own repository, and the template is free to be altered as the user sees fit. This provides the opportunity for creating domain-and/or modelling-approach-specific templates (Crüwell and Evans 2019).

- #todo need to explain a bit further what the hell i mean by ‘uncertain norms’. I think I can link the confusion from VEFMAP modellers about some concepts within the pre-registration document in here.

In order to address this confusion among norms, I state the expected norms that the template has been designed around here. However, I don't expect users to agree on this process. Because we have kept the template in a GitHub repository that can be *forked*, people are free to adapt the template (the items themselves) to suit whatever norm they see fit about how preregistrations should be created and used.

Using GitHub for Adaptive Preregistration

Benefits

- transparency
- existing tool already used by many modellers for both version control and collaborative project management of computational code for models

Challenges

Separate branches... Under significant time pressure (internal reporting) to do the analysis. Data cleaning and analysis of pilot data was happening on one branch while the preregistration was happening on another! Also meant that the fledge release system outlined in the guide did not quite work because you had to wait until each branch had been merged back into the main / master branch in order to bump the version.

****Critiques of GitHub as a platform for preregistration. and combining other tools to meet the minimum requirements of preregistration: timestamps, persistence, registry***

While there are some proponents of the use of version control systems for preregistration (e.g. Braga et al. 2023; Hofman et al. 2023), (Haroz 2022) make the compelling argument that GitHub (and potentially alternative version control systems) are inappropriate as preregistration platforms because it doesn't meet any of the three core requirements for a preregistration: timestamps, an indexed registry, and persistence. While we have described the use of GitHub timestamps in this study, (Haroz 2022) illustrate that timestamps can be manipulated on GitHub. Moreover, currently GitHub does not have a formal registry for indexing or searching for preregistrations hosted on GitHub, which means that there is the potential for *[[preregistration hacking]]*, wherein preregistrations may be deleted changed and reuploaded (Haroz 2022). Finally, GitHub does not meet the *persistence* requirement, there is no immutable unique identifier, like a digital object identifier (DOI), and GitHub is not guaranteed as an archive service in perpetuity *[[data sharing]]* / *[[data curation]]*.

While the use of GitHub in this study was to test it as a platform for implementing *Adaptive Preregistration* rather than *Preregistration* per se, given these shortcomings, we recommend pairing GitHub with other platforms for implementing Adaptive Preregistration.

Here we tested it as a system for managing and implementing *Adaptive Preregistration* rather than just preregistration in general. However, given these shortcomings, we recommend pairing the use of GitHub for Adaptive Preregistration with some additional tools to ensure that

these three minimum requirements of preregistration are met, we outline two complementary approaches here that meet the three core requisites of preregistration:

1. Archive the GitHub repository using Zenodo (CITE) at the first draft preregistration (V.1.0) and combine with GitHub tag and release strategy, such that each interim pre-registration is recorded as a new version / release on both GitHub and Zenodo . Deals with both the timestamped issue and the persistence issue by giving it a digital object identifier (**smithSoftwareCitationPrinciples2016?**), Furthermore, Because it is given a DOI, it is indexed and searchable (Hofman et al. 2023).
 - Add links to other papers suggesting Zenodo be combined with GitHub for data / code sharing #todo
2. But not a specifically preregistration registry. How to deal with the registry component? Can we add as component to OSF repo?? and then preregister? Yes. In fact, an alternative solution has been outlined by (Van Lissa et al. 2021), where they have used a similar approach to us with a version-controlled plain-text preregistration document containing the pre-analysis plan, and tagging and releasing this document on GitHub. However, they suggest a secondary optional step of uploading the preregistration to a dedicated preregistration server, without specifying what that server may be. We could use an ‘open-ended Registration’ template on the OSF (Hofman et al. 2023), and then link the document from GitHub component on the OSF project to the registry. Instructions for doing this are given in (Hofman et al. 2023).
 - WHAT ABOUT DIFFERENT VERSIONS OF THE PREREGISTRATION.. Can we make changes using the OSF?? DOUBLE CHECK THIS! Might not be able to use the registry here. #todo I think #MarkCall said there is the option to make changes now. Could we test with their staging platform?? Put the different versions on there?? But the template has changed.... Could just try it anyway as proof of concept. And note this as a limitation of the OSF templates approach (not being modular or flexible, as (Hofman et al. 2023) outline).
 - #todo CHECK THIS: Embedding the repository within an OSF repository will fix the timestamp issue from GitHub that (Hofman et al. 2023) outline??

Conclusions

Instead of throwing the baby out with the bathwater, in this study, we advocate that preregistration can and should be used in model-based research in ecology and conservation, but its particular form should reflect the norms and practices of the research context in which it is applied, in order to adequately restrict researcher degrees of freedom that may uniquely emerge within a domain-specific methodology.

This research provides a template and methodology for implementing adaptive preregistration of ecological models in either pure or applied settings and will facilitate the wider uptake of preregistration in ecology. Although we focus on ecology and conservation in this paper, the template is extensible to alternative fields and methods, and the proposed methodology for implementing adaptive preregistration can be applied in model-based re-search within science more broadly. Modellers need no longer cry “but, I can’t preregister my research.”

Software & Data Availability Statement

Author Contributions

CRedit

Supp Material

☐ Check out journal guidelines or papers #todo

SM1: Workshop Materials and Process Design

- Design process for the modelling workflow template for activity 1. The actual materials are in A1.
- Activity 2 structured discussion with prepared questions and talking points informed by recent debates about preregistration in the literature. Any literature review leading to this. Actual talking points in A1.
- Mention that we also ran the above in separate groups for designing a field-study preregistration template, but decided not to pursue further here.

SM?: Analysis of Workshop Outputs

“We revised the coding structure (inductively is this the word), according to patterns and themes identified across multiple personal workflows:” For example, we added the final phase (name for last phase), which aimed to capture the fact that analysts usually present the results of their modelling to clients, decision-makers or other stakeholders that must interpret the evidence and are responsible for making management decisions. In this way, we identified a common or generalised workflow from different types of applied ecological modelling projects that synthesised both (idealised norms and norms of practice, Pu et al., see other philosophy of sci refs).

A2: Preregistration Template

A3: Adaptive Preregistration Guide

A4: Environmental Flows Modelling Case Study

Study Background

Riparian vegetation is a critical ecosystem component for so many reasons (biodiversity, habitat, food, stability, filter, carbon source, aesthetic, etc.) but it is also severely degraded in many waterways all over the world. Vegetation clearing, livestock grazing, pollution, exotic species (flora and fauna) and alteration of flows are some of the most severe threats.

The impacts of river regulation on vegetation depend on the type and extent of flow alteration. Flows can increase or decrease total flow, be more or less variable, or have altered season flow timing. Environmental flows are designed to be manipulations of flow (usually flow releases from a reservoir) that aim to have specific and direct environmental benefits.

Evaluating the effects of environmental flows on waterways is difficult because of the need to evaluate the isolated impacts of specific flow events (short term response) as well as the impacts of successive flows (environmental or otherwise) within a broader flow regime (medium to long term response). Data collection at large temporal scales (annual or greater) prevents evaluation of any single flow event or component from the regime. Evaluating the effects of environmental flows in most cases is therefore only possible with repeated sampling at relatively short intervals. Evaluation also needs to be done with consideration of other major drivers of vegetation, such as rainfall, other flow attributes, livestock grazing, soil properties, exotic species and previous disturbance.

In south-eastern Australia, river regulation (including water extraction), livestock grazing and exotic species continue to have serious negative impacts on all major waterways, despite major government investment into waterway management. However, increased understanding of the impacts of major threats and the effectiveness of management is resulting in greatly improved management decisions. While many broad trends are known, critical details remain elusive and require concerted effort to address.

Study Aims

Study Methods

Data Collection

- ☐ 1 - 2 sentence overview, then point readers to details in [[project-VEFMAP|VEFMAP]] manuscript #todo #writing

Modelling Dataset

Variable Name	Unit	Definition	Variable Type	Variable Group
Species Vegetation Cover	Cover	Cover per-species, per sub-transect	Response	Vegetation Cover
Plant Functional Group Cover	Cover	Cover per-plant functional group, per-sub-transect	Response	Vegetation Cover
Sub-transect Vegetation Cover	Cover	Total cover (sum all species), per sub-transect	Response	Vegetation Cover
Species Richness	Counts	Number unique Species Functional Group, per sub-transect	Response	Species Richness
Days above baseflow / springfresh levels	Counts, Z-scale	Number of days per year (averaged over prev 3 water years)discharge > baseflow or springfresh levels,Z-scaled	Predictor	Flow Regime
Sub-transect zone	Factor	Zone in which a sub-transect occurred:- below baseflow level, - between baseflow and spring-fresh levels- Above spring-fresh level	Predictor	Flow Regime
Flow event	Factor	Timing of sampling:- before spring fresh- Between spring and summer fresh- After summer fresh	Predictor	Flow Event
Site Grazing Status	Factor	Grazing status of site:- Grazed- Ungrazed	Predictor	Other factors influencing spatio-temporal variation in cover/abundance

Variable Name	Unit	Definition	Variable Type	Variable Group
Species Functional Group	Factor	For each Origin group (native or exotic):- aquatic- emergent- riparian- terrestrial	Predictor	Other factors influencing spatio-temporal variation in cover/abundance
Site Identity	Factor	Name / Identity of the site	Random effect	Other factors influencing spatio-temporal variation in cover/abundance
Transect Identity	Factor	Identity of Nested within sites	Random effect	Other factors influencing spatio-temporal variation in cover/abundance
Sub-transect height	Metres above baseflow	Height of the sub-transect along bank	Random effect	Other factors influencing spatio-temporal variation in cover/abundance
Survey Year	Date, YYYY	Year in which the sample was recorded	Random effect	Other factors influencing spatio-temporal variation in cover/abundance

Variable Name	Unit	Definition	Variable Type	Variable Group
Waterbody	Factor	Waterbody in which the site is located	Random effect	Other factors influencing spatio-temporal variation in cover/abundance

Pilot Study

The final data set comprised XYZ observations, which presented a significant computational challenge. In addition, exploratory data analysis revealed considerable zero-inflation and potential over-dispersion of the response variables, particularly vegetation cover. We conducted a pilot study to determine the feasibility of different model structures and identify likely decision points given these challenges.

□ when did the EDA occur... during the pilot study? On a different data set? #help-wanted #jdyen

To address this, we defined a maximal model structure and tested this structure with a pilot data set (all observations from the Campaspe River from 2017–2020) to identify feasible model structures and likely decision points. We developed separate analyses for vegetation cover and species richness, with the same maximal model structure for both.

Initial model fitting assumed a zero-inflated Poisson distribution for both vegetation cover and species richness, with zero-inflated negative binomial distributions designated in cases where over-dispersion was severe. This approach does not introduce an upper bound on estimates of cover (e.g., specifying a maximum cover level of 100%). We used this approach for three reasons. First, overlapping species may generate cover estimates that exceed 100%. Second, parameters of bounded distributions are challenging to estimate when many values occur near the bounds. Third, bounded distributions require link functions (e.g., the logit function) that complicate the specification and interpretation of autoregressive models. We note that the Poisson distribution approximates the bounded binomial distribution in the limit of many trials (REF). The extent of zero-inflation and over-dispersion (and a model’s capacity to capture these patterns) was assessed using posterior predictive checks, which use values simulated from the fitted model to determine whether the model structure captures the observed distribution of the response variable (REF). ### Results

Pilot Study Results

The final maximal model was as follows:

```
response (cover or species richness) ~ cover in previous survey +  
                                     functional group × origin × flow metrics +  
                                     functional group × origin × zone × sampling period +  
                                     grazing +  
                                     (1 | site / transect) + (1 | sub-transect height) +  
                                     offset(number of points).
```

Vegetation cover in the previous survey was included only for analyses of cover and meant that models of vegetation cover examined how flow influenced the change in cover through time rather than absolute vegetation cover (akin to models of population growth rate used in the VEFMAP long-term fish monitoring project). Flow metrics are the (z-scaled) days above baseflow level and days above spring-fresh level in the three years preceding a survey and included both linear and quadratic associations[JDLY(1)]. Zone is the location of a sub-transect relative to baseflow or spring-fresh levels and sampling period is the timing of a survey relative to the delivery of spring and summer freshes. Sub-transect height records the level of a sub-transect in metres above baseflow level. Waterbody was excluded from the pilot study given its focus on a single waterway (the Campaspe River).

Results of the pilot study inform the subsequent analyses and are presented here. Models of vegetation cover at the species level did not converge in any instance. Models of vegetation cover aggregated to watering-response plant functional groups converged but required separate models of two-way interactions instead of the three- and four-way interactions included in the maximal model structure. Models of vegetation cover aggregated to structural groups (aquatic, emergent, riparian, terrestrial) converged but required removal of interactions with species origin and separate analyses of flow metrics and for each of the categorical flow terms (zone and sampling period). Models of vegetation species richness within watering-response functional groups converged but required separate models of two-way interactions instead of the three- and four-way interactions included in the maximal model structure. Models of vegetation species richness within structural groups converged but required removal of interactions with species origin and separate models for the two categorical flow variables (zone and sampling period). Therefore, subsequent analyses of vegetation cover and species richness focused on structural groups because these supported model structures closest to the maximal structure. The included interactions were re-assessed in models of the full data set because the inclusion of additional data was predicted to either support more-complex model structures or introduce additional heterogeneity (and, therefore, support less-complex model structures) than in the pilot study. This did not include re-assessment of the functional group classifications.

Posterior predictive checks indicated that a zero-inflated Poisson distribution reliably captured the distribution of observed vegetation cover and a Poisson distribution (without zero-inflation) reliably captured observed distributions of species richness (FIGS). Models of cover had excess

over-dispersion, but a zero-inflated negative binomial distribution did not substantially reduce unmodelled over-dispersion and resulted in worse characterisation of zero-inflation and the observed distribution of cover data (FIGS?). Therefore, subsequent models of vegetation cover and species richness assumed that data were derived from a Poisson distribution, with additional parameters for zero-inflation specific to each functional group in the case of vegetation cover.

Full Dataset

Appendices

A1 : Workshop Materials

- Activity 1 Template, and instructions / materials provided to workshop attendees. (insert link to OSF)
- Activity 2 structured discussion with prepared questions and talking points informed by recent debates about preregistration in the literature (insert link to OSF)

A2: Draft Preregistration Template

☐ Insert draft template

A3: Final preregistration Template

☐ Insert template

Alexandrov, G. A., D. Ames, G. Bellocchi, M. Bruen, N. Crout, M. Erechtkhoukova, A. Hildebrandt, et al. /03/01 2011. “Technical Assessment and Evaluation of Environmental Models and Software: Letter to the Editor.” *Environmental Modelling & Software* 26 (3): 328–36. <https://doi.org/10.1016/j.envsoft.2010.08.004>.

Alspaugh, Sara, Nava Zokaei, Andrea Liu, Cindy Jin, and Marti A. Hearst. 2019. “Futzing and Moseying: Interviews with Professional Data Analysts on Exploration Practices.” *IEEE Transactions on Visualization and Computer Graphics* 25 (1): 22–31. <https://doi.org/10.1109/TVCG.2018.2865040>.

“ATLAS.ti Mac.” 2019. ATLAS.ti Scientific Software Development GmbH.

Augusiak, Jacqueline, Paul J. Van den Brink, and Volker Grimm. 2014. “Merging Validation and Evaluation of Ecological Models to ‘Evaluation’: A Review of Terminology and a Practical Approach.” *Ecological Modelling* 280: 117–28. <https://doi.org/10.1016/j.ecolmodel.2013.11.009>.

- Bakker, Marjan, Coosje Lisabet Sterre Veldkamp, Marcel A. L. M. van Assen, Elise Anne Victoire Cromptoets, How Hwee Ong, Brian A. Nosek, Courtney K. Soderberg, David Thomas Mellor, and Jelte M. Wicherts. 2018. “Ensuring the Quality and Specificity of Preregistrations.” *PsyArXiv*. <https://doi.org/10.31234/osf.io/cdgyh>.
- Baldwin, J. R., J. B. Pingault, T. Schoeler, H. M. Sallis, and M. R. Munafò. 2022. “Protecting Against Researcher Bias in Secondary Data Analysis: Challenges and Potential Solutions.” *Eur J Epidemiol* 37 (1): 1–10. <https://doi.org/10.1007/s10654-021-00839-0>.
- Bennett, Neil D., Barry F. W. Croke, Giorgio Guariso, Joseph H. A. Guillaume, Serena H. Hamilton, Anthony J. Jakeman, Stefano Marsili-Libelli, et al. 2013. “Characterising Performance of Environmental Models.” *Environmental Modelling & Software* 40: 1–20. <https://doi.org/10.1016/j.envsoft.2012.09.011>.
- Benning, Stephen D., Rachel L. Bachrach, Edward A. Smith, Andrew J. Freeman, and Aidan GC Wright. 2019. “The Registration Continuum in Clinical Science: A Guide Toward Transparent Practices.” *Journal of Abnormal Psychology* 128 (6): 528.
- Betts, Matthew G., Adam S. Hadley, David W. Frey, Sarah J. K. Frey, Dusty Gannon, Scott H. Harris, Hankyu Kim, et al. -3-25 2021. “When Are Hypotheses Useful in Ecology and Evolution?” *Ecol Evol*, -3-25 2021. <https://doi.org/10.1002/ece3.7365>.
- Bodner, Korryn, Marie-Josée Fortin, and Péter K. Molnár. 2020. “Making Predictive Modelling ART: Accurate, Reliable, and Transparent.” *Ecosphere* 11 (6). <https://doi.org/10.1002/ecs2.3160>.
- Braga, Pedro Henrique Pereira, Katherine Hébert, Emma J. Hudgins, Eric R. Scott, Brandon P. M. Edwards, Luna L. Sánchez Reyes, Matthew J. Grainger, et al. 2023. “Not Just for Programmers: How GitHub Can Accelerate Collaborative and Reproducible Research in Ecology and Evolution.” *Methods in Ecology and Evolution* 14 (6): 1364–80. <https://doi.org/10.1111/2041-210X.14108>.
- Brudvig, Lars A. 2017. “Toward Prediction in the Restoration of Biodiversity.” *Journal of Applied Ecology* 54 (4): 1013–17. <https://doi.org/10.1111/1365-2664.12940>.
- Campbell, Harlan. 2021. “The Consequences of Checking for Zero-Inflation and Overdispersion in the Analysis of Count Data.” Edited by Robert B. O’Hara. *Methods in Ecology and Evolution* 12 (4): 665–80. <https://doi.org/10.1111/2041-210x.13559>.
- Connolly, S. R., S. A. Keith, R. K. Colwell, and C. Rahbek. 2017. “Process, Mechanism, and Modeling in Macroecology.” *Trends Ecol Evol* 32 (11): 835–44. <https://doi.org/10.1016/j.tree.2017.08.011>.
- Crüwell, Sophia, and Nathan J. Evans. 2019. “Preregistration in Complex Contexts: A Preregistration Template for the Application of Cognitive Models.” <https://doi.org/10.31234/osf.io/2hykx>.
- de Vos, M. G., S. J. C. Janssen, and L. G. J. Van Bussel Proceedings of the. 2011. “Are Environmental Models Transparent and Reproducible Enough.” In *19th International Congress on Modelling and Simulation International Congress on Modelling and Simulation*.
- Desbureaux, S. 2021. “Subjective Modeling Choices and the Robustness of Impact Evaluations in Conservation Science.” *Conserv Biol*, March. <https://doi.org/10.1111/cobi.13728>.
- Devezer, Berna, Danielle J. Navarro, Joachim Vandekerckhove, and Erkan Ozge Buzbas. 2020. “The Case for Formal Methodology in Scientific Reform,” April. <https://doi.org/10.1101/>

2020.04.26.048306.

- Dwork, Cynthia, Vitaly Feldman, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Aaron Roth. 2015. "The Reusable Holdout: Preserving Validity in Adaptive Data Analysis." *Science* 349 (6248): 636–38. <https://doi.org/10.1126/science.aaa9375>.
- Fidler, Fiona, Hannah Fraser, Michael A. McCarthy, and Edward T. Game. /04/15 2018. "Improving the Transparency of Statistical Reporting in Conservation Letters." *Conservation Letters* 11 (2): e12453–4. <https://doi.org/10.1111/conl.12453>.
- Fife, D. A., and J. L. Rodgers. 2022. "Understanding the Exploratory/Confirmatory Data Analysis Continuum: Moving Beyond the "Replication Crisis"." *Am Psychol* 77 (3): 453–66. <https://doi.org/10.1037/amp0000886>.
- Flournoy, J. C., N. Vijayakumar, T. W. Cheng, D. Cosme, J. E. Flannery, and J. H. Pfeifer. 2020. "Improving Practices and Inferences in Developmental Cognitive Neuroscience." *Dev Cogn Neurosci* 45 (October): 100807. <https://doi.org/10.1016/j.dcn.2020.100807>.
- Fraser, Hannah, Tim Parker, Shinichi Nakagawa, Ashley Barnett, and Fiona Fidler. 2018. "Questionable Research Practices in Ecology and Evolution." Edited by Jelte M. Wicherts. *PLoS One* 13 (7): e0200303. <https://doi.org/10.1371/journal.pone.0200303>.
- García-Díaz, P., T. A. A. Prowse, D. P. Anderson, M. Lurgi, R. N. Binny, and P. Cassey. 2019. "A Concise Guide to Developing and Using Quantitative Models in Conservation Management." *Conserv Sci Pract* 1 (2): e11. <https://doi.org/10.1002/csp2.11>.
- Gelman, Andrew, and Eric Loken. 2013. "The Garden of Forking Paths: Why Multiple Comparisons Can Be a Problem, Even When There Is No 'Fishing Expedition' or 'p-Hacking' and the Research Hypothesis Was Posited Ahead of Time." *Department of Statistics, Columbia University*.
- Getz, Wayne M., Charles R. Marshall, Colin J. Carlson, Luca Giuggioli, Sadie J. Ryan, Stephanie S. Románach, Carl Boettiger, et al. 2017. "Making Ecological Models Adequate." Edited by Tim Coulson. *Ecology Letters* 21 (2): 153–66. <https://doi.org/10.1111/ele.12893>.
- Gould, E. In prep. "Questionable Research Practices in Ecological Modelling," In prep.
- . 2019. "Questionable Research Practices in Non-Hypothesis Testing Research: Ecological Models for Conservation Decision-Making." In *Metascience Symposium*.
- Gould, Elliot, Christopher Jones S., J. D. L. Yen, Hannah Fraser S., Henry Wootton, Lyndsey Vivian, Megan Good, David Duncan H., Libby Rumpff, and Fiona Fidler. 2024. "EcoConsPreReg: Adaptive Preregistration for Model-Based Research in Ecology and Conservation." Zenodo. <https://doi.org/10.5281/zenodo.10807200>.
- Grimm, Volker, Jacqueline Augusiak, Andreas Focks, Béatrice M. Frank, Faten Gabsi, Alice S. A. Johnston, Chun Liu, et al. 2014. "Towards Better Modelling and Decision Support: Documenting Model Development, Testing, and Analysis Using TRACE." *Ecological Modelling* 280 (May): 129–39. <https://doi.org/10.1016/j.ecolmodel.2014.01.018>.
- Haroz, Steve. 2022. "Comparison of Preregistration Platforms." <https://doi.org/10.3122/osf.io/zry2u>.
- Hildebrandt, Mireille. 2018. "Preregistration of Machine Learning Research Design Against P-hacking." In *BEING PROFILED: COGITAS ERGO SUM*, edited by Irina Baraliuc Emre Bayamlioğlu. Amsterdam, Netherlands: Amsterdam University Press.
- Hoffmann, S., F. Schönbrodt, R. Elsas, R. Wilson, U. Strasser, and A. L. Boulesteix. 2021.

- “The Multiplicity of Analysis Strategies Jeopardizes Replicability: Lessons Learned Across Disciplines.” *R Soc Open Sci* 8 (4): 201925. <https://doi.org/10.1098/rsos.201925>.
- Hofman, Jake M., Angelos Chatzimparmpas, Amit Sharma, Duncan J. Watts, and Jessica Hullman. 2023. “Pre-Registration for Predictive Modeling.” arXiv. <https://arxiv.org/abs/2311.18807>.
- Houlahan, Jeff E., Shawn T. McKinney, T. Michael Anderson, and Brian J. McGill. 2016. “The Priority of Prediction in Ecological Understanding.” *Oikos* 126 (1): 1–7. <https://doi.org/10.1111/oik.03726>.
- Ioannidis, John P. A. 2022. “Pre-Registration of Mathematical Models.” *Mathematical Biosciences* 345 (March): 108782. <https://doi.org/10.1016/j.mbs.2022.108782>.
- Jakeman, A. J., R. A. Letcher, and J. P. Norton. /05/01 2006. “Ten Iterative Steps in Development and Evaluation of Environmental Models.” *Environmental Modelling & Software* 21 (5): 602–14. <https://doi.org/10.1016/j.envsoft.2006.01.004>.
- Lewis, Abigail S. L., Christine R. Rollinson, Andrew J. Allyn, Jaime Ashander, Stephanie Brodie, Cole B. Brookson, Elyssa Collins, et al. 2023. “The Power of Forecasts to Advance Ecological Theory.” *Methods in Ecology and Evolution* 14 (3): 746–56. <https://doi.org/10.1111/2041-210x.13955>.
- Liu, Yang, Tim Althoff, and Jeffrey Heer. 2020. “Paths Explored, Paths Omitted, Paths Obscured: Decision Points & Selective Reporting in End-to-End Data Analysis.” In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM. <https://doi.org/10.1145/3313831.3376533>.
- Lupia, Arthur, and Colin Elman. 2014. “Openness in Political Science: Data Access and Research Transparency.” *PS: Political Science & Politics* 47 (01): 19–42. <https://doi.org/10.1017/s1049096513001716>.
- MacEachern, Steven N., and Trisha Van Zandt. 2019. “Preregistration of Modeling Exercises May Not Be Useful.” *Comput Brain Behav* 2 (3-4): 179–82. <https://doi.org/10.1007/s42113-019-00038-x>.
- McDermott, M. B. A., S. Wang, N. Marinsek, R. Ranganath, L. Foschini, and M. Ghassemi. 2021. “Reproducibility in Machine Learning for Health Research: Still a Ways to Go.” *Sci Transl Med* 13 (586). <https://doi.org/10.1126/scitranslmed.abb1655>.
- Mertens, G., and A. M. Krypotos. 2019. “Preregistration of Analyses of Preexisting Data.” *Psychol Belg* 59 (1): 338–52. <https://doi.org/10.5334/pb.493>.
- Moallemi, Enayat A., Sondoss Elsayah, and Michael J. Ryan. 2019. “Strengthening ‘Good’ Modelling Practices in Robust Decision Support: A Reporting Guideline for Combining Multiple Model-Based Methods.” *Mathematics and Computers in Simulation*. <https://doi.org/10.1016/j.matcom.2019.05.002>.
- Munafò, Marcus R., Brian A. Nosek, Dorothy V. M. Bishop, Katherine S. Button, Christopher D. Chambers, Nathalie Percie du Sert, Uri Simonsohn, Eric-Jan Wagenmakers, Jennifer J. Ware, and John P. A. Ioannidis. 2017. “A Manifesto for Reproducible Science.” *Nature Human Behaviour* 1 (1): 1–9. <https://doi.org/10.1038/s41562-016-0021>.
- Nissen, Silas Boye, Tali Magidson, Kevin Gross, and Carl T. Bergstrom. 2016. “Publication Bias and the Canonization of False Facts.” *Elife* 5: e21451. <https://doi.org/10.7554/eLife.21451.001>.

- Nosek, Brian A., Charles R. Ebersole, Alexander C. DeHaven, and David T. Mellor. 2018. "The Preregistration Revolution." *Proceedings of the National Academy of Sciences of the United States of America* 115 (11): 2600–2606. <https://doi.org/10.1073/pnas.1708274114>.
- Parker, T., H. Fraser, and S. Nakagawa. 2019. "Making Conservation Science More Reliable with Preregistration and Registered Reports."
- Pavelin, K., J. A. Cham, P. de Matos, C. Brooksbank, G. Cameron, and C. Steinbeck. 2012. "Bioinformatics Meets User-Centred Design: A Perspective." *PLoS Comput Biol* 8 (7): e1002554. <https://doi.org/10.1371/journal.pcbi.1002554>.
- Peters, Gjalt-Jorn. 2023. "Prereg: Specify (Pre)Registrations and Export Them Human-And Machine-Readably."
- Prosperi, Mattia, Jiang Bian, Iain E. Buchan, James S. Koopman, Matthew Sperrin, and Mo Wang. 2019. "Raiders of the Lost HARK: A Reproducible Inference Framework for Big Data Science." *Palgrave Communications* 5 (1). <https://doi.org/10.1057/s41599-019-0340-8>.
- Pu, Xiaoying, and Matthew Kay. 2018. "The Garden of Forking Paths in Visualization: A Design Space for Reliable Exploratory Visual Analytics: Position Paper." In, 2018 IEEE evaluation and beyond-methodological approaches for visualization (beliv):37–45. IEEE. <https://doi.org/10.1002/0470033312>.
- Pu, Xiaoying, Licheng Zhu, Matthew Kay, and Frederick Conrad. 2019. "Designing for Pre-registration in Practice: Multiple Norms and Purposes."
- Roettger, Timo B. 2019. "Researcher Degrees of Freedom in Phonetic Research." *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 10 (1). <https://doi.org/10.5334/labphon.147>.
- Schmolke, Amelie, Pernille Thorbek, Donald L. Deangelis, and Volker Grimm. /08/01 2010. "Ecological Models Supporting Environmental Decision Making: A Strategy for the Future." *Trends in Ecology & Evolution* 25 (8): 479–86. <https://doi.org/10.1016/j.tree.2010.05.001>.
- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2017. "False-Positive Psychology." *Psychological Science* 22 (11): 1359–66. <https://doi.org/10.1177/0956797611417632>.
- Simonsohn, U., J. P. Simmons, and L. D. Nelson. 2020. "Specification Curve Analysis." *Nat Hum Behav*, July. <https://doi.org/10.1038/s41562-020-0912-z>.
- Srivastava, Sanjay. 2018. "Sound Inference in Complicated Research: A Multi-Strategy Approach." <https://doi.org/10.31234/osf.io/bwr48>.
- Tredennick, Andrew T., Giles Hooker, Stephen P. Ellner, and Peter B. Adler. 2021. "A Practical Guide to Selecting Models for Exploration, Inference, and Prediction in Ecology." *Ecology* 102 (6): e03336. <https://doi.org/10.1002/ecy.3336>.
- Van Lissa, Caspar J., Andreas M. Brandmaier, Loek Brinkman, Anna-Lena Lamprecht, Aaron Peikert, Marijn E. Struiksma, and Barbara M. I. Vreede. 2021. "WORCS: A Workflow for Open Reproducible Code in Science." Edited by Tobias Kuhn. *Data Science* 4 (1): 29–49. <https://doi.org/10.3233/DS-210031>.
- Walker, Warren E., Poul Harremoës, Jan Rotmans, Jeroen P. Van Der Sluijs, Marjolein BA Van Asselt, Peter Janssen, and Martin P. Krayen von Krauss. 2003. "Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support." *Integrated Assessment* 4 (1): 5–17.

- Wicherts, J. M., C. L. Veldkamp, H. E. Augusteijn, M. Bakker, R. C. van Aert, and M. A. van Assen. 2016. “Degrees of Freedom in Planning, Running, Analyzing, and Reporting Psychological Studies: A Checklist to Avoid p-Hacking.”
- Wood, Connor M., Zachary G. Loman, Shawn T. McKinney, and Cynthia S. Loftin. 2020. “Testing Prediction Accuracy in Short-Term Ecological Studies.” *Basic and Applied Ecology* 43: 77–85. <https://doi.org/10.1016/j.baae.2020.01.003>.
- Young, Cristobal, and Sheridan A. Stewart. 2021. “Multiverse Analysis: Advancements in Functional Form Robustness.” *Cornell University, Ithaca, NY. Unpublished Working Paper*.