Preregistration Template for Modelling in Ecology and Conservation

Elise Gould, Megan Good, Libby Rumpff, TBD:workshop collaborators

21 September, 2020

# About this template

This template has been collated from a combination of literature reviews of the ‘Good Modeling Practice’ and other modeling literature, as well as a collaborative workshop at DELWP/ARI. Chris Jones has kindly volunteered one component of the VEFMAP project investigating vegetation responses to environmental flows as a testing ground for the use of preregistration templates in applied ecological modeling research.

The template is largely structured around 6 “phases” in the process of developing, evaluating and analysing a model or models in ecological applications: Problem Formulation, Define Conceptual Model, Formalise and Specify Model, Model Validation and Evaluation, Model Analysis, Model Application. These have been defined and described to some extent throughout the template. This text (usually italicised) will eventually be converted to ‘help-text’, in the form-of mouse-over pop-ups, or else prompt-text within a text-box. Other parts of this text will be pulled into a supporting diagram describing the modelling workflow stages.

## Adaptive Preregistration

The following is an excerpt from my manuscript as at 21 September 2020. It describes the procedure for how this template is meant to be used within a modeling context. I hope to write this up as a guide (including screen-shots) that will accompany the template.

Adaptive Preregistration for model-based research using GitHub We propose leveraging both the version-control and the project management features of GitHub as a tool for implementing adaptive preregistration . When a document is version-controlled using git and GitHub, each version of a document is time-stamped, and is assigned its own unique document identifier, or commit-hash. Because GitHub documents are ‘time-stamped’ the genesis of the preregistration from one version to the next is apparent. Moreover, using the diff view of a document on GitHub, exactly how a document has changed between versions is clear - additions are coloured green, deletions are coloured red, and changes within a line are highlighted . Since the preregistration document would exist within the project repository containing the data and modelling and analysis code itself, any results from analyses that cause the analysis plan to change can be linked and referenced in the updated version of the plan — thus facilitating the process of model checking. By providing a method for contrasting the actual analysis undertaken against the registered analysis, this is a mechanism for allowing reviewers to properly evaluate the preregistration against the reported analysis. It thus provides a way of marking preregistered parts of a report from non-preregistered parts of the analysis.

How it works….

GitHub is used to track amendments to the preregistration — each time an additional question is answered, or an alteration is made to the preregistration, the change should be committed to the repository. Semantic versioning is used to document each version of the preregistration. Semantic versioning is an open-source software engineering practice used to document and and communicate the development of software, in a way that allows others to depend on it . We have created our own semantic versioning scheme to be used to specifically document the incremental version of the preregistration, in tandem with GitHub’s tag and release feature, each interim preregistration is tagged and released with its semantic version number.

All questions that can be answered should be completed before any data analysis or modelling proceeds. The preregistration should be completed until a) no further detail can be supplied or b) no additional template item can be answered until preliminary investigations or analyses are undertaken (Fig X, steps 1 - 3).

Shifting from planning to execution

Before shifting from planning to the execution of the interim preregistration, each complete response in the interim preregistration should have a corresponding GitHub issue created (Figure X, step 3, document). <define GitHub issue, <https://guides.github.com/features/issues/>>. Relevant discussion between analysts about how to proceed with the intended analysis or interpreting preliminary results are tracked within an issue’s thread. Each issue is assigned a unique number. Any code or analysis outputs, such as figures, tables or other files, should be committed and tagged with the issue number, so that all analysis addressing that task is tracked within the issue’s thread (Fig X, step 4, Do) <screen-cap?>. The URL for the GitHub issue should then be added to the relevant preregistration item, the preregistration document should then be committed with the changes, and the version number of the preregistration updated in a new release.

… And back again.

As the modelling proceeds, further detail on later phases of the analysis can be iteratively updated and preregistered. Where the results of model evaluation and analysis reveal that there are problems with the model, plans can be changed, and again the next phase can be preregistered. For example, if a researcher finds that assumptions are violated or other unexpected results occur that force a change to the planned analysis, the preregistration can be updated based on the findings of those analyses. When this occurs, the major version number of the preregistration document should be incremented (Fig X steps 5 - 6). Since the findings of the analyses are tracked within the relevant issue for that template item, and the issue is recorded in the preregistration itself, the trigger for the change in the plan is made explicitly clear.

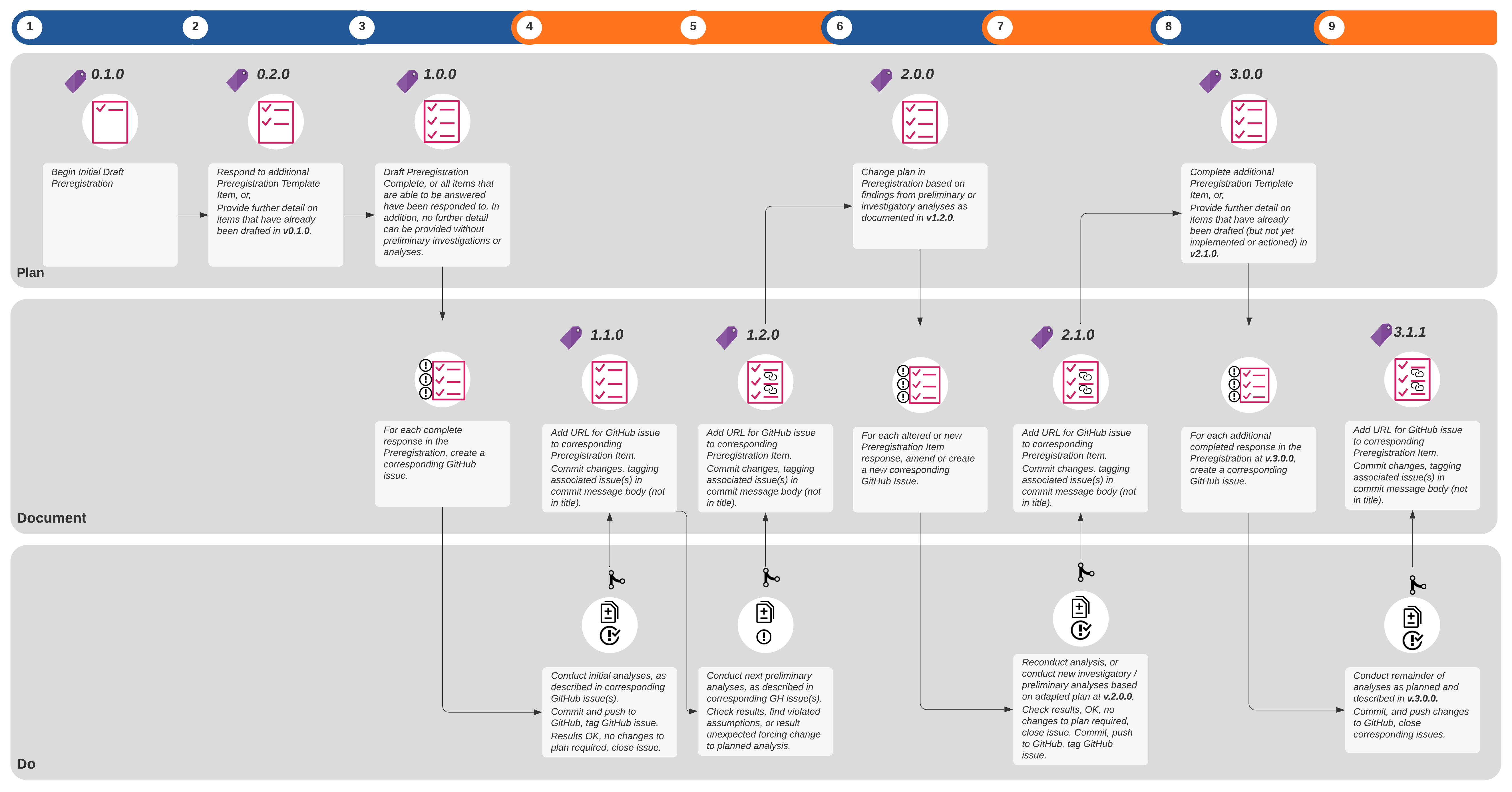
For each altered or new preregistration item, a the issue thread should either be updated, or a new issue created respectively (Fig X, step 6). Those analyses are then either reconnected based on the revised plan or the new investigatory analyses are then conducted (Fig X, step 7).

This process continues until the preregistration is complete, and the researcher can continue to execute the plan as it has been fully described in the final version of the preregistration document (Fig X, step 8 - 9).

Flexible Strategies, Preliminary Analyses or Investigations

When a researcher needs to conduct preliminary analyses or check parts of the data before committing to a particular decision about the model or analysis, they can specify a flexible strategy in their response to that preregistration item by:

* stating what needs to be known to move forward with the modeling and analysis
* explaining how the researcher will test this, and what parts of the data will be used in that analysis
* describing how the results will be interpreted, listing each potential decision and the analysis result that will trigger that decision, where possible.



Adaptive Preregistration Workflow - Research Workflow for implementing a ‘adaptive’ preregistration in a modelling context

# 1. Problem Formulation

*Specifies the decision-making context in which the model will be used or the intended scope and context of conclusions. Important components include the decision maker and stakeholders (including experts) and their view on: i) the nature of the problem or decision addressed and how the scope of the modeling tool fits within the (broader) context (i.e. model purpose; ii) the spatial and temporal scales relevant to the decision context; iii) specified desired outputs; iv) role and inclusion in model development and testing; v) whether they foresee unacceptable outcomes that need to be represented in the model (i.e. as constraints), and; vi) what future scenarios does the model need to account for (noting this may be revised later). It should also provide a summary of the domain of applicability of the model, and reasonable extrapolation limits (Grimm et al. 2014).*

## 1.1 Model Purpose

*Defining the purpose of the model is critical because the model purpose influences choices at later stages of model development (Jakeman, Letcher, and Norton 2006). Common model purposes in ecology include: gaining a better qualitative understanding of the target system, synthesising and reviewing knowledge, and providing guidance for management and decision-making (Jakeman, Letcher, and Norton 2006) (could also see Addison paper on the use of models). Note that modelling objectives are distinct from the analytical objectives of the model.*

### 1.1.1 Define model purpose and problem context

*Provide a clear statement of the modelling objectives and problem that the model seeks to illuminate. What is the purpose of the model/s? Ensure that you specify any focal taxa and study objectives, as well as any clients for whom the model is developed. Briefly outline the ecological problem and the decision-problem, including the decision-trigger and any regulatory frameworks relevant to the problem.*

## 1.2 Specify modelling context

*The scope of the model, including temporal and spatial resolutions are defined here (Mahmoud et al. 2009), and any limitations on model development analysis and flexibility should be outlined here (Jakeman, Letcher, and Norton 2006). Note that the modelling context is different from the problem context which are described in 1.3.1.*

### 1.2.1 Identify model interest group

*Identify interest groups in the modelling context. This includes clients, and end-users of the model. Who is the model for, who is involved in formulating the model? Who needs buy in? Describe the decision-making context in which the model will be used.*

***HF Comment: this is lovely, concise and informative***

### 1.2.2 Model Scope and Scale

*Determine the temporal and spatial scope of the model. Where is the boundary of the modelled system? Everything outside beyond the boundary and not crossing it is to be ignored within the domain of the model, and everything crossing the boundary is to be treated as external forcing (known/unknown), or else as model outputs (observed, or not, Jakeman et al. 2006)* **HF comment I dont understand the preceeding sentence, the concept of boundaries of the model is foreign to me. Also, is this question more to do with describing the scope and limitations of the data rather then the model?**\*. The choice of a model’s boundaries is closely linked to the choice of how finely to aggregate the behaviour within the model (Jakeman et al., 2006) - what is the intended scale and resolution of the model (temporal, spatial or otherwise)?

Explain how any key concepts or terms within problem or decision-making contexts, such as regulatory terms, will be operationalised and defined in a biologically meaningful way to answer the research question appropriately? (Should this last step go here or within define conceptual framework?. ***HF comment. I think it belongs with the conceptual framework***).

What is the intended domain of applicability of the model, what is the extent of acceptable extrapolations (Grimm)? This is relevant to model transferability… (similar to a COG statement..)\* A comment:This section feels a bit too big and confusing from my perspective and I do think you might need to move some of it to a different section to reduce the risk of misinterpretation. I’m just thinking about how I might fill this out for an imagined study. The literal boundaries of the study are usually straightforward, but the theoretic boundaries are a different and also big thing to try to conceptualise in one section. You could have a section for the spatial/temporal boundaries of the study, and another section on how far the results can be extrapolated based on the study design. This makes sense to me in how I design experiments, if I have too many confounding variables and not enough spatial and environmental replication then I’m less likely to make any broad claims. However, if I have controlled or experimentally manipulated factors then I feel safer making bigger generalisations about results.

### 1.2.3 Logistical Constraints

*Time-frame - When must the model by completed, e.g. to help make a decision?*

*What effort and resources are available for both developing the model and operating the model?* ***HF comment; as written this might be slightly confronting because you could interpret this to include writing explictly about budget constraints which might be confidential. Maybe this could be alleviated with with an example?***

*What degree of flexibility is required from the model? Might the model need to be quickly reconfigured to explore new scenarios or problems proposed by clients / managers / model-users?*

## 1.3.1 Explain analytical objectives (MAKE THIS A SEPARATE SECTION WHICH INCLUDES PERF MEASURES)

*How will the model be analysed, what analytical questions will the model be used to answer? Examples from ecological decision-making include: to compare the performance of alternative management actions under budget constraint (Fraser et al. 2017) to search for robust decisions under uncertainty (McDonald-Madden, Baxter, and Possingham 2008), to choose the conservation policy that minimises uncertainty [insert ref]. See other examples in Moallemi et al. 2019 & ask Libby and others.*

### 1.3.2 - Identify how the model will tell you about your analytic objectives

*The outcomes should speak directly to the analytical objectives identified in 1.3.1. Outcomes could be qualitative or quantitative. For example, you might be using your model in a scenario analysis to determine which management decision is associated with minimum regret or the highest likelihood of improvement.* To be honest, I don’t know how I might apply this to my work or any future works. Maybe I do not have enough knowledge of this type of research to apply it to my basic ecology research. But this isn’t to say it’s not important, I just wonder if there’s a way to make this accessible to non-desicion or modelling peeps. Like, how would I apply it to the woodland condition analysis?

## 1.4 Define candidate decisions and method of identifying relevant management actions

*Candidate decisions should be investigated and are specified a priori. Depending on the modelling context, they may be specified by stakeholders, model users or the analyst(Moallemi, Elsawah, and Ryan 2019). Describe the method used to identify relevant management actions.*

## 1.5 Specify scenarios

*Describe what processes you will use to elicit and identify relevant scenarios, e.g. literature review, structured workshops with stakeholders or decision-makers. Specify scenarios under which decisions are investigated. Scenarios should be set a priori (i.e. before the model is built, (Moallemi, Elsawah, and Ryan 2019)) and may be stakeholder-defined or driven by the judgement of the modeller or other experts (Mahmoud et al. 2009).*

# 2.0 Define Conceptual Model

*Conceptual models underpin the formal or quantitative model (Cartwright et al. 2016). The conceptual model describes the biological mechanisms relevant to the ecological problem and should capture basic premises about how the target system works, including any prior knowledge and assumptions about system processes. Conceptual models may be representeed in a variety of formats, such as influence diagrams, linguistic model block diagram or bond graphs, and these illustrate how model drivers are linked to both outputs or observed responses, and internal (state) variables (Jakeman, Letcher, and Norton, 2006).*

## 2.1 Choose elicitation and representation method

*Describe what method you will use to elicit or identify the conceptual model. Some common methods include interviews, drawings, and mapping techniques including influence diagrams, cognitive maps and Bayesian belief networks (Moon et al. 2019). (Libby, to provide link to any structured expert elicitation methods?).While it is difficult to decide and justify which method is most appropriate, however Moon et al. (2019) provide guidance addressing this methodological question. Finally, how do you intend on representing the final conceptual model? This will likely depend on the method chosen to elicit the conceptual model.*

## 2.2.1 Explain Critical Conceptual Design Decisions

*This step should list and explain the critical conceptual design decisions, including: spatial and temporal scales, selection of entities and processes, representation of stochasticity and heterogeneity, consideration of local versus global interactions, environmental drivers, etc. (Grimm et al. 2014). The influence of particular theories, concepts, or importantly, earlier models, should be explained and justified against alternative conceptual design decisions that might lead to alternative model structures (Grimm et al. 2014).*

## 2.2.2 Specify key assumptions and uncertainties underlying the model’s design.

* Describe how uncertainty and variation will be represented in this model. This includes both exogenous uncertainties affecting the system, parametric uncertainty in input data and structural / conceptual nonparametric uncertainty in the model (Moallemi, Elsawah, and Ryan 2019).\*

## 2.3 Identify predictor and response variables

*The identification and definition of primary model input variables should be driven by scenario definitions, and by the scope of the model described in the problem formulation phase (Mahmoud et al. 2009) – Identify and define system variables structures:*

* *What variables would support taking this action or making this decision?*
* *What additional variables may interact with this system (things we can’t control, but can hopefully measure)?*
* *What variables have not been measured, but may interact with the system (often occurs in field or observational studies)?*
* *What variables are indice or surrogate measures of variables that we cannot or have not measured?*
* *In what ways do we expect these variables to interact (model structures)?*

## 2.4 Define prior knowledge, data specification and evaluation

*Collect, process and prepare data available for parameterisation, determining model structure, and for scenario analysis.*

### 2.4.1 Collate available data sources that could be used to parameterise or structure the model

*Document the identity, quantity and provenance of any data that will be used to develop, identify and test the model. Describe how the data is arranged, in terms of replicates and covariates.*

**For pre-existing data: – delete as appropriate.**

* *For each dataset, is the data open or publically available? (Y/N)*
* *How can the data be accessed? Provide a link or contact as appropriate, indicating any restrctions on the use of data.*
* *Date of download, access, or future access:*
* *Describe the source of the data - what entity originally collected this data? (National Data Set, Private Organisational Data, Own Lab Collection, Other Lab Collection, External Contractor, Meta-Analysis, Expert Elicitation, Other).*
* *Codebook and meta-data. If a Codebook or other meta-data is available, please link to it here and / or upload the document(s).*
* *Prior work based on this dataset - Have you published / presented any previous work based on this dataset? Include any publications, conference presentations (papers, posters), or working papers (in-prep, unpublished, preprints) based ont this dataset you have worked on.* HF- Love this question
* *Unpublished Prior Research Activity - Describe any prior but unpublished research activity using these data. Be specific and transparent.*
* *Prior knowledge of the current dataset - Describe any prior knowledge of or interaction with the dataset before commencing this study. For example, have you read any reports or publications about this data?*

**Sampling Plan (For data you will collect) – delete as appropriate.**

* Data collection procedures - Please describe your data collection process, including how sites and transects or any other physical unit were selected and arranged. Describe any inclusion or exclusion rules, and the study timeline.
* Sample Size - Describe the sample size of your study.
* Sample Size Rationale - Describe how you determined the appropriate sample size for your study. It could include feasibility constraints, such as time, money or personnel.
* If sample size cannot be specified, specify a stopping rule - i.e. how will you decide when to terminate your data collection?

### 2.4.2 Data Processing and Preparation

*Describe any data preparation and processing steps, including manipulation of environmental layers, e.g. standardisation and geographic projection.*

*Describe how you will separate and distinguish between raw data, manipulated data, and outputs from modelling or any analyses of the model.*

### 2.4.3 Describe any data exploration or preliminary data analyses.

*In most modelling cases, it is necessary to perform preliminary analyses to understand the data and check that assumptions and requirements of the chosen modelling procedures are met. Data exploration prior to model fitting or development may include exploratory analyses to check for collinearity, spatial and temporal coverage, quality and resolution, outliers, or the need for transformations (Yates et al. 2018). Describe how you will summarise and explore your data, and explain what method you will use (graphical, tabular or otherwise) to represent your data and any analyses. Specify what you intend to find out or hope to learn from each data exploration analysis.*

### 2.4.4 Data Evaluation

*Describe how you will determine how reliable the data is for the given purpose. Ideally, model input data should be internally consistent across temporal and spatial scales and resolutions, and appropriate to the problem at hand (Mahmoud et al. 2009).* ### 2.4.5 Data Evaluation: Data reliability reporting ### 2.4.5. Data Evaluation: data reliability reporting

*Document any issues with data reliability. This is important because data quality and ecological relevance might be constrained by measurement error, inappropriate experimental design, and heterogeneity and variability inherent in ecological systems (Grimm et al. 2014).*

### 2.4.6. Data evaluation: data exclusion

*How will you determine what data, if any, will be excluded from your analyses? How will outliers be handled? Describe rules for identifying outlier data, and for excluding a site, transect, quadrat, year or season, species, trait, etc.*

### 2.4.7 Data Evaluation: Missing Data

*How will you identify and deal with incomplete or missing data?*

## 2.5 Conceptual model evaluation

*Describe how your conceptual model will be critically evaluated. Evaluation includes both the completeness and suitability of the overall model structure.*

*How will any simplifying assumptions be critically assessed (Augusiak, Van den Brink, and Grimm 2014)?*

*Explain whether this process will include consultation or feedback from a client, manager, or model user.*

# 3.0 Formalise and Specify Model

*In this section describe what quantitative methods you will use to build the model/s, explain how they are relevant to the client/manager/user’s purpose.*

## 3.1 Choose model class, framework and approach

*Describe what class or approach of model you will use and¬†explain how the choice of model class was informed by the analytical objectives of the model. Modelling approaches lie on a spectrum from correlative or phenomenological to mechanistic or process-based¬† (Yates et al. 2018); where correlative models use mathematical functions fitted to data to describe underlying processes, and mechanistic models explicitly represent processes and details of component parts of a biological system that are expected to give rise to the data (White and Marshall 2019).*

## 3.2 Choose model features and family

*All modelling approaches require the selection of model features, which conform with the conceptual model and data specified in previous steps (Jakeman 2006). The choice of model are determined in conjunction with features are selected. Usually difficult to change fundamental features of a model beyond an early stage of model development, so careful thought and planning here is useful to the modeller (Jakeman, 2006). However, if changes to these fundamental aspects of the model do need to change, please document how and why these choices were made.* HF- it’s not clear to me what a model feature is and I find this a little confusing. I wonder if we need the text describing the 3.2 heading when it’s stepped out in the subsections.

### 3.2.1 Explain how you will operationalise response variable(s)

*Specify how you will operationalise the response variables in the model. This should relate directly to the analytical and or management objectives specified during the problem formulation phase. Operationalisations could include:- the extent of a response- an extreme value- a trend- a long-term mean- a probability distribution- a spatial pattern- a time-series, - qualitative change, such as a direction of change or- the frequency, location, or probability of some event occuring. Provide a rationale for your choices, including why plausible alternatives were not chosen (Jakeman, 2006).* HF- this is great. Only slight concern is the deceptively difficult task that last sentence represents

### 3.2.2 Choose model family

*Specify which family of statistical distributions you will use in your model, and describe any transformations, or link functions. Justify your decision based on the purpose, objectives, prior knowledge and logistical constraints (Jakeman, 2006) specified in the problem formulation phase.*

*Include in your rational for selection, detail about which variables the model outputs are sensitive to, what aspects of their behaviour are important, and any associated spatial or temporal dimensions in sampling.*

### 3.2.3 Choose model features

*Specify which variables are included in the model, and the nature of their treatment (e.g. lumped/distributed, linear/non-linear, stochastic/deterministic, Jakeman, 2006).*

*Specify model structural features, including: - the functional form of interactions, - data structures - measures used to specify links, - any bins or discretisation of continuous variables.*

## 3.3 Choose approach for identifying model structure and parameters

*This refers to the process of determining the best/most efficient/parsimonious representation of the system at the appropriate scale of concern (Jakeman, 2006) that best meets the analytical objectives specified in the problem formulation phase. Approaches to finding model structure and parameters may be knowledge-supported, or data-driven (Boets et al. 2015). Model selection methods can include traditional inferential approaches such as unconstrained searches of a dataset for patterns that explain variations in the response variable, or use of ensemble-modelling methods (Barnard et al. 2019). Ensemble modelling procedures might aim to derive a single model, or a multi-model average (Yates et al. 2018). Refining actions to develop a model could include iteratively dropping parameters or adding them, or aggregating / disaggregating system descriptors, such as dimensionality and processes (Jakeman, 2006). Specify what approach and methods you will use to identify model structure and parameters.*

## 3.4 Choose estimation technique and performance criteria

*Before calibrating the model to the data, the performance criteria on which the calibration is judged are chosen. These criteria and their underlying assumptions should reflect the desired properties of the parameter estimates / structure (Jakeman, Letcher, and Norton 2006). For example, modellers might seek parameter estimates that are robust to outliers, unbiased, and yield appropriate predictive performance. Modellers will need to consider whether the assumptions of the estimation technique yielding those desired criteria are appropriate to the problem at hand. For integrated or sub-divided models, other considerations might include choices about where to disaggregate the model for parameter estimation; e.g. spatial sectioning (streams into reaches) and temporal sectioning (piece-wise linear models) (Jakeman, Letcher, and Norton 2006). Specifying performance criteria a priori is important because it involves pre-specifying how the fitted or quantitative model will be interpreted in advance, avoiding biases like confirmation bias, HARKing, and cherry-picking some performance tests from the full suite of tests undertaken.* HF- I think this paragraph is only relevant to performance criteria- doesn’t really explore estimation technique. As above though, I’m not sure that this paragraph is neessary given that it’s stepped out in the subheadings below

### 3.4.1 Choice of performance criteria

*Specify which suite of performance criteria you will use to judge the performance of the model. Examples include correlation scores, coefficient of determination, specificity, sensitivity, AUC, etcetera (Yates et al. 2018). Relate any underlying assumptions of each criterion to the desired properties of the model, and justify the choice of performance metric in relation into whether it is sensitive to the problem at hand. Explain how you will identify which model features or components are significant or meaningful.*

### 3.4.2 Choice of parameter estimation technique

HF- I think that it might be double dipping to include structure here as well as in 3.3 *Specify what technique you will use to estimate parameter values, and how you will supply non-parametric variables and/or data (e.g. distributed boundary conditions). For example, will you calibrate all variables simultaneously by optimising fit of model outputs to observations, or will you parameterise the model in a piecemeal fashion by either direct measurement, inference from secondary data, or some combination (Jakeman, 2006). Identify which variables were parameterised directly, such as by expert elicitation or prior knowledge. Specify which algorithm(s) you will use for any data-driven parameter estimation, including supervised, or unsupervised machine learning, decision-tree, K-nearest neighbour or cluster algorithms (Liu et al. 2018).*

## 3.5 Specify model assumptions and uncertainties

*Specify all assumptions and key uncertainties in the formal model. Describe what gaps exist between the model conception, and the real-world problem, what biases might this introduce and how might this impact any interpertation of the model outputs, and what implications are there on evaluating model-output to inform decisions?*

## 3.6 Specify formal model

*Once critical decisions have been made about the approach and method of model specification,¬†translate the conceptual model into the quantitative model. For data-driven and model-selection approaches that determine model structure and parameters, describe any initial model specifications and parameterisations, including for any tune-in parameters. (Should this go here or in model calibration?)*

# 4.0 Model Calibration, Model Fitting & Checking

## 4.1 Model calibration and validation scheme

*Describe how you will validate and check the calibration of the model. The model may be tested on data independent of those used to parameterise the model (external validation), or the model may be cross-validated on random sub-samples of the data used to parameterise the model (internal cross-validation) (Yates et al. 2018; Barnard et al. 2019). Justify your choice*

## 4.1.2 Describe validation data

*If partitioning data for cross-validation: describe the approach specifying the number of folds that will be created and the relative size of each. Describe how will you document and/or share the partitioned data such that the data partitioning and any subsequent modelling based on this partitioning can be computationally reproduced.*

*If using external data: describe any known differences between the training and validation datasets and the size of the validation dataset. Describe how will you document and/or share the external data such that any subsequent modelling can be computationally reproduced.*

## 4.2 Implementation verification

*What Quality Assurance measures will you take to verify the model has been correctly implemented? Tests could include syntax checking of code, and code reviews by peers. Checks for verification implementation should include i) thoroughly checking for bugs or programming errors, and ii) whether the implemented model performs as dictated by the model description (Grimm et al. 2014). Specifying up front quality assurance tests for implementation verification may help to avoid selective debugging.*

## 4.3 Model checking

*“Model Checking” goes by many names (“conditional verification”, “quantitative verification”, “model output verification” ), and refers to a series of analyses that assess a model’s performance in representing the system of interest (Conn et al. 2018). Model checking therefore heps diagnose assumption violations, and reveals where a model might need to be altered to better represent the data, and therefore system (Conn et al. 2018). Quantitative model checking diagnostics include goodness of fit, tests on residuals or errors, such as for heteroscedascity, cross-correlation, and autocorrelation (Jakeman, Letcher, and Norton 2006).*

### 4.3.1 Quantitative Model Checking

*Specify any diagnostics or tests you will use during model checking to assess a model’s performance in representing the system of interest. During this process, observed data, or data and patterns that guided model design and calibration is compared to model output in order to identify if and where there are any systematic differences. For each test, specify the criteria that will you use to interpret the outcome of the test in assessing the model’s ability to sufficiently represent the gathered data used to develop and parameterise the model.*

### 4.3.2 Qualitative Model Checking

*This step is largely informal and case-specific, but requires ‚ ‘face validation’ with model users / clients / managers who aren’t involved in the development of the model to assess whether the interactions and outcomes of the model are feasible an defensible (Grimm et al. 2014). This process could be called a ‚“laugh test” or a “pub test” and in addition to checking the model’s believability, it builds the client’s confidence in the model (Jakeman, Letcher, and Norton 2006). Face validation could include structured walk-throughs, or presenting descriptions, visualisations or summaries of model results to experts for assessment. Briefly explain how you will qualitatively check the model, and whether and how you will include users and clients in the process.*

# 5.0 Model Validation and Evaluation

HF- I think that this is covered in section 4 with descriptions of validation and checking. I’m not sure how this section is different.

*Model Validation & Evaluation - This phase consists of a suite of analyses that collectively inform the decision and whether and when a model is suitable to meet its intended purpose (Augusiak, Van den Brink and Grimm 2014). Errors in design and implementation of the model and their implication on the model output are assessed. Ideally independent data is used against the model outputs to assess whether the model output behaviour exhibits the required accuracy for the model’s intended purpose. The outcomes of these analyses build confidence in the model applications and increase understandingof model strengths and limitations. Model evaluation including model analysis and other testing should complement model checking. It should evaluate model checking, and consider over-fitting and extrapolation. The higher the proportion of calibrated, or uncertain parameters, the greater the risk that the model seems to work correctly [‚Ä¶] but for the wrong reasons" Evaluation thus complements model checking because we can rule out the chance that the model fits the calibration data well, but has not captured the relevant ecological mecahnisms of the system (Grimm). Evaluation of model outputs against external data in conjunction with the results from model checking provide information about the structural realism and therefore credibility of the model (Grimm et al. 2014).*

## 5.1 Model output corroboration

*Ideally, model outputs or predictions are compared to independent data and patterns that were not used to develop, parameterise, or verify the model. Testing against a dataset of response and predictor variables that are sptially and/or temporally independent from the training dataset¬†minimises the risk of artificially inflating model performance measures (Araújo et al. 2019).State whether you will corroborate the model outputs on external data, and please document any independent validation data in step <2.4>.It is preferable that any independent data used for model evaluation remains unknown to modellers during the process of model building, please describe the relationship modllers have to model validation data, will independent datasets be known to any modeller or analyst involved in the model building process?Although the corroboration of model outputs against an independent validation dataset is considered the ‘gold standard’ for showing that a model properly represents the internal organisation of the system ), model validation is not always possible because empirical experiments are infeasible or model users are working on rapid-response time-frames ‚Äî hence, why ecologists often model in the first place (Grimm, et al. 2014. Independent predictions can instead be tested on submodels. Alternatively, patterns in model output that are robust and seem characteristic of the system can be identified and evaluated in consultation with the literature or by experts to judge how accurate the model’s output is (Grimm et al.2014). If unable to evaluate the model outputs against independent data, explain why and explain what steps you will take to interrogate the model.*

## 5.2 Choose performance metrics and criteria

*Model performance can be quantified by a range of tests, including measures of agreement between predictions and independent observations, or estimates of accuracy, bias, calibration, discrimination refinement, resolution and skill (Araújo et al. 2019). Specify what performance measures you will use to evaluate the model and briefly explain how each test relates to different desired properties of a model’s performance.Spatial, temporal and environmental pattern of errors and variance can change the interpretation of model predictions and conservation decisions (Araújo et al. 2019), where relevant and possible, describe how you will characterise and report the spatial, temporal and environmental pattern of errors and variance. If comparing alternative models, specify what measures of model comparison or out-of-sample performance metrics will you use to find support for alternative models or else to optimise predictive ability. State what numerical threshold or qualities you will use for each of these metrics.*

## 5.3 Evaluate model assumptions

HF- I would put this up in section 4 and get rid of section 5 *Explain how you will demonstrate a lack of violation of or robustness to model assumptions. If you cannot perform quantitative tests, specify what theoretical justifications would justify a lack of violation of or robustness to model assumptions. If you cannot demonstrate or theoretically justify violation or robustness to assumptions, explain why not, and specify whether you will discuss assumption violations and their consequences for interpretation of model outputs.(Complete in follow-up iterations of the pre-registration, only if there are departures from assumptions as demonstrated in the planned tests above). The consequences of assumption violations on the interpretation of results should be assessed (Araújo et al. 2019). Specify how you will assess departures from assumptions. If assumption violations cannot be avoided, explain describe how you will explore their consequences.*

## 5.4 Model analysis

HF- I would put 5.4 and it’s sbsections up in section 4 and get rid of section 5. Also, as above, remove the paragraph under 5.4 *Uncertainty in models arises due to incomplete system understanding (which processes to include, or which interact), from imprecise, finite and sparese data measurements, and from uncertainty in input conditions and scenarios for model simulations or runs (Jakeman, Letcher, and Norton 2006). Non-technical uncertainties can also be introduced throughout the modellign process, such as uncertainties arising from issues in problem-framing, indeterminicies, and modeller / client values (Jakeman, Letcher, and Norton 2006).The purpose of model analysis is to prevent blind trust in the model by understanding how model outputs have emerged, and to ‚Äòchallenge’ the model by verifying whether the model is still believeable and fit for purpose if one or more parameters are changed (Grimm et al. 2014). Model analysis should increase understanding of the model behaviour by identifying which processes and process interactions explain characteristic behaviours of the model system. Model analysis typically consists of sensitivity analyses preceded by uncertainty analyses (Saltelli et al. 2019), and a suite of other simulation or other computational experiments. The aim of such computational experiments is to increase understanding of the model behaviour by identifying which processes and process interactions explain characteristic behaviours of the model system (Grimm et al. 2014).Uncertainty analyses and sensitivity analyses augment one another to draw conclusions about model uncertainty ‚Äî the overall model will be very uncertain if the parameters that are the most uncertain are the ones the model is most uncertain to., for instance.Because the results from a full suite of sensitivity analysis and uncertainty analysis can be difficult to interpret due to the number and complexity of causal relations examined (Jakeman, Letcher, and Norton 2006), it is useful for the analyst to relate the choice of analysis to the modelling context, purpose and analytical objectives defined in the problem formulation phase, in tandem with any critical uncertainties that have emerged during model development and testing prior to this point.*

### 5.4.1 Uncertainty Analyses

*Uncertainty can arise from different modeling techniques, response data and predictor variables (Araújo et al. 2019). Uncertainty analyses characterise the uncertainty in model outputs, and identify how uncertainty in model parameters affects uncertainty in model output, but does not identify which model assumptions are driving this behaviour (Grimm et al. 2014; Saltelli et al. 2019). Uncertainty analyses can include propagating known uncertainties through the model, or by investigating the effect of different model scenarios with different parameters and modelling technique combinations (Araújo et al. 2019), for example. It also includes characterising the output distribution, for example “by constructing it empirically from model output data points”. It could also include extracting summary statistics like the mean, median and variance from this distribution, and perhaps constructing confidence intervals on the mean (Saltelli et al. 2019).Please describe how you will characterise model and data uncertainties, as well as how you plan on summarising the results of these experiments with graphical, tabular or other devices. Relate your choice of analysis to the context and purposes of the model described in the problem formulation phase. For instance ‚Äúdiscrepancies between model output and observed output may be important for forecasting models, where cost, benefit, an risk over a substantial period must be gauged, but much less critical for decision-making or management models where the user may be satisfied with knowing that the predicted ranking order of impcats of alternative scenarios or management options is likely to be correct, with only a rough indication of their sizes" (Jakeman, Letcher, and Norton 2006).If the chosen modelling approach is able to explicitly articulate uncertainty due to data, measurements or baseline conditions, such as by providing estimates of uncertainty (typically in the form of probabilistic paraameter covariance, (Jakeman, Letcher, and Norton 2006)), specify which measure of uncertainty you will measure.*

### 5.4.2 Sensitivity Analyses

*Sensitivity analysis examines how uncertainty in model outputs can be apportioned to different sources of uncertainty in model input (Saltelli et al. 2019). Model sensitivity can be evaluated using 3 approaches: 1. deterministic sensitivity, stochastic sensitivity (variability in the model), or scenario sensitivity (effect of changes based on scenarios). Describe any sensitivity analyses you will conduct by specifying which parameters will be held constant, which will be varied, and the range and intervals of values over which those parameters will be varied. State the primary objective of each sensitivity analysis, for example, to identify which input variables contribute the most to model uncertainty so that these variables can be targeted for further data collection, or alternatively to identify which variables or factors contribute little to overall model outputs, and so can be ‚‘dropped’ from future iterations of the model (Saltelli et al. 2019).*

### 5.4.3 Other simulation experiments and robustness analyses

*Please describe any other simulation experiments and / or robustness experiments you will perform on the model.*

# 6.0 Model Application

HF- I feel that this belongs at the end of section 1 *Explain how you will operationalise any scenarios identified during the problem formulation phase, if not already done so.¬†For example, how will you operationalise any qualitative changes of interest, such as ‚Äòdeterioration’ or ‚Äòimprovement’?Specify any input conditions and relevant parameter values for initial environmental conditions and decision-variables under each scenario. Please describe any other relevant technical details of model application, such as methods for how you will implement any simulations or model projections. Specify what raw and transformed model outputs will you extract from the model simulations or projections.*

## 6.1 Choose performance criteria for scenario analysis

HF- I feel that this belongs at the end of section 1 *Describe how you will evaluate and distinguish the performance of alternative scenario outcomes. These measures should link back to the analytical objectives specified during the problem formulation phase, and might include for example, descriptive statistic measures, such as a model in the top 10th percentile of worst conditions, satisficing measures such as a minimum performance threshold for achieving some management objective, or perhaps regret measures, such as choosing a decision with minimum regret (Moallemi, Elsawah, and Ryan 2019)."Justify or otherwise explain how you chose these measures, and determined performance criteria in relation to the analytical objectives, modelling context, such as the risk attitudes of decision-makers and stakeholders within this system.*

## 6.2 Synthesis of model application results

*Specify how you will map, plot or otherwise display and synthesise the results of scenario and model analyses. Explain how will they support meaningful analysis and interpretation to aid in answering the analytical objectives and modelling context specified in the problem formulation phase.Explain how you will analyse the model analysis output to answer your analytical objectives. For instance, describe any trade-off or robustness analyses you will undertake to help evaluate and choose between different alternatives in consultation with experts or decision-makers.*

## 6.3 Consultation with Experts, Clients, Stakeholders

HF- I’m not sure how possible people will find doing this… and whether it is potentially problematic to ask people to because their clients will amost certainly change their mind about how they deal with the information at this point in a way that the researcher cant control. *How will you use the results of model application analysis to provide advice. and support to the clients / managers / stakeholders based on these findings? Describe how these results will be characterised and presented to experts, clients and stakeholders once the results from the modelling exercises have been synthesised. What role will these consultations have in determining the selection of management interventions, or future iterations of model development? For example, what key outliers or strongest examples of best or worst case scenarios should be presented to experts or decision-makers?(Do we need an item about how do we interpret the model outputs? Should there be some a priori decisions about how we interpret the model findings, and how does case 49 about the risk attitude inform how the analyst interprets the model findings?)*

# References

Araújo, MB, RP Anderson, A Márcia Barbosa, CM Beale, CF Dormann, R Early, RA Garcia, et al. 2019. “Standards for Distribution Models in Biodiversity Assessments.” *Sci Adv* 5 (1): eaat4858.

Augusiak, Jacqueline, Paul J Van den Brink, and Volker Grimm. 2014. “Merging Validation and Evaluation of Ecological Models to ‘Evaludation’: A Review of Terminology and a Practical Approach.” *Ecological Modelling* 280: 117–28.

Barnard, David M., Matthew J. Germino, David S. Pilliod, Robert S. Arkle, Cara Applestein, Bill E. Davidson, and Matthew R. Fisk. 2019. “Can’t See the Random Forest for the Decision Trees: Selecting Predictive Models for Restoration Ecology.” *Restoration Ecology*.

Boets, Pieter, Dries Landuyt, Gert Everaert, Steven Broekx, and Peter L M Goethals. 2015. “Evaluation and Comparison of Data-Driven and Knowledge-Supported Bayesian Belief Networks to Assess the Habitat Suitability for Alien Macroinvertebrates” 74: 92–103.

Conn, Paul B, Devin S Johnson, Perry J Williams, Sharon R Melin, and Mevin B Hooten. 2018. “A Guide to Bayesian Model Checking for Ecologists.” *Ecological Monographs* 9: 341–17.

Fraser, Hannah, Libby Rumpff, Jian D L Yen, Doug Robinson, and Brendan A Wintle. 2017. “Integrated Models to Support Multiobjective Ecological Restoration Decisions.” *Conservation Biology* 31 (6): 1418–27.

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: 129–39.

Jakeman, A J, R A Letcher, and J P Norton. 2006. “Ten Iterative Steps in Development and Evaluation of Environmental Models.” *Environmental Modelling & Software* 21 (5): 602–14.

Liu, Zelin, Changhui Peng, Timothy Work, Jean-Noel Candau, Annie DesRochers, and Daniel Kneeshaw. 2018. “Application of Machine-Learning Methods in Forest Ecology: Recent Progress and Future Challenges.” *Environmental Reviews* 26 (4): 339–50.

Mahmoud, Mohammed, Yuqiong Liu, Holly Hartmann, Steven Stewart, Thorsten Wagener, Darius Semmens, Robert Stewart, et al. 2009. “A Formal Framework for Scenario Development in Support of Environmental Decision-Making.” *Environmental Modelling & Software* 24 (7): 798–808.

McDonald-Madden, Eve, Peter W. J. Baxter, and Hugh P. Possingham. 2008. “Making Robust Decisions for Conservation with Restricted Money and Knowledge.” *Journal of Applied Ecology* 45 (6): 1630–8.

Moallemi, Enayat A., Sondoss Elsawah, 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*.

Saltelli, Andrea, Ksenia Aleksankina, William Becker, Pamela Fennell, Federico Ferretti, Niels Holst, Sushan Li, and Qiongli Wu. 2019. “Why so Many Published Sensitivity Analyses Are False: A Systematic Review of Sensitivity Analysis Practices.” *Environmental Modelling & Software* 114: 29–39.

White, Craig R, and Dustin J Marshall. 2019. “Should We Care If Models Are Phenomenological or Mechanistic.” *Trends in Ecology & Evolution* 34 (4): 276–78.

Yates, KL, PJ Bouchet, MJ Caley, K Mengersen, CF Randin, S Parnell, AH Fielding, et al. 2018. “Outstanding Challenges in the Transferability of Ecological Models.” *Trends Ecol. Evol. (Amst.)* 33 (10): 790–802.