

Bayesian Data Analysis and New Opportunities for Ecologists

Julia Mayr (7050054)

MSc Environmental Biology, Ecology and Natural Resource Management

Writing Assignment

April 2022

Supervised by: Prof. Dr. Merel Soons

Examined by: Dr. Kathryn Barry

Ecology and Biodiversity Research Group

Utrecht University

Abstract

Ecological systems are complex. Therefore, scientists rely on statistical tools to infer patterns and causation into their data. This has traditionally been done by employing a frequentist method, which defines probability as the relative frequency of the occurrence of a repeated event. The probability of observing similar or more extreme data than the collected dataset is then estimated given a null hypothesis. However, as new technology has emerged, more complex data have become available and conventional approaches to statistical inference have had limited success in capturing the complex structure of these datasets and therefore also in solving complicated modern issues such as climate change. A literature review shows that ecological research employing a Bayesian method of inference is gaining momentum. Importantly, in Bayesian statistics probability is defined as a degree of belief in the likelihood of an event to occur by incorporating uncertainty. Then, the probability of a hypothesis being true is evaluated conditional on available data and prior knowledge on parameters of interest. Furthermore, Bayesian inference allows to include not only uncertainty but also complex structures by allowing models to be built on multiple levels. Through a case study, the present work highlights the flexibility a Bayesian data analysis can provide, leading to more information rich results compared to conventional approaches. Additionally, the case study also underlines the importance of including multiple analytical approaches in order to reduce bias when determining the best possible explanation for the collected data and observed phenomena.

Layman's Summary

Ecology is a branch of biology, which includes the study of living organisms and how they interact with one another as well as their physical environment to form functional ecosystems. Therefore, it is widely accepted that ecological systems are complex. As a result, scientists rely on statistics to conclude patterns and causation from the data they collect. This has traditionally been done by employing a method called frequentist statistics. Frequentists define probability as the relative frequency of the occurrence of a repeated event. The probability of observing similar or more extreme data than the collected dataset is then estimated given a so-called null hypothesis. This null hypothesis could be described as a theory of no relationship between different variables. However, modern times have brought with them complicated issues, such as climate change, that involve large amounts of uncertainty but require urgent action. At the same time, new technologies have emerged, and more data have become available. In this setting, this conventional statistical approach has had limited success in either capturing the complex structure of these datasets or in solving these modern issues. In this study, I show how ecologists are starting to employ an alternative statistical approach called Bayesian statistics by reviewing the scientific literature. Then, I describe a case study where the frequentist method has encountered considerable limitations and show how adopting a Bayesian method instead can solve those issues and contribute to the acquisition of new knowledge. First, the literature review shows that ecological research has increasingly been shifting towards Bayesian statistics. Importantly, in Bayesian statistics, probability is defined as a degree of belief in the likelihood of an event to occur by incorporating uncertainty. This means that the probability of a hypothesis being true is evaluated based on available data together with background knowledge on variables of interest. Furthermore, Bayesian methods allow uncertainty, as well as large amounts of complex data to be included into its statistical models by allowing these models to be built on multiple levels. This makes Bayesian statistics much more flexible than frequentist statistics, leading to more information rich results. Yet, this also means that adopting this statistical method is time consuming for the researcher, given that it requires considerable rethinking of concepts, delving into new and more complex software, and using computationally expensive models. Nevertheless, the case study demonstrates that adopting a Bayesian approach may be worthwhile when a frequentist method cannot yield adequate results. But maybe even more importantly, the case study points out how adopting multiple statistical approaches might ultimately lead to the best and most objective explanation for the collected data.

1. Introduction

Motivated by the recognition that environmental change is a complex issue requiring urgent action (Clark, 2005; Hampton et al., 2013; Shukla et al., 2019), the field of ecology has seen a substantial shift in focus from small-scale and short-term to large-scale and long-term studies (Devictor & Bensaude-Vincent, 2016; Farley et al., 2018; Reichman et al., 2011). This shift has been facilitated amongst others by the emergence of high-throughput technologies that have allowed for a cost-effective collection of large amounts of data at high resolution (Nathan et al., 2022; Reuter et al., 2015). Other efforts include the digitization of species records, citizen-science initiatives, publicly available databases and data repositories, as well as large-scale research networks (Andrew et al., 2017; Devictor & Bensaude-Vincent, 2016; Farley et al., 2018). These advancements have led to unprecedented large amounts of ecological data being available at different spatial and temporal scales (Farley et al., 2018).

Being characterized by high natural variability, the complexity of ecological systems has long been accepted, even before the emergence of these technologies. Therefore, scientists in this field rely on statistics to interpret and conclude patterns and causation from their data (P. A. Stephens et al., 2007). There are two main methods of statistical inference used in science: frequentist and Bayesian methods. Both of those will be described briefly in turn, but for important terms and definitions, please refer to Table 1.

Table 1. Definitions of key terms used in this review. Information for this glossary was compiled from Ellison (2004), König & van de Schoot (2018), Stegmüller (2013), van de Schoot & Depaoli (2014), and Wagenmakers et al. (2008).

α Level	Often called significance level, it is used in frequentist hypothesis testing to delineate the probability of falsely rejecting a correct null hypothesis (i.e., committing a Type I error). The most commonly used significance level is 0.05.
Background Knowledge	Knowledge about population parameters available before having observed the collected data; it can be determined based on prior research, analyses of previous data, or expert opinions. This knowledge may be used to construct the prior distribution.
Bayesian Statistics	A statistical method that can be used to combine background knowledge of population parameters with current data to obtain estimates by means of the resulting posterior distribution; this method evaluates the probability of a hypothesis being true conditional on the available data, whereby probability is defined as a degree of belief in the likelihood of an event to occur by incorporating uncertainty.
Bayes' Theorem	A formula describing how to update prior probabilities for model parameters or hypotheses based on observed data. Specifically, prior distributions are updated through a multiplication with the data likelihood, yielding the full posterior distribution.
Confidence Interval	Frequentist confidence intervals are based on repeated sampling theory. Consequently, a 95% confidence interval is interpreted as 95 out of 100 replications of exactly the same experiment capturing the fixed but unknown regression coefficient. This confidence interval is constructed with reference to an assumed Gaussian sampling distribution.

Credible Interval	The Bayesian version of the confidence interval can be constructed without reference to a hypothetical sampling distribution. It can be interpreted as the (e.g., 95%) probability that the population parameter lies within the defined interval.
Frequentist Statistics	A class of statistics that estimates the probability of the observed data having occurred given a null hypothesis through point estimation; probability is defined as the relative frequency of the occurrence of a repeated event.
Informative	Used to describe a prior distribution where much background knowledge is available about model parameters. Due to large amounts of previous knowledge, the prior distribution is characterized by small variance values and hence high levels of precision.
Non-informative, flat, or vague	Used to describe a prior distribution where no prior knowledge is available about parameters of interest prior to observing the data. Flat prior distributions are characterized by very large variances.
<i>p</i> -value	In frequentist statistics, this is the probability of obtaining the observed, or more extreme results given that the null hypothesis is true.
Parameter	A fixed but unknown feature of the population that is estimated through a model either with frequentist or Bayesian methods.
Posterior	A probability distribution that is obtained once combining the prior and the likelihood in the Bayesian estimation process.
Prior	A probability distribution that can be used to capture the amount of (un)certainty in a population parameter. This distribution is then weighted by the sample data to obtain the posterior, which is used to make inference.
Weakly informative	Used to describe a prior distribution where limited prior knowledge is available about parameters of interest; the variance describing this prior distribution reflects the range of values the parameters may lie within.

Ecologists have most heavily relied on the frequentist method of inference built on null hypothesis significance testing, confidence intervals and *p*-values (P. A. Stephens et al., 2007). Under the frequentist view, probability is defined as the relative frequency of the occurrence of a repeated event. This method allows to estimate the probability for the sample data to occur given a particular hypothesis, by considering parameters to be measures of true quantities (Ellison, 2004). In this sense, a test statistic is computed to assess the consistency of the data with the null hypothesis of no difference among parameters of interest (Anderson et al., 2000; Wagenmakers et al., 2008). A cut-off level is then set to determine whether the null hypothesis is accepted or rejected (Anderson et al., 2000).

While the choice of using Bayesian or frequentist approaches has often been motivated by philosophical choice, some important criticism has emerged against the use of this statistical approach (Anderson et al., 2000; P. A. Stephens et al., 2007). The list of criticisms is comprehensive (Anderson et al., 2000), but one of the main issues that has been acknowledged about the reliance of ecologists on null hypothesis significance testing is related to poor interpretation stemming from poor understanding or improper application (V. E. Johnson, 2013; P. A. Stephens et al., 2007). These problems have mainly been attributed to the combination of two different concepts into one: the Fisherian *p*-value and the Neyman-Pearson α level, resulting in the *p*-value being falsely interpreted as a measure of evidence for the rejection of the null hypothesis as well as for statistical significance (Wagenmakers et al., 2008). Moreover, the

significance level is often chosen without being adapted based on theory or prior power tests and may therefore be regarded as an arbitrary level of significance (Anderson et al., 2000; V. E. Johnson, 2013). Hence, the reported statistical significance ought not necessarily be interpreted as an indicator of the biological relevance of the obtained result (Martínez-Abraín, 2008). It needs to be noted at this point, that rather than representing criticism of the method itself, the issue outlined above criticises the misuse of the frequentist method, which should be characterized by its objectivity and inductive nature. However, the misunderstanding and misuse of this statistical method may introduce subjectivity into the analysis insofar, that it introduces assumptions into the model that are not derived from the data alone (such as setting an arbitrary level of significance which was not adjusted for statistical power) (Ellison, 2004; Kaplan & Depaoli, 2013; Kruschke, 2013). Failure to adapt the significance level may in turn introduce bias towards finding statistically significant results when there are none (Hedges et al., 1999; V. E. Johnson, 2013). Furthermore, given that the frequentist approach views probability as the relative frequency of the occurrence of events, the null hypothesis of no difference may always be rejected, as long as the collected sample is large enough. Consequently, effects may be interpreted as statistically significant and biologically relevant, even if this significance is only an artifact of the collected data (Anderson et al., 2000; Martínez-Abraín, 2008). Similarly, no effect may be detected, just because sample sizes are too small (Hedges et al., 1999; Z. Zhang et al., 2007). Yet, the frequentist method has widely remained applied in many scientific disciplines, mainly because most students are trained in this approach only, and due to its relative ease of conduction thanks to the availability of easily accessible software packages (Clark, 2005; Kaplan & Depaoli, 2013; König & van de Schoot, 2018).

Nevertheless, as ecological research has advanced into the 21st century and the usefulness of other methods for statistical inference have been recognized, these traditional approaches to data analysis have had limited success in capturing true complexity (Clark, 2005). The increased complexity as well as large degrees of variation in both quality and uncertainty associated with the data obtained from new technologies merely underline the need for ecologists to adapt their approach to statistical analysis in order to improve inference and speed up scientific advancement (Farley et al., 2018). This stems from two important limitations associated with the frequentist approach. The first limitation derives from the fact that null hypothesis significance testing assumes no relevant prior information based on similar experiments or studies. This has led to numerous studies independently testing the same or similar null hypothesis, even though it might have been falsified before (Ellison, 2004). As a result, even though the confidence of ecologists in their results increases with the number of studies falsifying similar null hypotheses, new knowledge may be accumulated rather slowly (König & van de Schoot, 2018). The second limitation is introduced when many – often interacting – effects need to be specified for models attempting to describe

complex natural phenomena (e.g., climate change). This may lead to overfitting, which in turn might cause biased estimates and little generalizability beyond the analysed dataset (Clark, 2005; Stegmüller, 2013).

One potential solution to these problems could be a paradigm shift towards the use of a Bayesian approach to statistical inference (P. A. Stephens et al., 2007). Bayesian inference differs from frequentist inference in several ways: (1) it estimates the probability of a research hypothesis being true conditional on the sample data; (2) probability is defined as a degree of belief in the likelihood of an event by incorporating uncertainty; (3) estimates are calculated based on prior knowledge together with the collected data; and (4) not only data but also model parameters are treated as random, thereby accommodating for the inclusion of various sources of uncertainty (Clark, 2005; Ellison, 2004).

Being the cornerstone of the Bayesian approach, the use of weakly informative or informative priors in Bayesian analyses brings two important advantages with it. The first advantage is that Bayesian analyses are less dependent on sample size (Clark, 2003; van de Schoot & Depaoli, 2014; Z. Zhang et al., 2007). A Bayesian approach with informative priors might for instance be preferable for small datasets, as it improves estimates by narrowing the ranges of both standard errors and credible intervals around the mean (Z. Zhang et al., 2007). Furthermore, this inclusion of prior knowledge, which might be taken from the results of previous studies (e.g., systematic reviews, meta-analyses or similar studies), from theoretical models or preliminary data analyses allows new research to build on the knowledge gained from previous research rather than trying to independently prove a similar hypothesis (Clark, 2005; van de Schoot & Depaoli, 2014). The second advantage of including weakly informative priors is that they improve model convergence (Doll & Jacquemin, 2019; Kim et al., 2013). Given the iterative nature of Bayesian algorithms, these models need to converge on a stationary posterior distribution for their parameters. While this does not mean that convergence cannot be an issue in more traditional methods of data analysis (e.g., Y. Zhang et al., 2019), in Bayesian methods it can be improved through the incorporation of prior knowledge which can limit the model parameter space and speed up computations (van de Schoot & Depaoli, 2014; Z. Zhang et al., 2007)

Most importantly though, the real advantage of Bayesian statistics to tackle modern questions lies in what are often called hierarchical models (Clark, 2005). These models allow to account for complexity in the data by being built on multiple levels, thereby allowing scientists to include various variables without the danger of overfitting (Clark, 2005). As a result, these models also allow to include large and numerous datasets at different temporal and spatial scales while accounting for errors in variables, random effects and hidden variables (Clark, 2003). Therefore, these models may not only account for various sources of data, but they also may account for various sources of uncertainty. Some of the most commonly included sources of uncertainty include measurement and processing errors, uncertainties in the sampling design and model specification, or variability in the model parameters (Cressie et al., 2009). This distinguishing feature of

Bayesian analyses leads to more information-rich and precise estimation of parameters (König & van de Schoot, 2018). This flexibility also means that models can be estimated under a Bayesian approach which would otherwise not be assessable. In turn this can lead to scientific advances which would otherwise not have been achievable with a frequentist approach (Adkison, 2009; Clark, 2005; König & van de Schoot, 2018).

It appears that a Bayesian approach to inference has two main advantages: first, it would allow ecologists to overcome shortcomings of conventional data analysis approaches that lead to criticism; second, multi-level models should lend themselves well to utilize the full range and inherently complex structure of available data in order to gain a better insight into the urgent matters of modern times. However, this approach also brings some limitations and criticisms with it. Mainly, the Bayesian approach has been criticised for its subjectivity due to the incorporation of prior knowledge and beliefs into the analyses (Clark, 2005; Ellison, 2004). Undeniably, this may be problematic for reproducibility, particularly when the sources of or reasoning for this prior knowledge are not clearly stated (König & van de Schoot, 2018). Moreover, such complex models are typically computationally expensive, time consuming and often require some rethinking from the researcher about how scientific knowledge is acquired, updated and accumulated as well as exploring more complicated software (König & van de Schoot, 2018; McElreath, 2020).

Indeed, even though there is a growing number of researchers in this field using these methods (Wikle & Hooten, 2010), the proportion is still relatively low compared to the total number of publications (P. A. Stephens et al., 2007). Furthermore, due to the connotation of subjectivity, ecologists seem to be reluctant to incorporate weakly informative and informative priors into the analyses (Banner et al., 2020; Lemoine, 2019). This signifies, that even though they have been shifting towards a new method of inference, they have not been taking advantage of the full infrastructure that this method has to offer. To get an up-to-date overview over the state of application of Bayesian hierarchical models in the field of ecology and its advantages, this study is divided into two parts. First, by systematically reviewing the literature it attempts to answer the following questions: (1) How are ecologists implementing hierarchical Bayesian data analysis? And (2) are they using the full infrastructure that hierarchical Bayesian data analysis provides? Second, the advantages of a Bayesian approach to statistical inference are highlighted through a case study on the scale-dependence of the relationship between diversity and temporal stability.

2. Literature Review

The database Web of Science was used to search for studies within the field of ecology that have analysed their data under a Bayesian framework. However, given an immense surge in the interest of ecologists in the use of Bayesian analyses in recent years, this resulted in an unexpected number of publications. In addition, given the complexity of ecological problems in combination with the large number of different interfaces that have facilitated the implementation of Bayesian models (e.g. MCMCpack (A. D. Martin et al., 2011), rstanarm (Goodrich et al., 2020), and JAGS (Plummer, 2012)), a comparison of the used methods would have been challenging. Therefore, an alternative approach was utilized, which is described in more detail below.

2.1. Methodology

In order to facilitate the comparison of methodologies among studies, this review focused on those studies that relied on the R package *brms* (Bürkner, 2017) to conduct their statistical analyses. For this purpose, the Web of Science was searched with the following terms to find the citation for the *brms* R Package: `'((TS=(brms)) AND AU=(Bürkner)) AND PY=(2017)'`. This result was refined in multiple subsequent steps. First, the citations of the *brms* package were refined by restricting the document types to *'Articles'* in order to discard reviews and book chapters etc. Second, to restrict the resulting publications further to the field of ecology, the search was restricted to the *Web of Science Category 'Ecology'*. Finally, those results were refined by searching for the term *'hierarch* OR multilevel*'* within all fields to restrict the results to only those publications which employed a hierarchical Bayesian approach, leading to the group of articles being reviewed in the present paper.

The above-described search procedure has resulted in a manageable group of articles, which were scanned to determine whether they should be included in the review. If they were determined to be suitable for the needs of this review, their methods (and supplementary information if necessary) were scanned scrupulously to extract information on the reason the researchers opted for a Bayesian rather than a frequentist approach to data analysis as well as to extract which most common Bayesian features, such as the use of a prior distribution, a non-normal family or a Bayesian model selection approach were used in each article. These data were gathered and visualized in Microsoft Excel, version 2202.

2.2. Results & Discussion

The search for the citations of the R package *brms* (Bürkner, 2017) gave a total of 2053 returns, of which 1974 were articles. Of those 1974 articles for which researchers used the package *brms*, only 311 articles pertained to the field of ecology, and only 35 articles mentioned hierarchical models. However, a scan of these resulting articles revealed that five of those articles could not be used for this review: two publications were not openly accessible (Bonnell & Vilette, 2021; Burn et al., 2022); one publication was a statistical report (Leach et al., 2022) and did therefore not satisfy the criteria set for this review; one publication did not use *brms* for their analysis (McCabe et al., 2021); and the methods of one publication were ambiguous, so that data could not be extracted (Jakovlić et al., 2021).

2.1.1. Trends through time

Even if it cannot be assumed that all publications referring to the *brms* R package are indeed empirical studies implementing this package, the more than 100-fold increase of citations between 2017 ($n = 6$) and 2021 ($n = 855$) is by itself remarkable (Fig. 1a). It may well be interpreted as an appreciation from the scientific society for such readily available packages that make complex statistical analyses more easily applicable. Likewise, this trend remains not only true for the overall citations of this R package, but also when we focus on only those studies within the field of ecology (Fig. 1b). This shows that ecologists very well follow the general trends in science of implementing a Bayesian approach in an effort to improve statistical inference and accelerate scientific advancement in order to accommodate more complex questions. However, it also needs to be acknowledged, that the number of studies employing a hierarchical Bayesian framework remains low with a maximum of thirteen studies in 2021 (Fig. 1c). Indeed, when focusing on overall numbers as well as on numbers in 2021 which is the year in which all three categories are characterized by the largest number of publications, it becomes apparent that the field of ecology may lag compared to other fields given that ecological studies make up only approximately 16% of all publications. Furthermore, of this already relatively small proportion of ecological studies, only less than 10% have employed a hierarchical Bayesian approach.

This small proportion of studies employing the *brms* package that opted for a multilevel approach is rather surprising given that this package was specifically developed to implement Bayesian multilevel models (Bürkner, 2017). Yet, the trend of increasing numbers of publications implementing or discussing the Bayesian approach to inference in general may already reflect the efforts of the entire scientific community to cope with the always larger amounts of complex data that are becoming available.

Additionally, the fact that these trends remain the same for the field of ecology and the use of hierarchical models may also indicate that ecologists are starting to rely less on the conventional methods of statistical inference. Then, a more ready implementation of a hierarchical approach may only be a matter of time.

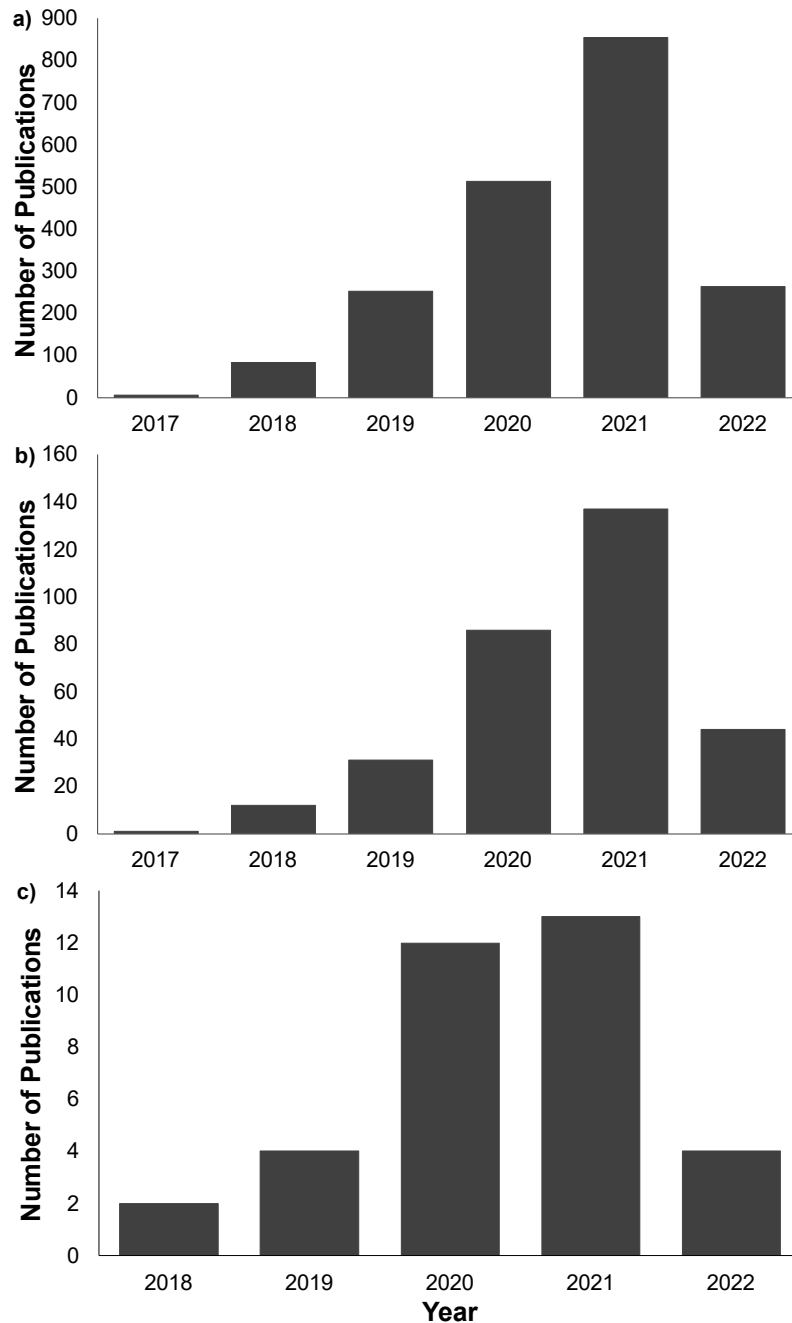


Figure 1. Number of publications through time citing the *brms* R Package (Bürkner, 2017). Panel a) shows all Publications in all fields (n=1974). Panel b) shows all those publications that fall under the Web of Science category “Ecology” (n=311). Panel c) shows those publications in the field of ecology that use a hierarchical Bayesian approach (n=35). Note the different scales of the axes on the different panels.

2.1.2. Trends in the use of Bayesian data analysis features

The choice between a frequentist and a Bayesian approach could long be attributed to a philosophical choice due to differences in how probability is viewed. Even though ecologists have long known of the Bayesian method of data analysis, they have criticised it for subjectivity through the incorporation of personal beliefs in the form of prior distributions (Ellison, 2004). Interestingly, a shift seems to have occurred, given that 90% of the articles reviewed here used weakly informative or informative prior distributions (Fig. 2). A fully Bayesian approach with informative priors has been advocated for studies with small sample sizes (e.g., van de Schoot & Depaoli, 2014; Z. Zhang et al., 2007). However, even though small sample sizes used to be a common problem in ecological studies (Hedges et al., 1999), this advantage of Bayesian analyses had seldomly been taken advantage of. Interestingly, of the articles reviewed in the present study, only one reported that the reason for using Bayesian analyses was to account for small sample sizes (Mumbanza et al., 2020). Rather, most articles opted for a Bayesian method – specifically a hierarchical Bayesian method – of inference to accommodate high levels of complexity in the data stemming from large volumes of high-resolution data (e.g., de Angeli Dutra et al., 2021; Giatzouzaki et al., 2022; Hertel et al., 2021; LoScerbo et al., 2020; Roycroft et al., 2020). Interestingly, of all articles mentioning complexity of their data as the main reason for the implementation of Bayesian methods, only one study incorporated complexity stemming from GPS data (Naidoo et al., 2018). Likewise, only one study justified the use of Bayesian methods for the incorporation of various sources of uncertainty into their analysis (Hoover et al., 2019). The most commonly mentioned source of complexity, referred to in as many as eight articles, stemmed from the incorporation of phylogenies into their analyses (e.g., Barrow et al., 2019; Craven et al., 2021; A. D. Martin et al., 2011; Nations et al., 2021). This demonstrates the versatility of this approach and also underlines that while small sample sizes might have been problematic for statistical inference in the past, this is the case anymore due to the availability of large amounts of data thanks to modern technology (Barneche et al., 2018; Hampton et al., 2013).

Furthermore, even when only some a priori information is available for model parameters (e.g., a parameter estimate is constraint between a minimum and a maximum value, or a parameter value cannot be negative) this limited knowledge might be worthwhile to include into the model in the form of a weakly informative prior (König & van de Schoot, 2018). In Bayesian analyses such a weakly informative prior can improve model convergence by limiting the model parameter space (van de Schoot & Depaoli, 2014; Z. Zhang et al., 2007). Indeed, multiple of the here reviewed articles report that the inclusion of weakly informative priors improved model convergence without significantly affecting parameter estimates (e.g., Barneche et al., 2018; Mumbanza et al., 2020; Naidoo et al., 2018).

However, the inclusion of prior distributions is not the only way to improve Bayesian model fit and parameter estimation. In *brms*, one can easily assign a distribution to the response variable by assigning a family function (Bürkner, 2017). Indeed, more than half of the reviewed studies even adopted a combination of (weakly) informative priors and non-normal model families to model the response variable in order to improve their inference (Fig. 2). Only in three articles, the adoption of a non-normal family was sufficient to ensure proper model convergence and a good fit of the model to the data even though vague default priors were used (Barrow et al., 2019; Davy et al., 2021; Snell Taylor et al., 2020).

Finally, it was assessed how often a comparative model selection approach was incorporated into the statistical analysis. In reality, model selection is a statistical approach separate from either frequentist or Bayesian methods of inference (J. B. Johnson & Omland, 2004; P. A. Stephens et al., 2007), but is often used as a complementary analysis (J. B. Johnson & Omland, 2004). This is because model selection simultaneously confronts multiple competing hypotheses with the observed data, whereby the best-fit model and therefore the best hypothesis can be identified (J. B. Johnson & Omland, 2004; Ward, 2008). Because this method explicitly acknowledges uncertainty in model selection it has been recognized mainly for its benefit of bias reduction towards a single hypothesis (P. A. Stephens et al., 2007). Given that in a Bayesian approach uncertainty is associated not only with estimates and model parameters but also with a given hypothesis (Ellison, 2004), this approach might be particularly beneficial when included into a Bayesian workflow. Indeed, here it was found that 50% of the reviewed articles implemented a model selection approach in combination with Bayesian data analysis in order to determine the model or set of parameters that was most consistent with the collected data (Fig. 2; e.g., Amirkhiz et al., 2021; Bailey & Moore, 2020; Barneche et al., 2018; Davy et al., 2021). Many different tools – called information criteria – are available for the evaluation of competing models, and which is the best criterion often depends on the purpose of a given study (Ward, 2008). For Bayesian models the preferred selection criteria appear to be the widely applicable information criterion (WAIC; Watanabe & Opper, 2010) as well as the leave-one-out cross-validation information criterion (LOOIC; Vehtari et al., 2015). Both of these criteria are calculated to estimate the out-of-sample predictive accuracy using within-sample fits (Vehtari et al., 2017). Leave-one-out cross-validation (LOO CV) predictive accuracy is estimated by re-fitting any given model to the same dataset by leaving out one datapoint at a time. *K*-fold CV is a similar but faster approach, by which the dataset is subdivided into a predefined number of so-called training sets. WAIC is again calculated in a similarly iterative way, hence of all available information criteria, these are the only ones that can truly be considered Bayesian (Vehtari et al., 2015). In fact, researchers adopting a Bayesian approach to inference appear to appreciate these fully Bayesian information criteria. Of those that implemented a model selection approach, 93% opted for one of these tools. Only in one study they opted for the Akaike Information Criterion (AIC; Nations et al., 2021).

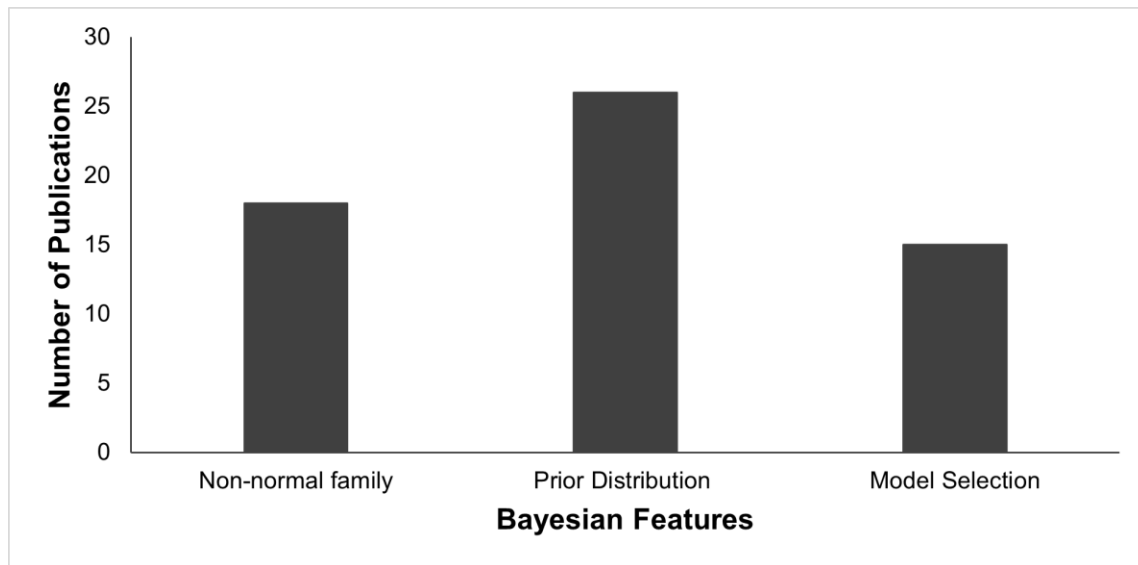


Figure 2. Three common features in Bayesian data analysis and the frequency at which they are being used in ecological publications that employ a hierarchical Bayesian approach using the *brms* package (Bürkner, 2017) (n=30).

Even though only about 20% of the reviewed articles have opted for a combination of all three identified common features of Bayesian data analysis, including all features may well improve statistical inference and a researcher's confidence into the obtained parameter and variable estimates (Krishnadas et al., 2021; J. S. Martin et al., 2020; Nations et al., 2021; Neilands et al., 2020; Vickruck et al., 2019; Wilms et al., 2021). In the following case study, I highlight how including all three of these features into a Bayesian workflow can be useful for hypothesis testing.

3. Scale-dependence of the diversity-stability relationship: a case study

3.1. Background Information

Biodiversity is being threatened at multiple spatial scales, i.e., globally as well as locally due to a suite of anthropogenic drivers such as land-use intensification, pollution, overexploitation, biological invasions, and climate change (Chase et al., 2019; Duraiappah et al., 2005; Newbold et al., 2015). In this context of global environmental and biodiversity change, the question of how biodiversity contributes to the maintenance of ecosystem functioning through time has become a fundamental challenge for ecologists

(Gonzalez et al., 2020; Isbell et al., 2017). Indeed, the importance of this question has led to many scientific studies attempting to quantify the relationship between biodiversity and the temporal stability of ecosystem functioning (e.g. Hautier et al., 2020; Wang et al., 2021; Wilcox et al., 2017; Zhang et al., 2018), hereafter simply called stability. These studies have been based on a hierarchical partitioning framework for temporal stability developed by Wang & Loreau (2014, 2016), which builds on Whittaker's (1972) concept of alpha, beta and gamma diversity. According to this framework, gamma stability is the product of alpha stability and spatial asynchrony. Here, alpha and gamma stability are the temporal stability of biomass production at the local and larger spatial scale respectively. Spatial asynchrony represents asynchronous fluctuations in biomass production among the different communities across space and through time. The different scales of diversity and stability are then linked so that alpha diversity contributes to gamma stability through alpha stability and beta diversity contributes to gamma stability through spatial asynchrony (Wang & Loreau, 2014, 2016). Yet, how area interacts with this relationship across hierarchies remains poorly known.

In a recent (unpublished) study, which investigated the scale-dependence of the relationship between species richness and the temporal stability of biomass production in a temperate grassland experiment across both hierarchical and spatial scales, some important new insights were gained. One of the main results was that even at the small spatial scales considered in the study, area can have considerable influence on the conclusions that can be drawn about the strength of the effect of diversity on stability. This finding supports previous results obtained by Zhang et al. (2018), and together, these studies call attention to account for the effect of scale when management decisions are made. This is important, because while management decisions are most often taken at regional and landscape scales (Gonzalez et al., 2020), the knowledge on the relationship between biodiversity and the stability of ecosystem functioning is based on theory and estimates made at local scales (Caballero-Vázquez & Vega-Cendejas, 2012), leading to a mismatch between the scale of knowledge and the scale of management.

Even though these results were remarkable, there were some important methodological limitations concerning the data analysis within the latest study (for the methodology, refer to *Appendix A3*), which in turn limit the confidence in the obtained results. Firstly, because experimental data was combined with landscape simulations, a large dataset with 12958 entries was obtained for this study. Furthermore, this practice of combining experimental data with simulation led to unbalanced data with unequal sample sizes among different areas. Since frequentist statistical inference was utilized without calculating effect sizes or adjusting the significance level, large sample sizes might have led to Type I errors of rejecting a correct null hypothesis (EFSA Scientific Committee, 2011). Secondly, the simulation in combination with the chosen metric of beta diversity introduced patterns into the data, whereby different planted diversity levels differed considerably in the ranges of both beta diversity and spatial asynchrony which could be obtained for each level. To account for that, linear mixed effects models were employed with planted species

diversity and area as random effects, even though these levels are not random. And lastly, in order to avoid singularity or convergence issues, and to allow for different slopes at different areas, both diversity and area were included as both fixed and random effects and hence had to be kept as continuous predictors, even though both were intended as factors.

This study indeed highlights, that as the quantity and complexity of the data increases, the frequentist approach becomes limiting. In an attempt to improve inference, a Bayesian approach was then employed, for which the statistical methods and results will be described below (note that the description of the methods used for the frequentist analysis can be found in *Appendix A3*, section 3).

3.2. Statistical Analysis

To determine the relationships between the different components of Wang & Loreau’s (2014; 2016) hierarchical partitioning framework, as well as the scale dependence of these relationships, a series of independent Bayesian analyses were performed. Generalized linear models were generated using the *brms* package in R (Bürkner, 2017), which relies on the Hamiltonian Monte Carlo (HMC) sampler Stan (Carpenter et al., 2017). Separate analyses were performed for every combination of explanatory and response variable (Table 1), leading to seven analyses, whereby they can be grouped into two distinct groups: 1) models which use alpha diversity as a covariate; and 2) models which use alpha diversity as a grouping variable.

Table 1. Various response variables and their associated explanatory variables associated with the models used in this study. Their pairing is based on the framework developed by Wang & Loreau (2014; 2016). Indirect explanatory variables are called such because they do not directly affect the associated response variable, but through another variable, i.e., beta diversity influences gamma stability indirectly through spatial asynchrony.

Response Variable	Explanatory Variable
Alpha Stability	Alpha Diversity
Spatial Asynchrony	Beta Diversity
Gamma Stability	Indirect: Alpha Diversity, Beta Diversity Direct: Alpha Stability, Spatial Asynchrony, Gamma Diversity

The first group was comprised of the two analyses with alpha stability and gamma stability as response variables and alpha diversity and area as explanatory variables. For each analysis, 12 models were compared to select the best-fit model among competing hypotheses of significant interactions or hierarchical structures:

1. a null model without interactions or random effects, where alpha diversity was used as a covariate and area was used as a factor
2. a model without random effects but with an interaction between alpha diversity and area
3. a hierarchical random intercept model without interactions, but with area as a grouping variable
4. a hierarchical random slopes model with area as a grouping variable and random slopes for alpha diversity
- 5-8. models 1-4 with a log-normal family in an attempt to improve model convergence as well as the fit of the model to the data; this family function was chosen based on previous knowledge of a lognormal distribution of the response variable stemming from an exploratory analysis as well as previous studies, which describe stability metrics as being log-normally distributed (Hautier et al., 2020; Wang et al., 2021; Wang & Loreau, 2014).
- 9-12. models 5-8 including a weakly informative Half-Cauchy(0,2) prior on all population-level (fixed) parameters to determine whether the inclusion of such a prior could further improve convergence and model fit. This prior was chosen because the response variables were limited to non-negative, non-zero, low continuous values resulting mainly from the simulation procedure;

The second group was comprised of the remaining five analyses, where the explanatory variables are used as covariates and area is used as a factor. Again, for each analysis 12 models were compared to select the best-fit model among competing hypotheses of significant interactions and/or different hierarchical structures:

1. a hierarchical null model without interactions and alpha diversity as a grouping variable
2. a hierarchical model with an interaction between the covariate and area, and alpha diversity as a grouping variable
3. a hierarchical random slopes model without interactions, alpha diversity as a grouping variable and random slopes for area
4. a hierarchical model with an interaction between the covariate and area, alpha diversity as a grouping variable and random slopes for area
- 5-8. models 1-4 with a log-normal family
- 9-12. models 5-8 including a weakly informative Half-Cauchy(0,2) prior on all population-level (fixed) parameters

For all models, the intercept parameter was removed, so that posterior distributions were directly estimated for each group (Bürkner, 2017). The models were run with 4 chains for 2000 iterations (1000 as

warm-up), and default Student-t priors for standard deviations of group-level (random) parameters. Where not specified, models were run with default priors for population-level parameters and a Gaussian distribution. To ensure there were no divergent transitions, the ‘*adapt_delta*’ parameter was increased to 0.999 and the ‘*max_treedepth*’ parameter was increased to 15 where necessary. Model convergence was assessed through effective sample sizes (ESS) the Gelman-Rubin diagnostic of $R\text{-hat} < 1.05$, and through visual inspection of traceplots (see *Appendix A2*). Model validity and fit to observed data were assessed through visual inspection of posterior predictive checks (see *Appendix A2*). Posterior distributions of the mean estimates of interest were compared using the *compare_levels* function in the R package *Tidybayes* (Kay, 2021)

A model selection approach was implemented to determine the model structure and combination of parameters that best explained each relationship. Within each analysis, models were compared using leave-one-out cross-validation information criterion (LOOIC) minimization (Vehtari et al., 2017) calculated with the *loo* package in R (Vehtari et al., 2021). All statistical analyses were performed in R 4.1.3. (R Core Team, 2022).

3.3. Results and Discussion

To keep the comparison between frequentist and Bayesian results manageable and as simple as possible as well as to highlight the benefit of including model selection into a Bayesian workflow, here I will focus on only one of the models. The selected model, where gamma stability (i.e., temporal stability at the largest spatial scale) is dependent on alpha diversity (i.e., species richness of the composite communities) and area well describes the key conclusion from the discussed study. The key conclusion being referred to here is that area significantly influences the conclusions that can be drawn about the relationship between diversity and stability.

From the selected linear mixed effects model (based on AIC) performed through a frequentist approach, it is suggested that even though gamma stability does not significantly increase with alpha diversity, area interacts significantly with alpha diversity to explain gamma stability (Table 2). The negative estimate associated with this interaction suggests that increasing area decreases the effect of planted species richness on gamma stability. Indeed, it can be seen that the slope of this relationship slightly flattens as area is increased (Fig. 3).

Table 2. Results from the linear mixed effects model performed on log-transformed data for gamma stability as the dependent variable. P-values are calculated automatically for models created with the ‘lme()’ function. Effects are significant at $P < 0.05$.

Fixed Effect	Estimate	Standard Error	DF	t-Value	p-Value	95% CI Lower	Upper
Alpha Diversity	0.1081	0.0194	1	5.5604	0.1133	-0.1389	0.3551
Area	0.0739	0.0139	12953	5.3089	0.0000	0.0466	0.1012
Alpha Diversity x Area	-0.0366	0.0077	12953	-4.7685	0.0000	-0.0517	-0.0216

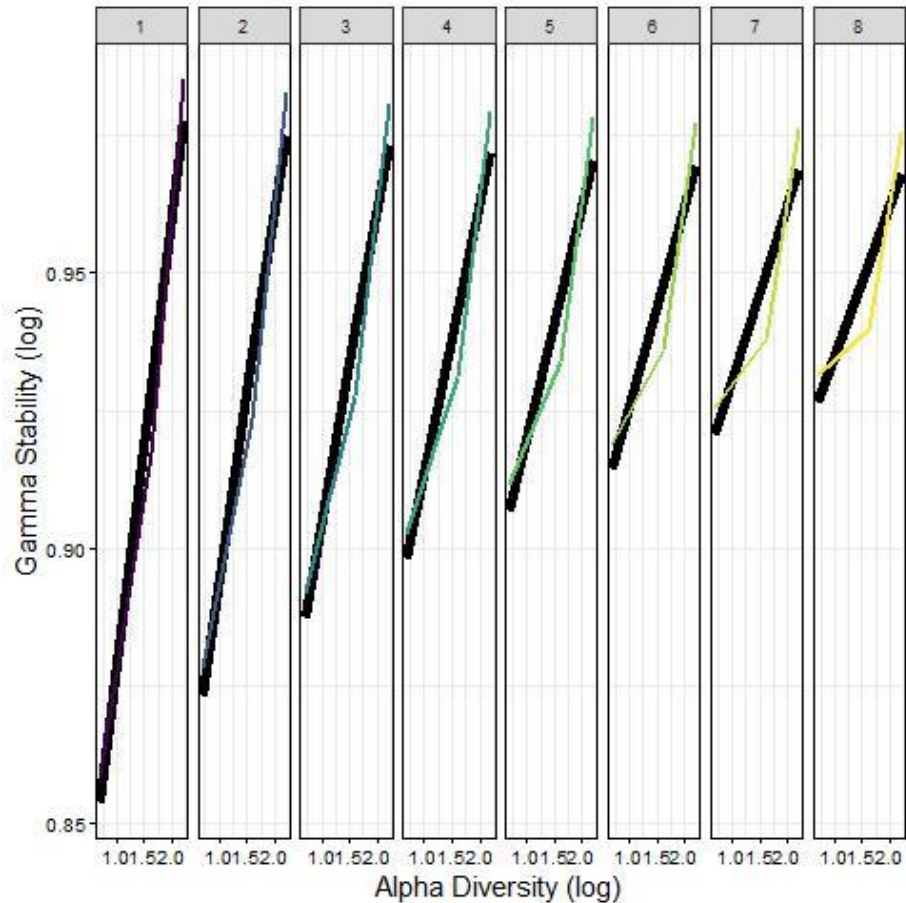


Figure 3. Results on the relationship between alpha diversity, gamma diversity and area obtained from a frequentist random intercept and slope model on the log-transformed data. Fixed effects are plotted in black. Colours represent the slopes at different areas. Facets represent different areas, ranging from one single to eight aggregated communities.

For the Bayesian counterpart, the LOOIC selected for a simple Bayesian model with alpha diversity as a covariate, area as a factor and an interaction between these two explanatory variables. The selected model fit with a log-normal family and a weakly informative Half-Cauchy prior showed excellent convergence ($R\text{-hat} = 1$) and $ESS > 1000$.

The result remains comparable to that of the frequentist analysis described above, with a significant interaction between alpha diversity and area. However, for the result from the Bayesian model we can see an even more pronounced decline in slope of the alpha diversity – gamma stability relationship as area is increased (Fig.4). Interestingly, through the Bayesian approach it could be discerned that area has a significantly negative effect on the alpha diversity – gamma stability relationship, but only at the largest two landscapes (Fig. 4), since the 95% CI for those interaction terms do not overlap with zero (see *Appendix A2 Table 1*). This result also highlights that as area is increased, the relationship changes due to an increase in the average estimate for gamma stability (Fig. 4, Fig. 5, *Appendix A2 Table 1*). Finally, pairwise contrasts show that at the smallest area – i.e., the single community of 1.2m² – the alpha diversity – gamma stability relationship is significantly different from all other areas (Fig. 5). Similarly, the largest area – i.e., at a landscape of eight aggregated communities and a size of 9.6m² – is significantly different from all other areas, except for the next largest area of seven communities. No difference was found in the diversity – stability relationship between areas of similar sizes (Fig. 5).

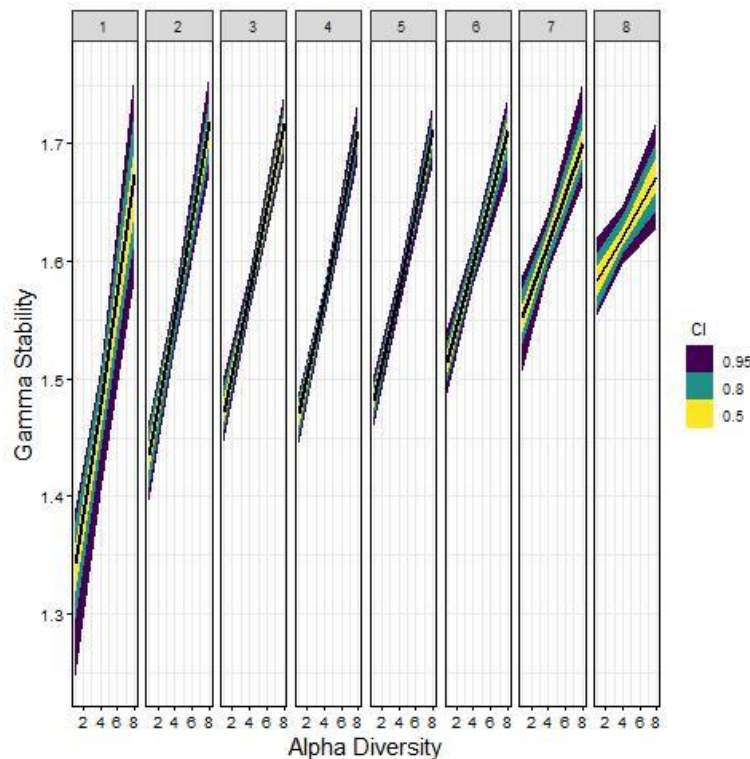


Figure 4. Results on the relationship between alpha diversity, gamma diversity and area obtained from a simple Bayesian ANCOVA. 50%, 80% and 95% credible intervals are displayed in different colours. Facets represent different areas, ranging from one single to eight aggregated communities.

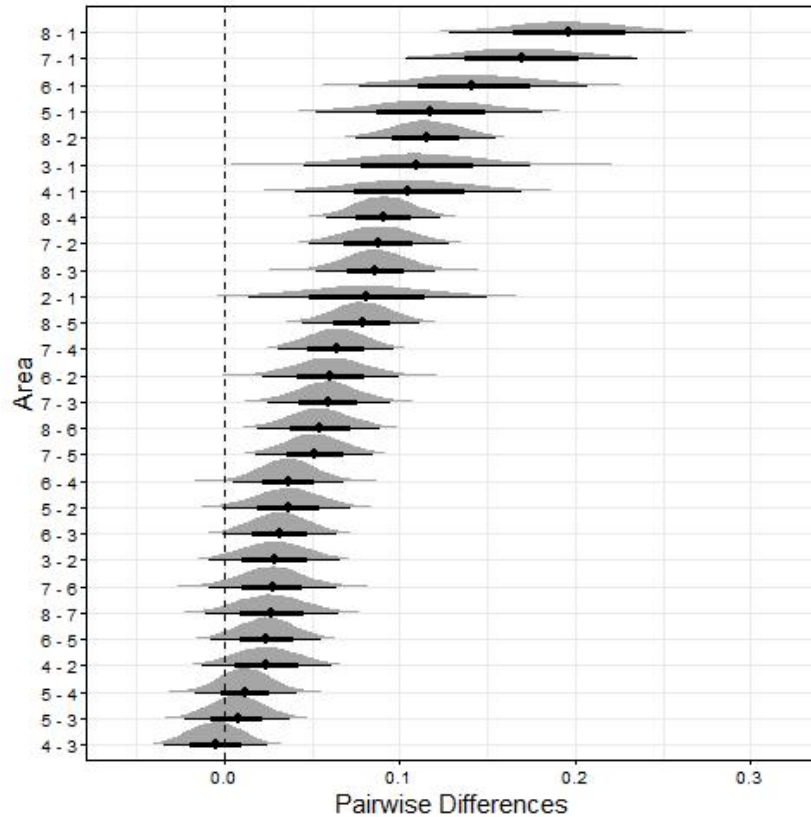


Figure 5. Density plots of contrasts between mean gamma stability values for the different area. Points are medians and narrow bars are 95% credible intervals. If the distributions of the differences do not overlap zero at the 95% credible level, then it can be said there is a significant difference between the two values.

On a first glance, it appears that the obtained results remain similar between a frequentist and a Bayesian approach. However, upon closer inspection, it becomes clear that there are some important differences. First, the mean estimates differ between analyses (Fig. 3, Fig. 4). One might attribute this to the fact that data were log-transformed for the frequentist analysis in order to meet the assumption of normality. However, a discrepancy remains when the result is visualized with the back-transformed data (see *Appendix A2*, Fig. 4). Therefore, one possible explanation that this underestimation of the frequentist analysis might be due to the fact that even though data were transformed to improve normality and variance functions were applied to adjust for heteroscedasticity, the models' assumptions were still not fulfilled, thus introducing bias (see *Appendix A2*, Fig. 3).

Moreover, because fixed effects needed to be kept as continuous variables to ensure model convergence and random effects cannot be compared in a frequentist analysis, one can only conclude that there is a significant interaction between area and alpha diversity, but estimates cannot be compared among areas. This is not a problem for the Bayesian analysis, where pairwise differences can be assessed and hence

it can be determined for which areas the relationship between alpha diversity and gamma stability was significantly different. This would also not have been possible had a hierarchical model been selected with area as a grouping variable, which further underlines the strength of Bayesian data analysis (McElreath, 2020). As a result, it could be said that the Bayesian analysis yielded similar results, but these results contained more information (König & van de Schoot, 2018) and their more intuitive interpretation provides more confidence to the researcher to make his/her claims.

Finally, this case study also shows the benefits of implementing a model selection approach complementary to the Bayesian analysis. Due to the limitations encountered when utilizing a frequentist approach (please refer to section 5.1) one might have directly implemented a more complex, hierarchical Bayesian model. Instead, a simple model was selected as the best-fit model for the observed data. This again shows, how important it is not to discard multiple possible hypotheses that could explain the patterns in the collected data. It also highlights how a Bayesian approach is less dependent on sample sizes – large or small (Barneche et al., 2018; Naidoo et al., 2018; van de Schoot & Depaoli, 2014; Z. Zhang et al., 2007) – so that one can confidently infer the deducted results without fear of having falsely rejected a correct null hypothesis.

4. Conclusion

This study has shown that even though it may require rethinking of some widely accepted concepts about probability, and be more computationally expensive and time consuming than the traditionally adopted frequentist method (König & van de Schoot, 2018), Bayesian inference has certain advantages. These include more flexibility, the ability to accommodate more complex data structures and a more intuitive interpretation of the results (Clark, 2005; Cressie et al., 2009; Ellison, 2004; König & van de Schoot, 2018). The growing number of ecological studies not only employing a Bayesian approach in general but more specifically a hierarchical Bayesian approach indicates that ecologists are increasingly appreciating the flexibility this method of inference has to offer, also thanks to the emergence of easily accessible and easily applicable software. This is especially important given that frequentist methods appear to be often misunderstood and misused, and confidence intervals are wrongly interpreted in a Bayesian way (Ellison, 2004; van de Schoot & Depaoli, 2014). However, statistics are only a tool that helps ecologist to infer patterns and causal relationships from their data. Therefore, no single approach is perfect or implies biological processes. Rigorous research design, the choice of reliable measures, appropriate analytic procedures and finally careful judgment of the obtained results are crucial to ensure the success of any

chosen method of inference (Kruschke et al., 2012; P. A. Stephens et al., 2007). Furthermore, both the review and case study presented here suggest that a combination of different approaches might be favourable to a single method approach in order to limit bias. In this sense, another important aspect about Bayesian inference is that p -values are forgone, and researchers are often more aware of the complexity of the ongoing computations and the need to carefully document sources of prior knowledge, leading to a higher degree of reproducibility.

5. Acknowledgments

I would like to thank my daily supervisor, Prof. Dr. Merel Soons, for her guidance and especially her comprehensive feedback on how to improve on earlier versions of this thesis. I would also like to express my gratitude for the help of Dr. Kathryn Barry whose advice and help during the Bayesian reanalysis was invaluable and without whom the realization of the thesis in its current form would not have been possible. Finally, I would also like to thank Dr. Yann Hautier for sharing his thoughts during the initial phases of this assignment. I would like to thank the three of them collectively for their guidance, emotional support and understanding during the entire time that I worked with them.

6. References

- Adams, F. V., Arseneau-Robar, T. J. M., Bonnell, T. R., Stead, S. M., & Teichroeb, J. A. (2021). Temporal patterns in the social network of core units in Rwenzori Angolan colobus monkeys: Effects of food availability and interunit dispersal. *Ecology and Evolution*, *11*(7), 3251–3263. <https://doi.org/10.1002/ece3.7274>
- Adkison, M. D. (2009). Drawbacks of complex models in frequentist and Bayesian approaches to natural-resource management. *Ecological Applications*, *19*(1), 198–205. <https://doi.org/10.1890/07-1641.1>
- Amirkhiz, R. G., Dixon, M. D., Palmer, J. S., & Swanson, D. L. (2021). Investigating niches and distribution of a rare species in a hierarchical framework: Virginia’s Warbler (*Leiothlypis virginiae*) at its northeastern range limit. *Landscape Ecology*, *36*(4), 1039–1054. <https://doi.org/10.1007/s10980-021-01217-7>
- Anderson, D. R., Burnham, K. P., & Thompson, W. L. (2000). Null Hypothesis Testing: Problems, Prevalence, and an Alternative. *The Journal of Wildlife Management*, *64*(4), 912–923. <https://doi.org/10.2307/3803199>
- Andrew, C., Heegaard, E., Kirk, P. M., Bässler, C., Heilmann-Clausen, J., Krisai-Greilhuber, I., Kuyper, T. W., Senn-Irlet, B., Büntgen, U., Diez, J., Egli, S., Gange, A. C., Halvorsen, R., Høiland, K., Nordén, J., Rustøen, F., Boddy, L., & Kauserud, H. (2017). Big data integration: Pan-European fungal species observations’ assembly for addressing contemporary questions in ecology and global change biology. *Fungal Biology Reviews*, *31*(2), 88–98. <https://doi.org/10.1016/j.fbr.2017.01.001>
- Bailey, C. J., & Moore, J. W. (2020). Resource pulses increase the diversity of successful competitors in a multi-species stream fish assemblage. *Ecosphere*, *11*(9). <https://doi.org/10.1002/ecs2.3211>
- Banner, K. M., Irvine, K. M., & Rodhouse, T. J. (2020). The use of Bayesian priors in Ecology: The good, the bad and the not great. *Methods in Ecology and Evolution*, *11*(8), 882–889. <https://doi.org/10.1111/2041-210X.13407>
- Barneche, D. R., Burgess, S. C., & Marshall, D. J. (2018). Global environmental drivers of marine fish egg size. *Global Ecology and Biogeography*, *27*(8), 890–898. <https://doi.org/10.1111/geb.12748>
- Barrow, L. N., McNew, S. M., Mitchell, N., Galen, S. C., Lutz, H. L., Skeen, H., Valqui, T., Weckstein, J. D., & Witt, C. C. (2019). Deeply conserved susceptibility in a multi-host, multi-parasite system. *Ecology Letters*, *22*(6), 987–998. <https://doi.org/10.1111/ele.13263>
- Bonnell, T. R., & Vilette, C. (2021). Constructing and analysing time-aggregated networks: The role of bootstrapping, permutation and simulation. *Methods in Ecology and Evolution*, *12*(1), 114–126. <https://doi.org/10.1111/2041-210X.13351>
- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1–28.
- Burn, D., Matthews, S., Pisapia, C., Hoey, A. S., & Pratchett, M. S. (2022). Changes in the incidence of coral injuries during mass bleaching across Australia’s Coral Sea Marine Park. *Marine Ecology Progress Series*, *682*, 97–109. <https://doi.org/10.3354/meps13935>
- Caballero-Vázquez, J. A., & Vega-Cendejas, M. E. (2012). Spatial patterns of diversity at local and regional scales in a tropical lagoon. *Neotropical Ichthyology*, *10*(1), 99–108. <https://doi.org/10.1590/S1679-62252012000100010>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, *76*(1). <https://doi.org/10.18637/jss.v076.i01>
- Chase, J. M., McGill, B. J., Thompson, P. L., Antão, L. H., Bates, A. E., Blowes, S. A., Dornelas, M., Gonzalez, A., Magurran, A. E., Supp, S. R., Winter, M., Bjorkman, A. D., Bruelheide, H., Byrnes, J. E. K., Cabral, J. S., Elahi, R., Gomez, C., Guzman, H. M., Isbell, F., ... O’Connor, M. (2019). Species richness change across spatial scales. *Oikos*, *128*(8), 1079–1091. <https://doi.org/10.1111/oik.05968>

- Clark, J. S. (2003). Uncertainty in population growth rates calculated from demography: The hierarchical approach. *Ecology*, 84(6), 1370–1381.
- Clark, J. S. (2005). Why environmental scientists are becoming Bayesians. *Ecology Letters*, 8(1), 2–14. <https://doi.org/10.1111/j.1461-0248.2004.00702.x>
- Conesa, M. À., Muir, C. D., Molins, A., & Galmés, J. (2020). Stomatal anatomy coordinates leaf size with Rubisco kinetics in the Balearic *Limonium*. *AoB PLANTS*, 12(1), plz050. <https://doi.org/10.1093/aobpla/plz050>
- Craven, D., Weigelt, P., Wolkis, D., & Kreft, H. (2021). Niche properties constrain occupancy but not abundance patterns of native and alien woody species across Hawaiian forests. *Journal of Vegetation Science*, 32(3), e13025. <https://doi.org/10.1111/jvs.13025>
- Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., & Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, 19(3), 553–570. <https://doi.org/10.1890/07-0744.1>
- Davy, C. M., Squires, K., & Zimmerling, J. R. (2021). Estimation of spatiotemporal trends in bat abundance from mortality data collected at wind turbines. *Conservation Biology*, 35(1), 227–238. <https://doi.org/10.1111/cobi.13554>
- de Angeli Dutra, D., Fecchio, A., Braga, É. M., & Poulin, R. (2021). Haemosporidian taxonomic composition, network centrality and partner fidelity between resident and migratory avian hosts. *Oecologia*, 197(2), 501–509. <https://doi.org/10.1007/s00442-021-05031-5>
- Devictor, V., & Bensaude-Vincent, B. (2016). From ecological records to big data: The invention of global biodiversity. *History and Philosophy of the Life Sciences*, 38(4). <https://doi.org/10.1007/s40656-016-0113-2>
- Doll, J. C., & Jacquemin, S. J. (2019). Bayesian Model Selection in Fisheries Management and Ecology. *Journal of Fish and Wildlife Management*, 10(2), 691–707. <https://doi.org/10.3996/042019-JFWM-024>
- Duraiappah, A. K., Naeem, S., Agardy, T., Ash, N. J., Cooper, H. D., Diaz, S., Faith, D. P., Mace, G., McNeely, J. A., Mooney, H. A., Oteng-Yeboah, A. A., Pereira, H. M., Polasky, S., Prip, C., Reid, W. V., Samper, C., Schei, P. J., Scholes, R., Schutysse, F., & Van Jaarsveld, A. (2005). *Ecosystems and human well-being: Biodiversity synthesis; a report of the Millennium Ecosystem Assessment*. World Resources Institute.
- EFSA Scientific Committee. (2011). Guidance on the risk assessment of the application of nanoscience and nanotechnologies in the food and feed chain. *EFSA Journal*, 9(5), 2140. <https://doi.org/10.2903/j.efsa.2011.2140>
- Ellison, A. M. (2004). Bayesian inference in ecology. *Ecology Letters*, 7(6), 509–520. <https://doi.org/10.1111/j.1461-0248.2004.00603.x>
- Farley, S. S., Dawson, A., Goring, S. J., & Williams, J. W. (2018). Situating Ecology as a Big-Data Science: Current Advances, Challenges, and Solutions. *BioScience*, 68(8), 563–576. <https://doi.org/10.1093/biosci/biy068>
- Gabry, J., & Mahr, T. (2020). R Package bayesplot: Plotting for Bayesian models. *R Package Version*, 1(2).
- Gade, M. R., Connette, G. M., Crawford, J. A., Hocking, D. J., Maerz, J. C., Milanovich, J. R., & Peterman, W. E. (2020). Predicted alteration of surface activity as a consequence of climate change. *Ecology*, 101(11), e03154. <https://doi.org/10.1002/ecy.3154>
- Giatzouzaki, D., Delhaye, G., & Meerts, P. J. (2022). Trait divergence between endemic plants of Aegean islands and their widespread congeners. *Plant Ecology*. <https://doi.org/10.1007/s11258-022-01224-x>
- Gonzalez, A., Germain, R. M., Srivastava, D. S., Filotas, E., Dee, L. E., Gravel, D., Thompson, P. L., Isbell, F., Wang, S., Kéfi, S., Montoya, J., Zelnik, Y. R., & Loreau, M. (2020). Scaling-up biodiversity-ecosystem functioning research. *Ecology Letters*, 23(4), 757–776. <https://doi.org/10.1111/ele.13456>
- Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). rstanarm: Bayesian applied regression modeling via Stan. *R Package Version*, 2(1).

- Gordijn, P. J., & O'Connor, T. G. (2021). Multidecadal effects of fire in a grassland biodiversity hotspot: Does pyrodiversity enhance plant diversity? *Ecological Applications*, *31*(6), e02391. <https://doi.org/10.1002/eap.2391>
- Hampton, S. E., Strasser, C. A., Tewksbury, J. J., Gram, W. K., Budden, A. E., Batcheller, A. L., Duke, C. S., & Porter, J. H. (2013). Big data and the future of ecology. *Frontiers in Ecology and the Environment*, *11*(3), 156–162. <https://doi.org/10.1890/120103>
- Hautier, Y., Zhang, P., Loreau, M., Wilcox, K. R., Seabloom, E. W., Borer, E. T., Byrnes, J. E. K., Koerner, S. E., Komatsu, K. J., Lefcheck, J. S., Hector, A., Adler, P. B., Alberti, J., Arnillas, C. A., Bakker, J. D., Brudvig, L. A., Bugalho, M. N., Cadotte, M., Caldeira, M. C., ... Wang, S. (2020). General destabilizing effects of eutrophication on grassland productivity at multiple spatial scales. *Nature Communications*, *11*(1), 5375. <https://doi.org/10.1038/s41467-020-19252-4>
- Hedges, L. V., Gurevitch, J., & Curtis, P. S. (1999). The meta-analysis of response ratios in experimental ecology. *Ecology*, *80*(4), 1150–1156.
- Hertel, A. G., Royauté, R., Zedrosser, A., & Mueller, T. (2021). Biologging reveals individual variation in behavioural predictability in the wild. *Journal of Animal Ecology*, *90*(3), 723–737. <https://doi.org/10.1111/1365-2656.13406>
- Hoover, A. L., Liang, D., Alfaro-Shigueto, J., Mangel, J. C., Miller, P. I., Morreale, S. J., Bailey, H., & Shillinger, G. L. (2019). Predicting residence time using a continuous-time discrete-space model of leatherback turtle satellite telemetry data. *Ecosphere*, *10*(3), e02644. <https://doi.org/10.1002/ecs2.2644>
- Isbell, F., Gonzalez, A., Loreau, M., Cowles, J., Díaz, S., Hector, A., Mace, G. M., Wardle, D. A., O'Connor, M. I., Duffy, J. E., Turnbull, L. A., Thompson, P. L., & Larigauderie, A. (2017). Linking the influence and dependence of people on biodiversity across scales. *Nature*, *546*(7656), 65–72. <https://doi.org/10.1038/nature22899>
- Jakovlić, I., Zou, H., Chen, J.-H., Lei, H.-P., Wang, G.-T., Liu, J., & Zhang, D. (2021). Slow crabs - fast genomes: Locomotory capacity predicts skew magnitude in crustacean mitogenomes. *Molecular Ecology*, *30*(21), 5488–5502. <https://doi.org/10.1111/mec.16138>
- Johnson, J. B., & Omland, K. S. (2004). Model selection in ecology and evolution. *Trends in Ecology & Evolution*, *19*(2), 101–108. <https://doi.org/10.1016/j.tree.2003.10.013>
- Johnson, V. E. (2013). Revised standards for statistical evidence. *Proceedings of the National Academy of Sciences*, *110*(48), 19313–19317. <https://doi.org/10.1073/pnas.1313476110>
- Kaplan, D., & Depaoli, S. (2013). Bayesian statistical methods. *Oxford Handbook of Quantitative Methods*, 407–437.
- Kay, M. (2021). *Tidybayes: Tidy data and geoms for Bayesian models. R package version 3.0.0.*
- Kim, S.-Y., Suh, Y., Kim, J.-S., Albanese, M. A., & Langer, M. M. (2013). Single and Multiple Ability Estimation in the SEM Framework: A Noninformative Bayesian Estimation Approach. *Multivariate Behavioral Research*, *48*(4), 563–591. <https://doi.org/10.1080/00273171.2013.802647>
- König, C., & van de Schoot, R. (2018). Bayesian statistics in educational research: A look at the current state of affairs. *Educational Review*, *70*(4), 486–509. <https://doi.org/10.1080/00131911.2017.1350636>
- Krishnadas, M., Sankaran, M., Page, N., Joshi, J., Machado, S., Nataraj, N., Chengappa, S. K., Kumar, V., Kumar, A., & Krishnamani, R. (2021). Seasonal drought regulates species distributions and assembly of tree communities across a tropical wet forest region. *Global Ecology and Biogeography*, *30*(9), 1847–1862. <https://doi.org/10.1111/geb.13350>
- Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*(2), 573–603. <https://doi.org/10.1037/a0029146>
- Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods*, *15*(4), 722–752.

- Lai, H. R., Chong, K. Y., Yee, A. T. K., Mayfield, M. M., & Stouffer, D. B. (2022). Non-additive biotic interactions improve predictions of tropical tree growth and impact community size structure. *Ecology*, *103*(2), e03588. <https://doi.org/10.1002/ecy.3588>
- Leach, C. B., Williams, P. J., Eisaguirre, J. M., Womble, J. N., Bower, M. R., & Hooten, M. B. (2022). Recursive Bayesian computation facilitates adaptive optimal design in ecological studies. *Ecology*, *103*(2), e03573. <https://doi.org/10.1002/ecy.3573>
- Lemoine, N. P. (2019). Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, *128*(7), 912–928.
- LoScerbo, D., Farrell, M. J., Arrowsmith, J., Mlynarek, J., & Lessard, J.-P. (2020). Phylogenetically conserved host traits and local abiotic conditions jointly drive the geography of parasite intensity. *Functional Ecology*, *34*(12), 2477–2487. <https://doi.org/10.1111/1365-2435.13698>
- Martin, A. D., Quinn, K. M., & Park, J. H. (2011). *MCMCpack: Markov chain monte carlo in R*.
- Martin, J. S., Ringen, E. J., Duda, P., & Jaeggi, A. V. (2020). Harsh environments promote alloparental care across human societies. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1933), 20200758. <https://doi.org/10.1098/rspb.2020.0758>
- Martínez-Abraín, A. (2008). Statistical significance and biological relevance: A call for a more cautious interpretation of results in ecology. *Acta Oecologica*, *34*(1), 9–11. <https://doi.org/10.1016/j.actao.2008.02.004>
- McCabe, J. D., Clare, J. D., Miller, T. A., Katzner, T. E., Cooper, J., Somershoe, S., Hanni, D., Kelly, C. A., Sargent, R., Soehren, E. C., Threadgill, C., Maddox, M., Stober, J., Martell, M., Salo, T., Berry, A., Lanzone, M. J., Braham, M. A., & McClure, C. J. W. (2021). Resource selection functions based on hierarchical generalized additive models provide new insights into individual animal variation and species distributions. *Ecography*, *44*(12), 1756–1768. <https://doi.org/10.1111/ecog.06058>
- McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.
- Mumbanza, F. M., Bauters, M., Kearsley, E., Boeckx, P., Lubini, C. A., & Verbeeck, H. (2020). Liana communities exhibit different species composition, diversity and community structure across forest types in the Congo Basin. *Biotropica*, *52*(4), 651–663. <https://doi.org/10.1111/btp.12787>
- Naidoo, R., Kilian, J. W., Du Preez, P., Beytell, P., Aschenborn, O., Taylor, R. D., & Stuart-Hill, G. (2018). Evaluating the effectiveness of local- and regional-scale wildlife corridors using quantitative metrics of functional connectivity. *Biological Conservation*, *217*, 96–103. <https://doi.org/10.1016/j.biocon.2017.10.037>
- Nakayama, R., Nakano, T., & Yusa, Y. (2020). Seasonal utilization patterns of two snail hosts by the epizoic limpet *Lottia tenuisculpta* (Gastropoda: Patellogastropoda). *Journal of Experimental Marine Biology and Ecology*, *530–531*, 151402. <https://doi.org/10.1016/j.jembe.2020.151402>
- Nathan, R., Monk, C. T., Arlinghaus, R., Adam, T., Alós, J., Assaf, M., Baktoft, H., Beardsworth, C. E., Bertram, M. G., Bijleveld, A. I., Brodin, T., Brooks, J. L., Campos-Candela, A., Cooke, S. J., Gjelland, K. Ø., Gupte, P. R., Harel, R., Hellström, G., Jeltsch, F., ... Jarić, I. (2022). Big-data approaches lead to an increased understanding of the ecology of animal movement. *Science*, *375*(6582), eabg1780. <https://doi.org/10.1126/science.abg1780>
- Nations, J. A., Mount, G. G., Morere, S. M., Achmadi, A. S., Rowe, K. C., & Esselstyn, J. A. (2021). Locomotory mode transitions alter phenotypic evolution and lineage diversification in an ecologically rich clade of mammals. *Evolution*, *75*(2), 376–393. <https://doi.org/10.1111/evo.14156>
- Neilands, P., Claessens, S., Ren, I., Hassall, R., Bastos, A. P. M., & Taylor, A. H. (2020). Contagious yawning is not a signal of empathy: No evidence of familiarity, gender or prosociality biases in dogs. *Proceedings of the Royal Society B: Biological Sciences*, *287*(1920), 20192236. <https://doi.org/10.1098/rspb.2019.2236>
- Newbold, T., Hudson, L. N., Hill, S. L. L., Contu, S., Lysenko, I., Senior, R. A., Börger, L., Bennett, D. J., Choimes, A., Collen, B., Day, J., De Palma, A., Díaz, S., Echeverria-Londoño, S., Edgar, M. J., Feldman, A., Garon, M., Harrison, M. L. K., Alhousseini, T., ... Purvis, A. (2015). Global effects of

- land use on local terrestrial biodiversity. *Nature*, 520(7545), 45–50. <https://doi.org/10.1038/nature14324>
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., & Maintainer, R. (2017). Package ‘nlme.’ *Linear and Nonlinear Mixed Effects Models, Version, 3*(1).
- Plummer, M. (2012). *JAGS Version 3.3. 0 user manual*. Lyon, France.
- R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Reichman, O. J., Jones, M. B., & Schildhauer, M. P. (2011). Challenges and Opportunities of Open Data in Ecology. *Science*, 331(6018), 703–705. <https://doi.org/10.1126/science.1197962>
- Reum, J. C. P., Holsman, K. K., Aydin, K. Y., Blanchard, J. L., & Jennings, S. (2019). Energetically relevant predator-prey body mass ratios and their relationship with predator body size. *Ecology and Evolution*, 9(1), 201–211. <https://doi.org/10.1002/ece3.4715>
- Reuter, J. A., Spacek, D. V., & Snyder, M. P. (2015). High-Throughput Sequencing Technologies. *Molecular Cell*, 58(4), 586–597. <https://doi.org/10.1016/j.molcel.2015.05.004>
- Roycroft, E. J., Nations, J. A., & Rowe, K. C. (2020). Environment predicts repeated body size shifts in a recent radiation of Australian mammals*. *Evolution*, 74(3), 671–680. <https://doi.org/10.1111/evo.13859>
- Shukla, P. R., Skeg, J., Buendia, E. C., Masson-Delmotte, V., Pörtner, H.-O., Roberts, D. C., Zhai, P., Slade, R., Connors, S., & van Diemen, S. (2019). *Climate Change and Land: An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems*.
- Snell Taylor, S., Umbanhowar, J., & Hurlbert, A. H. (2020). The relative importance of biotic and abiotic determinants of temporal occupancy for avian species in North America. *Global Ecology and Biogeography*, 29(4), 736–747. <https://doi.org/10.1111/geb.13064>
- Stachelek, J., Weng, W., Carey, C. C., Kemanian, A. R., Cobourn, K. M., Wagner, T., Weathers, K. C., & Soranno, P. A. (2020). Granular measures of agricultural land use influence lake nitrogen and phosphorus differently at macroscales. *Ecological Applications*, 30(8), e02187. <https://doi.org/10.1002/eap.2187>
- Stegmueller, D. (2013). How Many Countries for Multilevel Modeling? A Comparison of Frequentist and Bayesian Approaches. *American Journal of Political Science*, 57(3), 748–761. <https://doi.org/10.1111/ajps.12001>
- Stephens, E. Z., Murar, C. P., Tinker, D. B., & Copenhaver-Parry, P. E. (2019). Environmental Determinants of Recruitment Success of Subalpine Fir (*Abies lasiocarpa*) in a Mixed-Conifer Forest. *Western North American Naturalist*, 79(4), 481–495. <https://doi.org/10.3398/064.079.0402>
- Stephens, P. A., Buskirk, S. W., & del Rio, C. M. (2007). Inference in ecology and evolution. *Trends in Ecology & Evolution*, 22(4), 192–197. <https://doi.org/10.1016/j.tree.2006.12.003>
- van de Schoot, R., & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. *The European Health Psychologist*, 16(2), 75–84.
- Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P. C., & Paananen, T. (2021). Package: Loo. *Efficient Leave-One-Out Cross-Validation and WAIC for Bayesian Models*.
- Vehtari, A., Gelman, A., & Gabry, J. (2015). Efficient implementation of leave-one-out cross-validation and WAIC for evaluating fitted Bayesian models. *ArXiv Preprint ArXiv:1507.04544*.
- Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.1007/s11222-016-9696-4>
- Vickruck, J. L., Best, L. R., Gavin, M. P., Devries, J. H., & Galpern, P. (2019). Pothole wetlands provide reservoir habitat for native bees in prairie croplands. *Biological Conservation*, 232, 43–50. <https://doi.org/10.1016/j.biocon.2019.01.015>
- Wagenmakers, E.-J., Lee, M., Lodewyckx, T., & Iverson, G. J. (2008). Bayesian Versus Frequentist Inference. In *Bayesian Evaluation of Informative Hypotheses* (pp. 181–207). Springer New York. https://doi.org/10.1007/978-0-387-09612-4_9

- Wang, S., & Loreau, M. (2014). Ecosystem stability in space: α , β and γ variability. *Ecology Letters*, *17*(8), 891–901. <https://doi.org/10.1111/ele.12292>
- Wang, S., & Loreau, M. (2016). Biodiversity and ecosystem stability across scales in metacommunities. *Ecology Letters*, *19*(5), 510–518. <https://doi.org/10.1111/ele.12582>
- Wang, S., Loreau, M., de Mazancourt, C., Isbell, F., Beierkuhnlein, C., Connolly, J., Deutschman, D. H., Doležal, J., Eisenhauer, N., Hector, A., Jentsch, A., Kreyling, J., Lanta, V., Lepš, J., Polley, H. W., Reich, P. B., van Ruijven, J., Schmid, B., Tilman, D., ... Craven, D. (2021). Biotic homogenization destabilizes ecosystem functioning by decreasing spatial asynchrony. *Ecology*, *102*(6), e03332. <https://doi.org/10.1002/ecy.3332>
- Ward, E. J. (2008). A review and comparison of four commonly used Bayesian and maximum likelihood model selection tools. *Ecological Modelling*, *211*(1), 1–10. <https://doi.org/10.1016/j.ecolmodel.2007.10.030>
- Watanabe, S., & Opper, M. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, *11*(12).
- Whittaker, R. H. (1972). Evolution and Measurement of Species Diversity. *TAXON*, *21*(2–3), 213–251. <https://doi.org/10.2307/1218190>
- Wikle, C. K., & Hooten, M. B. (2010). A general science-based framework for dynamical spatio-temporal models. *Test*, *19*(3), 417–451.
- Wilcox, K. R., Tredennick, A. T., Koerner, S. E., Grman, E., Hallett, L. M., Avolio, M. L., La Pierre, K. J., Houseman, G. R., Isbell, F., Johnson, D. S., Alatalo, J. M., Baldwin, A. H., Bork, E. W., Boughton, E. H., Bowman, W. D., Britton, A. J., Cahill Jr., J. F., Collins, S. L., Du, G., ... Zhang, Y. (2017). Asynchrony among local communities stabilises ecosystem function of metacommunities. *Ecology Letters*, *20*(12), 1534–1545. <https://doi.org/10.1111/ele.12861>
- Wilms, T. J. G., Norðfoss, P. H., Baktoft, H., Støttrup, J. G., Kruse, B. M., & Svendsen, J. C. (2021). Restoring marine ecosystems: Spatial reef configuration triggers taxon-specific responses among early colonizers. *Journal of Applied Ecology*, *58*(12), 2936–2950. <https://doi.org/10.1111/1365-2664.14014>
- Zhang, Y., Bandyopadhyay, G., Topham, D. J., Falsey, A. R., & Qiu, X. (2019). Highly efficient hypothesis testing methods for regression-type tests with correlated observations and heterogeneous variance structure. *BMC Bioinformatics*, *20*(1), 185. <https://doi.org/10.1186/s12859-019-2783-8>
- Zhang, Y., He, N., Loreau, M., Pan, Q., & Han, X. (2018). Scale dependence of the diversity–stability relationship in a temperate grassland. *Journal of Ecology*, *106*(3), 1277–1285. <https://doi.org/10.1111/1365-2745.12903>
- Zhang, Z., Hamagami, F., Lijuan Wang, L., Nesselroade, J. R., & Grimm, K. J. (2007). Bayesian analysis of longitudinal data using growth curve models. *International Journal of Behavioral Development*, *31*(4), 374–383. <https://doi.org/10.1177/0165025407077764>

7. Appendix

7.1. Appendix A1: References used for the Literature Review

Author	Year	Title
Naidoo et al.	2018	Evaluating the effectiveness of local- and regional-scale wildlife corridors using quantitative metrics of functional connectivity
Barneche et al.	2018	Global environmental drivers of marine fish egg size
Reum et al.	2019	Energetically relevant predator-prey body mass ratios and their relationship with predator body size
Hoover et al.	2019	Predicting residence time using a continuous-time discrete-space model of leatherback turtle satellite telemetry data
Vickruck et al.	2019	Pothole wetlands provide reservoir habitat for native bees in prairie croplands
Barrow et al.	2019	Deeply conserved susceptibility in a multi-host, multi-parasite system
(E. Z. Stephens et al., 2019)	2019	Environmental Determinants of Recruitment Success of Subalpine Fir (<i>Abies lasiocarpa</i>) in a Mixed-Conifer Forest
Snell Taylor et al.	2020	The relative importance of biotic and abiotic determinants of temporal occupancy for avian species in North America
Mumbanza et al.	2020	Liana communities exhibit different species composition, diversity and community structure across forest types in the Congo Basin
Davy et al.	2020	Estimation of spatiotemporal trends in bat abundance from mortality data collected at wind turbines
Stachelek et al.	2020	Granular measures of agricultural land use influence lake nitrogen and phosphorus differently at macroscales
Bailey & Moore	2020	Resource pulses increase the diversity of successful competitors in a multi-species stream fish assemblage
Nakayama et al.	2020	Seasonal utilization patterns of two snail hosts by the epizoic limpet <i>Lottia tenuisculpta</i> (Gastropoda: Patellogastropoda)
Gade et al.	2020	Predicted alteration of surface activity as a consequence of climate change
LoScerbo et al.	2020	Phylogenetically conserved host traits and local abiotic conditions jointly drive the geography of parasite intensity
Roycroft et al.	2020	Environment predicts repeated body size shifts in a recent radiation of Australian mammals
Conesa et al.	2020	Stomatal anatomy coordinates leaf size with Rubisco kinetics in the Balearic <i>Limonium</i>
Neilands et al.	2020	Contagious yawning is not a signal of empathy: no evidence of familiarity, gender or prosociality biases in dogs
Martin et al.	2020	Harsh environments promote alloparental care across human societies
Hertel et al.	2021	Biologging reveals individual variation in behavioural predictability in the wild
Amirkhiz et al.	2021	Investigating niches and distribution of a rare species in a hierarchical framework: Virginia's Warbler (<i>Leiothlypis virginiae</i>) at its northeastern range limit
Adams et al.	2021	Temporal patterns in the social network of core units in Rwenzori Angolan colobus monkeys: Effects of food availability and interunit dispersal
Craven et al.	2021	Niche properties constrain occupancy but not abundance patterns of native and alien woody species across Hawaiian forests
Krishnadas et al.	2021	Seasonal drought regulates species distributions and assembly of tree communities across a tropical wet forest region
Wilms et al.	2021	Restoring marine ecosystems: Spatial reef configuration triggers taxon-specific responses among early colonizers
Nations et al.	2021	Locomotory mode transitions alter phenotypic evolution and lineage diversification in an ecologically rich clade of mammals
Gordijn & O'Connor	2021	Multidecadal effects of fire in a grassland biodiversity hotspot: Does pyrodiversity enhance plant diversity?
de Angeli Dutra et al.	2021	Haemosporidian taxonomic composition, network centrality and partner fidelity between resident and migratory avian hosts
Lai et al., 2022	2022	Non-additive biotic interactions improve predictions of tropical tree growth and impact community size structure
Giatzouzaki et al.	2022	Trait divergence between endemic plants of Aegean islands and their widespread congeners

7.2. Appendix A2: Diagnostic Plots and Additional Results

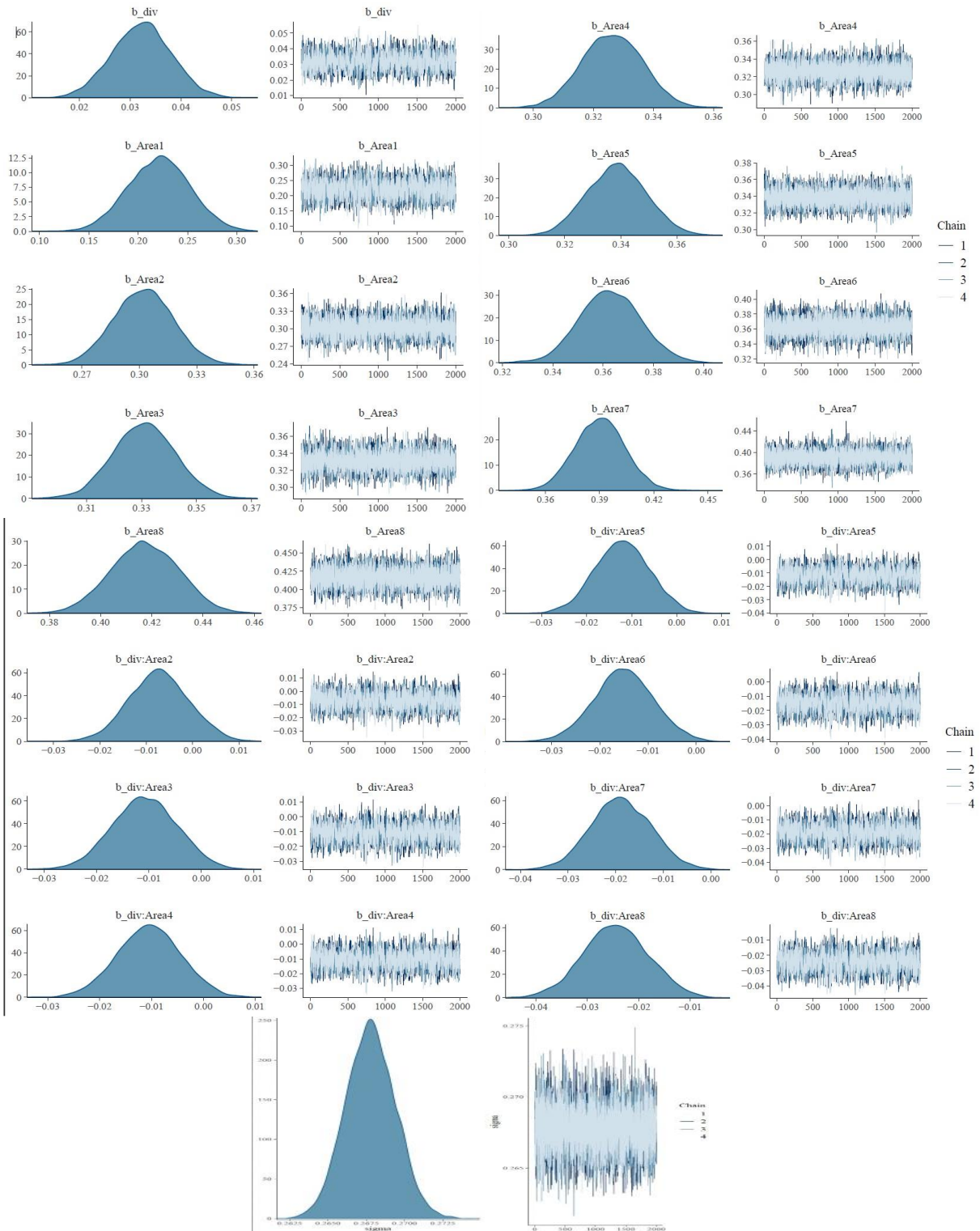


Figure 1. Trace plots for the visual assessment of chain convergence. These show that the selected model is exploring the full variation for its variables. Note that all trace plots are plotted excluding the warm-up iterations. All trace and density plots were generated using `plot()` function from the R package `brms` (Bürkner, 2017).

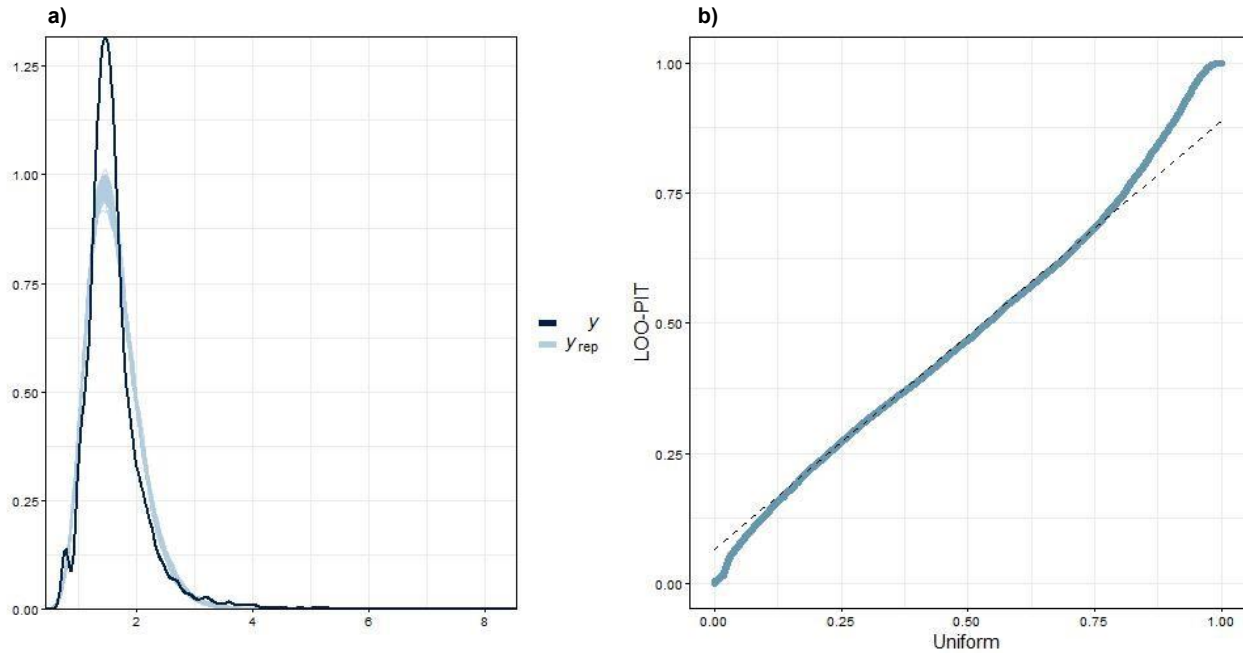


Figure 2. Posterior predictive checks show a satisfactory model fit. In a) the predictive distribution y_{rep} is compared to the observed data y . The similarity of the curves signifies the model captures the data well. Panel b) consists of a Leave One Out (LOO) Probability Integral Transform (PIT) marginal posterior predictive check. The points do not deviate considerably from the dotted line, therefore indicating the model is well-calibrated. All plots were generated using the `pp_check()` function of the R package *Bayesplot* (Gabry & Mahr, 2020).

Table 1. Statistical output for the selected Bayesian model. Estimates can be said to be significant when the 95% CI's do not overlap with 0.

Parameter	Estimate	SE	95% CI		R-hat	Bulk ESS	Tail ESS
			Lower	Upper			
Alpha Diversity	0.03	0.01	0.02	0.04	1.00	1746	3014
Area1	0.22	0.03	0.16	0.28	1.00	2113	3881
Area2	0.30	0.02	0.27	0.33	1.00	11780	5686
Area3	0.33	0.01	0.31	0.35	1.00	11337	5539
Area4	0.33	0.01	0.31	0.35	1.00	11964	5614
Area5	0.34	0.01	0.32	0.36	1.00	11211	5840
Area6	0.36	0.01	0.34	0.39	1.00	11130	5563
Area7	0.39	0.01	0.36	0.42	1.00	11425	5450
Area8	0.42	0.01	0.39	0.44	1.00	11271	5728
Alpha Diversity x Area2	-0.01	0.01	-0.02	0.01	1.00	2057	3441
Alpha Diversity x Area3	-0.01	0.01	-0.02	0.00	1.00	1898	3290
Alpha Diversity x Area4	-0.01	0.01	-0.02	0.00	1.00	1899	3329
Alpha Diversity x Area5	-0.01	0.01	-0.02	0.00	1.00	1855	3323
Alpha Diversity x Area6	-0.02	0.01	-0.03	0.00	1.00	1924	3412
Alpha Diversity x Area7	-0.02	-0.01	-0.03	-0.01	1.00	2010	3295
Alpha Diversity x Area8	-0.02	0.01	-0.04	-0.01	1.00	1938	3518

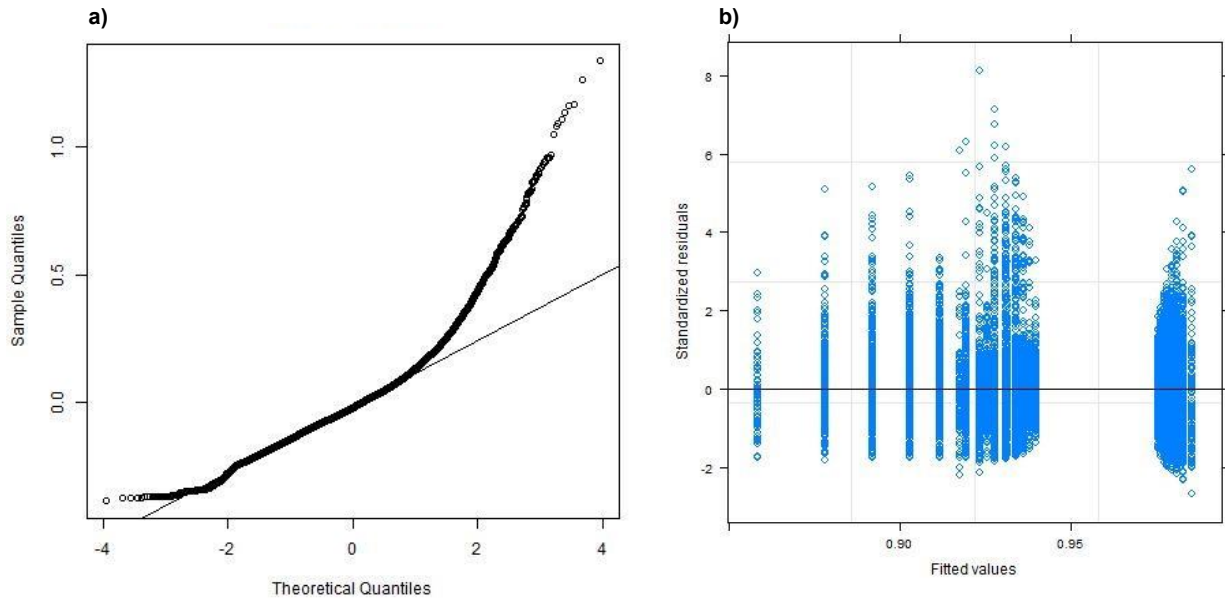


Figure 3. Diagnostic Plots for the best-fit linear mixed effects model generated using the ‘lme’ function from the R package ‘nlme’ (Pinheiro et al., 2017). In a) a QQ plot shows deviations from normality. In b) a heteroscedastic distribution of the residuals can be visualized, even after the most appropriate variance structure was selected for.

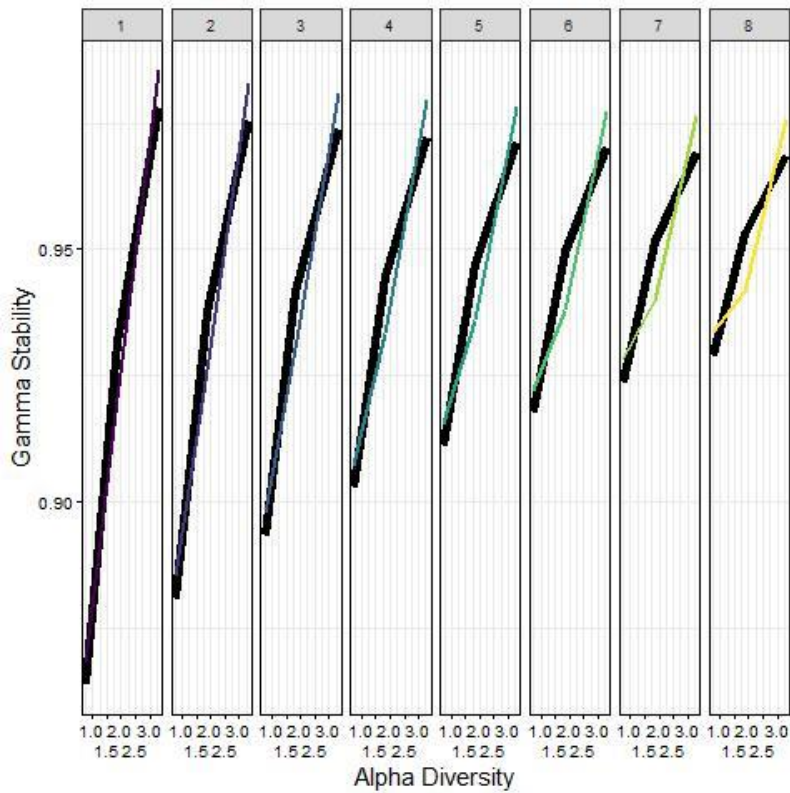


Figure 4. Results on the relationship between alpha diversity, gamma diversity and area obtained from a frequentist random intercept and slope model on the back-transformed. Fixed effects are plotted in black. Colours represent the slopes at different areas. Facets represent different areas, ranging from one single to eight aggregated communities.

7.3. Appendix A3: Full Methodology associated with the case study

To investigate the scale dependence and test for the effect of both of alpha and beta diversity on gamma temporal stability, both variables need to be separated from each other as well as from environmental conditions. For alpha diversity, this is achieved by using biodiversity experiments which are purposefully designed to manipulate alpha diversity under homogeneous environmental conditions (Hector, 2006). This means that spatial processes responsible for beta diversity in natural landscapes are not accounted for. Therefore, to be able to test also for the effect of beta diversity under similar levels of alpha diversity in these existing experiments, landscapes with varying levels of beta diversity need to be simulated (Pasari et al., 2013; Wang et al., 2021).

1. The Utrecht University Biodiversity and Climate Variability Experiment

Data from the Utrecht University Biodiversity and Climate Variability Experiment (UU BioCliVE) platform was utilized, where natural temperate grassland communities typically found in the Netherlands are replicated in experimental mesocosms (UU BioCliVE, n.d.). BioCliVE was established in 2017 at the Utrecht University Botanical Gardens (52°13'N, 5°29'E). The area surrounding the experiment is characterized by a mean annual temperature of 10.1 °C and a mean annual precipitation of 887 mm (1981-2010; KNMI, 2021).

1.1. Experimental Design

BioCliVE was designed to study the combined impacts of climate change (in the form of changing precipitation patterns) and biodiversity loss on ecosystem dynamics and ecosystem functioning. For that purpose, a total of 352 mesocosms were created using 1.2 m x 1.0 m x 1.15 m (l x b x h) bulk containers that were each filled with 1000 L of soil. To ensure appropriate drainage, this soil was composed of the following layers from bottom to top: (1) a 10 cm thick drainage layer of gravel followed by (2) a 70 cm thick layer of river sand and (3) a 25 cm thick topsoil layer, which is a perfect (1:1) mixture of soil and sand. For the top layer, the soil from a meadow with *Arrhenatherion elatioris* association was excavated from a nearby nature reserve in the *Rammelwaard* (52°10'N, 6°11'E). It is important to note here, that a single mesocosm represents a subplot, while four of them collectively represent one plot (Fig. 2). These plots were sown to construct communities with a species richness gradient of one, four, eight and twelve species, which were selected from a local species pool associated with the aforementioned vegetation type (Table 1). Seeds were sown at a density of 1800 seeds per mesocosm; for mixtures, the different species were sown at equal proportions.

Table 1. Target species sown in the BioCliVE platform. These species pertain to two functional groups.

Species	Family	Functional Group
<i>Arrhenaterum elatius</i>	Poaceae	Grasses
<i>Holcus lanatus</i>	Poaceae	
<i>Luzula campestris</i>	Juncaceae	
<i>Anthoxanthum odoratum</i>	Poaceae	
<i>Poa trivialis</i>	Poaceae	
<i>Festuca rubra</i>	Poaceae	
<i>Tragopogon pratensis</i>	Asteraceae	Forbs
<i>Rumex acetosa</i>	Polygonaceae	
<i>Veronica chamaedrys</i>	Plantaginaceae	
<i>Cardamine pratensis</i> *	Brassicaceae	
<i>Origanum vulgare</i>	Lamiaceae	
<i>Knautia arvensis</i>	Caprifoliaceae	

*replaced by *Ranunculus repens* (Family: Ranunculaceae) in 2020 due to negative growth

A total of 88 plots was arranged in two blocks with a varying degree of plant species diversity: in each block there are 12 monoculture plots and 32 polyculture plots. Of the polyculture plots, 12 contain four species, 12 contain eight species and 8 contain twelve species. Therefore, each species composition can be said to be replicated 8 times, except for the twelve-species plots, which are replicated 64 times in total (Fig. 2).



Figure 1. Experimental Design of the UU BioCliVE platform showing the setup in two blocks, each containing 44 plots of varying diversity levels.

1.2. Data Collection

Data on aboveground plant biomass has been collected for four years between 2018 and 2021. To maintain the compositional identity of each plot as well as the intended diversity gradient, weeds were removed manually thrice a year (in March, June, and September). For this purpose, weeds were defined as any species that were not intended for a particular plot. Because it is impossible to track whether intended species establish from seeds coming from within or among different subplots with the same species, the design allows for a certain degree of dispersal and thus interconnection among those communities.

Every year after the June weeding, when peak standing crop was reached, biomass of the sown species was collected. This was achieved by harvesting the inner 50 x 50 cm² (2018 – 2020) and 25 x 50 cm² (2021) of each plot by clipping the vegetation at a height of 7 cm. The harvested biomass was then sorted to species level before it was oven dried at 70°C for 48 hours and weighed.

2. Diversity and Stability Metrics in Simulated Landscapes

To investigate the relationships between species diversity, spatial asynchrony, area and the temporal stability of productivity, landscapes of varying sizes and with varying levels of beta diversity were simulated.

2.1. Landscape Simulations

Following the approach used by other studies investigating the effect of beta diversity on ecosystem functioning and stability (Pasari et al., 2013; Wang et al., 2021), our simulations were performed by randomly aggregating plots with the same level of species richness. The size of these landscapes ranged from one to eight communities (Fig. 3) and therefore from 1.2 to 9.6 m². The same levels of species richness were used to prevent the potential confounding effect of alpha diversity when testing for beta diversity (Wang et al., 2021). To reduce the size of the data, a subsample of 60 landscapes was taken for each area – beta diversity combination. A total of 12,958 landscapes were obtained with a continuous gradient of beta diversity by pooling subplots within each simulated landscape so that their species compositions were either the same, partially overlapping or completely different (Fig. 3; *see Appendix 1* Fig. A1). Due to this method, different landscapes may share some plots, but most importantly no plots were resampled for any given landscape. All simulations were performed in R 4.1.1 (R Core Team, 2021). Owing to the replacement of *C. pratensis* by *R. repens* in 2020 which resulted in the unavailability of the required temporal data on biomass production, all data on either species were excluded from our analyses.

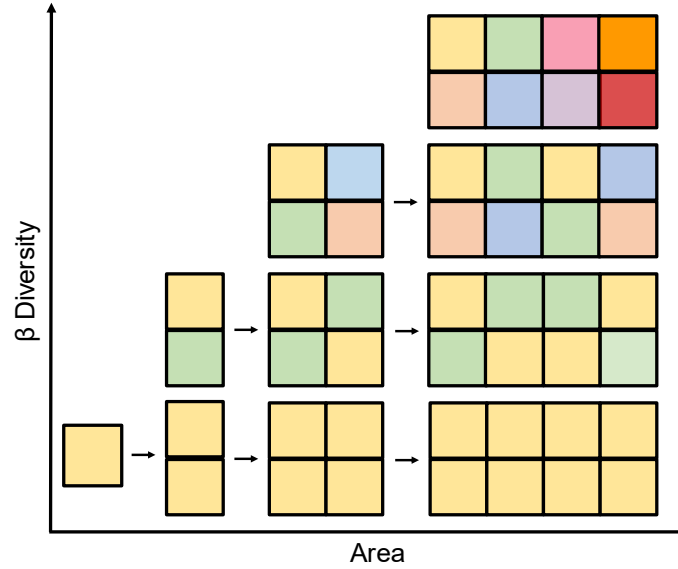


Figure 2. Graphical representation of the simulation procedure. Note that different colors represent different species compositions.

2.2. Calculating Diversity and Stability Metrics

Diversity and stability metrics were calculated by applying the ‘var.partition’ function developed by Wang et al. (2019) to data collected between 2018 and 2021. All measurements were computed for every landscape at both, community (alpha) and larger spatial (gamma) scales using presence-absence-based metrics of diversity. All calculations were performed to obtain one single value for each of the required metrics for every landscape.

Alpha diversity was calculated as the average number of species across each of the assembled communities through time. Similarly, gamma diversity was calculated as the average number of species for the entire landscape. Then, to represent the dissimilarity in species composition between the different communities within a landscape, beta diversity was calculated as the ratio of gamma diversity to alpha diversity (*sensu* Whittaker, 1972).

Ecosystem stability was defined as the temporal invariability of aboveground biomass productivity, calculated as the inverse of the coefficient of variation (Wang & Loreau, 2014, 2016; Wang et al., 2021). Accordingly, alpha stability and gamma stability were quantified for every landscape using the following formulae respectively:

$$\alpha_S = \frac{\sum_i \mu_i}{\sum_i \sqrt{v_{ii}}} \quad (1)$$

$$\gamma_S = \frac{\sum_i \mu_i}{\sqrt{\sum_{i,j} v_{ij}}} \quad (2)$$

where μ_i denotes the temporal mean and v_{ii} the temporal variance of community biomass in subplot i , and v_{ij} denotes the covariance in community biomass between subplot i and j . Then, the beta component of

stability, i.e., spatial asynchrony, could again be calculated as the ratio between gamma stability and alpha stability, yielding the following definition:

$$\omega = \frac{\sum_i \sqrt{v_{ii}}}{\sqrt{\sum_{i,j} v_{ij}}} \quad (3)$$

This definition of spatial asynchrony is also referred to as covariance-based asynchrony because it combines the correlation among plots as well as the variance within plots (Loreau & De Mazancourt, 2008). It is important to note, that even though there are also other measures of beta diversity and asynchrony, these multiplicative measures were chosen due to their consistency with the theoretical framework that we are testing (Wang & Loreau, 2014, 2016).

3. Statistical Analysis

To determine the relationships between the different components of the framework that is being examined, as well as the scale dependence of these relationships, a series of linear mixed effects models were performed. These models were generated using the ‘lme’ function from the R package ‘nlme’ (Pinheiro et al., 2017) and fitted with restricted maximum likelihood. For any of the possible response variables, one or more models were created with the associated explanatory variable (Table 2), area and planted alpha diversity as fixed effects, and area and planted alpha diversity as random effects. To visualize the relationship of each fixed effect with the response variable for the final models, mean marginal effects were extracted using the ‘ggeffect’ function from the R package ‘ggeffects’ (Lüdtke, 2018).

Table 2. Various response variables and their associated explanatory variables used in the series models used in this study. Their pairing is based on the framework developed by Wang & Loreau (2014; 2016). Indirect explanatory variables are called such because they do not directly affect the associated response variable, but through another variable, i.e., beta diversity influences gamma stability indirectly through spatial asynchrony.

Response Variable	Explanatory Variable
Alpha Stability	Alpha Diversity
Spatial Asynchrony	Beta Diversity
Gamma Stability	Indirect: Alpha Diversity, Beta Diversity Direct: Alpha Stability, Spatial Asynchrony, Gamma Diversity

All stability metrics were log-transformed before analyses to improve normality and to ensure the assumptions of the model were being met. To account for heteroscedasticity of the residuals, different variance structures were modelled. For each relationship, the best-fit model was determined following the protocol proposed by Zuur et al., (2009) and final model selection was based on Akaike Information Criterion (AIC) minimization.

4. References

- KNMI. (2021, 15 March 2019). Climate data De Bilt; temperature, precipitation, sunshine 1800-2014. Retrieved 30 November 2021 from <https://www.cbs.nl/en-gb/figures/detail/80370eng>
- Loreau, M., & de Mazancourt, C. (2008). Species synchrony and its drivers: neutral and nonneutral community dynamics in fluctuating environments. *The American Naturalist*, 172(2), E48-E66.
- Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software*, 3(26), 772.
- Pasari, J. R., Levi, T., Zavaleta, E. S., & Tilman, D. (2013). Several scales of biodiversity affect ecosystem multifunctionality. *Proceedings of the National Academy of Sciences*, 110(25), 10219-10222. <https://doi.org/10.1073/pnas.1220333110>
- Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., Heisterkamp, S., Van Willigen, B., & Maintainer, R. (2017). Package 'nlme'. *Linear and nonlinear mixed effects models, version*, 3(1).
- R Core Team. (2021). R Foundation for Statistical Computing. <https://www.R-project.org/>
- Wang, S., & Loreau, M. (2014). Ecosystem stability in space: α , β and γ variability. *Ecology Letters*, 17(8), 891-901. <https://doi.org/10.1111/ele.12292>
- Wang, S., & Loreau, M. (2016). Biodiversity and ecosystem stability across scales in metacommunities. *Ecology Letters*, 19(5), 510-518. <https://doi.org/10.1111/ele.12582>
- Wang, S., Loreau, M., De Mazancourt, C., Isbell, F., Beierkuhnlein, C., Connolly, J., Deutschman, D. H., Doležal, J., Eisenhauer, N., Hector, A., Jentsch, A., Kreyling, J., Lanta, V., Lepš, J., Polley, H. W., Reich, P. B., Van Ruijven, J., Schmid, B., Tilman, D., Wilsey, B., & Craven, D. (2021). Biotic homogenization destabilizes ecosystem functioning by decreasing spatial asynchrony. *Ecology*. <https://doi.org/10.1002/ecy.3332>
- Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). Mixed effects modelling for nested data. In *Mixed effects models and extensions in ecology with R* (pp. 101-142). Springer