REPOSITORY: <https://github.com/emilygsimmonds/Interspecific-interactions>

# Weak interspecific interactions are hard to estimate but have important consequences for population dynamics of interacting species

I read through the interactions paper. I think it's interesting, but have a couple of thoughts:

1. I think you'll have to need a good reason for just counting the times the signs are correctly estimated. If nothing else, I think you might need to put analyses with the estimates in the SI to justify it.

2. You don't really show that mis-estimating the interactions makes a difference. You show that if you change the interactions and keep everything else constant, you get different distributions. But in the (poorly) estimated models the parameters will probably compensate, so I think you'd need to show that simulations from the fitted models differ from simulations from the true model. I think this would be important whatever the result.

## Abstract

Species interactions are fundamental to shaping population and community dynamics. Despite their importance in nature, quantifying their strength is challenging. It is complicated by small effect sizes, confounding variables, and counting errors. Common tools used for identifying interactions from time series data are multivariate autoregressive models, usually of order 1. These models use population abundances at a previous time step to explain the abundance of each interacting species at the next time step. Formulating these as state-space models provides an avenue to quantify both the processes, or state, change in ‘true’ abundance through inter- and intra-specific interactions while accounting for error in the observation of this process (counting of individuals). Such models are intuitively appealing as they permit the inclusion of measurement error which will be present in virtually all real systems. However, explicitly modelling this extra layer of noise can make distinguishing the direction and strength of interactions more difficult. We use a simulation study to test whether multivariate autoregressive models can retrieve interaction strengths under four scenarios. We simulate time series for pairs of interacting species with different interaction strengths and types; predator-prey, competitive, and mutualist. We then use hierarchical models to retrieve the parameter values used to generate the time series. Our models are run in two Bayesian formats, MCMC and an approximate method; Nimble and INLA, respectively. We show that these models struggle to significantly distinguish the direction and strength of interspecific interactions due to weak interaction strengths. Weak interaction strengths are common in natural populations and could seem unimportant. We show that despite small effect sizes, interspecific interactions can still have a large impact of the population dynamics of the interacting species and a misidentification of the interaction strength or direction could have severe consequences for inference and prediction of population trends.

## Introduction

No populations occur entirely in isolation, instead they influence and are influenced by populations of other species. Species interactions shape communities and the populations that comprise them (Ovaskainen *et al.*, 2016). However, inferring their direction and strength from observational data can be challenging. Obtaining detailed data on when species interact and in what form that interaction takes is logistically difficult, if not impossible, to collect in most wild systems. Therefore, researchers must instead infer interactive effects from other sources of observational data. One such source is time series data of population counts. These data can be used to try and tease apart intra and interspecific effects as well as the influence of environmental drivers on population dynamics of a multitude of species. A typical model used to quantify species interactions from time series data are linked Gompertz equations with autoregressive terms (e.g. (Mutshinda, O’Hara and Woiwod, 2009, 2011,Stenseth *et al.*, 2015,Porzig *et al.*, 2016,Certain, Barraquand and Gårdmark, 2018)). Gompertz dynamics models have been used to explore community interactions across a range of different scenarios from quantifying community interaction strengths in macro-moths (Mutshinda, O’Hara and Woiwod, 2011) to exploring interaction strengths of land birds under rapid successional change (Porzig *et al.*, 2016) and quantifying how competition between two species shifts with a changing climate (Stenseth *et al.*, 2015).

While the Gompertz dynamics model can capture the processes driving changes in population abundance it should be noted that data on population counts are rarely perfect. Researchers will miss some individuals and double count others. There is an observation process that drives population counts in addition to the underlying processes driving population change. To jointly model these two processes, state-space formulations of hierarchical Bayesian models are good candidate for this. State-space models of joint time series have already been applied in community ecology. Mutshinda, O’Hara and Woiwod (2009) proposed a new formulation of Bayesian hierarchical model of species interactions that could tease apart the effects of environmental stochasticity from intra- and inter-specific interactions. Results from such models have typically found weak strengths of interactions in comparison to environmental forcing (Mutshinda, O’Hara and Woiwod, 2011, Porzig *et al.*, 2016). This work suggests that temporal variation in abundance is dominated by environmental forcing and interspecific interactions are weak. But, the ability of such models, with low sample sizes and high amounts of data noise, to quantify these interactions accurately and precisely has been seldom tested.

One study that did question the universal assumption of Gompertz dynamics was Certain, Barraquand and Gårdmark (2018). This study found that while multivariate autoregressive models of order 1 with Gompertz dynamics could return the sign of interspecific interactions, it could not reliably return net interaction strength, especially if true dynamics are actually non-linear. Therefore, despite often being the target quantity of community ecologists, estimating interaction strengths in natural systems appears to be difficult (Wolf, Novak and Gitelman, 2017). With assumed noisy data, many processes and relatively weak interaction strengths, this sets up a near impossible challenge for models. Increasing ease of application of modelling techniques designed to try and quantify interactions strengths and frequent lack of identification of any strong interactions (Mutshinda, O’Hara and Woiwod, 2011,Ovaskainen *et al.*, 2016,Porzig *et al.*, 2016), prompted us to test whether these results are a function of the models used or whether this says something about species interactions and their role in population dynamics more broadly.

Here, we conduct a simulation study to explore how these models perform and how often they can correctly retrieve estimates of interaction strength. Look at this in the context of four scenarios across three interaction types (predator-prey, mutualist, and competitive). Try two different Bayesian modelling approaches, look at results in terms of accuracy, precision, and consistency.

## Methodology

### Gompertz model

Gompertz dynamics were simulated on the log linear scale: following Mutshinda, O’Hara and Woiwod (2009, 2011) and Stenseth *et al.* (2015)

State process:

(1)

(2)

Observation process:

(3)

(4)

Where:

is the log abundance of species at time

is the intrinsic population growth rate

is the effect of species on itself (intraspecific competition) = , where is the natural log of the carrying capacity for species

is the effect of species on itself (intraspecific competition) = , where is the natural log of the carrying capacity for species

is the effect of species on species on the log linear scale (interspecific competition) = , where is the effect of species on species on the original scale

is the effect of species on species on the log linear scale (interspecific competition) = , where is the effect of species on species on the original scale

is a matrix of random error with rows and

= degrees of freedom ()

= 2

= a variance-covariance matrix

is the observed log population abundance for species at time

is variance of for both species

is the correlation in between the two species

and are the covariance between the two species calculated as

### Scope of the simulation study

This simulation study covered interspecific interactions between pairs of interacting species. We simulated time series of population counts for each species, called species and species . Our simulation covered three types of two-way interspecific interaction:

* Competitive (where species has a negative effect on the population size of species and species has a negative effect on the population size of species ), such as great tits (*Parus major*) and blue tits (*Cyanistes caeruleus*).
* Mutualist (where species has a positive effect on the population size of species and species has a positive effect on the population size of species ), such as plantain lilies and the European honeybee.
* Predator prey (where a predator species has a negative effect on the population size of the prey species , but the prey species has a positive effect on the population size of species ), such as the red fox and a rabbit.

We focused on interactions between only two species rather than a whole community in order to test our model formulations on minimal, illustrative examples.

Our simulation set up was designed to mimic the collection of population count data, therefore, we simulated both ‘true’ population abundance changes and an observation process to produce observed counts for each species. Further details of the simulation set up are provided below.

We made several simplifying assumptions about our simulated species:

* The only drivers of population abundance of each species are intra-specific competition, an inter-specific interaction, and environmental noise. We assume only the two focal species ( and ) are interacting.
* Environmental noise is normally distributed with a mean of zero and is correlated between the two interacting species.
* The observation error is Poisson distributed and not influenced by environmental variation.
* The species follow Gompertz dynamics.

### Testing the importance of interspecific interactions

We used two scenarios:

* Baseline: no interspecific interaction
* Interactions: competitive, mutualist, predator prey

We wanted to test the effect of different interaction types on observed population dynamics. All parameters other than the interspecific interaction effects, and ,were held constant. The population size estimates for both species and were unconstrained so could span from extinction to infinity.

For each scenario and interaction type, 100 paired time series were simulated each for 100 years with the first 50 being burn in (i.e. removed from the final analysis). The starting population size for each species was 100 individuals. The simulations were stochastic because although all parameters were fixed, was randomly generated for each simulation from a multivariate normal distribution. From each simulation, , , and were recorded for all values of .

|  |  |  |
| --- | --- | --- |
| **Parameters** | **Baseline** | **Interaction** |
|  | 1 | 1 |
|  | 4.6 | 4.6 |
|  | 4.6 | 4.6 |
|  | 0 | **0.5 (c),**  **-0.5 (m),**  **0.5 (pp)** |
|  | 0 | **0.5 (c),**  **-0.5 (m),**  **-0.5 (pp)** |
|  | -0.22 | **0.5 (c),**  **-0.5 (m),**  **0.5 (pp)** |
|  | 0 | **0.5 (c),**  **-0.5 (m),**  **-0.5 (pp)** |
|  | 0.05 | 0.05 |
|  | 0.7 | 0.7 |

### Testing the performance of state-space models

We explored four scenarios. The baseline scenario had equal interaction strength between the two species (i.e. species had the same strength of effect on species as species has on species ). The unequal scenario explored how models perform when one species has a stronger effect on the other species than vice versa. The increased noise scenario explored how well interactive effects can be retrieved when environmental noise is higher. The final scenario explored how model performance changes when the correlation () in environmental stochasticity between the species is altered. We tested both a decrease and an increase in correlation. Full details of the scenario parameters are shown in Table X. These scenarios were designed to represent situations that could be encountered with real time series data available to Ecologists. Across this study we assumed no *a priori* knowledge about the direction or strength of species interactions ( and , the variability of environmental stochasticity or the correlation in environmental stochasticity ( and ). However, we did assume some prior knowledge of plausible values for the intrinsic growth rate and the strength of the intra-specific competition ( and ). It is not likely to always be realistic that researchers would begin to model interspecific interactions between pairs of species with no prior knowledge of the interaction direction, however, we wanted to test model performance under a minimal information scenario. Such scenarios could be encountered for hard to observe species or when expanding the use of Gompertz dynamics models across whole communities. In both situations it would not always be possible to have a confident idea of interaction directions *a priori*.

For each scenario and interaction type, 100 paired time series were simulated each for 100 years with the first 50 being burn in (i.e. removed from the final analysis). The starting population size for each species was 100 individuals. The simulations were stochastic because although all parameters were fixed, was randomly generated for each simulation from a multivariate normal distribution. From each simulation, , , and were recorded for all values of .

In this study we explored four scenarios. Parameter values for each scenario were chosen to ensure stable population sizes with and , fluctuating between 3 and 1200 individuals. The values used are shown in Table X.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Parameters** | **Baseline** | **Unequal** | **Increased noise** | **Changed correlation in noise** |
|  | 1 | 1 | 1 | 1 |
|  | 4.6 (c),  3.9 (pp and m) | 4.6 (c),  3.9 (pp and m) | 4.6 (c),  3.9 (pp and m) | 4.6 (c),  3.9 (pp and m) |
|  | 4.6 (c and pp), 3.9 (m) | 4.6 (c and pp), 3.9 (m) | 4.6 (c and pp), 3.9 (m) | 4.6 (c and pp), 3.9 (m) |
|  | 0.5 (c),  -0.1 (m),  0.25 (pp) | **0.3 (c),**  **-0.1 (m),**  **0.1 (pp)** | 0.5, 0.3 (c),  -0.1, -0.1 (m),  0.3, 0.25 (pp) | 0.3 (c),  -0.1 (m),  0.3 (pp) |
|  | 0.5 (c),  -0.1 (m),  -0.25 (pp) | **0.1 (c),**  **-0.3 (m),**  **-0.3 (pp)** | 0.5, 0.1 (c),  -0.1, -0.3 (m), --0.25, -0.1 (pp) | 0.1 (c),  -0.3 (m),  -0.1 (pp) |
|  | 0.05 | 0.05 | **0.1** | 0.05 |
|  | 0.7 | 0.7 | 0.7 | **0.3, 0.5, 0.9** |

### Generating simulated time series

We simulate two processes, ‘true’ population change and an observation process.

The matrix formulation of the Gompertz equation (equations 1-2 above) was used to simulate paired ‘true’ population abundance time series. The observation process is then simulated with Poisson error around the ‘true’ population counts following equations 3 and 4.

### Statistical modelling of the simulated time series

To estimate the parameter values from the simulated time series we used two Bayesian hierarchical models run using the R-INLA and Nimble R packages (CITE). R-INLA is an approximate method using Laplace approximation. Nimble uses MCMC. By applying both approaches, we are able to assess whether method of sampling the posterior distribution impacts results.

Full code for running the simulations and fitting the models can be found on Zenodo (REF).

Both models took the same form as the simulation equations, therefore assuming that the study species population dynamics do follow Gompertz dynamics in this case.

Uninformative, but not flat, priors (see Table 2) were used to mimic a situation where the direction of an inter-specific interaction was not known *a priori*. All priors for interactive effects had a mean of 0, as has been applied in previous work on natural systems (Porzig *et al.*, 2016). Some constraints were imposed for biological realism, for example, parameter (the intrinsic grow rate) was constrained to be positive. A negative value for this parameter is not biologically plausible. (the carrying capacity) was also constrained to be positive, resulting in parameters and also being constrained to be positive.

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Prior R-INLA** | **Prior Nimble** |
|  | Normal distribution | Truncated normal distribution (min = 0, max = 10) |
|  | Normal distribution | Truncated normal distribution (min = 0, max = 10) |
|  | Normal distribution | Truncated normal distribution (min = 0, max = 10) |
|  | Normal distribution | Normal distribution |
|  | Normal distribution | Normal distribution |
|  | wishart2d prior (default in R-INLA) | Wishart distribution |

*Running information for the Nimble model*

Run with block sampling for correlated parameters. During initial trials of this model we found the posterior distributions of three parameters*,*  or , and or  were not independent. Therefore, a covariance matrix was set up and block samplers used for these parameters. Values for the variances and covariances were taken from the posterior samples of an initial run of the model for the Baseline scenario without block sampling. The same covariance structure was used for all subsequent model runs. The Nimble models were run for 10000 iterations with a burn in of 1000 and three chains. These running conditions were chosen to ensure convergence of the model, which was tested during set up using the Gelman diagnostic, calculated using the coda package (CITE). This diagnostic was not checked for every simulated dataset, it was assumed that the number of iterations was sufficient to achieve convergence under all simulated datasets.

### Evaluation of results

To evaluate the performance of our models for correctly estimating interactive effects we focused on the estimation of andFor these interactive effect parameters we assessed absolute error of estimates (mean posterior estimate – the true parameter value used during simulation), whether the direction of the effect was estimated clearly (whether the 95% credible interval spanned zero), and whether the sign was estimated correctly on average (was the mean posterior estimate of the same sign as the true parameter used for simulation).

## Results

### The impact of weak interactions

Results are shown below. It is just a plot of the distribution of all population sizes generated for each species across the 100 simulations (quite crude). I think it shows how different the results are for the different interaction types, even though the effects are relatively weak.

A picture containing timeline

Description automatically generated

Also summary tables of the population sizes.

Species i

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | Label |
| 19.94 | 77.08 | 98.71 | 105.61 | 125.95 | 387.96 | no interaction |
| 7.14 | 17.13 | 21.37 | 22.63 | 26.79 | 68.98 | competition |
| 223.64 | 6909.94 | 9440.36 | 10468.12 | 13017.51 | 39146.86 | mutualism |
| 58.74 | 192.87 | 252.32 | 270.22 | 323.23 | 1130.71 | predator prey |

Species j

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Min. | 1st Qu. | Median | Mean | 3rd Qu. | Max. | Label |
| 30.70 | 76.74 | 98.02 | 105.03 | 125.16 | 451.11 | no interaction |
| 6.35 | 17.15 | 21.84 | 23.02 | 27.27 | 68.99 | competition |
| 1862.63 | 6841.21 | 9444.04 | 10419.44 | 13021.04 | 37475.50 | mutualism |
| 2.18 | 5.12 | 6.36 | 6.67 | 7.85 | 19.94 | predator prey |

### The ability of state-space models to correctly retrieve interaction strengths

The overall results of our simulations showed that the signs of interactive effects were rarely recovered convincingly (credible intervals almost always spanned zero) (see SI Fig X). This was unchanged by any scenarios explored here. Under the baseline scenario, approximately half of the interactive effects were estimated in the wrong direction > 50% of the time, competitive interactions were estimated in the wrong direction half of the time and only the mutualist interactions using the Nimble model were consistently estimated in the correct direction. Some improvements were achieved under the different scenarios explored here (details below), however there was no clear pattern to these improvements and no scenario achieved universal success.

#### Baseline results

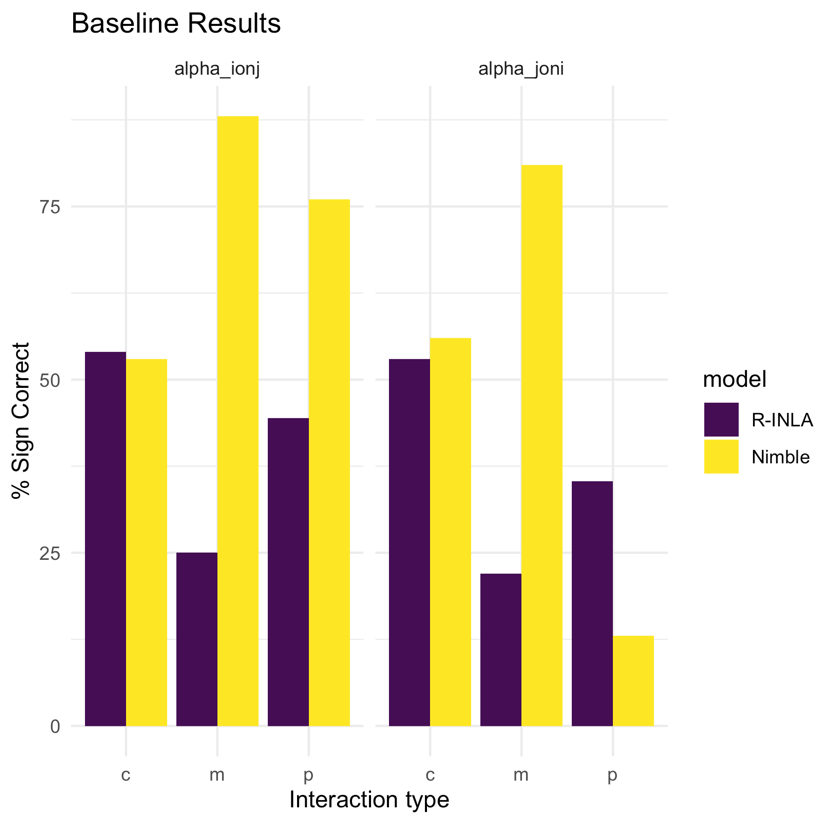
****

Figure 1: Results of baseline simulation. Bar chart shows percentage of results that returned mean posterior estimates of the correct sign for interactive parameters from 100 simulated datasets

Both competitive interactions were retrieved with the correct sign approximately 50% of the time for both the INLA and Nimble models (see Figure 1). The directions of the mutualist interactions were retrieved correctly >75% of the time using the Nimble model but < 25% of the time for the INLA model. Predator prey interactions showed a mixed picture. The Nimble model retrieved the correct sign of the species i on j interaction 75% of the time but the correct sign for the species j on i interaction only ~ 10 % of the time. In contrast, the INLA model did not retrieve either interaction in the correct direction > 50 % of the time. Overall, the Nimble model performed more consistently across the different interaction types. The INLA model only outperformed the Nimble model for the species j on i interaction for predator prey interaction types.

#### Unequal interactions

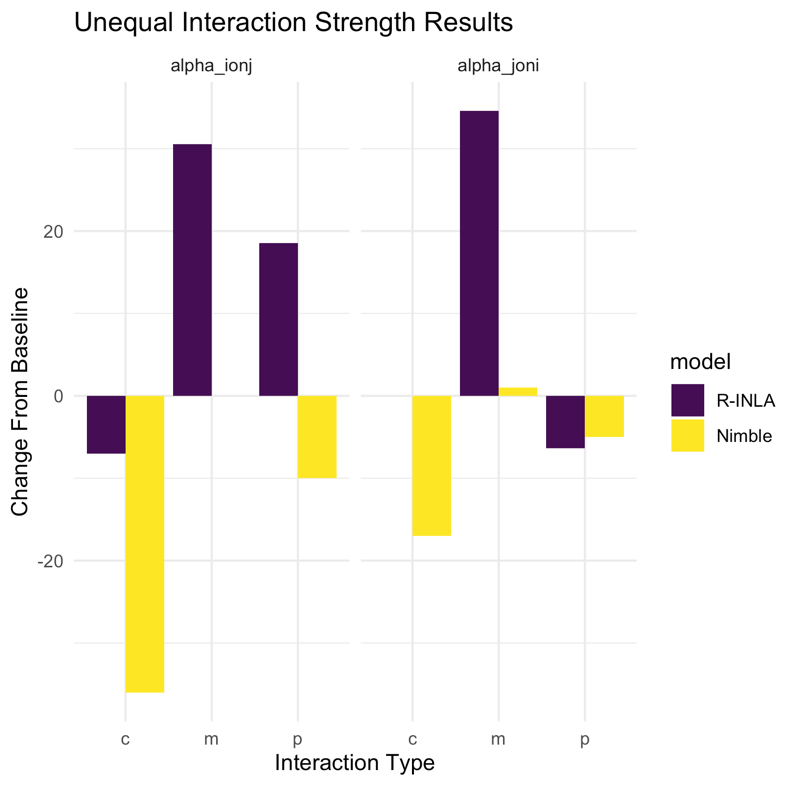
****

Figure 2: Results of unequal scenario simulations. Bar chart shows percentage of results that returned mean posterior estimates of the correct sign for interactive parameters from 100 simulated datasets

When interaction strengths are unequal the sign of the interaction parameters was retrieved less consistently for the Nimble model but more consistently for the R-INLA model. The competitive interactions using the Nimble model saw a particularly large shift with a decline in accuracy of > 20 percentage points for the species i on j interaction and 10 percentage points for the species j on i interaction respectively. The R-INLA model saw an improvement in sign accuracy for mutualist interactions when strengths were unequal (by more than 30 percentage points for both interactions). Predator prey interactions were less impacted for both models with an increase in accuracy for the species i on j interaction for the R-INLA model and only a slight decline in the percentage of times the sign was correctly retrieved for the species j on i interaction.

#### Increased environmental noise

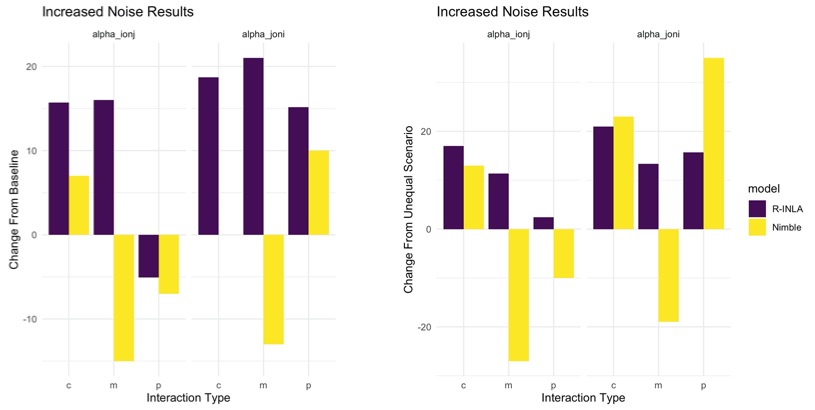


Figure 3: Results of increased variability in environmental noise scenario simulations. Panel A shows results for equal interaction strenghts, panel B shows results for unequal interaction strengths. Bar charts show percentage of results that returned mean posterior estimates of the correct sign for interactive parameters from 100 simulated datasets

Increasing environmental noise had no clear and consistent impact on the ability of either model to correctly retrieve the sign of interactive effects. Under this scenario there were improvements in the consistency of the retrieval of the sign of competitive interactions for both models and with both equal and unequal interaction strengths (see Figure 3) and for predator prey interaction of species j on i. The accuracy of the estimation of the sign of the predator prey interaction effects was improved for the Nimble model for the species j on i interaction under both equal and unequal interaction strengths however accuracy decreased for the species i on j interaction. A higher percentage of mutualist interactions were correctly estimated under equal and unequal interaction strengths for the R-INLA model. The accuracy of estimates of mutualist interactions was however reduced for the Nimble model for both scenarios of interaction strength.

#### Increased correlation in environmental effects

Changing the correlation in environmental noise for the two interacting species showed a threshold effect. When the correlation was changed from 0.3 to 0.5 results were largely unchanged and any changes were inconsistent. The general pattern of reducing the correlation was that the Nimble model performed less consistently and the R-INLA model performed more consistently. However increasing the correlation to 0.9 led to improvements in the estimation of all interaction effects (expect species i on j for the Nimble model and under a predator prey interaction). This led to most interactive effects being retrieved in the correct direction > 50 % of the time. There were specific improvements in the percentage of correctly estimated competitive and mutualist interactions for both model types (see Figure 5). An improvement in estimation of mutualist interactions under the highest correlation scenario was also seen for the INLA model.

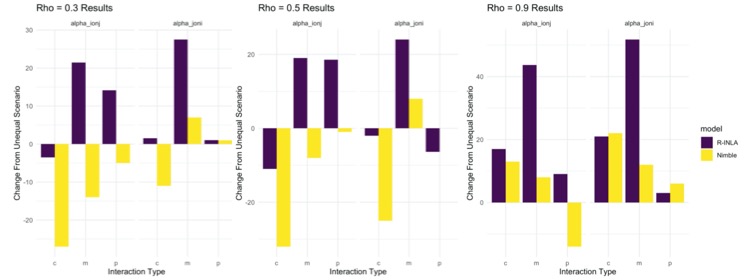


Figure 4: Results of increasing correlation in environmental noise scenario simulations. Panel A shows results for a correlation of 0.3, panel B shows results for a correlation of 0.5, panel C shows results for a correlation of 0.9. All results are for unequal interaction strengths. Bar charts show percentage of results that returned mean posterior estimates of the correct sign for interactive parameters from 100 simulated datasets

## Discussion

Suggested discussion points:

* Can clearly see from simulations that even weak interactive effects have a large impact on resulting population sizes. Therefore, they are ecologically important.
* However, models find them hard to retrieve because they are so weak.
* Highly simplified example here to demonstrate concept and even this struggles.
* Seems to fit with idea of weak interactions and hard to quantify. Based on model results, would assume that rarely were the interactions important. However, using interactions of this strength can have a large impact on population dynamics. So, focus on ‘statistical significance’ would be bad if you wanted to predict (either mathematically or postulate trends).
* Sign was not as well recovered as previous work (probably because interactions were not very strong). Did improve with higher correlation between species, so might work better on similar species as has been done in real work – often applied to many species of the same taxonomic group and same trophic level e.g. (Mutshinda, O’Hara and Woiwod, 2011,Bennett *et al.*, 2015,Stenseth *et al.*, 2015,Porzig *et al.*, 2016).
* Different model optimization methods seem to perform better under different conditions. Could use either, but be aware of limitations.
* Noise and strength of interactions do not seem so important. Correlation and interaction type is key. Could expect improvements if you know of interaction direction *a priori*. Recommend using in these contexts.

## Conclusion

* Recommendations: If using a naïve model with no strong priors then be careful of interpretation. Can still be useful models but do not expect to often find statistically significant results. Or accurate estimates of strength.
* But these weak interactions can still be important. But direction can be hard to determine as uncertainty is wide.
* Can use these models but maybe in context of where strong interaction expected and direction know rather than highly explorative across many species. More general point of don’t assume because it is not easily estimated that it is not important. Also, all of these ignore feedbacks between environmental change and the interaction – could be very important in future.

**Extra figures – rho was only returned as a single number for INLA it is not that Nimble failed to retrieve it correctly it just isn’t plotted**

**Chart, box and whisker chart

Description automatically generatedChart

Description automatically generatedChart

Description automatically generatedChart

Description automatically generatedChart

Description automatically generatedChart

Description automatically generatedChart, box and whisker chart

Description automatically generatedChart

Description automatically generatedChart, box and whisker chart

Description automatically generatedChart

Description automatically generatedChart, box and whisker chart

Description automatically generatedChart

Description automatically generatedChart, box and whisker chart

Description automatically generatedChart

Description automatically generated**