



**HAL**  
open science

# Non-random correlation of species dynamics in tropical tree communities

F. Jabot, T. Lohier

► **To cite this version:**

F. Jabot, T. Lohier. Non-random correlation of species dynamics in tropical tree communities. *Oikos*, 2016, 125 (12), pp.1733-1742. 10.1111/oik.03103 . hal-01852747

**HAL Id: hal-01852747**

**<https://hal.science/hal-01852747>**

Submitted on 2 Aug 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Non-random correlation of species dynamics in tropical tree communities

Franck Jabot<sup>1\*</sup> and Théophile Lohier<sup>1</sup>

<sup>1</sup>Irstea, UR LISC, Campus des Cézeaux 9, Avenue Blaise Pascal - CS 20085 63178 Aubière, France

\* Corresponding author: Franck Jabot

## Abstract

The importance of environmental stochasticity for tropical tree dynamics has been recently stressed by several studies. This has spurred the development of a “time-averaged neutral model” of community dynamics by Kalyuzhny and colleagues that extends the neutral model by incorporating environmental stochasticity. We here show that this framework can be used to assess the presence of non-random correlations between species dynamics. Indeed, the time-averaged neutral model makes the simplifying assumption that species responses to environmental variation are uncorrelated. We therefore propose to use this model as a null hypothesis against which observed community dynamics can be compared. This study makes five contributions. First, we describe a novel time-averaged neutral model of community dynamics that is close to, but more flexible than the one previously proposed by Kalyuzhny and colleagues. Second, we develop an inference method based on approximate Bayesian computation (ABC) and demonstrate the identifiability of the model parameters from community time series data. Third, we develop a test of the significance of environmental stochasticity, and a method to quantify its contribution to population variance. Fourth, we develop a test of non-random correlation between species dynamics. Fifth, we apply these developments to three datasets of tropical tree dynamics. We evidence both a strong contribution of environmental stochasticity to population variance in the three datasets, and a non-random correlation of species dynamics in one of them. We finally discuss the implications of these results for the modelling of tropical tree community dynamics.

## Introduction

The study of the assembly and dynamics of tropical tree communities has been largely influenced in the past fifteen years by neutral theory (Hubbell 2001, Rosindell et al. 2011). This theory considers as pivotal the role of demographic stochasticity in community dynamics, that is the random birth and death of tree individuals irrespective of the species they belong to. This theory has been shown to reproduce many community patterns including species abundance distributions (Volkov et al. 2003), beta-diversity (Condit et al. 2002) and phylogenetic imbalance (Jabot and Chave 2009). In contrast, when looking at community dynamics in community time series, several patterns cannot be reproduced by neutral models. In particular, the scaling of population size variance with respect to population size has been shown to be steeper in real communities than in neutral ones (Leigh 1981, Chisholm et al. 2014, Kalyuzhny et al. 2015).

The fact that populations of abundant species have a larger temporal variability than expected under demographic stochasticity indicates that the dynamics of conspecific individuals is correlated. One straightforward way to explain this correlated dynamics of conspecific individuals is to consider that they respond in the same manner to environmental temporal variability. This process has been named environmental stochasticity in the literature (Lande et al. 2003).

Kalyuzhny et al. (2015) made a critical contribution to the modelling of tropical tree community dynamics by showing that a community model incorporating both demographic and environmental stochasticity was able to reproduce a number of both static and dynamical patterns of the tropical tree community in Barro Colorado Island, including species abundance distribution, the scaling of population variance with population size, population fluctuation magnitude, and community temporal turn-over (Kalyuzhny et al. 2015). They called this model “the time-averaged neutral model”, since each individual has the same over-time-average prospect of birth and death irrespective of the species it

belongs to, although in contrast with a standard neutral model, individual birth rates differ among species at each time step.

In order to keep the model simple, they used a rather strong simplifying assumption: that species individual responses to environmental variability were uncorrelated (see methods). This simplifying assumption may appear doubtful. For instance, tree species differ in their sensitivity to drought (Engelbrecht et al. 2007), so that dry periods should benefit similarly to drought resistant species. And the same reasoning applies to any environmental component of species fitness (Hutchinson 1957). Still, an « effective » independence across species could result from the multi-dimensionality of environmental variation if species that respond similarly on one environmental axis do not co-vary on other axes. Recent analyses on tropical tree functional traits substantiate this possibility (Baraloto et al. 2010).

In addition, Kalyuzhny et al. (2015) also considered in their model that there was a temporal autocorrelation in environmental conditions (modelled through a within-species temporal autocorrelation in fecundity rates), making species fitness temporally autocorrelated. This *within*-species temporal autocorrelation indirectly produces some level of *between*-species correlation in their dynamics. It could therefore contribute to alleviating the potential model error due to the assumption of between-species independence.

The aim of this study is thus to apply a rigorous statistical framework to analyze tree community dynamics with environmental stochasticity, and to use this framework to assess both the contribution of environmental stochasticity to population variance and the presence of non-random correlations between species dynamics. In particular, we assess whether the model parameters can be reasonably well estimated from community time series data, notably the variance  $A$  and the temporal autocorrelation  $\rho$  of environmental stochasticity. We then develop a simple test of whether environmental stochasticity and within-species temporal autocorrelation are at play in a community using a model selection approach, and quantify their contributions to population variance using virtual communities.

Finally, we assess the presence of non-random correlations between species dynamics. These developments are applied to three datasets of tropical tree dynamics (Condit et al. 2006).

## Material and methods

### Modelling community dynamics with demographic and environmental stochasticity

We consider a model very similar to the one of Kalyuzhny et al. (2015) with two main differences. First, birth, immigration and death events are happening independently of each other at specified rates, so that there is no zero-sum rule and the community size can vary through time. Second, the temporal autocorrelation in environmental conditions is modelled for each species as a yearly draw from an autocorrelated lognormal distribution, rather than by independent draws every  $\tau$  years as in Kalyuzhny et al. (2015). These two modifications do not change the spirit of the model of Kalyuzhny et al. (2015). They were made to increase the generality of the model and to make it more flexible to accommodate observed temporal variations in community size (Table 1).

More precisely, in the model considered, during year  $t$ , each individual of species  $i$  present in the community 1) dies with a constant probability  $d$  independent of the species considered, and 2) produces descendants at a rate  $(1 - m) \times d \times f_i(t)$ , where  $m$  is the expected proportion of immigrants in the recruits, and the fecundity  $f_i(t)$  of species  $i$  during year  $t$  is drawn from a temporally autocorrelated lognormal distribution with mean 1, variance  $A$ , and within-species autocorrelation  $\rho$  on a log scale (i.e.,  $\text{cor}[\ln(f_i(t+1)), \ln(f_i(t))] = \rho$ ). Concretely, the number of descendants produced by an individual of species  $i$  during year  $t$  is drawn from a Poisson distribution with mean equal to  $(1 - m) \times d \times f_i(t)$ . This model of local community dynamics is thus neutral in the sense that every individual has the same *average over time* prospect of birth and death (Chave 2004), although birth rates differ among individuals at each time step. This is why Kalyuzhny et al. (2015) proposed to call this type of models, *time-averaged neutral models*, a terminology that we adopt here. The choice of the lognormal distribution to model among year and species variabilities in birth rates is a standard choice

for modelling environmental stochasticity (Lande et al. 2003). It is supported by previous analyses of tree demographic rates in tropical forests (Condit et al. 2006). We nonetheless compared this model choice to an alternative one based on a Weibull distribution that was shown to better fit bird dynamical properties in North America (Kalyuzhny et al. 2014b) and showed that the lognormal distribution provided a better fit to tropical tree data (Table S6). Importantly, alternative choices for the modelling of environmental stochasticity could be easily plugged into the approach detailed in this study.

On top of this local dynamics, immigration into the local community occurs at rate  $m \times d$ , so that the total number of immigrating individuals during year  $t$  is drawn from a Poisson distribution with mean equal to  $J(t) \times m \times d$ , where  $J(t)$  is the community size at time  $t$ . These immigrants are drawn from a neutral regional pool with regional diversity  $\theta$ , as in the neutral model of Hubbell (2001). Here, for computing efficiency, we model the species abundance distribution in the regional pool with a logseries distribution.

We additionally considered two degraded versions of this model: a model without temporal autocorrelation, in which  $\rho$  is equal to 0, and a model without environmental stochasticity in which  $A$  and  $\rho$  are equal to 0. We also considered an augmented version of the model incorporating a trend in community size. In this augmented version, we additionally consider that the community as a whole has a carrying capacity  $K$ , so that death rates during year  $t$  are no longer considered constant and equal to  $d$ , but they are equal to  $d \times J(t)/K$ . All the model versions considered are thus time-averaged neutral, except the one that does not include environmental stochasticity and is therefore neutral.

A commented R code to simulate this model of community dynamics with demographic and environmental stochasticity is provided in the supplementary material (Appendix S7).

## **Tropical tree data**

We make use of three datasets of tropical tree communities gathered by the Center for Tropical Forest Science, and published in Condit et al. (2006). The three 50 ha forest plots are located in Barro Colorado Island (BCI) in Panama, in Mudumalai in India, and in Pasoh in Peninsular Malaysia (Table 1). These three tropical tree communities have been censused more than three times, which is a minimum number of censuses to estimate both the variance and the temporal autocorrelation of environmental variability (BCI: 5 censuses, Mudumalai and Pasoh: 4 censuses, see Table 1). In these three forests, we will use all the tree individuals above 10 cm of diameter at breast height, which is a common arbitrary size threshold to distinguish adult trees from saplings (Condit et al. 2006).

### **Inference by ABC and a test of the presence of environmental stochasticity and temporal autocorrelation**

We developed an inference method for these models based on Approximate Bayesian Computation (ABC, Beaumont et al. 2002). ABC contains four main steps: 1) simulating a large number of times the model that is under study, with parameters drawn from prior distributions, 2) computing for each model simulation a set of summary statistics, 3) comparing these simulated summary statistics to the same set of statistics computed on the data, and 4) retaining the best fit simulations if their associated summary statistics are closer to the data summary statistics than an arbitrarily defined tolerance threshold. The parameter values used for these simulations form the posterior distribution of the parameters. A more detailed description of the ABC procedure for the present case study is given below.

Preliminary analyses revealed that successful inference could be performed with a set of 6 summary statistics: the average number of death events between two successive censuses  $N_{\text{death}}$  that is correlated with model parameter  $d$  (Fig. S1a), the average community size  $J$  that is correlated with  $K$  (Fig. S1f), the average species richness  $S$  across censuses that is correlated with  $m$  (Fig. S1b), the total number of species across censuses  $S_{\text{tot}}$  that is correlated with  $\theta$  (Fig. S1c), the average across species initially present of their standard

deviation of population sizes  $SD_{pop}$  that is correlated with  $A$  (Fig. S1d), and the average across species initially present of their temporal autocorrelation coefficients  $R(1)_{pop}$  that is correlated with  $\rho$  (Fig. S1e). This choice of summary statistics was largely inspired by the study of Kalyuzhny et al. (2015) and by previous works on the inference by ABC of neutral and non-neutral models (Jabot and Chave 2011). More precisely, some parameters jointly influence several summary statistics, notably  $m$  and  $\theta$  that both influence  $S$  and  $S_{tot.}$ , and  $A$  and  $\rho$  that both influence  $SD_{pop}$  and  $R(1)_{pop}$ . This will create correlations in the posterior distributions of these parameters, as we will see below.

We first performed a model selection through ABC, to assess whether all model parameters could be inferred, or whether there was some level of over-parameterization. For each dataset, we performed four sets of 20,000 ABC simulations with the four versions of the model: the full model with six parameters, and the three versions of the model without trend in community size: the models with environmental stochasticity with ( $\rho > 0$ ) and without ( $\rho = 0$ ) within-species autocorrelation, and the model without environmental stochasticity ( $A = 0$  and  $\rho = 0$ ). Uniform prior distributions were used:  $[0.01 ; 0.025]$  for  $d$ ,  $[0.001 ; 5]$  for  $A$ ,  $[0 ; 1]$  for  $\rho$ ,  $[\ln(0.001) ; \ln(0.1)]$  for  $\ln(m)$  and  $[\ln(1) ; \ln(10000)]$  for  $\ln(\theta)$ . Distinct uniform prior distributions were used for parameter  $K$  in the three datasets:  $[10000 ; 30000]$  for BCI,  $[1000 ; 15000]$  for Mudumalai, and  $[25000 ; 50000]$  for Pasoh. These prior distributions were chosen to encompass observed summary statistics in the three datasets. For each simulation, the model parameters were randomly drawn from the corresponding prior distributions. The community was initialized using the first census, and the community dynamics was simulated starting from this initial condition for the number of years separating the first from the last census of the studied dataset (20 years for the BCI dataset, 12 for Mudumalai and 13 for Pasoh). Note that this initialization procedure implies that parameter inference will be based solely on the community trajectory from this first census, as is classically done in community time-series analyses (e.g., Ives et al. 2003), and not on the likelihood of obtaining this initial community composition based on the model parameters evaluated. At the end of each simulation, the six summary statistics mentioned above were computed. They were normalized by their median absolute deviations across simulations (Csilléry et al. 2012). And the distance between the simulation and the data was

computed as the Euclidean distance in the 6-dimensional space of the summary statistics between the simulated and observed normalized summary statistics (Beaumont et al. 2002). Based on these four sets of ABC simulations for the four model versions, model selection was performed using the method of Beaumont et al. (2008) with a tolerance rate of 0.01 (meaning that the best  $0.01 \times 80000 = 800$  ABC simulations were retained), implemented in the R package “abc” (Csilléry et al. 2012). This method fits a multinomial logistic regression on the summary statistics based on the set of 800 retained simulations, and evaluates the four model posterior probabilities as the multinomial logistic prediction at the value of the summary statistics observed in the data. The idea of using model selection to assess the number of parameters that can be efficiently inferred within an ABC procedure was initially proposed in Lagarrigues et al. (2015). Here, this model selection approach is also a way to test for the presence of environmental stochasticity and within-species temporal autocorrelation: their presence will be supported if the models without these components are rejected.

We also assessed the quality of the fit of the selected models with a model checking procedure (Rubin 1984, Csilléry et al. 2010). This consists in assessing a posteriori that the summary statistics of the best-fit simulations are effectively close to the values observed in the data. For each summary statistics, we computed a p-value of model-data mismatch by computing the proportion of the 200 best-fit simulations that is equal or more extreme than the observed value. Since there is no a priori expectation of the direction of a potential misfit, this p-value must be lower than 0.025 to indicate a significant deviation at the 0.05 level. However, a correction for multiple testing must be used. We used here the Holm-Bonferroni correction.

### **Assessing the contribution of environmental stochasticity to population variance**

We quantified the respective influences of demographic and environmental stochasticities on observed mean population variances in the three datasets. For the BCI dataset, for which there was no trend in community size in the selected model (see results), we proceeded as

follows: we simulated community dynamics using model parameter values from the posterior distributions except that environmental variance  $A$  was fixed at 0. We then compared, for three species abundance classes (1-19, 20-199, >200), the median population variances  $VAR_{pop\_demo}$  obtained without environmental stochasticity to the observed median population variances  $VAR_{pop\_obs}$ . The contribution of demographic stochasticity to observed population variances was quantified as  $VAR_{pop\_demo} / VAR_{pop\_obs}$  and the contribution of environmental stochasticity was quantified as  $1 - VAR_{pop\_demo} / VAR_{pop\_obs}$ .

For the two other datasets for which there was a trend in community size in the selected model, we proceeded slightly differently. Indeed in these datasets, the trend in community size also contributes to the overall population variances. We also simulated community dynamics using model parameter values from the posterior distributions, but we performed two sets of simulations: one with environmental variance and the second without environmental variance ( $A$  fixed at 0). In these two sets of simulations, the fitted trend in community size was removed, using death rates independent of community size, with values drawn from the posterior distribution of parameter  $d$ . The contribution of demographic stochasticity to population variance was quantified as above as  $VAR_{pop\_demo} / VAR_{pop\_obs}$  while the contribution of environmental stochasticity to population variance was quantified as  $(VAR_{pop\_env} - VAR_{pop\_demo}) / VAR_{pop\_obs}$ , where  $VAR_{pop\_env}$  is the median population variance obtained in the simulation set with both demographic and environmental variances. The remaining part  $1 - (VAR_{pop\_env} / VAR_{pop\_obs})$  quantifies the contribution of the trend in community size to population variance.

### **Assessing non-random correlations between species dynamics**

Due to the short length of the time series considered (5 censuses in BCI, 4 in the two other datasets), direct tests of between-species correlations in their dynamics would have a low power. Besides, the *within*-species temporal autocorrelation incorporated in two of the four models considered can potentially generate some level of *between*-species correlations in their dynamics. Consequently, one needs to take this model component into account when

assessing the assumption of species independence in their responses to environmental conditions. We thus developed a test of the presence of non-random correlations between species dynamics, using a model checking procedure (Rubin 1984, Csilléry et al. 2010). We proceeded in two steps. First, a more precise parameter inference with the selected model was performed using the sequential ABC algorithm of Lenormand et al. (2013) within the R package “EasyABC” (Jabot et al. 2013). This procedure required 14,500 ABC simulations for the BCI dataset, 14,000 for Mudumalai, and 19,000 for Pasoh. Second, we sampled parameters from the obtained posterior distribution and computed, for each simulation below the tolerance threshold obtained by the sequential algorithm, the correlation matrix of the most abundant species with an initial abundance of at least 50 individuals. This correlation matrix was computed as follows: for two species  $i$  and  $j$ , the term  $(i,j)$  in the matrix was equal to the correlation coefficient  $Cor_{ij}$  between the two trajectories of population sizes across censuses of species  $i$  and  $j$ . 200 such correlation matrices were computed and they form the posterior predictive checks that can be compared to the correlation matrix computed from the data (Csilléry et al. 2010). Rather than comparing each correlation coefficient one by one, we compared a unique index: the average of the absolute values of the matrix coefficients  $|Cor_{pop}|$ .

We finally replicated the model checking analysis for the BCI dataset, using either the first three censuses or the first four censuses, to assess whether considering shorter time series might change our understanding of the correlations between species trajectories.

## Results

### **Parameter identifiability and a test for the presence of environmental stochasticity and within-species temporal autocorrelation**

Models with temporally autocorrelated environmental stochasticity were selected against the simpler nested models in the three datasets (Table 2), and their fits of the summary statistics were overall satisfactory (Table S2). This indicates that the model studied is not over-parameterized and that both environmental stochasticity and within-species temporal

autocorrelation are critical to explain the observed dynamics in the three tropical forests considered. In two of the three datasets (Mudumalai and Pasoh), a model further incorporating a trend in community size was selected. This result reminds that the standard zero-sum rule abundantly used since the work of Hubbell (2001) should not be seen as a panacea for modelling community dynamics, despite its attractive simplicity.

Selected models had an overall satisfactory fit of the six summary statistics used in the inference (Fig. 1, Table S2). For the BCI dataset, two p-values were below the threshold of 0.025 (those associated with the average  $S$  and total  $S_{\text{tot}}$  species richness). However, when applying the Holm-Bonferroni correction for multiple testing, they were no longer significant (since the corrected thresholds become equal to  $0.025/6 = 0.0041$  and to  $0.025/5 = 0.005$  for the two lowest p-values respectively). Still, this model checking procedure highlights that the fitted model tends to underestimate the average species richness  $S$ , while it tends to overestimate the total among censuses species richness  $S_{\text{tot}}$ . This model-data mismatch is due to a larger mortality rate of very rare species in the data compared to the one of more abundant species (data not shown), a pattern that could not be reproduced with the models studied. Interestingly, the selected model for BCI was also able to reproduce observed values of two additional summary statistics not included in the inferential procedure: the average Shannon's index  $H$  across censuses, and the average Bray-Curtis similarity  $BC$  between each census and the initial one (Fig. S3). It was also able to marginally reproduce the number of times  $N_{\text{trans}}$  that a population increase between two censuses is followed by a population decrease or the reverse (Fig. S3). These complementary results strengthen the demonstration of the overall good fit quality of the selected model regarding community-level static and dynamical properties. Similar results were obtained for the Mudumalai dataset (Table S2). For the Pasoh dataset, even the full model had a small but significant model-data mismatch for the average number of species  $S$  (Table S2). It is worth noting that for both BCI and Mudumalai, several models had good quality fits according to this model checking procedure (Table S2). The Bayesian model selection procedure still succeeded in clearly ranking these competing models (Table 2).

For the BCI dataset, estimations of the mortality rate  $d$  and of the temporal autocorrelation  $\rho$  had low uncertainty with narrow posterior distributions (Fig. 2a,c). The environmental variance  $A$  was also well estimated, although with a wider posterior distribution (Fig. 2c). The parameters of immigration  $m$  and of regional diversity  $\theta$  had wider posterior distributions (Fig. 2b), which is in line with previous studies (Etienne et al. 2006, Jabot and Chave 2009). Similar results were obtained with the two other datasets, with the additional result that the community carrying capacity parameter  $K$  also had a relatively narrow posterior distribution for these two datasets (Fig. S4-5).

We found larger environmental variances ( $A = 2.2$ , Fig. 1f) for BCI than in the study of Kalyuzhny et al. (2015) in which it was estimated at 0.8. This large environmental variance was also recovered in the Mudumalai dataset ( $A = 3$ ), but less so in Pasoh ( $A = 0.94$ , Fig. S4-5). We also found that parameters  $A$  and  $\rho$  were negatively correlated (Fig. 2c). This last result explains the discrepancy between our estimate of  $A$  and the one of Kalyuzhny et al. (2015) who considered a strong temporal autocorrelation in their model in which environmental variables were assumed constant during 10 years. Our analysis thus demonstrates that a larger environmental variance associated with a less strong temporal autocorrelation is more likely (Fig. 2c).

### **Assessing the contribution of environmental stochasticity to population variance**

Values of temporal autocorrelation were variable among the three datasets, varying from 0.60 at Pasoh to 0.68 at BCI and 0.92 at Mudumalai. The resulting contributions of environmental and demographic stochasticity to population variance were found to be consistent across the studied sites. In BCI, where no trend in community size was observed, demographic stochasticity was the main driver of population variance for rare species, and it was progressively surpassed by the influence of environmental stochasticity for more abundant species (Fig. 3a). In Mudumalai where the community size decreased by more than 16% in 12 years, the trend in community size explained the larger part of population variance (between 49 and 77% depending on the abundance class, Fig. 3b). The remaining

part of the population variance was again largely explained by demographic stochasticity for rare species, while environmental stochasticity was the main driver for more abundant species (Fig. 3b). In Pasoh where the community size increased by 6% in 13 years, the trend in community size also explained a large part of population variance (Fig. 3c). The remaining part was again mainly due to demographic stochasticity for rare species, and to environmental stochasticity for abundant ones (Fig. 3c).

### **Non-random correlations between species dynamics**

Observed correlation coefficients between population trajectories tended to be larger in absolute values than the ones computed on simulated community dynamics, with a larger number of strongly positively or negatively correlated population trajectories (Fig. 4a). This trend was significant for the BCI dataset ( $p < 0.005$ , Fig. 4b). This indicates that observed levels of correlations between species dynamics cannot be explained by the time-averaged neutral model in this dataset. Non-random correlations between species in their responses to environmental variability thus need to be invoked. In the two other datasets, the same test led to the same trends of stronger between-species correlations in empirical data than awaited under the time-averaged neutral model (Fig. 4c,d). These trends were however non-significant ( $p = 0.09$  in Mudumalai,  $p = 0.24$  in Pasoh, Fig. 4d). One possible explanation for this lack of significance for these two datasets may be that these communities were studied during a period of directional trend in community size which seems to be an important driver of the dynamics of these two communities (Fig. 3). Another possible explanation is the shorter length of the time series for these two datasets. Indeed, we found that deviations from model predictions were increasing with the number of censuses considered in the analyses for the BCI dataset (Fig. 5).

### **Discussion**

Several recent studies have proposed to model plant community dynamics with both demographic and environmental stochasticity (de Mazancourt et al. 2013, Kalyuzhny et al. 2015), following the seminal presentation of Lande et al. (2003). To further advance our

understanding of the consequences of environmental stochasticity on community assembly and dynamics, rigorous inference methods need to be developed for these models, as was done previously with neutral and non-neutral dynamical community models (Beeravolu et al. 2009, Jabot and Chave 2011). The present contribution provides a quantitative Bayesian inferential framework for slightly modified versions of the time-averaged neutral model recently proposed by Kalyuzhny et al. (2015).

We showed that it is possible to estimate both the variance  $A$  and the temporal autocorrelation  $\rho$  of environmental stochasticity from community time series of marked individuals (Fig. 2), and we provided a simple way to assess the importance of environmental stochasticity through a model selection technique. We recovered the results of Chisholm et al. (2014) and of Kalyuzhny et al. (2015) that environmental stochasticity was critical to understand the dynamics of tropical tree communities (Table 2). We further showed that environmental stochasticity was mostly important for abundant species, while rare species were mostly impacted by demographic stochasticity (Fig. 3). This may contribute to the fact that hyper-diverse tropical forests with many rare species often show neutral-like patterns (Jabot and Chave 2011, Kubota et al. 2015, Qiao et al. 2015).

Our analysis further revealed that environmental variance was larger and temporal autocorrelation in environmental conditions lower (Fig. 2) than previously found by Kalyuzhny et al. (2015). Still, the presence of relatively large year-to-year autocorrelation in species demographic rates, already evidenced by Feeley et al. (2011) in BCI, calls for a biological explanation. Indeed, climatic variability does not show large year-to-year autocorrelation (Lande et al. 2003), so that other sources of autocorrelation needs to be invoked. First, gap dynamics is likely to produce long standing (decadal) conditions favoring some species over others in tree communities, depending on species light requirements and their sensitivities to crowding (Shugart 1984, Uriarte et al. 2004). Second, the long generation time of trees and their capacity to store resources (Chapin et al. 1990, Poorter and Kitajima 2007) is likely to dampen their response to yearly environmental fluctuations and thus to increase the autocorrelation of the stochastic environmental variables. Third,

synchronizing processes such as mast fruiting (Norden et al. 2007) or extreme climatic events (Brando et al. 2014) may also tend to increase temporal autocorrelation. The inferred strong autocorrelation in stochastic environmental variables is thus biologically plausible. We must however recognize that our inference of autocorrelation strength present a non-negligible level of uncertainty (Fig. 2c, Fig. S4c, Fig. S5c) and that it is necessarily dependent on our (reasonable) model assumptions. Reported autocorrelation levels should thus be interpreted and used with caution.

Previous studies have investigated the signature and magnitude of environmental stochasticity in various types of ecological communities, using various mathematical approaches. For instance, model-based analyses of community time-series with likelihood (Ives et al. 2003) or Bayesian techniques (Mutshinda et al. 2009, 2011) have been proposed. These parameter-rich approaches require relatively long time-series so as to obtain reliable inferences and were consequently applied to a limited number of systems that have been intensively studied for a long time, with typically several dozens of censuses. Using these techniques, Mutshinda et al. (2009, 2011) demonstrated that environmental stochasticity exerted a much stronger influence on community dynamics than demographic stochasticity and interspecific interactions for a diverse range of animal communities. Similar conclusions were obtained using model-free analyses of dynamical patterns for both plant and animal communities (e.g., Houlahan et al. 2007, Chisholm et al. 2014, Kalyuzhny et al. 2014a,b). The main advantages of these last approaches are that 1) they are not restricted to massive datasets, and 2) they are more flexible, in that they do not rely on necessarily debatable model assumptions. They constitute in this sense attractive ways to design pertinent model assumptions that can be later tested with model-based approaches. Inferences based on time-averaged neutral models are thus complementary to these existing approaches, in that they are a model-based approach that is “parameter-economical”, and can thus be used on less massive datasets. They corroborated in this study the previous findings of the prevailing role of environmental stochasticity in shaping community dynamics, especially for abundant species (Fig. 3).

A critical simplifying assumption of the time-averaged neutral model is that species respond independently of each other to environmental variability. The second main goal of this study was thus to assess the validity of this strong assumption of species independence. We performed this assessment using a model checking procedure comparing the correlation matrix of species trajectories in both observed and best-fit simulated communities. This assumption was rejected for the BCI dataset ( $p < 0.005$ ), but not for the two other ones (Fig. 4). These results are unlikely to be biased by the (non-significant) model-data mismatch surrounding the dynamics of very rare species, since only abundant species (>49 individuals) were considered in the correlation matrix. These results show that non-random correlations between species dynamics need to be invoked to explain observed community dynamics, at least in some study sites. They nevertheless show that deviations from the predictions of the time-averaged neutral model are not systematic, neither very large. Indeed, at BCI, between-species correlation was found to be equal to 0.72 on average, a number only modestly larger than the predictions of the time-averaged neutral model that were located around 0.6 (Fig. 4b).

We then tried to understand what could explain this relatively modest failure of the model predictions on between-species correlations. We evidenced that the deviations from model predictions build over very long time, thereby making short time series insufficient to detect statistically significant deviations (Fig. 5). A more positive interpretation of these last results is that the simplifying assumption of species independence proposed by Kalyuzhny et al. (2015) seems to be sufficient for some communities, and up to twenty year-long temporal horizons.

An important question is whether one can use time-averaged neutral models in communities in which the assumption of independence is violated, as in BCI. The present work and the previous one of Kalyuzhny et al. (2015) shows that the model is still valid if one wants to make general predictions on community-wide dynamical patterns such as average population variability and its scaling with population abundance, or the respective contributions of demographic and environmental stochasticity to the observed community

dynamics. In contrast, this model is unlikely to make good predictions in such cases regarding between-species correlations in their demography.

More generally, the present contribution employs a quantitative statistical framework to analyze community time series with time-averaged neutral models. Being able to quantify environmental stochasticity in ecological communities from community time series opens exciting new questions for ecological research. Theoretical predictions could be developed on how environmental stochasticity might shape spatial and temporal patterns in metacommunities (Hubbell 2001, Chave et al. 2002, Ulrich 2004, Economo and Keitt 2010, May et al. 2015) and anthropogenic impacts on biodiversity (Solé et al. 2004, Gilbert et al. 2006, Hubbell et al. 2008), as well as how diversity indices may depart from null predictions incorporating environmental stochasticity (Münkemüller et al. 2012). A decade of research did not deplete the study of neutral models based on demographic stochasticity (Rosindell et al. 2011). This myriad of works now needs to be generalized to embrace the combined effect of demographic and environmental stochasticity.

As a concluding remark, it may sound abusive to use the terminology of “time-average neutrality” to designate a model including both demographic and environmental stochasticity, since what drives environmental stochasticity is classically referred to as niche processes (as opposed to neutral processes) in the ecological literature. Actually, the distinction between “strong” and “weak” versions of the equivalence assumption of the neutral theory was stressed early on during the maturation of this theory (Chave 2004), where “strong” neutrality refers to the strict fitness equivalence among individuals, while “weak” neutrality refers only to average fitness equivalence among individuals, as in time-averaged neutral models. A good reason for calling neutral these “weakly” neutral models is that individuals still behave neutrally from an evolutionary standpoint. Beyond this semantic discussion, the interesting originality of time-averaged neutral models is their way of merging what is commonly referred to as “niche” and “neutral” processes. Indeed, previously proposed syntheses of these two types of processes are commonly relying on the addition of neutral-based noise to niche-based processes that generate average fitness

differences among species (e.g., Tilman 2004, Gravel et al. 2006, Jabot and Chave 2011). In contrast, time-averaged neutral models integrate niche-based differences among species without breaking the weak version of the equivalence assumption among individuals (Alonso et al. 2007, Kalyuzhny et al. 2015).

## **Acknowledgements**

F.J. was funded by the Irstea Dynindic project. T.L. was funded by Irstea and the Regional Council of Auvergne. We thank the Subject Editor and three anonymous reviewers for their numerous suggestions that greatly improved this manuscript. We thank the volunteers, workers and data providers of the Center for Tropical Forest Science for having made their datasets publicly available.

## References

- Alonso, D. et al. 2007. Response to Benedetti-Cecchi: neutrality and environmental fluctuations. – *Trends Ecol. Evol.* 22: 232.
- Baraloto, C. et al. 2010. Decoupled leaf and stem economics in rain forest trees. – *Ecol. Lett.* 13: 1338-1347.
- Beaumont, M. A. et al. 2002. Approximate Bayesian computation in population genetics. – *Genetics* 162: 2025-2035.
- Beaumont, M. A. 2008. Joint determination of topology, divergence time, and immigration in population trees. – In : Renfrew, C. et al. (ed.), *Simulation, Genetics and Human Prehistory*, McDonald Institute Monographs, pp 134-154.
- Beeravolu, C. R. et al. 2009. Studying ecological communities from a neutral standpoint: a review of models' structure and parameter estimation. – *Ecol. Model.* 220: 2603-2610.
- Brando, P. M. et al. 2014. Abrupt increases in Amazonian tree mortality due to drought-fire interactions. – *Proc. Natl. Acad. Sci. USA* 111: 6347-6352.
- Chapin, F. S. et al. 1990. – The ecology and economics of storage in plants. – *Ann. Rev. Ecol. Sys.* 21: 423-447.
- Chave, J. 2004. Neutral theory and community ecology. – *Ecol. Lett.* 7: 241-253.
- Chave, J. et al. 2002. Comparing classical community models: theoretical consequences for patterns of diversity. *Am. Nat.* 159: 1-23.

- Chisholm, R. A. et al. 2014. Temporal variability of forest communities: empirical estimates of population change in 4000 tree species. – *Ecol. Lett.* 17: 855-865.
- Condit, R. et al. 2002. Beta-diversity in tropical forest trees. – *Science* 295: 666-669.
- Condit, R. et al. 2006. The importance of demographic niches to tree diversity. – *Science* 313: 98-101.
- Csilléry, K. et al. 2010. Approximate Bayesian computation (ABC) in practice. – *Trends Ecol. Evol.* 25: 410-418.
- Csilléry, K. et al. 2012. abc: an R package for approximate Bayesian computation (ABC). – *Methods Ecol. Evol.* 3: 475-479.
- De Mazancourt, C. et al. 2013. Predicting ecosystem stability from community composition and biodiversity. – *Ecol. Lett.* 16: 617-625.
- Economu, E. P. and Keitt, T. H. 2010. Network isolation and local diversity in neutral metacommunities. – *Oikos* 119: 1355-1363.
- Engelbrecht, B. M. et al. 2007. Drought sensitivity shapes species distribution patterns in tropical forests. – *Nature* 447: 80-82.
- Etienne, R. S. et al. 2006. Comment on "Neutral Ecological Theory Reveals Isolation and Rapid Speciation in a Biodiversity Hot Spot". – *Science* 311: 610-610.
- Feeley, K. J. et al. 2011. Directional changes in the species composition of a tropical forest. – *Ecology* 92: 871-882.

- Gilbert, B. et al. 2006. Can Neutral Theory Predict the Responses of Amazonian Tree Communities to Forest Fragmentation? – *Am. Nat.* 168: 304-317.
- Gravel, D. et al. 2006. Reconciling niche and neutrality: the continuum hypothesis. – *Ecol. Lett.* 9: 399-409.
- Houlahan, J. E. et al. 2007. Compensatory dynamics are rare in natural ecological communities. – *Proc. Natl. Acad. Sci. USA* 104: 3273-3277.
- Hubbell, S. P. 2001. *The unified neutral theory of biodiversity and biogeography.* – Princeton Univ. Press.
- Hubbell, S. P. et al. 2008. How many tree species are there in the Amazon and how many of them will go extinct? – *Proc. Natl Acad. Sci. USA* 105: 11498-11504.
- Hutchinson, G. E. 1957. Concluding remarks. – *Cold spring harbor symposium on quantitative biology.* 22: 415-427.
- Ives, A. R. et al. 2003. Estimating community stability and ecological interactions from time-series data. – *Ecol. Monog.* 73: 301-330.
- Jabot, F. and Chave, J. 2009. Inferring the parameters of the neutral theory of biodiversity using phylogenetic information and implications for tropical forests. – *Ecol. Lett.* 12: 239-248.
- Jabot, F. and Chave, J. 2011. Analyzing tropical forest tree species abundance distributions using a nonneutral model and through approximate Bayesian inference. – *Am. Nat.* 178: E37-E47.

- Jabot, F. et al. 2013. EasyABC: performing efficient approximate Bayesian computation sampling schemes using R. – *Methods Ecol. Evol.* 4: 684-687.
- Kalyuzhny, M. et al. 2014a. Temporal fluctuation scaling in populations and communities. – *Ecology* 95: 1701-1709.
- Kalyuzhny, M. et al. 2014b. Niche versus neutrality: a dynamical analysis. – *Am. Nat.* 184: 439-446.
- Kalyuzhny, M. et al. 2015. A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. – *Ecol. Lett.* 18: 572-580.
- Kubota, Y. et al. 2015. Non-neutrality in forest communities: evolutionary and ecological determinants of tree species abundance distributions. – *Oikos* doi:10.1111/oik.02232.
- Lagarrigues, G. et al. 2015. Approximate Bayesian computation to recalibrate individual-based models with population data: Illustration with a forest simulation model. – *Ecol. Model.* 306: 278-286.
- Lenormand, M. et al. 2013. Adaptive approximate Bayesian computation for complex models. – *Comput. Stat.* 28: 2777-2796.
- Lande, R. et al. 2003. Stochastic population dynamics in ecology and conservation. – Oxford Univ. Press.
- Leigh, E. G. 1981. The average lifetime of a population in a varying environment. – *J. Theor. Biol.* 90: 213-239.

- May, F. et al. 2015. Moving beyond abundance distributions: neutral theory and spatial patterns in a tropical forest. – *Proc. Roy. Soc. B.* 282: 20141657.
- Münkemüller, T. et al. 2012. From diversity indices to community assembly processes: a test with simulated data. – *Ecography* 35: 468-480.
- Mutshinda, C. M. et al. 2009. What drives community dynamics? – *Proc. Roy. Soc. B.* 276: 2923-2929.
- Mutshinda, C. M. et al. 2011. A multispecies perspective on ecological impacts of climatic forcing. – *J. Anim. Ecol.* 80: 101-107.
- Norden, N. et al. 2007. Mast fruiting is a frequent strategy in woody species of Eastern South America. – *Plos One* 2: e1079.
- Poorter, L. and Kitajima, K. 2007. Carbohydrate storage and light requirements of tropical moist and dry forest tree species. – *Ecology* 88: 1000-1011.
- Qiao, X. et al. 2015. A latitudinal gradient in tree community assembly processes evidenced in Chinese forests. – *Global Ecol. Biogeogr.* 24: 314-323.
- Rosindell, J. et al. 2011. The unified neutral theory of biodiversity and biogeography at age ten. – *Trends Ecol. Evol.* 26: 340-348.
- Rubin, D. B. 1984. Bayesianly justifiable and relevant frequency calculations for the applied statistician. – *Ann. Stat.* 12: 1151-1172.
- Shugart, H. H. 1984. A theory of forest dynamics: the ecological implications of forest succession models. – Springer-Verlag, New York.

- Solé, R. V. et al. 2004. Habitat fragmentation and biodiversity collapse in neutral communities. – *Ecol. Comp.* 1: 65-75.
- Tilman, D. 2004. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion and community assembly. – *Proc. Natl. Acad. Sci. USA* 101: 10854-10861.
- Ulrich, W. 2004. Species co-occurrences and neutral models: reassessing JM Diamond's assembly rules. – *Oikos* 107: 603-609.
- Uriarte, M. et al. 2004. A spatially explicit model of sapling growth in a tropical forest: does the identity of neighbours matter? – *J. Ecol.* 92: 348-360.
- Volkov, I. et al. 2003. Neutral theory and relative species abundance in ecology. – *Nature* 424: 1035-1037.

Table 1. Characteristics of the tropical tree datasets.

Site	Site area	Number of censuses used	Census date	Community size	Species richness
Colorado Island	50 ha	5	1985	20640	235
			1990	21176	227
			1995	21404	225
			2000	21148	225
			2005	20848	227
Mudumalai	50 ha	4	1988	15033	62
			1992	14028	63
			1996	13070	63
			2000	12574	61
Pasoh	50 ha	4	1987	26550	678
			1990	27659	666
			1995	29257	674
			2000	28279	671

Table 2. Posterior model probabilities for the three datasets. For each dataset, the probability of the selected model is in bold.

Site	BCI	Mudumalai	Pasoh
Dem. stochasticity ( $A = 0$ and $\rho = 0$ )	0	0	0
Dem. and env. stochasticity ( $\rho = 0$ )	0.03	0	0
Dem. and env. stoch. with autocorrelation	<b>0.63</b>	0.03	0.0002
Full model with trend in community size (K)	0.35	<b>0.97</b>	<b>0.9998</b>

Figure 1. Model checking plots for the BCI dataset and the selected model (see Table 2).

Histograms represent the simulated values of the various summary statistics for the 200 best-fit simulations. Observed values are depicted by the vertical grey lines, and the corresponding p-values are indicated.

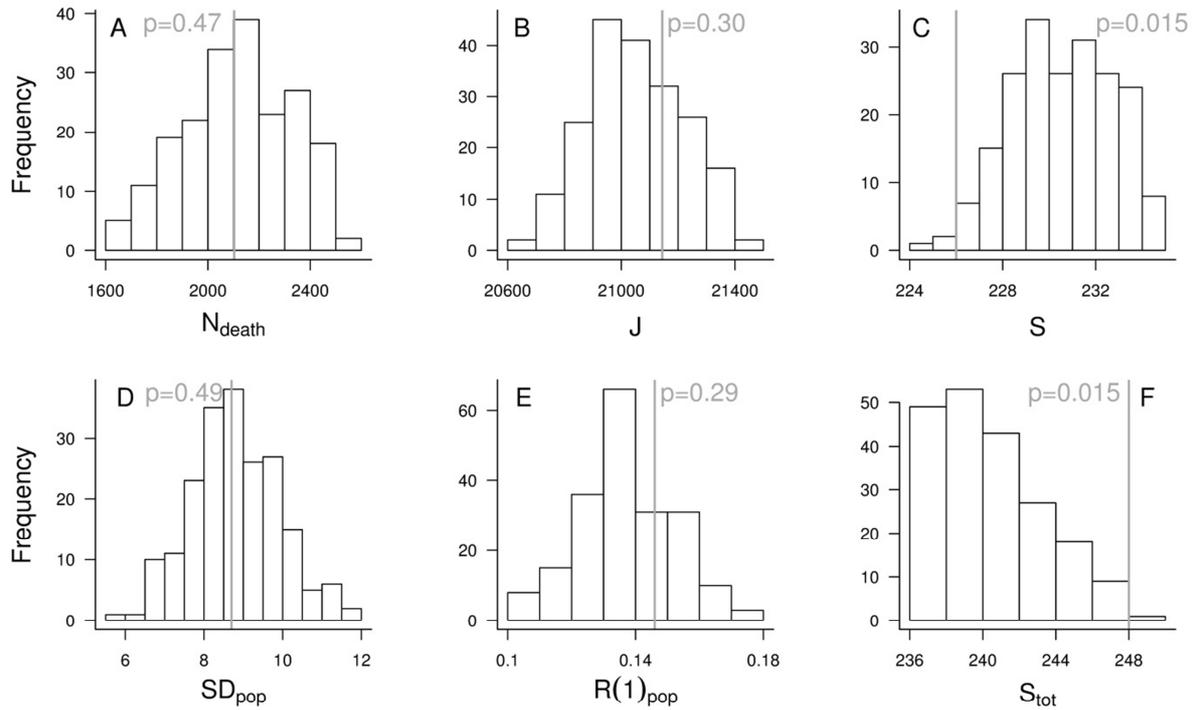


Figure 2. Posterior distributions for the BCI dataset. Panel A: marginal posterior distribution for parameter  $d$ . Panels B and C: joint posterior distributions for parameters  $(\theta, m)$  and  $(A, \rho)$  respectively, highlighting the correlations between these parameters. In these panels, points represent posterior samples obtained by ABC, the grey levels stand for posterior density, and the black lines are iso-density curves.

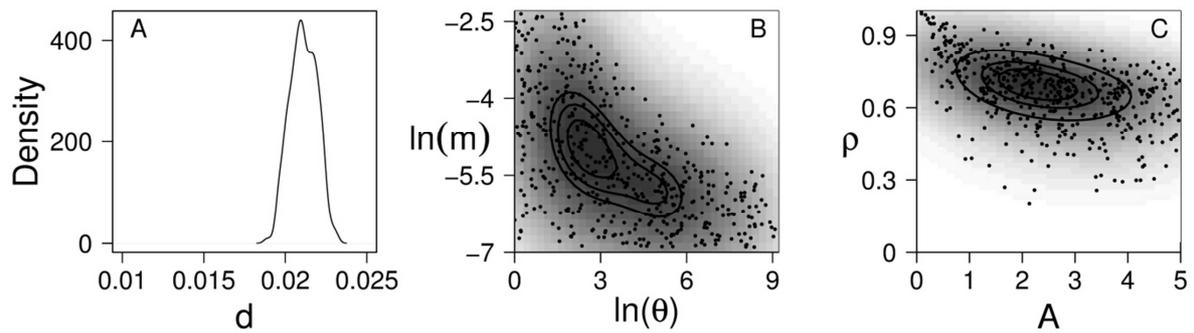


Figure 3. Contributions of demographic and environmental stochasticities to population variances. Panels A, B and C refers to the BCI, Mudumalai and Pasoh datasets respectively. The proportion of population variance due to demographic stochasticity is depicted in black, the one due to environmental stochasticity is in dark grey, and the variance due to the trend in community size is in light grey.

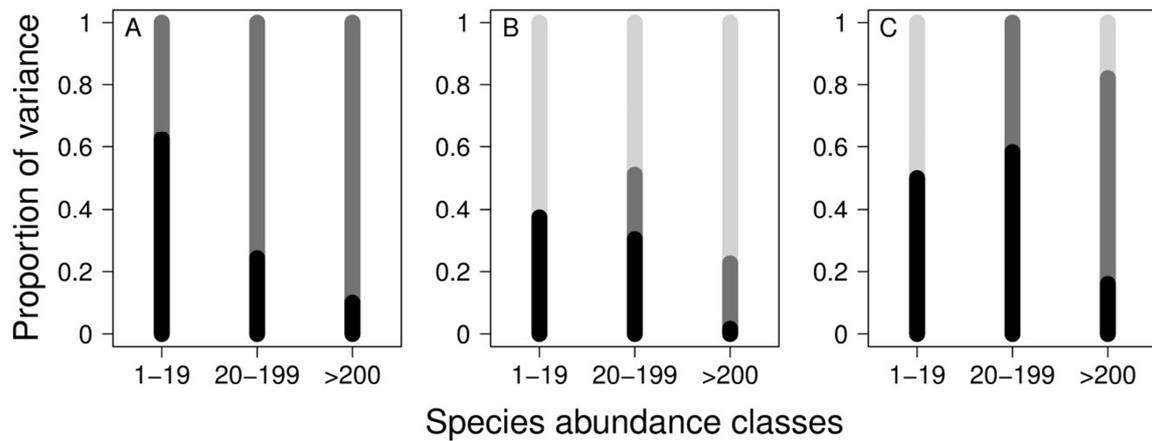


Figure 4. Non-random correlations between species dynamics. Panel A: Distribution of correlation coefficients in the BCI plot (grey) and in a typical simulation fitting the BCI plot (white). Panels B, C, D: Distributions of the average absolute values of correlation coefficients  $|\text{Cor}_{\text{pop}}|$  in the 200 best-fit simulations for the BCI, Mudumalai and Pasoh datasets. The vertical lines stand for observed values of this index in the three plots.

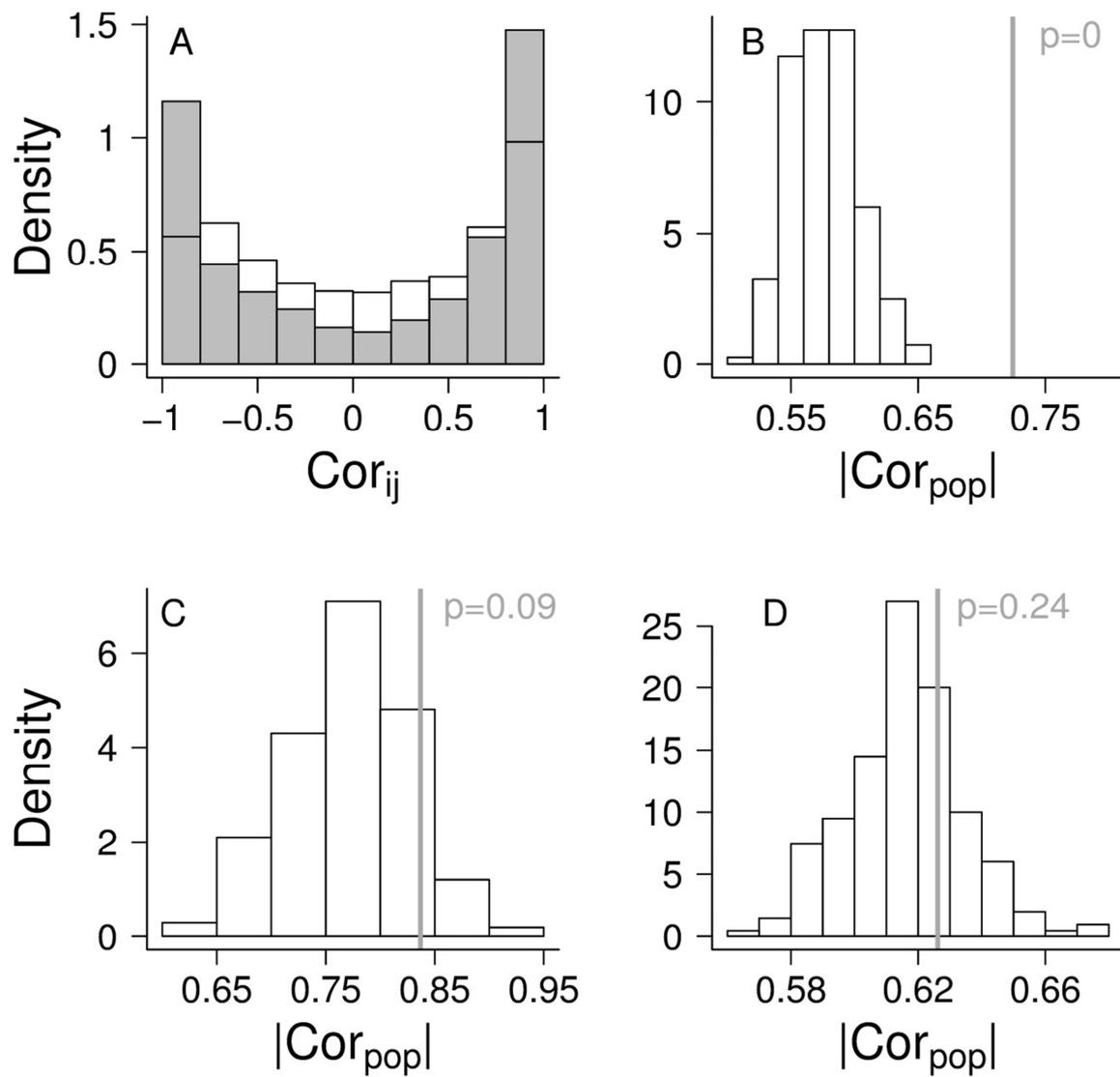


Figure 5. Lower model failure with shorter time series. Panels A, B: Distributions of the average absolute values of correlation coefficients  $|Cor_{pop}|$  in the 200 best-fit simulations in BCI, using either the first three or the first four censuses instead of the five censuses as in Fig. 4b. The vertical lines stand for observed values of this index in the plot.

