A peer-reviewed version of this preprint was published in PeerJ on 17 May 2017.

Allen KA, Bruno JF, Chong F, Clancy D, McClanahan TR, Spencer M, Żychaluk K. 2017. Among-site variability in the stochastic dynamics of East African coral reefs. PeerJ 5:e3290
https://doi.org/10.7717/peerj.3290

Among-site variability in the stochastic dynamics of East African coral reefs

Katherine A Allen [1,2], **John F Bruno** [3], **Fiona Chong** [1], **Damian Clancy** [4], **Tim R McClanahan** [5], **Matthew Spencer** [Corresp., 1], **Kamila Zychaluk** [6]

[1] School of Environmental Sciences, University of Liverpool, Liverpool, United Kingdom
[2] Institute of Integrative Biology, University of Liverpool, Liverpool, United Kingdom
[3] Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States
[4] School of Mathematical and Computer Sciences, Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, United Kingdom
[5] Wildlife Conservation Society, New York, United States
[6] Department of Mathematical Sciences, University of Liverpool, Liverpool, United Kingdom

Corresponding Author: Matthew Spencer
Email address: m.spencer@liverpool.ac.uk

Coral reefs are dynamic systems whose composition is highly influenced by unpredictable biotic and abiotic factors. Understanding the spatial scale at which long-term predictions of reef composition can be made will be crucial for guiding conservation efforts. Using a 22-year time series of benthic composition data from 20 reefs on the Kenyan and Tanzanian coast, we developed Bayesian vector autoregressive state-space models for reef dynamics, incorporating among-site variability, and quantified their long-term behaviour. We estimated that if there were no among-site variability, the total long-term variability would be approximately one third of its current value. Thus, our results showed that among-site variability contributes more to long-term variability in reef composition than does temporal variability. Individual sites were more predictable than previously thought, and predictions based on current snapshots are informative about long-term properties. Our approach allowed us to identify a subset of possible climate refugia sites with high conservation value, where the long-term probability of coral cover $\leq$ 0.1 was very low. Analytical results show that this probability is most strongly influenced by among-site variability and by interactions among benthic components within sites. These findings suggest that conservation initiatives might be successful at the site scale as well as the regional scale.

Among-site variability in the stochastic dynamics of East African coral reefs

Katherine A. Allen

School of Environmental Sciences, University of Liverpool, Liverpool, L69 3GP, UK *

John F. Bruno

Department of Biology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599-3300, USA †

Fiona Chong

School of Environmental Sciences, University of Liverpool, Liverpool, L69 3GP, UK ‡

Damian Clancy

School of Mathematical and Computer Sciences, Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, UK §

Tim R. McClanahan

Wildlife Conservation Society, 2300 Southern Boulevard, Bronx, NY 10460, USA ¶

Matthew Spencer

School of Environmental Sciences, University of Liverpool, Liverpool, L69 3GP, UK ‖

Kamila Żychaluk

Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 7ZL, UK **

*k.a.allen@liverpool.ac.uk. Current address: Institute of Integrative Biology, University of Liverpool, Liverpool L69 7ZB, UK

†jbruno@unc.edu

‡f.chong@student.liverpool.ac.uk

§d.clancy@hw.ac.uk

¶tmcclanahan@wcs.org

‖Corresponding author: m.spencer@liverpool.ac.uk

**kamila.zychaluk@liverpool.ac.uk

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

Abstract

Coral reefs are dynamic systems whose composition is highly influenced by unpredictable biotic and abiotic factors. Understanding the spatial scale at which long-term predictions of reef composition can be made will be crucial for guiding conservation efforts. Using a 22-year time series of benthic composition data from 20 reefs on the Kenyan and Tanzanian coast, we developed Bayesian vector autoregressive state-space models for reef dynamics, incorporating among-site variability, and quantified their long-term behaviour. We estimated that if there were no among-site variability, the total long-term variability would be approximately one third of its current value. Thus, our results showed that among-site variability contributes more to long-term variability in reef composition than does temporal variability. Individual sites were more predictable than previously thought, and predictions based on current snapshots are informative about long-term properties. Our approach allowed us to identify a subset of possible climate refugia sites with high conservation value, where the long-term probability of coral cover ≤ 0.1 was very low. Analytical results show that this probability is most strongly influenced by among-site variability and by interactions among benthic components within sites. These findings suggest that conservation initiatives might be successful at the site scale as well as the regional scale.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

Introduction

"Probabilistic language based on stochastic models of population growth" has been proposed as a standard way to evaluate conservation and management strategies (Ginzburg et al., 1982). For example, a stochastic population model can be used to estimate the probability of abundance falling below some critical level. Such population viability analyses are widely used, and may be reasonably accurate if sufficient data are available (Brook et al., 2000). In principle, the same approach could be used for communities, provided that a sufficiently simple model of community dynamics can be found.

A good candidate for such a model is the vector autoregressive model of order 1 or VAR(1) (Lütkepohl, 1993; Ives et al., 2003). This is a discrete-time model for the vector of log abundances of a set of species or groups, which includes environmental stochasticity and may include environmental explanatory variables. It makes the simplifying assumptions that inter- and intraspecific interactions can be represented by a linear approximation on the log scale, and that future abundances are conditionally independent of past abundances, given current abundances. Where possible, it is desirable to use a state-space form of the VAR(1) model, which also includes measurement error (Lindegren et al., 2009; Mutshinda et al., 2009).

Hampton et al. (2013) review applications of VAR(1) models in community ecology, which include studying the stability of freshwater plankton systems (Ives et al., 2003), designing adaptive management strategies for the Baltic Sea cod fishery (Lindegren et al., 2009), and estimating the contributions of environmental stochasticity and species interactions to temporal fluctuations in abundance of moths, fish, crustaceans, birds and rodents (Mutshinda et al., 2009). Recently, VAR(1) models have been applied to the dynamics of the benthic composition of coral reefs (Cooper et al., 2015; Gross and Edmunds, 2015), using a log-ratio transformation (Egozcue et al., 2003) rather than a log transformation, to deal with the constraint that proportional cover of space-filling benthic groups sums to 1.

Coral reefs are dynamic systems influenced by both deterministic factors such as interactions between macroalgae and hard corals (Mumby et al., 2007), and stochastic factors such as

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

temperature fluctuations (Baker et al., 2008) and storms (Connell et al., 1997), and are classic examples of non-equilibrium systems whose diversity is determined by both interspecific interactions and disturbance (Huston, 1985). In general, high coral cover is considered a desirable state for a coral reef, and there is some evidence that coral cover of at least 0.1 is important for long-term maintenance of reef function (Kennedy et al., 2013; Perry et al., 2013; Roff et al., 2015), although this threshold may differ among reefs dominated by different coral genera (Perry et al., 2015). Thus, coral cover of 0.1 might be an appropriate threshold against which to evaluate reef conservation strategies, and VAR(1) models can be used to estimate the probability of coral cover falling to or below this threshold (Cooper et al., 2015).

There is evidence for systematic differences in reef dynamics among locations. For example, on the Great Barrier Reef up to 2012, coral cover had declined more strongly at southern and central than at northern sites (De'ath et al., 2012), and in the U.S. Virgin Islands, VAR(1) models showed that sites differed in their sensitivity to disturbance and speed of recovery (Gross and Edmunds, 2015). Some sites in a region may therefore represent coral refugia, where reefs are either protected from or able to adapt to changes in environmental conditions (McClanahan et al., 2007). Alternatively, apparent differences among sites may simply be due to differences in recent disturbance history. Although it may be possible to associate differences in dynamics among sites with differences in environmental variables, it is also possible to treat among-site differences as another random component of a VAR(1) model. This will allow estimation of the relative importance of among-site variability and within-site temporal variability, which is important for the design of conservation strategies. If within-site temporal variability dominates, it will not be possible to identify good sites to conserve based on current status, while if among-site variability dominates, even a "snapshot" sample at one time point may be enough to identify good sites. Thus, for example, the reliability of among-site patterns from surveys at one time point, such as the relationship between benthic composition and human impacts on remote Pacific atolls (Sandin et al., 2008), depends on among-site variability dominating within-site temporal variability. Thus, even though a simple strategy based on a snapshot may turn out to be effective, it is not possible

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

to know this in advance of carrying out a more sophisticated analysis that treats the system as dynamic. Furthermore, since among-site variability will affect the probability of undesirable community composition (such as coral cover ≤ 0.1), conservation strategies that explicitly address among-site variability may be effective. As far as we know, the use of VAR(1) models to estimate spatiotemporal heterogeneity and identify refugia is novel, although other applications of VAR(1) models with random subject effects exist (e.g. Gorrostieta et al., 2012; Driver et al., 2016). Our approach differs from existing methods for identifying refugia (Keppel et al., 2012) in that it explicitly focuses on spatial variability in dynamics over ecological timescales, rather than on patterns that are static or vary only over much longer timescales. Furthermore, rather than differences in physical factors (West and Salm, 2003), we focus on differences in community dynamics.

Here, we develop a state-space VAR(1) model for regional dynamics of East African coral reefs, including random site effects and measurement error, and use it to answer four key questions about spatial and temporal variability. How important is among-site variability in the dynamics of benthic composition, relative to within-site temporal variability? How much variability is there among sites in the probability of low (≤ 0.1) coral cover? Which model parameters have the largest effects on the probability of low coral cover in the region? How informative is a single snapshot in time about the long-term properties of a site?

Methods

Data collection

Surveys of 20 spatially distinct reefs in Kenya and Tanzania (supporting information, Table A1, Figure A1) were conducted annually during the period 1991-2013 (generally in November or December prior to 1998, but January or February from 1998 onwards). Those in the north were typically fringing reefs, 100 m to 2000 m from the shore, while those in the south were typically smaller and more isolated patch reefs, further from the shore (McClanahan and Arthur, 2001). We

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

categorized reefs as either fished or unfished, although there was substantial heterogeneity within these categories, because some fished reefs were community management areas with reduced harvesting intensity (Cinner and McClanahan, 2015), and some unfished reefs had only recently been designated as reserves. Of the 20 reefs, 10 were divided into two sites separated by 20 m to 100 m, while the remaining 10 reefs comprised only one site. The selection of sites represents available data rather than a random sample from all the locations at which coral reefs are present in the geographical area (and all of the longest time series are from Kenyan fringing reefs). Thus, when we refer below to 'a randomly-chosen site' we strictly mean 'a site drawn at random from the population for which data could have been available.'

Each of the 30 sites was visited at least twice (data from sites visited once were omitted), with a maximum of 20 visits. A version of line-intercept sampling (Kaiser, 1983; McClanahan et al., 2001) was used to estimate reef composition. In total, 2665 linear transects were sampled across all sites and years, with between 5 and 18 transects (median 9) at each site in a single year. Transects were randomly placed between two points 10 m apart, but as the transect line was draped over the contours of the substrate, the measured lengths varied between 10 m and 15 m. Cover of benthic taxa was recorded as the sum of draped lengths of intersections of patches of each taxon with the line, divided by the total draped length of the line. Intersections with length less than 3 cm were not recorded. Taxa were identified to species or genus level, but for this study cover was grouped into three broad categories: hard coral, macroalgae and other (algal turf, calcareous and coralline algae, soft corals and sponges). Sand and seagrass were recorded, but excluded from our analysis, which focussed on hard substrate. The dynamics of a subset of these data were analyzed using different methods in Żychaluk et al. (2012).

Data processing

The three cover values form a three-part composition, a set of three positive numbers whose sum is 1 (Aitchison, 1986, Definition 2.1, p. 26). Standard multivariate statistical techniques are not appropriate for untransformed compositional data, due to the absence of an interpretable

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

covariance structure and the difficulties with parametric modelling (Aitchison, 1986, chapter 3). To avoid these difficulties, the proportional cover data were transformed to orthogonal, unconstrained, isometric log-ratio (ilr) coordinates (Egozcue et al., 2003). It is of course true that the model presented below for transformed data has an analogous model for untransformed data (Mateu-Figueras et al., 2011). However, working with transformed data allows us to use familiar methods.

The transformed data at site i, transect j, time t were represented by the vector $\mathbf{y}_{i,j,t} = [y_{1,i,j,t}, y_{2,i,j,t}]^T$, in which the first coordinate $y_{1,i,j,t}$ was proportional to the natural log of the ratio of algae to coral, and the second coordinate $y_{2,i,j,t}$ was proportional to the natural log of the ratio of other to the geometric mean of algae and coral (supporting information, section A1). The T denotes transpose: throughout, we work with column vectors. Note that both raw and transformed data are dimensionless.

The model

The true value $\mathbf{x}_{i,t} = [x_{1,i,t}, x_{2,i,t}]^T$ of the isometric log-ratio transformation of cover of hard corals, macroalgae and other at site i at time t was modelled by a vector autoregressive process of order 1 (i.e. a process in which the cover in a given year depends only on cover in the previous year), an approach used in other recent models of coral reef dynamics (Cooper et al., 2015; Gross and Edmunds, 2015). Unlike previous models, we include a random term representing among-site variation, and explicit treatment of measurement error (making this a state-space model). The full model is

$$\begin{aligned} \mathbf{x}_{i,t+1} &= \mathbf{a} + \boldsymbol{\alpha}_i + \mathbf{B}\mathbf{x}_{i,t} + \boldsymbol{\varepsilon}_{i,t}, \\ \boldsymbol{\alpha}_i &\sim \mathcal{N}(\mathbf{0}, \mathbf{Z}), \\ \boldsymbol{\varepsilon}_{i,t} &\sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \\ \mathbf{y}_{i,j,t} &\sim t_2(\mathbf{x}_{i,t}, \mathbf{H}, \nu), \end{aligned} \tag{1}$$

where all variables and parameters are dimensionless.

The column vector $\mathbf{a}$ represents the among-site mean proportional changes in $\mathbf{x}_{i,t}$ evaluated at

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

$\mathbf{x}_{i,t} = \mathbf{0}$. The column vector $\boldsymbol{\alpha}_i$ represents the amount by which these proportional changes for the ith site differ from the among-site mean, and is assumed to be drawn from a multivariate normal distribution with mean vector $\mathbf{0}$ and 2×2 covariance matrix $\mathbf{Z}$. The 2×2 matrix $\mathbf{B}$ represents the effects of $\mathbf{x}_{i,t}$ on the proportional changes, and can be thought of as summarizing intra- and inter-component interactions such as competition. The column vector $\boldsymbol{\varepsilon}_{i,t}$ represents random temporal variation, and is assumed to be drawn from a multivariate normal distribution with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$. We assume that there is no temporal or spatial autocorrelation in $\boldsymbol{\varepsilon}$, and that $\boldsymbol{\varepsilon}$ is independent of the among-site variation $\boldsymbol{\alpha}$.

The observed transformed compositions $\mathbf{y}_{i,j,t}$ vary around the corresponding true compositions $\mathbf{x}_{i,t}$ due to both small-scale spatial variation in true composition among transects within a site, and measurement error in estimating composition from a transect. We cannot easily separate these sources of variation because transects were located at different positions in each year, and there were no repeat measurements within transects. Observed log-ratio transformed cover $\mathbf{y}_{i,j,t}$ in the jth transect of site i at time t was assumed to be drawn from a bivariate t distribution (denoted by t_2) with location vector equal to the corresponding $\mathbf{x}_{i,t}$, and unknown scale matrix $\mathbf{H}$ and degrees of freedom ν, so that $\mathbf{y}_{i,j,t}$ has mean vector $\mathbf{x}_{i,t}$ if $\nu > 1$, and covariance matrix $\nu\mathbf{H}/(\nu - 2)$ if $\nu > 2$ (Lange et al., 1989). The bivariate t distribution can be interpreted as a mixture of bivariate normal distributions whose covariance matrices are the same up to a scalar multiple (Lange et al., 1989), and therefore allows a simple form of among-site or temporal variation in the distribution of measurement error or small-scale spatial variation, whose importance increases as the degrees of freedom decrease. Preliminary analyses suggested that it was important to allow this variation, because the model in Equation 1 fitted the data much better than a model with a bivariate normal distribution for $\mathbf{y}_{i,j,t}$ (supporting information, section A3).

We make the important simplifying assumptions that $\mathbf{B}$ is the same for all sites, and that the causes of among-site and temporal variation are not of interest. A separate $\mathbf{B}$ for each site, or even a hierarchical model for $\mathbf{B}$, would be difficult to estimate from the amount of data we have. It might be possible to explain some of the random temporal variation using temporally-varying

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

environmental covariates such as sea surface temperature, and some of the among-site variation using temporally constant covariates such as management strategies (Cooper et al., 2015). However, it is not necessary to do so in order to answer the questions listed at the end of the introduction, and keeping the model as simple as possible is important because parameter estimation is quite difficult. Furthermore, some of the relevant environmental variables may be associated with management strategies, making it difficult to separate the effects of environmental variation and management. For example, although some water quality variables were not strongly associated with protection status (Carreiro-Silva and McClanahan, 2012), unfished reefs were designated as protected areas due to their relatively good condition and are generally found in deeper lagoons with lower and more stable water temperatures than fished reefs (T. R. McClanahan, personal observation).

To understand the features of dynamics common to all sites, we plotted the back-transformations from ilr coordinates to the simplex of the overall intercept parameter $\mathbf{a}$ and the columns $\mathbf{a}_1$ and $\mathbf{a}_2$ of a matrix $\mathbf{A}$, which is related to $\mathbf{B}$ and describes the effects of current reef composition on the change in reef composition from year to year (Cooper et al., 2015). We plotted $\mathbf{A}$ rather than $\mathbf{B}$ because it leads to a simpler visualization of effects (supporting information, section A4). For example, a point lying to the left of the line representing equal proportions of coral and algae (the 1:1 coral-algae isoproportion line) corresponds to a parameter tending to increase coral relative to algae.

Parameter estimation

We estimated all model parameters and checked model performance using Bayesian methods implemented in the Stan programming language (Stan Development Team, 2015), as described in the supporting information (section A5). Stan uses the No-U-Turn Sampler, a version of Hamiltonian Monte Carlo, which can converge much faster than random-walk Metropolis sampling when parameters are correlated (Hoffman and Gelman, 2014). For most results, we report posterior means and 95% highest posterior density (HPD) intervals (Hyndman, 1996),

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

calculated in R (R Core Team, 2015). We showed using simulations that our estimated credible intervals had close to the specified coverage (section A5.2 and Figure A3).

Long-term behaviour

In the long term (as $t \to \infty$), the true transformed composition $\mathbf{x}^*$ of a randomly-chosen site will converge to a stationary distribution, provided that all the eigenvalues of $\mathbf{B}$ lie inside the unit circle in the complex plane (e.g. Lütkepohl, 1993, p. 10). If the eigenvalues of $\mathbf{B}$ are complex, the system will oscillate as it approaches the stationary distribution. Details of long-term behaviour are in the supporting information, section A6.

This stationary distribution is the multivariate normal vector

$$\mathbf{x}^* \sim \mathcal{N}(\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^* + \mathbf{Z}^*), \tag{2}$$

whose stationary mean $\boldsymbol{\mu}^*$ depends on $\mathbf{B}$ and $\mathbf{a}$, and whose stationary covariance is the sum of the stationary within-site covariance $\boldsymbol{\Sigma}^*$ (which depends on $\mathbf{B}$ and $\boldsymbol{\Sigma}$) and the stationary among-site covariance $\mathbf{Z}^*$ (which depends on $\mathbf{B}$ and $\mathbf{Z}$).

For a fixed site i, the value of $\boldsymbol{\alpha}_i$ is fixed and the stationary distribution is given by

$$\mathbf{x}_i^* \sim \mathcal{N}(\boldsymbol{\mu}_i^*, \boldsymbol{\Sigma}^*), \tag{3}$$

whose stationary mean $\boldsymbol{\mu}_i^*$ depends on $\mathbf{B}$, $\mathbf{a}$ and $\boldsymbol{\alpha}_i$, and whose stationary covariance matrix is $\boldsymbol{\Sigma}^*$. Note that $\mathbf{B}$, which describes intra- and inter-component interactions on an annual time scale, affects all the parameters of both stationary distributions, and therefore affects both within- and among-site variability in the long term. Also, the back-transformation of the stationary mean $\boldsymbol{\mu}^*$ of the transformed composition, rather than the arithmetic mean vector of the untransformed composition, is the appropriate measure of the centre of the stationary distribution (Aitchison, 1989).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

How important is among-site variability?

The covariance matrix of the stationary distribution for a randomly-chosen site (Equation 2) contains contributions from both among- and within-site variability. To quantify the contributions from these two sources, we calculated

$$\rho = \left(\frac{|\mathbf{\Sigma}^*|}{|\mathbf{\Sigma}^* + \mathbf{Z}^*|} \right)^{1/2}, \tag{4}$$

(supporting information, section A7), which is the ratio of volumes of two unit ellipsoids of concentration (Kenward, 1979), the numerator corresponding to the stationary distribution in the absence of among-site variation (or for a fixed site, as in Equation 3), and the denominator to the full stationary distribution of transformed reef composition in the region. The volume of each ellipsoid of concentration is a measure of the dispersion of the corresponding distribution. Thus ρ provides an indication of how much of the total variability would remain if all among-site variability was removed. A similar statistic was used by Ives et al. (2003) to measure the contribution of species interactions to stationary variability.

How much variability is there among sites in the probability of low coral cover?

For a given coral cover threshold κ, we define $q_{\kappa,i}$ as the long-term probability that site i has coral cover less than or equal to κ. This can be interpreted either as the proportion of time for which the site will have coral cover less than or equal to κ in the long term, or as the probability that the site will have coral cover less than or equal to κ at a random time, in the long term. We set $\kappa = 0.1$, which has been suggested as a threshold for a positive net carbonate budget, based on simulation models and data from Caribbean reefs (Kennedy et al., 2013; Perry et al., 2013; Roff et al., 2015). We calculated $q_{0.1,i}$ for each site numerically (supporting information, section A8). In order to determine whether differences in $q_{0.1,i}$ were related to current coral cover, we plotted $q_{0.1,i}$ against the corresponding sample mean coral cover for each site, over all transects and years. In order to

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

determine whether differences in $q_{0.1,i}$ had obvious explanations, we distinguished between fished and unfished reefs, and patch and fringing reefs. In order to determine whether there was strong spatial pattern in the probability of low coral cover, we calculated spline correlograms (Bjørnstad and Falck, 2001) for a sample from the posterior distribution of $q_{0.1,i}$ (supporting information, section A9).

Which model parameters have the largest effects on the probability of low coral cover?

For a given coral cover threshold κ, we define q_κ as the long-term probability that a randomly-chosen site has coral cover less than or equal to κ. This is equal to the expected long-term probability that coral cover is less than or equal to κ over the region, and can be calculated numerically (supporting information, section A8). To find the parameters with the largest effects on q_κ, we calculated its derivatives with respect to each model parameter. As above, we concentrated on $\kappa = 0.1$. However, we also compared results from $\kappa = 0.05$ and $\kappa = 0.20$. The probability q_κ is a function of 12 parameters: all four elements of $\mathbf{B}$; both elements of $\mathbf{a}$; elements σ_{11}, σ_{21} and σ_{22} of Σ; and elements ζ_{11}, ζ_{21} and ζ_{22} of $\mathbf{Z}$. The negative of the gradient vector of derivatives of q_κ with respect to these parameters describes the direction of movement through parameter space in which the probability of low coral cover will be reduced most rapidly, and the elements of this vector with the largest magnitudes correspond to the parameters to which q_κ is most sensitive. To understand why q_κ responds to each model parameter, note that q_κ depends on the parameters $\boldsymbol{\mu}^*$, Σ^* and $\mathbf{Z}^*$ of the stationary distribution (Equation 2), which are in turn affected by the model parameters. We therefore used the chain rule for matrix derivatives (Magnus and Neudecker, 2007, p.108) to break down the derivatives into effects of $\boldsymbol{\mu}^*$, Σ^* and $\mathbf{Z}^*$ on q_κ, and effects of model parameters on $\boldsymbol{\mu}^*$, Σ^* and $\mathbf{Z}^*$ (supporting information, section A10). We also calculated elasticities of q_κ with respect to each parameter, which measure the rate of relative change in q_κ with respect to relative change in the parameter (supporting information, section A11).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

How informative is a snapshot about long-term site properties?

In a stochastic system, how much can a "snapshot" survey at a single point in time tell us about the long-term behaviour of the system? For example, are differences among sites that appear to be in good and bad condition likely to be maintained in the long term? To make this question more precise, suppose that we draw a site at random from the region, and at one point in time, draw the true state of the site at random from the stationary distribution for the site. This scenario matches Diamond's definition of "natural snapshot experiments" as "comparisons of communities assumed to have reached a quasi-steady state" (Diamond, 1986). For simplicity, we assume that we can estimate the true state accurately (for example, by taking a large number of transects). To quantify how informative this is about the long term properties of the site, we computed the correlation coefficients between corresponding components of the true state at a given site at a given time and of stationary mean for that site (supporting information, section A12). If these correlations are high, then a snapshot will be informative about long term properties.

Results

Overall dynamics

At all sites, the back-transformed posterior mean true states from the model (e.g. Figure 1, grey lines) closely tracked the centres of the distributions of cover estimates from individual transects, although there was substantial among-transect variability at a given site in a given year (e.g. Figure 1, circles). Figure 1 shows two examples, and time series for all sites are plotted in the supporting information, Figures A6 to A35. There were also substantial differences in patterns of temporal change among sites. For example, Kanamai1 (Figure 1a-c), a fished site, had consistently low algal cover and no dramatic changes in cover of any component. In contrast, Mombasa1 (Figure 1d-f), an unfished site, had a sudden decrease in coral cover in 1998, and algal cover was high from 2007 onwards. As a result, Mombasa1 was unusual in that the current

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

estimate of true algal cover was well above the stationary mean estimate (Figure 1e: black circle at end of time series). For most other sites, current estimated true cover was close to the stationary mean (supporting information, Figures A6 to A35, black circles at ends of time series). The uncertainty in true states (Figure 1, grey polygons represent 95% highest posterior density (HPD) credible intervals) was higher during intervals with missing observations (e.g. 2008 in Figure 1). In general, uncertainty in true states (grey polygons) and stationary means (black bars at end of time series) was highest for sites with few observations (e.g. Bongoyo1, Figure A6).

The overall intercept parameter **a** (Figure 2, green), which describes the dynamics of reef composition at the origin (where each component is equally abundant) was consistent with the observed low macroalgal cover in the region (e.g. Figure 1b, e). The back-transformation of **a** lay close to the coral-other edge of the ternary plot, and slightly above the 1:1 coral-other isoproportion line. It therefore represented a strong year-to-year decrease in algae, and a slight increase in other relative to coral, at the origin.

Current reef composition acts on year-to-year change in composition (through matrix **A**) so as to maintain fairly stable reef composition. The first column $\mathbf{a}_1$ of **A**, which represents the effects of the transformed ratio of algae to coral on year-to-year change in composition, lay (when back-transformed) to the left of the 1:1 coral-algae isoproportion line, above the 1:1 other-algae isoproportion line, and below the 1:1 coral-other isoproportion line (Figure 2, orange). Thus, increases in algae relative to coral resulted in decreases in algae relative to coral and other, and increases in coral relative to other, in the following year. The second column $\mathbf{a}_2$ of **A**, which represents the effects of the transformed ratio of other to algae and coral on year-to-year change in composition, lay (when back-transformed) on the 1:1 coral-algae isoproportion line, below the 1:1 other-algae isoproportion line, and below the 1:1 coral-other isoproportion line (Figure 2, blue). Thus, increases in other relative to algae and coral resulted in little change in the ratio of coral to algae, but decreases in other relative to both coral and algae. Consistent with the above interpretation of year-to-year dynamics, every set of parameters in the Monte Carlo sample led to a stationary distribution, since both eigenvalues of **B** lay inside the unit circle in the complex

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

plane (supporting information, section A13). The magnitudes of these eigenvalues were smaller than those for a similar model for the Great Barrier Reef (Cooper et al., 2015), indicating more rapid approach to the stationary distribution. There was some evidence for complex eigenvalues of **B**, leading to rapidly-decaying oscillations in both components of transformed reef composition on approach to this distribution. This contrasts with the Great Barrier Reef, where there was no evidence for oscillations (Cooper et al., 2015).

How important is among-site variability?

There was substantial among-site variability in the locations of stationary means (Figure 3, dispersion of points). Stationary mean algal cover was always low, but there was a wide range of stationary mean coral cover. Although our primary focus is not on the causes of among-site variability, there was a tendency for most of the reefs with highest stationary mean coral cover to be patch reefs (Figure 3, circles). In the light of these observations, we experimented with a model in which reef type was included as an explanatory variable. Although the estimated effects of reef type were consistent with lower long-term probabilities of coral cover ≤ 0.1, including reef type did not improve the expected predictive accuracy of the model (F. Chong, unpublished results), probably because only 482 out of 2665 transects were from patch reefs, and all but one patch reefs had only very short time series (supporting information, Table A1). The stationary means did not clearly separate by management (Figure 3, open symbols fished, filled symbols unfished). The long-term temporal variability around the stationary means was also substantial (Figure 3, green lines). The ρ statistic (Equation 4), which quantifies the posterior mean contribution of within-site variability to the total stationary variability in reef composition in the region, was 0.29 (95% HPD interval $(0.20, 0.39)$), or approximately one third. Thus, while within-site temporal variability around the stationary mean was not negligible, among-site variability in the stationary mean was more important in the long term. As noted above, uncertainty in the location of the stationary means (Figure 3, grey dashed lines) was much higher for reefs with few observations than for reefs with many observations. Nevertheless, most parameters of the model are not

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

reef-specific, and data from reefs with few observations contribute to the estimation of these. For all three components of variability (within-site, among-site, and measurement error/small-scale spatial variability), variation in algal cover was larger than variation in coral or other. This can be seen in the shapes of the back-transformed unit ellipsoids of concentration (Figure 4: within-site, green; among-site, orange; measurement error and small-scale spatial variability, blue) which were all elongated to some extent along the 1:1 coral-other isoproportion line. This was similar to, but less extreme than, the pattern observed in the Great Barrier Reef (Cooper et al., 2015). The among-site ellipsoid almost entirely enclosed the within-site ellipsoid, consistent with the estimate above that within-site variability contributed only around one-third of the total stationary variability in reef composition. The large estimated measurement error/small-scale spatial variability component was consistent with the substantial observed variability in cover among transects at any given site and time (Figure 1, circles and supporting information, Figures A6 to A35, circles). The low estimated degrees of freedom ν for the bivariate t distribution of measurement error/small-scale spatial variability (posterior mean 2.99, 95% HPD interval $(2.64, 3.35)$) suggested that some aspect of the process leading to variation in measured composition among transects at a given site was varying substantially over space or time, although we cannot determine the mechanism.

How much variability is there among sites in the probability of low coral cover?

There was also substantial among-site variability in the probability of low coral cover. For a randomly-chosen site, the posterior mean probability of coral cover less than or equal to 0.1 ($q_{0.1}$) in the long term was 0.12 (95% credible interval $(0.04, 0.21)$). The corresponding site-specific probabilities $q_{0.1,i}$ varied from 8×10^{-5} to 0.52 but were low for most sites, with a strong negative relationship between probability of low coral cover and observed mean coral cover (Figure 5). There was no clear distinction between fished and unfished reefs (Figure 5, open symbols fished, filled symbols unfished). However, probability of low coral cover appeared to be systematically

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

lower on patch reefs, which were mainly in Tanzania (Figures 5 and A1, circles: median of posterior means 2×10^{-3}, first quartile 4×10^{-4}, third quartile 0.04) than on fringing reefs (Figures 5 and A1, triangles: median of posterior means 0.08, first quartile 0.04, third quartile 0.11). One site (Ras Iwatine) had a much higher probability of low coral cover than all others, and is relatively polluted compared to other sites in this study, due to high levels of nutrient effluent from a large hotel (T.R. McClanahan, personal observation).

There was little evidence for strong spatial autocorrelation in the probability of low coral cover, because the 95% envelope for the spline correlogram included zero for all distances other than 261 km to 322 km (supporting information, Figure A38). The general lack of strong spatial autocorrelation reflects the substantial variation in probability of coral cover less than or equal to 0.1 ($q_{0.1,i}$) among nearby sites, while the possibility of negative spatial autocorrelation at scales of around 300 km may reflect the generally low values of $q_{0.1,i}$ for Tanzanian patch reefs, separated from sites in the north of the study area with generally higher $q_{0.1,i}$ by approximately 300 km (Figure A1).

Which model parameters have the largest effects on the probability of low coral cover?

Both among-site variability and internal dynamics, particularly of other relative to algae and coral (component 2), were important in determining the probability $q_{0.1}$ of coral cover ≤ 0.1 in the region. Figure 6 shows the direction in parameter space along which the probability of low coral cover will reduce most rapidly (the estimated gradient vector of $q_{0.1}$ with respect to all the model parameters). The four parameters to which $q_{0.1}$ was most sensitive were (in descending order: Figure 6) ζ_{21} (among-site covariance between transformed components 1 and 2), b_{22} (effect of component 2 on next year's component 2), ζ_{22} (among-site variance of component 2), and b_{12} (effect of component 2 on next year's component 1). Although there was substantial variability among Monte Carlo iterations in the values of these derivatives, the rank order of magnitudes was fairly consistent (supporting information, Figure A39). All four most important parameters had

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

positive effects on $q_{0.1}$ (Figure 6), so reducing these parameters will reduce $q_{0.1}$. The effects of within-site temporal variability on the probability of low coral cover were relatively unimportant (Figure 6, derivatives of $q_{0.1}$ with respect to σ_{11}, σ_{21} and σ_{22} all had posterior means close to zero). The signs of the effects of each parameter on $q_{0.1}$, sensitivities for coral cover thresholds 0.05 and 0.2, and elasticities, are discussed further in the supporting information (sections A14 and A15).

How informative is a snapshot about long-term site properties?

For both components of transformed composition, a snapshot of reef composition at a single time on a randomly-chosen site will be informative about the stationary mean (correlations between true value at a given time and stationary mean: component 1 posterior mean 0.84, 95% HPD interval $(0.75, 0.91)$; component 2 posterior mean 0.82, 95% HPD interval $(0.73, 0.90)$). This is consistent with the negative relationship between long-term probability of coral cover ≤ 0.1 and observed mean coral cover (Figure 5). Thus, while long-term monitoring of East African coral reefs is important for other reasons, it should be possible to identify those with high conservation value (in terms of benthic composition) from a single survey.

Discussion

In the long term (as $t \to \infty$), among-site variability dominated within-site temporal variability in East African coral reefs. In consequence, the long-term probability of coral cover ≤ 0.1 varied substantially among sites. This suggests that it is in principle possible to make reliable decisions about the conservation value of individual sites based on a survey of multiple sites at one point in time, and to design conservation strategies at the site level. This was not the only possible outcome: if within-site temporal variability dominated among-site variability, among-site differences would be neither important nor predictable in the long term. Given the large positive effect of among-site variability on the long-term probability of coral cover ≤ 0.1, reducing

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

among-site variability in compositional dynamics may be an effective conservation strategy. The dominance of among-site variability has important implications for conservation. There was clear evidence for the existence of a stationary distribution of long-term reef composition in East Africa. The overall shape of this distribution (Figure 3) was similar to that estimated by Żychaluk et al. (2012) for a subset of the same data, using a different modelling approach. However, our new analysis shows that this distribution is generated by a combination of spatial and temporal processes, with substantial long-term differences among sites. Thus, the distribution in Żychaluk et al. (2012) may be a good approximation to the long-term distribution for a randomly-chosen site, but there will be much less variability over time in the distribution for any fixed site. In consequence, the sites having the highest long-term conservation value can be identified even from single-survey snapshots, and conservation strategies at the site scale may be possible. Furthermore, in cases where among-site variability in dynamics is dominant, it will be misleading to generalize from observations of a few sites to regional patterns (Bruno et al., 2009).

In our study, the sites with the highest long-term conservation value are those with very low long-term probabilities of coral cover ≤ 0.1 (Figure 5), a threshold chosen based on evidence that coral cover ≤ 0.1 is detrimental to reef persistence (Kennedy et al., 2013; Perry et al., 2013; Roff et al., 2015). Many of these sites are Tanzanian patch reefs, which may have maintained high coral cover despite disturbance because of local hydrography (McClanahan et al., 2007), and are priority sites for conservation, with high alpha and beta diversity (Ateweberhan and McClanahan, 2016). Thus, it seems likely that sites of high conservation value based on community dynamics may also be sites of high diversity resulting from a combination of physical factors and biological interactions (Huston, 1985; West and Salm, 2003). However, the absence of strong spatial autocorrelation in long-term probabilities of coral cover ≤ 0.1 suggests that it will be necessary to consider conservation value at small spatial scales, rather than simply to identify subregions with high conservation value. Similarly, Vercelloni et al. (2014) found that trajectories of coral cover on the Great Barrier Reef were consistent at the scale of km^2, but not at larger spatial scales. They argued that it would therefore be appropriate to focus management actions at the km^2 scale. Also,

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

it may be easier to persuade local communities to accept management at such scales than at larger scales (McClanahan et al., 2016).

A key result is that if we want to minimize the long-term probability $q_{0.1}$ that a randomly-chosen site has coral cover ≤ 0.1, we should minimize among-reef variability in dynamics, other things being equal. This is because the centre of the stationary distribution lies outside the set of compositions with coral cover ≤ 0.1 (Supporting Information, Section A14). Conversely, if the centre lay inside this set, then (other things being equal) maximizing among-site variability would minimize $q_{0.1}$. This result is very general, applying to any model of community composition which has a stationary distribution, for which increasing among-site variability increases stationary variability, and for any conservation objective based on a composition threshold. Note that even though our conservation objective is based on coral cover alone, and is therefore one-dimensional, it was strongly affected (through parameters such as ζ_{12} and b_{12}) by processes involving all three components of reef composition. Thus, our multivariate approach provides information that could not be obtained from a one-dimensional model.

Conservation strategies that might minimize among-site variability include distributing a fixed amount of human activity such as coastal development or fishing evenly, rather than concentrating it in a few locations. On the other hand, many conservation strategies will affect both the mean dynamics and the among-site variability in dynamics. For example, protecting the sites that are already in the best condition will tend to increase among-site variability, while moving the centre of the stationary distribution away from the set of compositions with coral cover ≤ 0.1.

Minimizing among-site variability in dynamics may conflict with other proposed conservation strategies. For example, it has been suggested that increased beta diversity is associated with lower temporal variability in metacommunities, for at least some taxa, and that regions of high beta diversity may therefore be priority regions for conservation (Mellin et al., 2014). It is likely that increased beta diversity will also be associated with increased among-site variability in dynamics, because different species are likely to have different population-dynamic characteristics. Hence, it may not always be possible to manage for both low among-site

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

variability in dynamics and high beta diversity. It is not yet clear which of these objectives is more important in general. Furthermore, there are many other plausible conservation objectives, such as maximizing the expected number of healthy sites or ensuring that at least one healthy site is present at all times, and different objectives can lead to very different optimal strategies (Game et al., 2008).

Our analyses were based on the long-term consequences of current environmental conditions, and may therefore not be relevant if environmental conditions change. For example, if changes in climate or local human activity altered the vector **a** so as to transpose the centre of the stationary distribution into the set with coral cover ≤ 0.1, then maximizing among-site variability would become the best strategy. Since declining coral cover trends have been observed at the regional level (e.g. Côté et al., 2005; De'ath et al., 2012), such a shift in the best strategy may occur. It is therefore better to view a stationary distribution under current conditions as a "speedometer" that tells us about the long-term outcome if these conditions were maintained, rather than as a prediction (Caswell, 2001, p. 30). Furthermore, our model did not include connectivity between sites. Although the absence of evidence for strong spatial autocorrelation in the probability of low coral cover (Supporting Information, Section A9 and Figure A38) suggests that connectivity is relatively unimportant for our analysis, it is possible that either current patterns of connectivity or future changes in these patterns may affect both the interpretation of stationary distributions and the optimal management strategy.

In conclusion, our analysis extends the broadly-applicable vector autoregressive approach to community dynamics (reviewed by Hampton et al., 2013) by quantifying random among-site variability in dynamics. This gives a new perspective on the long-term behaviour of the set of communities in a region, as a set of stationary distributions with random but persistent differences. The extent of these differences relative to temporal variability determines how predictable the behaviour of individual sites will be. Since these differences may be associated with differences in conservation value, probabilistic risk assessment based on this approach can be used to suggest conservation strategies at both site and regional scales. At site scales, our

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

approach can be used to identify potential coral refugia, while at regional scales, it can identify the parameters with most influence on conservation objectives.

Literature cited

Aitchison, J. (1986). *The statistical analysis of compositional data*. Chapman and Hall, London.

Aitchison, J. (1989). Measures of location for compositional data sets. *Mathematical Geology*, 21:787–790.

Ateweberhan, M. and McClanahan, T. R. (2016). Partitioning scleractinian coral diversity across reef sites and regions in the Western Indian Ocean. *Ecosphere*, 7(5):01243.

Baker, A. C., Glynn, P. W., and Riegl, B. (2008). Climate change and coral reef bleaching: an ecological assessment of long-term impacts, recovery trends and future outlook. *Estuarine, Coastal and Shelf Science*, 80:435–471.

Bjørnstad, O. N. and Falck, W. (2001). Nonparametric spatial covariance functions: Estimation and testing. *Environmental and Ecological Statistics*, 8:53–70.

Brook, B. W., O'Grady, J. J., Chapman, A. P., Burgman, M. A., Akçakaya, H. R., and Frankham, R. (2000). Predictive accuracy of population viability analysis in conservation biology. *Nature*, 404:385–387.

Bruno, J. F., Sweatman, H., Precht, W. F., Selig, E. R., and Schutte, V. G. W. (2009). Assessing evidence of phase shifts from coral to macroalgal dominance on coral reefs. *Ecology*, 90(6):1478–1484.

Carreiro-Silva, M. and McClanahan, T. R. (2012). Macrobioerosion of dead branching *Porites*, 4 and 6 years after coral mass mortality. *Marine Ecology Progress Series*, 458:103–122.

Caswell, H. (2001). *Matrix population models: construction, analysis, and interpretation*. Sinauer, Sunderland, MA, second edition.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

Cinner, J. E. and McClanahan, T. R. (2015). A sea change on the African coast? Preliminary social and ecological outcomes of a governance transformation in Kenyan fisheries. *Global Environmental Change*, 30:133–139.

Connell, J. H., Hughes, T. P., and Wallace, C. C. (1997). A 30-year study of coral abundance, recruitment, and disturbance at several scales in space and time. *Ecological Monographs*, 67(4):461–488.

Cooper, J. K., Spencer, M., and Bruno, J. F. (2015). Stochastic dynamics of a warmer Great Barrier Reef. *Ecology*, 96:1802–1811.

Côté, I. M., Gill, J. A., Gardner, T. A., and Watkinson, A. R. (2005). Measuring coral reef decline through meta-analyses. *Philosophical Transactions of the Royal Society Series B*, 360:385–395.

De'ath, G., Fabricius, K. E., Sweatman, H., and Puotinen, M. (2012). The 27-year decline of coral cover on the Great Barrier Reef and its causes. *Proceedings of the National Academy of Sciences of the USA*, 109:17995–17999.

Diamond, J. (1986). Overview: laboratory experiments, field experiments, and natural experiments. In Diamond, J. and Case, T. J., editors, *Community ecology*, pages 3–22. Harper & Row, New York.

Driver, C. C., Oud, J. H. L., and Voelkle, M. C. (2016). Continuous time structural equation modelling with R package `ctsem`. *Journal of Statistical Software*, in press.

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., and Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology*, 35(3):279–300.

Game, E. T., McDonald-Madden, E., Puotinen, M. L., and Possingham, H. P. (2008). Should we protect the strong or the weak? Risk, resilience, and the selection of marine protected areas. *Conservation Biology*, 22(6):1619–1629.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

564 Ginzburg, L. R., Slobodkin, L. B., Johnson, K., and Bindman, A. G. (1982). Quasiextinction
565 probabilities as a measure of impact on population growth. *Risk Analysis*, 21:171–181.

566 Gorrostieta, C., Ombao, H., Bédard, P., and Sanes, J. N. (2012). Investigating brain connectivity
567 using mixed effects vector autoregressive models. *NeuroImage*, 59:3347–3355.

568 Gross, K. and Edmunds, P. J. (2015). Stability of Caribbean coral communities quantified by
569 long-term monitoring and autoregression models. *Ecology*, 96:1812–1822.

570 Hampton, S. E., Holmes, E. E., Scheef, L. P., Scheuerell, M. D., Katz, S. L., Pendleton, D. E., and
571 Ward, E. J. (2013). Quantifying effects of abiotic and biotic drivers on community dynamics
572 with multivariate autoregressive (MAR) models. *Ecology*, 94(12):2663–2669.

573 Hoffman, M. D. and Gelman, A. (2014). The No-U-Turn Sampler: Adaptively setting path
574 lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, 15:1351–1381.

575 Huston, M. (1985). Patterns of species diversity on coral reefs. *Annual Review of Ecology and*
576 *Systematics*, 16:149–177.

577 Hyndman, R. J. (1996). Computing and graphing highest density regions. *The American*
578 *Statistician*, 50(2):120–126.

579 Ives, A. R., Dennis, B., Cottingham, K. L., and Carpenter, S. R. (2003). Estimating community
580 stability and ecological interactions from time-series data. *Ecological Monographs*,
581 73(2):301–330.

582 Kaiser, L. (1983). Unbiased estimation in line-intercept sampling. *Biometrics*, 39(4):965–976.

583 Kennedy, E. V., Perry, C. T., Halloran, P. R., Iglesias-Prieto, R., Schönberg, C. H. L., Wissah, M.,
584 Form, A. U., Carricart-Ganivet, J. P., Fine, M., Eakin, C. M., and Mumby, P. J. (2013).
585 Avoiding coral reef functional collapse requires local and global action. *Current Biology*,
586 23:912–918.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

587 Kenward, M. G. (1979). An intuitive approach to the MANOVA test criteria. *Journal of the Royal*
588 *Statistical Society Series D*, 28(3):193–198.

589 Keppel, G., Van Niel, K. P., Wardell-Johnson, G. W., Yates, C. J., Byrne, M., Mucina, L., Schut,
590 A. G. T., Hopper, S. D., and Franklin, S. E. (2012). Refugia: identifying and understanding safe
591 havens for biodiversity under climate change. *Global Ecology and Biogeography*, 21:393–404.

592 Lange, K. L., Little, R. J. A., and Taylor, J. M. G. (1989). Robust statistical modeling using the *t*
593 distribution. *Journal of the American Statistical Association*, 84:881–896.

594 Lindegren, M., Möllmann, C., Nielsen, A., and Stenseth, N. C. (2009). Preventing the collapse of
595 the Baltic cod stock through an ecosystem-based management approach. *Proceedings of the*
596 *National Academy of Sciences of the USA*, 106(34):14722–14727.

597 Lütkepohl, H. (1993). *Introduction to multiple time series analysis*. Springer-Verlag, Berlin, 2nd
598 edition.

599 Magnus, J. R. and Neudecker, H. (2007). *Matrix differential calculus with applications in*
600 *statistics and econometrics*. John Wiley & Sons, Chichester, third edition.

601 Mateu-Figueras, G., Pawlowsky-Glahn, V., and Egozcue, J. J. (2011). The principle of working
602 on coordinates. In Pawlowsky-Glahn, V. and Buccianti, A., editors, *Compositional data*
603 *analysis: theory and applications*, pages 31–42. John Wiley & Sons, Chichester.

604 McClanahan, T. R. and Arthur, R. (2001). The effect of marine reserves and habitat on
605 populations of East African coral reef fishes. *Ecological Applications*, 11(2):559–569.

606 McClanahan, T. R., Ateweberhan, M., Muhando, C. A., Maina, J., and Mohammed, M. S. (2007).
607 Effects of climate and seawater temperature variation on coral bleaching and mortality.
608 *Ecological Monographs*, 77(4):503–525.

609 McClanahan, T. R., Muthiga, N. A., and Abunge, C. A. (2016). Establishment of community

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

managed fisheries' closures in Kenya: early evolution of the tengefu movement. *Coastal Management*, 44:1–20.

McClanahan, T. R., Muthiga, N. A., and Mangi, S. (2001). Coral and algal changes after the 1998 coral bleaching: interaction with reef management and herbivores on Kenyan reefs. *Coral Reefs*, 19:380–391.

Mellin, C., Bradshaw, C. J. A., Fordham, D. A., and Caley, M. J. (2014). Strong but opposing β-diversity-stability relationships in coral reef fish communities. *Proceedings of the Royal Society of London Series B*, 281:20131993.

Mumby, P. J., Hastings, A., and Edwards, H. J. (2007). Thresholds and the resilience of Caribbean coral reefs. *Nature*, 450:98–101.

Mutshinda, C. M., O'Hara, R. B., and Woiwod, I. P. (2009). What drives community dynamics? *Proceedings of the Royal Society of London Series B*, 276:2923–2929.

Perry, C. T., Murphy, G. N., Graham, N. A. J., Wilson, S. K., Januchowski-Hartley, F. A., and East, H. K. (2015). Remote coral reefs can sustain high growth potential and may match future sea-level trends. *Scientific Reports*, 5:18289.

Perry, C. T., Murphy, G. N., Kench, P. S., Smithers, S. G., Edinger, E. N., Steneck, R. S., and Mumby, P. J. (2013). Caribbean-wide decline in carbonate production threatens coral reef growth. *Nature Communications*, 4:1402.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Roff, G., Zhao, J.-X., and Mumby, P. J. (2015). Decadal-scale rates of reef erosion following El Niño-related mass coral mortality. *Global Change Biology*, 21:4415–4424.

Sandin, S. A., Smith, J. E., DeMartini, E. E., Dinsdale, E. A., Donner, S. D., Friedlander, A. M., Konotchick, T., Malay, M., Maragos, J. E., Obura, D., Pantos, O., Paulay, G., Richie, M.,

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

Rohwer, F., Schroeder, R. E., Walsh, S., Jackson, J. B. C., Knowlton, N., and Sala, E. (2008). Baselines and degradation of coral reefs in the Northern Line Islands. *PLoS ONE*, 3(2):e1548.

Stan Development Team (2015). Stan: A C++ library for probability and sampling, version 2.7.0.

Vercelloni, J., Caley, M. J., Kayal, M., Low-Choy, S., and Mengersen, K. (2014). Understanding uncertainties in non-linear population trajectories: a Bayesian semi-parametric hierarchical approach to large-scale surveys of coral cover. *PLoS ONE*, 9(11):e110968.

West, J. M. and Salm, R. V. (2003). Resistance and resilience to coral bleaching: implications for coral reef conservation and management. *Conservation Biology*, 17(4):956–967.

Żychaluk, K., Bruno, J. F., Clancy, D., McClanahan, T. R., and Spencer, M. (2012). Data-driven models for regional coral-reef dynamics. *Ecology Letters*, 15:151–158.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

Figure legends

Figure 1. Time series of cover of hard corals, macroalgae and other at two of the 30 sites surveyed: Kanamai1 (fished, a-c) and Mombasa1 (unfished, d-f). Circles are observations from individual transects. Grey lines join back-transformed posterior mean true states from Equation 1, and the shaded region is a 95% highest posterior density interval. The back-transformed stationary mean composition for the site is the black dot after the time series and the bar is a 95% highest posterior density interval.

Figure 2. Posterior distributions of the back-transformed overall intercept $\mathbf{a}$ (green), effect $\mathbf{a}_1$ of component 1 (proportional to log(algae/coral)) on year-to-year change (orange), and effect $\mathbf{a}_2$ of component 2 (proportional to log(other/geometric mean(algae,coral)) on year-to-year change (blue).

Figure 3. Stationary among- and within-site variation in benthic composition. Grey points: back-transformed stationary means for each site (open circles fished patch, filled circles unfished patch, open triangles fished fringing, filled triangles unfished fringing, posterior means of of stationary means). Grey dashed curves: back-transformed unit ellipsoids of concentration representing uncertainty in stationary means (calculated using sample covariance matrices from Monte Carlo iterations). Green solid curves: back-transformed unit ellipsoids of concentration representing within-site stationary variation (calculated using posterior mean within-site covariance matrix).

Figure 4. Back-transformed unit ellipsoids of concentration for stationary within-site covariance $\mathbf{\Sigma}^*$ (green), stationary among-site covariance $\mathbf{Z}^*$ (orange), and measurement error/small-scale spatial variation $\nu\mathbf{H}/(\nu-2)$ (blue). In each case, 200 ellipsoids drawn from the posterior distribution are plotted, centred on the origin.

Figure 5. Long-term probability of coral cover less than or equal to 0.1 at each site against mean observed coral cover across all years. Circles are patch reefs and triangles are fringing reefs. Open symbols are fished reefs and shaded symbols are unfished. Vertical lines are 95% highest posterior density intervals.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

Figure 6. Elements of the gradient vector of partial derivatives of the long-term probability of coral cover less than or equal to 0.1 with respect to elements of the **B** matrix (effects of transformed composition in a given year on transformed composition in the following year), the **a** vector (overall intercept, representing among-site mean proportional changes in transformed composition at the origin), the covariance matrix of random temporal variation $\boldsymbol{\Sigma}$, and the covariance matrix of among-site variability **Z**. For each parameter, the dot is the posterior mean and the bar is a 95% highest posterior density credible interval. For the covariance matrices, the elements σ_{12} and ζ_{12} are not shown, because they are constrained to be equal to σ_{21} and ζ_{21} respectively. The horizontal dashed line is at zero, the no-effect value.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

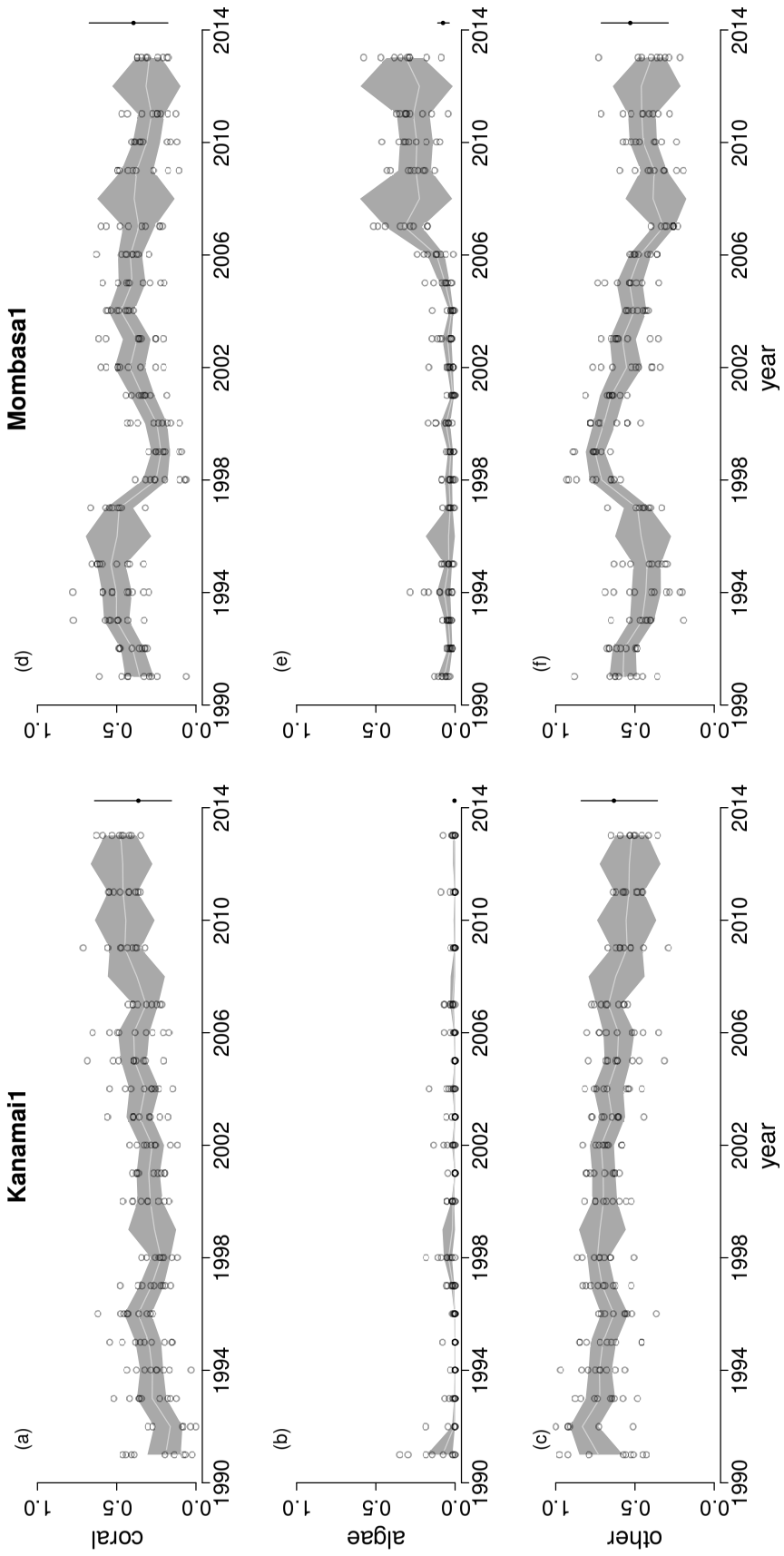

Figure 1:

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

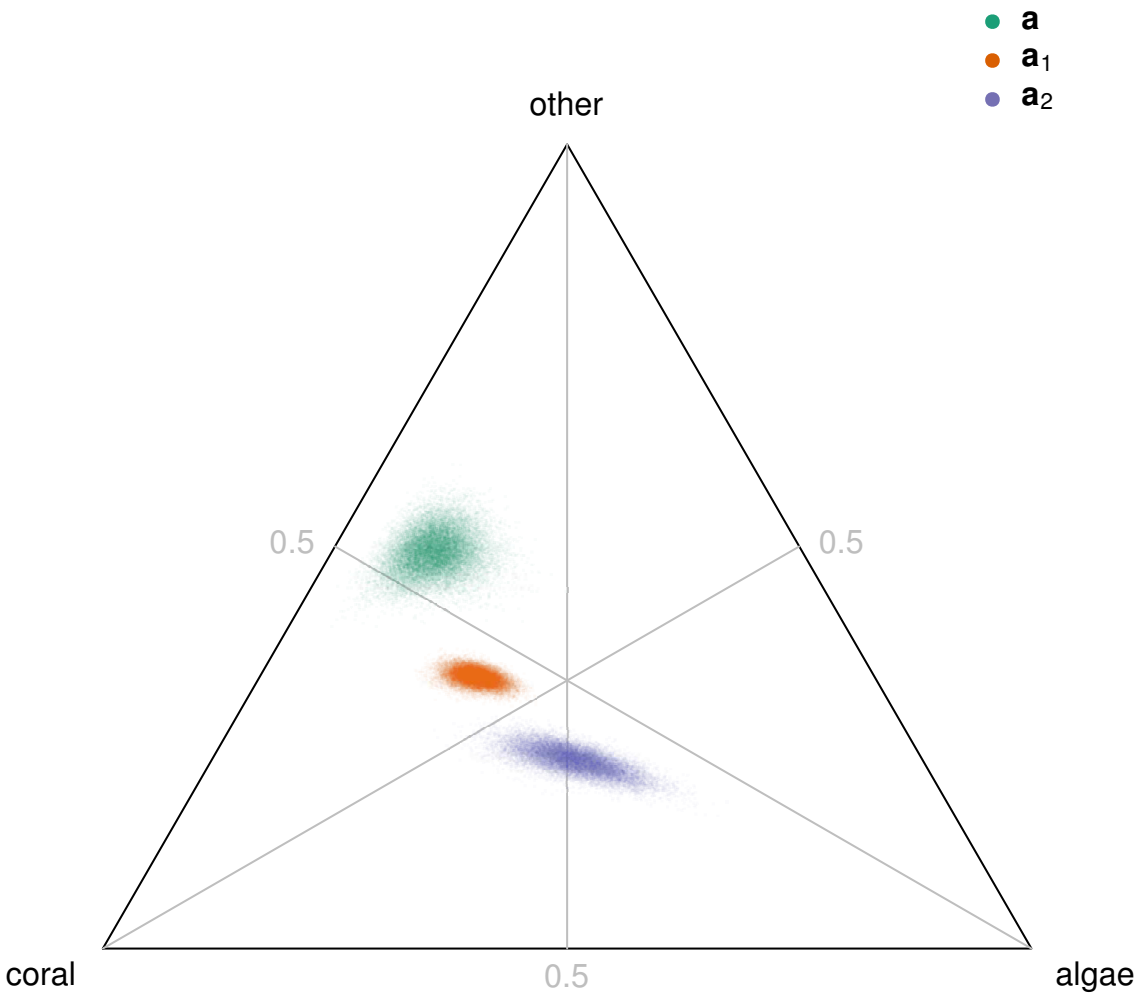

Figure 2:

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

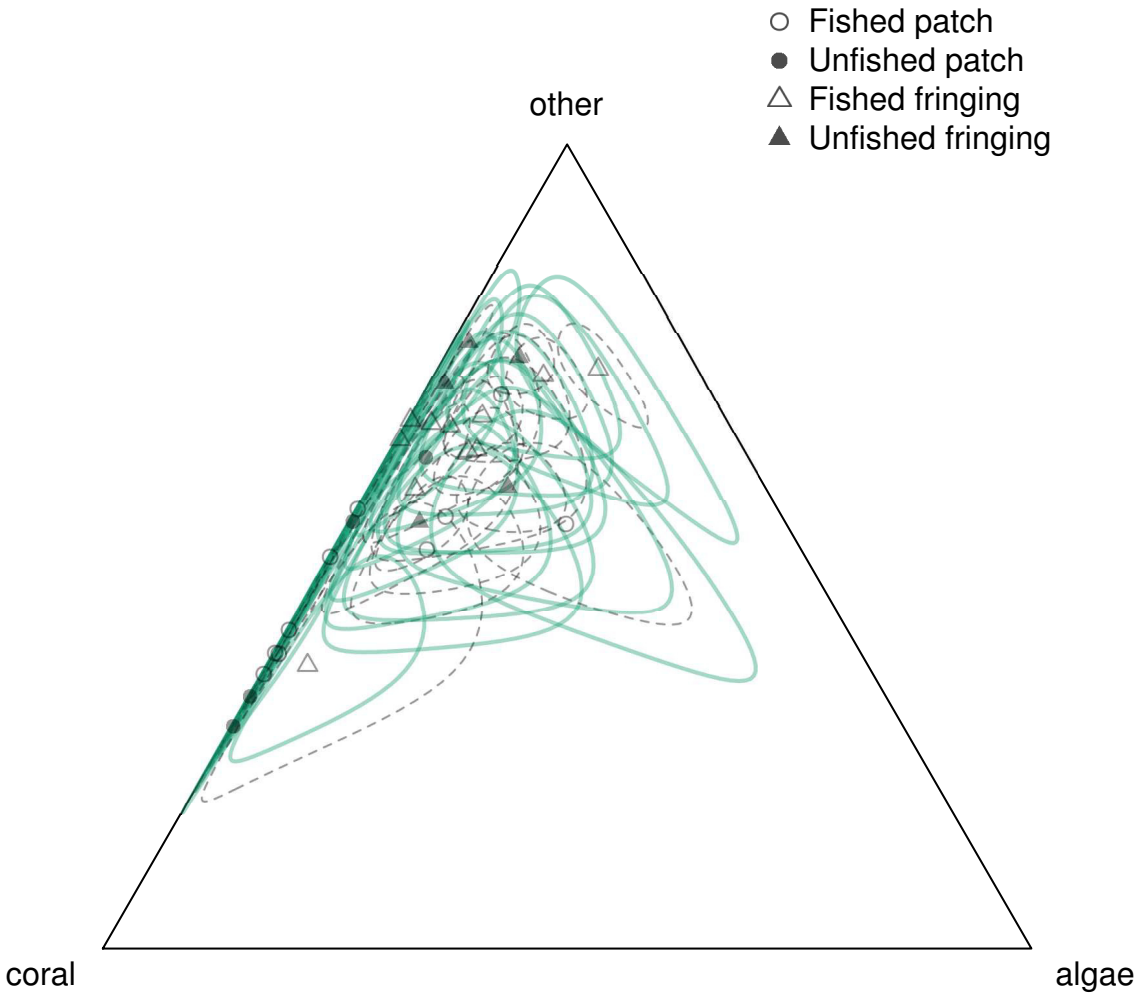

Figure 3:

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

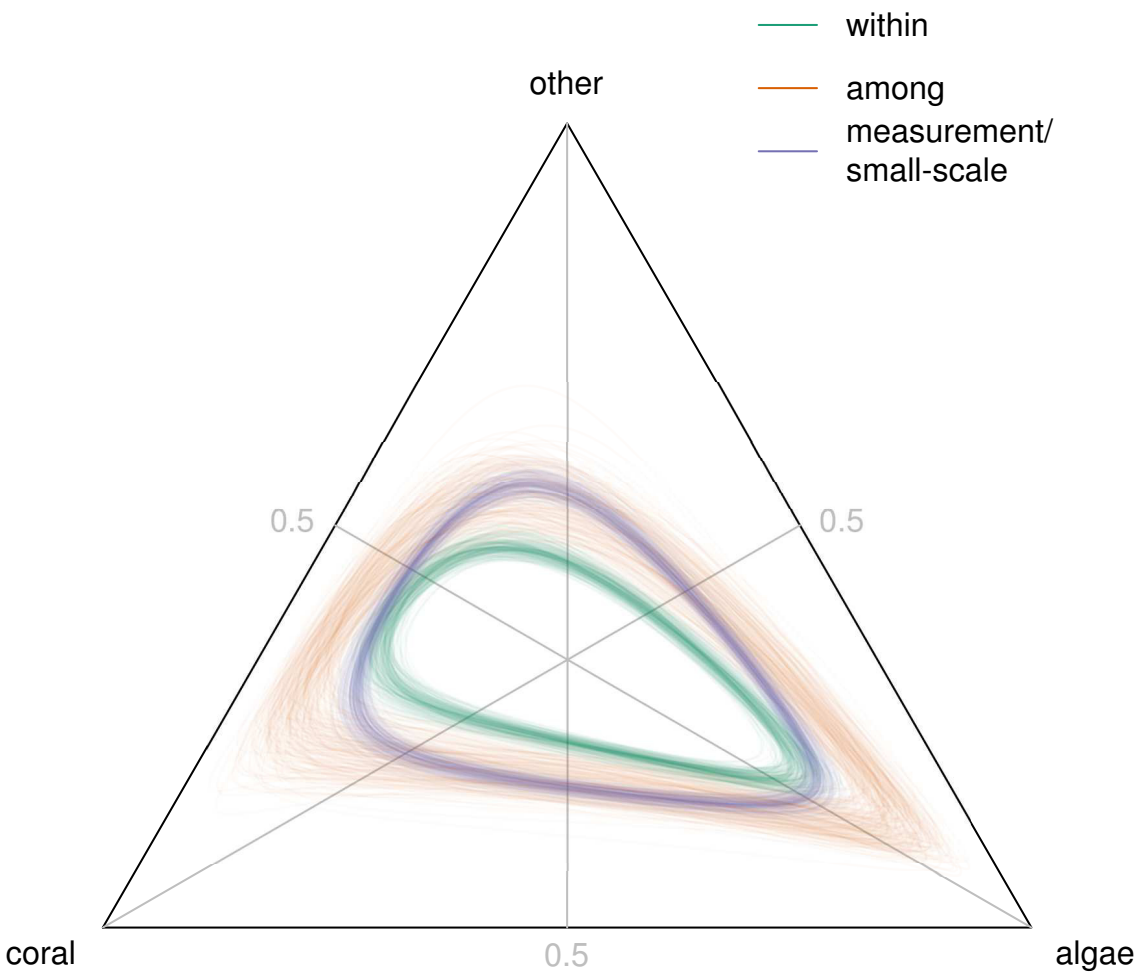

Figure 4:

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

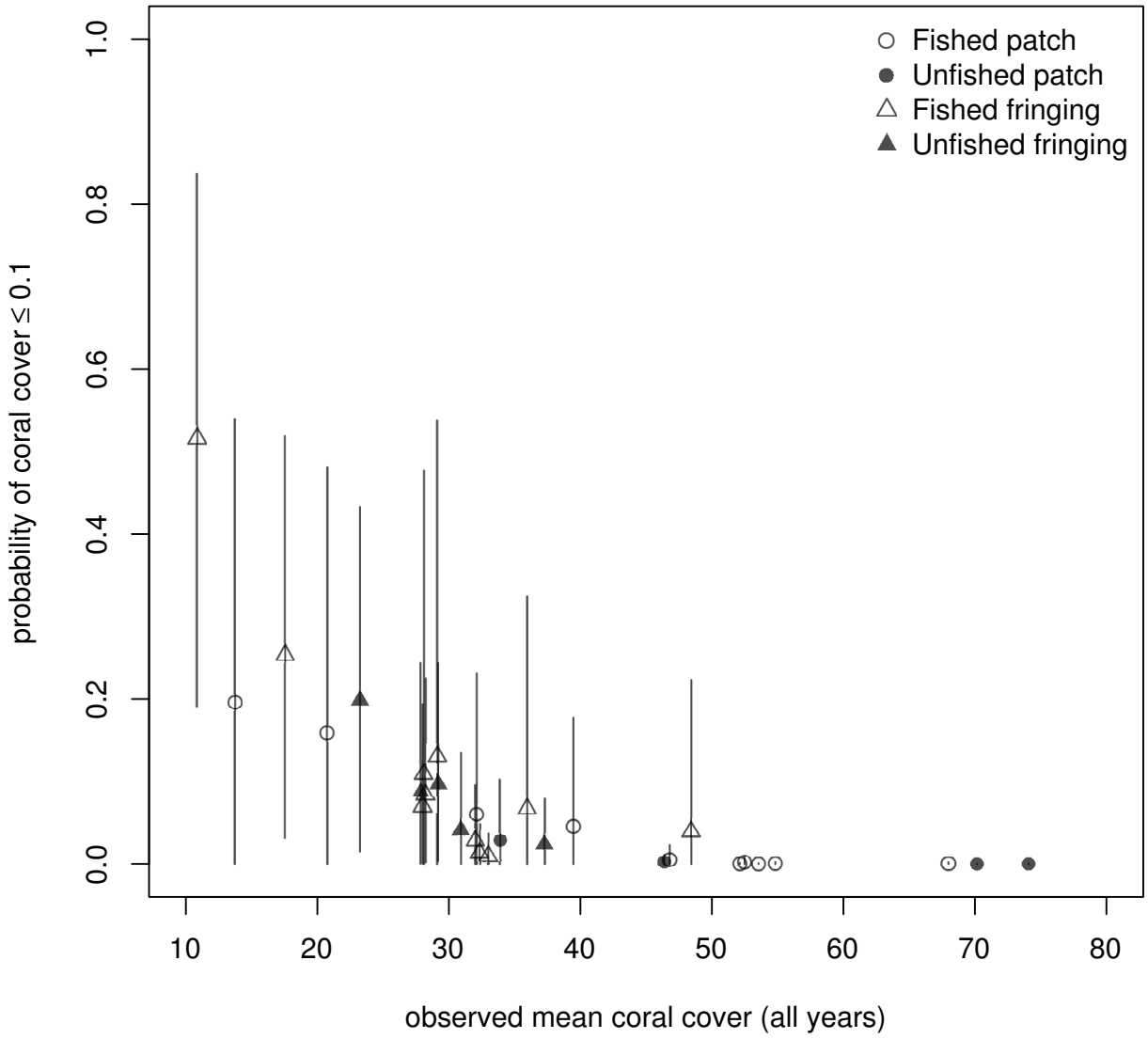

Figure 5:

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017

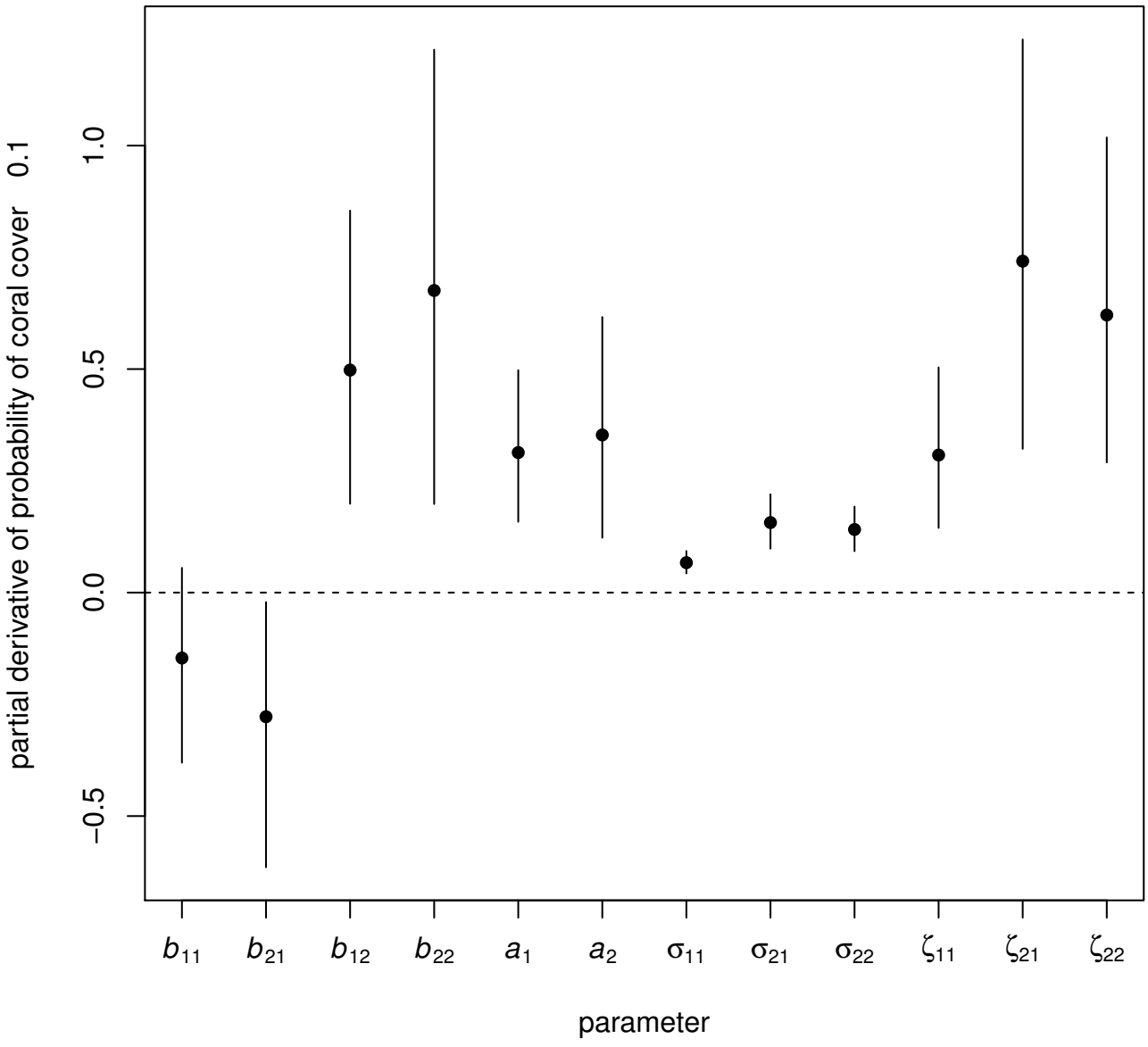

Figure 6:

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2687v1 | CC BY 4.0 Open Access | rec: 3 Jan 2017, publ: 3 Jan 2017