Appendix A: Detailed description of data, MAR modeling and data transformation

Data collection

Thirty photoquadrats (1 x 1 m) were recorded annually at Tektite and Yawzi Point, and 108-240 photoquadrats (0.5 x 0.5 m) were recorded at the RS (the RS sample size was increased in 2000 with the application of digital photography). Percentage cover of each group was determined using the software CPCe (Kohler & Gill 2006) with 200 randomly located dots on each image. A map of all study locations can be found in Edmunds (2013). Annual cover composition at each of the six sites that comprise the RS data are shown in Fig. A1.

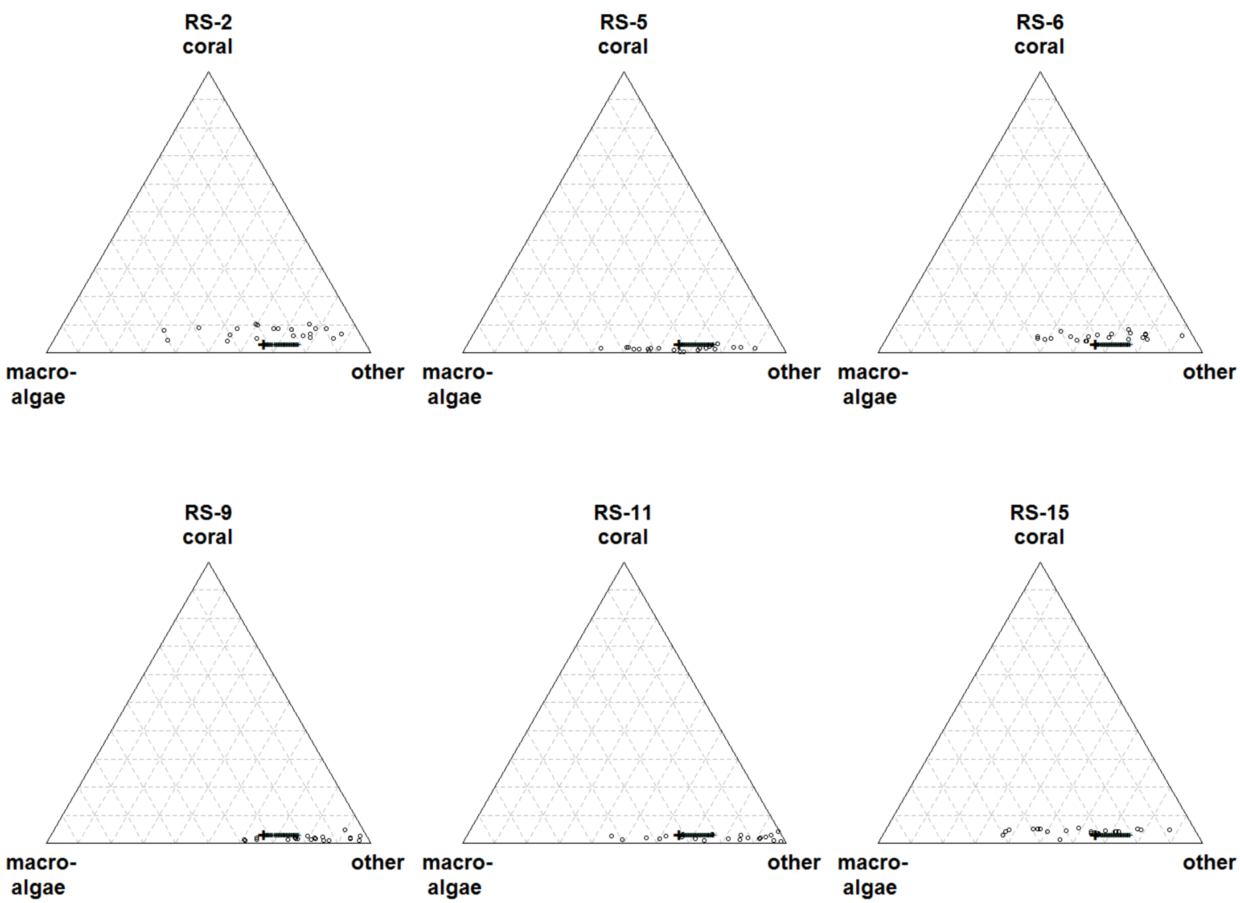

Figure A1. Composition of coral cover, macroalgal cover, and 'other' for the six sites that comprise the RS data. Site labels correspond to designations from Edmunds (2013). Large plus (+) symbols show the metric center of the quasi-stationary distribution for 2012, and small plus symbols trace how this mean has changed from 1992 – 2012. Open circles symbols show annual compositions.

Vectors and matrices in MAR model

For the cover analysis, the vectors and matrices in the MAR model (eq. 1) have the following forms:

$$\mathbf{x}_t = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_t ;\ \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix};\ \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix};\ \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix};\ \mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix};\ \mathbf{u}_t = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_t ;\ \mathbf{e}_t = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}_t . \qquad (1)$$

For the RS habitat, there is a unique intercept vector **a** for each of the six sites. For the taxonomic analysis, the vectors and matrices in the MAR model (eq. 1) have the following forms:

$$\mathbf{x}_t = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}_t ;\ \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \\ a_6 \end{bmatrix};\ \mathbf{B} = \begin{bmatrix} b_{11} & 0 & 0 & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & b_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & b_{66} \end{bmatrix};\ \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \\ c_{51} & c_{52} \\ c_{61} & c_{62} \end{bmatrix};$$
$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \\ z_6 \end{bmatrix};\ \mathbf{u}_t = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_t ;\ \mathbf{e}_t = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix}_t . \qquad (2)$$

For the RS habitat, there is a unique intercept vector **a** and a unique trend vector **z** for each of the six sites. The only parameters that are shared between the cover and taxonomic analyses are the mean vector $\boldsymbol{\mu}_{\mathrm{u}}$ and variance matrix $\boldsymbol{\Sigma}_{\mathrm{u}}$ for the environmental covariates. (This is because the values of the environmental covariates are the same for both cover and taxonomic analysis.) All other model parameters have separate values for the cover and taxonomic analysis.

Data transformation

In notation, if we write the proportional cover of coral, macroalgae, and 'other' as p_1, p_2, and p_3, respectively, then the corresponding isometric log-ratio (ilr) coordinates are

$$x_1 = \frac{1}{\sqrt{2}} \ln\left(\frac{p_1}{p_2}\right), \quad x_2 = \frac{2}{\sqrt{6}} \ln\left(\frac{\sqrt{p_1 p_2}}{p_3}\right) . \qquad (3).$$

With a change in sign, this is the same transformation used by Cooper et al. (*in press*). In short, ilr coordinates are orthogonal contrasts of the log proportions; results on the proportion scale do not depend on the particular set of contrasts chosen. Here, x_1 quantifies the difference between coral vs. macroalgae cover, and x_2 quantifies the difference between the geometric mean of coral and macroalgae cover vs. 'other'. This particular set of contrasts is based on a sequential binary partition (Egozcue and Pawlowsky-Glahn 2011; see their formula for 'balances' in their section 2.4). As Cooper et al. (*in press*) note, an ilr transformation is a sensible transformation for community compositions, because exponential growth of all components of the composition results in linear dynamics on the ilr scale (Egozcue et al. 2003). Time series of cover composition on the ilr-transformed scale are shown in fig. A2.

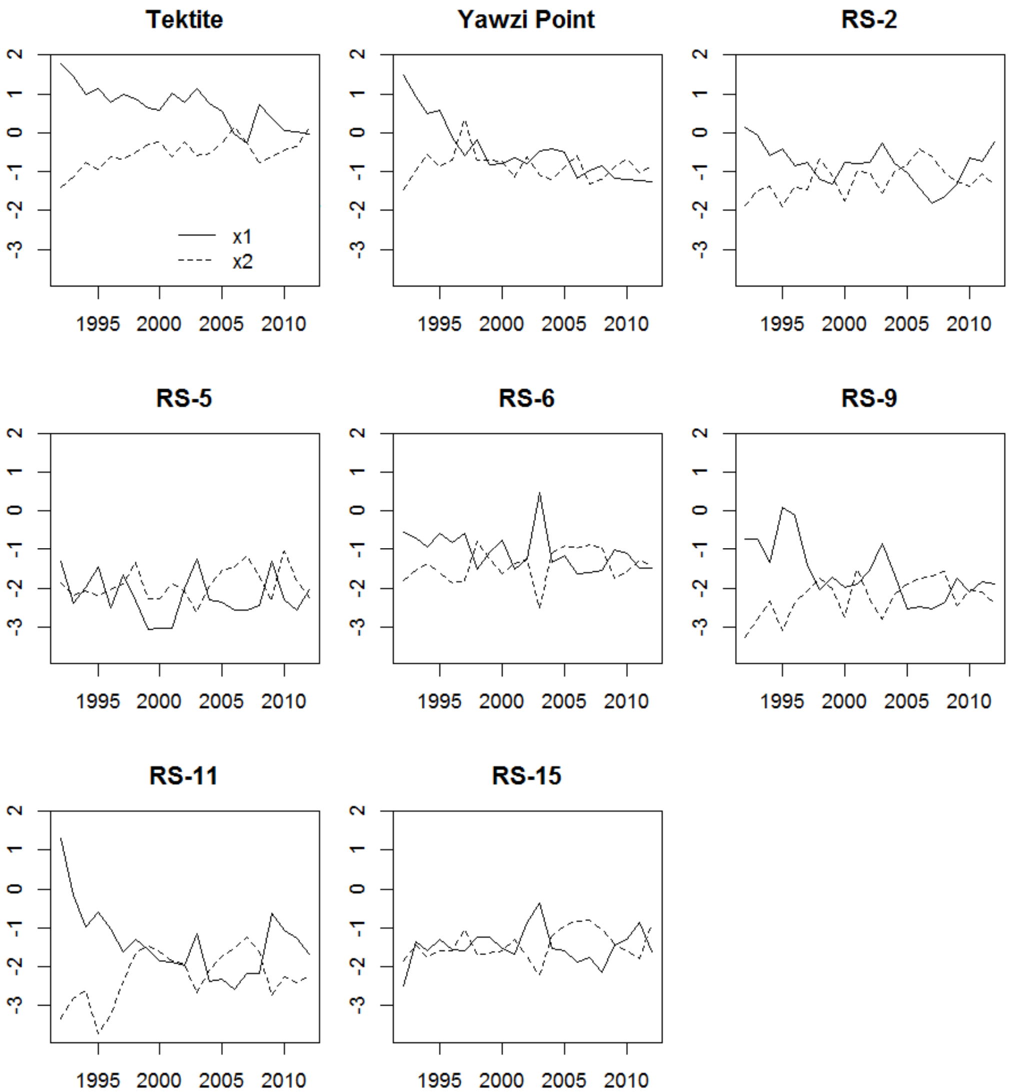

Figure A2. Cover composition in ilr coordinates for Tektite, Yawzi Point, and the 6 random sites.

To convert results back to the native proportion scale, write the cover proportions as the 3-vector **p**, write the ilr coordinates as the 2-vector **x**, and write the ilr transformation as *g*(), such that $\mathbf{x}=g(\mathbf{p})$ and $\mathbf{p}=g^{-1}(\mathbf{x})$. Applying the inverse transformation g^{-1} to $\boldsymbol{\mu}_\mathbf{x}$

$$\boldsymbol{\mu}_\mathbf{p}=g^{-1}(\boldsymbol{\mu}_\mathbf{x}) \tag{4}$$

yields the so-called "metric center" of the composition, which Aitchison (1989) and Pawlowsky-Glahn and Egozcue (2001) have argued provides the best measure of center for a composition. A linear approximations for the variance of the stationary distribution on the proportion scale (denoted $\mathbf{\Sigma}_{\mathbf{p}}$) is simply

$$\mathbf{\Sigma}_{\mathbf{p}} \approx \nabla g^{-1}\left(\boldsymbol{\mu}_{\mathbf{x}}\right)\mathbf{\Sigma}_{\mathbf{x}}\nabla g^{-1}\left(\boldsymbol{\mu}_{\mathbf{x}}\right)^{T} \ . \tag{5}$$

Expressions for eq. (4)-(5) on the proportion scale follow from the chain rule of calculus:

$$\frac{d\boldsymbol{\mu}_{\mathbf{p}}}{d\boldsymbol{\mu}_{\mathbf{u}}} \approx \nabla g^{-1}\left(\boldsymbol{\mu}_{\mathbf{x}}\right)\left(\mathbf{I}-\mathbf{B}\right)^{-1}\mathbf{C}; \quad \frac{d\boldsymbol{\mu}_{\mathbf{p}}}{dt^{*}} \approx \nabla g^{-1}\left(\boldsymbol{\mu}_{\mathbf{x}}\right)\left(\mathbf{I}-\mathbf{B}\right)^{-1}\mathbf{z}; \tag{6}$$

$$\frac{d\mathrm{Vec}\left(\mathbf{\Sigma}_{\mathbf{p}}\right)}{d\mathrm{Vec}\left(\mathbf{\Sigma}_{\mathbf{u}}\right)} \approx \left(\nabla g^{-1}\left(\boldsymbol{\mu}_{\mathbf{x}}\right)\otimes\nabla g^{-1}\left(\boldsymbol{\mu}_{\mathbf{x}}\right)\right)\left(\mathbf{I}-\mathbf{B}\otimes\mathbf{B}\right)^{-1}\left(\mathbf{C}\otimes\mathbf{C}\right) \ . \tag{7}$$

Literature cited in Appendix A

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

Cooper, J.K., M. Spencer, and J. F. Bruno. Stochastic dynamics of a warmer Great Barrier Reef. Ecology, *in press*.

Edmunds, P. J. 2013. Decadal-scale changes in the community structure of coral reefs of St. John, US Virgin Islands. Marine Ecology Progress Series 489:107–123.

Egozcue, J. J., and V. Pawlowsky-Glahn. 2011. Basic concepts and procedures. Pages 12–28 *in* V. Pawlowsky-Glahn and A. Buccianti, editors. Compositional data analysis: theory and applications. John Wiley and Sons New York, NY.

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

Kohler, K. E., and S. M. Gill. 2006. Coral Point Count with Excel extensions (CPCe): A Visual Basic program for the determination of coral and substrate coverage using random point count methodology. Computers & Geosciences 32:1259–1269.

Pawlowsky-Glahn, V., and J. J. Egozcue. 2001. Geometric approach to statistical analysis on the simplex. Stochastic Environmental Research and Risk Assessment 15:384–398.