Appendix S1 - Details of spatio-temporal modelling

Contents

Spatial modelling describes the spatial dependence in the response due to neighborhood relations as defined by the First Law of Geography [Tobler, 1970]. A few studies have demonstrated the importance of spatial structure on coral reef data from digital image surveys using a towed-diver [Aston et al., 2019, Ford et al., 2021] and small unmanned aerial systems [Levy et al., 2018]. These studies employed the Moran's I statistic [Moran, 1950] to detect the presence of spatial autocorrelation. However, this approach cannot be used to interpolate spatial patterns on locations without observations. In this case, additional analyses such as kriging are required [Cressie, 2015]. Spatial modelling allows estimation of spatial patterns while considering the effects of various drivers in species responses. It also enables interpolation over a continuous spatial field [Lindgren et al., 2011].

Justification for the Beta probability distribution

In this paper, we expressed estimates of coral cover as a percentage. In order to obtain these estimates, geo-referenced photoquadrats were processed using machine-learning algorithms [Gonzalez-Rivero et al., 2020]. A sample of 50 points within each image were classified, so that abundances of branching, plate and massive coral types are integers between 0 and 50. One option for modelling this data would be a binomial distribution with $n = 50$. However, the choice of 50 sample points is essentially arbitrary. Therefore, the variance of the binomial distribution would not accurately reflect our uncertainty about the true underlying coral cover. Instead, we transformed these counts into proportions between 0 and 1, which are modelled using a beta distribution with mean μ and precision parameter ϕ. Under this parameterization, the variance is given by:

$$Var(Y) = \frac{\mu(1-\mu)}{1+\phi},$$

whereas if we used the binomial distribution for counts Z, the variance would be:

$$Var(Z) = ny(1-y),$$

where $n = 50$ and y is a proportion between 0 and 1, as in the equation above. These two options are therefore essentially equivalent, except that the additional parameter ϕ gives us more flexibility to account for unexplained variation in the coral cover.

Another way to model this data is using Dirichlet distributions corresponding to a multivariate generalisation of the Beta distribution [Douma and Weedon, 2019]. Indeed, proportions of coral abundance are relative to each other. Composition data requires special attention for modelling that is beyond the scope of this paper [Chong and Spencer, 2018, Vercelloni et al., 2020]. At the time of writing, spatio-temporal models using Dirichlet distributions are not developed. Another consideration is that the proportions of the three coral types do not sum up to 1, as required by the Dirichlet. Rather, we would need to introduce additional categories for other benthic communities classified by the machine-learning algorithms.

Detection of spatial and temporal autocorrelation

The detection of spatial and temporal autocorrelation is based on model residuals from a beta regression model with years and habitats as fixed effects. Positive and significant spatial autocorrelation (expected values $<$ observed values and p-value < 0.05) is detected for the three coral types (Table S11). Temporal autocorrelation is significant for branching and plate corals only (Table S12). R packages DHARMa [Hartig, 2019] and glmmTMB [Brooks et al., 2017] were used to perform these tests.

Table S11: DHARMa Moran's I test for spatial autocorrelation.

Coral form	observed value	expected value	standard deviation	p-value
Branching	0.19	-0.003	0.003	$<$2.2e-16
Plate	0.16	-0.003	0.004	$<$2.2e-16
Massive	0.14	-0.004	0.005	$<$2.2e-16

Table S12: Durbin-Watson test for temporal autocorrelation.

Coral type	expected value	p-value
Branching	0.67	0.008
Plate	0.76	0.01
Massive	2.43	0.44

Statistical and computational details

To improve computational tractability, we aggregate the observations into sub-sites. We also employ the Gaussian Markov random field (GMRF) representation of the spatial Gaussian process as a stochastic partial differential equation (SPDE), using the method of Lindgren et al. [2011] and Lindgren and Rue [2015], as implemented in the R package INLA. This involves discretising the spatial domain into a lattice using a constrained Delaunay triangulation. The triangulation nodes for the SPDE were based on the observed data locations and boundaries were constrained using the Heron Reef shapefile [Roelfsema et al., 2013] with an edge of 0.005 (Fig. S11).

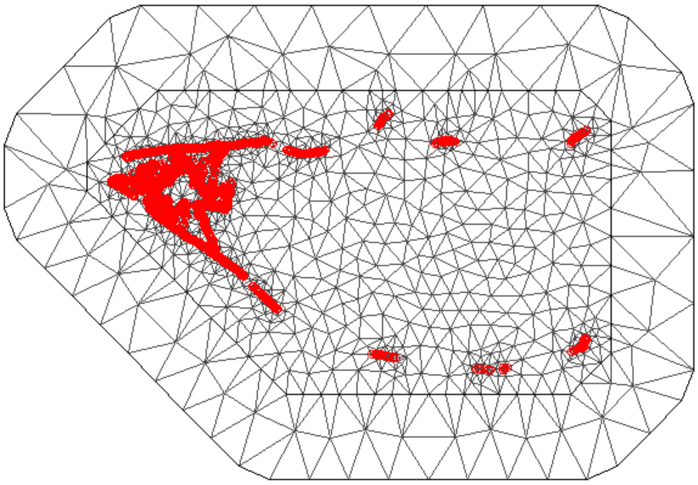

Figure S11: Mesh used to discretise the spatial domain across Heron Reef. Red dots show the observed data locations.

We used penalised complexity (PC) priors [Simpson et al., 2017, Fuglstad et al., 2019] for the precision parameter ϕ of the Beta distribution, as well as for the temporal autocorrelation parameter ω. We also placed a joint PC prior on the spatial range parameter, ℓ, and the marginal standard deviation parameter, σ, of the Matérn SPDE. The independent subsite-level random effects V_i were modelled using a normal distribution with mean zero and a precision parameter τ. The default priors in INLA for τ (log-Gamma distribution), as well as for the fixed effects of the intercept β_0 and of the geomorphic zones β_1 (normal distributions) were used. The precision parameter ϕ is estimated as a single value for all observations.

Model choice

A total of six model formulation was tested on the three coral groups. Differences between the model lay within the formulation of the spatio-temporal random effects with and without time nested into the SPDE model. Different types of correlation between years were tested to formulate the temporal random effects including autoregressive (ar1), independent (iid), random walk (rw1) and exchangeable. For the three coral forms, the best model formulation (lowest DIC values) was time nested in the SPDE model with an ar1 correlation between years (Figure S12).

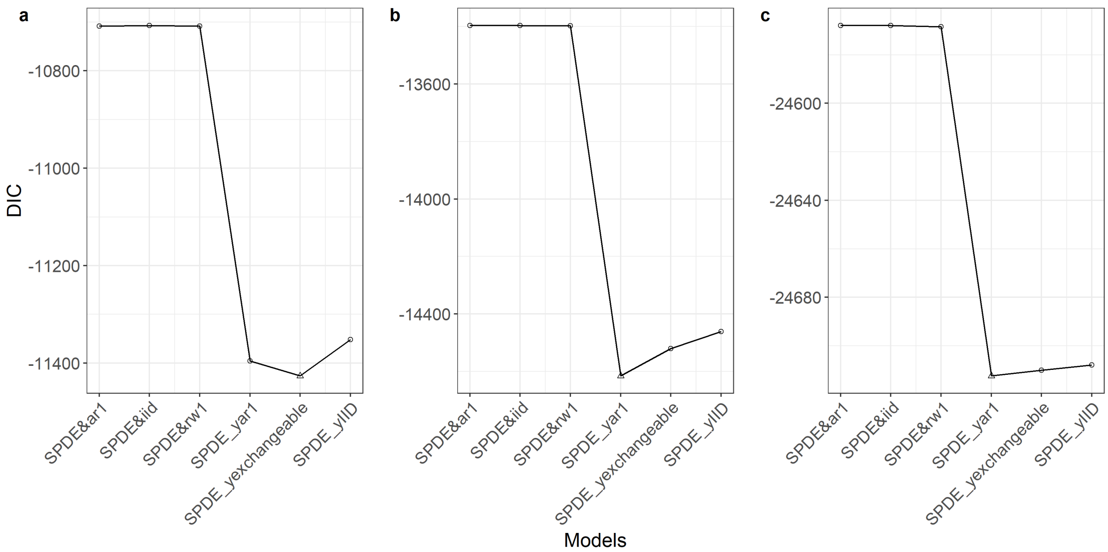

Figure S12: DIC values associated with different model formulations for a) branching, b) plate and c) massive corals. The symbol "&" means that the time effect was separated from the SPDE model, "_" signifies that the temporal random effects were nested in the SPDE model.

Model fit and residuals

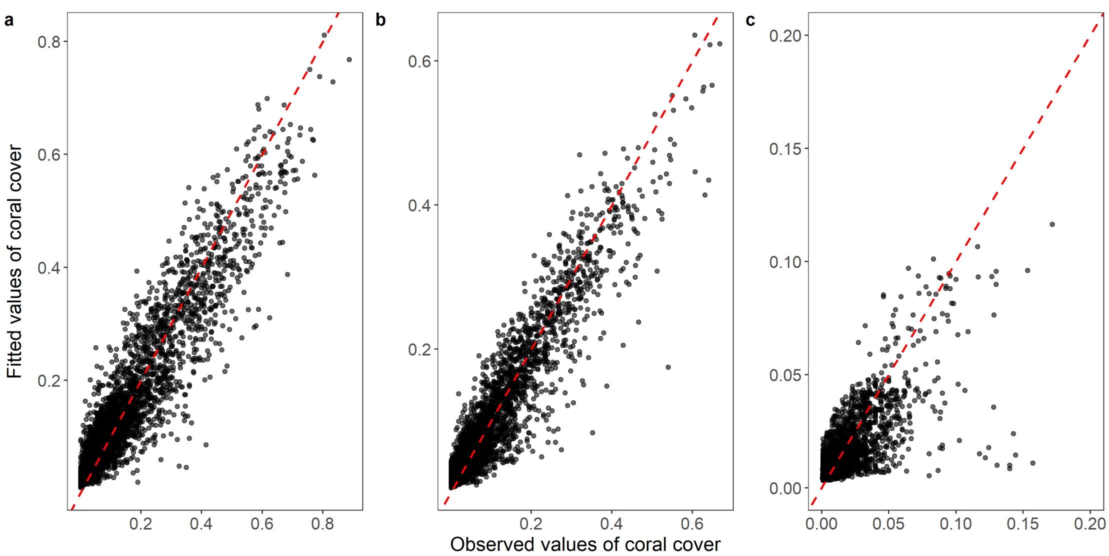

Figure S13: Model fit with a) branching, b) plate and c) massive corals.

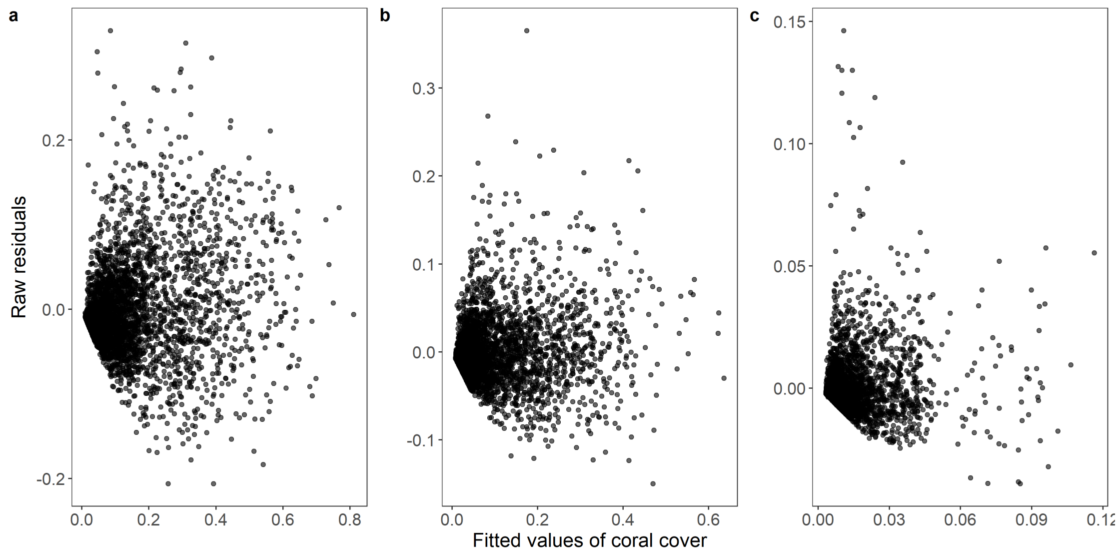

Figure S14: Model residuals with a) branching, b) plate and c) massive corals.

Spatial effects

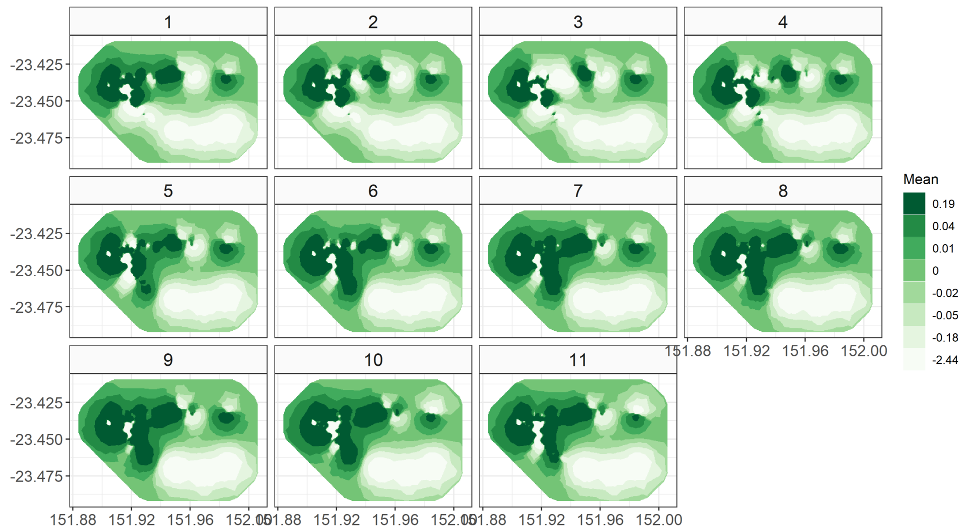

Figure S15: Spatial field estimated from the model for branching corals. Note that the values are on the transformed logit scale.

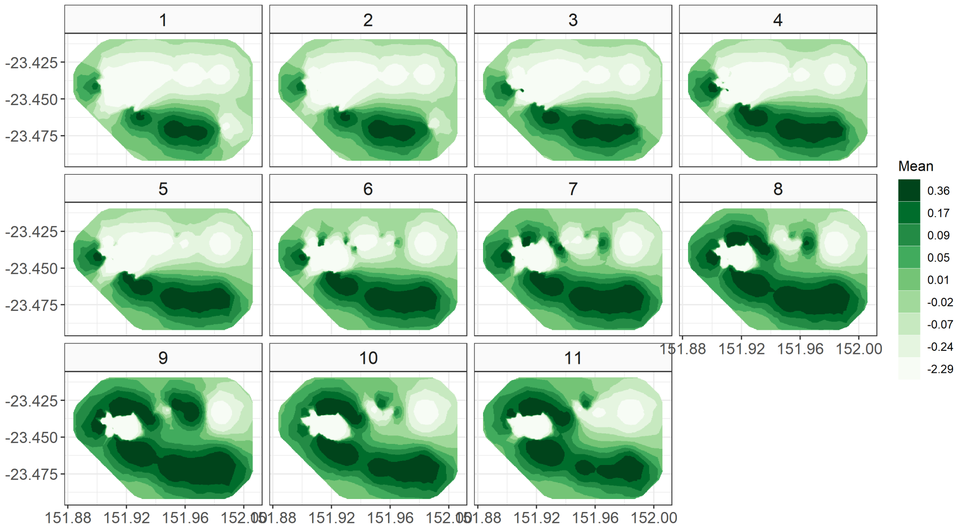

Figure S16: Spatial field estimated from the model for plate corals. Note that the values are on the transformed logit scale.

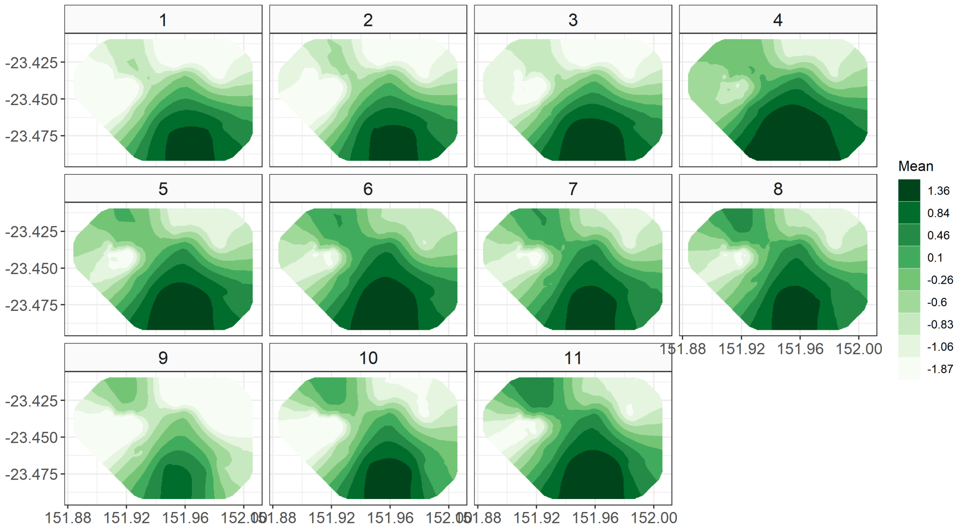

Figure S17: Spatial field estimated from the model for massive corals. Note that the values are on the transformed logit scale.

Summary of spatio-temporal random effects

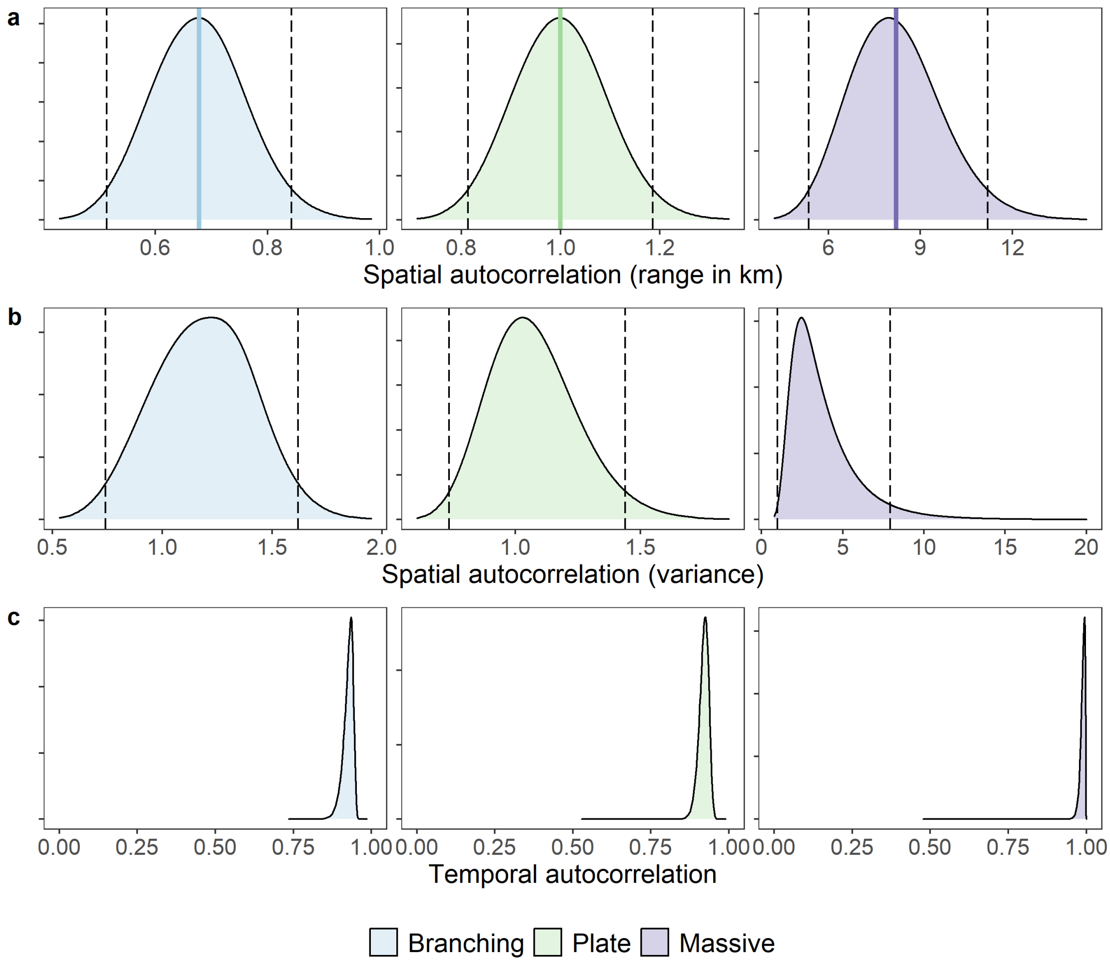

Figure S18: Distributions of the model parameters associated with the spatial and temporal random effects of the model.

Bibliography

Eoghan A Aston, Gareth J Williams, JA Mattias Green, Andrew J Davies, Lisa M Wedding, Jamison M Gove, Jean-Baptiste Jouffray, Timothy T Jones, and Jeanette Clark. Scale-dependent spatial patterns in benthic communities around a tropical island seascape. *Ecography*, 42(3):578–590, 2019.

Mollie E Brooks, Kasper Kristensen, Koen J Van Benthem, Arni Magnusson, Casper W Berg, Anders Nielsen, Hans J Skaug, Martin Machler, and Benjamin M Bolker. glmmtmb balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R journal*, 9(2):378–400, 2017.

Fiona Chong and Matthew Spencer. Analysis of relative abundances with zeros on environmental gradients: a multinomial regression model. *PeerJ*, 6:e5643, 2018.

Noel Cressie. *Statistics for spatial data*. John Wiley & Sons, 2015.

Jacob C Douma and James T Weedon. Analysing continuous proportions in ecology and evolution: A practical introduction to beta and dirichlet regression. *Methods in Ecology and Evolution*, 10(9):1412–1430, 2019.

Helen V Ford, Jamison M Gove, Andrew J Davies, Nicholas AJ Graham, John R Healey, Eric J Conklin, and Gareth J Williams. Spatial scaling properties of coral reef benthic communities. *Ecography*, 44(2):188–198, 2021.

Geir-Arne Fuglstad, Daniel Simpson, Finn Lindgren, and Håvard Rue. Constructing priors that penalize the complexity of Gaussian random fields. *Journal of the American Statistical Association*, 114(525):445–452, 2019. doi: 10.1080/01621459.2017.1415907.

Manuel Gonzalez-Rivero, Oscar Beijbom, Alberto Rodriguez-Ramirez, Dominic EP Bryant, Anjani Ganase, Yeray Gonzalez-Marrero, Ana Herrera-Reveles, Emma V Kennedy, Catherine JS Kim, Sebastian Lopez-Marcano, et al. Monitoring of coral reefs using artificial intelligence: A feasible and cost-effective approach. *Remote Sensing*, 12(3):489, 2020.

Florian Hartig. Dharma: residual diagnostics for hierarchical (multi-level/mixed) regression models. *R package version 0.2*, 4, 2019.

Joshua Levy, Cynthia Hunter, Trent Lukacazyk, and Erik C Franklin. Assessing the spatial distribution of coral bleaching using small unmanned aerial systems. *Coral Reefs*, 37(2):373–387, 2018.

Finn Lindgren and Håvard Rue. Bayesian spatial modelling with R-INLA. *Journal of Statistical Software*, 63(19):1–25, 2015. doi: 10.18637/jss.v063.i19.

Finn Lindgren, Håvard Rue, and Johan Lindström. An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498, 2011. doi: 10.1111/j.1467-9868.2011.00777.x.

Patrick AP Moran. Notes on continuous stochastic phenomena. *Biometrika*, 37(1/2): 17–23, 1950.

Chris Roelfsema, Stuart Phinn, Stacy Jupiter, James Comley, and Simon Albert. Mapping coral reefs at reef to reef-system scales, 10s–1000s km2, using object-based image analysis. *International Journal of Remote Sensing*, 34(18):6367–6388, 2013.

Daniel Simpson, Håvard Rue, Andrea Riebler, Thiago G Martins, and Sigrunn H Sørbye. Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, 32(1):1–28, 2017. doi: 10.1214/16-STS576.

Waldo R Tobler. A computer movie simulating urban growth in the detroit region. *Economic geography*, 46(sup1):234–240, 1970.

Julie Vercelloni, Benoit Liquet, Emma V Kennedy, Manuel González-Rivero, M Julian Caley, Erin E Peterson, Marji Puotinen, Ove Hoegh-Guldberg, and Kerrie Mengersen. Forecasting intensifying disturbance effects on coral reefs. *Global change biology*, 26 (5):2785–2797, 2020.