Studying coral reef patterns in UAE waters using panel data analysis and multinomial logit and probit models

Haïfa Ben-Romdhane¹,²*, Taha B.M.J. Ouarda², Prashanth Marpu³, Anbiah Rajan⁴, Ibrahim Bugla⁵, Richard J.O. Perry⁶, Hosni Ghedira⁴

¹ Khalifa University, Masdar Campus, Abu Dhabi, United Arab Emirates
² INRS-ETE, National Institute of Scientific Research, Quebec City, Canada
³ Environment Agency-Abu Dhabi, Abu Dhabi, United Arab Emirates

ARTICLE INFO

Keywords:
Coral reef patterns
United Arab Emirates
Panel data analysis
Multinomial logit
Probit models

ABSTRACT

Like coral reefs around the world, the reefs of the United Arab Emirates (UAE) are facing global climate change and associated threats. The coasts and islands that flank Abu Dhabi host an important number of corals that should be the focus of conservation actions. Well-designed conservation and management plans require efficient monitoring systems that include understanding coral reef patterns. To understand some of these patterns; coral cover data, satellite-derived and in-situ water quality parameters from nine key reef environments in the UAE from 2011 to 2014 to model coral patterns were used. The objectives were to model coral patterns and realistically predict coral damage intensity with changing environmental variables. Coral damage cover models were defined and estimated for the coral damage cover. Effects of environmental factors were estimated, and predictions of coral damage intensity were presented with changing factors. Main findings, based on the studied data, showed that nutrient enrichment, a proxy for anthropogenic pressure, and salinity are the most influential factors to induce coral damage in UAE waters. Furthermore, results demonstrated that the probability of severe damage increases with decreasing water oxygenation and with increasing temperature, light, salinity, acidity and nutrient levels. The defined and estimated predictions accounted for corals’ behavioural aspects, across individual reefs and over time. This approach is more appropriate than estimation predictions that just account for historic trends. Nevertheless, there are, probably, many components within the model framework that can be expanded and/or improved as more information become available. An extended dataset will enable a means to independently validate the defined models and test other modelling approaches. Continually increasing the in-situ and remote sensing data sizes, spatially and temporally, defines a long-term priority.

1. Introduction

Corals face threats from many sources such as those associated with global warming and climate change (Hughes et al., 2003). Bleaching and mortality are manifestations of excessive and/or long exposure to these threats. Bleaching in cnidarian/dinoflagellate endosymbiosis translates into a symbiosis breakdown. This is known to result from physiological damage to animal host cells and/or symbionts (Douglas, 2003; Hoegh-Guldberg, 1999). Corals can survive a bleaching event, but they are under more stress and are subject to mortality (McClanahan, 2004). Corals have shown some adaptive responses in adjusting to changes in environmental and climatic conditions (Hughes et al., 2003; LaJeunesse et al., 2010). The coral’s (animal) hosts are capable of rapidly adjusting to rises in sea-surface temperatures (SST) and light fluxes following two hypothesised mechanisms; by 1) switching to a more thermally tolerant symbiotic partner (Buddemeier and Fautin, 1993; Oliver and Palumbi, 2011) and/or by 2) reducing the stress on their symbionts. The latter can be done by means of produced pigments, enzymes, and acquired amino acids (Baird et al., 2009; Dove et al., 2006; Dove et al., 2008; Fitt et al., 2009; Lesser and Farrell, 2004; Middlebrook et al., 2008) (e.g., fluorescent pigments (Matz et al., 1999; Salih et al., 2006; Salih et al., 2000) and mycosporine-like amino acids (Shick and Dunlap, 2002)). In addition, some carnivorous corals may increase their effective carnivory under stressful conditions (Grottoli et al., 2006). This change in diet reduces demands on symbionts and allows the latter to allocate more energy to defend their own algal cells (Baird et al., 2009). The animal host as well as the microalgal symbionts are therefore, both implied in determining the coral’s response to

* Corresponding author at: PO Box 127788, Abu Dhabi, United Arab Emirates.
E-mail address: haifa.benromdhane@ku.ac.ae (H. Ben-Romdhane).

https://doi.org/10.1016/j.ecolind.2019.106050

Received 12 March 2019; Received in revised form 22 December 2019; Accepted 28 December 2019

1470-160X/ © 2019 Elsevier Ltd. All rights reserved.
environmental and climatic changes. The present study considers the coral holobiont (the entire assemblage: host and symbiont) to be the appropriate unit to determine the susceptibility of corals to climatic and environmental stressors.

Historical environmental conditions shape populations and communities, as well as their response to disturbances. As bleaching progresses, for instance, it is changing the response of individual reefs to thermal stress by progressively altering the species mix on the reef via the loss of the more thermo-sensitive corals. Evolution by natural selection can occur rapidly in some situations (Carroll et al., 2007). Assessing and monitoring the same coral individuals over time is important in order to, qualitatively and quantitatively, determine this evolutionary process. This study focuses on observations collected from nine key coral reef environments flanking the waters of Abu Dhabi, United Arab Emirates (UAE).

The life history of UAE corals suggests there is a capacity for corals harbouring Symbiodinium thermophilum symbiont to tolerate harsh environmental conditions and extreme events (Hume et al., 2013, 2015, 2016). It is, however, an as-yet undefined capacity to adapt to changing climates and extreme events for UAE corals in general, e.g. corals hosting symbionts other than Symbiodinium thermophilum. Such ability for UAE corals can be to tolerate and/or acclimate to extreme environmental conditions (Ben-Romdhane et al., 2018). It can also be their ability to recover after climatic disturbances such as those witnessed in the late 1990s and early 2000s associated with the 1997–98 and 2009–10 strong El Niño seasons that have caused massive bleaching and mortality across UAE corals (Rieg, 2002, 2003; Wilkinson, 2000). Maintaining the ability to exhibit recovery trajectories after disturbance, or coral resilience, is a central and defining goal in preserving coral reef ecosystems (Mumby and Steneck, 2008).

Coral damage as a response to environmental stress is induced by a variety of factors, acting alone or in synergy (Brown, 1997; Brown and Dunne, 2016; Eakin et al., 2009; Hoegh-Guldberg, 1999; Vega Thurber et al., 2014). Most coral bleaching and mortality events were associated with anomalous sea temperatures; positive anomalies typically (Atwood et al., 1988; Brown et al., 1996; Hoegh-Guldberg, 1999; Hughes et al., 2003). The second main factor associated with coral bleaching is solar irradiance, both its visible (PAR: 400–700 nm) and ultraviolet (UVR: 290–400 nm) portions of the spectrum (Lesser and Farrell, 2004). Water quality effect in modulating the corals’ thermal bleaching limits was highlighted by Wooldridge (2009). Ban et al. (2014) reviewed multiple-stressor interaction studies on coral reefs. They stated that interaction of irradiance and temperature on corals has been the subject of most research (33% of the total reviewed studies) than any other combination of stressors, with an important number of studies reporting a synergistic effect on coral symbiont photosynthetic performance. West and Salm (2003) reviewed the environmental factors that are likely to correlate with resistance and resilience to coral bleaching. They listed: cloud cover, temperature variability, turbidity, light absorption, wind, high wave energy, upwelling and proximity to deep water. They also listed broad size and species distributions and history of corals surviving bleaching events as indirect indicators of bleaching tolerance. Maina et al. (2008) used SST, wind velocity, surface current velocity, UV irradiance, PAR and chlorophyll-a concentration. Guinotte and Buddemeier (2008) commented that Maina et al. (2008) omitted the acidity variable and drew attention to the need for models designed to incorporate the synergistic effects of numerous environmental changes occurring in concert (e.g. higher ocean temperatures plus lower aragonite saturation states). Increasing acidity was stated to be a significant coral stressor (Anthony et al., 2011; Anthony et al., 2008; Doney et al., 2009; Kleypas et al., 1999), Middlebrook et al. (2010) found out that the rate of heat accumulation is an important variable that was not accounted for using the duration and intensity of thermal anomalies determined as degree heating weeks. Salinity was stated to be a stressful factor when abruptly fluctuating: negatively (Brown, 1997; Kerswell and Jones, 2003) or positively (Ferrier-Pages et al., 1999) depending on the coral environment. UAE marine environment is more likely to be negatively affected by high salinity levels due to the presence of important desalination capacities located on the western margin of the Arabian Gulf (Uddin et al., 2011; Wilson et al., 2002). Dissolved oxygen is used by corals during respiration and oxygen depletion were stated to limit coral respiration (Dodd et al., 2007; Revsbech et al., 1995). Dissolved oxygen concentrations were shown to interact with seasonal change in temperature, irradiance and nutrient enrichment (Langdon and Atkinson, 2005; Suzuki et al., 1996). Nutrients associated with increased anthropogenic pressure, importantly chlorophyll, nitrogen and phosphorus compounds, have been defined to increase the susceptibility of corals to bleaching (Wiedenmann et al., 2013), diseases (Bruno et al., 2003; Voss and Richardson, 2006) and decline (Hallock and Schlager, 1986). Nominated environmental factors expected to drive UAE coral damage variation were defined based on findings from the reviewed studies and based on their availability. Variable selection and parametrisation are explained in 2.2.

Understanding patterns of corals’ response to environmental factors is required to develop a decision support system which can facilitate sustainable corals. Modelling paves the way to such understanding (Mumby, 2016). Most models, using natural and anthropogenic factors, attempt to explain coral response such as bleaching, mortality, acclimation and evolution (Donner et al., 2005; Gailani et al., 2016; Hendeel et al., 2001; Hoare and Tsokos, 2009; Hoegh-Guldberg, 1999; Holmes and Johnstone, 2010; Hughes et al., 2003; Liu et al., 2006; Maina et al., 2008; Meesters et al., 1998; Nelson et al., 2016; Sheppard, 2003; Skirving et al., 2006; Ware et al., 1996; Wooldridge and Done, 2004). Unfortunately, a modest number of studies have attempted to model coral reef processes including coral bleaching in UAE waters (Paparella et al., 2019; Purkis and Riegl, 2005; Riegl and Purkis, 2009; Riegl et al., 2011). The limitation is often due to the nature and availability of environmental data. The latter can limit the success and usefulness of coral reef models, as well. While models of reef processes, traditionally, require precise and regular data, frequent and regular observations on coral reefs are lacking (Meesters et al., 1998). Often, valuable information is non-quantitative (Silvert, 1997). The imprecise nature of ecological interrelationships and the subjectivity of field observers make the modelling of coral reef ecological processes complicated (Bosserman and Ragade, 1982; Silvert, 1997).

To provide more insights into coral damage (that includes coral bleaching, disease and death) and its changing intensity as a function of environmental variables in the UAE waters, the present work employs statistical analysis and modelling to assess two main objectives. It first, defines the main environmental factors that affect UAE corals and accurately estimates their interactions with coral cover. It, then, realistically predicts coral damage intensity with changing environmental variables. To the best knowledge of the authors, this is the first study to approach corals using panel data analysis, and multinomial logit and probit models. In addition, it is the first to use wide-ranging coral assessment data (from nine stations) and water quality monitoring data (from seven stations) combined, from along the Abu Dhabi coast over a period of four years, thus, supposed to be significantly representative of Abu Dhabi and UAE corals. This will help further understand UAE corals-specific ability factors such as their tolerance, acclimation and resilience to environmental conditions and climatic disturbances.

2. Methods

2.1. Dataset and study area

Current metrics for predicting coral status including bleaching, e.g. National Oceanic and Atmospheric Administration’s Coral Reef Watch products, might apply well for certain regions, e.g., Eakin et al. (2010) and Liu et al. (2003), and apply less well for others (Krug et al., 2013). In regions where current metrics do not apply well, alternative
predictive approaches and modelling are required.

The Environment Agency – Abu Dhabi (EAD) is the predominate observer of UAE corals. Yearly observations started as early as 2006 and became more frequent (twice a year) in 2010. These observations covered nine key reefs along UAE coasts. Coral damage records; including bleaching, coral disease and death, started in 2011.

EAD’s monitoring programme has measured marine water quality parameters in several locations throughout UAE waters, starting as early as January 2008 in Port Zayed. In 2011, additional stations augmented the programme. This study examines temporal and spatial correspondence between the coral cover observations and information from the water quality monitoring. The used dataset included in situ observations from all coral stations and seven of the water quality monitoring stations, in addition to satellite-derived data. The study area covered the coral reef environments that are flanking Abu Dhabi coast (Fig. 1). The study period covered four years between 2011 and 2014.

The cross-sectional data analysis or the time-series data analysis of the data provided by EAD were limited and imprecise if considered separately. Several advantages, presented here, are realised by restructuring the data containing time series and individual observations; i.e. by blending the inter-individual differences and intra-individual dynamics of these data. Such hierarchical structure corresponds to panel data and allows for greater degrees of freedom and more sample variability than cross-sectional data or time-series data alone (Hsiao, 2007). This allows a higher capacity of analysis and modelling to capture the complexity of the variables’ behaviour, as it provides more accurate modelling inference by improving the efficiency of the model’s predicted estimates.

Panel data analysis proliferated since 1986, after the publication of Hsiao (1986) and Balestra and Nerlove (1966). The proliferation marked the econometrics society above all with more available data and because of the additional capacity for modelling the complexity of human behaviour panel data provided (Hsiao, 2014). Individuals, represented here by the monitored coral reefs, are part of the same ecosystem when considering Abu Dhabi waters as one environment. In this regard, panel data analysis are proven to generate more accurate predictions for individual outcomes by pooling the data rather than generating predictions of individual outcomes using the data on the individual in isolation (Hsiao, 2007; Hsiao et al., 1993, 1989).

Furthermore, panel data analysis simplifies the modelling for computation and statistical inference. The panel data analyses rely on the inter-individual differences to reduce the collinearity between current and lag variables to estimate unrestricted time-adjustment patterns (Hsiao, 2007). They also remedy the violation of the normality condition that is invoked by non-stationary data (due to the temporal feature of the data in question). The independence of cross-sectional observation units (different reefs) can invoke the central limit theorem across cross-sectional units to demonstrate that the limiting distributions of many estimators remain asymptotically normal (Binder et al., 2005; Im et al., 2003; Levin et al., 2002; Phillips and Moon, 1999). In addition, panel data analysis is suitable as a solution to the coral observations, being relatively limited in size; ordered logistic (logit) (McCullagh,
1980) and probability-unit (probit) (Bliss, 1934) models are proposed after discretisation of coral observations.

The multinomial logit and probit models were selected for their simplicity and interpretability (Long and Freese, 2006). In fact, logit and probit models are widely used to predict disease severity and mortality outcomes (Biondo et al., 2000; Boyd et al., 1987; Homser and Lemeshow, 2000; Kologlu et al., 2000; Le Gall et al., 1993; Marshall et al., 1995). The discretisation of coral observations and use of logit and probit models in this study complement the interpretation of coral data analysis by providing further insights into UAE corals’ ecology by predicting categorical coral damage outcomes given their environmental state-variables.

2.2. Variable selection and parametrisation

The selection of the analytical tools and parameters were based on the understanding of the ecological and physiological mechanisms reviewed in several studies such as Ware et al. (1996), Meesters et al. (1998), Hoegh-Guldberg (1999), Maina et al. (2008), Holmes and Johnstone (2010), Buddemeier et al. (2011), Browne (2012), Krug et al. (2013), Evenhuis et al. (2015), Gailani et al. (2016), Nelson et al. (2016), Chen et al. (2018), Stuart-Smith et al. (2018) and Beetham et al. (2018). Stepwise estimation was then, performed using several logics including backward- and forward-selection as well as backward- and forward-stepwise searches (Bendel and Affifi, 1977). Forward-selection starts with fitting the dependent variable on a constant while adding the most-significant variable and re-estimating the model at every step. A backward-selection model of damaged coral cover was performed using different search logic steps. During backward-selection, the full model was fitted on all proposed explanatory variables, the least-significant term is removed and the model is re-estimated. Each variable is treated as its own term and thus considered separately. Backward stepwise involves starting off in a backward selection approach and then potentially adding back variables if they later appear to be significant. It, thus, fits the full model on all proposed explanatory variables, removes the least-significant terms at every step and re-estimates the model. Repeatedly, it reconSIDers all dropped variables (except the most recently dropped), adds the most-significant term and re-estimates the model. During forward selection, the probability to enter terms was first set to 0.99 so that all of the variables would enter and their order of entry is observed. During backward selection, the probability to remove terms was, first, set to 0.01 so that all of the variables except the last one would be removed and their order of removal is observed. During backward stepwise, two separate significance levels were chosen for deletion from and adding to the model. The adding significance has to be smaller $p$ value (more stringent) than the dropping significance. The stepwise estimation helped distinguishing seawater temperature, photosynthetically active radiation, ultraviolet radiation, salinity, pH, dissolved oxygen, nitrite and phosphate as statistically significant terms in explaining the damaged coral cover variable.

2.2.1. Seawater temperature (SWT)

Seawater temperatures (SWT, °C) were recorded using HOBO® Water Temp Pro v2 loggers installed at each coral monitoring station for EAD for their coral reef assessment and monitoring programme. The loggers record hourly temperatures averaged for the month of observation, as used here for analysis, i.e. for each time-period.

2.2.2. Photosynthetically active radiation (PAR)

Photosynthetically active radiation (PAR, (E m⁻² s⁻¹)) data were National Aeronautics and Space Administration (NASA)/s Moderate-resolution Imaging Spectroradiometer (MODIS) - Aqua data downloaded from Ocean Productivity online data repository (http://www.science.oregonstate.edu/ocean.productivity/index.php). PAR were Level-3 9 km spatially binned MODIS cloud-corrected incident daily data, reprocessing 2014.0, that defined the quantum energy flux from the sun in the spectral range of 400–700 nm (NASA, 2016). The data were retrieved in HDF format for all time-periods between 2011 and 2014 before being spatially subset to the study area.

2.2.3. Ultraviolet radiation (UVR)

UVR increases that were proved to affect the host tissues and algal symbionts are between 290 and 310 nm; those which will increase in the event of decreases in stratospheric ozone over equatorial regions (Lesser and Lewis, 1996). Those UVR increases were stated by Lesser and Lewis (1996) to have impacts on: growth, reproduction and occurrence of bleaching for shallow water corals. Daily global maps of UVR relevant to general DNA damage (256–370 nm) were retrieved from the Tropospheric Emission Monitoring Internet Service (http://www.temis.nl/index.php). UVR expressed in kJ m⁻², represented the integration of the erythemal UV index, as derived from satellite observations, from sunrise to sunset, with a time step of 10 min (Van Geffen et al., 2005). The integration takes the cloud cover into account and thus leads to an estimate of the daily erythemal UV dose. Daily UVR data were monthly averaged for the study time-periods before being spatially subset to the study area.

2.2.4. Salinity

Salinity data (ppt or %o) were in-situ data collected by EAD using Hydrolab® DSS water quality multi-probe at several stations for marine water quality parameter monitoring. Data were spatially interpolated using the Inverse Distance Weighting (IDW) method, (Shepard, 1968), to match salinity measurements to coral monitoring stations.

2.2.5. Acidity (pH)

pH data (unit less) represented EAD’s in-situ data collected using Hydrolab® DSS water quality multi-probe at the marine water quality parameter monitoring several stations. Similar to salinity, pH data were spatially interpolated using IDW to match measurements to the coral monitoring stations.

2.2.6. Dissolved oxygen concentration (DO)

Dissolved oxygen data (mg L⁻¹) represented EAD’s in-situ data recorded using Hydrolab® DSS water quality multi-probe at the marine water quality parameter monitoring stations. Dissolved oxygen data were also spatially interpolated using IDW to match coral monitoring station locations.

2.2.7. Nitrite and phosphate

Water samples were collected in-situ by EAD at each marine water quality monitoring station and analysed for nitrite (as NO₂⁻) and phosphate (as PO₄³⁻) both in μg L⁻¹ among other nutrients. Nitrite and phosphate were analysed using the American Public Health Association (APHA) 4500 NO₂ (B) and P (C) test methods, respectively.

2.2.8. Coral cover data

During EAD’s coral monitoring and assessment, coral and other benthic classes were identified with reference to their geographic locations and cover proportions. Bleaching and morbidity are highly and positively correlated with coral mortality (Glynn and D’croz, 1990; McClanahan, 2004; Peters, 2015). Corals in decline often have high fleshy macroalgae biomass that inversely correlates with coral cover (Bruno et al., 2009; Cummings et al., 2018; Hughes, 1994; McCook et al., 2001). This provided, among others, a good indicator of ecosystem degradation. Coral tissues’ aspect was another ecological indicator. Bleached tissue may appear white (translucent) or pale (discoloured) (Fig. 2. b). In bleached corals, the polyp tissues can still be seen above the skeleton. Diseased corals would present impairments to a coral’s normal condition as a sign of stress (Fig. 2. c). There exist several diseases and conditions. Most observed indicators of diseased corals include black band-, white band- and, mostly, yellow band diseases. Black band disease’s indicator is a discrete, concentric or linear,
dark black or maroon band or bacterial mat at the interface of live tissue and dead skeleton. White band disease is distinguished by a band of white, exposed skeleton or bleached tissue separating healthy coral tissue from exposed or algae covered skeleton. Yellow band disease has a small yellow or white ring surrounding dead tissue. Non-living parts of a coral in which the corallite structures are often no longer present or covered over by organisms such as algae, sponges and sediments, were the indication of the coral’s death (Fig. 2. d). Proportions of bleached, diseased and dead corals were referred to as damaged corals. The latter represented the focus of the current study in an attempt to understand and describe corals’ susceptibility through establishing relations between coral vulnerability, manifested in damaged corals, and environmental stressors.

Corals located around the monitoring stations are, throughout this study, referred to as entities or individual reefs. The term “individual” is referring to the different studied reefs monitored and modelled for coral cover. The spatial autocorrelation or spatial dependence among these reefs is not considered as the spatial proximity among observational units and the numeric similarity among their values were not studied. Times of observations are referred to as time periods or dates. The dependent variable is the percentage of damaged coral cover recorded across the nine stations (individual reefs) from 2011 to 2014. Highest recorded amounts of damaged coral cover were observed in Al Yasat over the studied period. The damaged coral cover seems to have a positive trend across the studied period with highest value in October 2011; that being one year after 2009-10 El Niño (Riegl et al., 2011) (Fig. 3). The independent variables are the marine water quality variables (Appendix A). The independence of the environmental variables is, initially, investigated using numerical methods such as the Pearson’s correlation coefficient (Benesty et al., 2009) and further examined for variables that exhibited relatively high correlations.

2.2.9. Data characteristics and variation types

Properties and quality of panel data are crucial in panel data analysis as they significantly influence the model selection (Park, 2011). Data were arranged to create balanced panel data, i.e. all reef entities have measurements in regular time periods (summer and winter of each year from 2011 to 2014). The same entities (coral monitoring stations) are observed for each period.

Park (2011) provides a checklist of data features to be examined in order for the data set to be panel data in a statistical sense and not only by being physically arranged so. Park (2011)’s checklist was followed to evaluate and ensure the qualities of the dataset. The data size was also checked to be appropriate for panel data analysis. Theoretically, Granger (1999) and Park (2011) addressed panel samples’ size and concluded that the number of entities and/or time periods should not be too small (Type I error problem) or too large (Type II error problem). Practically, model panel data analysis cases are, in average, such as constructed by Caves et al. (1984), using six individuals and fifteen times which lead to ninety sample points (Park, 2005, 2011). The data set of concern has fewer observations; nine individual reefs by eight time periods; that is seventy-two sample points. It is here argued that although the dataset size is relatively below the average, it should not discourage modelling attempts as data limitation is a common problem in marine environmental studies and coral research in particular (Guinan et al., 2009; Maina et al., 2008). Finally, the data set was verified to be strongly balanced, i.e. each panel contains the same number of observations. Following confirmation of the commensurability of the data set, panel data modelling analyses fixed and/or random effects embedded in the panel data of concern.

In order to perform panel data analysis, the dataset was declared cross-sectional (across coral monitoring stations, n = 9 reef entities) and time-dependent (half-yearly during four years of monitoring, T = 8 time periods). Formulation of data variation types are expressed in Table 1. Overall statistics are ordinary statistics that are
between statistics are computed based on individual reefs’ or entities’ statistics, while within statistics summarise the time periods’ statistics regardless of the entities.

The dependent variable has more within variation (15.52%), i.e. variation over time, than between variation (from one location to the next) (13.81%). Similarly, variations of water quality variables were larger over time than across monitoring stations, e.g. 4.85 vs. 1.51 °C for SWT, 4.38 vs. 1.76 E m⁻² s⁻¹ for PAR, 1.14 vs. 0.83 ppt for salinity and 0.14 vs. 0.06 for pH. The lowest percentage of coral damage was observed at Al Dhabiyah in April 2011, while the highest was observed at Al Yasat in October 2014. The latter corresponded to high SWT and highest PAR conditions (34.29 °C and 59.60 E m⁻² s⁻¹). The within variation (over time) of the damaged coral cover ranged from −11.75 and 62.61%, which is not to say that any station has observed negative damage. The within damage cover refers to the deviation from each station’s average, and naturally, some of those deviations must be negative.

### 2.3. Modelling coral cover distribution

#### 2.3.1. Modelling coral cover spatio-temporal distribution using pooled and individual-specific effects models

Panel data modelling is about examining entity- (or individual) specific and/or time effects in order to isolate unobserved heterogeneity. Entity-specific and/or time effects are either fixed effects (FE) or random effects (RE). Fixed effect models examine the variability of the model’s intercept across individual reefs or time periods. Random effect models investigate differences in the model’s error variance components across individual reefs or time periods (Hsiao, 2014). Statisticians recommend to start with a simpler model (Kennedy, 2008; Park, 2011). The strategy, here, is to first examine the poolability of the panel data in hand through ordinary least squares (OLS) model. Investigation on unobserved heterogeneity to determine entity- and/or time period effect using fixed effect and random effect models follows.

Interpretation of the results will complement the estimated parameters from the model that fits the data well and produces statistically significant individual regressors. Equations for panel data models and estimators are expressed in Appendix B. Panel data estimators are several, and their characteristics differ. Estimators that are consistent and efficient are preferred. Consistency is based on the law of large numbers. Efficiency is based on minimum variance.

When an estimator is consistent, more observations will tend to provide more precise and accurate estimates. If the true model is the pooled model and the regressors are uncorrelated with the error terms, POLS regressor is consistent. Whereas, if the true model is fixed effects, then POLS regressor is inconsistent. The between estimator uses the between variation only. POLS and RE estimators are more efficient than between estimator. The within or fixed effects estimator uses time-demeaned variables thus providing the individual-specific deviations of variables from their time-averaged values. It is an OLS estimation of the time-demeaned dependent variable on the time-demeaned regressors.

### Table 1

Panel data variation types.

<table>
<thead>
<tr>
<th>Variation of the dependent variable and regressors</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual mean</td>
<td>( \pi = \frac{1}{T} \sum \Sigma x_i ) (1)</td>
</tr>
<tr>
<td>Overall mean</td>
<td>( \pi = \frac{1}{N} \sum \Sigma x_i ) (2)</td>
</tr>
<tr>
<td>Overall variation (variation over time and individual reefs)</td>
<td>( x_i - \pi ) (3)</td>
</tr>
<tr>
<td>Overall variance</td>
<td>( s_i^2 = \frac{1}{N-1} \sum \Sigma (x_i - \pi)^2 ) (4)</td>
</tr>
<tr>
<td>Between variation (variation between individual reefs)</td>
<td>( \pi - \pi ) (5)</td>
</tr>
<tr>
<td>Between variance</td>
<td>( s_i^2 = \frac{1}{N-1} \sum \Sigma (\pi_i - \pi)^2 ) (6)</td>
</tr>
<tr>
<td>Within variation (variation within individual reef over time)</td>
<td>( x_i - \pi ) (7)</td>
</tr>
<tr>
<td>Within variance</td>
<td>( s_i^2 = \frac{1}{T-1} \sum \Sigma (x_i - \pi)^2 ) (8)</td>
</tr>
</tbody>
</table>

Fig. 3. Variation of damaged coral cover from 2011 to 2014 within Abu Dhabi waters.
FE estimator always gives consistent estimates, however, they may not be the most efficient. The first-differences estimator uses first-differences variables, i.e. it uses the one-period changes for each individual. It is an OLS estimation of the one-period changes of the dependent variable on the one-period change in the regressors. One observation vanishes for each individual because of first-differencing. Random effects estimator is an OLS estimation of the transformed model (Appendix B). RE estimates are a weighted average of the between and within estimates. RE estimator is consistent and most efficient under the random effects model. However, it is inconsistent if the appropriate model is FE model. Choosing between fixed and random effects models and between POLS and panel data models is crucial. A summary of this approach is given in Fig. 4.

2.3.2. Predicting probabilities of categorical coral cover distribution using ordered logit and probit models

Estimated interaction terms are useful to infer how environmental factors affect the coral cover distribution. These interaction terms are used, in addition to pooled and individual-specific effects, in non-linear models, such as logistic (logit) and probability-unit (probit) models (Table 2). In ordinal logit and probit models, the coral cover dependent variable has three categories \( J = 3 \) in meaningful order (Table 3). An example of categorical coral cover distribution is given in Fig. 5.

![Fig. 4. Modelling coral cover spatio-temporal distribution using pooled and individual-specific effects models.](image)

![Fig. 5. The probability of an observed significant coral damage \( y = 2 \) is the highlighted area under the curve between the pair of cutpoints \( \tau_1 = 30\% \) and \( \tau_2 = 50\% \).](image)

Table 2

<table>
<thead>
<tr>
<th>Ordered logit and probit models and CDFs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>The probability that observation ( i ) will select alternative ( j )</td>
</tr>
<tr>
<td>( P_j = P(y_i = j) ) (11)</td>
</tr>
<tr>
<td>( P_j = P(y_i &lt; j \leq y_i \leq j) ) (12)</td>
</tr>
<tr>
<td>( P_j = F(y_i - \beta \lambda) - F(y_i - \beta \lambda) ) (13)</td>
</tr>
</tbody>
</table>

FE estimator always gives consistent estimates, however, they may not be the most efficient. The first-differences estimator uses first-differences variables, i.e. it uses the one-period changes for each individual. It is an OLS estimation of the one-period changes of the dependent variable on the one-period change in the regressors. One observation vanishes for each individual because of first-differencing. Random effects estimator is an OLS estimation of the transformed model (Appendix B). RE estimates are a weighted average of the between and within estimates. RE estimator is consistent and most efficient under the random effects model. However, it is inconsistent if the appropriate model is FE model. Choosing between fixed and random effects models and between POLS and panel data models is crucial. A summary of this approach is given in Fig. 4.

2.3.2. Predicting probabilities of categorical coral cover distribution using ordered logit and probit models

Estimated interaction terms are useful to infer how environmental factors affect the coral cover distribution. These interaction terms are used, in addition to pooled and individual-specific effects, in non-linear models, such as logistic (logit) and probability-unit (probit) models (Table 2). In ordinal logit and probit models, the coral cover dependent variable has three categories \( J = 3 \) in meaningful order (Table 3). An example of categorical coral cover distribution is given in Fig. 5.
The index model for the latent coral damage cover variable is as follows.

\[ y_i^* = \hat{\beta} x_i + u_i \]  

(9)

\[ y_j = j \text{ if } \tau_{j-1} < y_i^* \leq \tau_j \text{ for } j = 1 \text{ to } J \]  

(10)

where \( i \) is the observation, \( u \) is a random error and the cutpoints \( \tau_i \) through \( \tau_{J-1} \) are estimated.

The marginal effect of an increase in an environmental variable \( x_r \) on the probability of selecting alternative \( j \) is the following.

\[ \frac{\partial P_j}{\partial x_{ri}} = [F(\tau_{j-1} - \hat{\beta} x_i) - F(\tau_{j-1} - \hat{\beta} x_i)]\hat{\beta}_r \]  

(16)

The probability of an observed outcome for a given value of \( x \) is formulated as:

\[ Pr(y = j | x) = F(\tau_j - \hat{\beta} x) - F(\tau_{j-1} - \hat{\beta} x) \]  

(17)

and predicted probabilities can be estimated using the following formula.

\[ \hat{P}_j(y = j | x) = F(\hat{\tau}_j - \hat{\beta} x) - F(\hat{\tau}_{j-1} - \hat{\beta} x) \]  

(18)

with cumulative probabilities computed as:

\[ \hat{P}_j(y \leq j | x) = F(\hat{\tau}_j - \hat{\beta} x) \]  

(19)

### 3. Results and discussion

#### 3.1. Environmental impact on coral distribution and the unobserved ability of corals

**3.1.1. Estimated effects of environmental variables on coral damage cover**

The estimation suggests, among others, that an increase in SWT is associated with higher coral damage for all estimators. It also suggests that an increase in dissolved oxygen concentrations is associated with lower coral damage for all estimators. An increase in one dissolved oxygen unit will lead to less 9% coral damage. In addition, it is estimated that an increase in one dissolved oxygen unit for one individual reef above its own average will lead to less 10% and 9% coral damage (FE and RE). The results also show that within variation (of individual reefs over time) is well explained by FE and RE estimators. The models' estimated 39% of variation proportion to be explained by individual-specific term. The estimated effects of each environmental variable on coral damage cover are detailed in Appendix C.

**3.1.2. Individual-specific effects**

The main advantage of individual-specific effects modelling is that it can better detect and measure effects that cannot be simply observed in pooled regression. Subsequent to fixed effects model estimation, the unobserved ability of an individual (tolerance) \( \xi_i \) that affects coral response to environmental variables could be recovered from the remaining variation in the dependent variable that cannot be explained by the independent variables such that in:

\[ \hat{\xi}_i = \hat{\tau}_i - \hat{\beta} \hat{x}_i \]  

(20)

The representation of the individual-specific effects is given in Fig. 6.
then, investigated using Cook’s distance (Cook’s D) (Cook, 1977, 1979) and difference in fits (DFITS) measures (Belsley et al., 1980). These measures scale differently, however, both combine information on the residual and leverage thus providing similar results. The observations are compared against a conventional cut-off point estimated to be $4/(NT)$ (i.e. 0.06) for Cook’s D and $2/\sqrt{k/(NT)}$ (i.e. 0.67) for DFITS. Cook’s D and DFITS measures for Al Dhabiyah – April 2011 observations are by far the largest (0.3 and 1.6, respectively). DFBETA measure (Rethemeyer, 2007) was then considered to assess the change in coefficients based on observations taken at stations and time periods worthy of further investigation. DFBETA assesses the specific impact of an observation on the regression coefficients. DFBETA values in excess of $|2/\sqrt{NT}|$ (i.e. 0.24) are further investigated. It is often advised that highly influential observations are dropped from the regression model. The change of the model’s coefficients due to the deletion of these observations was assessed for each environmental variable. In this case, due to spatial and temporal dependence of the observations, influential and usual observations detected using these measures might be individual-specific, e.g. attributed to inconsistency in observations, as they might be associated to actual exceptional climatic events such as El Niño event. It was therefore, opted to highlight those points and discuss them without dropping them from the model.

3.1.3.1.2. Diagnostics on normality of residuals. A kernel density estimate (Silverman, 1981) of POLS model’s residuals was plotted against true normal density distribution (Fig. 8. a). Probability plots, e.g. standardised normal probability (P-P) and quantile–quantile probability (Q-Q) plots were also used to assess the POLS model’s residuals’ normality (Wilk and Gnanadesikan, 1968). Q-Q plot (Fig. 8. b) shows a slight deviation from normal at the upper tail, as could be seen in Fig. 8. a. There seems to be a minor but insignificant deviation of POLS residuals from normality. It, thus, can be accepted that POLS residuals are close to a normal distribution. Furthermore, numerical tests for testing normality; Shapiro-Wilk (W) (Shapiro and Wilk, 1965) and inter-quartile range (IQR) (Hamilton, 1992) tests were performed. W test calculates $p$ value that is based on the assumption that the distribution is normal. In this case, $p$ value was large (0.66868), indicating the hypothesis that residuals are normally distributed cannot be rejected. IQR assumes symmetry of the distribution. It reports severe outliers that are either 3 inter-quartile ranges below the first quartile or 3 inter-quartile ranges above the third quartile. IQR indicated the presence of zero severe outliers, which is evidence of the non-rejection of normality at 5% significance level. It was reported one mild outlier, which is very common in samples of any size (Hamilton, 1992). No severe outliers and a fair symmetric distribution of the residuals are additional indications on the normality of POLS residuals.

3.1.3.1.3. Diagnostics on homoscedasticity of residuals. Homogeneity of variance of the residuals is one of the main assumptions of the OLS regression (Jarque and Bera, 1980). Graphical and numerical methods are used to detect heteroscedasticity of the residuals. In Fig. 8. c, a pattern of data points getting narrower towards the right end is an indication of heteroscedasticity. Numerical tests, White test (White, 1980) and Breusch-Pagan test (Breusch and Pagan, 1979), were conducted to confirm this observed heteroscedasticity. Both tests assess the null hypothesis that the variance of the residuals is homogenous. $p$ values obtained from White test and Breusch-Pagan (0.3328 and 0.1759, respectively) were significant enough to not reject the null hypothesis of homoscedasticity.

3.1.3.2. Hausman test. As RE estimator is most efficient, Hausman test needs to support this property of the estimator. Otherwise, FE model is to be used. Hausman test is used to test whether there is a significance difference between FE and RE estimators. The test is chi-square distributed with degrees of freedom equal to the number of regressors. Hausman test was significant prevailing the use of FE model.

3.1.3.3. Breusch-Pagan Lagrange Multiplier test and consistency of POLS estimators. Breusch-Pagan Lagrange Multiplier test is a test for the random effects model based on the OLS residual. It tests whether $\sigma^2_\epsilon$ or equivalently $\text{cor}(u_\epsilon, u_\epsilon)$ is significantly different from zero. Here, the LM test is insignificant, thus, POLS model prevails RE model. The use of panel data modelling was still useful as it allowed the coral cover’s observation across individual reefs and compared it against its observation over time. As the true model is the pooled model, POLS estimators are consistent if regressors are uncorrelated with residual terms. The correlation of the residuals with the regressors were calculated to be zero in all cases and were confirmed graphically to be inexistent (Appendix D).
3.2. Severity of coral damage as a function of environmental variables

3.2.1. Estimated effects of environmental variables on damage severity

Using the data of concern, the following model was estimated.

$$\Pr(damage \ level = j | x) = F(t_j - \hat{\beta}x) - F(t_{j-1} - \hat{\beta}x)$$  \hspace{1cm} (22)

where

$$\hat{\beta}x = \beta_{SST} \cdot SST + \beta_{PAR} \cdot PAR + \beta_{UVR} \cdot UVR + \beta_{Salinity} \cdot Salinity + \beta_{pH} \cdot pH + \beta_{Dissolved\ oxygen} \cdot Dissolved\ oxygen + \beta_{Nitrite} \cdot Nitrite + \beta_{Phosphate} \cdot Phosphate$$  \hspace{1cm} (23)

The coefficients were estimated using ordered logit and ordered probit models. The models' estimated coefficients are presented in Appendix D. In both models, the null hypothesis that all of the coefficients associated with independent variables are simultaneously equal to zero was rejected at 95% confidence level. The intercept cuts represent the cut-points or thresholds introduced in Section 2.3.2. The ordered logit and probit models with three damage levels ($j$ alternatives) have one set of coefficients with two ($j - 1$) intercepts. The threshold parameters (intercept cuts) are different from each other so the number of categories of coral damage is reasonable. McFadden's $R^2$ or Pseudo-$R^2$, also known as the “likelihood-ratio index”, assesses the predictive strength of the logit and probit models (Hu et al., 2006). As a rule of thumb, Pseudo-$R^2$ ranging from 0.2 to 0.4 indicates very good model fit (Hensher and Stopher, 1979).

The ordered logit and probit models have three sets of marginal effects. The sign of parameters shows whether the latent coral damage categorical variable ($y^*$) increases with the environmental variables (independent variables). The coral status is worse with higher SWT, PAR, UVR, salinity and nutrient enrichment. The Coral damage is mitigated with higher pH (basicity) and dissolved oxygen. The magnitude of the coefficients is different by a scale factor of about 1.8 between the probit and logit models, therefore, the magnitude of the coefficients cannot be interpreted. For example, the ratio of the logit to probit coefficient is 1.78 for salinity and 1.74 for pH and dissolved oxygen. The difference in coefficient estimates reflects the differing scaling of the ordered logit and ordered probit models. The exception to the 1.8 ratio was for SWT and PAR. Values of z-statistics are very similar since they are not affected by the scaling.

3.2.2. Marginal change in degree of severity of coral damage cover

The marginal effects of environmental variables for different levels of coral damage are represented in Appendix F. The marginal effects of each environmental variable on the different coral damage alternatives sum up to zero. The marginal effects are interpretable such that each unit increase in the environmental variables increases or decreases the probability of selecting alternative $j$ by the marginal effect expressed as a percent. For instance, a one-unit increase in Salinity is associated with being 13% less likely to be in the indicative damage status and 7% more likely to be in the strong damage level. A one-unit increase in Dissolved oxygen, for instance, is associated with being 30% more likely to be in lower level of damage and 16% less likely to be in the higher level of damage.
damage categories. The marginal effects for the probit model are similar to those of the logit model. The marginal effect of an independent variable, as expressed in Eqn 16, is the derivative; that is, the slope of the prediction function, which, by default, is the probability of success following logit or probit estimations. As the slope of a function can be greater than one, margins can be estimated to be greater than one. It was the case for the marginal effect of pH for indicative damage Appendix F. Conditional marginal effect of pH for the indicative damage level; that is the marginal effect of pH where all covariates are set to fixed values, exceeds 1 for pH values greater than 8.02 (Fig. 9. a). The marginal effect of pH for the indicative damage level, considered by its own, shows a change in the marginal effect greater than 1 (the slope of the tangent line at pH = 8.02) (Fig. 9. b).

3.2.3. Predicted probabilities of coral damage level

The logit and probit predicted probabilities between categorical predictor variables and categorical outcomes were calculated (Appendix G). The predicted probabilities for the indicative and severe damage categories tend to be less than 20%, with the majority of predictions for the middle category, falling between 20% and 80% (Appendix G). The mean predicted probabilities are very close to the probabilities of the actual coral damage cover variable presented in Table 4.

Predictions were then computed at specific, substantively more informative values. To do so, predictions were determined for individual reefs with a particular set of characteristics at a time. The predicted probabilities for individual reefs were examined for individual reefs with the following characteristics.

### Table 4

<table>
<thead>
<tr>
<th>Damage category/ level</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Logit predicted probabilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indicative</td>
<td>0.373333</td>
<td>0.394131</td>
<td>0.0000183</td>
<td>0.9999979</td>
</tr>
<tr>
<td>Strong</td>
<td>0.320011</td>
<td>0.262241</td>
<td>0.0000020</td>
<td>0.7746989</td>
</tr>
<tr>
<td>Severe</td>
<td>0.306656</td>
<td>0.367798</td>
<td>0.0000000</td>
<td>0.9988372</td>
</tr>
<tr>
<td>Probit predicted probabilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indicative</td>
<td>0.37551</td>
<td>0.394716</td>
<td>0.0000000</td>
<td>1.0000000</td>
</tr>
<tr>
<td>Strong</td>
<td>0.315665</td>
<td>0.2542</td>
<td>0.0000000</td>
<td>0.7430087</td>
</tr>
<tr>
<td>Severe</td>
<td>0.308825</td>
<td>0.368226</td>
<td>0.0000000</td>
<td>0.9999418</td>
</tr>
</tbody>
</table>

- High temperature and light ($>\text{SWT} = 29^\circ\text{C}$ and $>\text{PAR} = 52.05\text{ E m}^{-2}\text{s}^{-1}$)
- High radiation ($>\text{PAR} = 52.05\text{ E m}^{-2}\text{s}^{-1}$ and $>\text{UVR} = 2.71\text{ kJ m}^{-2}$)
- High salinity ($>\text{Salinity} = 44.03\text{ ppt}$)
- High acidity ($<\text{pH} = 8.12$)
- Low oxygenation ($<\text{Dissolved oxygen} = 4.81\text{ mg L}^{-1}$)
- High nutrient enrichment ($>\text{Nitrite} = 4.7\text{ µg L}^{-1}$ and $>\text{Phosphate} = 178.5\text{ µg L}^{-1}$)

3.2.3.1. Prediction of coral damage level under high temperature and high light conditions. The predicted probabilities of the three categories of damaged coral level are given for changing SWT values (Fig. 11. a). The vertical line marks the average SWT (28.99972 °C). The probability of severe damage increases with increasing SWT. The probability of indicative damage decreases most rapidly with increasing SWT.
With changing PAR values (Fig. 11 b), the probability of severe damage increases with increasing PAR. The probability of indicative damage decreases more rapidly with increasing PAR. PAR values as high as 57.70 E m⁻² s⁻¹ are more likely to cause severe damage.

The photic zone that supports coral reefs is generally vertically well-mixed (Riegl and Purkis, 2012; Sheppard et al., 2010). George et al. (2005) state that thermoclines rarely develop in Abu Dhabi open waters, even when air temperatures reach 50 °C during long summer days. The observed SWT can, thus, be discussed in light of reported SST. Indeed, SST thresholds reported for coral bleaching by (Shuail et al., 2016) in Saadiyat (34.48 °C), Ras Ghanadah (34.55 °C) and Dalma (35.05 °C), and for significant mortality by Riegl and Purkis (2009) in the Arabian Gulf (35.5 °C). These values were not captured for the studied time periods and stations (Fig. 10). This might explain the probabilities that did not define a SWT threshold to cause severe damage (Fig. 11 a), unlike predicted damage probabilities with changing PAR (Fig. 11 b).

3.2.3.2. Prediction of coral damage level under high UV solar radiations. The predicted probabilities of severe damaged coral level increase with increasing UVR (Fig. 11 c). The vertical line marks the average UVR (2.710278 kJ m⁻²). The probability of indicative damage decreases more rapidly with increasing UVR. UVR levels higher than 2.7 kJ m⁻² are more likely to cause severe damage to UAE coral habitats.

It is to be noted that over periods of months, the coral photosystem acclimates to changing light quickly (obtaining fast acclimation within 5–10 days) (Anthony and Hoegh-Guldberg, 2003). The defined amount of light (PAR and/or UVR) that causes coral damage is, therefore, to be continuously be monitored.

3.2.3.3. Prediction of coral damage level under high salinity. Indicative coral damage decreases very rapidly with increasing salinity levels (Fig. 11 d). While salinity levels of 42.3 ppt are more likely to cause strong damage to the studied coral reefs, salinity levels higher than 44.5 ppt are liable to cause severe damage.

3.2.3.4. Predictions of coral damage level under high acidity. As stated by several studies (Anthony et al., 2008, 2011; Doney et al., 2009; Kleypas et al., 1999), increasing acidity represents a significant coral stressor. In this study, high acidity levels, pH levels lower than 8.05 are revealed to be inducers of severe coral damage (Fig. 11 e). As basicity increases, coral damage level decreases. Probabilities of different damage levels change at the same rate.

3.2.3.5. Predictions of coral damage level under low oxygenation. Dissolved oxygen concentrations lower than 4.8 mg L⁻¹ are likely to cause severe damage to the studied coral reef environments (Fig. 11 f). Indicative damage level changes more rapidly with changing dissolved oxygen concentrations. Seawater oxygenation seems to reduce the likelihood of severe damage occurrence.

3.2.3.6. Predictions of coral damage level under high nutrient enrichment. Higher nutrient enrichment is more likely to cause severe coral damage (Fig. 11 g and h). Nitrite concentrations higher than 14 μg L⁻¹ are more likely to cause severe damage to the studied reefs (Fig. 11 g). The effects of phosphate enrichment affect the coral health more significantly than nitrite enrichment (Fig. 11 g and Fig. 11 h).

Fig. 10. SWT variation and reported SST thresholds: bleaching (Shuail et al., 2016) and mortality (Riegl and Purkis, 2009).
Phosphate concentrations are more likely to be detrimental when phosphate concentrations exceed 210 μg L\(^{-1}\) \(\text{(Fig. 11. h).}\)

### 3.2.4. Model testing and post-estimation analysis\(^1\)

#### 3.2.4.1. Proportional odds assumption

Proportional odds assumption (also known as parallel regression assumption for binary logit and probit models) takes on the decomposition of the ordered logit (or probit) model into \(J-1\) binary logit (or probit) models assuming that the slope coefficients are identical across each regression. This would mean that the lowest versus all higher categories of the response variable are the same as those that describe the relationship between the next lowest category and all higher categories; and so on \(\text{(Long and Freese, } 2014\text{).}\)

Proportional odds assumption is important as it helps investigating the main model’s residuals and detecting influential observations \(\text{(Section 3.2.4.2).}\) It needs, however, to be proven valid. To the degree that the proportional odds assumption holds, the coefficients \(\beta_1, \beta_2, \ldots, \beta_{J-1}\) should be close. This is tested using an approximate Likelihood Ratio (LR) test \(\text{(Wolfe and Gould, } 1998\text{), essentially. LR test compares the log likelihood from the main logit model to that obtained from pooling \(J-1\) binary logit models while making an adjustment for the correlation between the binary outcomes defined by } y \leq j; \text{ the null hypothesis being that the coefficients are equal across categories. The results from LR test reported a non-significant } p \text{ value } = 0.1541 \text{ indicating that the proportional odds assumption cannot be rejected at 95% confidence level. However, LR test is an omnibus test that the coefficients for all variables are simultaneously equal. It does determine if the coefficients for some variables are identical across the binary equations while coefficients for other variables differ \(\text{(Long and Freese, } 2014\text{).}\) To this end, proportional odds assumption’s validity for the damage level ordered logit model is also tested using Wald test \(\text{(Brant, } 1990\text{).}\) It tests the parallel regression assumption for each variable individually. The estimated coefficients from \(J-1\) binary regressions are given in Appendix H. Wald test results of an omnibus test for the entire model reports the non-significant \(p \text{ value } = 0.523\) indicating that the proportional odds assumption has not been violated at 95% confidence level. However, results from testing the assumption for each of the independent variables in the main model revealed a violation of this assumption for the dissolved oxygen variable \(p \text{ value } = 0.036\) \(\text{(Appendix H).}\) In an ideal parallel regression assumption case, each probability curve differs only in being shifted to the left or right as coefficients are presumably equal for each equation. Predicted probabilities of dissolved oxygen’s effects on the damage level for the two binary models \(\text{(Fig. I.1) were compared to the variable with most parallel equations (highest } p \text{ value } > \chi^2; \text{ that is SWT \(\text{(Fig. I.2).}\) The numerical \(p \text{ value } = 0.036\) and graphical \(\text{(Fig. I.1) inspections of the violation of proportional odds assumption seems to be minor, however, further investigations were performed. These investigations involved a post-estimation analysis that performs five tests: likelihood ratio test, score test, Wald test, Wolfe-Gould test, and Brant test, \(\text{(Buis and Williams, } 2013\text{), in order to compare the ordered}}\)

---

\(^1\) In order to avoid redundancy and since results from ordered logit and probit models were very similar, model testing and post-estimation analysis are reported for the logit model only.
logit model with the fully generalised ordered logit model.

Results from this post-estimation analysis are given in Table 5. Akaike’s information criterion (AIC) (Akaike, 1974) and Bayesian information criterion (BIC) (Schwarz, 1978) both indicate that the ordered logit is the better-fitting model among the two. Based on all these findings, it is decided to ignore the violation of the proportional odds assumption manifested by dissolved oxygen assumption and proceed with further post-estimation analyses.

### 3.2.4.2. Diagnostics on residuals and influential observations. Hosmer and Lemeshow (2004) suggest the application of the binary models to the $J - 1$ cumulative probabilities of the ordered logistic model. This approach requires the dichotomisation of the model. Using Eq. Eqn(25) and based on the proportional odds assumption, the predicted probabilities in the ordered logit model can be expressed as:

$$
\Pr(y \leq 1 | x) = F(q_1 - \hat{\beta}x)
$$

$$
\Pr(y \leq 2 | x) = F(q_2 - \hat{\beta}x)
$$

(24)

The original ordinal dependent variable was decomposed into two binary variables ($y_1$ and $y_2$) coded as tabulated below. Binary logits for $y_1$ and $y_2$ using the same independent environmental variables were estimated. Standardised residuals: Pearson and deviance residuals, were generated for each estimated logit. Pearson residual defines the standardised difference between the observed frequency and the predicted frequency. Pearson residual has the form:

$$
\rho_i = \frac{y_i - \hat{y}_i}{\hat{S}_i}
$$

(25)

$\rho_i$ measures the relative deviations between the observed and fitted values.

Deviance residual measures the disagreement between the maxima of the observed and the fitted log likelihood functions. Logit uses the maximal likelihood principle, therefore, the objective is to minimise the sum of deviance residuals (similarly to OLS regression, where the objective is to minimise the sum of squared residuals).

Deviance residual is defined as:

$$
d_i = -2(logL(\hat{y}_i, y) - logL(y_i, y))
$$

(26)

where $L(\hat{y}, y)$ denotes the maximum value of the logit likelihood function and $\hat{y}_i$ denotes the estimated parameters for the saturated logit.

Leverages are measured using the diagonal of the hat (or projection) matrix (Pregibon, 1981). The hat matrix projects the vector of observations, $y_i$, onto the vector of predictions, $\hat{y}_i$. Plots of Pearson, deviance residuals, and leverages are given for each binary logit. Two types of plots are considered for each of Pearson, deviance and leverage statistics: the plots of the statistics against the predicted values, and the plots of these statistics against the index identifier (also called index plots). These two types of plots convey the same information. On the first type of plots, data are indexed by station identifier. The same data points are indexed by date identifier on index plots, making it convenient to track observations by location and time of record for further interpretation.

The examination of Pearson residuals and deviance residuals identified observations recorded at Saadiyat in April 2013, Al Dhibiyah in October 2014 and Bu Tinah in October 2012 to have high Pearson and deviance residuals. The examination of leverages identified records taken at Makasib in April 2012, Bu Tinah in April 2012 and in October 2014, Barakah in October 2012, Saadiyat in October 2011, in October 2012 and October 2013, Ras Ghanadhah in October 2014, Al Yasat in April 2011, and Dalma in April 2013 to be influential.

Influential observations (high residuals and leverages) detected using POLS are supposed to be more correct since influential observations defined using ordered logit were based on an approach evaluating an approximation to the model estimated (the coefficients of the binary models differed from those estimated in the ordinal model). There were, yet, observations detected to be influential using POLS as well as ordered logit models. These were observations recorded at Saadiyat coral monitoring station in April 2013 (detected to have high residuals), in addition to observations taken at Saadiyat monitoring station in October 2012, Ras Ghanadhah in October 2014, Barakah in October 2012, and Dalma in April 2013 (detected to have high leverages in ascending order).

### 4. Conclusion

Survival of corals in extreme environments can be enhanced through increased dependence on heterotrophic feeding (Grottoli et al., 2006), a community shift to more resistant and tolerant species (Done, 1999) or the switching to more temperature-tolerant symbionts (Little et al., 2004). Modelling is useful in examining the ecological impact of such adaptation (Ware et al., 1996). Data availability for the current study is limited in extent; hence, the choice for panel data analysis, as well as logistic and probit models over time-series analysis, and forecasting models. This work has endeavoured to provide solutions to challenges such as data shortage. Through the developed panel data analysis, it was possible to elucidate robust relationships from a relatively limited data set. The assessment analysis indicates that nutrient enrichment, a proxy for anthropogenic pressure, and salinity are the most influential factors in inducing coral damage in UAE waters. These findings can be useful for coral reef managers. Ordered logit and probit models were used to generate predictions of coral damage intensity given a set of state variables. For most environmental factors, the thresholds for significant coral damage were attained. However, for some, such as water temperature, the required values were not captured for the studied time periods and stations. Nevertheless, the model simulated the effect of increasing temperatures and expressed this in the generated predictions. Predictive ecological models are known to be data-hungry (Wisz et al., 2008). Despite the limited sample size, predictions were coherent and substantively interpretable. Predictions that are more realistic require feeding the model with fresh environmental data on a regular basis, as this would enable a more efficient coral reef management.

The defined and estimated predictions accounted for patterns observed in corals across individual reefs and over time. This approach is more appropriate than estimation predictions that just account for historic trends. The models developed here may not be definitive coral reef ecosystem models since there are, probably, many components within the model framework capable of expansion and/or improvement as more information become available. However, by achieving the objectives stated previously, an additional tool to support local and regional coral studies of population ecology and behaviour is believed to be founded. An extended dataset will enable a means to independently validate the defined models and test other modelling approaches. Continually increasing the in-situ and remote sensing data sizes, spatially and temporally, defines a long-term priority. This may also require the addition of new variables. Colony size measurement and morphology/geometry description, for instance, could be more
informative and add greater precision and substance on solar irradiance effects on UAE corals. Indeed, greater morphological complexity due to bigger colony size can result in more self-shading, leading to less light absorption (Enrıquez et al., 2005; Stambler and Dubinsky, 2005). Integrating observations of the atmospheric medium and its properties by including the effect of aerosols in addition to sea fluid flow and sedimentation monitoring may also help the more effective study of reef hydrodynamics and sediment transport. It is anticipated that the integration of these variables will empower more efficient management approaches.

Moreover, objective and more consistent characterisation of coral cover, such as bleaching observations, are crucial in defining resilience, tolerance and adaptation responses. They are also needed for distinguishing anomalous bleaching from regularly occurring bleaching; that is a natural acclimatisation phenomenon (Suggett and Smith, 2011). As in fact, mis- (under)-estimation of the potential resistance and resilience of reefs against environmental change ultimately limits (depreciates) the validity and effectiveness of reef management policies and practices.

Lastly, it is essential to mention that statistical and quantitative evidence of the interactions of variables, requires piecing together, along with the biological and ecological interactions within the context of the system being studied (Dunne, 2010). In addition to collecting in-situ light measurements to validate associated results, coral sampling and biological analyses are also recommended to be integrated to ensure most accurate results.

Acknowledgments

The authors would like to express their gratitude to the Editor in Chief, Dr. João Carlos Marques, and one anonymous reviewer whose comments helped to improve the quality of the paper, considerably. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendix A. Loess scatterplot smoothing of damage coral cover on environmental predictors. $R^2 = 0.9460$.

A local regression (loess), (Jacoby, 2000), of the damaged coral cover on the environmental variables based on a method by Hastie and Tibshirani (1990) is represented here. It expresses the multivariate structure in the data as a sum of different bivariate relationships, each of which expresses the impact of an independent variable on the dependent variable with the effects of the other independent variables removed. Loess is a useful non-parametric graphical tool for initial data exploration. It is non-parametric in the sense that the fitting technique does not require an a priori specification of the relationship between the dependent and independent variables (Jacoby, 2000).
### Appendix B: Panel data models, estimators and implications.

<table>
<thead>
<tr>
<th>Model</th>
<th>Estimator</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pooled OLS (P-OLS)</td>
<td>( y_{it} = \alpha + \beta u + u_{i} ) (27)</td>
<td>( \hat{y}<em>{it} = \alpha + \hat{\beta} u</em>{it} + (u_{i} - \alpha + \epsilon_{it}) ) (28)</td>
</tr>
<tr>
<td>Between</td>
<td>( \bar{\gamma} = \alpha + \beta \bar{\pi} + u_{i} ) (29)</td>
<td>( \hat{\gamma}<em>{i} = \alpha + \hat{\beta} \bar{\pi}</em>{i} ) (30)</td>
</tr>
<tr>
<td>Fixed effects (-FE)</td>
<td>( y_{it} = \alpha + \beta x_{it} + u_{i} ) (31)</td>
<td>( nT ) (POLS and RE estimators are more efficient)</td>
</tr>
<tr>
<td>First differences</td>
<td>( y_{it} - \bar{\gamma}<em>{i} = \beta (x</em>{it} - \bar{x}) + (u_{i} - \alpha) ) (32)</td>
<td>( nT ) (individual-specific effect can be recovered)</td>
</tr>
<tr>
<td>Random effects (RE)</td>
<td>( y_{it} = \hat{\alpha} + \hat{\beta} (x_{it} - \bar{x}) + v_{it} ) ( \bar{\gamma}<em>{i} = \bar{x}</em>{i_{-1}} + \bar{\beta} (x_{i_{-1}} - \bar{x}) + v_{i_{-1}} ) (34)</td>
<td>( n(T-1) )</td>
</tr>
</tbody>
</table>

### Appendix C: Estimated effects of environmental variables on coral damage cover for each model

<table>
<thead>
<tr>
<th>Damage cover</th>
<th>OLS</th>
<th>Between</th>
<th>Within or Fixed Effects</th>
<th>First differences</th>
<th>Random Effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWT</td>
<td>0.7525753* (2.26)</td>
<td>8.606165</td>
<td>0.964748* (3.17)</td>
<td>1.10054* (4.42)</td>
<td>0.7525753* (2.26)</td>
</tr>
<tr>
<td>PAR</td>
<td>0.7134898 (1.61)</td>
<td>−5.120748</td>
<td>0.6933973 (1.66)</td>
<td>0.7613713 (1.66)</td>
<td>0.7134898 (1.61)</td>
</tr>
<tr>
<td>UVR</td>
<td>2.051944 (0.80)</td>
<td>−7.020989</td>
<td>−0.3071922 (−0.12)</td>
<td>−2.610667 (−0.83)</td>
<td>2.051944 (0.80)</td>
</tr>
<tr>
<td>Salinity</td>
<td>2.921125* (3.72)</td>
<td>5.733294</td>
<td>1.659293* (1.94)</td>
<td>2.610375* (2.93)</td>
<td>2.921125* (3.72)</td>
</tr>
<tr>
<td>pH</td>
<td>−37.48543* (−4.70)</td>
<td>−145.4454</td>
<td>−33.80257* (−4.55)</td>
<td>−38.48543* (−4.70)</td>
<td>−37.48543* (−4.70)</td>
</tr>
<tr>
<td>Nitrite</td>
<td>0.052* (6.08)</td>
<td>0.029253</td>
<td>0.0423441* (5.28)</td>
<td>0.041648* (5.90)</td>
<td>0.0519938* (6.08)</td>
</tr>
<tr>
<td>Phosphate</td>
<td>184.646* (2.43)</td>
<td>1137.966</td>
<td>187.7073* (2.50)</td>
<td>0.041648* (5.90)</td>
<td>184.646* (2.43)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.08277</td>
<td>0.8285</td>
<td>0.8285</td>
<td></td>
<td>0.8285</td>
</tr>
</tbody>
</table>

Value of t-statistics in parentheses.

* indicates significant at 5% level.
Appendix D. Correlation between POLS model’s residuals and estimators

Appendix E. Ordered coral damage model’s coefficients

<table>
<thead>
<tr>
<th></th>
<th>Ordered logit coefficients</th>
<th>Ordered probit coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWT</td>
<td>0.0349698 (0.38)</td>
<td>0.0172681 (0.32)</td>
</tr>
<tr>
<td>PAR</td>
<td>0.0718588 (0.44)</td>
<td>0.0310594 (0.33)</td>
</tr>
<tr>
<td>UVR</td>
<td>1.846671 (1.7)</td>
<td>1.080979 (1.72)</td>
</tr>
<tr>
<td>Salinity</td>
<td>0.9590286* (3.08)</td>
<td>0.5380324 (3.2)</td>
</tr>
<tr>
<td>pH</td>
<td>−7.543817* (−2.78)</td>
<td>−4.332615 (−2.76)</td>
</tr>
<tr>
<td>Dissolved oxygen</td>
<td>−2.257213* (−2.18)</td>
<td>−1.293781 (−2.2)</td>
</tr>
<tr>
<td>Nitrite</td>
<td>0.0465054 (1.17)</td>
<td>0.0254804 (1.09)</td>
</tr>
<tr>
<td>Phosphate</td>
<td>0.0131090* (4.2)</td>
<td>0.0074127 (4.5)</td>
</tr>
<tr>
<td>Intercept cut1</td>
<td>−19.25756</td>
<td>−12.14099</td>
</tr>
<tr>
<td>Intercept cut2</td>
<td>−15.10327</td>
<td>−9.857208</td>
</tr>
<tr>
<td>Pseudo-R²</td>
<td>0.5468</td>
<td>0.5495</td>
</tr>
</tbody>
</table>

Value of z-statistics in parentheses.
* indicates significant at 5% level.
### Appendix F. Marginal effects of environmental variables for different levels of coral damage

<table>
<thead>
<tr>
<th>Environmental Variable</th>
<th>Marginal effects for indicative damage (%)</th>
<th>Marginal effects for significant damage (%)</th>
<th>Marginal effects for severe damage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ordered logit</td>
<td>Ordered probit</td>
<td>Ordered logit</td>
</tr>
<tr>
<td>SWT</td>
<td>−0.0046872 (−0.38)</td>
<td>−0.0045846 (−0.32)</td>
<td>0.0022191 (0.37)</td>
</tr>
<tr>
<td>PAR</td>
<td>−0.0096316 (−0.45)</td>
<td>−0.00828462 (−0.33)</td>
<td>0.0045601 (0.44)</td>
</tr>
<tr>
<td>UVR</td>
<td>−0.2475191 (−1.41)</td>
<td>−0.2869985 (−1.57)</td>
<td>0.1171876 (0.91)</td>
</tr>
<tr>
<td>Salinity</td>
<td>−0.1285437 (−2.48)</td>
<td>−0.1428477 (−2.92)</td>
<td>0.0608588 (1.17)</td>
</tr>
<tr>
<td>pH</td>
<td>1.011138* (2.30)</td>
<td>1.150304* (2.53)</td>
<td>−0.4787217 (−1.19)</td>
</tr>
<tr>
<td>Dissolved oxygen</td>
<td>0.3025462* (2.05)</td>
<td>0.3434973* (2.21)</td>
<td>−0.14324 (−1.15)</td>
</tr>
<tr>
<td>Nitrite</td>
<td>−0.0062334 (−1.11)</td>
<td>−0.006765 (−1.08)</td>
<td>0.0029512 (0.89)</td>
</tr>
<tr>
<td>Phosphate</td>
<td>−0.001756* (−3.12)</td>
<td>−0.0019681* (−3.89)</td>
<td>0.0008314 (1.25)</td>
</tr>
</tbody>
</table>

Value of z-statistics in parentheses.

* indicates significant at 5% level.

### Appendix G. Predicted probabilities for the coral damage categories
### Appendix H. Estimated coefficients from after the dichotomisation of the main logit model into two binary logit regressions

<table>
<thead>
<tr>
<th>Variable</th>
<th>$y &gt; 1$</th>
<th>$y &gt; 2$</th>
<th>$\chi^2$ Statistic</th>
<th>$p &gt; \chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWT</td>
<td>0.086</td>
<td>0.104</td>
<td>0.01</td>
<td>0.937</td>
</tr>
<tr>
<td></td>
<td>(0.68)</td>
<td>(0.52)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PAR</td>
<td>0.001</td>
<td>0.093</td>
<td>0.02</td>
<td>0.882</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UVR</td>
<td>1.677</td>
<td>4.656</td>
<td>0.40</td>
<td>0.527</td>
</tr>
<tr>
<td></td>
<td>(1.47)</td>
<td>(1.00)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salinity</td>
<td>0.941</td>
<td>1.4</td>
<td>0.34</td>
<td>0.562</td>
</tr>
<tr>
<td></td>
<td>(2.18)</td>
<td>(1.94)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pH</td>
<td>−10.926</td>
<td>−7.587</td>
<td>0.21</td>
<td>0.644</td>
</tr>
<tr>
<td></td>
<td>(−2.45)</td>
<td>(−1.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissolved oxygen</td>
<td>−0.08</td>
<td>−9.038</td>
<td>4.38</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>(−0.05)</td>
<td>(−2.16)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nitrite</td>
<td>0.039</td>
<td>0.192</td>
<td>0.92</td>
<td>0.338</td>
</tr>
<tr>
<td></td>
<td>(0.73)</td>
<td>(1.24)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phosphate</td>
<td>0.016</td>
<td>0.028</td>
<td>0.93</td>
<td>0.334</td>
</tr>
<tr>
<td></td>
<td>(2.82)</td>
<td>(2.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>39.196</td>
<td>11.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.23)</td>
<td>(0.34)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Appendix I. Inspections of the violation of proportional odds assumption using conditional marginal effects

#### Appendix I.1. Effects of Dissolved oxygen on damage level in each binary logit models

![Conditional marginal effects of dissolved oxygen with 95% CL](image)

**Conditional marginal effects of dissolved oxygen with 95% CL**

**Pr(Binary logit 1 (y<2))**  
**Pr(Binary logit 2 (y<3))**

#### Appendix I.2. Effects of SWT on damage level in each binary logit models

![Conditional marginal effects of SWT](image)

**Conditional marginal effects of SWT**

**Pr(Binary logit 1 (y<2))**  
**Pr(Binary logit 2 (y<3))**

### Appendix I. Inspections of the violation of proportional odds assumption using conditional marginal effects

HB conceived the work. AR, IB and HB acquired the data. HB analysed and interpreted the data for the work. AR, IB and HB have contributed to the acquisition and interpretation of the data. HB drafted the work. RP, TO, PM and HG have revised the work for important intellectual content. HB, TO, PM, AR, IB, RP and HG have given their final approval to the version to be published.


The National Academy of Sciences, 201601910.


Radiation Monitoring Based on GOME and SCIAMACHY (14), Envisat & ERS Symposium.


