1-1-2012

A comparison of spatially explicit and classic regression modelling of live coral cover using hyperspectral remote-sensing data in the Al Wajh lagoon, Red Sea

S Hamylton

University of Wollongong, shamylto@uow.edu.au

A comparison of spatially explicit and classic regression modelling of live coral cover using hyperspectral remote-sensing data in the Al Wajh lagoon, Red Sea

Abstract

Live coral is a key component of the Al Wajh marine reserve in the Red Sea. The management of this reserve is dependent on a sound understanding of the existing spatial distribution of live coral cover and the environmental factors influencing live coral at the landscape scale. This study uses remote-sensing techniques to develop ordinary least squares and spatially lagged autoregressive explanatory models of the distribution of live coral cover inside the Al Wajh lagoon, Saudi Arabia. Live coral was modelled as a response to environmental controls such as water depth, the concentration of suspended sediment in the water column and exposure to incident waves. Airborne hyperspectral data were used to derive information on live coral cover as a response (dependent) variable at the landscape scale using linear spectral unmixing. Environmental controls (explanatory variables) were derived from a physics-based inversion of the remote-sensing dataset and validated against field-collected data. For spatial regression, cases referred to geographical locations that were explicitly drawn on in the modelling process to make use of the spatially dependent nature of coral cover controls. The transition from the ordinary least squares model to the spatially lagged model was accompanied by a marked growth in explanatory power (R 2 = 0.26 to 0.76). The theoretical implication that follows is that neighbourhood context interactions play an important role in determining live coral cover. This provides a persuasive case for building geographical considerations into studies of coral distribution.

Keywords

spatially, explicit, classic, modelling, live, coral, cover, hyperspectral, regression, remote, comparison, sensing, data, al, wajh, lagoon, red, sea

Disciplines

Life Sciences | Physical Sciences and Mathematics | Social and Behavioral Sciences

Publication Details

Hamylton, S. (2012). A comparison of spatially explicit and classic regression modelling of live coral cover using hyperspectral remote-sensing data in the Al Wajh lagoon, Red Sea. International Journal of Geographical Information Science, 26 (11), 2161-2175.

A comparison of spatially explicit and classic regression modelling of live coral cover using hyperspectral remote sensing data in the Al Wajh lagoon, Red Sea

Abstract

Live coral is a key component of the Al Wajh marine reserve in the Red Sea, and the management of this reserve is dependent on a sound understanding of the existing spatial distribution of live coral cover and the environmental factors influencing live coral at the landscape scale. The present study uses remote sensing techniques to develop ordinary least squares and spatially lagged autoregressive explanatory models of the distribution of live coral cover inside the Al Wajh lagoon, Saudi Arabia. Live coral was modelled as a response to environmental controls such as water depth, the concentration of suspended sediment in the water column and exposure to incident waves. Airborne hyperspectral data were used to derive information on live coral cover as a response (dependent) variable at the landscape scale using linear spectral unmixing. Environmental controls (explanatory variables) were derived from a physics-based inversion of the remote sensing dataset and validated against field-collected data. For spatial regression, cases referred to geographical locations that were explicitly drawn on in the modelling process to make use of the spatially dependent nature of coral cover controls. The transition from the ordinary least squares model to the spatially lagged model was accompanied by a marked growth in explanatory power (R^2=0.26 to R^2=0.76). The theoretical implication that follows is that neighbourhood context interactions play an important role in determining live coral cover. This provides a persuasive case for building geographical considerations into studies of coral distribution.

Keywords: *Spatial regression, Saudi Arabia, spatial autoregression, spatial autocorrelation, live coral cover*

1 Introduction

Coral reefs underpin tropical coastal ecosystems through the provision of ecological services (e.g., mangrove and seagrass growth promotion, structural habitat complexity for fish) and goods (e.g., primary production to support fish and invertebrate populations, calcification) (Côté and Reynolds, 2006). To sustain these goods and services, marine protected areas have been proven a highly effective conservation measure for coral reefs (Roberts et al., 2003). At the heart of marine protected area planning is the need to understand both the existing spatial distribution of live coral and the environmental factors influencing their distribution at the landscape scale (10 – 100 km^2) (Sobel & Dahlgren, 2004; Almany et al., 2009). One of the key challenges to the development of this understanding is the paucity of biophysical datasets available in these frequently large but remote environments. Recent increases in the accuracy, precision and affordability of geospatial technologies (GIS, GPS and remote sensing) provide new opportunities for mapping and modelling live coral cover. Such technologies yield geographically-referenced datasets that allow mapping and modelling exercises to be conducted in a spatially explicit manner. This allows reef managers to quantify spatial patterning in benthic communities, determine optimal sampling strategies for monitoring ecological health and avoid the incorporation of redundancy into datasets (which in turn violates statistical assumptions about the geographical independence of benthic communities across reefs) (Haining, 2003).

Mapping the distribution of live coral cover has largely been made possible through the development of optical satellite and airborne as well as acoustic remote sensing technology and the associated refinement of image processing routines for application to marine environments. Airborne hyperspectral remote sensing campaigns acquire imagery of the requisite spatial and spectral detail to accurately resolve live coral while accounting for the influence of the overlying atmospheric and water column layers on light transfer (Klonowski et al, 2007). The rich content of hyperspectral datasets allows their manipulation to retrieve information on water quality, bathymetry and benthic cover using physics-based inversion techniques (Brando et al., 2009; Hedley et al 2009), spectral unmixing (Goodman and Ustin, 2007), optimization and semi-analytical techniques (Lee et al. 1999; Wettle et al. 2006). Such mapping exercises yield spatially continuous, landscape scale datasets on the distribution of live coral that can be used as a foundation for modelling exercises that further our understanding of the relationship between live coral cover and local environmental influences.

Spatial modelling can be defined as an assemblage of empirical techniques in which a clear association is maintained and exploited between quantitative data and the spatial coordinates that locate them (Chorley 1972). Defined in this way, the application of spatial modelling has largely developed through the establishment of spatially explicit rule sets for defining segments of objectbased image analysis techniques (Benfield et al. 2007) and the use of spatial metrics to quantify spatial patterning on reefs (Le Drew et al. 2000; Phinn et al. 2003; Purkis et al. 2007). In terms of inferential modelling that seeks to explain or predict observable patterns in live coral cover, classic (spatially implicit) statistical approaches have commonly been employed at the landscape scale, such as ordinary least squares regression (Harborne et al, 2006) and generalised additive models (Garza Perez, 2004). Such approaches do not account for the inherent spatial structure of ecosystems (Fortin and Dale, 2005) that is manifest on a coral reef as a result of the autocorrelated distribution of the environmental characteristics that determine coral survival.

The objective of this study is to use hyperspectral remote sensing techniques to implement and compare two different multivariate regression models that seek to explain the spatial distribution of live coral cover inside a lagoon at the landscape scale. A wide variety of controls could potentially influence the proportion of live coral cover inside the Al Wajh lagoon. These include, but are not limited to, water depth, wave power, suspended sediment concentration, the frequency and intensity of high energy storm events, the availability of antecedent platform and suitable substrate for larval settlement (for a comprehensive summary of environmental controls of coral distribution, see Done, 2011). These controls operate across a range of scales and while some are subject to local fluctuations that produce interrelationships, synergies and feedbacks, others (e.g. salinity) can be considered uniform across the extent of the study area and treated as constant terms. Of these variables, water depth, wave power and suspended sediment concentration were selected for the models because they have been suggested as determinants of coral community structure inside the Al Wajh lagoon (Sheppard et al. 1992; De Vantier 2000). They also exhibit variation at the scale of the study area and information on these variables can be derived at the landscape scale across the study area using GIS and remote sensing techniques.

A key aim of this study is to establish which type of regression modelling is most appropriate for explaining the distribution of live coral inside the Al Wajh lagoon. One model uses ordinary least squares regression while a second introduces a spatially lagged autoregressive term to build a spatial component into the model. The null hypothesis for these models is that none of the variables have any influence on the distribution of live coral cover inside the Al Wajh lagoon.

2 Methodology

2.1 Study Area

The Al Wajh Bank is situated along the north-eastern part of the Red Sea coastline of the Kingdom of Saudi Arabia (Figure 1), it is the most extensive of a series of reef platforms that comprise a reserve network designated by the National Commission for Wildlife Conservation and Development in 2000. The modelling exercise aimed to develop an understanding of the environmental controls on live coral distribution to inform reserve management. It was applied to a sub-area of interest at the northern end of the lagoon which traversed environmental gradients of water depth, suspended sediment concentration and wave exposure (see inset box on Figure 1).

The barrier reef system is comprised of a continuous line of reefs stretching for approximately 100 km and separated by several narrow (< 200 m width) channels. The outer edge of the bank lies approximately 26 km offshore and runs parallel to the shoreline for approximately 50 km before curving landward to enclose the reef system around a central lagoon (Fig. 1). The depth of the lagoon floor ranges from 30-60 m, becoming progressively shallower towards the coastline that comprises an alluvial sandy plain. The present living reefs, both along the barrier and inside the lagoon, have developed during the past 6000 years as Holocene sea levels have risen on top of topographic highs formed by earlier reef structures (Sheppard et al. 1992; De Vantier 2000). The shelf inside the barrier supports a range of islands and associated reef formations including platform or patch reefs, lagoon pinnacles, reticulate reef systems, submerged reef ridges and cay reefs.

[Figure 1 here]

2.2 Methodology

The methodological components of this study can be subdivided into two sections: *i.* the derivation and validation of variables using remote sensing techniques and *ii.* The construction and comparison of two different types of regression model (classic and spatially explicit) for live coral cover.

[Figure 2 here .]

2.3 The derivation and validation of variables using remote sensing techniques

2.3.1 Acquisition of airborne hyperspectral imagery

Hyperspectral data inside the Al Wajh barrier were acquired on 9th May 2008 using an AISA Eagle imaging sensor mounted to a Cessna seaplane. The AISA Eagle instrument measured 128 contiguous spectral bands from 400 to 994 nm at a spectral and spatial resolution of 5 nm and 1 m respectively. The image covered approximately 20 km^2 (1.5 km wide by 13 km in length) and was located along the northern coast of the inner Wajh Barrier.

2.4 Derivation of information on explanatory variables: Water depth, suspended sediment concentration and wave exposure

2.4.1 Water depth and suspended sediment concentration

An atmospheric correction was carried out on the raw hyperspectral imagery using the fast-line-of-sight atmospheric analysis of spectral hypercubes (FLAASH) module™ within the software environment for visualising images (ENVI) 4.5. Standard atmospheric water column amounts were calculated for a tropical atmosphere with a maritime aerosol model to represent the boundary layer above oceans, accounting for sea spray (Cooley et al. 2002).

A semi-analytical optimization model was used to simultaneously derive bathymetry, water optical properties and subsurface remote sensing reflectance from the atmospherically corrected hyperspectral image. The semi-analytical model algorithm was based on quasi-single-scattering theory (Gordon, 1994), and was implemented through a series of simulations that populated parameters to estimate subsurface remote sensing reflectance from surface remote sensing reflectance (Lee et al. 1998 and 1999). To perform the optimisation it was necessary to impose a series of constraints on input parameters, the derivation of which are outlined by Goodman et al. (2008), who describe the application of this approach to a coral reef environment (Table 1). For the purpose of this analysis, the model was applied to coral spectra to yield information on bottom albedo, the particle-backscattering coefficient (from which a measure of suspended sediment could be derived) and water depth.

Table 1. Constraints employed for optimisation of the semi-analytical inversion model, as defined in Goodman (2008).

2.42 Field measurement of water depth and suspended sediment

A dataset of 188 bathymetric readings across the study area was collected using a Norcross X single beam bathymetric sounder in conjunction with the water sampling for validating the output bathymetry from the semi-analytical model. The suspended sediment concentration (SSC) was measured *in-situ* by extracting 50 water samples of 200 mL volume from transects ran perpendicular to the coastline across the coastal shelf. Sample collection was timed to coincide with acquisition of the airborne remotely sensed imagery and extractions were taken from just below the wave base at a depth of 1 m using a length of piping with a pre-rinsed sample bottle attached to the end of it. The location of each sample was recorded using a dGPS (accuracy < 1 m).

Suspended sediment was measured from the field samples in a laboratory using filtration methods. For estimation of SSC across the study area, the dataset of fifty water samples was divided randomly so that 25 of the samples could be used to establish a simple power relationship between the particle backscattering coefficient (derived from the semi-analytical optimization modelling) and suspended sediment. This relationship was then used to predict suspended sediment concentration across the study area using ArcGIS Model Builder. The remaining 25 samples were used to test the accuracy of this relationship once it had been extrapolated over the study area. This validation proceeded by plotting the locations of the field samples taken and comparing suspended sediment measured in the laboratory with that modelled from the remote sensing image.

2.4.3 Wave Exposure model

To estimate wave exposure, the fetch-based method of Ekebom et al. (2003) was employed using linear wave theory to estimate incident power on the basis of fetch and wind power statistics, with bathymetric information incorporated to account for the influence of refraction and shoaling (for further details of method see Hamylton, 2011b). A 30 m grid was placed over the study area and the radiating lines extension tool in ArcView (Jenness 2006) was used to generate 8 lines of length 30 km, spaced 45 degrees apart, originating from each grid point. All fetch-limited lines (i.e., those intersecting an overlaid coastline shapefile) were trimmed at the point of intersection with the coastline. Polyline lengths were then calculated and input as fetch distances from each direction into the linear wave transform model.

Data on the speed and frequency of direction from which winds blew in the study area were extracted from the Indian Ocean volume of the *Marine Climatic Atlas of the World* (United States Navy 1995) for input into the linear wave transform model. This atlas reported wind speed and frequency data from a meteorological station at 10m above sea level approximately 50 km north of the study site located on Bahrein Island, Saudi Arabia (26^{o}16'N, 50^{o} 37'E). Data were averaged across a time period that spanned from January 1991-October 1995.

Fetch lengths and wind data were input into the significant wave height and wave period equations which were used to calculate wave energy from linear wave theory (Ekebom et al. 2003; Hamylton, 2011b). As the study area was inside an enclosed lagoon, fetch-limited equations were employed for each cardinal and subcardinal direction and summed to provide an overall measure of exposure at each grid point, which was then interpolated to a continuous surface of 1 m resolution.

2.5 Derivation of information on the Dependent variable: live coral cover

2.5.1 Field sampling of image spectra and coral community surveys

Field spectra of four benthic coverages (live coral, dead coral, macroalgae and sand) were collected for input to the spectral unmixing algorithm using a TRIOS™ Ramses ARC sensor. These coverages were representative of the community components falling inside the study area on habitat maps previously prepared by the Japanese International Cooperation Agency through interpretation of aerial photography. The spectrometer measured light in the wavelength range 300 - 920 nm, with an optical resolution of ~5 nm (Datentechnik GmbH, 2004). Underwater measurements were taken across an integration time of 63 ms with 50 replications collected for each benthic coverage within each of five sample sites. Average endmember spectra for each target were smoothed for the elimination of high frequency noise (Savitzky-Golay, 1964) and interpolated to yield reflectance at 1 nm intervals with a cubic spline (Karpouzli *et al.,* 2004).

Additional coral community records were collected in the form of six detailed 20 x 2 m phototransects established across a range of inshore - offshore and sheltered – exposed locations. This methodology yielded 20 photographs per transect line, i.e., 120 photographs overall, each of which were visually assessed for percentage of live coral cover using Coral Point Count with single random point specification (Kohler and Gill 2006).

2.5.2 Spectral unmixing of the hyperspectral imagery

A brief summary of the unmixing routine applied to the hyperspectral imagery is provided here as a detailed description has been published elsewhere (Hamylton, 2011a). Pre-processing steps included atmospheric and water column correction (see section 2.4.1), geometric correction and data subsetting via multiple discriminant function analysis. Multiple discriminant function analysis was applied to the collected field spectra to define an optimal subset of wavelengths for resolving benthic coverages and spectral unmixing was performed on this subset to decompose the reflectance of the materials with different spectral properties inside the ground field of view of a single pixel (1 x 1 m resolution) (Kruse et al. 1993). On the basis of the image reflectance for each pixel and the field collected spectra of the individual benthic coverages, the proportions of the individual elements falling inside each pixel were derived by solving a set of simultaneous linear equations. The linear mixture model assumed that, for a given wavelength, the total number of photons reflected from a single pixel and detected by the sensor was a linear function of the reflectance of the individual components and the fractional area of the pixels they cover:

$$r_x = \sum_{j=1}^{n} a_{xi} f_j + e_x \qquad \text{Equation 1}$$

where r_x = reflectance of a given pixel in the xth of z spectral bands

n = the number of mixture components

f_j = the jth fractional component in the makeup of r_x

a_{xj} = the reflectance of mixture component j in spectral band x

e_x = the difference between the pixel reflectance and that computed from the model.

Unmixing accuracy was assessed using a combination of the root mean square error model and comparison against the field data collected using the phototransects. The overall root mean square error was calculated as the difference between the reflectance measured by the sensor and that computed from the unmixing algorithm, this was averaged for each waveband independently. Comparisons against field data proceeded via a linear regression between the actual proportion (as estimated from the phototransect mosaic) and the estimated proportion (from spectral unmixing). The output image depicting the derived spatial patterns of abundance for live coral across the study area was treated as a representation of the response variable for input into the regression models.

2.6 The construction and comparison of classic and spatially explicit regression models for live coral cover.

The spatial structure of the coral coverage dataset was explored by converting the unmixed coral cover layer to a point file and computing the local Geary's C statistic as a measure of spatial autocorrelation between all pairs of points. A semivariogram was generated to determine the optimum sampling grid size at which there was no spatial dependence between the data points and therefore no internal redundancy. An exponential model was found to best fit the dataset with spatial dependency reaching a sill at 30 m distance between points. To represent coral cover (the dependent variable of the model), a 30m grid of coverage values was therefore extracted by overlaying a grid of points spaced 30 m apart and taking an average value of a 3 by 3 cell window from the unmixed coral cover layer (1m resolution) as input to the regression modelling exercise. Corresponding values were extracted for each of the explanatory variables (water depth, suspended sediment concentration and wave exposure) at each grid point location, each of which represented a data case.

Two regression procedures were carried out using the derived data cases in the software GeoDa (Anselin, 2003). These were ordinary least squares (classic) regression and spatially lagged autoregression. After confirmation that the raw data complied with the assumptions of regression, the two types of regression analysis were carried out in sequence to measure the proportion of variation in coral cover accounted for by each model. In the second regression model, spatial structure was included via the introduction of a spatially lagged autoregressive term as an explanatory variable. This approach drew explicitly on the location of each individual case through the construction of a spatial weights matrix ($w_{(i,j)}$) that expressed for each data case whether or not other cases belonged to its neighbourhood, such that $w_{ij}=1$ when i and j were neighbours and $w_{ij}=0$ otherwise (Anselin and Bera 1998). The values of the dependent variable at neighbouring locations were therefore introduced into the standard regression equation:

$$\mu_{(i)} = \beta_0 + \beta_1 X_{1(i)} + \beta_2 X_{2(i)} + \beta_3 X_{3(i)} + \rho \sum_{j \in N(i)} w(i,j) Y(j) + e_{(i)} \qquad i = 1, \ldots, n \qquad \text{Equation 2}$$

where n = the number of points or areas

$X_1 - X_3$ are the explanatory variables,

e(i) = independent, normally distributed error term

β_0 to β_k = coefficients estimated using the model.

ρ = a parameter associated with the interaction effect.

To estimate the spatial autoregressive terms in the spatial lag model, all cases and the spatial weights matrix were input into a maximum likelihood procedure that generated consistent estimates of ρ and β. A distinguishing feature of the likelihood for linear regression parameters with a spatial autoregressive component is a Jacobian term of the form $|I - \rho W|$, an evaluation of which was carried out based on the characteristic polynomial of the spatial weights matrix, W, to maximise the likelihood function of this term. This approach was originally suggested by (Ord 1975) and was developed into an efficient computer algorithm in the software GeoDa (Smirnov and Anselin 2001). After each regression analysis, diagnostics were recorded (including the Moran's I statistic, t-test, and measures of fit) and the spatial distribution of model residuals was mapped. A model building approach was taken whereby a range of independent variables were employed in the initial runs, with analysis of the t-statistic providing justification for retaining some variables and excluding others future runs. For example, both phytoplankton backscatter and dissolved organic matter were taken out of the model after initial runs as they did not make a statistically significant contribution to the performance of the model.

3 Results

3.1 Derivation of information on explanatory variables: Water depth, suspended sediment concentration and wave exposure

The bathymetric map revealed that water depths inside the study area ranged between 0.2m above reef patches and 30m inside the channel towards the northern end of the study area. These closely approximated the 188 values measured *in-situ* with a bathymetric sounder (R^2 0.95). In the broader context of the Al Wajh reef system, the deep channel towards the north of the study area leads to a large opening in the northern barrier wall, one of only two sites of water exchange between the lagoon and outside ocean waters. The shallower areas of the study site coincided with the platform in the north, the ridge network and the tops of the patches in the south.

Suspended sediment values measured from the water samples extracted inside the lagoon ranged between 5 and 73 mg L^{-1}. The distribution indicated that suspension of sediments coincided with shallower areas. The association between the particle backscatter coefficient estimated from the imagery and sediment content of the water was strong (R^2 0.91 based on the 25 samples).

The wave power model distribution was elevated over the ridge towards the north of the study area immediately below an opening in the Wajh Bank. The majority of the study area was fetch-limited, being surrounded by the Wajh Bank to the west and the mainland to the east. However, one small area in the north of the study area is non fetch-limited in a northerly direction. Power levels ranged between 2 and 699 Jm^{-3} throughout the study area.

3.2 Derivation of information on the Dependent variable: live coral cover

The two hundred and fifty reflectance spectra collected showed considerable variability between the spectra of the different benthic coverages, each of which had their own unique reflectance curve. The airborne dataset was reduced from 128 bands to 27 discriminant functions composed of reflectance and first order derivative spectra, as identified by the multiple discriminant function analysis. For the field sites where the coral community was sampled via phototransects, the cover of live coral ranged from 30-74% (Figure 3).

[Figure 3 here]

On the spectrally unmixed output coverage, white areas that indicated high coral cover coincided with coral that was visible on the three band pseudocolour image (Fig. 4a) and the overall root mean square error was low (<0.01). Estimates of live coral cover correlated strongly with field assessments (R^2 0.89) and were elevated in three general areas. Firstly, to the north of the study area around the periphery of the shallow bank (although not across the shallow top of this, an area which is exposed at low tide). Secondly, several prominent ridges of high live coral cover stood out among the network across the centre of the study site. Thirdly, areas of interspersed high coral cover were present in conjunction with the patches in the south of the study site.

[Figure 4 **here**]

3.3 The construction and comparison of classic and spatially explicit regression models for live coral cover.

All of the input variables were significant and the ordinary least squares and spatially lagged regression models explained 26 % and 76 % respectively of the variation in live coral cover inside the study area. For both models, water depth was negatively correlated and suspended sediment and wave exposure were positively correlated with live coral cover. Suspended sediment had the highest t-statistic in both cases, which was notably higher in the ordinary least squares model, with depth and wave power contributing less explanatory power. Nonetheless, the t-test values suggested that each variable was significant ($p<0.001$) and it follows that their contribution to the overall live coral coverage model was valuable, providing a statistical justification for rejecting the null hypothesis. The test for multicolinearity revealed minimal association between these distinct explanatory variables of the dataset. The residuals from the ordinary least squares regression model displayed strong positive spatial structure, which was corroborated by the Moran statistic (Table 2). For the spatially lagged model, weak negative autocorrelation was apparent.

Table 2 Summary of results and diagnostics for the two types of regression.

4 Discussion

The moderate T-statistic for water depth was not in agreement with other coral reef studies which identify this as a key determinant of coral cover (Done, 2011; Kleypas et al. 1999). This is perhaps because of its status as an indirect variable, or environmental proxy, in marine environments. Potential controlling variables for which depth could act as a surrogate include temperature, light availability and degree of atmospheric exposure. These may mask or altogether counteract each other by exerting opposing influences on live coral cover. Processes may also interact in a non-linear manner along a depth gradient to cancel each other out in terms of their effects on live coral coverage. For example, coral cover may be highest at a depth where the mechanical disturbance caused by wave interaction is moderate at an intermediate disturbance level (Aronson and Precht 1995). Such a pattern could not be captured in a regression model.

The concentration of suspended sediment explained the highest proportion of variation, with higher concentrations associated with greater proportions of live coral cover. Although the presence of sediments is generally an impediment to coral survival because of abrasion and smothering, they are less likely to stress corals when strong currents are present (Rogers, 1990). Fine material (<0.15 mm diameter) rarely settles in waters of velocity 25 cms^{-1}, rather it stays uniformly entrained throughout the fluid (Komar 1976). Wajh lagoon sediments (which were consistently found to be <0.15 mm in diameter) likely stay suspended in shallower water of elevated velocity at a concentration too low to impede photosynthesis. Furthermore, the association of food particulates that favour coral growth such as zooplankton and dissolved organic matter with suspended sediment might benefit heterotrophic corals that feed directly from the water column (Johannes et al. 1970).

Wave power explained the least amount of variation in live coral cover, likely because of a trade-off between the constructive and destructive influence of water movement on coral. While circulation replenishes food and oxygen provision and removes metabolic waste products (Birkeland 1996), it also presents a mechanical stress whereby shallow benthic communities must withstand the force of breaking waves to persist (Massel 1996).

In the presence of spatial dependence, the initial ordinary least squares model inflated the goodness of fit measure and underestimated the standard error, increasing the likelihood of a Type I error (Cliff and Ord 1981). Failure to include spatial autocorrelation in the specification meant that some of the effect due to interaction would have been allocated to the existing covariates, particularly those with a similar spatial structure to the response variable. Respecification to incorporate a neighbourhood context effect operating through a spatially lagged expression of the response variable itself allowed this to be addressed. This neighbourhood context effect might be underpinned by either ecological factors, such as coral community reproduction, geomorphological ones, such as the presence of antecedent platform. In the Red Sea, endogenous influences could include a relatively short planktonic life cycle phase of around 35-40 days (Rinkevich and Loya 1979) and structural support provided by the existing structure of primary reef framework (Goreau 1959). Over longer timescales this latter influence may be perpetuated by regional variability of eustatic sea level, which spreads alluvial material from adjacent mountain ranges smothering reef and encouraging contemporary corals to grow on the elevated platforms of their Pleistocene counterparts (Shaked and Genin, 2011; Hayward 1982). Scaling up to the interaction of multiple corals, ecological processes such as the spread of disease and competition for light are known to have a characteristic spatial structure (Fortin and Dale, 2005). The action of any of these influences would associate the presence of nearby live corals on the reef with existing live coral coverages, as demonstrated by the autoregressive model.

The study exemplifies the degree to which hyperspectral data can be manipulated to support spatially-explicit modelling in coral reef environments. Extended coverage of the electromagnetic spectrum underpinned much of the modelling process with different dimensions of this dataset to providing critical information on water depth, suspended sediment concentration and coral cover. Unmixing algorithms that treated the data as spectrally continuous yielded outputs at the *ratio* level of measurement (i.e. a continuous map of the percentage of live coral cover across the study area). This added versatility to the modelling process by extending the range of statistical techniques available for realising explanatory power through the model. The value of introducing a spatial component was demonstrated for a number of reasons, including *i*. identification of an appropriate sampling scale for model development, *ii*. use of spatially lagged information (i.e., from a neighbouring site) on the response variable itself to increase explanatory model power, and *iii*. detection of spatial dependence (autocorrelation) in the model. Nonetheless, each of the image processing steps from which the dependent and explanatory variables were derived (pre-procesing, inversion, unmixing etc.) introduced an element of uncertainty into the models applied. While validation and accuracy assessment exercises permitted comparison of model outputs with values observed *in-situ*, an awareness of the cumulative influence of uncertainty along the analysis chain, for example, error in inversion and unmixing closure, is important. The study presented here could profitably be improved by a further error propagation or sensitivity analysis (Schott 2007).

Conclusion

A key aim of this study was to establish which type of regression modelling is most appropriate for explaining the distribution of live coral inside the Al Wajh lagoon. To do so, it is useful to distinguish between determinants that reflect endogenous interaction between the sites and those that respond to some other exogenous variable. Assessing the relative contribution of effects caused by a reaction to external forces and effects that are a reaction to neighbouring individuals determines the appropriateness of the model specified. When external forces are the major influence, a classic ordinary least squares regression model is appropriate, whereas interactive effects suggest a need for a model with a spatially dependent covariance structure (Hamylton, 2011c; Cliff and Ord 1981). Transition to a model that incorporated spatial dependence was accompanied by a marked growth in explanatory power. The theoretical implication that follows is that neighbourhood interactions play a more important role than previously thought. This invites greater consideration of explanatory variables that reflect interaction between sites, providing a persuasive case for explicitly building geographical considerations into studies of coral distribution.

Acknowledgements

This work would not have been possible without the generous support of the Khaled bin Sultan Living Oceans Foundation. Dr Tom Spencer at the University of Cambridge is thanked for guidance on the manuscript and Dr Rainer Reuter, University of Oldenburg and the National Coral Reef Institute, Nova Southeastern University are thanked for assistance with fieldwork.

6.0 References

Almany G.R., Connolly S.R., Heath D.D., Hogan J.D., Jones G.P., McCook L.J., Mills M., Pressey R.L., Williamson D.H.. Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs. *Coral Reefs*, 2009; 28, 339–351.

Anselin L., Bera A. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah A, Giles DE (Eds), *Handbook of applied economic statistics*. New York: Marcel Dekker Press.

Anselin L. (2003) GeoDa^{TM} 0.9 User's Guide. Illinois: Centre for Spatially Integrated Science.

Aronson R.B., Precht W.F. (1995). Landscape patterns of reef coral diversity: A test of the intermediate disturbance hypothesis, *Journal of Experimental Biology and Ecology,* 192, 1-14.

Benfield S.L., Guzman H.M., Mair J.M., Young J.A.T. (2007). Mapping the distribution of coral reefs and associated sublittoral habitats in Pacific Panama: a comparison of optical satellite sensors and classification methodologies. *International Journal of Remote Sensing,* 28, 5047-5070.

Birkeland C. (1996). Life and death of coral reefs. New York: Chapman and Hall.

Brando, VE, Anstee, JM, Wettle, M, Dekker, AG, Phinn, SR & Roelfsema, C (2009) A physics based retrieval of quality assessment of bathymetry from suboptimal hyperspectral data*, Remote Sensing of Environment*, 113, 755-700.

Chorley R.J.E. (1972). Spatial analysis in geomorphology. London: Methuen Press.

Cliff A.D., Ord J.K. (1981). Spatial processes: models & applications. London: Pion.

Cooley T., Anderson G.P., Felde G.W., Hoke M.L., Ratkowski A.J, Chetwynd J.H., Gardner J.A., Adler-Golden S.M., Matthew M.W., Berk A., Bernstein L.S., Acharya P.K., Miller D., Lewis P (2002). FLAASH, A MODTRAN-4 based atmospheric correction algorithm, its application and validation. Proceedings of the Geoscience and Remote Sensing Symposium, 2002, Toronto, Canada, 3, 1414 – 1418.

Côté, IM, Reynolds, JD (2006) Coral Reef Conservation, Cambridge University Press, Cambridge

Datentechnik GmbH (2004). Ramses Hyperspectral Radiometer Manual Release 1.0,

De Vantier L. (2000). Final report on the study on coastal marine habitats and biological inventories in the northern part of the Red Sea coast in the Kingdom of Saudi Arabia. Report to the National Commission for Wildlife Conservation and Development, Riyadh, Saudi Arabia, Japan International Cooperation Agency.

Done, T. (2011) Corals: Environmental Controls on Growth. In: Hopley D (ed) Encyclopedia of Modern Coral Reefs. Springer, Dordrecht, The Netherlands, p 281.

Doxoran, D., Froidefond, J., Lavendar, S. & Castaing, P. (2002). Spectral signature of highly turbid waters; application with SPOT data to quantify suspended particulate matter concentrations. *Remote Sensing of Environment,* 81*,* 149-161.

Ekebom J., Laihonen P., Suominen T. (2003) A GIS-based step-wise procedure for assessing physical exposure in fragmented archipelagos. *Estuarine and Coastal Shelf Science,* 57*,* 887-898

Fortin M-J., Dale M.R.T. (2005). Spatial analysis: a guide for ecologists. Cambridge: Cambridge University Press.

Garza-Perez, J.R., Lehmann, A., Arias-Gonzalez, J.E. (2004). Spatial prediction of coral reef habitats: integrating ecology with spatial modelling and remote sensing, *Marine Ecology Progress Series,* 269, 141-152.

Gordon, H.R. (1994). Modeling and simulating radiative transfer in the ocean. Chapter 1 in *Ocean Optics*, R.W. Spinrad, K.L. Carder and M.J. Perry (Eds), Oxford: Oxford University Press.

Goodman, J & Ustin, S.L. (2007) Classification of benthic composition in a coral reef environment using spectral unmixing, *Journal of Applied Remote Sensing,* 1, 011-501.

Goodman, J., Lee Z.P. & Ustin, S. (2008). Influence of atmospheric and sea-surface corrections on retrieval of bottom depth and reflectance using a semi-analytical model: a case study in Kaneohe Bay, Hawaii. *Applied Optics,* 47, 1-11

Goreau T.F. (1959). The ecology of Jamaican coral reefs. I. Species composition and zonation. *Ecology* 40, 67-90

Hamylton, S. (2011a). Estimating the coverage of coral reef benthic communities in the Al Wajh, Red Sea from airborne hyperspectral remote sensing data: Multiple discriminant function analysis and linear spectral unmixing. *International Journal of Remote Sensing,* 32: 9673 - 9690.

Hamylton, S. (2011b). The use of remote sensing and linear wave theory to model local wave energy around Alphonse Atoll, Seychelles. *Estuarine, Coastal and Shelf Science,* 95, 349-358

Hamylton, S (2011c) The Al Wajh Bank reef system, Saudi Arabia, Red Sea, pp. 28 – 37. In: Harris, P & Baker, E (Eds.) (2009) Atlas of Seafloor Geomorphology as Habitat, Elsevier, Chatswood, Australia. 600pp.

Haining, R. (2003) Spatial Data Analysis: Theory and practice. Cambridge: Cambridge University Press.

Harborne A.R., Mumby P.J., Zychaluk K., Hedley J.D., Blackwell P.G. (2006). Modeling the beta diversity of coral reefs. *Ecology* 87, 2871-2881.

Hayward A.B. (1982). Coral reefs in a classic sedimentary environment: Fossil (Miocene, SW Turkey) and Modern (Recent, Red Sea) analogues. *Coral Reefs,* 1, 109-114.

Hedley, J, Roelfsema, C, & Phinn, SR (2009_) Efficient radiative transfer model inversion for remote sensing applications. *Remote Sensing of Environment*, 113, 2527 – 2532.

Jenness J. (2006). Radiating Lines and Points (rad_lines.avx) Extension for ArcView 3.x, v. 1.1. Jenness Enterprises.

Johannes R.E., Coles S.L., Kuenzel N.T. (1970). The role of zooplankton in the nutrition of some scleractinian corals. *Limnologocal Oceanography* 15, 579-586

Karpouzli E., Malthus, T.J.M. & Place C.J. (2004). Hyperspectral discrimination of coral reef benthic communities in the western Caribbean. *Coral Reefs* 24, 124-134.

Kleypas J.A., Buddemeier R.A., Archer D., Gattuso J.P., Langdon C., Opdyke B.N. (1999). Geochemical consequences of increased atmospheric carbon dioxide on coral reefs. *Science* 284, 118–120.

Klonowski, W.M., Fearns, CS and Lynch, MJ (2007) Retrieving key benthic cover types and bathymetry from hyperspectral imagery, *Journal of Applied Remote Sensing,* 1, 11 – 50.

Kohler, K.E. and Gill, S.M. (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 and Geosciences*, 32, 1259-1269

Komar P.D. (1976). Beach processes and sedimentation. New Jersey: Prentice-Hall.

Kruse F.A., Lefkoff A.B., Boardman J.B., Heidebrecht K.B., Shapiro A.T., Barloon P.J., Goetz A.F.H. (1993). The Spectral Image Processing System (SIPS) - Interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment,* 44, 145–163.

Lee, Z. K. Carder, C. D. Mobley, R. Steward, & Patch, J. (1998). Hyperspectral Remote Sensing for Shallow Waters. 1. A semi-analytical model. *Applied Optics,* 37, 6329-6338.

Lee, Z. K. Carder, C. D. Mobley, R. Steward, & Patch, J. (1999). Hyperspectral Remote Sensing for Shallow Waters. 2. Deriving Bottom Depths and Water Properties by Optimization. *Applied Optics,* 38, 3831-3843.
LeDrew E.F., Holden H., Wulder M.A., Derksen C., Newman C. (2004). A spatial statistical operator applied to multidate satellite imagery for identification of coral reef stress. *Remote Sensing of Environment,* 91, 271-279.
Massel S.R. (1996). Ocean surface waves: their physics and prediction. , Singapore and London: World Scientific.
Ord J.K. (1975). Estimation methods for models of spatial interaction. *Journal of American Statistical Association,* 70, 120-126.
Phinn S.R., Neil, D.T., Joyce K.E. (2000). Coral Reefs: A multi-scale approach to monitoring their composition and dynamics. 30th International Symposium on Remote Sensing of Environment, Honolulu Sheraton, Hawai'I, 2672-2674.
Purkis S.J., Kohler K.E, Riegl B.M., Rohmann S.E. (2007). The statistics of natural shapes in modern coral reef landscapes. *Journal of Geology,* 115, 493–508.
Rinkevich B., Loya Y. (1979). The reproduction of the Red Sea coral *Stylophora pistillata*. 1. Gonads and planulae. *Marine Ecology Progress Series*, 1, 133-144.
Rogers C.S. (1990). Responses of coral reefs and reef organisms to sedimentation. *Marine Ecology Progress Series,* 62, 185-202.
Savitzky, A. & Golay, M.J.E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry,* 36*,* 1627–1639.
Schott, J.R. (2007) Remote sensing: The image chain approach, Second Edition. Oxford: Oxford University Press.
Shaked, Y., Genin, A. (2011) Red Sea and Gulf of Aqaba. In: Hopley D (ed) Encyclopedia of Modern Coral Reefs. Springer, Dordrecht, The Netherlands, p 839.
Sheppard, C. R. C. (1982) Coral Populations on Reef Slopes and Their Major Controls *Marine Ecology Progress Series*, 7, 83-115.
Sheppard C.R.C., Price A., Roberts C. (1992). Marine ecology of the Arabian region: patterns and processes in extreme tropical environments. London: Academic Press.
Smirnov O., Anselin L. (2001). Fast maximum likelihood estimation of very large spatial autoregressive models: A characteristic polynomial approach. *Computational Statistics and Data Analysis* 35, 301–319.
Sobel J, Dahlgren C. 2004. *Marine reserves: a guide to science, design and use*. Island Press: Washington DC.
US Navy (1995). U.S. Navy marine climatic atlas of the world. Volume III: Indian Ocean. Published by direction of the Chief of Naval Operations. U.S. Government Printing Office, Washington, DC
Veron J.E.N., Pichon M. (1976). Scleractinia of Eastern Australia. Part I. Families Thamnasteriidae, Astrocoeniidae, Pocilloporidae. *Australian Institute of Marine Science Monograph Series,* 1, 1-86
Wettle, M & Brando, V.E. (2006) SAMBUCA – Semi-Analytical model for Bathymetry, Unmixing and Concentration assessment, CSIRO Land and Water Science Report No. 22/06. CSIRO.

Figure list

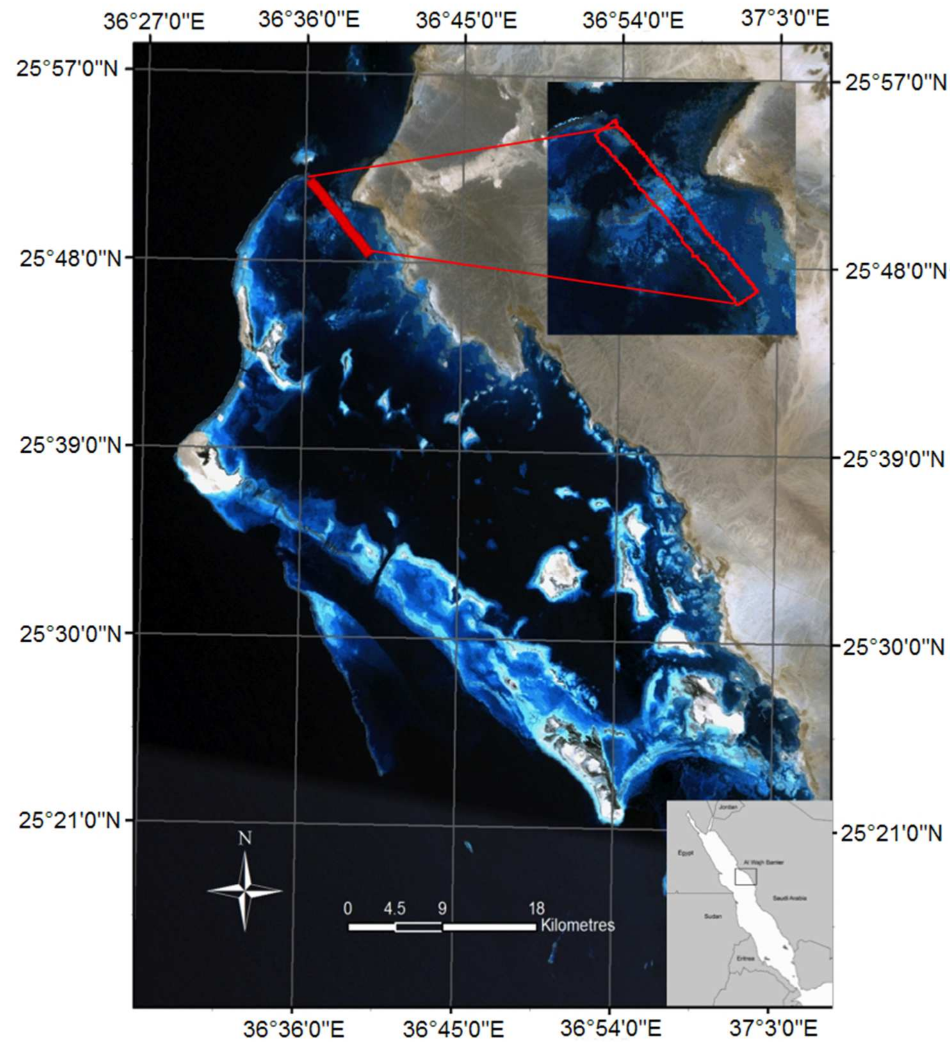
distribution of the modelled values for the three explanatory variables: (c) Bathymetry, (d) Wave power, and (e) Suspended sediment concentration.

Table 1. Constraints employed for optimisation of the semi-analytical inversion model, as defined in Goodman (2008).

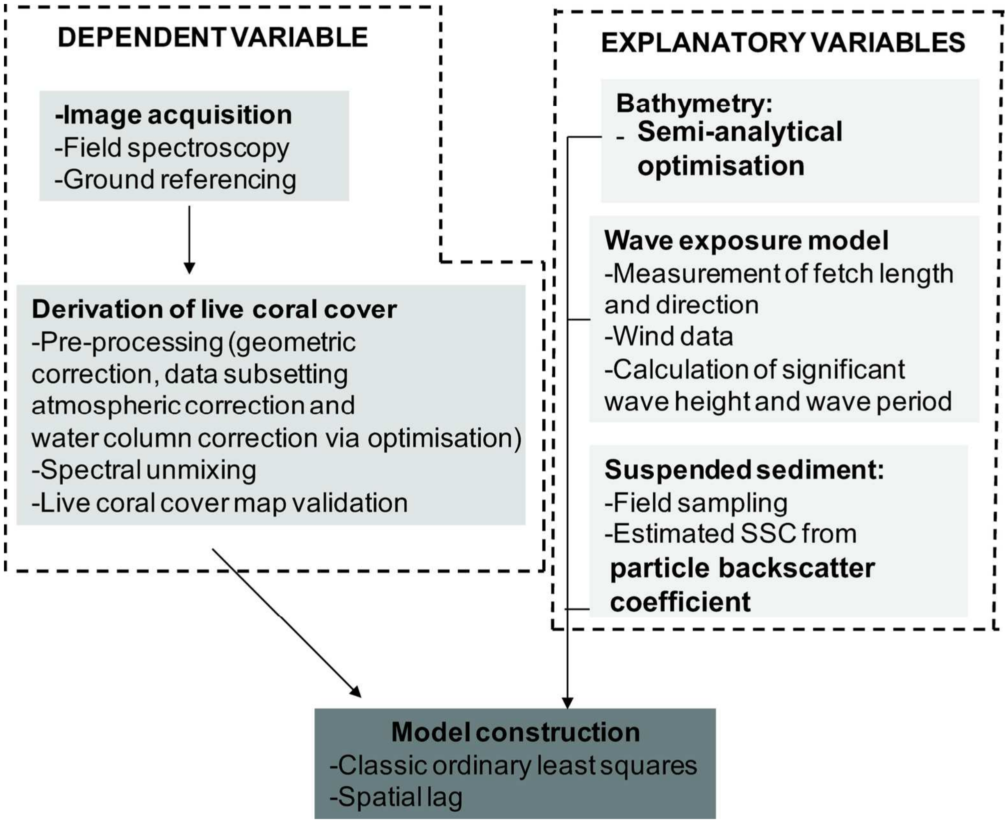
Parameter	Constraint
P (m^{-1}) is the phytoplankton absorption coefficient at 440 nm	$0.005 \leq P \leq 0.5$
G (m^{-1}) = absorption coefficient for gelbstoff and detritus at 440 nm	$0.002 \leq G \leq 3.5$
BP (m) particle-backscattering coefficient	$0.001 \leq BP \leq 0.5$
B is the bottom albedo at 550 nm	$0.01 \leq B \leq 0.6$
H (m) is the bottom depth	$0.2 \leq H \leq 33.0$

Table 2 Summary of results and diagnostics for the two types of regression.

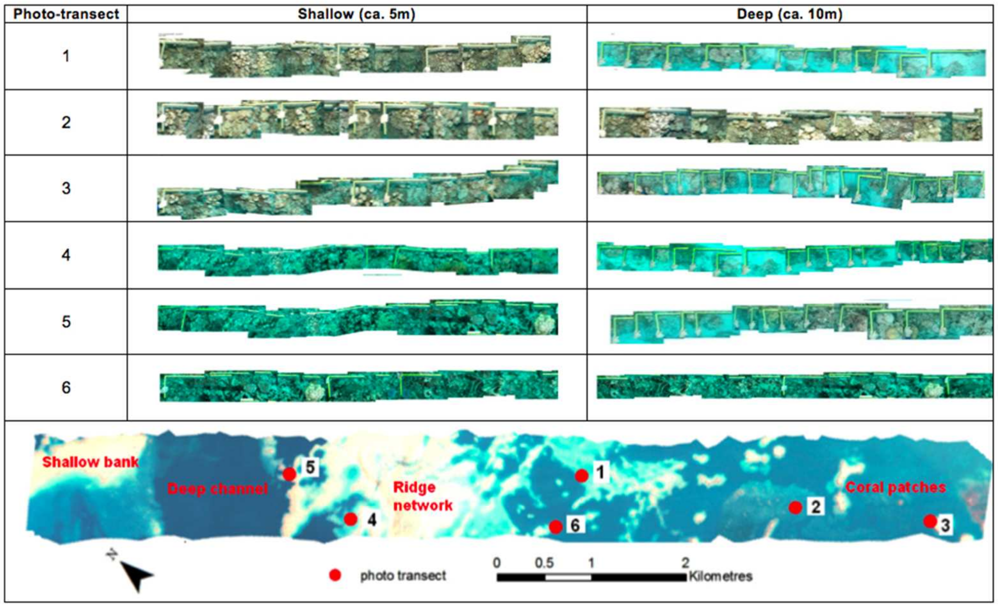
CLASSIC ORDINARY LEAST SQUARES REGRESSION			
R^2 (adjusted value)		0.27 (0.26)	
Moran's I of residuals		0.73	
Variable	β Coefficient	Standard error	t-statistic
Depth	-0.55	0.03	-12.31 ($p<0.001$)
Suspended sediment	0.96	0.03	27.78 ($p<0.001$)
Wave power	0.03	0.02	14.74 ($p<0.001$)
SPATIAL MODEL			
R^2		0.76	
Moran's I of residuals		-0.14	
Variable	β Coefficient	Standard error	t-statistic
Depth	-0.088	0.02	-3.67 ($p<0.001$)
Suspended sediment	0.168	0.02	8.80 ($p<0.001$)
Wave power	0.040	0.01	4.15 ($p<0.001$)



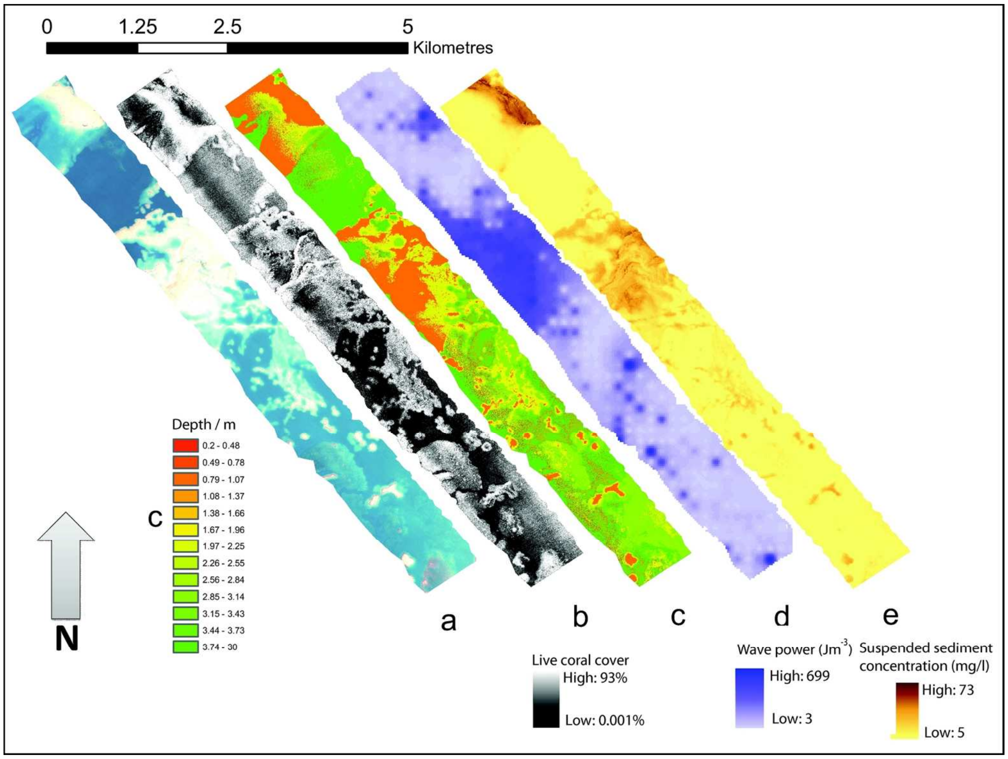
Landsat TM image of the Al Wajh Bank, Saudi Arabia, Red Sea (25°39′N, 34°45′E) and the location of the study site (upper inset) and the Al Wajh Bank in the Red Sea (lower inset).
225x254mm (96 x 96 DPI)



Schematic overview of the construction process for the live coral cover model at Al Wajh.
162x132mm (300 x 300 DPI)



Phototransects used for validating benthic estimations derived from the spectral unmixing algorithm, one shallow and one deep transect per site. Locations plotted on the RGB image composite of the study area
322x195mm (72 x 72 DPI)



a) Hyperspectral colour composite imagery of the study area (RGB wavebands at 767, 519 and 403nm) b) Gray scale unmixed image output depicting the abundance of coral, white areas indicate areas of high coral cover, c-e) Spatial distribution of the modelled values for the three explanatory variables: iii. Bathymetry, iv. Wave power, and v. Suspended sediment concentration.
157x118mm (300 x 300 DPI)