Modelling the spatial distribution of three marine fish species in the southern Benguela

Abstract

Understanding the spatial distribution of species in relationship to climatic and environmental variables is key to conservation and management of important species, as their distribution might change under climate change and variability. Based on presence absence data from scientific trawl surveys, this study used Generalized Additive Model (GAM) and Krigging with External Drift (KED) statistical techniques to determine the spatial distribution of three marine fish species of commercial interest: *Merluccius capensis*, *Merluccius paradoxus*, and *Thyrsites atun*, on the West and South coasts of South Africa. The modelled distributions reflect the previously determined range and habitats of the two species of hake and are in accordance with the common knowledge on the biology of the two species. Presence-absence modelling found depth to be the main factor for explaining hake distribution on both coasts. For the West coast an interaction between sea surface temperature and chlorophyll-a combined with depth as a factor was found to provide the best model. On the South coast depth was the only factor retained. The models for *M. capensis* and *M. paradoxus* are potentially useful in mapping and determining future distributions based on environmental factors. The model obtained for the spatial distribution of *T. atun* has a lower explanatory power than those of the two hake species.

Chris Wilkes

Honours Thesis, University of Cape Town, 2013

Email: WLKCHR007@myuct.ac.za

Supervisors: Yunne-jai Shin and Dawit Yemane

Introduction

Global climate change has the potential to alter the spatial distribution of species by affecting and altering key environmental drivers of fish population dynamics (Walther *et al*, 2002). A warmer ocean might result in spatial shifts and contractions for many species, potentially leading to the extinction of some vulnerable species that live at the 'edge' of the thermal plane, such as corals (Rahmstorf, 2002; Meehl *et* al, 2007). The need to understand how species react to environmental changes is critical in understanding how to preserve and protect species and ecosystems.

Mapping and determining the distribution of plants and animals has long been a central tenet of ecology (Guisan and Zimmerman, 2000). Understanding where a species occurs is a fundamental part of understanding how a species interacts with the ecological system around it. By understanding this ecosystem interaction conservation management is given key information for protecting and promoting the health of ecosystems (Pearson and Dawson, 2003).Technological advancements have allowed ecologists and biologists to determine potential suitable habitats and probable occurrence through the use of statistical modelling (Guisan *et al*, 2002). The era of 'big data' means that ecologists and biologists no longer need to solely rely on determining the mechanistic relationship between a species' distribution and the various factors that determine what type of habitat is actually suitable; an often time- and resource- consuming endeavour (Guisan and Zimmerman, 2000, Robinson et al, 2011).

The modelling of species' environmental niches, which links species spatial distribution to multiple environmental variables, can provide insights into the effects of climate change on past and future distributions of species (Guisan and Zimmerman, 2000, Araujo and Guisan, 2006). This could prove useful in determining which species are more sensitive to the effects of climate change and provide rational information to conservationists and fisheries managers. The approach can for example directly support management decision in establishment of marine protected areas

The Benguela and Agulhas current systems envelope southern Africa. The upwelling system of the Benguela current is particularly productive, whilst the Agulhas system has a higher level of biodiversity (Shannon and Nelson, 1988; Wardell-Johnson, 2000; Lutjeharms, 2006, Griffiths et al, 2010). The Agulhas Current in particular has undergone rapid warming in the last 30 years due to climate change (Rouault *et al*, 2009). This warming has implications for the diversity and productivity in these two systems. The impact on fish populations could be due to several factors; for example, ocean warming could weaken the current systems which many species rely on for the advection of eggs and larvae to favourable nursery grounds (Hutchings *et al*, 2002; Rouault *et al*, 2009).

The distribution of many species is constrained by oceanic currents like the Benguela current (Griffiths *et al*, 2005). Other environmental variables like sea surface temperature (SST) and chlorophyl-a (chl-a), which can be used as an indication of primary production, are highly dependent on wind stress and ocean currents (Enriquez and Friehe, 1995; Rykaczewski and Checkley, 2008). With the potential for wind and currents to be affected by climate change it is important to understand how these factors affect the distribution of marine species (Walther *et al*, 2002; Edwards and Richardson, 2004). Range shifts and the expansion or contraction of ranges of key ecosystem species could have important consequences on the functioning of the ecosystem as a whole (Perry et al, 2005).

Changes in wind speed in the Benguela upwelling zone could change the amount of nutrient-rich water brought to the surface and thus hamper productivity (Meehl *et* al, 2007). These localized effects of global climate change mean that it is important to understand the factors that influence the distribution and range of the species that inhabit these waters (Perry et al, 2005). Statistical environmental niche models allow the relationship between fish spatial distribution and environment to be quantified without determining the mechanistic relationship that exists between species distribution and environmental variables (Guisan and Zimmerman, 2000). We propose to use this approach to determine the environmental niche of three fish species that are both ecologically and commercially important in the southern Benguela ecosystem: the two species of Cape hakes, the shallow-water *Merluccius capensis*, the deep-water *Merluccius paradoxus*, and snoek, *Thyrsites atun*.

The two species of Cape hake differ only slightly in their morphology, such as differing number of vertebrae, the colouration of the gill rakers and the shape of the pectoral fins and otoliths (Smith, 2003). The geographical ranges of Cape hakes overlap to some extent but there are generally notable differences in their spatial distribution. *M. capensis* is found up to 400 m depth and *M. paradoxus* occurs between 150m – 900 m. *M. paradoxus* is more abundant on the West coast of southern Africa. Bathymetry is hypothesized to be the main reason for the difference in distribution between these two very similar species (Botha, 1985; Payne et al, 1987; Payne and Punt, 1995).

The West coast has a larger shelf area between 200 m and 600 m. The West coast and South coast species are thought to constitute the same stocks rather than separate stocks as was previously believed (Payne et al, 1987).There has also been some speculation that the hakes are migratory. If the two hakes each constitutes one stock around southern Africa rather than two, any shift in the distribution of the hake due to environmental changes need to be carefully measured (Burmeister, 2005). Due to its commercial value and previous overexploitation, the hake fishery is strictly controlled with total allowable catches (TAC) each year (Griffiths, 2000). This makes the understanding of the spatial distribution of the hakes a very important source of information in conservation and sustainable exploitation.

The diets of both species of Cape hakes consist of small crustaceans, fish, cephalopods (Botha, 1980). The importance of each prey varies seasonally as food availability changes. *M. capensis* is also known to predate heavily on juvenile *M. paradoxus*. This was long thought to be intraspecific cannibalism until the two species were identified. *M. capensis* can travel down to deeper depths where it can feed on the juvenile *M. paradoxus (*Payne and Punt, 1995). The Cape hakes are an important mid-level predator for the Benguela ecosystem (Payne et al, 1987).

Thyrsites atun, commonly known as snoek, is an important commercial and ecological component of the Benguela and Agulhas ecosystems. It has a pan-global distribution in the southern hemisphere and forms an important part of the commercial fishery of many countries, including South Africa, Australia and Chile. It occurs from the surface of the ocean to a depth of roughly 500 m. It is a fairly large pelagic predator growing up to 200 cm in length. It feeds on a wide variety of smaller fishes, as well as small crustaceans and cephalopods. Its movements are nomadic and are thought to be dictated by the availability of food sources (Griffiths, 2002).

This study will aim to explore the relationship between the two species of Cape hake, shallow-water and deep-water, and snoek with environmental variables. This approach aims to illuminate the effect that climatic and physical factors have on the range and distribution of these three important species. This knowledge can contribute to sound and sustainable management practices for these exploited species within a context of global change.

Methods

Study area

This study deals with the modelling of the spatial distribution of the two hake species (*Merluccius capensis* and *Merluccius paradoxus*) and snoek (*Thyrsites atun*) off the coast of South Africa (Figure 1, map of the study area).The bathymetry of the West Coast is characterised by a large fairly shallow shelf area, known as the Orange Shelf, which extends at depth of between 200 m and 400 m from the coast to a distance of approximately 200 km where it sharply drops off to deeper depths. The northern area of the South Coast has a much narrower continental shelf area some 50 km to 100 km at the most, notwithstanding the Agulhas Bank which is a large shelf area to the south of Cape Agulhas. The warm Agulhas current sweeps past the south coast and many jets and smaller offshoot currents are formed in this area which flow in a generally westward direction (Lutjeharms, 2006). The West Coast is characterised by the Benguela upwelling system which brings cold, nutrient-rich bottom

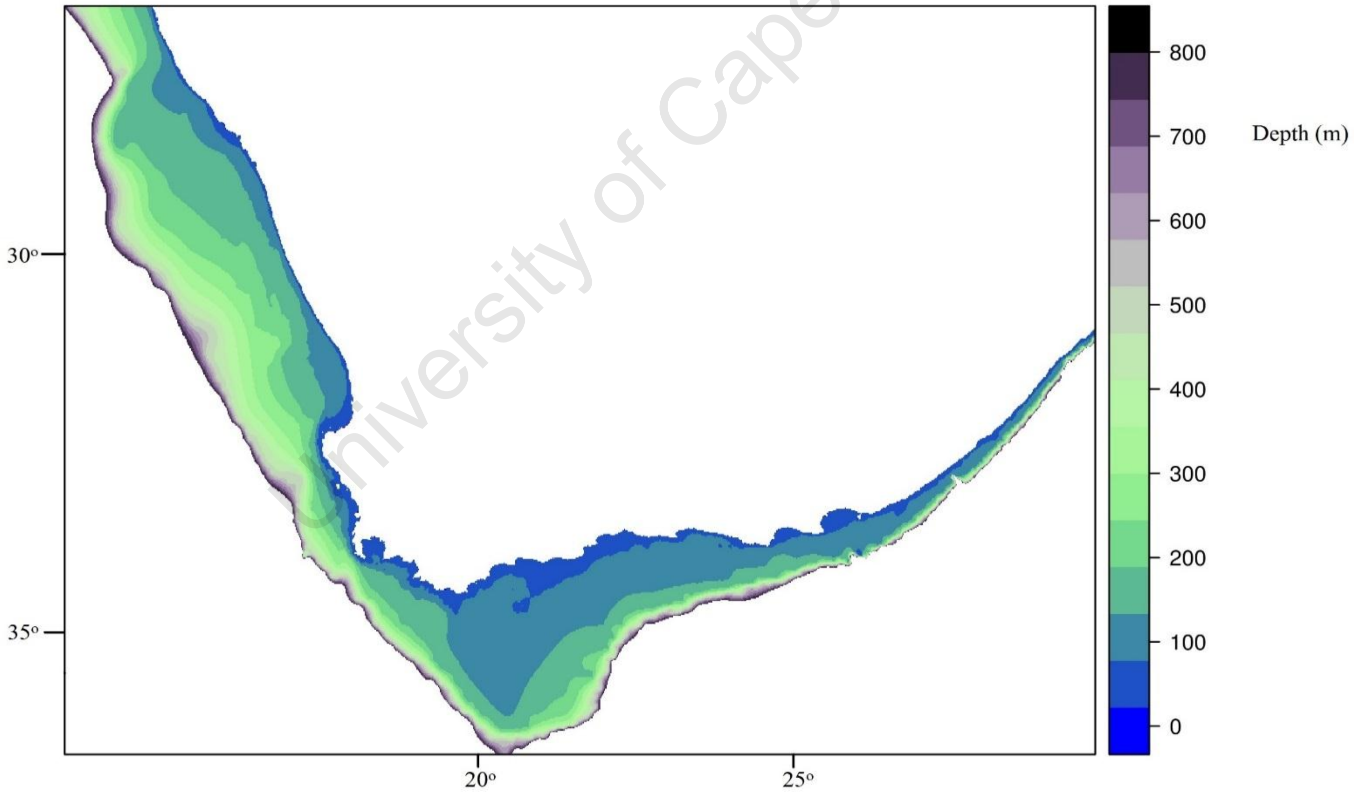


Figure 1. South African coastline. The study area where demersal trawls were conducted. The scale-bar shows the depth in meters.

water to the coastal region. This allows for a high level of primary production and ultimately a highly productive marine ecosystem (Shannon and Nelson, 1996).

Data source

The data used in this study come from the routine demersal trawl sampling off both the West and South coast of South Africa. The data for this analysis were obtained from surveys conducted by Marine and Coastal Management (MCM), a branch of the Department of Environmental Affairs and Tourism (DEAT). The surveys were conducted twice a year on the West and South coasts of South Africa.

The survey on the West coast is conducted during the summer season whereas that on the South coast is conducted during the autumn season. Trawl locations were sampled in a semi-random manner using 5x5 minute grid cells, with survey lines randomly distributed along the shelf in the depth range 30-500 m. The duration of the trawls was standardized to 30 minutes in all surveys, with a towing speed over the bottom of approximately three knots (1 knot = 1.852 km h^{-1}). In some instances, due to unfavourable seafloor conditions (i.e. unsuitable for demersal trawling) or exceptionally large catches, tows were hauled in before the 30 min duration. In these cases the catch was standardized to 30 min tow duration.

The trawl equipment used is as follows: a 55-m German otter trawl with 75 mm mesh cod-end fitted with a 35-mm mesh liner, a rope-wrapped chain footrope and 1,500-kg WV otterboards. Mouth width of the trawl is assumed to be constant at 26 m. For this study the catch rate is defined as the biomass caught per standard trawl.

The environmental variables used are satellite-derived SST and Chl-a, obtained from the Department of Environment Affairs and Tourism (DEAT). These variables were chosen because they are often a good indication of primary productivity (Morin et al, 1999). Data were only available for the time period 1999-2011with no data for 2000 and 2001. This is reflected in the modelled maps.

Data Analysis

Modelling distribution

There are three common approaches that are used to understand and determine the spatial distribution of fish populations: coupled bio-physical models, experimental ecology coupled with mechanistic models and statistical correlative models.

Bio-physical models coupled to hydrodynamic model represent the life cycle of the populations and explicitly models the impact of environmental variables on the population via different paths (effect on recruitment, effect on availability, timing and type of food resource, direct effect on physiological rates). These models can then be used to understand the influence of different environmental variables in shaping the current distribution of the population and how future changes in the environmental condition translate to changes in the distribution;

Experimental ecology coupled with mechanistic models rely on the determination of optimal environmental conditions from experiments and incorporation of this information into mechanistic models to simulate the current and future spatial distribution of fish.

Statistical correlative models show the underlying correlation between the observed distribution of populations and the prevailing environmental conditions, as determined by sets of environmental variables. These models are used to study potential changes in distribution with changes in the environment.

Of the three approaches, the most commonly used approach is correlative statistical modelling. There are a number of statistical models that fall under this approach. These include Generalized Linear Model (GLM), Generalized Additive Model (GAM), Classification Tree Analysis (CTA), Generalized Boosted Regression (GBR), Random Forest (RF), Artificial Neural Network (ANN) among others. For the purpose of this study the distribution of the three species is modelled using a GAM based on presence/absence data. In GAM the predictors are assumed to affect the response variable through additive and unrestrictive smooth functions of the predictors and their interaction (Hastie and Tibshirani 1990). GAMs generally assume that the mean response (μ) is related to the p predictor variables ($x_1, \dots, x_p$) by the following relationship:

$$g(\mu) = \alpha + \sum_{j=1}^{P} f_j \left(X_j\right) \qquad \text{Eq.1}$$

where $g(\mu)$ is the link function, in this case the logit link function. This is done to transform the response variable in order to linearize the relationship between the response and sets of predictors. Doing this also ensures that the predicted probability of occurrence will always fall between 0 and 1. The logit link function links the mean μ of the random variable y, to the additive function of the predictors and defines the relationship between the response and the additive predictor; α is the intercept term, f_j is the unspecified smoothing function, and the logit link is:

$$logit(P) = log\left(\frac{P}{1-P}\right) \qquad \text{Eq.2}$$

The probability of occurrence is then calculated as follows:

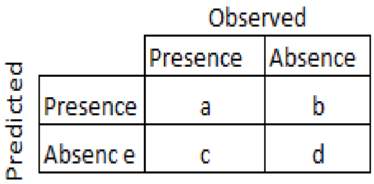
$$P = \frac{e^{g(x)}}{1+e^{g(x)}} \qquad \text{Eq.3}$$

Sets of models ranging in complexity from the simplest to the most complex were compared. The variables used were depth (meters), sea surface temperature (^{o}C), and chlorophyll-a (mg.l^{-1}). The simplest models only used one variable to predict distribution; the most complex used an interaction between all three as a basis for prediction. The model with the lowest Akaike Information Criterion (AIC) score was selected as the best performing model.

Probability thresholds of presence can be selected arbitrarily (usually probability of 0.5) but could also be selected by optimizing certain objectives (overall accuracy, higher value of presence, higher value of absence, etc.). In this study thresholds were selected by optimizing the higher overall accuracy of the model (i.e. probability thresholds that maximize the matches between predicted presence-absence and observed presence-absence).

It is important to assess the performance of the species distribution model applied to predict presence and absence of the population modelled. There are few commonly used model performance statistics for classification type models these include: Area Under the Curve (AUC) of the Receiver Operating Characteristics (ROC) curves, Kappa, True Skills Statistics (TSS), overall Accuracy, Sensitivity, and Specificity. In this study model performance was assessed using the AUC (also known as threshold independent performance criteria as it is computed across ranges of probability of occurrence), Kappa and TSS. Both Kappa and TSS were computed at thresholds optimized to maximize the overall accuracy of the model. The performance statistics are computed based on the confusion matrix (also referred to as error matrix) (see table 1.).

Table 1. Confusion matrix used to evaluate the performance of the logistic GAM in predicting the distribution of the three species considered in this study. *a* is the number of trawl stations where the model correctly predicted presence, *b* is the number of stations where the species was absent but the model predicted presence, *c* is the number of stations where the species was present but the model predicted absence, *d* is the number of stations where the species was not found and the model correctly predicted absence.



		Observed	
		Presence	Absence
Predicted	Presence	a	b
	Absence	c	d

$$Sensitivity = \frac{a}{a + c}$$

$$Specificity = \frac{d}{b + d}$$

$$TSS = Sensitivity + Specificity - 1$$

$$Kappa = \frac{\left(\frac{a+d}{n}\right) - \frac{(a+b)(a+c)+(c+d)(d+b)}{n^2}}{1 - \frac{(a+b)(a+c)+(c+d)(d+b)}{n^2}}$$

$n = (a + b + c + d)$ is the total number of observations.

Modelling spatial distribution of catch rates

In the previous section the distribution of the three species was modelled. In this section spatial distribution of catch rate was modelled.

In addition to presence-absence modelling, density of the three fish species was modelled using GAM and geostatistical model universal kriging, also known as Krigging with External Drift (KED) These two methods used catch abundance or density rather than presence-absence as in the previous model. KED, also known as Universal Krigging, was developed to model large-scale trends in spatial data (Fortin and Dale, 2005). Krigging determines the best combination of weights to interpolate values for the unsampled locations, by minimizing variance derived from the spatial covariance in the data according to linear regressions.

The variability of spatially referenced observations $Z(s)$ at spatial location s is modelled as a sum of trend and residual components and takes the following form:

$$Z(s) = \sum_{j=0}^{p} X_j(s)\beta_j + e(s) \qquad \text{Eq. 4}$$

where $X_j(s)$, is the predictor variable j, p is the number of predictor variables, β_0 is usually the intercept term and $X_0(s) \equiv 1$, β is a vector of unknown regression coefficients, $e(s)$ is the vector of residuals (Pebesma, 2004). The best unbiased linear prediction of a response variable in a location $Z(S_0)$ is given by:

$$\hat{Z}(s_0) = x(S_0)\hat{\beta} + \acute{v}V^{-1}\left(Z(s) - X\hat{\beta}\right) \qquad \text{Eq. 5}$$

where $x(S_0)$ is the row of the predictors X that correspond to $Z(S_0)$, with $\hat{\beta}$ the unbiased estimate of the linear coefficient which is obtained, using generalized least square estimates, as follows:

$$\hat{\beta} = \left(\acute{X}V^{-1}X\right)\acute{X}V^{-1}Z(s) \qquad \text{Eq. 6}$$

with v' is formulated as

$$v = \left(cov\left(Z(S_0), Z(S_1)\right), \ldots, cov\left(Z(S_0), Z(S_n)\right)\right)' \qquad \text{Eq. 7}$$

and $cov(.,.)$ representing the covariance.

The covariance is modelled by fitting a theoretical covariance function which takes one of the various set of theoretical functions (spherical, exponential, nugget, power, gaussian, linear, etc.). The fitted theoretical variograms can include one or the sum of many of basic variogram models, each of which can have its own zonal or geometric anisotropy parameters. To solve this system of linear algebraic equations, the sum of the weights is constrained to equal 1 so that there are more equations than unknown parameters to estimate (Fortin and Dale, 2005).

The specific form of the GAM model applied to model distribution of catch rates.

$$CatchRate(C) = s(depth) + s(longitude, latitude) + e \qquad \text{Eq. 8}$$

Where the *s(depth)* refers to effect of depth modelled by the spline smoother *s*, *s(longitude, latitude)* refers to the effect of geographic location modelled as smoothed surface and e is the error term which is assumed to be normally distributed.

The performance, mainly predictive performance, of the geostatistical model using KED and GAM was compared based on three metrics that are based on the residual (the mean prediction error/mean error, root mean squared prediction error, and inter-quartile range). These metrics were obtained after conducting Leave One Out Cross Validation (LOOCV) for both the KED and GAM. Generally model with smaller values in all of the above metrics is preferred. LOOCV was performed for each technique. That is, for each time period (including the overall study period) for all the three countries, each station was randomly left out of the data set prior to estimation, and the predicted estimate of species richness subsequently obtained for the location of the omitted station was compared with the true value for the station. Three performance statistics were used to cross-compare the residuals (true value - estimated value) obtained by applying LOOCV for each technique, namely the root mean squared error (RMSE) of the residuals, the inter-quartile range (IQR) and the mean of the residuals (ME). For the first two statistics, low values are desirable and indicate greater reliability, while for ME, values as close as possible to 0 are desirable.

The three performance statistics from the LOOCV were computed as follows:

$$ME = \frac{1}{k}\sum_{i=1}^{k}\left[\hat{Z}(S_k) - \acute{Z}(S_k)\right] \qquad \text{Eq. 9}$$

$$RMSE = \sqrt{\frac{1}{k}\sum_{i=1}^{k}\left[\hat{Z}(S_k) - \acute{Z}(S_k)\right]^2} \qquad \text{Eq. 10}$$

$$IQR = (RS_1 - RS_2) \qquad \text{Eq. 11}$$

Where ME - is the mean error, RMSE is the mean square prediction error, k is the total number of cross validation points (the total number of trawl stations), $\hat{Z}(S_k)$ estimated values of the response variables at location S_k, $\acute{Z}(S_k)$ is values of the response variable at validation location S_k, RS_1 is the 1st of the residuals and RS_2 is the 3rd quartile of the residuals.

All the statistical analysis, logistic GAM, Gaussian GAM, and geostatistics were conducted using the statistical analysis environment R (R Development Core Team, 2012) and the following packages: gstat (Pebesma 2004), SoDA (Chambers 2012), and lattice (Sarkar 2008) for geostatistics, and mgcv for GAM (Wood 2004, 2006)

Results

Table 1. Analysis of several GAM models fitted to the distribution data for the West coast. HKP denotes the models used for M. paradoxus, HKC for M. capensis, and SNOK denotes T. atun. For each of the three species a GAM model using all three environmental variables (depth, chl-a, SST) was found to be the best performing with the lowest AIC score.

Models	AIC	dAIC	R.DF	M.deviance	R.deviance	P.deviance	Species	Coast
M7: HKP > 0 ~ s(sst, chl) + s(depth)	500.6101	0	1117.594	773.3637499	453.7989587	63.02047352	*M. paradoxus*	West
M3: HKP > 0 ~ s(depth)	591.4651	90.85497	1137.814	642.070385	585.0923236	52.32153654	*M. paradoxus*	West
M6: HKP > 0 ~ s(sst, chl) + s(Dist)	774.1151	273.505	1116.956	501.1363223	726.0263863	40.83699079	*M. paradoxus*	West
M5: HKP > 0 ~ s(sst) + s(chl) + s(Dist)	777.4936	276.8835	1124.757	482.1545694	745.0081392	39.29019078	*M. paradoxus*	West
M4: HKP > 0 ~ s(Dist)	803.2343	302.6242	1133.734	438.4613274	788.7013812	35.72968151	*M. paradoxus*	West
M2: HKP > 0 ~ s(chl)	1026.348	525.7379	1131.118	220.5780442	1006.584664	17.97463716	*M. paradoxus*	West
M1: HKP > 0 ~ s(sst)	1131.294	630.6835	1133.711	110.4478584	1116.71485	9.000261958	*M. paradoxus*	West
M7: HKC > 0 ~ s(sst, chl) + s(depth)	336.0638	0	1130.422	652.6576643	314.9083362	67.453555	*M. capensis*	West
M3: HKC > 0 ~ s(depth)	355.0092	18.94535	1137.282	619.9935644	347.5724361	64.07765094	*M. capensis*	West
M6: HKC > 0 ~ s(sst, chl) + s(Dist)	754.5193	418.4555	1104.364	286.317749	681.2482515	29.59154713	*M. capensis*	West
M5: HKC > 0 ~ s(sst) + s(chl) + s(Dist)	768.6265	432.5627	1122.254	236.4313129	731.1346876	24.43567806	*M. capensis*	West
M4: HKC > 0 ~ s(Dist)	790.7111	454.6472	1132.834	193.1875866	774.378414	19.96634715	*M. capensis*	West
M2: HKC > 0 ~ s(chl)	847.823	511.7591	1134.698	132.3472436	835.2187569	13.67836856	*M. capensis*	West
M1: HKC > 0 ~ s(sst)	901.2962	565.2323	1132.24	83.78996963	883.7760309	8.659871222	*M. capensis*	West
M7: SNOK > 0 ~ s(sst, chl) + s(depth)	1124.177	0	1116.307	435.0233466	1074.791611	28.81302403	*T. atun*	West
M3: SNOK > 0 ~ s(depth)	1215.587	91.40999	1136.221	303.7859518	1206.029006	20.12074064	*T. atun*	West
M6: SNOK > 0 ~ s(sst, chl) + s(Dist)	1354.152	229.9745	1109.368	218.9265304	1290.888427	14.50022265	*T. atun*	West
M5: SNOK > 0 ~ s(sst) + s(chl) + s(Dist)	1355.825	231.6474	1119.567	196.8563535	1312.958604	13.0384424	*T. atun*	West
M1: SNOK > 0 ~ s(sst)	1448.075	323.8972	1131.523	80.69358278	1429.121375	5.344600832	*T. atun*	West
M4: SNOK > 0 ~ s(Dist)	1448.31	324.133	1132.055	79.39446044	1430.420497	5.258555696	*T. atun*	West
M2: SNOK > 0 ~ s(chl)	1468.759	344.5815	1135.222	52.61154147	1457.203416	3.484635068	*T. atun*	West

Table 2. Analysis of several GAM models fitted to the distribution data for the South coast. HKP denotes the models used for *M. paradoxus*, HKC for *M. capensis*, and SNOK denotes *T. atun*. For *M. paradoxus* and *T. atun* a GAM model using all three environmental variables (depth, chl-a, SST) was found to be the best performing with the lowest AIC score. For *M .capensis* the model using only depth had the lowest AIC score.

Models	AIC	dAIC	R.DF	M.deviance	R.deviance	P.deviance	Species	Coast
M7: HKP > 0 ~ s(sst, chl) + s(depth)	281.9202402	0	1422.60574	1290.71154	227.1317198	85.035891	M. paradoxus	South
M3: HKP > 0 ~ s(depth)	304.1380434	22.21780317	1444.113342	1225.478531	292.3647279	80.73814762	M. paradoxus	South
M6: HKP > 0 ~ s(sst, chl) + s(depth)	588.7333716	306.8131314	1419.354919	990.4000497	527.4432096	65.25048246	M. paradoxus	South
M5: HKP > 0 ~ s(sst) + s(chl) + s(Depth)	638.3344437	356.4142035	1435.171334	909.1661481	608.6771112	59.89855293	M. paradoxus	South
M4: HKP > 0 ~ s(Depth)	1024.158847	742.2386071	1440.173791	513.3368303	1004.506429	33.82014757	M. paradoxus	South
M1: HKP > 0 ~ s(sst)	1067.232504	785.3122634	1443.653194	463.3043686	1054.538891	30.52386113	M. paradoxus	South
M2: HKP > 0 ~ s(chl)	1158.442595	876.5223544	1442.477945	374.4447738	1143.398485	24.66952839	M. paradoxus	South
M3: HKC > 0 ~ s(depth)	161.8201287	0	1447.073877	337.4083917	155.9678826	68.38764028	M. capensis	South
M7: HKC > 0 ~ s(sst, chl) + s(depth)	164.6046479	2.784519244	1444.928898	338.91383	154.4624444	68.69277011	M. capensis	South
M5: HKC > 0 ~ s(sst) + s(chl) + s(Depth)	259.67643	97.85630133	1433.162615	267.3746134	226.001661	54.19283969	M. capensis	South
M6: HKC > 0 ~ s(sst, chl) + s(Depth)	264.0221966	102.2020679	1437.683466	253.9871451	239.3891293	51.47939986	M. capensis	South
M4: HKC > 0 ~ s(Depth)	365.688092	203.8679633	1442.032402	143.6233793	349.7528951	29.11031331	M. capensis	South
M1: HKC > 0 ~ s(sst)	381.2237971	219.4036684	1442.811271	126.5299363	366.8463381	25.64572779	M. capensis	South
M2: HKC > 0 ~ s(chl)	435.7429583	273.9228296	1447.999529	61.63425845	431.7420159	12.4923434	M. capensis	South
M7: SNOK > 0 ~ s(sst, chl) + s(depth)	955.7045568	0	1421.897338	238.0081914	899.4992319	20.9236605	T. atun	South
M5: SNOK > 0 ~ s(sst) + s(chl) + s(Depth)	956.8962684	1.191711552	1438.840866	202.9294238	934.5779996	17.83983292	T. atun	South
M6: SNOK > 0 ~ s(sst, chl) + s(Depth)	957.007358	1.302801207	1430.601869	219.2963274	918.211096	19.2786722	T. atun	South
M4: SNOK > 0 ~ s(Depth)	1033.688117	77.98356014	1443.243793	117.3317213	1020.175702	10.31480928	T. atun	South
M3: SNOK > 0 ~ s(depth)	1063.044972	107.3404149	1443.605028	87.25239563	1050.255028	7.670490217	T. atun	South
M1: SNOK > 0 ~ s(sst)	1086.533708	130.8291508	1445.457184	60.05934813	1077.448075	5.279908236	T. atun	South
M2: SNOK > 0 ~ s(chl)	1091.621935	135.9173781	1445.151924	55.58164017	1081.925783	4.886266149	T. atun	South

For the West Coast the best performing model for all three species was the one consisting of an interaction between SST and Chl-a, with depth as a factor (table 1). Depth was found to be the most important factor for both coasts. Although all three variables were used for the West coast models, the AIC score for the depth-only model for *M. capensis* was only slightly higher than the model using all three variables. For the South coast depth was the only factor that was used to model the distribution of *M. capensis* (table 2) although the AIC score for a 3 variable model was only slightly higher. The *M. paradoxus* and *T. atun* for the South coast models used all three variables.

Table 3. KED performance statistics for West coast model distributions. The performance statistics are detailed in the Methods section.

	M. paradoxus	*M. capensis*	*T. atun*
Kappa	0.920245399	0.952673094	0.769500438
Sens	0.747126437	0.802325581	0.765778401
Spec	0.971590909	0.979360165	0.775700935
TSS	0.718717346	0.781685747	0.541479336
AUC	0.96326193	0.972160223	0.83916517
threshold	0.42	0.47	0.45

Models for *M. paradoxus* and *M. capensis* have high Kappa values whereas *T. atun* has a much lower value. The TSS score for *T. atun* is also low at 0.54.

Table 4. KED performance statistics for South coast model distributions

	M. paradoxus	*M. capensis*	*T. atun*
Kappa	0.970344828	0.986896552	0.867586207
Sens	0.983259912	0.796610169	0.99681782
Spec	0.923809524	0.994967649	0.025906736
TSS	0.907069436	0.791577819	0.022724556
AUC	0.994499685	0.94519246	0.817634717
threshold	0.545	0.555	0.54

Models for *M. paradoxus* and *M. capensis* have high Kappa values whereas *T. atun* has a higher value than that of the West coast. The TSS score for *T. atun* is extremely low at 0.023, indicating very low explanation for this model.

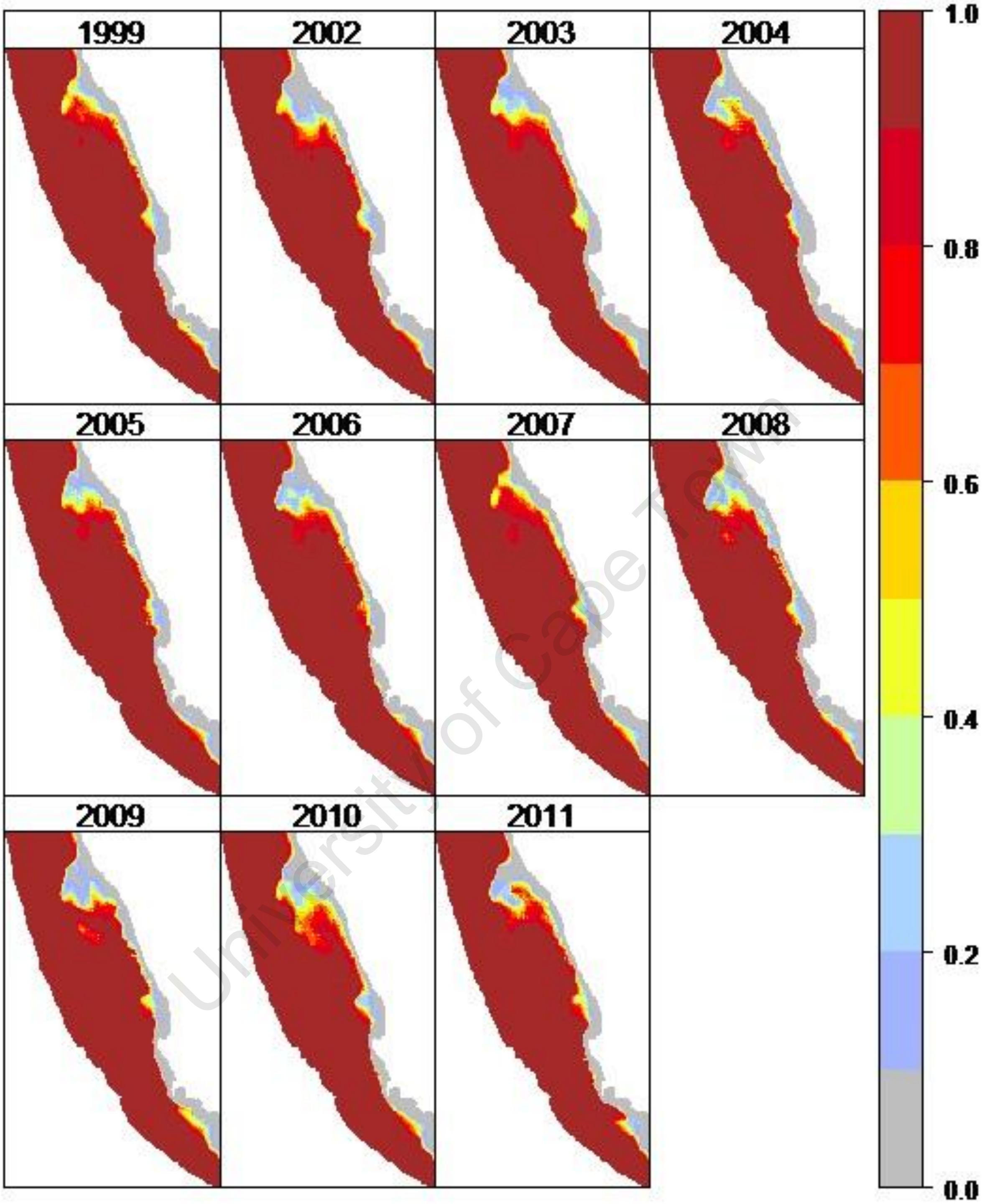


Figure 1. GAM modelled probability of presence of *M. paradoxus* on the West Coast based on presence/absence data using all three environmental variables, depth, SST and chl-a. Scale-bar indicates the probability of *M. paradoxus* being present.

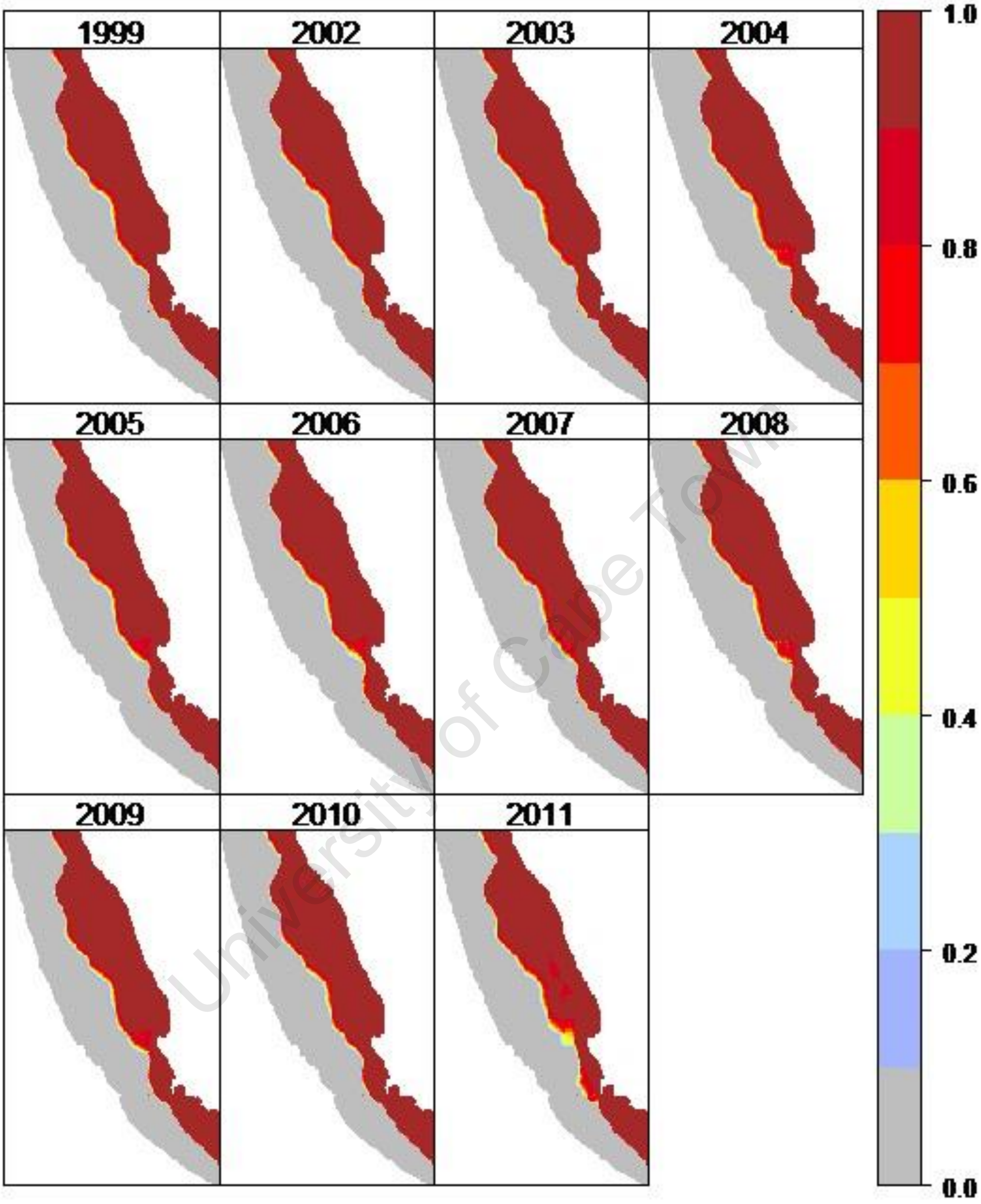


Figure 2. GAM modelled distribution of *M. capensis* on the West Coast based on presence/absence data using all three environmental variables, depth, SST and chl-a.. Scale-bar indicates probability of *M. capensis* being present.

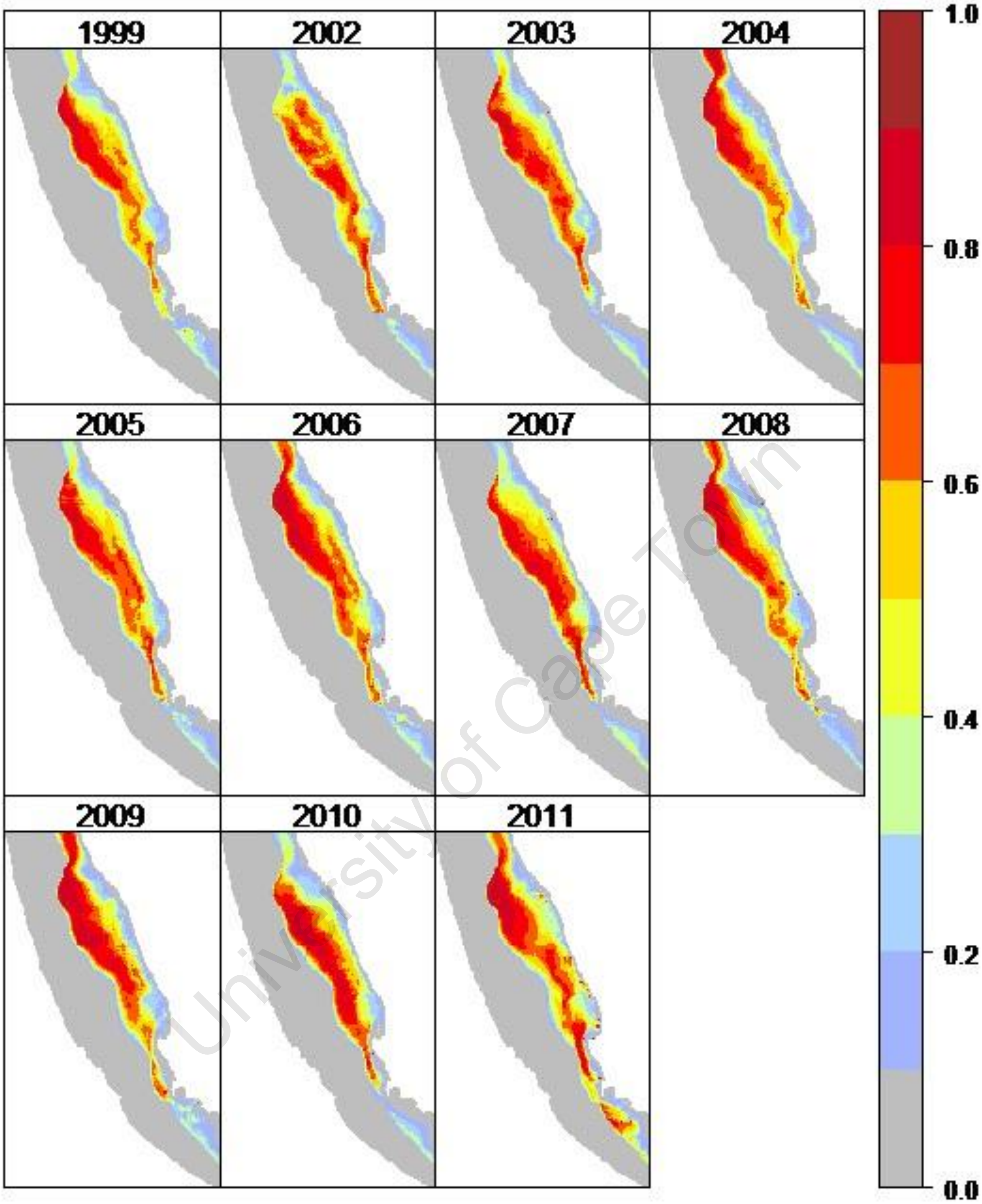


Figure 3. GAM modelled distribution of *T. atun* on the West Coast based on presence/absence data using all three environmental variables, depth, SST and chl-a.. Scale-bar indicates probability of *T. atun* being present.

There is a high probability, 0.9-1.0 of *M. paradoxus* being found in deeper waters off the West Coast (figure 1). The probability drops sharply nearer the coast. This distribution shows a strong fidelity to depth with a narrow band of changing probability at the edges of the distribution. Adjacent to the coast there is a probability of 0.0-0.1 of *M. paradoxus* being present. There is some interannual variability apparent in the modelled distribution which is most evident in the northern region of the map closest to the coast, with a distinct change in the probability of *M. paradoxus* being present. In 1999 there is a relatively small area of low probability extending into the deeper regions of the West coast. This area increases several fold by 2002 and stays roughly the same size until 2007 when it almost disappears. In 2008-2011 this area begins to extend again and stretches southwards into the previously high area of probability. There is also high variability around the southern part of the study area, particularly around the False Bay area but this is far less pronounced.

The modelled distribution for *M. capensis* on the West coast (figure 2) shows the same fidelity to depth as *M. paradoxus* in figure 1. The probability of *M. capensis* being found near the coast is between 0.9-1.0. This distribution is uniform in the near shore areas. There is a very sharp transition from very high probability to very low probability which closely reflects the bathymetry of the region (See figure 1 in Methods section). The range *of M. capensis* overlaps that of *M. paradoxus*. There is very little interannual variation except for a small region near Cape Columbine. This region shows a decrease in probability from 2004-2006 and from 2008-2009. 2011 sees a sharper contraction and a larger area of lower probability for this small region.

The distribution of *T. atun* (figure 3) is patchier than that of *M. paradoxus* and *M. capensis*. The highest probability of occurrence lies at the edge of the continental slope, midway between the shallow coastal waters and deeper offshore waters. The probability of occurrence is higher in the north and tapers off closer to Cape Point. South of Cape Point there is a very low probability of occurrence except for 2011 where there is a large region which has a probability of 0.9-1.0.

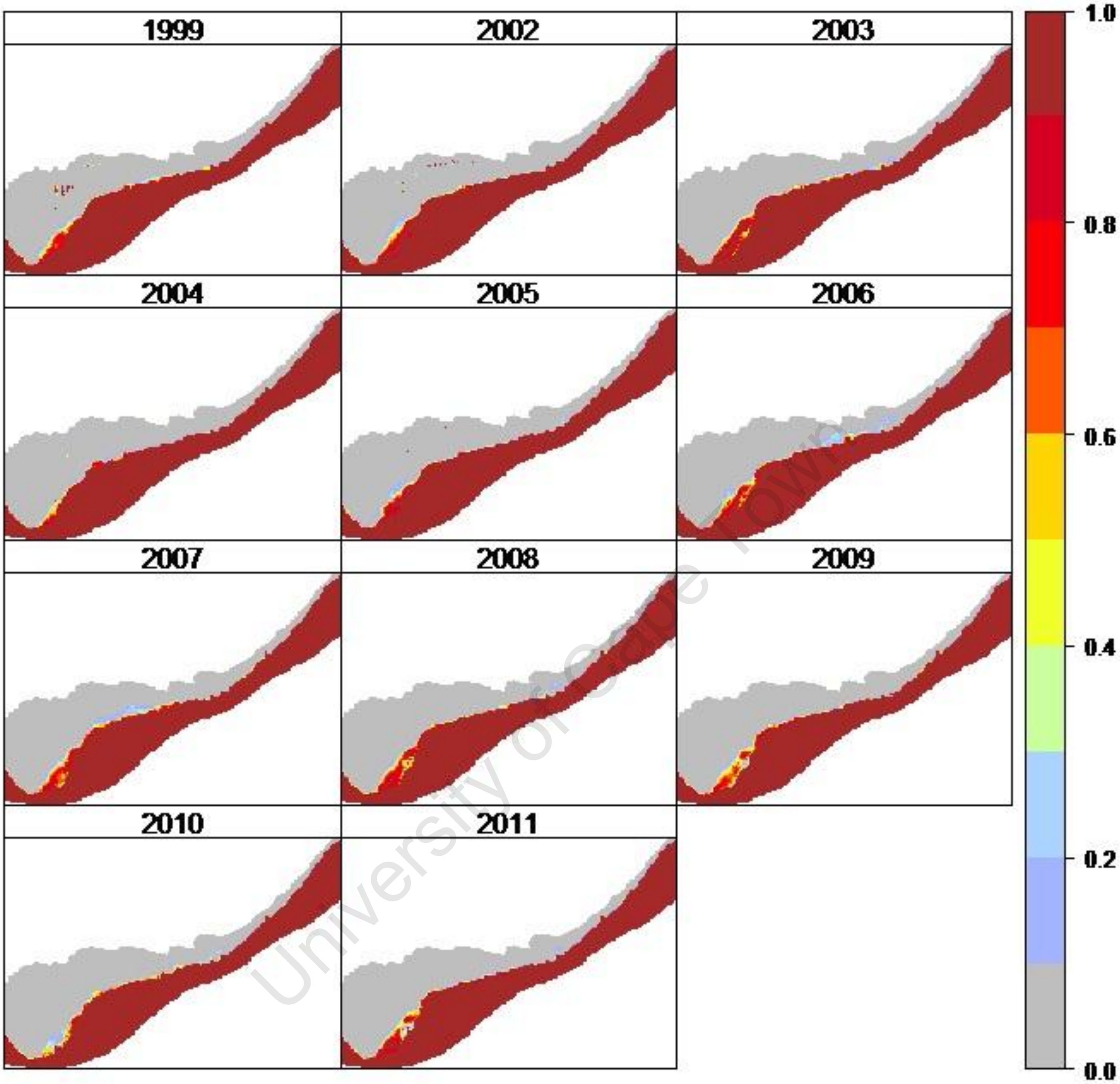


Figure 4. GAM modelled distribution of *M. paradoxus* on the South Coast based on presence/absence data using all three environmental variables, depth, SST and chl-a. Scale-bar indicates probability of *M. paradoxus* being present.

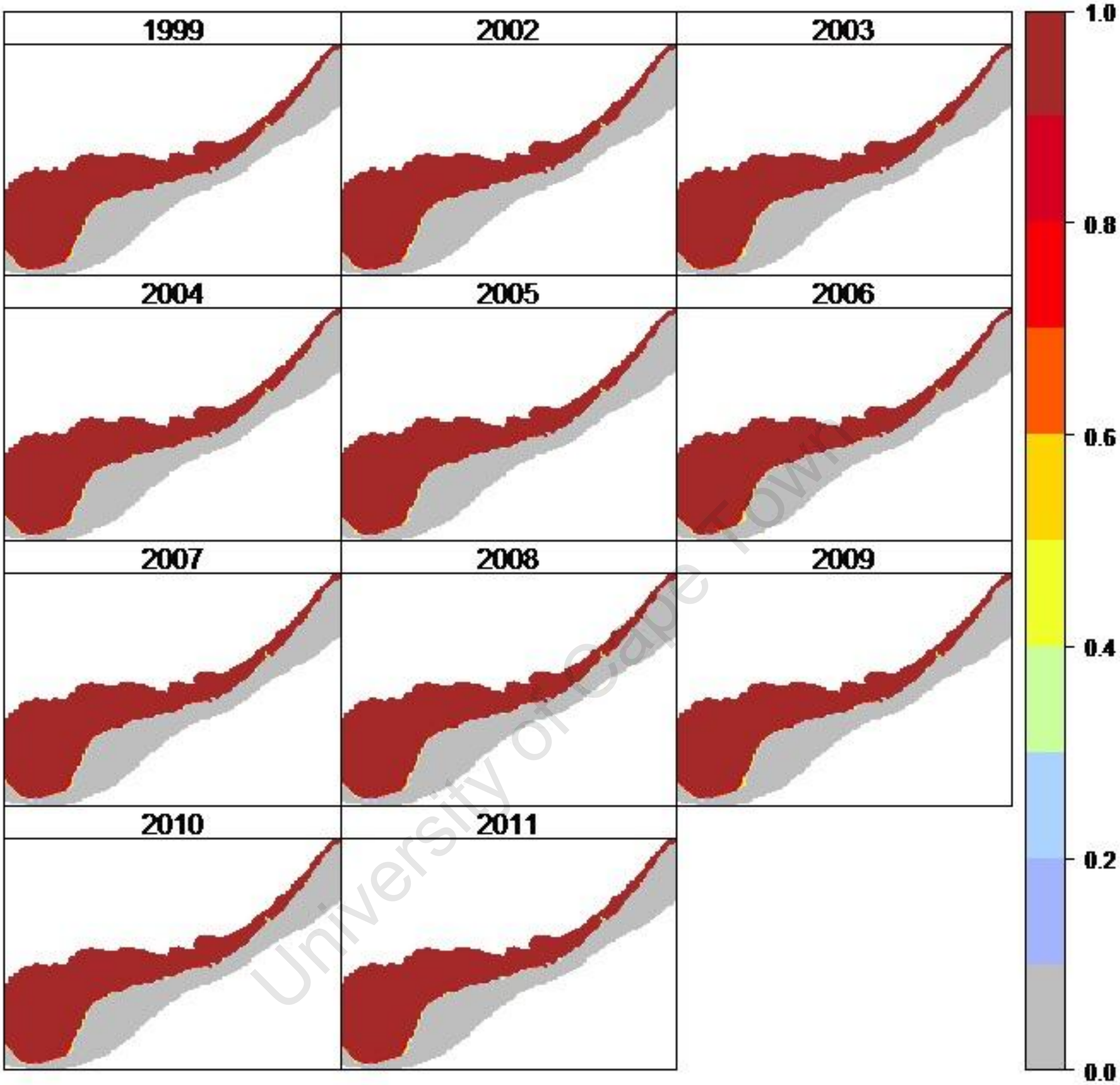


Figure 5. GAM modelled distribution of *M. capensis* on the South Coast based on presence/absence data using only depth as a factor. Scale-bar indicates probability of *M. capensis* being present.

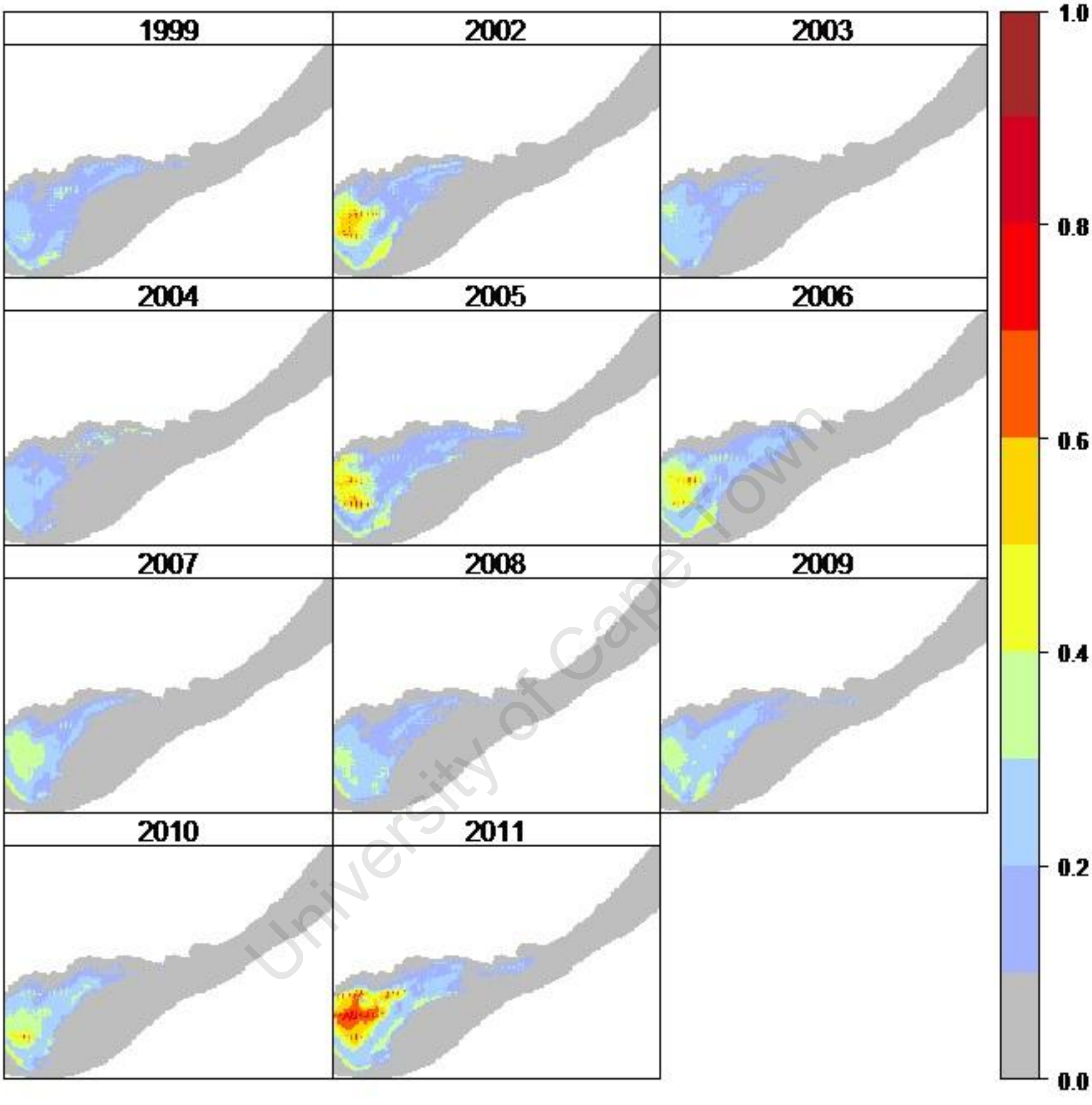


Figure 6. GAM modelled distribution of *T. atun* on the South Coast based on presence/absence data using all three environmental variables, depth, SST and chl-a. Scale-bar indicates probability of *T. atun* being present.

For the South coast (figure 4) *M. paradoxus* is present only in the deeper waters with a very low probability of this species being found in the shallower coastal waters. There is a very narrow transition zone where the species goes from a high probability of being found to a very low probability of being found. The high probability zone very closely follows the bathymetry of this region, with *M. paradoxus* staying in the deeper waters. There is very little interannual variation.

M. capensis (figure 5) has a high probability of occurrence in the near shore area. For this distribution the only factor in the model was depth. There is very little overlap with the M. paradoxus in this region. The high probability zone closely matches the bathymetry of this area. As with *M. paradoxus* there is very little interannual variation and a sharp decline in probability of occurrence at the edges of the distribution. North of Algoa Bay there is 0.0-0.1 probability of occurrence.

T. atun (figure 6) has, for the most part a low probability of being found on the South Coast. The region with the highest probability of *T. atun* being found is in the South-west region. There is some interannual variation evident with 2002, and 2005-2006 showing probability of 04.-05, with the highest probability of T. atun being present seen in 2011. This corresponds to the higher probability seen in the southern region of figure 3 in 2011.

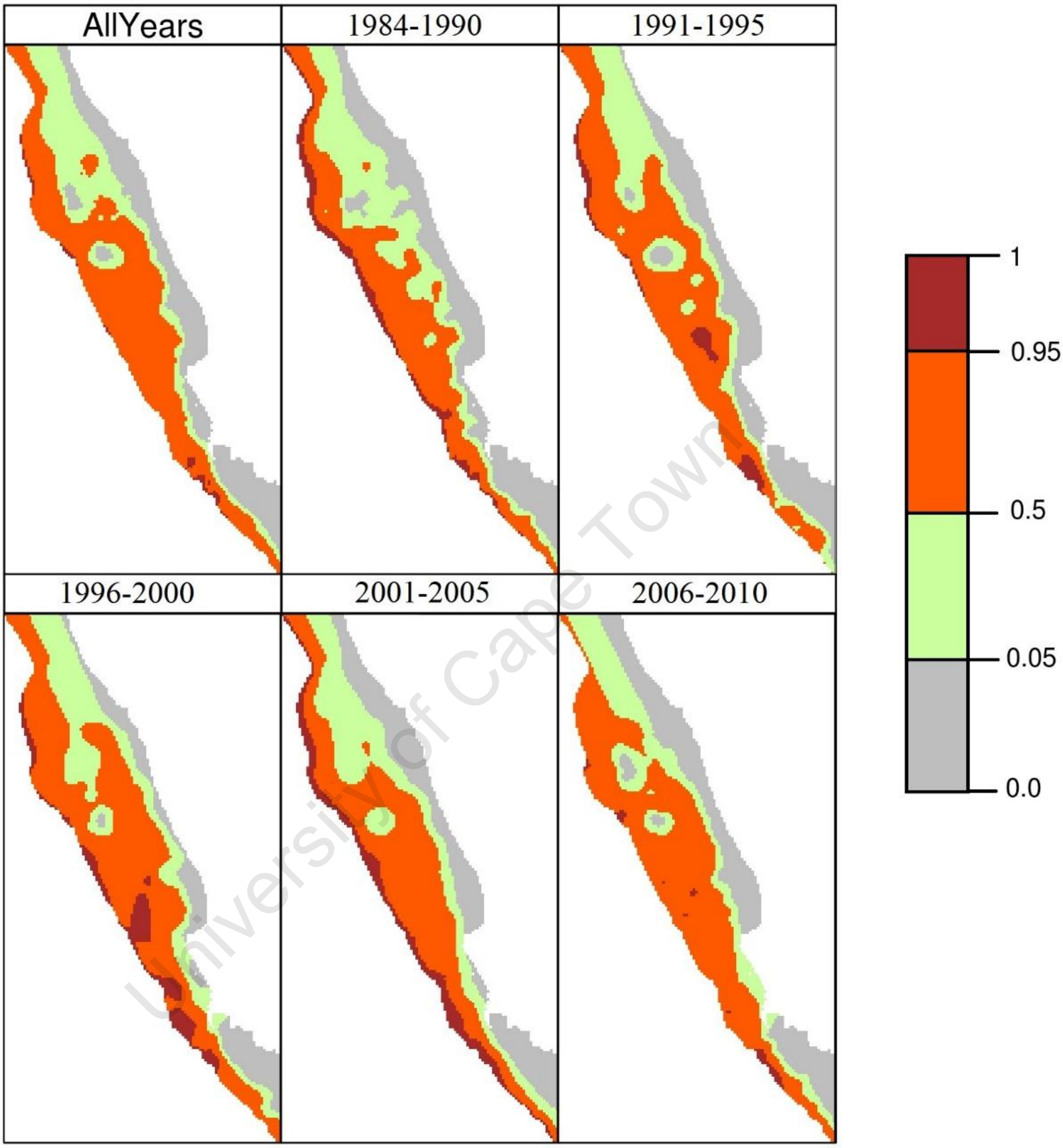


Figure 7. *M. paradoxus* KED modelled density as a proportion of the maximum catch recorded for the west coast. Red denotes 95%-100% of the maximum density, orange denotes 50%-95%, green 5%-50%, grey 0%-5%. The data were divided into five time periods. 1984-1990, 1991-1995, 1996-2000, 2001-2005, 2006-2010. The frame with heading AllYears is the model distribution for data from all five time periods.

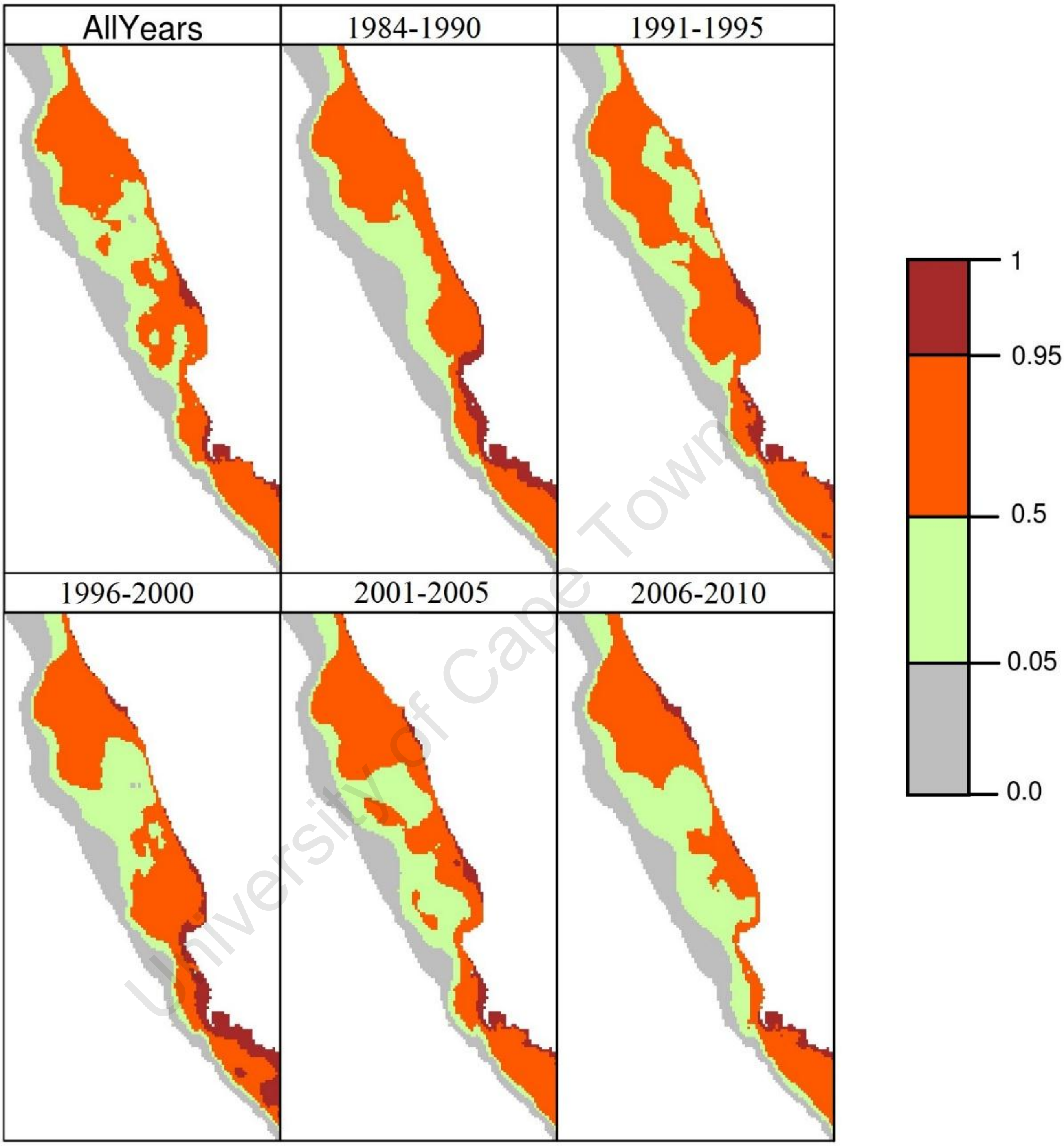


Figure 8. *M. capensis* KED modelled density as a proportion of the maximum catch recorded for the west coast. Red denotes 95%-100% of the maximum density, orange denotes 50%-95%, green 5%-50%, grey 0%-5%. The data were divided into five time periods. 1984-1990, 1991-1995, 1996-2000, 2001-2005, 2006-2010. The frame with heading AllYears is the model distribution for data from all five time periods.

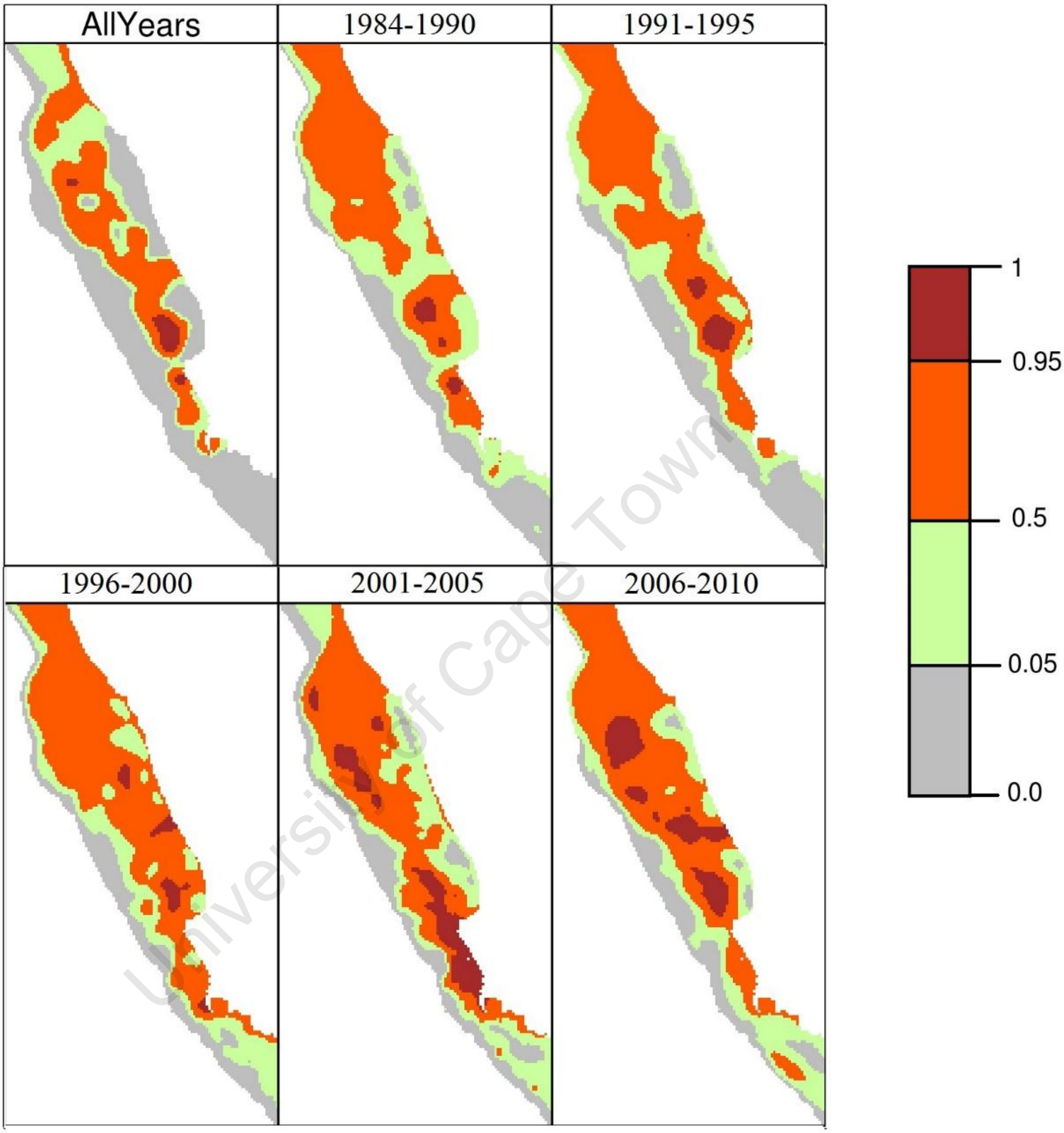


Figure 9. *T. atun* KED modelled density as a proportions of the maximum density recorded for the West coast. Red denotes 95%-100% of the maximum density, orange denotes 50%-95%, green 5%-50%, grey 0%-5%. The data were divided into five time periods. 1984-1990, 1991-1995, 1996-2000, 2001-2005, 2006-2010. The frame with heading AllYears is the model distribution for data from all five time periods.

There is some variation in the KED modelled density of *M. paradoxus* (figure 7) based on catch density but there are no distinct shifts in distribution over the five time periods. The most obvious change is at the western edge of the distribution where the density changes from 0.5-1.0 over the 5 time periods. For the period 2006-2010 the density around Cape Point increases above 0 for the first time in this area and reaches 0.05-0.5.

M. capensis (figure 8) has a high modelled density around Cape Point and the surrounds which remains consistently high throughout the five time periods, only changing slightly in the range of this area of high density. The shallower waters are where the density is highest across the coast with the density dropping as the depth increases. There are few areas where high density of the two species overlap, with these areas occurring in the intermediate depths. For the most part the high densities do not overlap and are complimentary rather than co-occurring.

There is a slight shift northwards in the model for the highest density of *T. atun* over time (figure 9). Also in the later time periods the distribution increases indicating higher densities of *T. atun* and possibly a range expansion. The centre of the modelled density appears to be in the intermediate waters north of Cape Columbine. This reflects the model illustrated in figure where the highest probability of presence was in the intermediate waters north of Cape Columbine. Over the five time periods there is also an increase in density in a southerly direction past Cape Point.

Discussion

The spatial distribution for the two hake species overlap in the West coast to a large degree (figures 1 and 2) and overlap very little in the South coast (figures 4 and 5). Studies of age and size structure of *M. paradoxus* on the South Coast indicate that there are mostly only adults present (Botha, 1985, von der Heyden, 2007). *M. paradoxus* are believed to migrate to the South Coast to spawn. Eggs and larvae are then carried back towards the West Coast by strong jet currents (Hutchings et al, 2002,). These eggs and larvae then develop into juveniles on the intermediate shelf area off the West Coast (Garavelli et al, 2012). This means that almost all of the *M. paradoxus* individuals on the South coast are adults (Botha, 1986; von der Heyden et al, 2007) The presence-absence models showing that depth is a major factor in the distribution of the cape hakes confirm what previous studies have found in that the two species are closely associated with depth (Botha, 1985, Burmeister, 2001).

Juveniles of *M. paradoxus* are almost exclusively found on the West Coast. *M. capensis* are known to predate on juveniles of *M. paradoxus* (Botha, 1980; Roel and McPherson, 1988). The movement by *M. capensis* to deeper waters is reflected in figure 2 by the large overlap in the distribution of the two hakes. This overlap is not evident in the South Coast. Therefore the lack of overlap evident on the South coast could be due to the lack of juvenile *M. paradoxus* meaning that *M. capensis* would not move to deeper water to feed on the juveniles and remain in the shallower areas instead (Roel and Macpherson, 1988). But this could be construed as a rather deterministic view of ecology, with another hypothesis being predation by *M. capensis* on juvenile *M. paradoxus* is determined by opportunity (the ranges of the two species overlap therefore predation will occur) rather than by purpose.

However, the distribution is strictly dictated by the bathymetry of the region; the Shallow-water hake stay near shore in the shallow water, and the Deep-water hake are found in deeper water where the continental shelf begins to drop. The narrow shelf region between 200m and 600m could be the reason for the model using depth for the South coast to be the best performing model. The narrow shelf would mean that the depth increases sharply and without there being a reason for *M. capensis* to move deeper, such as predation of juvenile *M. paradoxus*, there would be little overlap of the two hake species.

The distribution for *T. atun* (see figures 3 and 6) is more variable. Previous studies have indicated that adult specimens of *T. atun* are mostly found in water 150 m to 550m deep (Dudley, 1987; Griffiths, 2002). This is reflected in the West Coast distribution (fig 3) which is highest in the northern regions but drops significantly further south, closer to the Agulhas bank. According to the models' results, the South coast has a fairly low probability of finding this species except in 2011 (see fig 6), when *T. atun* has a high probability of occurrence in the southern region of the South coast. The low overall explanation of the variability seen in the two modelled distributions of *T. atun* is most likely due to their nomadic nature (Griffiths 2002; McPherson and Jetz, 2007). The very low explanation of the model (table 4) is indicative of this. *T. atun* is a voracious predator of sardine (*Sardinops sagax*) and anchovy (*Engraulis japonicus*) and will typically move to where these prey items are most readily available (Wickens et al, 1998; Griffiths, 2002). Future modelling of *T.atun* should include the potential distribution of sardines and anchovy in concord with environmental variables such as SST and Chl-a. This could improve the accuracy and reliability of this modelling approach. The variability seen in the models generated for *T. atun* for this study are not reliable enough to base any type of management scenario upon but they do lay the groundwork for future attempts at modelling this species (See McPherson and Jetz, 2007, for a more comprehensive discussion of the effect a species' ecology has on modelling its distribution).

The KED modelling used catch abundance data in order to estimate density and model species spatial distributions. These models generally performed better than the models that used presence/absence data (table 3). Burmesiter (2001) found that density of *M. capensis* was highest above 27^{o}S which is contrary to our study in that density increased further south until it peaked around the Cape Point Peninsula (see fig 12). In the models' output, *M. paradoxus* shows a reduction in density near the edge of its distribution in the last time period (see figure 11). *M. capensis* also shows a drop in density in the False Bay region for the last time period (see figure 12). These two drops in density could be due to overfishing or competition with other predators, such as seals (*Arctocephalus pusillus*), for available food resources, most likely a combination of the two (Wickens et al, 1998, Griffiths et al, 2005, Griffiths et al, 2010. Both these explanations could benefit from closer study and analysis of catch data in the future (e.g. time period 2011-2015) with a comparison of the total landings for commercial species such as sardine and anchovy.

There is greater variation in the modelled distribution of *T. atun* (figure 13). This is probably due to the species being nomadic and driven primarily by the distribution and availability of its food resource (Crawford et al, 1990). Modelling distribution of species such as this will always have a larger degree of variation than species that have a more predicable home range, like the two hake species (McPherson and Jetz, 2007)..*T. atun* spawns on the continental shelf on the west coast and the eggs and larvae are transported north of Cape Columbine where they mature (Griffiths, 2002) therefore a study of *T. atun* occurrence and density would benefit from a more northerly focus, i.e. the northern Benguela system off Namibia and Angola. The lower density seen on the south coast could be due to *T. atun* only moving to the south coast when prey availability is higher there, rather than to breed as is the case for *M. paradoxus* (Botha, 1985; Griffiths, 2002).

These results highlight an important aspect of how climate change can affect the distribution of a species. In the environmental niche models, the two hake species were constrained primarily by depth rather than any other environmental factors. This means that in the face of climate and environmental change, the physiological rates of hake, their food, and predation success might be affected. The possible consequence is that hake could potentially be more vulnerable to climate change if they do not move and are constrained by their fidelity to a particular depth or depth range (Burmeister, 2001). Species like *T. atun*, which are not constrained by depth and favour a nomadic lifestyle, can move to wherever the food is most plentiful but are still constrained by prey items, such as sardine and anchovy, being affected by these environmental variables (Griffiths, 2002; Edwards and Richardson, 2005). Changes in food supply will mean that the suitable habitat niche for these species might change and shrink, leaving a greatly reduced population. This potential for a reduced area of suitability needs to be carefully considered by fisheries managers in light of potential climate change in the future coupled with overfishing (Worm et al, 2009).

There are some drawbacks to the data used for this study. Deeper trawl points below 500m were rare and trawls further afield were fewer than those closer to shore. This means that there is more data available for the shallower inshore waters. This would result in the data being more accurate with regards to *M. capensis*, leaving potential for sampling bias to creep into the models. This would not be a problem in the presence-absence models but the density models may have been affected. The nomadic nature if *T.atun* also means that sampling this species is difficult, as witnessed by the high variability and low explanation of the models for this species. For *T. atun* a tagging programme combined with survey trawls would greatly increase the resolution and accuracy of distribution models.

GAM and KED modelling provide ecologists and fisheries managers a powerful and useful tool in predicting species distribution without the need for exhaustive studies on species-specific mechanistic responses to environmental variables. With climate change comes the potential for species distribution to change (Pearson and Dawson, 2003). This approach could be applied to future climate models and thus be able to aid in the prediction of the future range and distribution of important marine species. As such, this approach has potential to support decision-making in conservation and fisheries management.

Acknowledgments

Many thanks go to my supervisors, Yunne-jai Shin and Dawit Yemane, for their help and input over many long hours, their supreme patience with late night emails, and close supervision. Thanks also go to Rob Anderson and the staff at DAFF who graciously endured my questions during my time spent working there.

References

Ando, A., Camm, J., Polasky, S., & Solow, A. (1998). Species distributions, land values, and efficient conservation. *Science, 279*(5359), 2126-2128.

Antilla, L. (2005). Climate of scepticism: US newspaper coverage of the science of climate change. *Global environmental change*, *15*(4), 338-352.

Araujo, M. B., &Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, *33*(10), 1677-1688.

Botha, L. (1973). Migrations and spawning behaviour of the Cape hakes. *S. Afr. Shipp. News Fish. Ind. Rev*, *28*(4), 62-63.

Botha, L. (1980). *The biology of the Cape hakes Merluccius capensis Cast. and M. paradoxus Franca in the Cape of Good Hope area* (Doctoral dissertation, Stellenbosch: Stellenbosch University).

Botha, L. (1985). Occurrence and distribution of Cape hakes Merluccius capensis Cast. and M. paradoxus Franca in the Cape of Good Hope area.*South African Journal of Marine Science*, *3*(1), 179-190.

Botha, L. (1986). Reproduction, sex ratio and rate of natural mortality of Cape hakes Merluccius capensis Cast. and M. paradoxus Franca in the Cape of Good Hope area. *South African Journal of Marine Science*, *4*(1), 23-35.

Burmeister, L. M. (2001). Depth-stratified density estimates and distribution of the Cape hake Merluccius capensis and M. paradoxus off Namibia deduced from survey data, 1990–1999. *South African Journal of Marine Science*, *23*(1), 347-356.

Burmeister, L. (2005). Is there a single stock of Merluccius paradoxus in the Benguela ecosystem?. *African Journal of Marine Science*, *27*(1), 23-32.

Chambers, J.M (2012). SoDA: Functions and Examples for "Software for DataAnalysis". R package version 1.0-4. http://CRAN.R-project.org/package=SoDA

Crawford, R. J. M., Underhill, L. G., & Venter, J. D. (1990). Handline catches and stock identity of snoek Thyrsites atun off South Africa and Namibia. *South African Journal of Marine Science*, *9*(1), 95-99.

Dudley, S. F. J. (1987). *Snoek Thyrsites atun in South African waters: aspects of its biology, distribution and fishery* (Doctoral dissertation, University of Cape Town).

Edwards, M., & Richardson, A. J. (2004). Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature*, *430*(7002), 881-884.

Enriquez, A. G., & Friehe, C. A. (1995). Effects of wind stress and wind stress curl variability on coastal upwelling. *Journal of Physical Oceanography*, *25*(7), 1651-1671.

Fortin, M-J., and Dale, M. R. T. 2005. Spatial analysis: A Guide for Ecologists. Cambridge University Press 380 pp.

Garavelli, L., Grüss, A., Grote, B., Chang, N., Smith, M., Verley, P.& Lett, C. (2012). Modeling the dispersal of Cape hake ichthyoplankton. *Journal of plankton research*, *34*(8), 655-669.

Griffiths, C. L., Van Sittert, L., Best, P. B., Brown, A. C., Clark, B. M., Cook, P. A., ... & Gordon, J. D. M. (2005). Impacts of human activities on marine animal life in the Benguela: a historical overview. *Oceanogr. Mar. Biol. Annu. Rev.*, *42*, 303-392.

Griffiths CL, Robinson TB, Lange L, Mead A (2010). Marine Biodiversity in South Africa: An Evaluation of Current States of Knowledge. PLoS ONE 5(8): e12008. doi:10.1371/journal.pone.0012008

Griffiths, M. H. (2000). Long-term trends in catch and effort of commercial linefish off South Africa's Cape Province: snapshots of the 20th century. *South African Journal of Marine Science*, *22*(1), 81-110.

Griffiths, M. H. (2002). Life history of South African snoek, *Thyrsites atun* (Pisces: Gempylidae): a pelagic predator of the Benguela ecosystem. *FISHERY BULLETIN-NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION*, *100*(4), 690-710.

Guisan, A., & Zimmermann, N. E. (2000). Predictive habitat distribution models in ecology. *Ecological modelling*, *135*(2), 147-186.

Guisan, A., Edwards Jr, T. C., & Hastie, T. (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological modelling*, *157*(2), 89-100.

Hastie TJ and Tibshirani RJ (1990) Generalized additive models. Monographs on statistics and applied probability, 43. Chapman and Hall, London

Hoegh-Guldberg, O., Mumby, P. J., Hooten, A. J., Steneck, R. S., Greenfield, P., Gomez, E., &Hatziolos, M. E. (2007). Coral reefs under rapid climate change and ocean acidification. *Science*, *318*(5857), 1737-1742.

Hutchings, J. A., & Myers, R. A. (1994). What can be learned from the collapse of a renewable resource? Atlantic cod, Gadus morhua, of Newfoundland and Labrador. *Canadian Journal of Fisheries and Aquatic Sciences*, *51*(9), 2126-2146.

Hutchings, L., Beckley, L. E., Griffiths, M. H., Roberts, M. J., Sundby, S., & Van Der Lingen, C. (2002). Spawning on the edge: spawning grounds and nursery areas around the southern African coastline. *Marine and Freshwater Research*, *53*(2), 307-318.

Lutjeharms, J. R. (2006). The Agulhas current. *Afr. J. Mar. Sci*, *28*, 729-732.

McPherson, J., & Jetz, W. (2007). Effects of species' ecology on the accuracy of distribution models. *Ecography*, *30*(1), 135-151.

Meehl, G. A., Stocker, T. F., Collins, W. D., Friedlingstein, P., Gaye, A. T., Gregory, J. M., ...& Zhao, Z. C. (2007). Global climate projections. *Climate change*, 747-845.

Morin, A., Lamoureux, W., & Busnarda, J. (1999). Empirical models predicting primary productivity from chlorophyll a and water temperature for stream periphyton and lake and ocean phytoplankton. *Journal of the North American Benthological Society*, 299-307.

Pauly, D., Christensen, V., Guénette, S., Pitcher, T. J., Sumaila, U. R., Walters, C. J., & Zeller, D. (2002). Towards sustainability in world fisheries.*Nature*, *418*(6898), 689-695.

Payne, A. I. L., Rose, B., & Leslie, R. W. (1987). Feeding of hake and a first attempt at determining their trophic role in the South African west coast marine environment. *South African Journal of Marine Science*, *5*(1), 471-501.

Payne, A. I., & Punt, A. E. (1995). Biology and fisheries of South African Cape hakes (*M. capensis* and *M. paradoxus*). In *Hake* (pp. 15-47). Springer Netherlands.

Pebesma, E.J., 2004. Multivariable geostatistics in S: the gstat package. Computers & Geosciences, 30: 683-691. R Development Core Team (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Pearson, R. G., & Dawson, T. P. (2003). Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful?.*Global ecology and biogeography*, *12*(5), 361-371.

Rademeyer, R. A., Butterworth, D. S., & Plagányi, É. E. (2008). Assessment of the South African hake resource taking its two-species nature into account. *African Journal of Marine Science*, *30*(2), 263-290.

Rahmstorf, S. (2002). Ocean circulation and climate during the past 120,000 years. *Nature*, *419*(6903), 207-214.

Roel, B. A., & Macpherson, E. (1988). Feeding of Merluccius capensis and M. paradoxus off Namibia. *South African Journal of Marine Science*, *6*(1), 227-243.

Robinson, L. M., Elith, J., Hobday, A. J., Pearson, R. G., Kendall, B. E., Possingham, H. P., & Richardson, A. J. (2011). Pushing the limits in marine species distribution modelling: lessons from the land present challenges and opportunities. *Global Ecology and Biogeography*, *20*(6), 789-802.

Rouault, M., Penven, P., & Pohl, B. (2009). Warming in the Agulhas Current system since the 1980's. *Geophysical Research Letters*, *36*(12), L12602.

Rykaczewski, R. R., & Checkley, D. M. (2008). Influence of ocean winds on the pelagic ecosystem in upwelling regions. *Proceedings of the National Academy of Sciences*, *105*(6), 1965-1970.

Sarkar, Deepayan (2008) Lattice: Multivariate Data Visualization with R. Springer, New York. ISBN 978-0-387-75968-5

Shannon, L. V., Crawford, R. J. M., Brundrit, G. B., & Underhill, L. G. (1988). Responses of fish populations in the Benguela ecosystem to environmental change. *Journal du Conseil: ICES Journal of Marine Science*, *45*(1), 5-12.

Shannon, L. V., & Nelson, G. (1996). The Benguela: large scale features and processes and system variability. In *The South Atlantic* (pp. 163-210). Springer Berlin Heidelberg.

Smith, J. L. B. (2003). *Smiths' sea fishes*. Struik. Cape Town.

Thomas, A. C., Carr, M. E., & Strub, P. T. (2001). Chlorophyll variability in eastern boundary currents. *Geophysical Research Letters*, *28*(18), 3421-3424.

von der Heyden, S., Lipinski, M. R., & Matthee, C. A. (2007). Mitochondrial DNA analyses of the Cape hakes reveal an expanding, panmictic population for *Merluccius capensis* and population structuring for mature fish in *Merluccius paradoxus*. *Molecular phylogenetics and evolution*, *42*(2), 517-527.

Walther, G. R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J., &Bairlein, F. (2002). Ecological responses to recent climate change. *Nature*, *416*(6879), 389-395.

Wardell-Johnson, G. (2000). Biodiversity and Conservation in Namibia into the 21st Century. *Population–Development–Environment in Namibia*, 17.

Wickens, P. A., Japp, D. W., Shelton, P. A., Kriel, F., Goosen, P. C., Rose, B., ... & Krohn, R. G. (1992). Seals and fisheries in South Africa—competition and conflict. *South African Journal of Marine Science*, *12*(1), 773-789.Wood, S.N. (2006) Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC.

Wood, S.N. (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models. Journal of the American Statistical Association. 99:673-686.

Worm, B., Hilborn, R., Baum, J. K., Branch, T. A., Collie, J. S., Costello, C., ...& Zeller, D. (2009). Rebuilding global fisheries. *Science*, *325*(5940), 578-585.