Vol. 321: 267–281, 2006 | MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser | Published September 8

Variation in demersal fish species richness in the oceans surrounding New Zealand: an analysis using boosted regression trees

J. R. Leathwick[1,*], J. Elith[2], M. P. Francis[3], T. Hastie[4], P. Taylor[1]

[1]National Institute of Water and Atmospheric Research, PO Box 11115, Hamilton, New Zealand
[2]School of Botany, University of Melbourne, Parkville 3010, Victoria, Australia
[3]National Institute of Water and Atmospheric Research, Private Bag 14901, Wellington, New Zealand
[4]Sequoia Hall, Stanford University, Stanford, California 94305-4065, USA

ABSTRACT: We analysed relationships between demersal fish species richness, environment and trawl characteristics using an extensive collection of trawl data from the oceans around New Zealand. Analyses were carried out using both generalised additive models and boosted regression trees (sometimes referred to as 'stochastic gradient boosting'). Depth was the single most important environmental predictor of variation in species richness, with highest richness occurring at depths of 900 to 1000 m, and with a broad plateau of moderately high richness between 400 and 1100 m. Richness was higher both in waters with high surface concentrations of chlorophyll *a* and in zones of mixing of water bodies of contrasting origins. Local variation in temperature was also important, with lower richness occurring in waters that were cooler than expected given their depth. Variables describing trawl length, trawl speed, and cod-end mesh size made a substantial contribution to analysis outcomes, even though functions fitted for trawl distance and cod-end mesh size were constrained to reflect the known performance of trawl gear. Species richness declined with increasing cod-end mesh size and increasing trawl speed, but increased with increasing trawl distance, reaching a plateau once trawl distances exceed about 3 nautical miles. Boosted regression trees provided a powerful analysis tool, giving substantially superior predictive performance to generalized additive models, despite the fitting of interaction terms in the latter.

KEY WORDS: Demersal fish · Species richness · Boosted regression trees · Statistical model

Resale or republication not permitted without written consent of the publisher

INTRODUCTION

Apart from its inherent scientific interest, understanding the distributional patterns of individual species, communities and ecosystems is essential to robust conservation management. Such information is required not only for the identification of priority sites for protection (Ward et al. 1999), but also for monitoring the impacts of human activities, particularly in systems subject to recurrent harvesting of natural resources as in the oceans (e.g. Colloca et al. 2003). Unfortunately, comprehensive inventories of species distributions are rarely available when conservation-management issues require resolution over extensive geographic areas. This is particularly so in the oceans where sampling at depth over large areas is not only difficult and expensive, but is often also constrained by lack of taxonomic expertise (Solbrig 1991). Several approaches have been developed to overcome this lack of comprehensive data, including the analysis of existing data to identify zones of high species richness (e.g. Ponder et al. 2001, Shackell & Frank 2003) and the identification of species whose status can be used as indicators of wider ecosystem health (e.g. Diaz et al. 2004). Alternatively, abiotic data are used either to construct frameworks for management (e.g. T. H. Snelder et al. 2006) or to predict the distributions of biological properties from scattered samples (e.g. Ferrier et al. 2002).

*Email: j.leathwick@niwa.co.nz

© Inter-Research 2006 · www.int-res.com

Geographic variation in species richness has long been explored in both terrestrial (Rohde 1992, Huston 1994) and marine settings (Grassle & Maciolek 1992, Rex et al. 1993, Roy et al. 1998, Gray 2001, 2002), and it has become the subject of increased interest with recognition of the global imperative for biodiversity conservation. In marine studies, the search for evidence of declining species richness with progression from equatorial to polar environments, a change that would parallel terrestrial patterns, has been a dominant theme (Gray 2001). However, although evidence of declining diversity with increasing latitude has been found in several Northern Hemisphere studies, mostly of benthic organisms (Stehli et al. 1967, Rex et al. 1993, Roy et al. 1998), results from Southern Hemisphere studies have been much less convincing (Clarke 1992, Gray et al. 1997). There is also conflicting evidence about relationships between species richness and depth. Although Levington (1995) argues for a general increase in species richness with depth, reaching a maximum about the continental slope and declining thereafter, results from quantitative studies of fish species richness are generally inconsistent except for their demonstration of declining richness in abyssal waters (McClatchie et al. 1997). Similarly, after reviewing evidence for trends of marine benthic diversity along depth gradients, Gray (2001) concluded 'there is no clear trend in increasing species richness from coasts to deep sea'. Here we present results of an analysis of relationships between fish species richness and a comprehensive set of functionally based environmental predictors, using an extensive set of trawl data collected from the oceans surrounding New Zealand. Our aims were to model species richness with a method capable of revealing important ecological relationships, while also producing a map of predicted species richness that could be used in conservation planning.

Boosted regression trees

The majority of our analyses in this study are carried out using the relatively new statistical technique of gradient-boosted regression trees (Friedman 2001), sometimes referred to as stochastic gradient boosting. Along with other model-averaging (ensemble) methods, this differs fundamentally from conventional regression based techniques such as generalised additive models (GAM – Hastie & Tibshirani 1990). Whereas the latter seek to fit the single most parsimonious model that best describes the relationship between a response variable and some set of predictors, ensemble methods fit a large number of relatively simple models whose predictions are then combined to give more robust estimates of the response. In boosted regression trees (BRT) each of the individual models consists of a simple classification or regression tree, i.e. a rule-based classifier that partitions observations into groups having similar values for the response variable, based on a series of binary rules (splits) constructed from the predictor variables (Hastie et al. 2001). The boosting algorithm uses an iterative method for developing a final model in a forward stage-wise fashion, progressively adding trees to the model, while re-weighting the data to emphasize cases poorly predicted by the previous trees. A BRT model can therefore be seen as a regression model in which each of the individual model terms is a simple regression tree (Friedman et al. 2000).

Advantages offered by a BRT model include its ability to accommodate both different types of predictor variables and missing values, its immunity to the effects of extreme outliers and the inclusion of irrelevant predictors, and its facility for fitting interactions between predictors (Friedman & Meulman 2003). Fitting of interaction effects is controlled by varying the size of the individual regression trees. Where the individual tree terms consist of a single rule constructed using just 1 predictor variable, no interaction effects are fitted, and the final model is likely to approximate closely one fitted using any conventional regression technique that allows the fitting of non-linear responses, e.g. a GAM. However, where the individual trees consist of 2 or more rules, the function fitted for any one predictor may vary depending on the value taken by another predictor, with the potential complexity of these interaction effects increasing as the size of the individual tree terms increases.

While these features of a BRT model make it a potentially powerful tool for analysing complex ecological datasets, it also poses a number of challenges for which we demonstrate possible solutions. In particular, trees can continue to be added to a BRT model until, eventually, all observations are perfectly explained, i.e. the model becomes over-fitted to the training data. Given that our objective is to produce a model having a high level of generality (Hastie et al. 2001), some procedure is required to identify an optimal number of trees that maximises the ability of a model to make accurate predictions to new, independent sites while avoiding excessive model complexity. Care is also required with BRT models where the tree size is 2 or greater, because of their capacity to automatically fit interactions between predictor variables. Given that such effects are only fitted where required by the data, and given the complexity of a BRT model, the contribution of these interaction effects can be difficult to detect. In addition, care is required when interpreting the functions fitted for predictor variables, as their shapes

can vary dramatically depending on the values taken by other predictors.

Given its relatively recent advent, there are only a few published examples of the use of boosting (Friedman & Meulman 2003, Kuhnert et al. 2003, Lawrence et al. 2004), particularly with ecological data (Cappo et al. 2005, Kawakita et al. 2005, Elith et al. 2006). Because of this we also fit parallel GAM models, where feasible, to allow evaluation of the comparative performance of BRT. GAMs were chosen for comparison because they are regularly used in ecology (e.g. Guisan & Zimmerman 2000), due to their ability to fit non-linear relationships between a response variable and its predictors. This is generally advantageous when analysing the complex relationships typically found in ecological datasets (e.g. Olden & Jackson 2002).

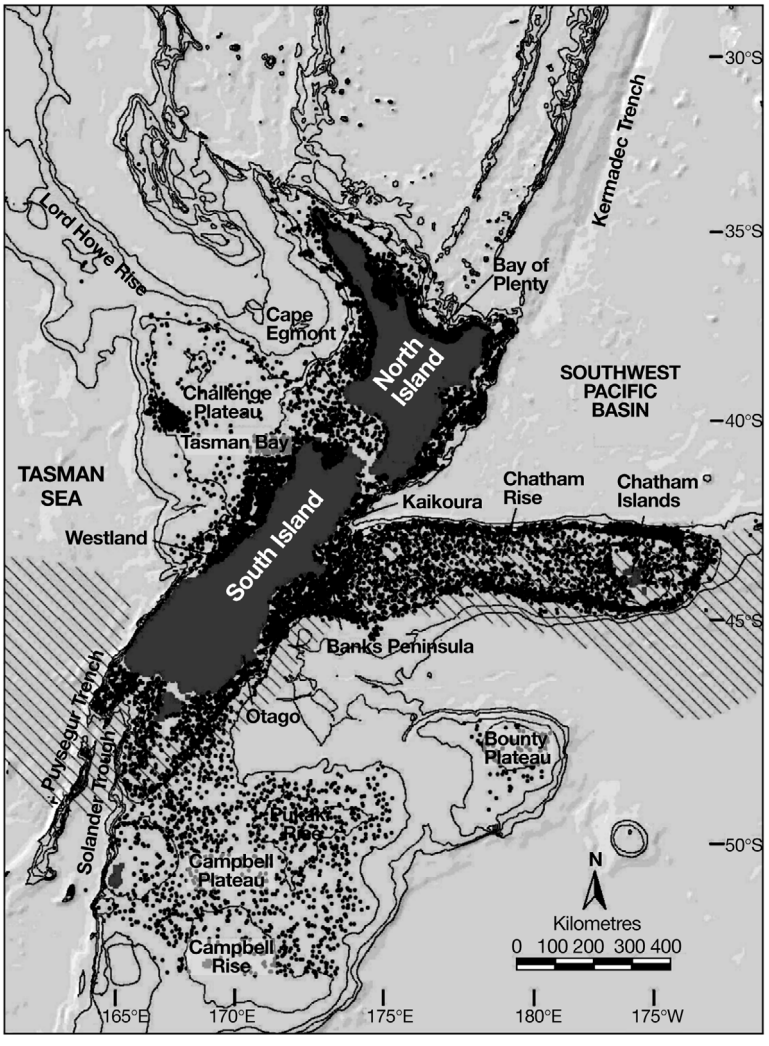

Fig. 1. Broad features of the oceans surrounding New Zealand. Bathymetric contours are shown only to a depth of 2000 m. The 500, 1000, 1500 and 2000 m isobaths are given. Locations of trawl sites are indicated by dots, and the approximate position of the subtropical front (STF) is shown by diagonal hatching

STUDY AREA AND METHODS

Study area. Although New Zealand has a relatively small land area, with its offshore islands it extends across a wide latitudinal range (~30 to 55° S, Fig. 1), and the oceans that surround it encompass a diverse range of environmental conditions (e.g. Heath 1985, Bradford-Grieve et al. 1991, in press). The dominant feature of these waters is the Subtropical Front (STF), which separates warm, saline and nutrient-poor waters of subtropical origin in the north, from colder, low-salinity but nutrient-rich waters of subantarctic origin to the south. The STF is deflected to the south from its normal latitudes by the New Zealand landmass, but returns to the north immediately east of the South Island, and then to the east along the Chatham Rise. Strong current flows occur along this front, particularly along the Southland coast, and form several relatively stable gyres mostly to the east of New Zealand (Bradford-Grieve et al. in press). The continental shelf surrounding New Zealand is generally narrow, but extensive submarine plateaus occur to the northwest and southeast. The most prominent and economically important of these is the Chatham Rise, which extends east from Banks Peninsula to the Chatham Islands and forms a bathymetric anchor for the STF. Deeper abyssal waters occur close to the south-western coast of the South Island along the Puysegur Trench, and to the northeast of the North Island along the Kermadec Trench. Descriptions of the demersal fish assemblages in these waters are contained in Bull et al. (2001), Kendrick & Francis (2002), Beentjes et al. (2002) and Francis et al. (2002)

Data. Species richness data used in this analysis were drawn from an extensive collection of research bottom trawls carried out over the period from 1979 to 1997 (Francis et al. 2002) and sampling most of the waters around New Zealand (Fig. 1). To minimise the effect of variation among vessels and gear types in species catchability, only data from 3 research vessels were included (RVs 'James Cook', 'Kaharoa' and 'Tangaroa'). A total of 16 946 records were used in the analysis, after discarding a number of records either that lacked associated attribute data or for which there

was a substantial mismatch between the recorded trawl depth and the average depth in that general location, suggesting inaccurate geo-referencing. Abundances (catch weights) were recorded in the database for all demersal species occurring in at least of 1% of all trawls, with a total of 126 species recorded overall, including both commercial and non-commercial species. Species richness was calculated for each trawl by tallying the number of species recorded as present, which we interpret as a measure of point or alpha-diversity as defined by Whittaker (1972).

Eight environmental predictors (Table 1) were selected for their likely functional relevance to variation in the distributions of individual fish species, and hence species richness. Selection of variables was influenced in part by results from previous analyses of fish species richness, including one New Zealand study (McClatchie et al. 1997). The average depth of each trawl (*AvgDepth* in Table 1) was included as a surrogate for the environmental changes that occur with increasing depth, i.e. increasing pressure, decreasing light and temperature, and variation in salinity. The average depth across all trawls was 537 m, but the distribution of values was bimodal, with many coastal trawls, a second peak of trawls at around 800 to 900 m, and very few trawls deeper than 1500 m.

Table 1. Predictors used in the analyses. GLM: Generalised Linear Model. n miles: nautical miles

Variable	Derivation	Mean (range)
AvgDepth	Average depth as recorded during trawling	537 m (5–1700)
TempResid	Residuals from a GLM relating temperature to depth using natural splines	0.0°C (–5.3–3.9)
SalResid	Residuals from a GLM relating salinity to depth and temperature using natural splines	0.0 psu (–0.28–0.28)
Chla	Satellite-image based estimate of chl *a* concentrations	0.579 ppm (0–4.87)
SstGrad	Spatial gradient of mean annual sea-surface temperature	0.11°C km^{-1} (0–0.52)
TidalCurr	Maximum, depth-averaged tidal current velocity	0.19 m s^{-1} (0–1.6)
OrbVel	Mean orbital velocity, derived from a wave climatology – log_{10}-transformed after adding a value of 1	3.59 m s^{-1} (1–38.9) (after transformation)
Slope	Seabed slope derived from bathymetric layer	0.96° (0–13.3)
Speed	Average trawl speed	3.2 knots (0.2–5.7)
Dist	Trawl distance	2.43 n miles (0.1–26.4)
CodEndSize	Mesh size of the trawl codend	75.0 mm (38–140)

Estimates of the average temperature and salinity on an annual basis were derived for each trawl location from the CSIRO Atlas of Regional Seas (CARS – Ridgeway et al. 2001). As this provides estimates for half-degree grid cells at fixed depth intervals, we extracted the relevant depth profile for each trawl site and used a spline interpolation routine to estimate the temperature and salinity at the specific depth at which trawling was carried out. To avoid problems in fitting the subsequent richness model arising from strong correlations among depth, temperature and salinity, we applied transformations to both temperature and salinity estimates. First, we adjusted our temperature estimates for depth by fitting a non-linear regression (Fig. 2a) describing the average relationship between depth and temperature. We then used the residuals from this regression as a predictor (*TempResid*), these indicating for any site the deviation from the average temperature expected at its depth. Positive values indicate waters of subtropical origin and occur at depths down to approximately 700 to 800 m to the west and north of New Zealand, but in the east occur only as far south as the northern flanks of the Chatham Rise (see Fig. 1). Negative values indicate cool waters of subantarctic origin and are widespread east of the southern South Island and on the southern flanks of the Chatham Rise, also to depths of around 800 m. Similarly, we fitted a regression relating salinity to both depth (Fig. 2b) and temperature, and we used the residuals from this (*SalResid*) to describe variation in salinity, given the depth and temperature at any site. Negative values indicate lower salinity than expected given the depth and temperature and occur mostly at inshore sites around the western and southern South Island and the south-eastern North Island, while positive values occur both in shallow waters around the Chatham and subantarctic islands, and in deep southern waters around the margins of the Campbell Plateau.

Because the distributions of many fish species are likely to be influenced by overall ecosystem productivity (e.g. Bakun 1996), we overlaid the locations of trawl sites onto satellite-image-derived layers describing concentrations of chlorophyll (chl) *a* and sea-surface-temperature gradients. Estimates of chl *a* concentration (*Chla*), which gives a broad indication of primary productivity, were derived from remotely sensed data from 6 visible wavebands, collected between September 1997 and July 2001 by the Sea-Viewing Wide-Field-of-view Sensor (SeaWiFS) (Murphy et al. 2001). Values in oceanic waters typically range between 0.1 and 0.8 ppm, while those in many coastal waters are inflated by the confounding effects of suspended sediments, mainly of terrestrial origin. The data layer

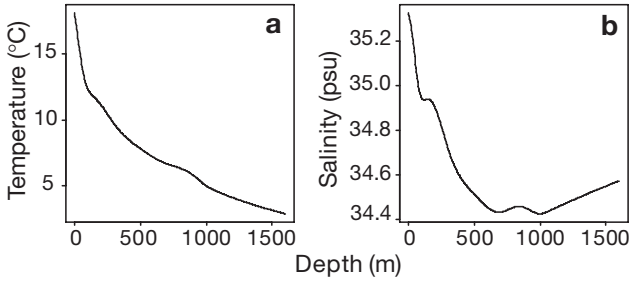

Fig. 2. Average relationships between (a) depth and temperature and (b) depth and salinity, as predicted from non-linear regressions fitted to data from all trawl sites

describing variation in sea-surface-temperature gradients (*SstGrad*) indicates locations in which frontal mixing of different water bodies is occurring. It was derived from estimates of the mean annual sea-surface temperature as measured from imagery with a spatial resolution of approximately 9 km, averaged over the period from 1993 to 1997 (Uddstrom & Oien 1999). These values were smoothed, and the magnitude of the spatial gradient for each grid cell was calculated by centred differencing.

Two variables were included to assess the effects of variation in more local scale mixing, predominantly in coastal waters. The first of these (*TidalCurr*) describes maximum depth-averaged flows from tidal currents and was calculated from a tidal model for New Zealand waters (Walters et al. 2001). The second (*OrbVel*) estimates the average mixing at the sea floor as a consequence of orbital wave action, and it was calculated from a wave climatology derived from a 20 yr hindcast (1979 to 1998) of swell-wave conditions in the New Zealand region (Gorman et al. 2003). Bed orbital velocities were assumed to be zero where the depth is greater than 200 m. Finally, the sea-floor slope (*Slope*) was calculated from a 1 km bathymetry layer using standard GIS routines, and it was included to allow testing of the suggestion by McClatchie et al. (1997) that there is a positive association between slope and fish diversity. Three variables were used to describe the trawl characteristics. These were the average trawl speed (*Speed*), the distance towed (*Dist*), and the cod-end mesh size (*CodendSize*).

Model fitting and evaluation. All GAM models were fitted in S-Plus (v6.1, Insightful) assuming a Poisson error distribution. All predictors were fitted using smooth terms with 4 degrees of freedom, and each predictor was tested for possible simplification of the fitted function or exclusion from the model. When an initial model was fitted, the functions for *Dist* and *CodendSize* were both clearly inconsistent with our knowledge of the behaviour of trawl gear, i.e. maximum species richness was predicted to occur at intermediate values for both variables. By contrast, we expect a monotonic decrease in richness with progression to coarse mesh sizes, as these allow smaller species greater chance of escape, and a monotonic increase in richness with increasing trawl distance, gradually reaching a plateau at longer distances. Closer inspection of the raw data indicated that analysis outcomes for these 2 predictors were being confounded by the patchy spatial and environmental distribution of some mesh sizes and distances. In particular, almost all trawls using very fine mesh sizes were in shallow northern waters, where species richness is generally lower than at comparable depths further south; almost no trawls were taken in these waters with coarse mesh sizes. Similarly, the majority of longer trawls were undertaken in deep environments, where both the total catch and species richness is generally low. We therefore refitted the model, specifying the terms for these variables so that the function fitted for distance was constrained to allow only a monotonic increase, and that for cod-end mesh size was constrained to a monotonic decrease. A second model was then fitted in which interaction terms were added using a forwards-step-wise procedure, with candidate terms consisting of those pair-wise interactions that were frequently fitted in a BRT model allowing for first-order interactions (see below).

All BRT models were fitted in R (v2.0.1, www.R-project.org; R Development Core Team 2004) using the 'gbm' library (Ridgeway 2004). Fitting a BRT model requires specification of 2 main parameters. The learning rate controls the rate at which model complexity is increased, with smaller values resulting in the fitting of a larger number of trees, each of lower individual influence and generally giving superior predictive performance in the ensemble model (Friedman 2001). The size (number of splits) of the individual trees is controlled by a parameter termed the interaction depth in the 'gbm' library. A value for this parameter of 1 indicates that each tree consists of a single node or decision rule (a decision 'stump'), a depth of 2 indicates that 2 nodes are used in each tree, allowing for 2-way interactions, and so on. As use of trees with 2 or more nodes only results in the fitting of interaction effects if required by the data, we view use of the term 'interaction depth' as potentially misleading, as even large trees are capable of fitting simple additive effects. We therefore prefer to describe this parameter as setting the 'tree size' of the model. Other settings were left at the defaults recommended in 'gbm'.

Three BRT models were fitted using a learning rate of 0.05 and with tree sizes of 1, 2 and 5 respectively. In each of these models, the fitted function for *Dist* was constrained to allow only a monotonic increase, and that for *CodendSize* was constrained to a monotonic

decrease, as for the GAM model. Ten-fold cross-validation was used to identify the optimal number of trees to use for each model and to subsequently assess the predictive performance of both GAM and BRT models (Hastie et al. 2001). The importance of predictor variables in BRT models was evaluated using a script in 'gbm' that calculates the contribution to model fit attributable to each predictor, averaged across all trees (Friedman 2001, Friedman & Meulman 2003). Purpose-written scripts were used to graph the fitted functions from both the GAM and non-interaction BRT models, with bootstrap re-sampling used to calculate confidence intervals around these fitted functions. We also wrote scripts to identify important interactions between predictors in those BRT models fitted with a depth of 2 and 5, and to calculate and graph values predicted in relation to major variables, while other variables were either held constant or varied in steps.

Predictions in geographic space were made in R using a set of spatial data describing the environmental attributes of cells on a 4 km grid across the waters surrounding New Zealand. Cells with depths greater than 1600 m were excluded, as were those with latitudes greater than 54°S, for which satellite-image-based data were not available. Predictions were formed using a script available in 'gbm', with values for predictors describing trawl characteristics set at their respective means in the trawl database. To obtain an estimate of the error associated with these predictions, we took repeated bootstrap samples, to which we fitted a BRT model and used this to make a separate prediction for the spatial data. Once these had been accumulated we identified the 5- and 95-percentile values for each grid cell as an estimate of the confidence intervals around our predictions. A detailed description of all methods used in the fitting and evaluation of BRT models is provided in Appendix 1.

RESULTS

Non-interaction models of species richness

Species richness averaged 12.7 across all trawls, and it ranged from a minimum of 0 to maximum of 38. All variables were retained as significant terms in the GAM model relating species richness to environment and trawl characteristics, and the non-interaction BRT model also made use of all variables. Contributions of predictors to model fit in the non-interaction BRT model were greatest for trawl distance and depth, followed by chl *a* concentration, temperature, and sea-surface-temperature gradient (Table 2). Comparable statistics were not available for the GAM model. Comparison of the respective performance statistics for these models indicated that the BRT model had greater predictive power, explaining 6% more deviance than the GAM model when making predictions for independent sites (Table 3a).

Table 2. Contributions of predictor variables to boosted regression tree (BRT) models relating demersal fish species richness to environment and trawl characteristics, using varying tree sizes. Variables are sorted in order of decreasing contribution, averaged across the 3 models

Variable	Tree size			Average
	1	2	5	
Dist	32.6	24.8	24.6	27.3
AvgDepth	29.6	24.4	21.5	25.2
Chla	8.7	10.1	11.1	10.0
TempResid	5.0	10.4	12.6	9.3
SstGrad	7.8	6.9	5.7	6.8
CodendSize	5.9	6.2	6.3	6.1
Speed	3.7	5.8	6.2	5.2
SalResid	3.7	5.5	4.9	4.7
TidalCurr	1.4	2.2	2.8	2.1
Slope	0.7	2.3	2.6	1.9
OrbVel	0.9	1.3	1.7	1.3

Relationships fitted by the non-interaction GAM and BRT models, both of which can be displayed as univariate functions, were broadly similar (Figs. 3 & 4). They indicated that greatest variation in species richness occurred with change in depth, trawl distance, trawl speed and salinity. More muted variation occurred in relation to temperature, salinity and chl *a*.

Table 3. Predictive performance of GAM (generalised additive model) and BRT models relating demersal fish species richness to environment and trawl characteristics. Table values indicate, for each model method: degree of complexity (1 = no interaction, 2 = 10 pair-wise interactions for GAM model, and a tree size of 2 for BRT model, and 5 = tree size of 5 for BRT models only); the number of trees fitted (boosted models); the mean residual deviance of the model; the mean residual deviance and its SE, calculated using 10-fold cross-validation repeated 5 times (Appendix 1); and the cross-validated proportion of the total deviance explained (D^2). The mean deviance for a null model is 2.971

Method	Complexity	No. of trees	Model residual deviance	Cross-validated residual deviance (SE)	D^2
(a) GAM	1	–	1.630	1.637 (0.074)	0.45
BRT	1	1312	1.524	1.558 (0.021)	0.48
(b) GAM	2	–	1.420	1.456 (0.063)	0.51
BRT	2	3476	1.142	1.281 (0.014)	0.57
(c) BRT	5	1252	1.000	1.195 (0.021)	0.60

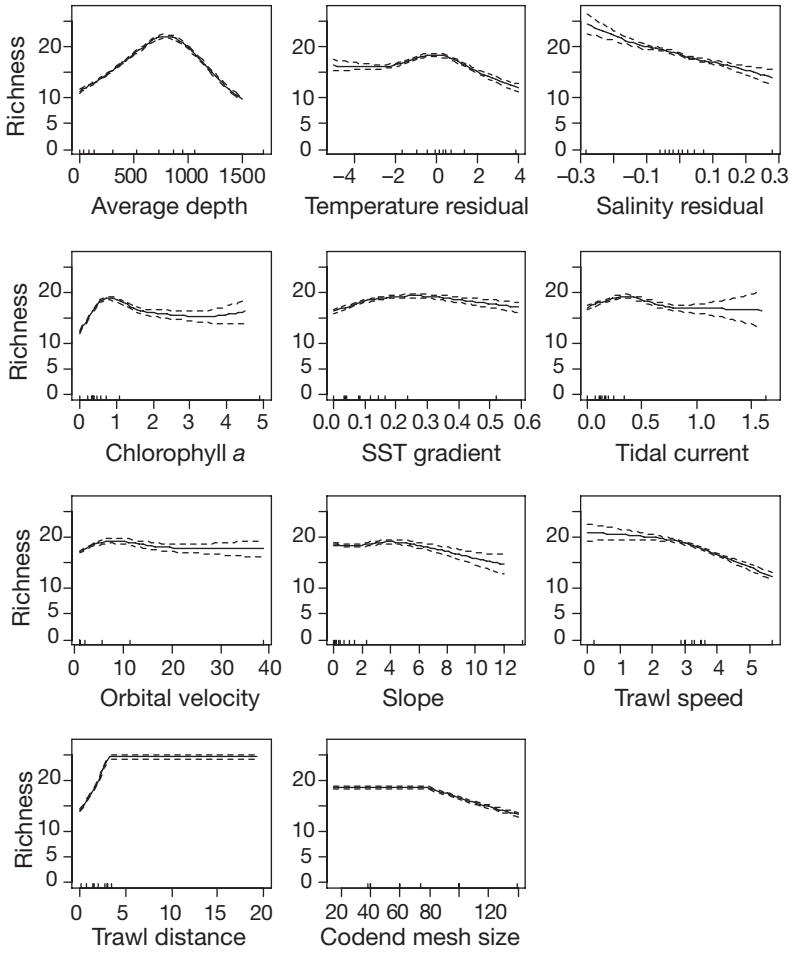

Fig. 3. Variation in demersal fish species richness predicted by a non-interaction GAM model using environment and trawl characteristics as predictors. Values were predicted for each variable, holding values for all other variables at their mean value for the trawl dataset. Dashed lines indicate the 5 to 95% confidence intervals for the predicted values, estimated from predictions made from 1000 models fitted to bootstrap samples of the trawl dataset. Ticks at the base of each plot indicate deciles of the predictor variable. See Table 1 for units

Predicted species richness had an approximately bell-shaped relationship with depth, with high values occurring over a wide range of intermediate depths (ca. 400 to 1100 m) that are of average temperature and less saline than expected, given their depth. High species richness was predicted for waters with chl *a* concentrations of around 0.8, a value typical of oceanic sites of high primary productivity, e.g. along the Chatham Rise. A second peak of richness at sites with chl *a* concentrations of 4 to 5 should be regarded with caution, given the problems with suspended sediments described earlier, a feature that is reflected in the wide confidence limits. Higher richness was predicted for waters with moderate spatial gradients in sea-surface temperature, i.e. where mixing of different water masses is occurring, mostly along the STF. Species richness was predicted to show a slight decline as tidal currents and orbital velocities increase. While both models showed similar predictions of declining species richness with increasing trawl speed, they differed subtly in the nature of the constrained functions fitted for *Dist* and *CodendSize*.

Interaction models of species richness

Addition of simple interaction terms improved both the deviance explained and the predictive performance of the GAM and BRT models by 13 and 19%, respectively, compared to the non-interaction models (Table 3b), indicating the importance of interactions between predictor variables in explaining variation in species richness in this dataset. While only second-order interactions could be added to the GAM model, expansion of the BRT model to allow a tree size of 5 brought about a further increase in performance compared with the boosted model using a tree size of 2 (Table 3c). Given this superior performance, we focus on results from the BRT interaction model with a tree size of 5 for the remainder of this paper.

With progression to a tree size of 5, the contributions of predictor variables altered subtly (Table 2), with trawl distance and depth declining in importance and the contributions of several other variables increasing, particularly those for chl *a* and temperature. These latter changes presumably reflected the more frequent inclusion of these variables in the more complex individual regression trees fitted by this interaction model. The strongest interaction effects involved the predictors *AvgDepth* and *TempResid*, *AvgDepth* and *Chla*, *AvgDepth* and *SstGrad*, and *AvgDepth* and *CodendSize*. Although relationships between species richness, environment and trawl characteristics predicted by the depth 5 BRT model were broadly similar to those predicted from the non-interaction models (e.g. Fig. 4), they showed greater subtlety, which reflected the inclusion of interaction effects. In general, richness was predicted to increase with depth from about 15 species at 100 m to a maximum of about 22 at depths of around 1000 m but declined steeply thereafter (Fig. 5). However, the fitting of interaction effects resulted in

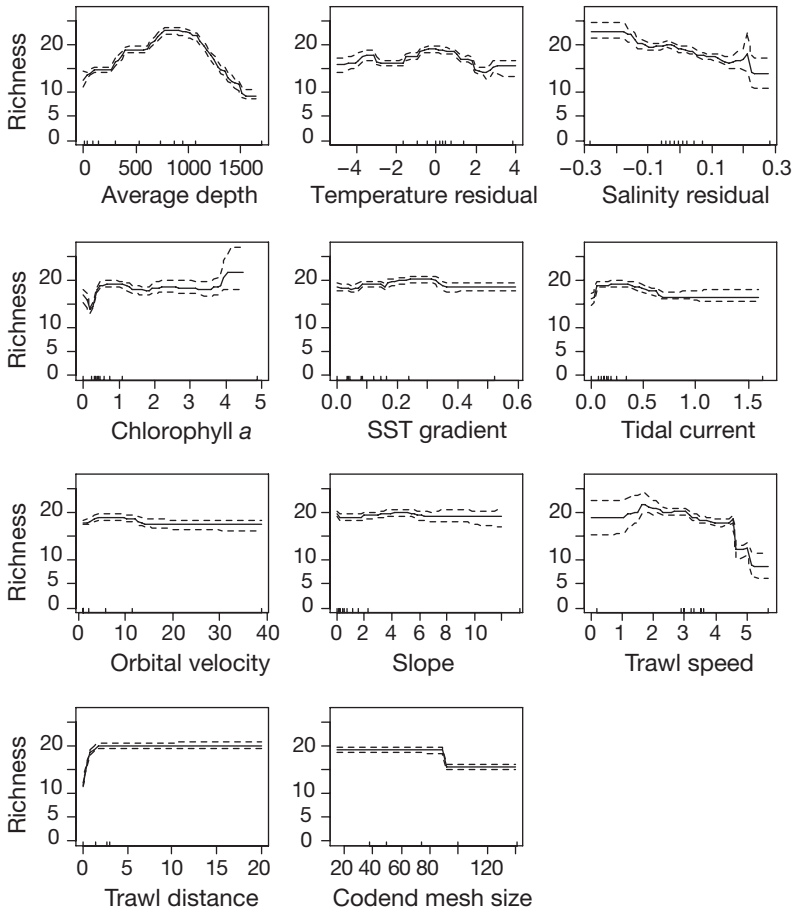

Fig. 4. Variation in demersal fish species richness predicted by a BRT model with a tree size of 1, and using environment and trawl characteristics as predictors. Values were predicted for each variable, holding values for all other variables at their mean value for the trawl dataset. Dashed lines indicate the 5 to 95% confidence intervals for the predicted values, estimated from predictions made from 1000 models fitted to bootstrap samples of the trawl dataset. Ticks at the base of each plot indicate deciles of the predictor variable. See Table 1 for units

the relationship between species richness and depth varying with temperature – for deep trawls, richness predicted for cold waters was approximately 20% lower than for trawls from waters with average or warm temperatures, but in shallow waters this trend reverses, with richness predicted to be up to 50% higher in cool (southern) waters than in warm (northern) waters (Fig. 5a). Similarly, species richness was predicted to be higher in waters of lower-than-expected salinity (Fig. 5b), where it also showed a more pronounced peak of maximum richness at depths of around 800 m. Predicted species richness also increased with increasing chl *a* concentrations (Fig. 5c), although interaction effects resulted in this response being muted in shallow waters. Finally, richness was predicted to be higher in areas of mixing of different water bodies, as indicated by high values for *SstGrad* (Fig. 5d), but again this was pronounced only in depths greater than 400 m.

The constrained function fitted for trawl distance by this model indicated a gradual increase in richness up to a distance of about 3 km, after which it effectively reached a plateau (Fig. 6a). Similarly, richness was predicted to remain relatively constant across a range of finer mesh sizes, but was predicted to be lower with mesh sizes of 100 mm of more (Fig. 6b). Highest species richness was associated with intermediate trawl speeds (Fig. 6c), and the decline in richness with increasing speed is greater in deeper than in shallow water.

Spatial predictions of species richness

Spatial predictions derived from the BRT model with a tree size of 5 and using environment and trawl characteristics as predictors, are shown in Fig. 7a, along with an estimate of uncertainty (the width of the 5 to 95% confidence intervals estimated using bootstrap re-sampling, Fig. 7b). While the predicted species richness ranged from 3.9 to 33 species per trawl, estimated confidence interval widths were 3 or less over approximately 80% of the area for which predictions are made, but reached moderate levels of uncertainty (3 to 5) both in shallow waters around the South Island and southern North Island coast, and towards the shelf slope in the north. Wide confidence intervals occurred mostly in offshore regions that are inadequately sampled (cf. Figs. 1 & 7b), e.g. steep slopes around the margins of the Campbell and Bounty plateaus, on the Challenger Plateau, off the northeast North Island, and along the Kermadec and Norfolk ridges.

In geographic terms, highest species richness was predicted to occur along the northern flanks of the Chatham Rise and around the northern end of the Solander Trough (Fig. 7a). These are locations that combine depths of 800 to 1000 m with high primary productivity associated with the mixing of subtropical and subantarctic waters along the subtropical front. High richness was also predicted for Tasman Bay, and a narrow strip of water around the continental slope off the coast of Westland, Otago, and from Kaikoura north along the east coast of New Zealand to the Bay of Plenty. Moderately high species richness was predicted for large areas with depths between 500 m and

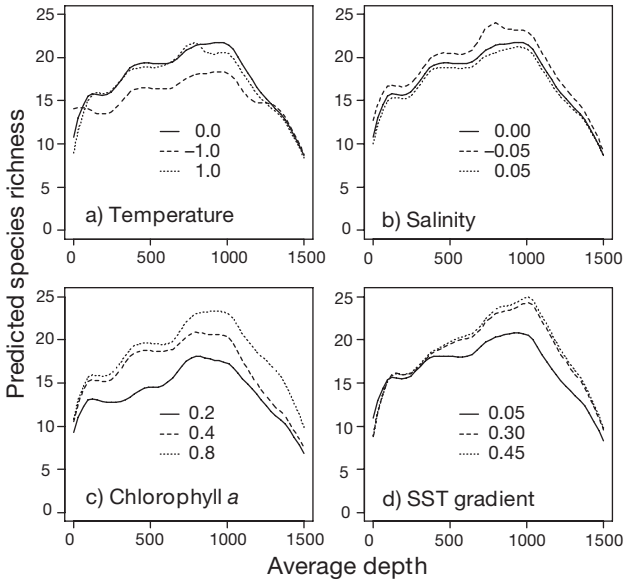

Fig. 5. Variation in demersal fish species richness in relation to depth, given variation in other individual environmental variables, as predicted from a BRT model with a tree size of 5. Values for variables held constant were set at their mean for the trawl dataset. Individual figure legends indicate step-wise changes in the varied environmental predictor. See Table 1 for units

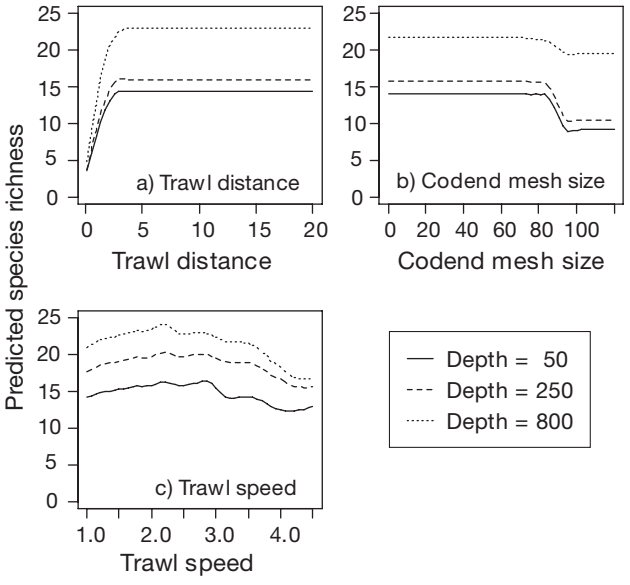

Fig. 6. Predicted variation in demersal fish species richness in relation to trawl parameters, given variation in other individual environmental variables, as predicted from a BRT model with a tree size of 5. Values for predictors other than those explicitly varied are held constant at their mean for the trawl data set. Horizontal-axis labels indicate step-wise changes in the varied environmental predictor. See Table 1 for units

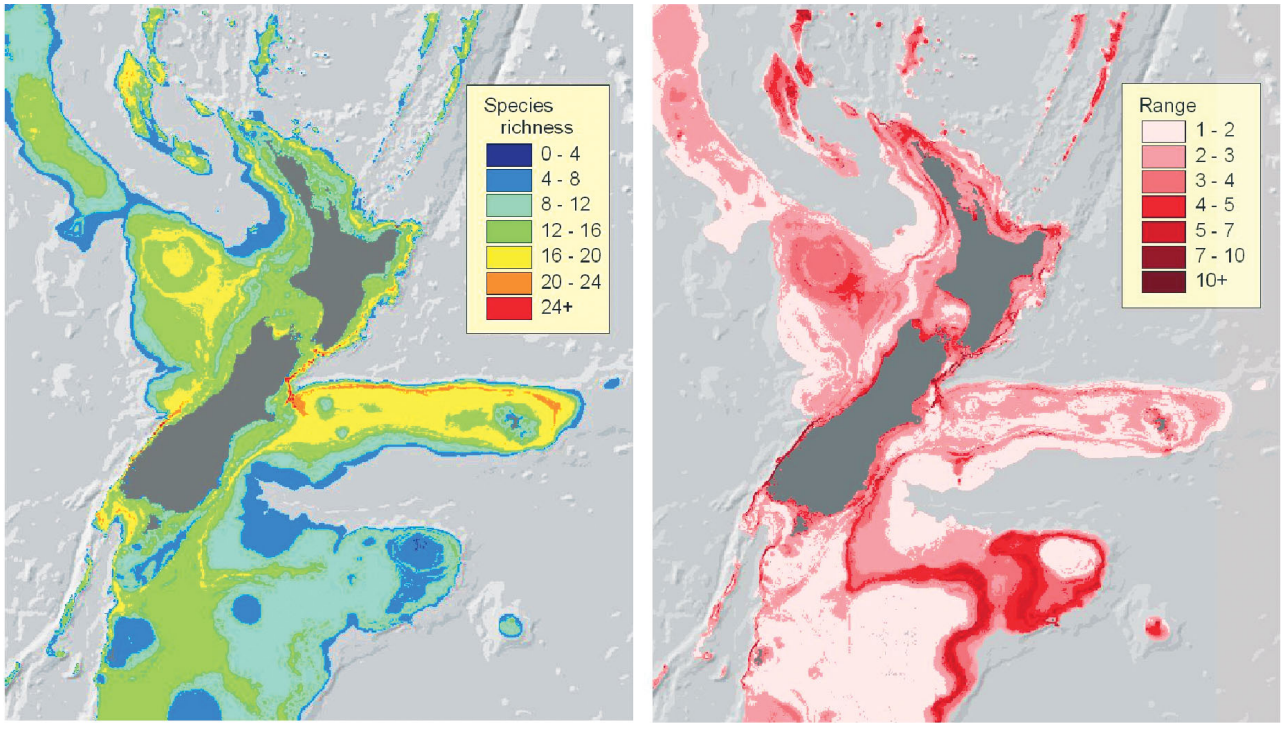

Fig. 7. Demersal fish species richness as predicted from a BRT model using environment and trawl characteristics as predictors, and fitted with a tree size of 5. Confidence intervals were estimated from predictions made from 1000 models fitted to bootstrap samples of the trawl dataset. (a) Predicted species richness, and (b) 5 to 95% confidence interval

1000 m on the Chatham Rise and Challenger Plateau, but richness was predicted to decrease with progression to greater depths, averaging a little over 6 species per trawl at depths of 1500 m across a wide geographic range. However, in the shallow waters of the continental shelf there was a little more variation, with lower richness predicted for waters north of about 38° S than for the continental-shelf waters around the east and south of the South Island. Richness predicted for the shallow waters around the subantarctic islands was low, while the shallow waters around the Chatham Islands were predicted to have similar richness to those around central New Zealand.

DISCUSSION

Links between species richness and environment

Our results indicate that there is a high level of predictability in the relationship between demersal fish species richness and environment , with depth, factors related to productivity, and temperature (both as temperature anomalies and as SST gradient) identified as the most important predictors. Depth is the strongest predictor of species richness, with predicted richness peaking at mid depths (800 to 1000 m) before tailing off in abyssal waters, a pattern that is consistent both with results from several site-specific studies elsewhere (e.g. Colloca et al. 2003) and with the more general description of Levington (1995). However, at any given depth there is a strong positive association between species richness and primary productivity, with maximum richness concentrated in the zone of high productivity associated with the mixing of waters of subtropical and subantarctic origins along the STF. Surface concentrations of chl *a* had the second highest environmental contribution to overall model outcomes (ignoring trawl characteristics). These are closely linked to primary productivity, and there are also strong functional links between surface primary productivity and biological activity at the sea floor through the episodic deposition of particulate material. This is demonstrated in studies both in the waters surrounding New Zealand, including the Chatham Rise (e.g. Nodder et al. 2003), and in other global locations (e.g. Honjo et al. 1995, Beet et al. 2001). In addition, declines in the amount of particulate matter reaching the sea floor in abyssal regions have been suggested as a likely explanation of the declines in species richness there (Gray 2002).

Regional-scale variation in salinity played a more muted role, as indicated by the smaller contributions of the residuals that describe departures from the overall average relationship with depth. Depth-independent variation in temperature and salinity mostly occurs in waters of shallow-to-moderate depth away from the STF, and these variables provide broad discrimination between waters of subtropical and subantarctic origins. Results indicate that richness is generally lower in warmer and/or more-saline waters than in cooler less-saline waters, provided that the latter occur in reasonable proximity to zones of water mixing. Similar differences in regional-scale richness were described for meso-pelagic organisms in the waters north and south of the STF by Robertson et al. (1978).

A positive association is also indicated between species richness and zones of mixing of water bodies of contrasting origins, with the variable describing sea-surface-temperature gradients making the fourth-highest contribution to model outcomes. In functional terms we interpret this as reflecting the important role these zones play in the concentration of nutrients, productivity and food resources (e.g. Bakun 1996). However, currents also play a role in the horizontal displacement of organic particles during their vertical descent from the surface to the sea floor, which may result in the deposition of food resources at locations away from sites of high surface primary production (Nodder et al. 2003).

Relationships between slope and species richness suggested by McClatchie et al. (1997) are not supported by the results from our analysis. In fact, variables describing slope, tidal currents and wave-induced seabed stress had a low contribution overall, and our models consistently predicted a decline in species richness at sites with steeper slopes. However, this latter result should be regarded with caution, as few samples were available from areas with higher slopes, and variation in substrate has been shown to be an important correlate of varying fish diversity in other studies (e.g. Kendall et al. 2004). In addition, slope probably has only limited ability to act as a surrogate for habitat variation, and while it is easily calculated, inclusion of a variable explicitly describing variation in seabed conditions would have been preferable had reliable data been available across the entire study region.

Spatial variation in species richness

Our analyses revealed large-scale spatial variation in demersal fish species richness in the New Zealand region (Fig. 7). To the extent that comparisons are possible, the patterns we observed were consistent with those from previous, smaller-scale studies in the same region. This is perhaps not surprising given that the earlier studies used some of the same research trawl data that we used. McClatchie et al. (1997) analysed shelf and slope richness from 80 to 898 m depth and 43 to 54° S (Chatham Rise to Campbell Plateau). They found that species-richness 'hotspots' were concen-

trated on the Chatham Rise (particularly the north Chatham Rise) and 'coldspots' were concentrated on the Campbell Plateau. They also noted that richness increased with depth to reach a maximum at 500 to 1000 m. In a further study focusing solely on the Chatham Rise, Bull et al. (2001) observed that mean species richness peaked at 550 to 800 m on the north Chatham Rise, and was lower on the south Chatham Rise, and in depths between 200 and 550 m. Beentjes et al. (2002) reported conflicting depth-related trends in species richness between summer surveys (richness increased with depth) and winter surveys (richness decreased with depth) of Canterbury Bight, but as their study covered limited latitudinal and depth ranges (43 to 46° S and 10 to 400 m respectively), the observed patterns may reflect small-scale seasonal migrations.

In a larger-scale study covering latitudes 35 to 47° S, Tracey et al. (2004) compared species richness between seamounts and adjacent areas of lower-relief seabed. Seamounts showed increasing fish species richness with increasing latitude, with the 3 highest mean richness values being recorded for Chatham Rise seamounts (2 sites) and Puysegur Bank (near Solander Trough, see Fig. 1); both these areas were identified in our study as having high species richness. Tracey et al. (2004) also reported that richness was higher at the adjacent sites than on the seamounts themselves, but their analysis was confounded by other uncontrolled variables, including depth and distance towed.

Our results therefore agree well with those previously reported, but we extend the analysis of fish species richness in the New Zealand region to span 25 degrees of latitude (29 to 54° S) with a high degree of spatial resolution, taking into account a wide range of environmental and operational variables. Nevertheless, our predictions of richness in some areas having few or no trawl stations (notably the submarine ridges north of 34° S and steep slopes around the margins of the Campbell and Bounty plateaus) have wide confidence intervals and require validation.

Use of environmental versus geographic predictors

Our focus on use of environmental predictors for this study contrasts with practices adopted in a number of other recent studies of demersal fish distribution, in which geographic variables such as latitude and longitude are used as predictors, often in combination with environment. This was prompted by the understanding of the relative utility of environmental and geographic predictors developed in terrestrial plant ecology (e.g. Austin 2002), which demonstrates that the contribution of geographic predictors, such as latitude and elevation, is largely derived from their correlations with more proximate physical drivers of biological phenomena. However, we suggest that the utility of latitude as a geographic proxy in marine studies is likely to be low because of its generally lower levels of correlation with environmental variation. While latitudinal gradients in environment, and particularly irradiance, clearly play a significant role at the ocean surface, correlations between latitude and environment generally decrease with progression to greater depths. Here, spatial variation in environmental conditions is frequently complicated by oceanographic processes such as the long-distance movement of water bodies of contrasting temperature and salinity, and the collision of these water bodies produces marked environmental discontinuities (e.g. Bradford-Grieve et al. in press). As a consequence, latitudinal sorting becomes increasingly blurred with progression to greater depths because of the increasing degree of disconnection with surface environmental conditions, and in abyssal waters environmentally homogeneous conditions frequently prevail across wide latitudinal ranges.

We argue therefore that analytical approaches based on functionally relevant environmental factors are crucial in developing a better understanding of geographic variation in species richness, particularly in the Southern Hemisphere, where oceanic circulation is less impeded by extensive landmasses and marked environmental discontinuities are common. However, we also acknowledge that evolutionary influences on variation in species richness are more likely to operate primarily in geographic rather than in environmental space, and these may be important, particularly in analyses conducted over wider geographic ranges than that for this analysis. The detection of such effects is likely to require the development of analyses that allow for a careful partitioning of the relative roles of both environment and geography.

Variation with trawl characteristics

Various approaches have been adopted in other studies to accommodate the effects of differences in the fishing characteristics between vessels, including segregation of data by vessel (e.g. McClatchie et al. 1997), use of fishing power coefficients and/or categorical descriptors of gear type (e.g. Muetter & Norcross 2002), and use of generalised linear mixed models (GLMM) to adjust for the effects of between-vessel differences (e.g. Cooper et al. 2004). Our approach was relatively simplistic compared to these latter two, and ideally we would have calculated the area swept for each trawl, but this was not feasible given the amount of missing data for key trawl parameters. Even though this was not possible, our use of predictors describing trawl dis-

tance, cod-end mesh size and trawl speed contributed substantially to the analysis outcome, with the first of these variables explaining nearly 25% of the variation in the most successful model of species richness.

Our initial analyses, which gave results for cod-end mesh size and trawl distance that were clearly inconsistent with the behaviour of trawl sampling, highlight the care required in analysing large datasets assembled from disparate sources. Inspection of the geographic distribution of trawls using various mesh sizes and distances indicated that these discrepancies are much more likely to have resulted from the very uneven distribution of variation in these predictors with respect to both environment and geography. While regularizing the models to allow only the fitting of monotonic functions reduced the total amount of deviance explained (data not presented), we argue that it allowed a more accurate description of the relationship between richness and environment.

Including a variable describing the year in which trawling occurred would have also been desirable, particularly given the potential to use such an analysis to assess both the long-term impacts of sustained harvesting and the impacts of environmental variation associated with factors such as the El Niño–Southern Oscillation, which has a substantial effect on some aspects of the oceans around New Zealand (e.g. Livingston 2000). However, this was frustrated by 2 factors. First, there is marked variation in the spatial sampling by trawls in different years; systematic coverage is never achieved in any particular year, and several regions have been intensively sampled in only 1 or a few years. Second, an exploratory model fitted using year as a predictor indicated a slight but gradual increase in richness with time, a result that we attribute not only to the greater frequency of trawls in deeper waters in later years, but also to an increase in taxonomic knowledge of demersal fish and a greater interest in non-commercial species. Such an effect was also noted by Shackell & Frank (2003) in their analysis using a trawl database in which sampling extended over a lengthy period. While this result does not preclude future use of trawl data to monitor changes in fish species richness, it highlights the need for consistency of data collection in any ongoing trawl surveys likely to be used for long-term monitoring.

Analytical considerations

Results from our analysis provide a clear demonstration of the ability of BRT to outperform substantially conventional regression models such as GAMs. The progressive improvement in the relative performance of BRT as the size of the individual trees is increased indicates that several factors contribute to this improvement. First, the performance gains in BRT models fitted with a tree size of 1 (= no interaction effects) suggests that this method has greater flexibility in describing data complexities than in a GAM. This probably reflects the effectiveness of the strategy used in boosting, i.e. fitting successive models that are progressively adapted to explain cases poorly predicted by the preceding models, compared with the approach used in a GAM of fitting a single most parsimonious model. As a caution though, we note that the discrepancies between these 2 models may also reflect the greater ease with which a monotone function could be specified in the BRT model, as standard GAM and boosted models fitted without such restrictions matched each other in predictive performance much more closely. Second, the greater performance gains in the BRT model fitted with a tree size of 2 compared with the interaction GAM model suggest that boosting offers flexibility in interaction fitting that is far more practical than the fitting of interactions in a GAM. A wide range of interactions were automatically identified and fitted, whereas the manual fitting of interactions in the comparable GAM model was both tedious and computationally constrained. Finally, the BRT model using a tree size of 5 delivered further improvements in predictive performance, fitting a wide array of interactions in a manner not achievable with a GAM or similar model. Results from this model suggested a continued rise in ecological interpretability, particularly with respect to the relationship between species richness and depth, which varied depending on values taken by other variables.

CONCLUSION

Our analysis indicates that, while there are strong associations between species richness and depth, high species richness is also associated with areas of high primary productivity, as indicated both by surface chl *a* concentrations and zones of mixing of water bodies of contrasting origin associated with the Subtropical Front. Use of results such as these as a baseline for longer-term monitoring of the status of New Zealand's oceanic resources is feasible, but it would probably require consideration of the effects of inter-annual variation in environment, as well as long-term means. Boosted regression trees appear to offer considerable performance gains over conventional regression techniques, and a large part of this gain is attributable to their capability for fitting interactions among predictor variables. However, because of their tendency to overfit, care should be exercised both in fitting such models and in reporting on their success.

Acknowledgements. The trawl data used in this study came from the New Zealand Ministry of Fisheries' research trawl database (*trawl*). The data were extensively groomed and extended with other attributes to generate another database (*fish_comm*) during a NIWA study of New Zealand fish-assemblage composition funded by the Foundation for Research Science and Technology (FRST). We thank the Ministry of Fisheries for access to the data. G. Ridgeway kindly gave advice on the use of his GBM package. T. H. Snelder and M. A. Weatherhead facilitated access to the environmental data layers, and G. McBride critically reviewed the manuscript.

LITERATURE CITED

Austin MP (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. Ecol Model 157:101–118

Bakun A (1996) Patterns in the ocean: ocean processes and marine population dynamics. California Sea Grant College System. University of California, La Jolla, CA

Beentjes MP, Bull B, Hurst RJ, Bagley NW (2002) Demersal fish assemblages along the continental shelf and upper slope of the east coast of the South Island, New Zealand. NZ J Mar Freshw Res 36:197–223

Beet BJ, Malzone MG, Narayanaswamy BE, Wigham BD (2001) Temporal variability in phytodetritus and megabenthic activity at the seabed in the deep Northeast Atlantic. Prog Oceanogr 20:349–368

Bradford-Grieve JM, Lewis KB, Stanton BR (1991) Advances in New Zealand oceanography, 1967–91. NZ J Mar Freshw Res 25:429–441

Bradford-Grieve J, Probert K, Lewis K, Sutton P, Zeldis J, Orpin A (in press) New Zealand shelf region. In: Robinson A, Brink H (eds) The sea, Vol 14: The global coastal ocean: interdisciplinary regional studies and syntheses, Chap 36. Harvard University Press, Cambridge, MA

Bull B, Livingston ME, Hurst RJ, Bagley NW (2001) Upper-slope fish communities on the Chatham Rise, New Zealand, 1992–99. NZ J Mar Freshw Res 35:795–815

Cappo M, De'ath G, Boyle S, Aumend J, Olbrich R, Hoedt F, Perna C, Brunskill G (2005) Development of a robust classifier of freshwater residence in barramundi (*Lates calcarifer*) life histories using elemental ratios in scales and boosted regression trees. Mar Freshw Res 56:713–723

Clarke A (1992) Is there a latitudinal diversity cline in the sea? Trends Ecol Evol 7:286–287

Colloca F, Cardinale M, Belluscio A, Ardizzone G (2003) Pattern of distribution and diversity of demersal assemblages in the central Mediterranean Sea. Estuar Coast Shelf Sci 56:469–480

Cooper AB, Rosenberg AA, Stefánsson G, Mangel M (2004) Examining the importance of consistency in multi-vessel trawl survey design based on the U.S. west coast groundfish bottom trawl survey. Fish Res 70:239–250

Diaz RJ, Solan M, Valente RM (2004) A review of approaches for classifying benthic habitats and evaluating habitat quality. J Environ Manage 73:165–181

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman and Hall, London

Elith J, Graham CH, NCEAS Modeling Group (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151

Ferrier S, Watson G, Pearce J, Drielsma M (2002) Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. Biodivers Conserv 11:2275–2307

Francis, MP, Hurst RJ, McArdle BH, Bagley NW, Anderson OF (2002) New Zealand demersal fish assemblages. Environ Biol Fish 65:215–234

Friedman JH (2001) Greedy function approximation: the gradient boosting machine. Annal Stat 29:1189–1232

Friedman JH, Meulman JJ (2003) Multiple additive regression trees with application in epidemiology. Stat Med 22: 1365–1381

Friedman JH, Hastie T, Tibshirani R (2000) Additive logistic regression: a statistical view of boosting. Annal Stat 28:337–407

Gorman RM, Bryan KR, Laing AK (2003) A wave hindcast for the New Zealand region — deep water wave climate. NZ J Mar Freshw Res 37:589–612

Grassle JF, Maciolek NJ (1992) Deep-sea species richness: regional and local diversity estimates from quantitative bottom samples. Am Nat 139:313–341

Gray JS (2001) Marine diversity: the paradigms in patterns of species richness examined. Sci Mar 65(Suppl 2):41–56

Gray JS (2002) Species richness of marine soft sediments. Mar Ecol Prog Ser 244:285–297

Gray JS, Poore GCB, Ugland KI, Wilson RS, Olsgard F, Johannessen Ø (1997) Coastal and deep-sea benthic diversities compared. Mar Ecol Prog Ser 159:97–103

Guisan A, Zimmerman NE (2000) Predictive habitat distribution models in ecology. Ecol Model 135:147–186

Hastie T, Tibshirani RJ (1990) Generalized additive models. Chapman and Hall, London

Hastie T, Tibshirani R, Friedman JH (2001) The elements of statistical learning: data mining, inference, and prediction. Springer-Verlag, New York

Heath RA (1985) A review of the physical oceanography of the seas around New Zealand – 1982. NZ J Mar Freshw Res 19:79–124

Honjo S, Dymond J, Collier R, Manganini SJ (1995) Export production of particles to the interior of the equatorial Pacific Ocean during the 1992 EQPAC experiment. Deep-Sea Res II 42:831–870

Huston M (1994) Biological diversity: the coexistence of species in changing landscapes. Cambridge University Press, Cambridge

Kawakita M, Minami M, Eguchi S, Lennert-Cody CE (2005) An introduction to the predictive technique AdaBoost with a comparison to generalized additive models. Fish Res 76: 328–343

Kendall MS, Christensen JD, Caldow C, Coyne C, Jeffrey C, Monaco ME, Morrison W, Hillis-Starr Z (2004) The influence of bottom type and shelf position on biodiversity of tropical fish inside a recently enlarged marine reserve. Aquat Conserv 14:113–132

Kendrick TH, Francis MP (2002) Fish assemblages in the Hauraki Gulf. NZ J Mar Freshw Res 36:699–717

Kuhnert PM, Mengersen K, Tesar P (2003) Bridging the gap between different statistical approaches: an integrated framework for modelling. Int Stat Rev 71:335–368

Lawrence R, Bunn A, Powell S, Zambon M (2004) Classification of remotely sensed imagery using stochastic gradient boosting as a refinement of classification tree analysis. Remote Sensing Environ 90:331–336

Levington JS (1995) Marine biology. Function, biodiversity, ecology. Oxford University Press, New York

Livingston ME (2000) Links between climate variation and the year class strength of New Zealand hoki (*Macruronus novaezelandiae*) Hector. NZ J Mar Freshw Res 34:55–69

McClatchie S, Millar RB, Webster F, Lester PJ, Hurst R, Bagley N (1997) Demersal fish community diversity off New Zealand: is it related to depth, latitude and regional surface phytoplankton? Deep-Sea Res I 44:647–667

Muetter FJ, Norcross BL (2002) Spatial and temporal patterns in the demersal fish community on the shelf and upper slope regions of the Gulf of Alaska. Fish Bull 100: 559–581

Murphy RJ, Pinkerton MH, Richardson KM, Bradford-Grieve JM, Boyd P (2001) Phytoplankton distributions around New Zealand derived from SeaWiFS remotely-sensed ocean colour data. NZ J Mar Freshw Res 35: 343–362

Nodder SD, Pilditch CA, Probert PK, Jall JA (2003) Variability in benthic biomass and activity beneath the Subtropical Front, Chatham Rise, SW Pacific Ocean. Deep-Sea Res I 50:959–985

Olden JD, Jackson DA (2002) A comparison of statistical approaches for modeling fish species distributions. Freshw Biol 47:1976–1995

Ponder WF, Carter GA, Flemons P, Chapman RR (2001) Evaluation of museum collection data for use in biodiversity assessment. Conserv Biol 15:648–657

R Development Core Team (2004) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rex MA, Stuart CT, Hessler RR, Allen JA, Sanders HL, Wilson GDF (1993) Global-scale patterns of species diversity in the deep-sea benthos. Nature 365:636–639

Ridgeway G (2004) gbm: generalized boosted regression models. R package, version 1.3-5. Available at: http://www.i-pensieri.com/gregr/gbm.shtml

Ridgway KR, Dunn JR Wilkin JL (2001) Ocean interpolation by four-dimensional least squares—application to the waters around Australia. J Atmos Ocean Technol 19:1357–1375

Robertson DA, Roberts PE, Wilson JB (1978) Mesopelagic faunal transition across the Subtropical Convergence east of New Zealand. NZ J Mar Freshw Res 12:295–312

Rohde K (1992) Latitudinal gradients in species diversity: the search for the primary cause. Oikos 65:514–527

Roy K, Jablonski D, Valentine JW, Rosenburg G (1998) Marine latitudinal diversity gradients: tests of causal hypotheses. Proc Natl Acad Sci USA 95:3699–3702

Shackell NL, Frank KT (2003) Marine fish diversity on the Scotian Shelf, Canada. Aquat Conserv 13:305–321

Snelder TH, Leatherwick JR, Dey K, Weatherhead MA and 11 others (2006) Development of an ecological marine classification in the New Zealand region. Environ Manag (in press)

Solbrig OT (ed) (1991) From genes to ecosystems: a research agenda for biodiversity. IUBS, Paris

Stehli FG, McAlester AL, Helsey CE (1967) Taxonomic diversity of recent bivalves and some implications for geology. Geol Soc Am Bull 78:455–466

Tracey DM, Bull B, Clark MR, Mackay KA (2004) Fish species composition on seamounts and adjacent slope in New Zealand waters. NZ J Mar Freshw Res 38:163–182

Uddstrom MJ, Oien NA (1999) On the use of high-resolution satellite data to describe the spatial and temporal variability of sea surface temperatures in the New Zealand region. J Geophys Res 104:20729–20751

Walters RA, Goring DG, Bell RG (2001) Ocean tides around New Zealand. NZ J Mar Freshw Res 35:567–579

Ward TJ, Venderklift MA, Nicholls AO, Kenchington RA (1999) Selecting marine reserves using habitat and species assemblages as surrogates for biological diversity. Ecol Appl 9:691–698

Whittaker RH (1972) Evolution and the measurement of species diversity. Taxon 21:213–251

Appendix 1. Robust fitting and evaluation of BRT models

This appendix describes methods developed to allow the robust fitting of BRT models, reflecting both the challenges posed by their propensity to over-fit and the potential gains they offer given their automatic fitting of interactions between predictor variables. Several of these techniques reflect our belief that evaluation of model performance using the data used to fit the model is misleading. Even though many models can be made to perform well on their training data, the danger is that they over-fit to specific features of that data that lack applicability in a wider sample, degrading model performance when predicting to new sites. We prefer to assess model performance using independent sites, and this can be achieved using a number of strategies. Partitioning the data into separate modelling and evaluation subsets is one alternative, although it involves a loss of information, particularly with smaller datasets. While we used this for initial testing, for most of our model fitting and evaluation we used *k*-fold cross-validation (e.g. Hastie et al. 2001) or bootstrap re-sampling (e.g. Efron & Tibshirani 1993). These alternative approaches allow the use of all available information, while using subsets of the data to estimate model performance when predicting to independent sites.

Choice of learning rate

To establish a suitable value for the learning rate used in fitting BRT models, we carried out an initial evaluation of the relationship between learning rate and model predictive performance with a script that used standard options in the 'gbm' library functions. We achieved this by fitting a series of models to randomly selected subsets of 70% of the available data, with trees successively added until no further improvement in prediction could be detected when predicting to the 30% of the data that were withheld. Models were fitted with learning rates of 0.5, 0.1, 0.05, 0.01 and 0.005, with the number of trees fitted increasing steadily as the learning rate was decreased. Results indicated a progressive improvement in prediction performance as the learning rate decreased from 0.5 to 0.05, with the latter value typically resulting in 800 to 1000 trees being fitted. Use of learning rates smaller than 0.05 not only brought about minimal improvement in predictive performance, but also increased substantially the computational requirements.

Setting model complexity

Because of the risks of over-fitting when using BRT, we explored a number of options for identifying the optimal number of trees to include in a model, including options provided as part of the 'gbm' library. The method we identified as the most consistent and computationally efficient was based on a *k*-fold cross-validation procedure described by Hastie et al. (2001, Chap. 7), which we implemented using a purpose written script.

In this procedure, the data available for model fitting were first divided into 10 mutually exclusive subsets, selected using a randomisation procedure. Ten models were then fitted simultaneously, each using a different subset of the total

Appendix 1 (continued)

data (= training data) containing 90% of the total data set, and formed by omitting 1 of the 10 subsets. A step-wise procedure was then used to gradually increase the complexity of all 10 models, typically by adding groups of 100 trees. At each step, predictions were formed from each model for the 10% subset of data omitted from its training data (= evaluation data), and the residual deviance was calculated to compare the goodness of fit between the model predictions and the species richness as recorded in the evaluation data.

Results typically indicated an initial decline in prediction error (residual deviance) as more trees were added, but with most models a point was eventually reached where, even though the training error continued to decline, the prediction error would begin to rise as the model became excessively adapted to the training data, i.e. over-fitting occurred. At this point the mean prediction errors and their standard errors (estimated from the 10 subsets) were plotted as a function of the number of trees fitted (Fig. A1), and this graph was used to determine the lowest number of trees giving a prediction error equal to or less than 1 SE above the best model (see Hastie et al. 2001). This number of trees was then used in a model fitted to the entire dataset with the required learning rate and tree size.

Assessing model performance

The performance of both GAM and BRT models was evaluated using *k*-fold cross-validation to estimate their predictive ability with new data. Using a procedure similar to that used to estimate the optimal tree size for the BRT models, we wrote a script in which the input data were divided into 10 mutually exclusive subsets that were omitted in turn. At each iteration, a model was fitted to the retained data, predictions were made for the omitted data, and the residual deviance was calculated as a measure of the correspondence between measured and predicted richness. The mean and standard error were then calculated for each of these 10 estimates of predictive performance. Because results from *k*-fold cross-validation can vary depending on the random selection of points for the folds, this procedure was repeated 5 times for each model, and overall means were calculated for the mean prediction error and its standard error.

Display of fitted functions

Relationships between species richness and environment fitted by the both the GAM and BRT models were displayed by plotting the fitted relationship for each individual predictor. Values for plotting were calculated by setting values for all but 1 predictor to their mean. Predictions were then formed for points along the range of the remaining variable using a purpose written script. As a BRT model provides no estimate of the confidence intervals around these fitted functions, we estimated these by taking 1000 bootstrap samples of the input data, i.e. a sample of equivalent size to the trawl dataset, but selected randomly with replacement A GAM or BRT model was fitted to each sample, and predictions were formed for each predictor and accumulated. Five and 95 percentile values were calculated for points along the range of each function from the accumulated values. The complexity of interactions fitted by the BRT models with tree size greater than 1 made display of their fitted relationships more challenging, as these can vary depending on the values assigned to other predictors. We therefore wrote our own scripts in R to calculate and graph values predicted in relation to major variables, while other variables were either held constant or varied in steps.

Detection and interpretation of interactions

For BRT models with tree size greater than 1, we assessed the magnitude of interaction effects using a purpose-written script that examined the relationship between the model predictions and all possible pair-wise combinations of predictors. This was achieved by selecting each possible pair-wise combination of predictors in turn. For each pair of predictors, 2 variables (x_1, x_2) were created that consisted of values at constant intervals along the ranges of the 2 predictors, and predictions on the linear predictor scale (y') were calculated from the BRT model for all possible combinations of these. In making these predictions, values for the remaining predictors were set at their mean for the dataset. We then used a linear model to relate these predicted values to the values of the 2 marginal variables, i.e. $y' \sim x_1 + x_2$, with the 2 predictor variables fitted as factors. Where the predicted values are formed by a purely additive combination of the 2 predictors, this regression object will have zero residual variance. However, as stronger interaction effects for the 2 predictors are fitted in the BRT model, so the variance in y' left unexplained by the test linear model increases. Thus the amount of residual variance in the test linear model can be used as a direct indication of the strength of the interaction effect fitted by the BRT model for that pair of predictors.

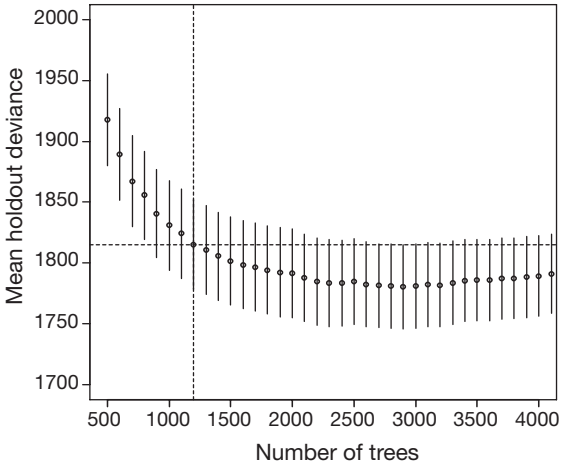

Fig. A1. Relationship between model complexity and predictive error for a boosted regression tree model relating species richness to environment and trawl characteristics calculated using a 10-fold cross-validation procedure. Circles indicate the mean predictive error averaged across 10 iterations for each level of model complexity, with standard errors indicated by vertical lines. Dashed horizontal line indicates the minimum mean predictive error plus 1 SE, and the vertical dashed line indicates the model complexity with predictive error equal to the minimum predictive error plus 1 SE

Editorial responsibility: John S. Gray (Contributing Editor), Oslo, Norway

Submitted: August 22, 2005; Accepted: February 28, 2006
Proofs received from author(s): August 25, 2006