Title: Continental-scale hotspots of pelagic fish abundance inferred from commercial catch records

Article type: Research paper

Authors: Phil J. Bouchet, Jessica J. Meeuwig, Zhi Huang, Tom B. Letessier, Scott L. Nichol, M. Julian Caley, Reg A. Watson

Correspondence:

Phil J. Bouchet (corresponding author)
Centre for Marine Futures & UWA Oceans Institute
School of Biological Sciences, University of Western Australia, 6009 Crawley, Australia.
Mailing address: Centre for Marine Futures M092, The University of Western Australia, 35 Stirling highway, 6009 Crawley, WA (Australia).
Tel: +61 (8) 6488 3590
Email: phil.bouchet@uwa.edu.au
Jessica J. Meeuwig
Centre for Marine Futures & UWA Oceans Institute
School of Biological Sciences, University of Western Australia, 6009 Crawley, Australia.
Email: jessica.meeuwig@uwa.edu.au
Zhi Huang
Geoscience Australia, GPO Box 378, 2601 Canberra, Australia.
Email: Zhi.Huang@ga.gov.au
Tom B. Letessier
Institute of Zoology, Zoological Society of London (ZSL), NW1 4RY London, United Kingdom & Centre for Marine Futures, The University of Western Australia, 6009 Crawley, Australia.
Email: Tom.Letessier@IOZ.ac.uk
Scott L. Nichol
Geoscience Australia, GPO Box 378, 2601 Canberra, Australia.
Email: Scott.Nichol@ga.gov.au

M. Julian Caley

School of Mathematical Sciences, Queensland University of Technology, 4000 Brisbane, Queensland, Australia & Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers.

Email: julian.caley@gmail.com

Reg A. Watson

Institute for Marine and Antarctic Studies, University of Tasmania, 7001 Taroona, Tasmania, Australia & Centre for Marine Socioecology, Private Bag 129, Hobart TAS 7001, Australia

Email: R.A.Watson@utas.edu.au

Keywords: Submarine canyons, marine protected areas, fisheries, pelagic fish, hotspot, topography, geomorphometrics, spatial modelling, random forest, catch rate standardisation.

Running title: Mapping pelagic fish hotspots around Western Australia

Corresponding author: Phil Bouchet

Word count: Abstract (290 words); Main text (5,281 words).

Reference count: References (90).

ABSTRACT

Aim: Protected areas have become pivotal to the modern conservation planning toolbox but a limited understanding of marine macroecology is hampering their efficient design and implementation in pelagic environments. We explored the respective contributions of environmental factors and human impacts in capturing the distribution of an assemblage of commercially valuable, large-bodied, open-water predators (tunas, marlins, and mackerels).

Location: Western Australia.

Time period: 1997-2006.

Major taxa studied: Pelagic fishes.

Methods: We compiled ten years of commercial fishing records from the *Sea Around Us Project* and derived relative abundance indices from standardised catch rates while accounting for confounding effects of effort, year, and gear type. We used these indices to map pelagic hotspots over a 0.5°-resolution grid, and built random forests to estimate the importance of 33 geophysical, oceanographic and anthropogenic predictors in explaining their locations. We additionally examined the spatial congruence between these hotspots and an extensive network of marine reserves, and determined whether patterns of co-occurrence deviated from random expectations using null model simulations.

Results: (1) We identify several pelagic hotspots off the coast of Western Australia. (2) Geomorphometrics explained up to 50% of the variance in relative abundance of pelagic fishes, and submarine canyon presence ranked as the most influential variable in the North bioregion. Seafloor complexity, geodiversity, salinity, temperature variability, primary production, ocean energy, current regimes and human impacts were also identified as important predictors. (3) Spatial overlap between hotspots and marine reserves was limited, with most high-abundance areas primarily found in zones where anthropogenic activities are subject to few regulations.

Main conclusions: This study reveals geomorphometrics as valuable indicators of the distribution of mobile fish species and highlights the relevance of harnessing static topography as a key element in any blueprint for ocean zoning and spatial management.

INTRODUCTION

The past decades have seen unprecedented changes in the abundance of marine living resources. Despite divergent perspectives on the present status and future prospects of the world's fisheries (Daan *et al.*, 2011; Froese *et al.*, 2012), overexploitation since the 1950s is widely accepted as a catalyst of modern declines in elasmobranchs (sharks, skates and rays) and teleosts globally (Letessier *et al.*, 2017). Although efforts to rebuild fisheries are now underway in several ecosystems (Worm *et al.*, 2009), their success to date remains relatively modest as insufficient control on current exploitation rates often precludes recovery, even for resilient stocks that may have adapted to moderate levels of extractive pressure (Neubauer *et al.*, 2013).

In this context, a new generation of multilateral environmental treaties has emerged to reverse the large-scale erosion of biodiversity. The United Nations Convention on Biological Diversity (CBD, http://www.cbd.int/) is perhaps one of the best-known examples, and presently bind 196 countries to take legislative and policy action to tackle the loss of threatened species. A key element of the convention's strategic plan is Aichi Target 11, which commits to expand the coverage of marine reserves (MRs) to at least 10% of the world's ocean by 2020. While the ecological and socio-economic merits of MRs are well established in coastal systems (e.g. Angulo-Valdés & Hatcher, 2010; Kerwath *et al.*, 2013), their utility in pelagic environments remains more contentious, in part due to the perception that mobile species require protection over too large a geographic space to be logistically, politically or financially practical to implement or enforce (Kaplan *et al.*, 2010). However, such assumptions rarely account for heterogeneous population structuring (Grewe *et al.*, 2015), partial migration, residency, site fidelity, philopatry (Chapman *et al.*, 2015), evolutionary reductions in mobility (Mee *et al.*, 2017), and predictable aggregative behaviour in upper-trophic level organisms (e.g. Kessel *et al.*, 2014). There is now growing consensus that even mobile predators with extensive home ranges such as seabirds (Young *et al.*, 2015), cetaceans (Gormley *et al.*, 2012), turtles (Scott *et al.*, 2012), fishes (Kerwath *et al.*, 2009) and sharks (White *et al.*, 2017) can benefit from spatial closures (Breen *et al.*, 2015), provided mortality rates do not rise disproportionately in adjacent unprotected waters as a result of effort displacement (Powers & Abeare, 2009). This is especially true where reserves strategically encompass core habitat areas or sites in which key life-cycle events perennially occur (Hooker *et al.*, 2011).

Sparse ecological data in remote offshore waters (i.e. beyond the territorial sea baseline, both within EEZs and beyond; Webb *et al.*, 2010) make the direct identification of critical habitat difficult, and call for robust surrogates of biological diversity to predict, delineate and prioritise candidate sites for zoning (McArthur *et al.*, 2010). Complex topography has been recognised as a determinant of wildlife dynamics across numerous taxa (Bouchet *et al.*, 2015), which suggests that locating protected areas over sites of rugged terrain could yield a range of conservation gains (Harris & Whiteway, 2009; Michael *et al.*, 2014). Worm et al. (2003), Morato et al. (2008) and Morato *et al.* (2010) illustrated this possibility in the open ocean by showing that North Atlantic and Pacific seamounts were important centres of taxonomic richness of special interest for the management of threatened vertebrates. Nevertheless, whether static topography could be widely used as a key element in ocean planning remains to be investigated in other ocean basins and for different types of geomorphologies. Submarine canyons, for instance, are prominent and commonly occurring physical seabed features throughout the world's oceans. While their role as drivers of primary productivity, plankton abundance and benthic biomass is well documented (Fernandez-Arcaya *et al.*, 2017), our understanding of their importance to pelagic megafauna remains nascent, and somewhat skewed towards marine mammals (e.g. Moors-Murphy, 2014).

Here, we combine long-term landings and fishing effort datasets from the *Sea Around Us Project* (SAUP, http://www.seaaroundus.org) with a recent reclassification of Australian submarine canyons (Huang *et al.*, 2014) to (1) determine the location of abundance hotspots for a suite of commercially important predatory pelagic fishes within the western part of Australia's exclusive economic zone (hereafter 'wEEZ'); (2) examine associations between these hotspots and an array of abiotic variables including seabed topography on a continental scale; and (3) assess their spatial congruence with a subset of Australia's 3.1 million km^2 national network of Commonwealth Marine Reserves (http://www.environment.gov.au/topics/marine/marine-reserves).

METHODS

Fish and environmental data

Annual commercial fishing records for the wEEZ were extracted from the quality-checked databases compiled by SAUP, but did not include the recent catch reconstructions that capture estimates of recreational and illegal fishing (Pauly & Zeller, 2016). These data are global, readily accessible (http://www.seaaroundus.org/data/), transparent, vetted through extensive peer-review, and were adopted as the only spatially resolved dataset available that currently allowed analyses at macro-ecological scales. Landings (in tonnes) originated primarily from yearly catch reports (corrected for discarded bycatch) produced by the Food and Agriculture Organization (FAO) (Watson *et al.*, 2005; see Appendix S1 in Supplementary Information). Effort statistics were obtained from an array of public domain sources (Watson *et al.*, 2013), and standardised to a common unit of vessel engine power and operation time (kilowatt sea days, *kwsd*), following Anticamara *et al.* (2011). Both catch and effort were disaggregated into a grid of 0.5° (latitude) x 0.5° (longitude) spatial cells using a rule-based algorithm, a technical description of which is provided elsewhere (Watson *et al.*, 2004; Watson *et al.*, 2013). The full dataset used here spanned the period 1950-2006 and comprised 5,640,222 entries of 111 species and higher taxa, representing a total catch of 3.11 million tonnes extracted over 3.35 million km^2 of the eastern Indian Ocean between 93-129°E and 8-39°S.

We also collated a biophysical dataset composed of 51 variables computed over the same half-degree square cells as the fisheries data (Tables 1 & S1 in Supporting Information). Most of these variables were geomorphometrics (measures of seabed complexity, n=20) assembled from existing archives curated by Geoscience Australia, including a newly revised digital catalogue of submarine canyons (Huang *et al.*, 2014). Recognising that mesoscale hydrographic features attract mobile megafauna (Scales *et al.*, 2014) and that human activities can affect spatial patterns in biodiversity by driving changes in species composition and abundances (Navarro *et al.*, 2015), the dataset also comprised a number of oceanographic factors (n=14; Tables 1 & S1) as well as indices of cumulative anthropogenic threats (n=3; Tables 1 & S1).

Data processing

Our stepwise approach to data preparation and analysis is illustrated in Fig. 1, and a brief description of each step follows below (see Appendix S1 for full details). Analyses were carried out in R 3.3.1 and Matlab 2012a.

Gear matching

Catch and effort were estimated independently. Landing records were segregated by species and fishing method, and were well-resolved both spatially and temporally. In contrast, the effort data were incomplete and reported on the basis of gear rather than taxa. To guarantee compatibility between the two datasets, we re-allocated every fishing gear type to one of five discrete and mutually exclusive categories, namely gillnet (GIL), seine (SEI), trawl (TRW), line (LIN), and miscellaneous (OTH) (Fig. S2). We used this classification key as a common denominator to pair all landings with a corresponding value of effort for any combination of *year x grid cell* using a purpose-built Microsoft Access query.

Filtering

We filtered the data to generate a time series spanning the period 1997-2006, excluding all demersal or bentho-pelagic species caught, and all grid cells situated outside the continental wEEZ (Appendix S1 and Table S2). Catch records for unidentified species were discarded, and those only reported in families and genera were reapportioned to confirmed species in proportion to their relative contribution to the total family- or genus-specific catch. Finally, gears contributing less than 5% of total landings were omitted, effectively removing all benthic fishing from our study area. Clear outliers (see Appendix S1 for definition), including particularly small coastal grid cells and exaggerated estimates of effort, were also excluded (Table S2).

Geographic partitioning

To control for spatial non-stationarity, we subdivided the consolidated data into four contiguous bioregions: North, Gascoyne, West and South. These reflected broadly homogeneous environmental conditions and biological assets (Fig. S3), consistent with management boundaries recognised by the Western Australian Department of Fisheries (http://www.fish.wa.gov.au/).

Imputation of missing effort data

Although the filtered landings achieved full temporal and spatial coverage with a positive catch value for each *grid cell x year* combination, available effort estimates were incomplete and required reconstructing in locations where none existed (Fig. S4). We performed gap-filling on the effort matrix using the *smoothn* package (http://www.biomecardio.com/matlab/smoothn.html) introduced by Garcia (2010), which applies penalised regression to smooth evenly-gridded data in multi-dimensional space. The algorithm was executed on a 3D space-time cube (*year x latitude x longitude*) of effort values, summed across fishing gear types due to data scarcity in some grid cells (Fig. S5). After imputation, effort predictions were re-allocated to all gears proportionally to their original usage rate in (i) each *cell x year* when cell-level data existed or (ii) the bioregion as a whole when cell-level data were missing.

Catch rate standardisation

The use of fisheries-dependent data for ecological inference hinges on standardising the catch-per-unit-effort (CPUE) to control for confounding elements that may affect catchability (Maunder & Punt, 2004). We constructed lognormal generalised linear models (GLMs) to standardise CPUE values for each bioregion (Fig. S6) across *year*, *fishing gear category*, grid cell, and *species body weight* (as a proxy for size-mediated gear selectivity, Appendix S1). A subset of nine plausible models was assessed and competing formulations ranked based on their second-order (corrected) Akaike's information criterion scores (AICc, Table 2). We forced the inclusion of *grid cell* as a factor in all models, but its interaction with *year* was not considered owing to the paucity of records on a per year and cell basis. Following Francis (1999), the β coefficients of the *grid cell* term were scaled to canonical form to guarantee invariance to the choice of baseline level. Scaled values were taken as relative fish abundance indices and calculated as: $\beta_i' = \frac{\beta_i}{\bar{\beta}}$, where $\bar{\beta}$ is the geometric mean of all β_s, defined as $\bar{\beta} = \sqrt[n_j]{\prod_{j=1}^{n_j} \beta_{ij}}$.

Random forests

We gauged quantitative relationships between predictor variables and abundance indices with random forests (Breiman, 2001), using the *party* R package (Strobl *et al.*, 2009) to construct 2,000 unbiased conditional inference trees of fish abundance (on the log scale) in each bioregion (Strobl *et al.*, 2007). For comparison, we developed models from both the full environmental dataset and a reduced set comprising geomorphometrics only. We also assessed the importance of each predictor through unconditional permutation (Strobl *et al.*, 2008), and determined how many top-ranking variables should be designated as truly ‘important’ by implementing a modified version of the recursive selection protocol described by Sabatia and Burkhart (2014) (see Appendix S1 for full details).

Hotspot detection

Bartolino *et al.* (2011b) proposed cumulative relative frequency distributions (CRFDs) as a simple way of delineating spatial hotspots in a continuous parameter of interest (e.g. density, diversity). The method is objective and preferable to the more traditional Getis-Ord Gi* or Moran’s I, which rely heavily on a priori user-selected settings (Bartolino *et al.*, 2011a). We obtained CRFD curves by plotting the relative values of fish abundance indices in each bioregion against their own frequency distribution. The majority of CRFDs derived from biological data should increase monotonically and approach an upper asymptote, such that tangents to the curve can then be interpreted as rates of accumulation, with high-density areas (hotspots) characterised by tangent slopes less than 45°. Instead of empirically approximating tangents from pairs of adjacent points, we fitted local polynomial regressions (LOESS smoothers, obtained from the R package fANCOVA; Wang, 2010) to the CRFDs (Fig. S7), with smoothing span widths chosen according to the bias-corrected AICc. This approach is less sensitive to noise and enables a more robust detection of fish hotspots that reflects fundamental ecological signals rather than data idiosyncrasies.

Spatial overlap between MRs and hotspots

We matched hotspots with MRs and measured their spatial congruence according to Jaccard's similarity coefficient J (Real & Vargas, 1996). We developed null models to determine the probability of obtaining these patterns by chance, with the null expectation that hotspots could occur anywhere within each bioregion. Following Leroux *et al.* (2007), this was achieved by randomly selecting without replacement the same number of grid cells as identified hotspots, calculating J, and reiterating this permutation 10,000 times. We then compared the simulated distribution of Jaccard indices to the corresponding observed values *via* two-sample Mann-Whitney-Wilcoxon tests. The same steps were followed to quantify overlap with MNPs.

Uncertainty propagation

Estimates of variance for relative abundance indices, predictor rankings, hotspot thresholds and Jaccard values were acquired by applying a non-parametric bootstrapping procedure, which resampled the CPUE data randomly and with replacement 100 times.

RESULTS

Fish landings and abundance indices

The catch data consisted of 23 pelagic species (22 teleosts, one elasmobranch), of which 12 are highly migratory (Annex I of the 1982 Convention on the Law of the Sea, Table S3). Mackerels and tunas dominated, with six species (*Katsuwonus pelamis*, skipjack tuna; *Thunnus maccoyii*, southern bluefin tuna; *Thunnus albacares*, yellowfin tuna; *Thunnus obesus*, bigeye tuna; *Scomberomorus commerson*, narrow-banded Spanish mackerel; and *Trachurus declivis*, greenback horse mackerel) making up nearly 75% of all landings over the ten years of the study. Northern fisheries contributed nearly twice as much to total catches (37%) as those operating in other bioregions (18% Gascoyne, 23% West, 22% South).

The GLM formulation that minimised the AICc contained a three-way interaction between *year, fishing gear category,* and *species body weight* (Table 2). Bootstrap catch rate standardisation models explained between 45% and 50% of the deviance (mean adjusted R^2 and 95% percentile confidence interval — North: 0.5 [0.49 - 0.5]; Gascoyne: 0.44 [0.43 - 0.45]; West: 0.48 [0.46 - 0.49]; South: 0.45 [0.44 - 0.47]. Fish abundance was highest in the North bioregion (spatial regional mean ± SD: 15.7 ± 54.1), followed by the South (2.6 ± 4.7), Gascoyne (1.7 ± 1.4) and West (1.6 ± 2.5) in that order (Fig. S8). Model prediction uncertainty, expressed as the coefficient of variation of bootstrap values, showed a reverse pattern, being largest in the South (mean regional CV ± SD: 0.36 ± 0.07), relative to the West (0.31 ± 0.05), Gascoyne (0.27 ± 0.05) and North (0.23 ± 0.06) (Fig. S9).

Pelagic hotspots

Numerous hotspots of fish abundance, and clusters thereof, were consistently identified, with a high frequency of inclusion in bootstrap resamples (Fig. 2). These included areas adjacent to the Scott reefs (14.1°S, 121.8°E), the edge of the Exmouth Plateau and the Argo-Rowley terrace (17°S, 117°E), offshore waters northwest of Barrow Island (20.8°S, 115.4°E), the length of Ningaloo Reef peninsula (22.5°S, 113.5°E) south to Shark Bay (26°S, 114°E), Rottnest Island and the Perth canyon (31.9°S, 115.1°E), Geographe Bay (33.6°S, 115.3°E), and the vicinity of Cape Leeuwin (35°S, 115.1°E), as well as along a significant stretch of the southern coastline from the Bremer Basin (35.5°S, 119.5°E), east to the western half of the Great Australian Bight (32°S, 129°E).

Random forest models built on the full set of biophysical predictors explained between 24-70% of the out-of-bag variance (R^2) in fish abundance (Appendix S1 and Table 3). Geomorphometrics accounted for more than half of this percentage in all cases, and although reduced models suffered a loss of predictive power, performance remained satisfactory, particularly in the North (R^2 in excess of 50%). The number and order of predictors selected as important also varied spatially (Fig. 3 and Table S4). For example, canyon presence in neighbouring cells (CANadj) was the top-ranked parameter in the

North (bootstrap mean rank ± SD: 1.06 ± 0.28), closely followed by fractal dimension (FRD; 2.28 ± 0.55). In contrast, the Gascoyne bioregion was clearly oceanography-driven, with sea surface temperature variability (SSTstd; 1.0 ± 0.0) and annual mean primary production (PP; 2.27 ± 0.51) emerging as the most influential pair. Likewise, the West was dominated by PP (1.79 ± 0.89), ocean energy (L3; 2.01 ± 0.98) and regional circulation regimes (L2; 2.45 ± 0.92), whereas in the South, L3 (1.0 ± 0.0), dominant geomorphic features (FEATdom; 2.0 ± 0.0) and east-west current velocity (CUREW; 3.62 ± 0.59) were most important (Fig. 3). Despite this variation in variable importance, measures of static topographic complexity, including canyon attributes, comprised an average of 34% of the top 10 splitting variables and 43% of all predictors retained after variable selection. Despite lower mean ranks, some metrics exhibited particularly high selection frequencies, e.g. geodiversity (FEATcount; 100% in the North and 75% in the West), the contour index (CI; 91% in the West and the South), and to a lesser extent canyon orientation (CANorient; 58% in the South) or canyon distribution (CANadj; 56% in the West). A small number of predictors including CANadj, SAL, PP, or SSTstd were recurrently important throughout the entire wEEZ. By contrast, the frequency of chlorophyll peak index (FCPI) was of relatively trivial importance, as retained in 17% of bootstrap runs on average in only three of the four bioregions. Two human presence indicators (Hi and Him) were especially dominant (mean rank less than 10 and frequency above 80%) in the North and the South bioregions.

Spatial overlap

Hotspots occupied a mean area of 198,017 (95% percentile CI: [153,398 – 239,884]) km^2 across bioregions (n=77 [60 - 94] cells, i.e. 8.4% of total) and were most abundant in the North (30% [18 - 40] of the cumulative number of hotspots in the entire wEEZ) but most prevalent in Gascoyne (12% [7 - 17] of the number of grid cells in the bioregion). In comparison, MR coverage was 796,110 km^2 (n=316 cells) and was more prominent in the North (32% of the cumulative number of reserve cells in the entire wEEZ) but widest in the West (39% of the number of grid cells in the bioregion). Congruence was highest in the North bioregion, although Jaccard scores remained generally very low everywhere (< 0.2), suggesting that a substantial proportion of the MR network as a whole does not intersect hotspots (Fig. 4). All Mann-Whitney-Wilcoxon tests were significant (null permutation models, $p < 0.01$), and hotspots coincided with reserves less frequently than expected by chance alone in all bioregions but the North. Overlap with marine national parks was substantially lower than with the wider MRs in both the Gascoyne and West bioregions. This trend was only marginally reversed in the North and South.

DISCUSSION

Our study offers quantitative insights into the occurrence patterns of an assemblage of highly mobile, pelagic predatory fishes in the eastern Indian Ocean. Prediction maps revealed several fish hotspots along the northwestern, southwestern and southern continental shelves of Western Australia, broadly consistent with findings from previous research. For example, tuna and billfish species richness peaks in similar parts of the North and Gascoyne bioregions (Trebilco *et al.*, 2011), and analyses of movement behaviour in tiger sharks (*Galeocerdo cuvier*, Ferreira *et al.*, 2015) and pygmy blue whales (*Balaenoptera musculus brevicauda*, Double *et al.*, 2014) indicate that pelagic hotspots coincide with home range cores and areas of distinctly higher occupancy times in both species (Fig. S11). These congruent spatial patterns suggest a potential common basis to hotspot formation across multiple taxa (Bouchet *et al.*, 2015). Higher relative abundance was also inferred at a number of discrete sites, confirming their importance to marine megafauna (Fig. S8). These included the seasonally productive Bremer basin, a foraging ground for white sharks (*Carcharodon carcharias*) and killer whales (*Orcinus orca*), the Albany canyon group, a favourable habitat for sperm whales (*Physeter macrocephalus*) (Johnson *et al.*, 2016), the Recherche Archipelago, the Rowley Shoals, the Ashmore, Scott, Cartier and Ningaloo reefs (Sleeman *et al.*, 2007) or to a lesser degree the waters off Dirk Hartog Island (Letessier *et al.*, 2013).

Importantly, we add to a growing body of literature demonstrating that mobile top predators congregate at discrete and perennial sites (Graham *et al.*, 2016). As international support for expanding the world's marine reserve coverage continues to rise (Singleton & Roberts, 2014), the hotspot concept may become particularly appropriate in guiding long-term MR placement and focusing research attention and resources on regions of persistently high ecological value for mobile species, whilst conferring maximum conservation benefits per dollar invested. Such information will be crucial as designing reserves is notoriously difficult and constrained by the costs of sampling vast volumes of ocean (Letessier *et al.*, 2017) as well as by the necessity to accommodate a broad gamut of socio-economic and geopolitical interests. Thus far, most extant pelagic MRs have been established opportunistically without reliance on well-defined scientific criteria (Roberts, 2000) and/or residually where there is little perceived conflict with resource users (Devillers *et al.*, 2015). In many information-poor settings such as offshore waters, surrogate-based approaches may thus be the only viable option for improvement (McArthur *et al.*, 2010). If so, the main difficulty will then lie in identifying not only a universally accepted operational definition of what constitutes a hotspot, an exercise so far thwarted by mixed interpretations (Bouchet *et al.*, 2015), but also reliable proxies that can predict hotspot locations and possibly their change through time. Both will require the creation of standardised operating procedures that reconcile the numerous ways in which hotspots are measured (Marchese, 2015), importance scores calculated and predictors pruned (Szymczak *et al.*, 2016).

Methodological biases notwithstanding, conflicting results from empirical studies have fuelled controversy about the application of abiotic surrogates to conservation planning scenarios (c.f. Stevens & Connolly, 2004 and Rees *et al.*, 2014). Such disputes likely stem from unresolved questions regarding the spatial and temporal stability of surrogate relationships or the effects of data quality and availability on indicator performance (Mellin, 2015). However, as efforts to test the validity of explanatory variables continue to be made, it will become easier to draw generalisations and identify those that perform consistently better across ecosystems, regions and scales (Sutcliffe *et al.*, 2015). Geomorphometrics have generally remained broadly unvalidated proxies of oceanic biodiversity (though see Worm *et al.*, 2003; Morato *et al.*, 2010) in part because large portions of the seafloor are yet to be mapped and the majority of seamounts, submarine canyons and other prominent features around the world are still poorly explored (Huvenne & Davies, 2014). This suggests that weak correlations between static topography and predator hotspots may, at least historically, more likely reflect sporadic and uneven sampling than the absence of genuine wildlife-habitat relationships.

Syntheses are also beginning to emerge that highlight the value of using geodiversity to prioritise areas for biological conservation (Beier *et al.*, 2015). Here, we have confirmed the value of geomorphometry by showing that it can be a good predictor of fish abundance on a macro-ecological scale, and highlighting several indices of topographic complexity with consistent relationships with pelagic biodiversity. That said, not all geomorphometrics were equally important. For instance, reflecting Huang *et al.* (2014)'s observations that the Australian margin is both physically and morphologically heterogeneous, submarine canyon metrics were outperformed in some bioregions but not others. We see two possible explanations for this. Firstly, the formation and maintenance of open-ocean hotspots may demand a suite of interacting oceanographic and bio-physical forces that are not associated with all canyons or topographic features (e.g. upwellings, fronts, eddies, physical retention of prey; Hazen *et al.*, 2013). Secondly, some canyons may only provide favourable conditions for pelagic fish species episodically. If the latter is true, the relatively coarse temporal and spatial resolution of our data may not have been sufficient to reveal such variable relationships. This could be the case for canyon heads, which were not retained as important predictors of hotspots in our analysis but which are frequently reported to be sustained by cyclical upwelling events (Rennie *et al.*, 2009). Similarly, the Oceanic Shoals Commonwealth Marine Reserve (11.5°S, 128.5°E) was here identified as a cold spot of fish abundance despite records of seasonally elevated pelagic diversity (Nichol *et al.*, 2013).

In order to fully explore the hotspot spectrum, and more robustly prioritise candidate areas for protection, a conceptual shift is warranted whereby hotspots are no longer defined in purely geographical terms, but are rather mapped in the four dimensions of latitude, longitude, time and depth, with explicit evaluations of their levels of intra/inter-annual variability (Santora & Veit, 2013). Despite being essential to long-term resource management, such evaluations of hotspot persistence are seldom attempted in practice (Santora & Sydeman, 2015). A natural way of quantifying the consistency of hotspots through time is to determine the frequency with which a given site exceeds a chosen hotness level. Piacenza *et al.* (2015) proposed a 'universal' threshold in the form of a mean value across years, whereas Santora and Veit (2013) recommended that one standard deviation above a grand regional mean (> 1SD) be used as a common cut-off to align with previous work on remotely-sensed patterns in ocean colour (Suryan *et al.*, 2012) and krill and whale distributions (Santora & Reiss, 2011). Though pragmatic, neither approach was feasible in this study and future efforts should be made to characterise hotspot variability based on finer-scale time series data.

Assessing the processes underpinning the environmental preferences of wildlife species is a major challenge in the pelagic realm. The trophodynamics and habitat usage of tunas, billfishes and their relatives are complex, dynamic and niche-dependent such that species occupying temperate or tropical eco-regions may exhibit contrasting tolerance for, and responses to, similar environmental signals (Arrizabalaga *et al.*, 2015). Such non-stationarity (illustrated in Fig. S10) may be mediated, and further complicated by, biotic interactions of varying intensity and direction across latitudes (Schemske *et al.*, 2009). This may explain why temperature, kinetic energy, oxygen concentration and salinity are often seen as important predictors of biogeographic range but a mechanistic understanding of their influence is often missing in the literature. For example, the role of salinity, a consistently high-scoring variable in all bioregions (Table S4), in determining the occurrence of pelagic species, is particularly obscure albeit some evidence exists that haline fronts may be indirectly linked with reproductive success (e.g. Alvarez-Berastegui *et al.*, 2014), prey density and therefore favourable foraging areas (e.g. Maury *et al.*, 2001). Remotely-sensed measurements of ocean colour and their derivatives, such as FCPI, are more readily interpretable, but may lack explanatory power (Table S4) if the target organisms sit several trophic levels above primary producers (Grémillet *et al.*, 2008), track productivity at depth by following the deep scattering layer or chlorophyll maxima and/or time lags occur between chlorophyll peaks and resource availability for consumers (Navarro *et al.*, 2015).

Few studies have investigated the relative contribution of anthropogenic factors in controlling the distribution of ocean wildlife (Navarro *et al.*, 2016). Distributional shifts caused by climate change or overfishing have been documented in some species (Fromentin *et al.*, 2014), but relationships with cumulative, distal factors can prove complex and taxon-specific (Navarro *et al.*, 2015). As such, disentangling the respective effects of oceanographic conditions, migratory behaviour, density-dependence, exploitation history and population structure on habitat selection remains a significant challenge. Here, the prevalence of human impact measures (Hi and Him), especially in the north of the wEEZ, is unsurprising as this region boasts a diverse array of active and productive industries (e.g. petroleum, tourism, shipping, defence) earmarked for accelerated growth in coming decades. Further research is however needed to define the strength, directionality and persistence of their impacts on individual species at both population and organism levels. Comprehensive syntheses of contemporary pressures and their trends should greatly facilitate this process.

Against a backdrop of limited global marine protection (Costello & Ballantine, 2015) and a failure of conventional management to halt declines in ocean health, support for the implementation of place-based conservation strategies such as marine reserves is rapidly growing. In the past, MRs have largely mirrored the static frameworks that proved successful on land, yet recognition of the more fluid and three-dimensional nature of the pelagic realm has prompted calls for more complex dynamic ocean management (DOM) approaches that can harness near real-time data to provide adaptive and flexible responses to changes in the distribution and behaviour of species, habitats and resource users alike (Maxwell *et al.*, 2015). Although a valid and pragmatic concept, successful examples of its execution remain few (but see Dunn *et al.*, 2016) and restricted to developed countries with sufficient logistical capacity and financial means to make DOM a viable option. That said, with access to technology rapidly expanding, DOM should become increasingly feasible in the future. This does not imply that static reserves are ill-suited to mobile organisms. Indeed, mounting evidence suggests that even migratory species can benefit from static closures as (i) protecting part of an animal's range or life cycle contributes to reducing overall mortality, (ii) pelagic species are not necessarily as far-ranging as previously believed (White *et al.*, 2017) and tend to aggregate around predictable bathymetric and hydrographic features, (iii) their rates of residency, philopatry or site fidelity have generally been underestimated, (iv) static MRs ought to be easier to enforce and therefore more likely to bear tangible benefits in species conservation, and (v) neglecting the potential for evolution of individual movement rates has compromised expectations of MR effectiveness (Mee et al., 2017). In this context, Australia recently declared a national network of MRs that occupies more than a third of its entire jurisdiction (ca. 3.1 million km^2). This areal coverage is exceptional by international standards, however, the reserve system in its current form provides low levels of protection equality across habitats and bathymetric classes (Barr & Possingham, 2013). Our analysis demonstrates (despite some caveats, see Appendix S1) that other natural assets, namely hotspots of mobile predatory wildlife, are also significantly under-represented. Whilst the declaration of the network has been a milestone in Australia's ecosystem-based approach to conservation, work remains to be done to ensure the framework in place is ecologically coherent and enables rapid progress towards the new target set by the IUCN 2014 Sydney World Parks Congress to have at least 30% of ocean environments afforded strict protection within the next fifteen years. Of course, marine reserves are just one conservation tool and are not a blanket solution to the problem of declining fish populations and biodiversity loss (Allison *et al.*, 1998). They can, however, complement management efforts focused on setting and enforcing sustainable exploitation levels, controlling illegal fishing, mitigating pollution, capping fleet capacities, decreasing reliance on destructive gear, and reducing bycatch rates. We also recognise that the mission statement of the world's protected area portfolio has expanded far beyond the original objectives the first reserves were created to fulfil in the early 1900s. Today, MRs are not only promoted as a means of preserving iconic wildlife/seascapes, but also to help bolster national economies, increase tourism, support the livelihood of local communities, alleviate poverty, replenish depleted stocks, and provide resilience in the face of environmental change. The relevance of MRs for pelagic species will therefore need to be balanced against these and numerous other goals.

Here, we have provided empirical evidence that geomorphometrics can be meaningful proxies of macro-ecological patterns in pelagic marine species, a notion long suspected to be true but seldom comprehensively tested (with some exceptions; e.g. Worm *et al.*, 2003; Morato *et al.*, 2010). We also reinforced the notion that landing statistics can be instructive in a biogeographical context (Reygondeau *et al.*, 2012), provided they are handled with care, transparency, and with a thorough understanding of their limitations in addressing specific questions. Appropriate use of these data is critical as they provide some of the most spatially and temporally extensive information available for marine organisms, making them relevant as inputs in spatial planning within the data-limited pelagic ocean. Moreover, we demonstrate that significant opportunities to advance existing conservation frameworks await within offshore waters. The establishment of a global “hotspot repository” (Hazen *et al.*, 2013), to which this study contributes, is an essential step in developing an effective and flexible system of ocean management.

ACKNOWLEDGEMENTS

We are grateful to the *Sea Around Us Project* team and to Daniel Pauly in particular for making the fisheries landing data available for this study. We would like to acknowledge Scott Foster, Damien Garcia, Johnathan Kool, Timothy Leslie, Jarod Santora and Carolin Strobl for their guidance at various stages of the data assembly and analysis, and we also thank two anonymous reviewers whose constructive comments considerably improved the quality of this manuscript. PJB was supported by the Marine Biodiversity Hub through the Australian Government's National Environmental Research Program (NERP), administered by the Department of the Environment. NERP Marine Biodiversity Hub partners include the Institute for Marine and Antarctic Studies, University of Tasmania; CSIRO, Geoscience Australia, Australian Institute of Marine Science, Museum Victoria, Charles Darwin University and the University of Western Australia. PJB was also the recipient of a scholarship for international research fees (SIRF) during the course of this work. RAW acknowledges funding support from the Australian Research Council Discovery project support (DP140101377). ZH and SLN publish with permission of the Chief Executive Officer, Geoscience Australia.

SUPPLEMENTARY MATERIALS

Appendix S1. Supplementary methods.

Appendix S2. Supplementary figures.

Figure S1. Correlation matrix between the retained environmental variables.

Figure S2. Fishing gear classification key.

Figure S3. Bioregions recognised by the WA Department of Fisheries.

Figure S4. Patterns of missing data across the fishing effort time series.

Figure S5. Fishing effort imputation in the North bioregion.

Figure S6. Diagnostic plots from the catch rate standardisation GLMs.

Figure S7. Cumulative frequency distribution curves used for hotspot identification.

Figure S8. Predicted spatial patterns in fish relative abundance.

Figure S9. Uncertainty in predicted patterns in fish relative abundance.

Figure S10. Example partial dependence plots from random forest models.

Figure S11. Cross-taxonomic comparison of spatial hotspots.

Appendix S3. Supplementary tables.

Table S1. Predictor variables used in models of pelagic fish hotspots.

Table S2. Filtering statistics for the Sea Around Us Project catch data.

Table S3. Biological and life-history parameters of pelagic fishes.

Table S4. Summary of predictor rankings in each bioregion.

Appendix S4. Calculation of the frequency of chlorophyll peak index (FCPI).

BIOSKETCH

Phil Bouchet is a postdoctoral researcher interested in the use of statistical models to better understand, and robustly predict, patterns in the abundance, distribution and diversity of marine predators in support of conservation planning decisions throughout the pelagic ocean. He works on a range of vertebrate species (cetaceans, sharks, fishes), at different spatial scales (local, regional, global) and across a number of biogeographical areas (temperate to tropical).

REFERENCES

Allison, G.W., Lubchenco, J. & Carr, M.H. (1998) Marine reserves are necessary but not sufficient for marine conservation. *Ecological applications*, **8**, S79-S92.

Alvarez-Berastegui, D., Ciannelli, L., Aparicio-Gonzalez, A., Reglero, P., Hidalgo, M., López-Jurado, J.L., Tintoré, J. & Alemany, F. (2014) Spatial scale, means and gradients of hydrographic variables define pelagic seascapes of bluefin and bullet tuna spawning distribution. *PloS one*, **9**, e109338.

Angulo-Valdés, J.A. & Hatcher, B.G. (2010) A new typology of benefits derived from marine protected areas. *Marine Policy*, **34**, 635-644.

Anticamara, J.A., Watson, R., Gelchu, A. & Pauly, D. (2011) Global fishing effort (1950–2010): trends, gaps, and implications. *Fisheries Research*, **107**, 131-136.

Arrizabalaga, H., Dufour, F., Kell, L., Merino, G., Ibaibarriaga, L., Chust, G., Irigoien, X., Santiago, J., Murua, H. & Fraile, I. (2015) Global habitat preferences of commercially valuable tuna. *Deep Sea Research Part II: Topical Studies in Oceanography*, **113**, 102-112.

Barr, L.M. & Possingham, H.P. (2013) Are outcomes matching policy commitments in Australian marine conservation planning? *Marine Policy*, **42**, 39-48.

Bartolino, V., Maiorano, L. & Colloca, F. (2011a) Frequency distribution curves and the identification of hotspots: Response to comments. *Population Ecology*, **53**, 603-604.

Bartolino, V., Maiorano, L. & Colloca, F. (2011b) A frequency distribution approach to hotspot identification. *Population Ecology*, **53**, 351-359.

Beier, P., Hunter, M.L. & Anderson, M. (2015) Special Section: Conserving Nature's Stage. *Conservation Biology*, **29**, 613-617.

Bouchet, P.J., Meeuwig, J.J., Salgado Kent, C.P., Letessier, T.B. & Jenner, C.K. (2015) Topographic determinants of mobile vertebrate predator hotspots: Current knowledge and future directions. *Biological Reviews*, **90**, 699-728.

Breen, P., Posen, P. & Righton, D. (2015) Temperate marine protected areas and highly mobile fish: A review. *Ocean & Coastal Management*, **105**, 75-83.

Breiman, L. (2001) Random forests. *Machine Learning*, **45**, 5-32.

Chapman, D.D., Feldheim, K.A., Papastamatiou, Y.P. & Hueter, R.E. (2015) There and back again: A review of residency and return migrations in sharks, with implications for population structure and management. *Annual Review of Marine Science*, **7**, 547-570.

Costello, M.J. & Ballantine, B. (2015) Biodiversity conservation should focus on no-take Marine Reserves: 94% of Marine Protected Areas allow fishing. *Trends in ecology & evolution*, **30**, 507-509.

Daan, N., Gislason, H., Pope, J.G. & Rice, J.C. (2011) Apocalypse in world fisheries? The reports of their death are greatly exaggerated. *ICES Journal of Marine Science: Journal du Conseil*, **68**, 1375-1378.

Devillers, R., Pressey, R.L., Grech, A., Kittinger, J.N., Edgar, G.J., Ward, T. & Watson, R. (2015) Reinventing residual reserves in the sea: are we favouring ease of establishment over need for protection? *Aquatic Conservation: Marine and Freshwater Ecosystems*, **25**, 480-504.

Double, M.C., Andrews-Goff, V., Jenner, K.C.S., Jenner, M.-N., Laverick, S.M., Branch, T.A. & Gales, N.J. (2014) Migratory movements of pygmy blue whales (*Balaenoptera musculus brevicauda*) between Australia and Indonesia as revealed by satellite telemetry. *PLoS ONE*, **9**, e93578.

Dunn, D.C., Maxwell, S.M., Boustany, A.M. & Halpin, P.N. (2016) Dynamic ocean management increases the efficiency and efficacy of fisheries management. *Proceedings of the National Academy of Sciences*, **113**, 668-673.

Fernandez-Arcaya, U., Ramirez-Llodra, E., Aguzzi, J., Allcock, A.L., Davies, J.S., Dissanayake, A., Harris, P., Howell, K., Huvenne, V.A.I., Macmillan-Lawler, M., Martín, J., Menot, L., Nizinski, M., Puig, P., Rowden, A.A., Sanchez, F. & Van den Beld, I.M.J. (2017) Ecological role of submarine canyons and need for canyon conservation: A review. *Frontiers in Marine Science*, **4**

Ferreira, L.C., Thums, M., Meeuwig, J.J., Vianna, G.M.S., Stevens, J., McAuley, R. & Meekan, M.G. (2015) Crossing latitudes—Long-distance tracking of an apex predator. *PLoS ONE*, **10**, e0116916.

Francis, R.I.C.C. (1999) The impact of correlations in standardised CPUE indices. In: *New Zealand Fisheries Assessment Research Document 99/42*, p. 31, Wellington, New Zealand.

Froese, R., Zeller, D., Kleisner, K. & Pauly, D. (2012) What catch data can tell us about the status of global fisheries. *Marine Biology*, **159**, 1283-1292.

Fromentin, J.M., Reygondeau, G., Bonhommeau, S. & Beaugrand, G. (2014) Oceanographic changes and exploitation drive the spatio-temporal dynamics of Atlantic bluefin tuna (Thunnus thynnus). *Fisheries Oceanography*, **23**, 147-156.

Garcia, D. (2010) Robust smoothing of gridded data in one and higher dimensions with missing values. *Computational Statistics & Data Analysis*, **54**, 1167-1178.

Gormley, A.M., Slooten, E., Dawson, S., Barker, R.J., Rayment, W., du Fresne, S. & Bräger, S. (2012) First evidence that marine protected areas can work for marine mammals. *Journal of Applied Ecology*, **49**, 474-480.

Graham, F., Rynne, P., Estevanez, M., Luo, J., Ault, J.S. & Hammerschlag, N. (2016) Use of marine protected areas and exclusive economic zones in the subtropical western North Atlantic Ocean by large highly mobile sharks. *Diversity and Distributions*, **22**, 534-546.

Grémillet, D., Lewis, S., Drapeau, L., Van Der Lingen, C.D., Huggett, J.A., Coetzee, J.C., Verheye, H.M., Daunt, F., Wanless, S. & Ryan, P.G. (2008) Spatial match–mismatch in the Benguela upwelling zone: Should we expect chlorophyll and sea-surface temperature to predict marine predator distributions? *Journal of Applied Ecology*, **45**, 610-621.

Grewe, P.M., Feutry, P., Hill, P.L., Gunasekera, R.M., Schaefer, K.M., Itano, D.G., Fuller, D.W., Foster, S.D. & Davies, C.R. (2015) Evidence of discrete yellowfin tuna (Thunnus albacares) populations demands rethink of management for this globally important resource. *Scientific Reports*, **5**, 16916.

Harris, P.T. & Whiteway, T. (2009) High seas marine protected areas: Benthic environmental conservation priorities from a GIS analysis of global ocean biophysical data. *Ocean & Coastal Management*, **52**, 22-38.

Hazen, E.L., Suryan, R.M., Santora, J.A., Bograd, S.J., Watanuki, Y. & Wilson, R.P. (2013) Scales and mechanisms of marine hotspot formation. *Marine Ecology Progress Series*, **487**, 177-183.

Hooker, S.K., Hyrenbach, K.D., Corrigan, C., Polovina, J.J. & Reeves, R.R. (2011) Making protected area networks effective for marine top predators. *Endangered Species Research*, **13**, 203-218.

Huang, Z., Nichol, S.L., Harris, P.T. & Caley, M.J. (2014) Classification of submarine canyons of the Australian continental margin. *Marine Geology*, **357**, 362-383.

Huvenne, V.A.I. & Davies, J.S. (2014) Towards a new and integrated approach to submarine canyon research. *Deep Sea Research Part II: Topical Studies in Oceanography*, **104**, 1-5.

Johnson, C.M., Beckley, L.E., Kobryn, H., Johnson, G.E., Kerr, I. & Payne, R. (2016) Crowdsourcing modern and historical data identifies sperm whale (*Physeter macrocephalus*) habitat offshore of south-western Australia. *Frontiers in Marine Science*, **3**

Kaplan, D.M., Chassot, E., Gruss, A. & Fonteneau, A. (2010) Pelagic MPAs: the devil is in the details. *Trends in Ecology & Evolution*, **25**, 62-63.

Kerwath, S.E., Winker, H., Götz, A. & Attwood, C.G. (2013) Marine protected area improves yield without disadvantaging fishers. *Nature Communications*, **4**, 2347.

Kerwath, S.E., Thorstad, E.B., Naesje, T.F., Cowley, P.D., Økland, F., Wilke, C. & Attwood, C.G. (2009) Crossing invisible boundaries: The effectiveness of the Langebaan Lagoon Marine Protected Area as a harvest refuge for a migratory fish species in South Africa. *Conservation Biology*, **23**, 653-661.

Kessel, S.T., Chapman, D.D., Franks, B.R., Gedamke, T., Gruber, S.H., Newman, J.M., White, E.R. & Perkins, R.G. (2014) Predictable temperature-regulated residency, movement and migration in a large, highly mobile marine predator (Negaprion brevirostris). *Marine Ecology Progress Series*, **514**, 175-190.

Leroux, S.J., Schmiegelow, F.K. & Nagy, J.A. (2007) Potential spatial overlap of heritage sites and protected areas in a boreal region of northern Canada. *Conservation Biology*, **21**, 376-386.

Letessier, T., Kemp, K., Gollock, M., Groves, L., Bouchet, P., Chapuis, L., Vianna, G., Koldewey, H. & Meeuwig, J. (2013) Assessing pelagic fish and shark populations: The application of demersal techniques to the mid-water. *Methods in Oceanography*, **8**, 41-55.

Letessier, T.B., Bouchet, P.J. & Meeuwig, J.J. (2017) Sampling mobile oceanic fishes and sharks: Implications for fisheries and conservation planning. *Biological Reviews*, **92**, 627-646.

Marchese, C. (2015) Biodiversity hotspots: A shortcut for a more complicated concept. *Global Ecology and Conservation*, **3**, 297-309.

Maunder, M.N. & Punt, A.E. (2004) Standardizing catch and effort data: A review of recent approaches. *Fisheries Research*, **70**, 141-159.

Maury, O., Gascuel, D., Marsac, F., Fonteneau, A. & Rosa, A.-L.D. (2001) Hierarchical interpretation of nonlinear relationships linking yellowfin tuna (Thunnus albacares) distribution to the environment in the Atlantic Ocean. *Canadian Journal of Fisheries and Aquatic Sciences*, **58**, 458-469.

Maxwell, S.M., Hazen, E.L., Lewison, R.L., Dunn, D.C., Bailey, H., Bograd, S.J., Briscoe, D.K., Fossette, S., Hobday, A.J., Bennett, M., Benson, S., Caldwell, M.R., Costa, D.P., Dewar, H., Eguchi, T., Hazen, L., Kohin, S., Sippel, T. & Crowder, L.B. (2015) Dynamic ocean management: Defining and conceptualizing real-time management of the ocean. *Marine Policy*, **58**, 42-50.

McArthur, M.A., Brooke, B.P., Przeslawski, R., Ryan, D.A., Lucieer, V.L., Nichol, S., McCallum, A.W., Mellin, C., Cresswell, I.D. & Radke, L.C. (2010) On the use of abiotic surrogates to describe marine benthic biodiversity. *Estuarine, Coastal and Shelf Science*, **88**, 21-32.

Mee, J.A., Otto, S.P. & Pauly, D. (2017) Evolution of movement rate increases the effectiveness of marine reserves for the conservation of pelagic fishes. *Evolutionary Applications*, n/a-n/a.

Mellin, C. (2015) Abiotic surrogates in support of marine biodiversity conservation. *Indicators and Surrogates of Biodiversity and Environmental Change* (ed. by D. Lindenmayer, P. Barton and J. Pierson), p. 216. CSIRO PUBLISHING.

Michael, P.E., Jahncke, J. & Hyrenbach, K.D. (2014) Relative influence of static and dynamic features on black-footed albatross (*Phoebastria nigripes*) habitat use in central California Sanctuaries. *Fisheries Oceanography*, **23**, 18-31.

Moors-Murphy, H.B. (2014) Submarine canyons as important habitat for cetaceans, with special reference to the Gully: A review. *Deep Sea Research Part II: Topical Studies in Oceanography*, **104**, 6-19.

Morato, T., Hoyle, S.D., Allain, V. & Nicol, S.J. (2010) Seamounts are hotspots of pelagic biodiversity in the open ocean. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 9707-9711.

Morato, T., Varkey, D.A., Damaso, C., Machete, M., Santos, M., Prieto, R., Santos, R.S. & Pitcher, T.J. (2008) Evidence of a seamount effect on aggregating visitors. *Marine Ecology Progress Series*, **357**, 23-32.

Navarro, J., Coll, M., Cardador, L., Fernández, Á.M. & Bellido, J.M. (2015) The relative roles of the environment, human activities and spatial factors in the spatial distribution of marine biodiversity in the Western Mediterranean Sea. *Progress in Oceanography*, **131**, 126-137.

Navarro, J., Cardador, L., Fernández, Á.M., Bellido, J.M. & Coll, M. (2016) Differences in the relative roles of environment, prey availability and human activity in the spatial distribution of two marine mesopredators living in highly exploited ecosystems. *Journal of Biogeography*, **43**, 440-450.

Neubauer, P., Jensen, O.P., Hutchings, J.A. & Baum, J.K. (2013) Resilience and recovery of overexploited marine populations. *Science*, **340**, 347-349.

Nichol, S.L., Howard, F.J.F., Kool, J., Stowar, M., Bouchet, P., Radke, L., Siwabessy, J., Przeslawski, R., Picard, K., Alvarez de Glasby, B., Colquhoun, J., Letessier, T. & Heyward, A. (2013) Oceanic Shoals Commonwealth Marine Reserve (Timor Sea) Biodiversity Survey. In: *Post-survey report GA0339/SOL5650*, p. 112. Geoscience Australia

Pauly, D. & Zeller, D. (2016) Catch reconstructions reveal that global marine fisheries catches are higher than reported and declining. *Nature Communications*, **7**, 10244.

Piacenza, S.E., Thurman, L.L., Barner, A.K., Benkwitt, C.E., Boersma, K.S., Cerny-Chipman, E.B., Ingeman, K.E., Kindinger, T.L., Lindsley, A.J., Nelson, J., Reimer, J.N., Rowe, J.C., Shen, C., Thompson, K.A. & Heppell, S.S. (2015) Evaluating temporal consistency in marine biodiversity hotspots. *PLOS ONE*, **10**, e0133301.

Powers, J.E. & Abeare, S.M. (2009) Fishing effort redistribution in response to area closures. *Fisheries Research*, **99**, 216-225.

Real, R. & Vargas, J.M. (1996) The probabilistic basis of Jaccard's index of similarity. *Systematic biology*, **45**, 380-385.

Rees, M.J., Jordan, A., Price, O.F., Coleman, M.A. & Davis, A.R. (2014) Abiotic surrogates for temperate rocky reef biodiversity: Implications for marine protected areas. *Diversity and distributions*, **20**, 284-296.

Rennie, S., Hanson, C.E., McCauley, R.D., Pattiaratchi, C., Burton, C., Bannister, J., Jenner, C. & Jenner, M.N. (2009) Physical properties and processes in the Perth Canyon, Western Australia: Links to water column production and seasonal pygmy blue whale abundance. *Journal of Marine Systems*, **77**, 21-44.

Reygondeau, G., Maury, O., Beaugrand, G., Fromentin, J.M., Fonteneau, A. & Cury, P. (2012) Biogeography of tuna and billfish communities. *Journal of Biogeography*, **39**, 114-129.

Roberts, C.M. (2000) Selecting marine reserve locations: optimality versus opportunism. *Bulletin of Marine Science*, **66**, 581-592.

Sabatia, C.O. & Burkhart, H.E. (2014) Predicting site index of plantation loblolly pine from biophysical variables. *Forest Ecology and Management*, **326**, 142-156.

Santora, J. & Veit, R. (2013) Spatio-temporal persistence of top predator hotspots near the Antarctic Peninsula. *Marine Ecology Progress Series*, **487**, 287-304.

Santora, J.A. & Reiss, C.S. (2011) Geospatial variability of krill and top predators within an Antarctic submarine canyon system. *Marine Biology*, **158**, 2527-2540.

Santora, J.A. & Sydeman, W.J. (2015) Persistence of hotspots and variability of seabird species richness and abundance in the southern California Current. *Ecosphere*, **6**, 1-19.

Scales, K.L., Miller, P.I., Hawkes, L.A., Ingram, S.N., Sims, D.W. & Votier, S.C. (2014) On the front line: Frontal zones as priority at-sea conservation areas for mobile marine vertebrates. *Journal of Applied Ecology*, **51**, 1575-1583.

Schemske, D.W., Mittelbach, G.G., Cornell, H.V., Sobel, J.M. & Roy, K. (2009) Is there a latitudinal gradient in the importance of biotic interactions? *Annu. Rev. Ecol. Evol. Syst.*, **40**, 245-269.

Scott, R., Hodgson, D.J., Witt, M.J., Coyne, M.S., Adnyana, W., Blumenthal, J.M., Broderick, A.C., Canbolat, A.F., Catry, P., Ciccione, S., Delcroix, E., Hitipeuw, C., Luschi, P., Pet-Soede, L., Pendoley, K., Richardson, P.B., Rees, A.F. & Godley, B.J. (2012) Global analysis of satellite tracking data shows that adult green turtles are significantly aggregated in Marine Protected Areas. *Global Ecology and Biogeography*, **21**, 1053-1061.

Singleton, R.L. & Roberts, C.M. (2014) The contribution of very large marine protected areas to marine conservation: Giant leaps or smoke and mirrors? *Marine pollution bulletin*, **87**, 7-10.

Sleeman, J., Meekan, M., Wilson, S., Jenner, C., Jenner, M., Boggs, G., Steinberg, C. & Bradshaw, C. (2007) Biophysical correlates of relative abundances of marine megafauna at Ningaloo Reef, Western Australia. *Marine and Freshwater Research*, **58**, 608-623.

Stevens, T. & Connolly, R.M. (2004) Testing the utility of abiotic surrogates for marine habitat mapping at scales relevant to management. *Biological Conservation*, **119**, 351-362.

Strobl, C., Hothorn, T. & Zeileis, A. (2009) Party on! A new, conditional variable importance measure for random forests available in the party package. *The R Journal*, **1**, 14-17.

Strobl, C., Boulesteix, A.-L., Zeileis, A. & Hothorn, T. (2007) Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics*, **8**, 1-21.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T. & Zeileis, A. (2008) Conditional variable importance for random forests. *BMC Bioinformatics*, **9**, 1-11.

Suryan, R.M., Santora, J.A. & Sydeman, W.J. (2012) New approach for using remotely sensed chlorophyll a to identify seabird hotspots. *Marine Ecology Progress Series*, **451**, 213-225.

Sutcliffe, P., Klein, C., Pitcher, C. & Possingham, H. (2015) The effectiveness of marine reserve systems constructed using different surrogates of biodiversity. *Conservation Biology*, **29**, 657-667.

Szymczak, S., Holzinger, E., Dasgupta, A., Malley, J.D., Molloy, A.M., Mills, J.L., Brody, L.C., Stambolian, D. & Bailey-Wilson, J.E. (2016) r2VIM: A new variable selection method for random forests in genome-wide association studies. *BioData Mining*, **9**, 7.

Trebilco, R., Halpern, B.S., Flemming, J.M., Field, C., Blanchard, W. & Worm, B. (2011) Mapping species richness and human impact drivers to inform global pelagic conservation prioritisation. *Biological Conservation*, **144**, 1758-1766.

Wang, X.-F. (2010) *fANCOVA: Nonparametric analysis of covariance*.

Watson, R., Kitchingman, A., Gelchu, A. & Pauly, D. (2004) Mapping global fisheries: sharpening our focus. *Fish and Fisheries*, **5**, 168-177.

Watson, R., Alder, J., Kitchingman, A. & Pauly, D. (2005) Catching some needed attention. *Marine Policy*, **29**, 281-284.

Watson, R.A., Cheung, W.W.L., Anticamara, J.A., Sumaila, R.U., Zeller, D. & Pauly, D. (2013) Global marine yield halved as fishing intensity redoubles. *Fish and Fisheries*, **14**, 493-503.

Webb, T.J., Vanden Berghe, E. & O'Dor, R. (2010) Biodiversity's big wet secret: The global distribution of marine biological records reveals chronic under-exploration of the deep pelagic ocean. *PLoS ONE*, **5**, e10223.

White, T.D., Carlisle, A.B., Kroodsma, D.A., Block, B.A., Casagrandi, R., De Leo, G.A., Gatto, M., Micheli, F. & McCauley, D.J. (2017) Assessing the effectiveness of a large marine protected area for reef shark conservation. *Biological Conservation*, **207**, 64-71.

Worm, B., Lotze, H. & Myers, R. (2003) Predator diversity hotspots in the blue ocean. *Proceedings of the National Academy of Sciences*, **100**, 9884-9888.

Worm, B., Hilborn, R., Baum, J.K., Branch, T.A., Collie, J.S., Costello, C., Fogarty, M.J., Fulton, E.A., Hutchings, J.A., Jennings, S., Jensen, O.P., Lotze, H.K., Mace, P.M., McClanahan, T.R., Minto, C., Palumbi, S.R., Parma, A.M., Ricard, D., Rosenberg, A.A., Watson, R. & Zeller, D. (2009) Rebuilding global fisheries. *Science*, **325**, 578-585.

Young, H.S., Maxwell, S.M., Conners, M.G. & Shaffer, S.A. (2015) Pelagic marine protected areas protect foraging habitat for multiple breeding seabirds in the central Pacific. *Biological Conservation*, **181**, 226-235.

TABLES

Table 1. Summary of explanatory variables. Only the top 15 random forest predictors from each bioregion are shown (Fig. 3). See Table S1 in the Supplementary Information for the full list and links to data sources. Geomorphometrics encompass both "bathymetric / topographic" and "geologic" parameters. d.u.: 'dimensionless unit'.

VARIABLE	NAME	UNIT	STATE	ECOLOGICAL INTERPRETATION
Bathymetric / topographic				
CI	Contour index	d.u.	Static	Vertical circulation and mixing, eddy formation, prey refugia, prey entrapment
CRS	Cross-sectional curvature	rad.m^{-1}	Static	Vertical circulation and mixing, eddy formation, prey refugia, prey entrapment
FRD	Fractal dimension	d.u.	Static	Vertical circulation and mixing, eddy formation, prey refugia, prey entrapment
LSRI	Land surface ruggedness index	d.u.	Static	Vertical circulation and mixing, eddy formation, prey refugia, prey entrapment
RUG	Rugosity	d.u.	Static	Vertical circulation and mixing, eddy formation, prey refugia, prey entrapment
Geologic				
CAN	Presence/absence of one or more submarine canyons	d.u.	Static	Breeding/spawning habitat, migration cue, prey availability
CANpercent	Percentage of grid cell area	%	Static	Breeding/spawning habitat, migration cue, prey availability

	occupied by submarine canyons			
CANhead	Presence/absence of one or more canyon heads	d.u.	Static	Productivity (upwelling), food availability, feeding ground
CANadj	Number of adjacent cells containing submarine canyons	d.u.	Static	Population connectivity, larval dispersal
CANdepth	Maximum canyon depth within a cell	m	Static	Prey availability, breeding/spawning habitat
FEATcount	Number of distinct geomorphic features within a cell	d.u.	Static	Prey and habitat diversity
FEATdom	Dominant geomorphic feature class	d.u.	Static	Prey availability, breeding/spawning habitat
Oceanographic				
CUREW	East-west current velocity	$m.s^{-1}$	Dynamic	Nutrient inputs, oxygenation, enhanced productivity, larval drift and juvenile recruitment
CURNS	North-south current velocity	$m.s^{-1}$	Dynamic	Nutrient inputs, oxygenation, enhanced productivity, larval drift and juvenile recruitment
FFD	Daily sea surface temperature frontal frequency	%	Dynamic	Food availability, migration cue
L2	Regional circulation regimes	d.u.	Dynamic	Eddy formation, enhanced primary and secondary production
L3	Ocean energy	d.u.	Dynamic	Prey availability, breeding habitat, feeding success, larval growth rates
MIX	Mixed layer depth	m	Dynamic	Prey availability, physiological tolerance, oxygen availability

PP	Annual mean primary production	mg $C.m^{-2}.d^{-1}$	Dynamic	Prey availability
PPstd	Standard deviation of annual mean primary production	mg $C.m^{-2}.d^{-1}$	Dynamic	Prey availability
SAL	Annual mean salinity at the surface	PSU	Dynamic	Prey availability, physiological tolerance, hatching rate
SSTstd	Standard deviation of sea surface temperature	°C	Dynamic	Spawning cue, breeding habitat, metabolic stress
WAT	Water mass at surface	d.u.	Dynamic	Prey availability, physiological tolerance
Anthropogenic				
Hi	Mean cumulative human impact	d.u.	Dynamic	Mortality, sub-lethal disturbance, displacement
Hir	Range of cumulative human impacts	d.u.	Dynamic	Mortality, sub-lethal disturbance, displacement
Him	Maximum cumulative human impact	d.u.	Dynamic	Mortality, sub-lethal disturbance, displacement

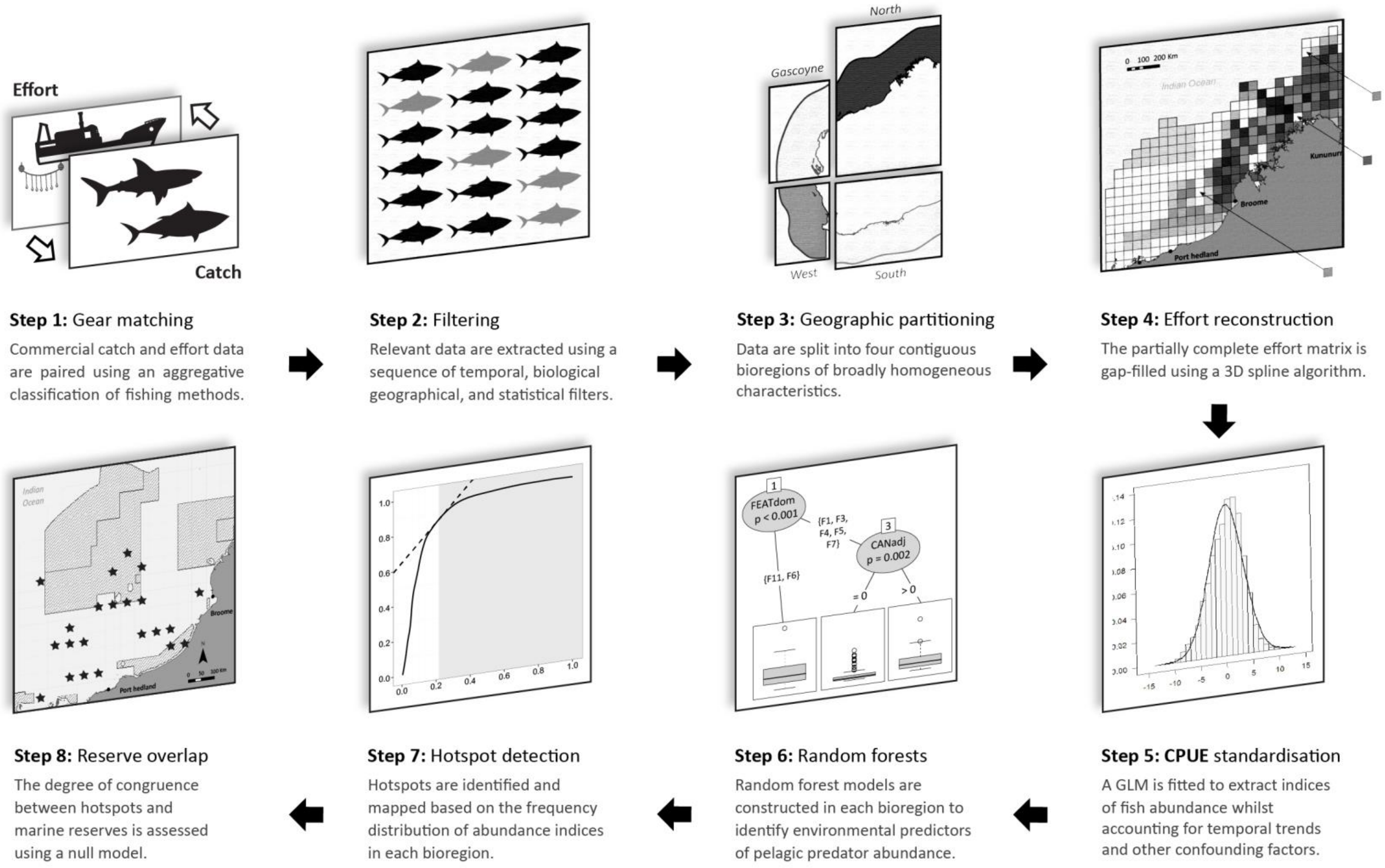
Table 2. Model selection summary from the GLM standardisation of pelagic fish catch rates. The best model (with minimum second-order Akaike Information Criterion AICc and maximum adjusted R^2) is shown in bold. ΔAICc quantifies the difference in AICc score between the current and best models. α is the intercept, ε is the residual variation, log(E) and log(A) are offset terms for fishing effort and water surface area respectively, and the standard notation ‘x’ represents all covariate main effects and their associated interactions. Weights are the species-specific values reported in Table S3.

Model formulation	**North**		**Gascoyne**		**West**		**South**	
	ΔAICc	**Adj. R^2**	**ΔAICc**	**Adj. R^2**	**ΔAICc**	**Adj. R^2**	**ΔAICc**	**Adj. R^2**
log(Catch) ~ α + Cell + Year + log(E) + log(A) + ε	23,976	0.355	9372.1	0.204	10,611	0.235	9,911	0.203
log(Catch) ~ α + Cell + Year + Gear + log(E) + log(A) + ε	10,541	0.456	4687.1	0.347	6,445	0.356	8,411	0.256
log(Catch) ~ α + Cell + Year + Gear + Weight + log(E) + log(A) + ε	4,928	0.459	3063.7	0.347	3,997	0.369	1936.6	0.384
log(Catch) ~ α + Cell + Year + Gear + Weight + Dport + log(E) + log(A) + ε	4,928	0.459	3063.7	0.347	3,997	0.369	1936.6	0.384
log(Catch) ~ α + Cell + Year + Gear x Weight + log(E) + log(A) + ε	3,739	0.468	2838.1	0.354	3604.9	0.380	1249.3	0.404
log(Catch) ~ α + Cell + Year + Gear x Dport + log(E) + log(A) + ε	10,389	0.457	3940.2	0.368	5,326	0.385	8,281	0.261
log(Catch) ~ α + Cell + Year x Dport + Gear + Weight + log(E) + log(A) + ε	4,478	0.463	2425.5	0.365	3613.2	0.380	1785.1	0.389
log(Catch) ~ α + Cell + Year x Gear + Weight + Dport + log(E) + log(A) + ε	3,182	0.472	767.8	0.409	956.3	0.445	988.3	0.412
log(Catch) ~ α + Cell + Year x Gear x Weight + log(E) + log(A) + ε	**0**	**0.493**	**0**	**0.428**	**0**	**0.467**	**0**	**0.440**

Table 3. Predictive accuracy of the random forest models of pelagic fish abundance. Values represent bootstrap means (with associated 95% percentile confidence intervals). Performance is evaluated on the out-of-bag data (see Appendix S1). RMSE stands for the root mean squared error, and R^2 represents the percentage of variance explained.

Input	**Performance metric**	**Bioregions**			
		North	Gascoyne	West	South
All variables	RMSE	1.58 (1.54 - 1.61)	1.0 (0.93 - 1.05)	0.77 (0.73 - 0.83)	1.07 (1.02 - 1.11)
All variables	R^2	0.70 (0.69 - 0.71)	0.44 (0.39 - 0.48)	0.24 (0.19 - 0.29)	0.5 (0.46 - 0.54)
Geomorphometrics	RMSE	1.93 (1.9 - 1.96)	1.16 (1.1 - 1.21)	0.81 (0.76 - 0.87)	1.23 (1.17 - 1.27)
Geomorphometrics	R^2	0.51 (0.5 - 0.52)	0.19 (0.16 - 0.24)	0.17 (0.12 - 0.23)	0.33 (0.3 - 0.37)

FIGURES



Step 1: Gear matching

Commercial catch and effort data are paired using an aggregative classification of fishing methods.

Step 2: Filtering

Relevant data are extracted using a sequence of temporal, biological geographical, and statistical filters.

Step 3: Geographic partitioning

Data are split into four contiguous bioregions of broadly homogeneous characteristics.

Step 4: Effort reconstruction

The partially complete effort matrix is gap-filled using a 3D spline algorithm.

Step 5: CPUE standardisation

A GLM is fitted to extract indices of fish abundance whilst accounting for temporal trends and other confounding factors.

Step 6: Random forests

Random forest models are constructed in each bioregion to identify environmental predictors of pelagic predator abundance.

Step 7: Hotspot detection

Hotspots are identified and mapped based on the frequency distribution of abundance indices in each bioregion.

Step 8: Reserve overlap

The degree of congruence between hotspots and marine reserves is assessed using a null model.

Figure 1. Graphic representation of the grooming and analysis of the *Sea Around Us* Project (SAUP) data.

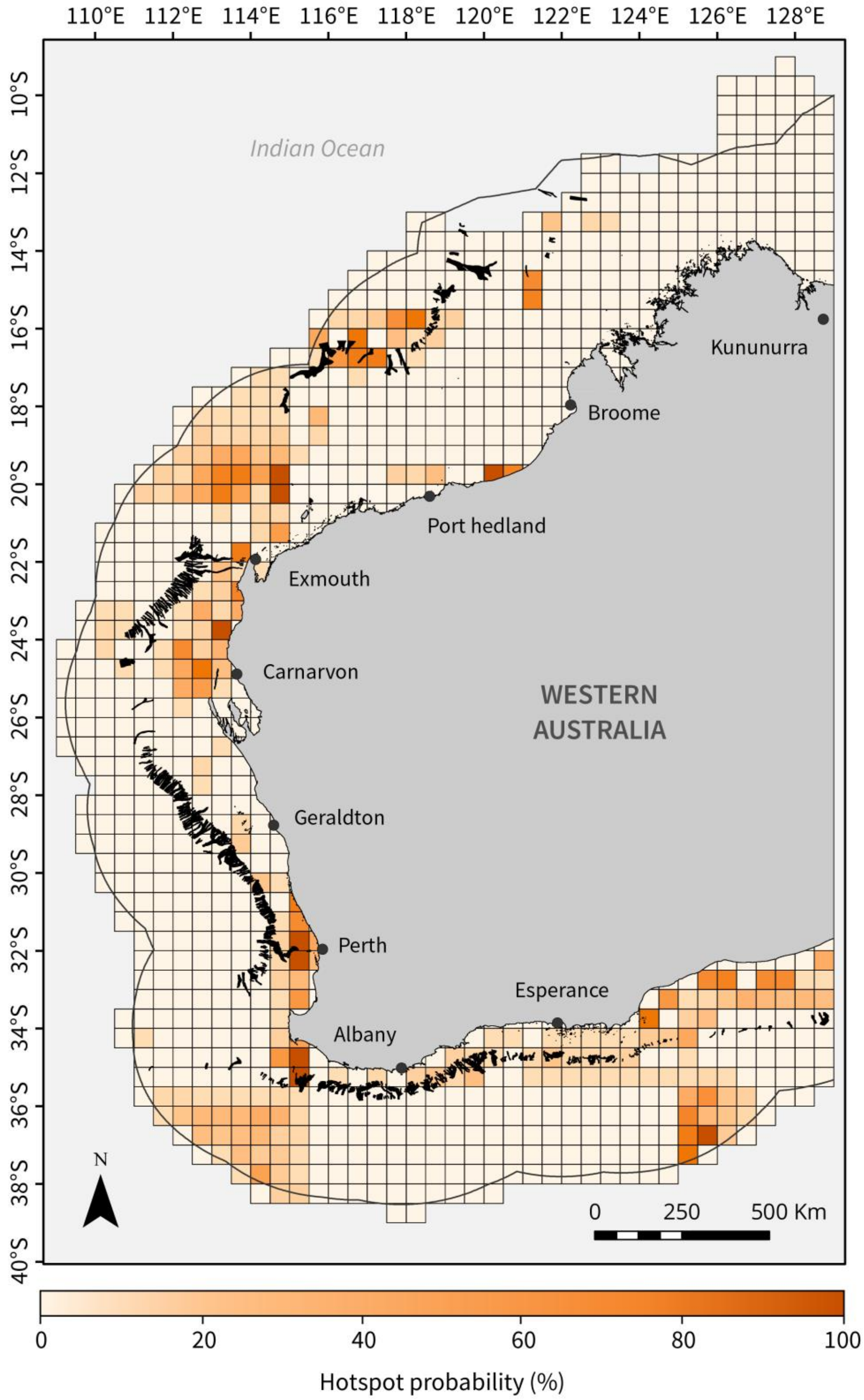


Figure 2. Pelagic fish hotspots derived from the SAUP data. Hotspot probability was derived as the frequency with which each grid cell was selected as a hotspot across n=100 bootstrap iterations, with darker tones denoting higher values. Submarine canyons (Huang *et al.*, 2014) are overlaid in black.

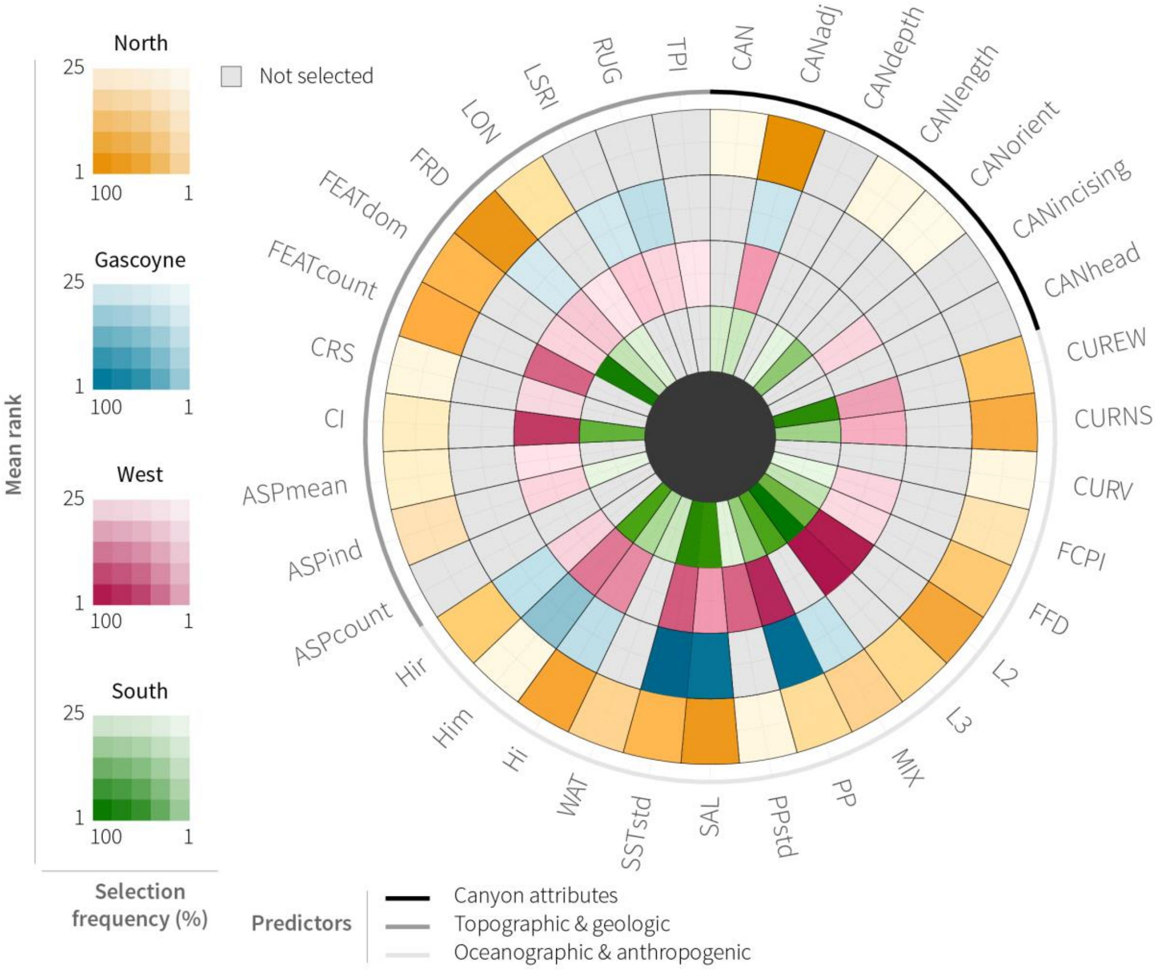


Figure 3. Summary of predictor importance in the random forest models. Each bioregion is assigned a different colour scheme and position on the wheel, from the outer (North) to the inner rings (South). Darker tones indicate predictors that both ranked highly and consistently (across bootstrap resamples), as per the bivariate key. Predictors that were not retained following variable selection (see Appendix S1 for details) are shown in grey. Canyon attributes are identified in black, topographic variables in dark grey and all remaining predictors in light grey (left). A full list of variable abbreviations is provided in Table S1.

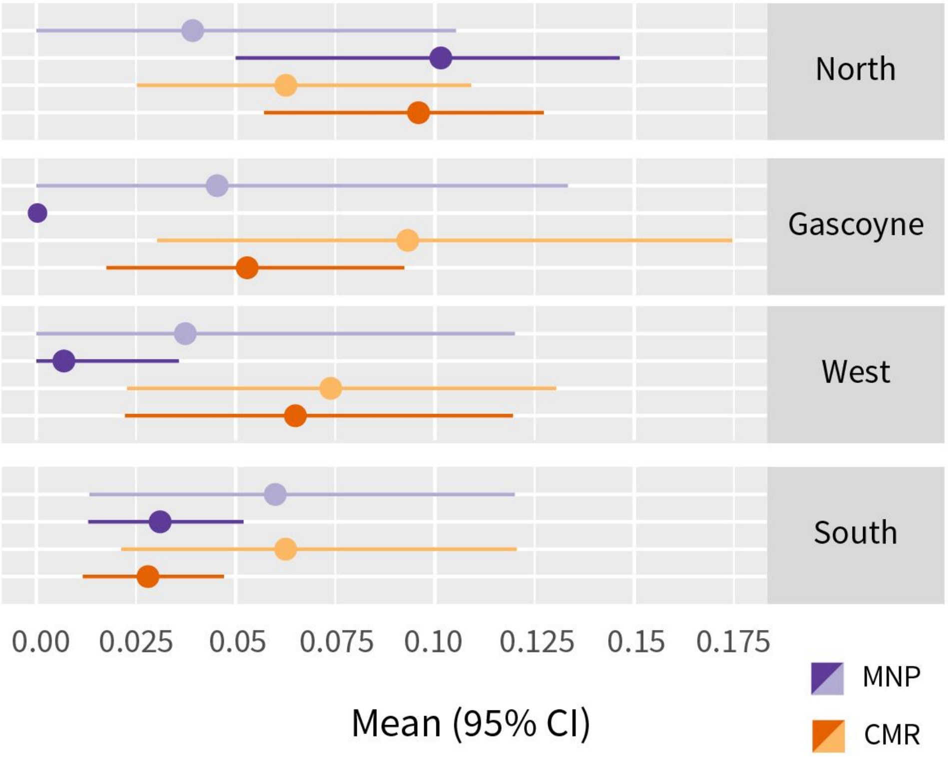


Figure 4. Congruence between pelagic fish hotspots and marine reserves. Overlap is measured as the Jaccard similarity coefficient J, which ranges from 0 (no overlap) to 1 (complete overlap). Lighter and darker colours capture the distribution of values under random null models (n=10,000) and empirical bootstrap resamples (n=100) respectively. Mean values are shown as filled circles, and their associated 95% percentile confidence intervals (CIs) as solid lines. CMR: Commonwealth Marine Reserves (all zones), MNP = Marine National Parks (no-take zones only).