Received: 17 October 2023 | Revised: 22 December 2023 | Accepted: 30 January 2024
DOI: 10.1002/aqc.4114

RESEARCH ARTICLE

WILEY

Independent statistical validation of the New Zealand Seafloor Community Classification

Fabrice Stephenson[1,2] | Jordi Tablada[3] | Ashley A. Rowden[4,5] | Richard Bulmer[6] | David A. Bowden[4] | Shane W. Geange[3]

[1]School of Science, University of Waikato, Hamilton, New Zealand
[2]School of Natural and Environment Sciences, Newcastle University, Newcastle upon Tyne, UK
[3]New Zealand Department of Conservation, Wellington, New Zealand
[4]National Institute of Water & Atmospheric Research, Wellington, New Zealand
[5]School of Biological Sciences Wellington, Victoria University Wellington, Wellington, New Zealand
[6]Tidal Research, Auckland, New Zealand

Correspondence
Fabrice Stephenson, School of Science, University of Waikato, Hamilton, New Zealand.
Email: fabrice.stephenson@ncl.ac.uk

Funding information
New Zealand Department of Conservation, Grant/Award Number: NOF-BIO-317

Abstract

1. The New Zealand Seafloor Community Classification (NZSCC) is a national-scale numerical community classification which depicts compositional turnover of 1716 taxa (demersal fish, reef fish, benthic invertebrates and macroalgae) classified into 75 groups representing seafloor communities. To ensure the continual use of the NZSCC for spatial planning and reporting, a robust maintenance framework must be set in place; key to this is being able to assess the ability of the classification to represent (discriminate between) different seafloor communities.
2. Here we describe an approach for validating the NZSCC using temporally independent evaluation data for demersal fish and benthic invertebrates (the latter sampled via a different method), which identifies whether the NZSCC represents different seafloor communities (i.e., assesses classification strength), evaluates the underlying statistical model, and considers heterogeneity in environmental coverage and statistical uncertainty. Additionally, the availability of abundance estimates for these evaluation datasets provides an opportunity to test whether the NZSCC—which was developed using presence-absence data—can reflect abundance-weighted seafloor communities.
3. The ANOSIM global R values (measuring classification strength) were 0.53 and 0.46 (and significant at the 1% level) for demersal fish and benthic invertebrates, respectively, indicating that the NZSCC groups define biologically distinctive environments.
4. The proportion of significant inter-group differences were very high (95% and 97% for demersal fish and benthic invertebrates, respectively) suggesting NZSCC groups were distinct from each other in their taxonomic composition.
5. There were positive relationships between the evaluation datasets and the underlying statistical model. There was no evidence of these relationships being affected by the statistical uncertainty of the NZSCC.
6. NZSCC model validation metrics using abundance evaluation data were also moderately high (albeit lower than for presence-absence for invertebrates) suggesting that the NZSCC, can at least in part, represent variation in abundance-

weighted communities. Results presented here suggest that the existing NZSCC is currently fit-for-purpose for informing management decisions.

KEYWORDS

abundance; demersal fish, benthic invertebrates; gradient Forest model; NZSCC; presence-absence; spatial management, biodiversity; validating

1 | INTRODUCTION

Marine habitats and ecosystems are under increasing pressure from human activities that can threaten biodiversity, ecosystem functioning and the delivery of ecosystems services (Lade et al., 2020; Smale et al., 2019). Target 3 of the Kunming-Montreal Global Biodiversity Framework has called for at least 30 per cent of marine and coastal areas to be conserved in ecologically representative, well-connected and equitably governed systems of marine protected areas (MPAs), and other effective area-based conservation measures. As referenced in Target 3, one of the principal attributes of effective MPA network design is 'representativeness'—the inclusion of representative marine habitats and communities that, through their protection in the MPA network, will help ensure the preservation of the overall biodiversity of the area of interest (Day et al., 2012). To design a representative MPA network, it is helpful to develop biological classifications that represent biodiversity patterns and identify environmental and/or ecological units that can be predicted to contain distinct representations of biodiversity (Costello, 2009). Where biological classifications exist, they can be used to design ecologically representative MPA networks by ensuring either a minimum proportion of each classification group is included in protected areas; a minimum proportion of transitional boundaries between classification groups (which can aggregate high diversity and density of species) are included in protected areas; or protected areas are replicated within each classification group to help ensure representation of biological heterogeneity (Stephenson et al., 2021). Although the successful application of these approaches necessitates understanding the performance of the biological classification in predicting the spatial patterns of communities (Dixon-Bridges et al., 2014), few environmental classifications have had their performance evaluated nor assessed numerically (Stephenson et al., 2021; e.g., Stevens & Connolly, 2004; Sutcliffe et al., 2015).

The New Zealand Seafloor Community Classification (NZSCC, Figure 1) is a national-scale community classification depicting the compositional turnover of 1716 taxa (demersal fish, reef fish, benthic invertebrates and macroalgae) classified in 75 groups representing seafloor communities (Petersen et al., 2021; Stephenson et al., 2021; Stephenson et al., 2022). The NZSCC relies on both biotic and physical/environmental data to represent the interconnection between biological assemblages and the environment. The underpinning statistical model for the NZSCC is a community-based (multivariate) modelling method called Gradient Forests (GF, Ellis et al., 2012; Pitcher et al., 2012). GF use species distribution data to control the selection, weighting, and transformation of (spatially explicit) environmental predictors to maximize their correlation with species compositional turnover and establish where along the range of environmental gradients important compositional changes occur (Ellis et al., 2012). These transformed environmental layers (representing species compositional turnover) can then be (hierarchically) classified into discrete classification groups that capture spatial variation in species composition and turnover (Stephenson et al., 2021, 2022). Using GF models, the NZSCC was developed and tuned using biotic data from demersal fish, reef fish, benthic invertebrates, and macroalgae. The final 75 group classification (referred to hereafter as the NZSCC) aims to represent seafloor communities of these four biotic groups (Petersen et al., 2021).

To support the interpretation of the NZSCC and its application in MPA network design and marine spatial planning, two spatially explicit uncertainty estimates were generated: environmental coverage and standard deviation of the predicted species compositional turnover. The environmental coverage provides an indication of the parts of the environmental space that, for example, contain many samples—meaning we can be more confident of the relationships and the predictions for compositional turnover and NZSSC groupings in such areas (and conversely, we may be less confident of NZSCC groupings in areas within environmental space that have fewer samples - for a more detailed description of the methods and outputs of the NZSCC, see Stephenson et al., 2021, 2022). The uncertainty estimates of community compositional turnover provide an important indication of the variability in the GF modelling estimates. That is, where there is more variability in the GF estimates, the resulting predicted classification will be less certain. Together, these uncertainty estimates provide complementary measures of uncertainty to be considered by managers (Stephenson et al., 2022).

The development of the NZSCC represents a key step in improving the information base underpinning the design of representative and well-connected MPAs in New Zealand, and ultimately delivering on New Zealand's contribution to Target 3 of the Kunming-Montreal Global Biodiversity Framework. Descriptions of biodiversity patterns provided by the NZSCC are intended to inform the expansion of New Zealand's MPA network by ensuring that a minimum proportion of each classification group is included in MPAs (or other effective area-based conservation measures) that are integrated into the wider seascape and that MPAs and other effective area-based conservation measures are replicated within each

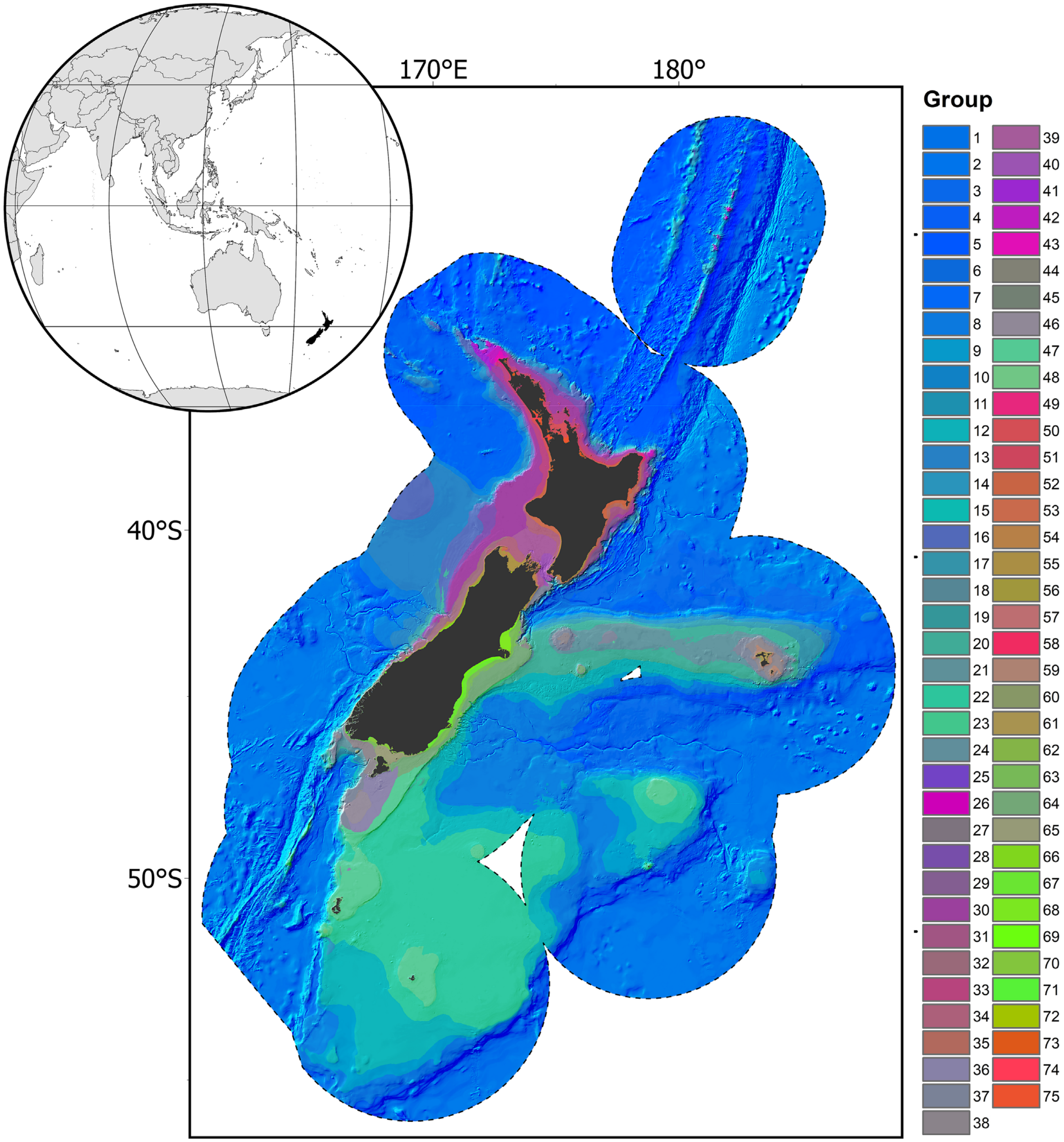

FIGURE 1 The New Zealand Seafloor Community Classification (75 groups) adapted from Stephenson et al. (2022). Colours broadly correspond to similarities/differences in predicted compositional turnover (i.e., similar colours represent similar communities). The NZSCC is freely available for visualization and download from the New Zealand Department of Conservation's Marine Data Portal available at https://doc-marine-data-deptconservation.hub.arcgis.com/.

classification group. For example, the NZSCC has been used to evaluate the design of bottom trawl closures within the Hauraki Gulf Marine Park (Bennion et al., 2023). However, to support the use of the NZSCC by central government, local and regional councils, and other parties interested in the design and implementation of MPAs and other effective area-based conservation measures in New Zealand, a robust maintenance framework for the NZSCC is required. This framework should set out when and how to update the NZSCC to ensure it remains the best available information for informing decision-making.

A key component of a maintenance framework for the NZSCC is being able to assess the ability of the classification to represent (discriminate between) different seafloor communities. This ability is important because understanding the performance of the classification in describing biodiversity patterns is key in both understanding if and how the classification should be used to inform management decisions (i.e., how much confidence to place in the classification), and determining if the classification should be updated. Should the representation of different seafloor communities by the NZSCC be less than desirable, or additional data becomes available to rerun the classification, then it will be important to consider how and when the NZSCC could be updated.

Stephenson et al. (2022) assessed the NZSCC's ability to discriminate across classification levels using the biological data included in the classification in an analysis of similarities test (ANOSIM) (i.e., using internal data validation). This test showed that the 75-group classification explained the most variation with the fewest number of groups based on data used to develop the models. However, validation of statistical models should ideally be undertaken with independent validation data (Friedman et al., 2001). In this study, we validate the NZSCC using independent evaluation data for demersal fish and benthic invertebrates (i.e., data that was not used to develop the classification). Additionally, the availability of abundance estimates for these evaluation datasets provides the added opportunity to test whether the NZSCC—which was developed using presence-absence data—can reflect abundance-weighted seafloor communities. Should the NZSCC reflect patterns in communities accounting for abundance this would provide additional utility for managers since the understanding of spatial patterns in species' abundances facilitates the identification of the most important areas for marine protection (Tokeshi, 1993).

2 | METHODS

2.1 | Biological data

To validate the NZSCC, temporally independent evaluation data from across the area covered by the NZSCC were collated for two of the four biotic groups used to generate the NZSCC: demersal fish and benthic invertebrates. Although the validation used here were not used in the development of the NZSCC, some of these data may have been collected in the same spatial location (including for another biotic group or using another sampling method) and therefore may not be completely spatially independent. However, given that we assume that the majority of these data are independent, we refer to this data as 'independent evaluation data' throughout.

Demersal fish biomass data were extracted from the New Zealand Ministry for Primary Industry's database *trawl* (for dates 01/01/2017 – 01/01/2022, MPI rep log 14760). Only samples using bottom trawls, undertaken within the New Zealand Territorial Sea and Exclusive Economic Zone—hereafter referred to as the New Zealand marine environment—were retained for analysis. Demersal fish biomass (kg) was converted to presence-absence. As a second step to assess whether the NZSCC could also represent variation in abundance-weighted community data, demersal fish biomass was standardized by trawl swept area (kg km^{-2}, with area calculated as the product of fishing gear door width and the distance fished for each trawl). The final demersal fish evaluation dataset contained presence-absence and standardized biomass for 268 species from 4099 bottom trawls (Figure 2a).

Benthic invertebrate densities (number of individuals per 1000 m^2) from seafloor imagery data collected using National Institute of Water & Atmospheric Research's Deep Towed Imaging System (DTIS) were collated from multiple surveys between 2007 and 2020 (Anderson et al., 2023, 2020; Bowden et al., 2019, and references therein). Only samples contained within the New Zealand marine environment were retained. No further data grooming of this dataset was necessary because these data were collated, groomed and the taxonomy standardized as part of a New Zealand Department of Conservation project (Anderson et al., 2023). The final benthic invertebrate evaluation dataset contained presence-absence and standardized densities for 74 taxa from 735 DTIS transects (Figure 2b).

For simplicity, we refer to standardized demersal fish biomass and benthic invertebrate densities as *abundance* from this point forward.

2.2 | Statistical validation

Validation of the NZSCC using demersal fish and benthic invertebrate presence-absence data was undertaken using two different approaches.

First, the classification strength of the 75-group NZSCC was assessed. This assessment was undertaken separately for demersal fish and benthic invertebrates following the approach described in Stephenson et al. (2022) based on methods developed by Snelder et al. (2007) and Bowden et al. (2011). Briefly, classification strength was assessed using an analysis of similarities test (ANOSIM) (Clarke & Warwick, 2001) of the multivariate presence-absence taxonomic data tagged with NZSCC groups based on spatial location. The classification strength is measured as the global R statistic, which was calculated as the difference in ranked biological similarities arising from all pairs of replicate sites between different groups, and the average of all rank similarities within groups, adjusted by the total number of sites. Global-R is equal to 1 if all replicates within groups are more like each other than any replicates from different groups and is approximately 0 if there is no group structure. Significance levels of the ANOSIM statistics were tested with a randomization procedure based on the null hypothesis of no group structure. All ANOSIM analyses were undertaken in R v4.0.3 (R Core Team, 2020) using the *Vegan* (v 2.6) package (Oksanen et al., 2022). Only groups with >5 unique occurrences were included in the analysis as per Stephenson et al. (2022).

Second, the relationship between community composition (using presence-absence data) and predicted compositional turnover

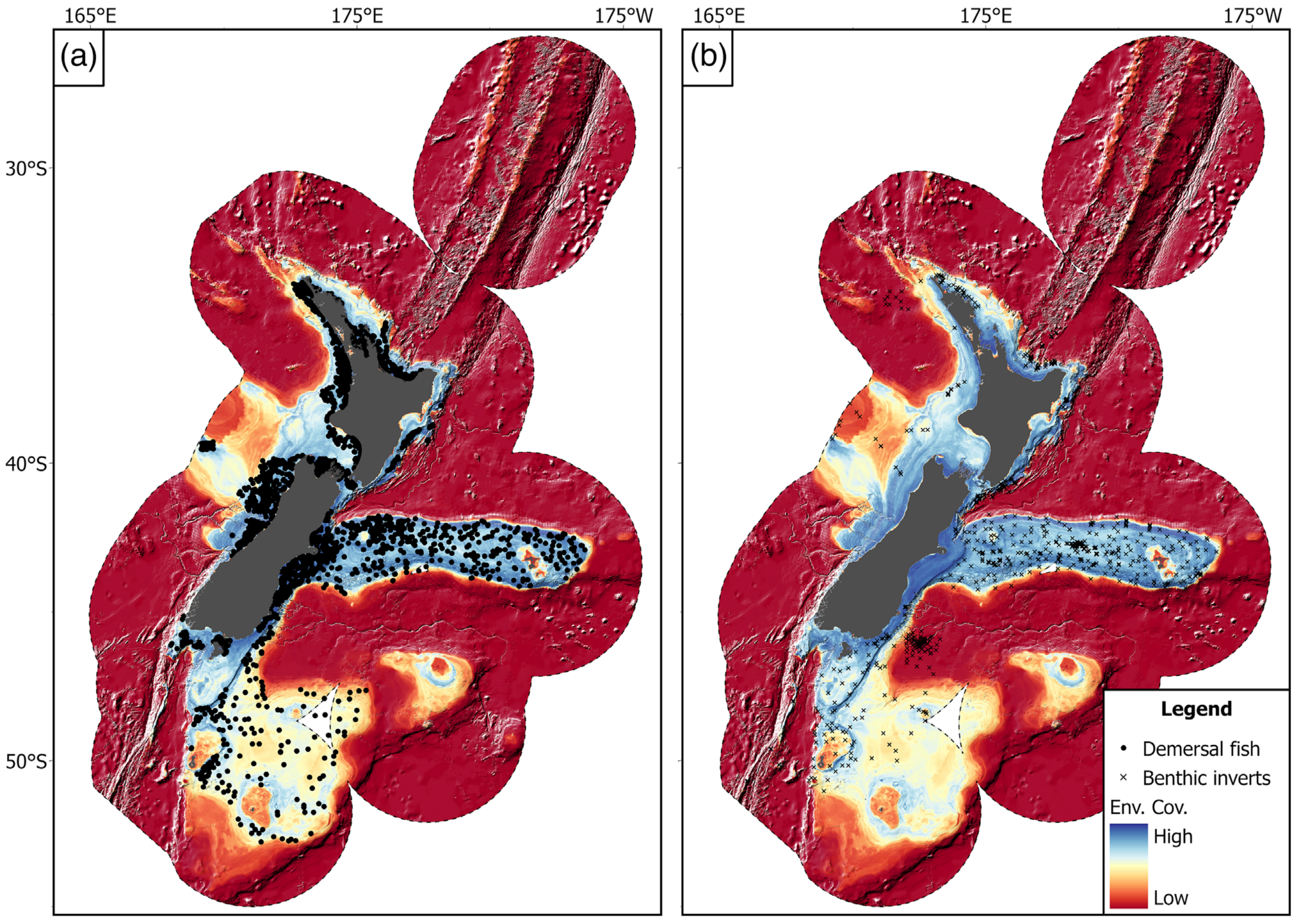

FIGURE 2 Location of independent evaluation data for demersal fish (a) and benthic invertebrates (b) overlaid on the NZSCC environmental coverage layer. Environmental coverage depicts the predicted confidence that can be placed in the predictions of compositional turnover underpinning the NZSCC based on the number and location of the biotic records used to train the model. Values range from low (i.e., no samples in the dataset with those environmental conditions: low confidence in predictions) to high (i.e., many samples with those environmental conditions: high confidence in predictions) within the New Zealand marine environment.

(outputs from the GF model) was explored. This assessment was undertaken separately for demersal fish and benthic invertebrates following modified methods described in Stephenson et al. (2018). Briefly, extended biological dissimilarities (De'ath, 1999) were calculated among all sample locations (using the functions *vegdist* and *stepacross* implemented in the *Vegan* package in R) with the Bray–Curtis dissimilarity measure and shortest dissimilarity set to 0.8. That is, extended biological dissimilarities were only calculated when Bray–Curtis dissimilarity values were greater than 0.8. The use of extended biological distances improves ordinations with high beta diversity, that is, when there are many sites with no species in common (De'ath, 1999). Correlations (Mantel test, implemented in the *Vegan* package) were then calculated between these biological dissimilarities and the equivalent distances in the predicted compositional turnover from the NZSCC (Euclidean distance calculated using the multivariate estimate of transformed environmental variable layers—termed environmental distance here). Relationships between extended biological dissimilarities in the evaluation data and the paired environmental distances from the NZSCC were then visualized in scatter plots. Because of the size of the dissimilarity matrices, a random subset of 50,000 points was selected for these plots (approximately 7% of the total number available for demersal fish and 20% for benthic invertebrates; multiple iterations of random subsets were generated and plotted with no visible differences in the relationship). Finally, to explore whether the relationship between the community composition and predicted compositional turnover was affected by the statistical uncertainty from the GF modelling that underpins the NZSCC, each point in the scatterplot was coloured according to the environmental coverage, and separately, the standard deviation of the predicted species compositional turnover.

Finally, to explore whether the NZSCC represents variation in abundance-weighted communities, both validation approaches were repeated but using the abundance estimates of demersal fish and benthic invertebrates. All analyses and R code are available in Github (see data availability statement).

3 | RESULTS

3.1 | Description of evaluation data

3.1.1 | Demersal fish

Demersal fish evaluation data (4099 bottom trawls) occurred across a broad range of water depths (10–1310 m) and 49 of the 75 NZSCC groups; 41 of these groups having more than 5 trawl samples (the minimum required for the ANOSIM analysis). These groups were 9, 12, 13, 16, 17, 18, 20, 21, 22, 23, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 39, 40, 45, 47, 50, 51, 52, 53, 54, 55, 56, 57, 63, 64, 65, 67, 68, 69, 70, 73, and 75. See Petersen et al. (2021) for descriptions of these groups. Bottom trawls occurred primarily in parts of the environmental space that would be considered well represented in the NZSCC ($n = 3718$ with environmental coverage > 0.5). That is, in areas with high environmental coverage which would indicate a higher confidence in spatial predictions from the GF models. However, some bottom trawls ($n = 360$) occurred in areas of moderate environmental coverage (0.1–0.5) and a few ($n = 21$) in areas of low environmental coverage (<0.1).

3.1.2 | Benthic invertebrates

Benthic invertebrate evaluation data (735 DTIS transects) occurred across a broader range of water depths than demersal fish (47–2989 m), but due to the lower number of samples, these occurred across a narrower range of NZSCC groups ($n = 32$ for benthic invertebrates) and fewer groups had more than 5 samples ($n = 22$). The groups with sufficient data to assess classification strength using benthic invertebrate evaluation data were: 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 20, 21, 22, 28, 30, 36, 37, 47 and 64. See Petersen et al. (2021) for descriptions of these groups. Benthic invertebrate data were more evenly spread across the environmental coverage than the demersal fish data were ($n = 507$ in areas with high environmental coverage; $n = 117$ in areas with moderate environmental coverage; and $n = 111$ areas with low environmental coverage).

3.2 | Statistical validation of the NZSCC: Presence-absence

There were sufficient data to assess the discriminatory power of 55% and 29% of the NZSCC groups using demersal fish and benthic invertebrate evaluation data, respectively (Table 1). Given the large extent of the New Zealand marine environment, it is unlikely that evaluation data would cover all NZSCC groups (with sufficient replication) without a targeted sampling programme. To put this into context, the large dataset used to develop the NZSCC (630,997 records across the four biotic groups occurring at 39,766 unique locations) only covered 76% of NZSCC groups for the demersal fish and 91% of NZSCC groups for the benthic invertebrates (Stephenson et al., 2022).

The global ANOSIM R values were 0.53 for demersal fish and 0.46 and benthic invertebrates, and both were significant at the 1% level, indicating that the NZSCC groups define biologically distinctive environments as assessed by completely independent evaluation data (Table 1). For context, the global R values for the NZSCC, as assessed by use of the internal training data by Stephenson et al. (2022) as part of the development of the NZSCC, was somewhat higher for demersal fish (R value: 0.72) but lower for benthic invertebrates (R value: 0.25).

Of the NZSCC groups with sufficient independent evaluation data, the proportion of significant inter-group differences were very high (95% and 97% for demersal fish and benthic invertebrates, respectively—Table 1) suggesting that NZSCC groups were distinct from each other in their taxonomic composition. Pairwise differences between the 75 groups of the NZSCC for demersal fish and benthic invertebrates using presence-absence and abundance data are provided in Data S1. Group codes (columns and rows) are colour coded to match the colours used in Figure 1 and Stephenson et al. (2022) and broadly represent similarities (or differences) in predicted compositional turnover; similar colours representing similar communities. Of particular importance when interpreting these pairwise plots is identifying groups which are not significantly different from other groups that are expected to be very different in their species composition. For example, Group 36 was not significantly different to a wide variety of other groups using the demersal fish presence-absence data, noting that Group 36 shares similar environmental conditions to most groups (i.e., it occurs in the middle of the transformed environmental space) but in contrast discriminates well when using the benthic invertebrate presence-absence dataset (Data S1).

The relationship between community composition and predicted compositional turnover was positive for both demersal fish (Mantel $r = 0.66$) and benthic invertebrates (Mantel $r = 0.32$) (Figure 3a,c). That is, as community composition increased in dissimilarity (extended Bray–Curtis dissimilarity) so too did the environmental distance (i.e., the dissimilarity in the predicted community compositions from the NZSCC). There were no obvious patterns (clustering) in

TABLE 1 Results of the pairwise analysis of similarities test (ANOSIM) analysis of the NZSCC using presence-absence evaluation data for demersal fish and benthic invertebrates.

Biotic group	Proportion of NZSCC groups >5 unique occurrences	Proportion of significant inter-group differences	ANOSIM R value
Demersal fish	0.55	0.95	0.53
Benthic invertebrates	0.29	0.97	0.46

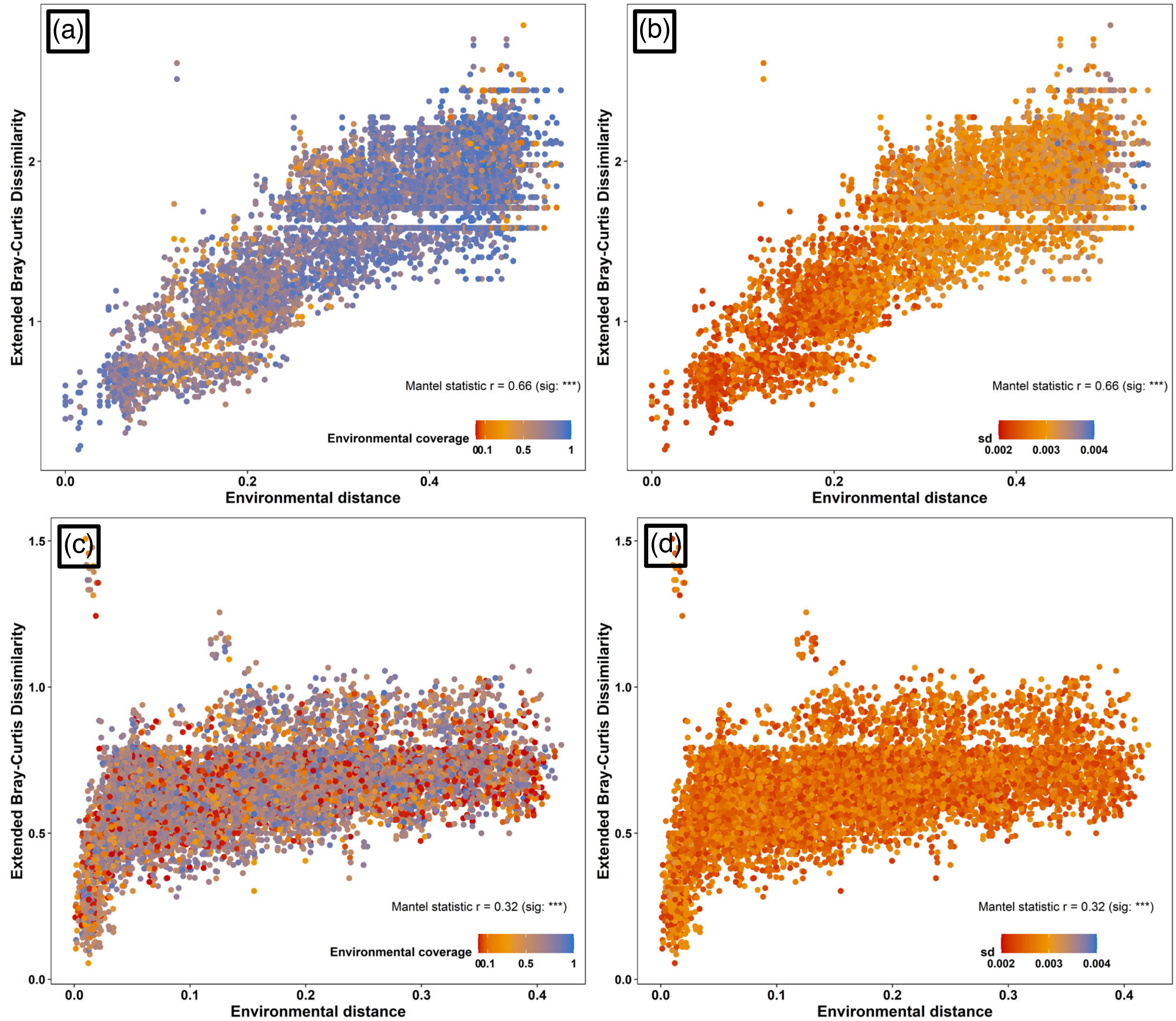

FIGURE 3 Extended Bray–Curtis dissimilarity against distance in transformed environmental space (Euclidean distance) of randomly sampled presence-absence samples (n = 50,000) for demersal fish (a, b) and benthic invertebrates (c, d). Colours represent measures of uncertainty: environmental coverage (a, c) and standard deviation of mean compositional turnover (b, d). Separations of points at Bray–Curtis dissimilarity values at 0.8 and 1.6 show where dissimilarity values were extended (following methods by De'ath, 1999—see methods for greater detail).

environmental coverage values for either demersal fish or benthic invertebrates (Figure 3a,c). If the relationship between community compositions and predicted compositional turnover were affected by the uncertainty associated with low sampling in certain environments (i.e., low environmental coverage), then we would expect to see red/orange points distributed on the edges of the point cloud in Figure 3a,c. Similarly, there was no pattern in the distribution of the standard deviation of mean compositional turnover across the benthic invertebrate community composition and predicted compositional turnover (Figure 3d). However, there appeared to be a slight pattern of increasing standard deviation of mean compositional turnover for samples with increasing demersal fish compositional dissimilarity and increasing environmental distance (Figure 3b). That is, there was greater variability in the prediction of environmental distance at higher values, noting that this variability was very low compared with the mean prediction (i.e., maximum standard deviation of the mean was 0.004 compared with a predicted intergroup distance of more than 0.4) (Figure 3c).

3.3 | Statistical validation of the NZSCC: Abundance

Discriminatory power of the NZSCC groups (global ANOSIM R values) using abundance data was similar to those using presence-absence (albeit somewhat lower for the benthic invertebrates, Table 2). Global R values were 0.53 and 0.38 (both significant at the 1% level) for demersal fish and benthic invertebrates, respectively, indicating that

TABLE 2 Results of the pairwise analysis of similarities test (ANOSIM) analysis of the NZSCC using standardized abundance evaluation data for demersal fish and benthic invertebrates.

Biotic group	Proportion of groups >5 unique occurrences	Proportion of significant inter-group differences	ANOSIM R value
Demersal fish	0.55	0.97	0.53
Benthic invertebrates	0.29	0.97	0.38

the NZSCC groups define biologically distinctive environments, even when using abundance data (Table 2). Finally, there was a higher proportion of significant inter-group differences when using abundance data than when compared with presence-absence data (97% for both demersal fish and benthic invertebrates, Table 2 and Data S1).

There was a positive relationship between abundance-weighted community compositions and predicted compositional turnover for both demersal fish and benthic invertebrates (Mantel $r = 0.59$ and 0.23, respectively, Figure 4a,c). This relationship was not as strong as when using presence-absence data, particularly for the benthic invertebrate abundance-weighted data where a much more variable relationship was observed (Figure 4c). Given that the abundance-weighted data and the presence-absence data are from the same locations, the same patterns of uncertainty were observed when using the weighted-abundance estimates as those already described in the previous section (Figure 4a–d).

4 | DISCUSSION

Here, we present two methods for validating the 75-group NZSCC environmental classification, which consider heterogeneity in environmental coverage and statistical uncertainty, and which used independent evaluation data for demersal fish and benthic invertebrates. The model validations did not seem to be affected by either statistical uncertainty measures of the NZSCC (environmental coverage and standard deviation of mean compositional turnover). Validation of the NZSCC provided promising results both in terms of the ability of the GF underlying the NZSCC to reflect patterns of compositional turnover and of the 75-group classification to reflect differences in community composition. In addition, similar (albeit slightly weaker for invertebrate evaluation data) patterns were also apparent when using abundance-weighted community data, illustrating the usefulness of the NZSCC to support management decisions by providing robust information on communities (since in reality species' abundance are key determinants of communities; Tokeshi, 1993).

4.1 | Statistical considerations

The analysis presented here represents an independent validation for two of the four biotic groups informing the NZSCC, using a (relatively) modest sample number (and a subset of taxa). The moderate/high R-values presented here describing the performance of the NZSCC in defining biologically distinctive groups as assessed by independent evaluation data (global ANOSIM R values of 0.46 and 0.53 for benthic invertebrates and demersal fish, respectively) are particularly encouraging. Previous evaluations of numeric seafloor classifications developed for the New Zealand marine environment—the New Zealand Marine Environment Classification (MEC; Snelder et al., 2007) and the Benthic Optimised Marine Environment Classification (BOMEC; Leathwick et al., 2012)—using independent evaluation data suggested that neither of these classifications were able to provide adequate spatial distribution or discrimination of benthic habitats and faunal assemblage composition (at least on the Chatham Rise and Challenger Plateau), with mean R values of less than 0.06 (Bowden et al., 2011). Here we provide further evidence (as in, e.g., Stephenson et al., 2018; Thomson et al., 2014) that the GF modelling approach is an effective approach for summarising spatial variation in species composition and turnover, including when combining models for several biotic groups (i.e., demersal fish, reef fish, benthic invertebrates and macroalgae).

The validation approach presented here provides the first step in assessing the predictive power of the NZSCC, and in determining whether, where, and for which biotic groups the classification may benefit from additional biological data. An important next step would be to extend the approach to other biotic groups and other datasets (i.e., repeat the analysis with macroalgal and reef fish datasets, when suitable data become available). To facilitate this step, the R code is provided in an online repository (Github) allowing routine re-assessment of NZSCC model fits as and when data becomes available. Future model validations could benefit from data that span the full environmental gradients found in New Zealand's marine environment to further explore the ability of NZSCC to represent communities in low or unsampled areas. The NZSCC may be better able to discriminate between communities in well sampled areas (i.e., areas where there is high environmental coverage) compared with areas that have been poorly sampled. It would therefore be useful to ensure, when possible, independent validation data comes from areas that encompasses the full breadth of environmental coverage and statistical uncertainty values. Ideally, independent data would be available for all four biotic groups (demersal fish, reef fish, benthic invertebrates and macroalgae) across all NZSCC groups where they are expected to occur (e.g., macroalgae do not occur below the photic zone and therefore will not be present in all NZSCC groups) and distributed across the full range of values in the uncertainty layers. A stratified approach for future collection of evaluation data could include samples from all NZSCC groups which, within each group,

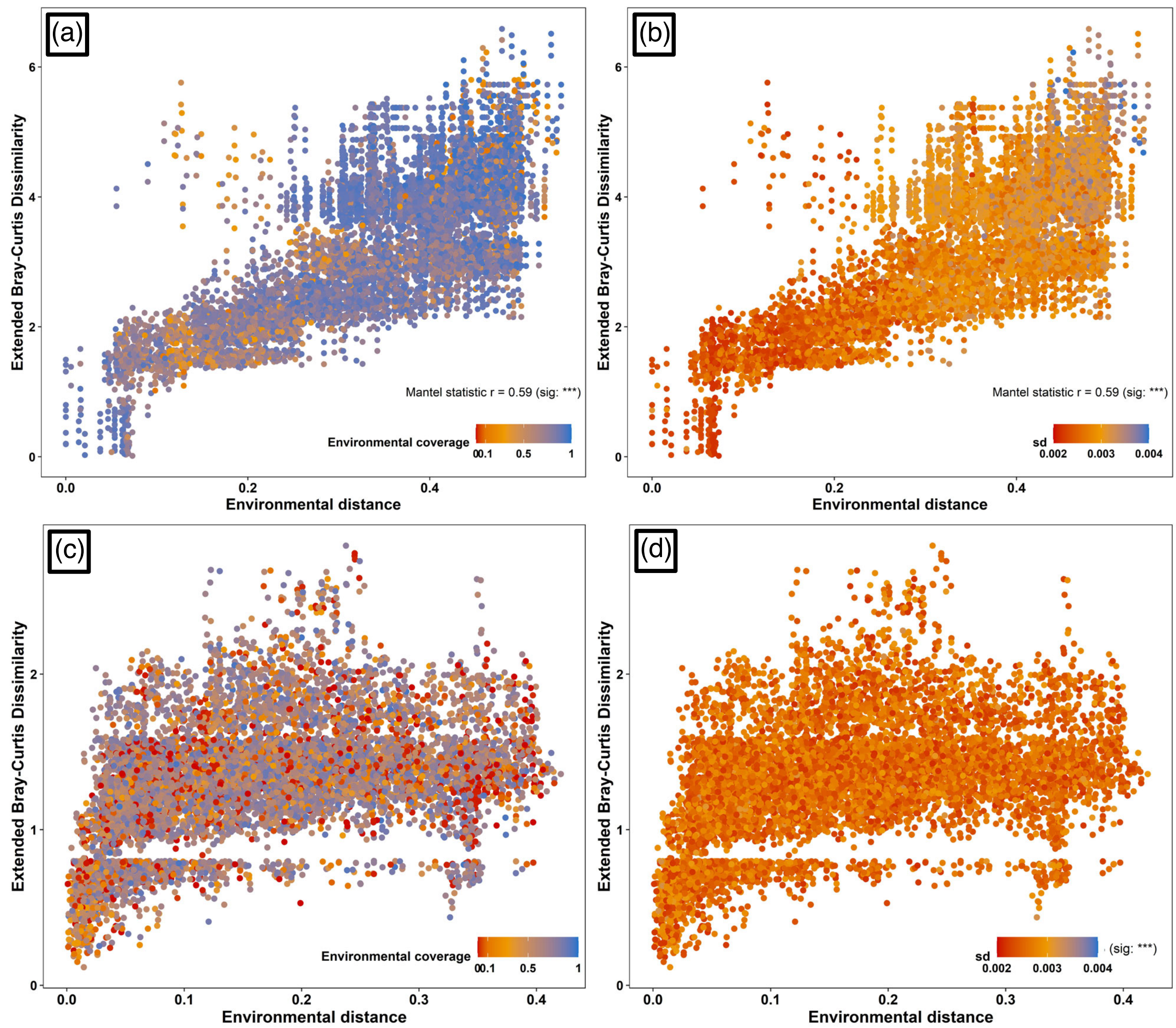

FIGURE 4 Extended Bray–Curtis dissimilarity against distance in transformed environmental space (Euclidean distance) of random abundance samples ($n = 10{,}000$) for demersal fish (a, b) and benthic invertebrates (c, d). Colours represent measures of uncertainty: environmental coverage (a, c) and standard deviation of mean compositional turnover (b, d). Separations of points at Bray–Curtis dissimilarity values at 0.8 and 1.6 show where dissimilarity values were extended (following methods by De'ath, 1999—see methods for greater detail).

span the gradient of both spatially explicit uncertainty estimates (standard deviation of the predicted species compositional turnover and environmental coverage).

4.2 | Implications for conservation and management

One key strength of classification-based approaches is that they can be created at various hierarchical (nested) levels of group-detail, for example, as a 9-group bioregionalization (Stephenson et al., 2023), to the 75-group community classification (Stephenson et al., 2022), a feature that makes them particularly useful when they need to be applied at differing spatial scales (national to regional to local scales) (Stephenson et al., 2020). The model validation approach developed here can equally be used to assess the classification strength at different spatial scales (and for different classification levels), including for defining classifications with a higher number of groups (200+ groups) which may be more appropriate for regional scale management planning (nested within broader scale classifications), particularly for inshore areas where there is greater heterogeneity in environmental conditions and biological communities. In addition, despite several other national/regional environmental classifications developed and used to support spatial planning internationally, e.g., EUNIS (Galparsoro et al., 2012); bioregions for the Southwest Pacific (Beger et al., 2020) or bioregionalization of the Ross Sea (Sharp et al., 2010), few have had their ability to discriminate communities evaluated, nor assessed numerically. Despite using different methods to develop these classifications, the evaluation approach to assess differences in community composition (or other classification grouping) described here could be applicable to these other classifications and locations.

The evaluation of the predictive power of the NZSCC, combined with the periodic collation of new biological data and updating spatially explicit environmental variables allows, if necessary, the generation of a new NZSCC. The exact number of additional (newly collated) biotic samples that would provide improved predictive power to the NZSCC is unknown. However, when the NZSCC has poor evaluation scores using new samples (i.e., repeating the evaluation process presented here), or if new samples provide additional information for rare species (i.e., when the addition of samples results in new taxa included in the GF models) or occur in previously low or unsampled parts of the New Zealand marine environment [i.e., locations with low (<0.1) to moderate (0.1–0.5) environmental coverage], it will be important to further explore whether these additional samples would affect estimates of compositional turnover. In the first instance, the approaches described here can be used to assess if the new data indicates the NZSCC is performing well, or if the data provides any important additional coverage that could improve model performance. Where the approaches described here indicate potential improvements in estimates of compositional turnover, new GF models should be generated and the same assessment approaches used to evaluate their predictive ability (e.g., by retaining approximately 10% of independent data evenly spread across geographic space). Where they perform better than the existing NZSCC model, they can be carried forward as an update to the NZSCC. In this way, the approaches described here can be used to iteratively evaluate the performance of the NZSCC and identify when it should be updated to ensure management decisions are based on best available information. For instance, shifts in the composition of the classification's groupings resulting from incorporation of new biological data could suggest the effectiveness of previously implemented conservation measures may need to be re-evaluated. That is, where the conservation objective of a MPA network is to protect a representative range of seafloor communities, the representativeness of the MPA network may need to be re-evaluated if updates to the classification results in changes to the classification groupings. Where this results in a decrease in the effectiveness of a MPA network in protecting a representative range of seafloor communities, a management decision could be taken to adjust MPA boundaries or implement additional MPAs to achieve representativity objectives. Given past rates of data collection in New Zealand, a reasonable timeframe for assessing the NZSCC could be every 5–10 years. However, the implementation of the classification evaluation approach to date, suggests that the existing NZSCC is currently fit-for-purpose for informing management decisions, and provides a templar for other regions/countries exploring the development of environmental classifications to support spatial planning.

5 | SUPPLEMENTARY MATERIALS

The excel file "Supplementary Materials—Pairwise differences.xls" provides pairwise differences between the 75 groups of the NZSCC for demersal fish and benthic invertebrates using presence-absence and abundance data. Group codes (columns and rows) are colour coded to match the colours used in Figure 1 and Stephenson et al. (2022) and broadly represent similarities (or differences) in predicted compositional turnover; similar colours representing similar communities. Cells labelled as "ns" represent non-significant differences in community composition (p-value > 0.05); *p-value 0.01–0.05; **p-value 0.001–0.01; ***p-values < 0.001.

AUTHOR CONTRIBUTIONS

Fabrice Stephenson: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; software; supervision; visualization; writing—original draft preparation; writing—review and editing. **Jordi Tablada**: Conceptualization; data curation; funding acquisition; software; validation; writing—original draft preparation; writing—review and editing. **Ashley A. Rowden**: Conceptualization; methodology; writing—review and editing. **Richard Bulmer**: Conceptualization; supervision; writing—review and editing. **David A. Bowden**: Conceptualization; methodology; writing—review and editing. **Shane W. Geange**: Conceptualization; data curation; funding acquisition; supervision; writing—original draft preparation; writing—review and editing.

ACKNOWLEDGMENTS

This work was funded by the New Zealand Department of Conservation (contract number: NOF-BIO-317). We thank the New Zealand Department of Conservation (DOC), Fisheries New Zealand (FNZ) and the National Institute of Water and Atmospheric Research (NIWA) for providing the DTIS imagery data (which was collated as part of DOC project POP2021-02) and the demersal bottom trawl data (*trawl* database). Specifically, we thank Owen Anderson, Lyndsey Holland and Richard O'Driscoll for input on the data used in the analyses.

CONFLICT OF INTEREST STATEMENT

The authors have no relevant financial or non-financial interests to disclose.

DATA AVAILABILITY STATEMENT

The data generated in this research will be shared on reasonable request to the corresponding author. The R code used in this research and data for demersal fish are available in the GitHub open repository: https://github.com/Fabrice-Stephenson/Independent_statistical_validation_of_the_NZSCC.

ORCID

Fabrice Stephenson https://orcid.org/0000-0002-9500-5204

REFERENCES

Anderson, O.F., Pallentin, A., Bowden, D.A., Chin, C., Davey, N., Eton, N. et al. (2020). *Quantifying benthic biodiversity—phase II: a factual voyage report from RV Tangaroa voyage TAN2004 to Campbell plateau 17 may–7 June 2020*. New Zealand Aquatic Environment and Biodiversity Report, No. 264, p. 40.

Anderson, O., Schnabel, K., Bowden, D., Davey, N. & Hart, A. (2023). *Identification of protected coral hotspots using species distribution modelling*. Report Prepared by NIWA for Project POP2021-02, Conservation Services Programme, Department of Conservation.

Beger, M., Wendt, H., Sullivan, J., Mason, C., LeGrand, J., Davey, K. et al. (2020). National-scale marine bioregions for the Southwest Pacific. *Marine Pollution Bulletin*, 150, 110710. https://doi.org/10.1016/j.marpolbul.2019.110710

Bennion, M., Brough, T., Leunissen, E., Morrison, M., Hillman, J., Hewitt, J.E. et al. (2023). Exploring the use of spatial decision-support tools to identify trawl corridors in the Hauraki gulf Marine Park. *New Zealand Aquatic Environment and Biodiversity Report*, No. 306, 101.

Bowden, D.A., Anderson, O., Escobar-Flores, P., Rowden, A. & Clark, M. (2019). Quantifying benthic biodiversity: using seafloor image data to build single-taxon and community distribution models for Chatham rise, New Zealand. *Aquatic Environment and Biodiversity Report*, No. 235, 67.

Bowden, D.A., Compton, T., Snelder, T. & Hewitt, J. (2011). *Evaluation of the New Zealand marine environment classifications using ocean survey 20/20 data from Chatham rise and challenger plateau*. Ministry of Fisheries.

Clarke, K.R. & Warwick, R.M. (2001). *Change in marine communities: an approach to statistical analysis and interpretation*. 2nd edition, Plymouth: PRIMER-E.

Costello, M.J. (2009). Distinguishing marine habitat classification concepts for ecological data management. *Marine Ecology Progress Series*, 397, 253–268. https://doi.org/10.3354/meps08317

Day, J., Dudley, N., Hockings, M., Holmes, G., Laffoley, D.D.A., Stolton, S. et al. (2012). *Guidelines for applying the IUCN protected area management categories to marine protected areas*. IUCN.

De'ath, G. (1999). Extended dissimilarity: a method of robust estimation of ecological distances from high Beta diversity data. *Plant Ecology*, 144 (2), 191–199. https://www.jstor.org/stable/20050827

Dixon-Bridges, K., Hutchings, P. & Gladstone, W. (2014). Effectiveness of habitat classes as surrogates for biodiversity in marine reserve planning. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 24 (4), 463–477. https://doi.org/10.1002/aqc.2377

Ellis, N., Smith, S.J. & Pitcher, C.R. (2012). Gradient forests: calculating importance gradients on physical predictors. *Ecology*, 93, 156–168. https://doi.org/10.1890/11-0252.1

Friedman, J., Hastie, T. & Tibshirani, R. (2001). *The elements of statistical learning*. Springer series in statistics, New York.

Galparsoro, I., Connor, D.W., Borja, Á., Aish, A., Amorim, P., Bajjouk, T. et al. (2012). Using EUNIS habitat classification for benthic mapping in European seas: present concerns and future needs. *Marine Pollution Bulletin*, 64. https://doi.org/10.1016/j.marpolbul.2012.10.010

Lade, S.J., Steffen, W., de Vries, W., Carpenter, S.R., Donges, J.F., Gerten, D. et al. (2020). Human impacts on planetary boundaries amplified by earth system interactions. *Nature Sustainability*, 3(12), 119–128. https://doi.org/10.1038/s41893-019-0454-4

Leathwick, J., Rowden, A., Nodder, S., Gorman, R., Bardsley, S., Pinkerton, M. et al. (2012). A benthic-optimised marine environment classification (BOMEC) for New Zealand waters. In: *New Zealand aquatic environment and biodiversity report*; ISSN 1176–9440, Wellington: Ministry of Fisheries.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R. et al. (2022). *Package 'vegan'*. Community Ecology Package. Version 2, 1–295.

Petersen, G., Stephenson, F., Brough, T. & Rowden, A. (2021). *Seafloor community classification: group descriptions*. NIWA Report Prepared for the Department of Conservation (DOC), July 2020.

Pitcher, R.C., Lawton, P., Ellis, N., Smith, S.J., Incze, L.S., Wei, C.L. et al. (2012). Exploring the role of environmental variables in shaping patterns of seabed biodiversity composition in regional-scale ecosystems. *Journal of Applied Ecology*, 49(3), 670–679. https://doi.org/10.1111/j.1365-2664.2012.02148.x

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. In: R foundation for statistical computing. Vienna, Austria.

Sharp, B., Parker, S., Pinkerton, M., Breen, B., Cummings, V., Dunn, A. et al. (2010). Bioregionalisation and spatial ecosystem processes in the Ross Sea region. Working Groupon ecosystem monitoring and Management-10/30. *Commission for the Conservation of Antarctic Marine Living Resources*.

Smale, D.A., Wernberg, T., Oliver, E.C., Thomsen, M., Harvey, B.P., Straub, S.C. et al. (2019). Marine heatwaves threaten global biodiversity and the provision of ecosystem services. *Nature Climate Change*, 9, 306–312. https://doi.org/10.1038/s41558-019-0412-1

Snelder, T.H., Leathwick, J.R., Dey, K.L., Rowden, A.A., Weatherhead, M.A., Fenwick, G.D. et al. (2007). Development of an ecologic marine classification in the New Zealand region. *Environmental Management*, 39, 12–29. https://doi.org/10.1007/s00267-005-0206-2

Stephenson, F., Bulmer, R., Leathwick, J., Brough, T., Clark, D., Greenfield, B. et al. (2021). *Development of a New Zealand seafloor community classification (SCC)*NIWA report prepared for Department of Conservation (DOC). Hamilton: National Institute of Water & Atmospheric Research. https://www.doc.govt.nz/globalassets/documents/conservation/marine-and-coastal/marine-protected-areas/development-of-new-zealand-seafloor-community-classification.pdf

Stephenson, F., Leathwick, J.R., Francis, M.P. & Lundquist, C.J. (2020). A New Zealand demersal fish classification using gradient Forest models. *New Zealand Journal of Marine and Freshwater Research*, 54(1), 60–85. https://doi.org/10.1080/00288330.2019.1660384

Stephenson, F., Leathwick, J.R., Geange, S.W., Bulmer, R.H., Hewitt, J.E., Anderson, O.F. et al. (2018). Using gradient forests to summarize patterns in species turnover across large spatial scales and inform conservation planning. *Diversity and Distributions*, 24(11), 1641–1656. https://doi.org/10.1111/ddi.12787

Stephenson, F., Leathwick, J.R., Geange, S., Moilanen, A., Pitcher, C.R. & Lundquist, C.J. (2021). Species composition and turnover models provide robust approximations of biodiversity in marine conservation planning. *Ocean and Coastal Management*, 212. https://doi.org/10.1016/j.ocecoaman.2021.105855

Stephenson, F., Rowden, A.A., Brough, T., Petersen, G., Bulmer, R.H., Leathwick, J.R. et al. (2022). Development of a seafloor community classification for the New Zealand region using a gradient Forest approach. *Frontiers in Marine Science*, 8. https://doi.org/10.3389/fmars.2021.792712

Stephenson, F., Rowden, A.A., Tablada, J., Tunley, K., Brough, T., Lundquist, C.J. et al. (2023). A seafloor bioregionalisation for New Zealand. *Ocean and Coastal Management*, 242, 106688. https://doi.org/10.1016/j.ocecoaman.2023.106688

Stevens, T. & Connolly, R.M. (2004). Testing the utility of abiotic surrogates for marine habitat mapping at scales relevant to management. *Biological Conservation*, 119(3), 351–362. https://doi.org/10.1016/j.biocon.2003.12.001

Sutcliffe, P.R., Klein, C.J., Pitcher, C.R. & Possingham, H.P. (2015). The effectiveness of marine reserve systems constructed using different surrogates of biodiversity. *Conservation Biology*, 29(3), 657–667. https://www.jstor.org/stable/24483098

Thomson, R.J., Hill, N.A., Leaper, R., Ellis, N., Pitcher, C.R., Barrett, N.S. et al. (2014). Congruence in demersal fish, macroinvertebrate, and macroalgal community turnover on shallow temperate reefs. *Ecological Applications*, 24(2), 287–299. https://www.jstor.org/stable/24432146

Tokeshi, M. (1993). Species abundance patterns and community structure. In: *Advances in Ecological Research*: Elsevier, pp. 111–186.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Stephenson, F., Tablada, J., Rowden, A.A., Bulmer, R., Bowden, D.A. & Geange, S.W. (2024). Independent statistical validation of the New Zealand Seafloor Community Classification. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 34(3), e4114. https://doi.org/10.1002/aqc.4114