

Predictive habitat suitability modeling of *Nereocystis luetkeana* in San Juan County, WA

Emily Rose Bews

A thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Marine Affairs

University of Washington

2026

Committee:

Terrie Klinger

Tarang Khangaonkar

Program Authorized to Offer Degree:

School of Marine and Environmental Affairs

© Copyright 2026

Emily Rose Bews

University of Washington

Abstract

Predictive habitat suitability modeling of *Nereocystis luetkeana* in San Juan County, WA

Emily Rose Bews

Chair of the Supervisory Committee:

Terrie Klinger

School of Marine and Environmental Affairs

Nereocystis luetkeana (bull kelp) is a foundational component of nearshore ecosystems in the inland waters of Washington State and holds deep cultural significance for Coast Salish peoples. This study presents the first habitat suitability models focused specifically on *N. luetkeana* in San Juan County, Washington, developed in response to the need for improved spatial tools to support kelp monitoring and restoration planning in this subbasin. Two binomial Generalized Linear Mixed Effects Model (GLMM) spatial models were fit to 888 training points (293 presence, 595 absence) derived from 2024 Samish Indian Nation kelp canopy data, using the sdmTMB package in R. Oceanographic predictor variables were derived from standard (~1km) and high-resolution (~180m) configurations of the Salish Sea Model (SSM and SSM_{HR} respectively). The habitat suitability models (HSMs) constructed in this study were validated through spatial block cross-validation and independent validation using 2025 snorkel survey data (n=345). Effective fetch and slope were the dominant positive predictors of *N. luetkeana* habitat

in both the standard resolution habitat suitability model (HSM_{STD}) and high-resolution habitat suitability model (HSM_{HR}), and mobile substrate the strongest negative predictor, together reflecting the dependence of *N. luetkeana* on stable hard substrate in wave-exposed environments. HSM_{HR} resolved four additional predictor associations not clearly defined in HSM_{STD} and produced a substantially lower spatial random field standard deviation ($\sigma_0 = 9.10$ vs 13.97), indicating finer-resolution hydrodynamic predictor variables helped captured spatial structure that standard resolution variables could not resolve. Cross-validation area under the curve (AUC) score was 0.918 for HSM_{HR} and 0.895 for HSM_{STD}; independent validation AUC was 0.698 for HSM_{HR} and 0.722 for HSM_{STD}, with documented spatial and temporal mismatch between training and validation datasets that suggest these values underestimate true predictive performance. At a 0.5 probability threshold, HSM_{STD} predicted 2,241 ha and HSM_{HR} predicted 1,705 ha of suitable habitat across the 0-12m MLLW depth domain. These results demonstrate that the inclusion of higher-resolution hydrodynamic predictor data meaningfully improves habitat suitability predictions, particularly for site-level applications. The resulting suitability maps provide a spatial framework for prioritizing field surveys and identifying candidate restoration sites for Tribal and state resource managers.

Table of Contents

List of Figures	i
List of Tables	ii
Acknowledgements.....	iii
Introduction.....	1
Methods.....	4
Study Area.....	4
Bathymetric Data.....	5
Oceanographic Predictor Variables.....	6
Substrate and Fetch Predictor Variables	10
Presence/Absence Data and Predictor Extraction	10
Predictor Selection and Multicollinearity Assessment.....	12
Model Fitting, Variable Response, and Suitability Predictions	14
Model Validation.....	15
Results.....	18
Model Fit and Fixed Effects.....	18
Predictor Variable Response Curves.....	23
Predicted Habitat Distribution and Area Estimates.....	26
Model Validation.....	30
Discussion	35
Ecological Interpretation of Significant Predictors	35
Model Performance Comparison.....	39
Prediction Output Comparison and Uncertainty Estimates.....	42
Limitations.....	44
Conclusion	45
Appendix 1. Shoreline barrier feature repair locations	48
Appendix 2. Correlation matrix for HSM _{STD} and HSM _{HR} full predictor candidate sets	50
Appendix 3. Model comparisons under different predictor configurations.....	52
Appendix 4. Orcas island cross-validation diagnostics	53
Appendix 5. Distribution of predictor values at <i>N. luetkeana</i> presence and absence points	54
References.....	55

List of Figures

Figure 1. Model regression coefficients	22
Figure 2. Partial dependence response curves for key variables	25
Figure 3. Suitability prediction maps	27
Figure 4. Suitable habitat area estimates maps	28
Figure 5. Prediction uncertainty estimates maps	29
Figure 6. Spatial block cross-validation AUC	30
Figure 7. Independent validation AUC	32
Figure 8. Comparison of predictor variable distributions between models	33
Figure 9. Receiver operating characteristic (ROC) curves	34

List of Tables

Table 1. Seasonal periods for environmental covariates	7
Table 2. Final HSM _{STD} and HSM _{HR} predictor variables used in modeling	14
Table 3. Fixed-effect regression coefficients for all model predictors	20
Table 4. Summary of model parameters	21
Table 5. Summary of model discrimination performance (AUC values)	32

Acknowledgements

This research was made possible with the help of many wonderful people and institutions. First and foremost, I would like to thank my advisor, Dr. Terrie Klinger, for her mentorship and endless support. Terrie has helped me grow as a researcher and writer, and I am deeply grateful for the knowledge and kindness she shared with me, as well as her thoughtful guidance throughout this process. I would also like to thank my committee member, Dr. Tarang Khangaonkar, for his expertise and assistance in utilizing data from the Salish Sea Model. Not only were the resources he shared invaluable to this research, but his encouragement and excitement for this project propelled my exploration of the Salish Sea Model and the development of this work. Additionally, I would like to thank Dr. Tom Mumford; his knowledge and support helped guide this research, and his passion for ecology and phycology has been an inspiration throughout my journey.

Several others provided technical expertise and data that were essential to this research, including Samish Indian Nation, specifically Sophia Ammons and Todd Woodard; Salish Sea Modeling Center, specifically Su Kyong Yun; Washington State Department of Natural Resources, specifically Gray McKenna, Dr. Danielle Claar, and Tyler Cowdrey; and the Seattle Aquarium, specifically Megan Williams and Dr. Zachary Randell. I am truly grateful for the support and involvement of these individuals. I would also like to thank the Northwest Straits Foundation for the Caroline Gibson Scholarship Award that supported me during this research, along with the professional development opportunities provided by the foundation.

I would like to thank Friday Harbor Laboratories for providing funding and resources that helped support my research. Many researchers and staff at FHL provided guidance, planning assistance, and field support, specifically, Dr. Katie Dobkowski, Bob Oxborrow, Dr. Olivia

Graham, Dr. Megan Dethier, Pema Kitaeff, Jennifer Ortiz, Grace Leuchtenberger, and Meredith Dirkman. There were also many kind individuals and families that supported my field data collection by providing access to survey sites and offering boating support, specifically, Kari Koski (who also served as an invaluable buddy on many of my field surveys) and Doug McCutchen, Dr. George and Peggy Hunt, Dr. Trish Morse, Suzanne Regan and family, Dr. Jennifer Burnaford, Juan Lopez and family, Drs. Ana Mari Cauce and Susan L. Joslyn, and Dr. Kenneth Sebens.

In addition to those who directly contributed to my education and research, I would like to thank my wonderful friends who have been with me during this process. I would like to thank my family, my sister Alison, my mother Judy, and my father Alan for their love and encouragement that brought me to this point. Finally, I would like to thank my partner, Chase. His care, patience, humor, and positive spirit were an anchor for me during this time, and I consider myself so very lucky to have such a dedicated partner by my side. This study would not have been possible without him, along with all the people who contributed to my work in one way or another, and I am deeply grateful to all the people I have crossed paths with while completing this research.

Introduction

Kelp forests are among the most productive and ecologically important ecosystems in temperate marine environments globally, creating complex three-dimensional habitat that supports diverse food webs, drives nutrient and carbon cycling, and provides critical coastal protection (Eger et al., 2023; Hollarsmith et al., 2022; Krumhansl & Scheibling, 2012; Steneck et al., 2002; Teagle et al., 2017). Large brown algae in the order *Laminariales* form the structural foundation of these systems, and kelp forests shape the distribution and abundance of diverse assemblages of fish, invertebrates, and other marine organisms (Hauser et al., 2006; Pérez-Matus et al., 2025; Teagle et al., 2017). Kelp forests are experiencing widespread declines globally due to compounding stressors including ocean warming, marine heatwaves, trophic disturbances, and anthropogenic activity (Filbee-Dexter et al., 2020; Norderhaug et al., 2020; Smale, 2020; Straub et al., 2019).

In the inland waters of Washington State, *Nereocystis luetkeana*, commonly known as bull kelp, is the primary floating canopy-forming kelp species and as such serves as a foundational component of nearshore ecosystems throughout the region (Hollarsmith et al., 2022). *N. luetkeana* is an annual species that is characterized by a biphasic life history in which the large sporophyte alternates with a microscopic, overwintering gametophyte. New sporophytes germinate in late winter and early spring, reach peak canopy development over summer, and release spores in late summer and fall before the sporophyte senesces and is ultimately dislodged from the seafloor by seasonal storms (Springer et al., 2007). This annual life history makes *N. luetkeana* populations sensitive to interannual variability in environmental conditions, as successful reproduction and recruitment depend on favorable conditions across multiple seasonal windows, from winter conditions experienced by the microscopic gametophyte to summer

conditions affecting mature sporophyte growth, spore release in late summer, and subsequent sporophyte recruitment success in late winter and early spring (García-Reyes et al., 2022).

In addition to its ecological importance, *N. luetkeana* is also of deep cultural significance to the Coast Salish peoples, who have stewarded the lands and waters of the Pacific Northwest since time immemorial. Kelp has long been woven into the cultural identity, traditional practices, and ecological knowledge of indigenous communities throughout the region (Naar, 2020). In Washington State, Samish Indian Nation and other Coast Salish tribes are critical leaders and stewards of nearshore ecosystems, and the health of *N. luetkeana* populations is directly relevant to tribal subsistence, cultural practices, and resource management responsibilities.

N. luetkeana has experienced significant declines in distribution and abundance throughout Washington waters in recent decades, driven by compounding stressors associated with climate change and anthropogenic disturbance including warming ocean temperatures, altered nutrient dynamics, and habitat degradation (Berry et al., 2021; Calloway et al., 2020). However, changes in *N. luetkeana* abundance are highly variable by subbasin, reflecting the distinct environmental conditions across the complex geography of the Salish Sea (Calloway et al., 2020). Disparities in monitoring further complicate assessment (G. E. McKenna & Claar, 2026). To address challenges related to kelp loss in this context, researchers, resource managers, and conservation practitioners have developed a collaborative management framework focused on monitoring *N. luetkeana* populations, identifying stressors, quantifying long-term trends, and guiding conservation and restoration efforts throughout the region (Calloway et al., 2020).

The San Juan Island region represents an important yet understudied subbasin within this management context. This region currently lacks sufficient long-term monitoring data to definitively characterize trends in kelp abundance, although there is growing concern among

resource managers and indigenous communities that floating kelp has declined around the islands, based on management reports and indigenous and community knowledge (Samish Indian Nation, 2024). However, the spatial and temporal extent of any changes remains poorly quantified at this time. An annual monitoring program has recently been established in this region using fixed-wing aerial surveys to capture imagery of *N. luetkeana* canopy. Aerial surveys are conducted by Washington State Department of Natural Resources, with imagery digitized into bed extent polygons by Samish Indian Nation GIS staff (Samish Indian Nation GIS, 2024). This represents the most comprehensive spatial dataset of kelp distribution currently available to quantify trends in *N. luetkeana* distribution and persistence in the region.

The availability of this extensive canopy data is especially useful for habitat suitability modeling. Habitat suitability models can be used to predict the distribution of suitable habitat for a species by relating species occurrence data to environmental covariates to identify the conditions associated with species presence or absence (Bekkby & Moy, 2011; Gregr et al., 2019). These models are useful tools for exploring the potential distribution of a species, quantifying habitat quality, and prioritizing areas for monitoring and restoration investment (Questad et al., 2014). Several habitat suitability models for *N. luetkeana* have been developed in recent years for other regions of Washington State (G. McKenna et al., 2022; Williams et al., 2024), but no previous modeling effort has focused specifically on habitat suitability for *N. luetkeana* in the San Juan Islands.

This study addresses that gap by developing and comparing two binomial habitat suitability models for *N. luetkeana* in San Juan County that use predictors derived from two configurations of the Salish Sea Model: the standard and high spatial resolutions (interpolated at 1km and 180m respectively). Three research questions guided this work:

1. What is the distribution of suitable habitat for *N. luetkeana* in San Juan County, Washington based on environmental factors?
2. Which environmental factors are most strongly associated with *N. luetkeana*?
3. Does finer spatial resolution of hydrodynamic predictor data improve habitat suitability predictions compared with standard resolution data, in terms of spatial detail, estimated suitable habitat area, and predictive accuracy?

Methods

Two habitat suitability models (HSMs) were developed to compare the effect of hydrodynamic predictor resolution on predicted bull kelp habitat distribution within the study region. The first model (HSM_{STD}) used standard resolution hydrodynamic variables (temperature, current velocity, and salinity) derived from the Salish Sea Model (SSM), while the second (HSM_{HR}) used same variables derived from the high-resolution version of the Salish Sea Model (SSM_{HR}). Spatial operations and analyses were completed in ArcGIS Pro (version 3.6.0) and R (version 4.5.1).

Study Area

San Juan County encompasses the San Juan Islands, an archipelago within the waters of Washington State that is situated at the convergence of three major water masses in the Salish Sea: the Strait of Juan de Fuca, which connects the inland sea to the Pacific Ocean to the west; the Strait of Georgia, which carries Fraser River outflow southward from British Columbia; and Puget Sound which lies to the south. This convergence creates a mosaic of highly dynamic oceanographic conditions across a relatively small geographic area of the Salish Sea. Combined with strong tidal currents and complex bathymetry, there is substantial spatial variability in

temperature, salinity, nutrient availability, and hydrodynamic energy in this region across fine spatial scales.

The nearshore environments throughout the San Juan Islands support diverse marine assemblages, with *N. luetkeana* serving as an important component of the subtidal ecosystem (Christiaen et al., 2023). The rocky shorelines, strong currents, and nutrient-rich waters provide conditions generally favorable for the establishment and growth of kelp, though the high spatial variability in oceanographic conditions across this region means that kelp distribution is structured by local environmental dynamics.

The availability of detailed biological data and high-resolution oceanographic outputs from the Salish Sea Model, along with the ecological and restoration management relevance of the region, made San Juan County a well-suited study system for evaluating habitat suitability modeling approaches for *N. luetkeana*.

Bathymetric Data

Bathymetric data for the study region was sourced from two digital elevation models: the Coastal National Elevation Database (CoNED) at 1m resolution covering the San Juan Islands, and the NOAA NCEI Continuously Updated Digital Elevation Model (CUDEM) at 3m resolution (Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder, 2014; Washington State Department of Natural Resources, 2025). CUDEM was used to supplement coverage across a minor spatial gap in the northwest corner of the study extent. Both rasters were aligned to EPSG:2856 (NAD83 HARN/Washington State Plane North, meters) and the CUDEM raster was resampled to 1m resolution using bilinear interpolation before merging the DEMs in R, with CoNED values taking priority where coverage overlapped. The merged 1m DEM was used to derive slope (degrees, gradient method) and rugosity (Terrain

Ruggedness Index), although rugosity was ultimately dropped from the model following multicollinearity analyses. All three rasters were saved at 1m resolution for predictor extraction at kelp presence/absence points.

Oceanographic Predictor Variables

Seasonal oceanographic predictor variables were derived from 2023-2024 hindcasts of the standard resolution Salish Sea Model (SSM) and high-resolution Salish Sea Model (SSM_{HR}) using University of Washington's Hyak supercomputing cluster (Khangaonkar et al., 2018; Premathilake & Khangaonkar, 2022). The Salish Sea Model is a three-dimensional hydrodynamic and water quality model of the Salish Sea developed and maintained by the Salish Sea Modeling Center at the University of Washington Tacoma in collaboration with Pacific Northwest National Laboratory (Khangaonkar et al., 2018). The model is built on an unstructured triangular grid; variables are either computed at vertices (nodes) and at the center of the triangular cells (elements) across 10 vertical sigma-coordinate layers to model conditions throughout the water column. SSM is at a standard resolution with a mean node spacing of roughly 2km in the San Juan Islands, interpolated to 1km resolution for predictor extraction, approximately half the mean node spacing following standard practice for spatial interpolation. SSM_{HR} uses the same structure as SSM but at a finer resolution, with mean node spacing of approximately 180m in the nearshore environments of the San Juan Islands. A Python script was used to spatially subset model nodes within the study region and generate seasonal summaries of relevant variables from the hindcast solution files. Seasons were defined by ecological relevance to *N.luetkeana* life history: the temporal range was from the winter prior to canopy development to the following summer (Table 1).

Table 1. Seasonal periods used in data summarization as they relate to the life history periods of *Nereocystis luetkeana*

Seasonal Average	Months	Biological and Ecological Significance
Winter	October - February	Conditions influencing the microscopic gametophyte and sexual reproduction
Spring	March - June	Conditions influencing the developing juvenile sporophytes
Summer	April - September	Conditions influencing mature sporophytes

Hydrodynamic variables (HYD) extracted from SSM and SSM_{HR} for testing included temperature, salinity, current speed ($\sqrt{u^2 + v^2}$), and bed shear stress (TAUC). These variables were selected based on their effects on *N. luetkeana* growth and distribution (Eckman et al., 2003; Fales et al., 2023; Farrugia Drakard et al., 2025; Weigel et al., 2023). Seasonal summaries were computed as means of daily extremes where ecologically appropriate: surface temperature was summarized as the average of daily maximum values; bottom temperature as the average of daily mean values; spring surface salinity as the average of daily minimum values; and current speed and bed shear stress as averages of daily maximum values. Water quality variables were extracted from the SSM water quality model (WQM). High-resolution WQM solutions are not available in SSM_{HR}, so standard resolution WQM variables were shared between both habitat suitability models. Variables initially selected for nutrient variable candidates included nitrate (NO₃), ammonium (NH₄), and dissolved inorganic nitrogen (DIN), derived from NO₃ and NH₄. Variables initially selected to represent light availability or water clarity were total suspended solids (TSS), turbidity (TURB), essential diffuse attenuation coefficient (KESS), and photosynthetically active radiation at depth (PARAD). All water quality variables were summarized as averages of daily means. As both HSMs shared water quality variables of the same resolution, the models constructed in this study were tested for differences related to the use of SSM HYD versus SSM_{HR} HYD variables.

Prior to interpolation, the full set of candidate variables were assessed for collinearity and spatial variability to address initial redundancies among potential predictors. SSM hydrodynamic and water quality node summaries were joined by node ID, inspected for missing values, and a Pearson correlation matrix was computed using pairwise complete observations and visualized as a correlogram. Variable pairs with $|r| > 0.7$ were flagged as potentially collinear and evaluated for redundancy. TSS, TURB, and KESS were highly correlated ($r > 0.93$), reflecting their shared dependence on suspended particulate matter. TSS was retained over other variables as it represents a direct measurement of suspended particles relevant to both light availability and substrate quality for spore settlement. PARAD was excluded due to minimal spatial variation ($>50\%$ of nodes = 0 mol photons $\text{m}^{-2} \text{day}^{-1}$, mean = 0.0006 mol photons $\text{m}^{-2} \text{day}^{-1}$). DIN and NO_3 were perfectly correlated ($r = 1.0$), and NO_3 was selected for modeling over DIN as nitrate is the primary, biologically relevant nitrogen source for kelps.

Spatial Interpolation

Node-level seasonal summaries for SSM hydrodynamic, SSM water quality, and SSM_{HR} hydrodynamic variables were exported from R to CSV format with projected coordinates in EPSG:2856 for spatial interpolation in ArcGIS Pro. Spline with Barriers interpolation was selected for this spatial operation to account for land boundaries during surface fitting, preventing through-land interpolation.

A barrier line feature was derived from an island shoreline polygon feature from San Juan County GIS with a 200m inward buffer applied to retain nearshore nodes positioned slightly landward of the shoreline in the hydrodynamic model grid. Manual vertex editing was performed to restore land barrier continuity at locations where the buffer operation severed narrow land connections across several islands in the study area (Appendix 1). All manual edits were

restricted to reestablishing connections present in the original shoreline dataset, and none of the selected SSM or SSM_{HR} nodes fell within any repaired polygon areas. The final count of nodes used for the interpolation SSM (HYD and WQM) exports contained 509 of 522 nodes within the water domain, and SSM_{HR} (HYD) contained all 20,654 nodes. The remaining 13 SSM nodes fell outside the interpolation domain due to their position slightly inland on land features.

Output cell size was determined from nearest-neighbor distance statistics on the final node datasets. For SSM_{HR}, node spacing decreased with proximity to shore, so a cell size of 180m was selected based on average nearest-neighbor distance (mean = 181m) among nodes within 1km of the shoreline (the area containing 96.4% of mapped *N. luetkeana* bed polygons). For SSM variables, a cell size of 1,000m was selected as approximately half the average nearest-neighbor distance (~2km) across the full node range. All variables were interpolated with a smoothing factor of 0.

Raster Validation and Preprocessing

The interpolated rasters were loaded in R and validated for coordinate reference system consistency (EPSG:2856). Physical clamping was applied to constrain Spline with Barriers outputs to the range of observed node values from seasonal summaries, preventing extrapolation artifacts at domain edges. Current speed and bed shear stress were additionally clamped to a lower bound of zero as physically constrained magnitude variables; small negative values in both SSM and SSM_{HR} seasonal outputs were artifacts of interpolation overshoot. Log-transformation of bed shear stress was applied after clamping as $\log(\text{tauc} + 0.0001)$ to avoid $\log(0)$ and compress the right-skewed distribution. All rasters were resampled to 5m resolution using bilinear interpolation and masked to the 0-12m MLLW depth domain.

Substrate and Fetch Predictor Variables

Benthic substrate stability was classified using an adaptation of the Dethier habitat classification system (Dethier, 1990, 2014). This data is available within a line feature that assigns shoreline segments a stability class based on the dominant substrate type. The original Dethier classes present within the study extent were: Stable (bedrock and consolidated rock), Semi-stable (gravel and cobble), Mobile (sand, mud, and organic), and Stabilized by Vegetation. This fourth class was collapsed into the Mobile substrate category due to only three occurrences in the study extent, all with zero kelp presence, to avoid complete separation of this sparse class in the model. Stable substrate was used as the reference level in all model formulas.

Effective fetch was extracted from the Washington State ShoreZone polyline feature dataset (Washington State Department of Natural Resources, 2024). This variable measures a directionally weighted distance across open water that characterizes wind-driven wave exposure potential for each shoreline segment. Both substrate class and effective fetch were assigned to points by joining the nearest point feature to the respective shoreline segments. Median assignment distance was approximately 88m (mean 197m). Approximately 21% of points were assigned from segments more than 250m away, reflecting the coarser spatial resolution of the ShoreZone inventory relative to the 5m prediction grid in areas distant from mapped shoreline features.

Presence/Absence Data and Predictor Extraction

Depth Domain Selection

The modeling domain was restricted to 0-12m MLLW based on the extent of kelp canopy in 2024. Of 3,461,468 one-meter cells intersecting 2024 Samish kelp canopy polygons, 99% fell within 0-12m MLLW. Of the 1% of cells extending beyond 12m, 75% fell between 12-14m and

25% beyond 14m, consistent with minor edge effects as a result of misalignment between digitized canopy boundaries and modeled bathymetry. Random sampling of presence points from within kelp polygons also yielded zero points beyond 12m, providing additional confirmation that ecologically meaningful canopy area is concentrated primarily at depths shallower than 12m. This domain restriction was motivated in part by preliminary modeling results indicating that depth dominated predictions when tested across a broader depth extent of 0-30m, obscuring the contribution of other environmental predictors. Restricting the domain to the ecologically supported 0-12m range ensures that depth operates as a contextually appropriate covariate rather than an artifact of the prediction domain.

Presence Point Generation

Kelp presence points were generated by random sampling within digitized 2024 kelp canopy polygons (Samish Indian Nation GIS, 2024). The imagery captured in aerial surveys and used for canopy digitization was collected in July and August of 2024. A target sample size of 300 presence points was drawn and later filtered to retain only complete training points with values for all predictor variables, yielding 293 valid presence points after removal of points with missing predictor values at the intertidal boundary.

Absence Point Generation

Absence points were generated using a depth-stratified sampling design to address the higher concentration of training presence points in shallow water. Depth bins were defined at three, equal-width 4m intervals: 0-4m, 4-8m, and 8-12m. Absence proportions were derived from the presence point depth distribution: 51.5% of presence points fell in the 0-4m bin, 38.6% in the 4-8m bin, and 9.9% in the 8-12m bin. Absence proportions were set to 50%, 40%, and 10%

respectively to reflect this distribution. A 2:1 absence to presence ratio was used, yielding a target of 600 absence points.

Absence points were sampled from within the intersection of each depth bin's bathymetric mask, the 2024 WADNR aerial survey footprint buffered by 1km, and a 25m exclusion buffer around all kelp canopy polygons. The 25m exclusion buffer was used to reflect the localized spore dispersal range of *N. luetkeana* and prevents absences from being placed directly next to or nearby locations likely to support kelp (Burnett et al., 2024). The survey footprint constraint ensures absence locations are spatially consistent with the actual coverage of aerial survey flight paths (GeoTerra, Inc., 2024).

Following removal of four points at the intertidal boundary with missing oceanographic predictor values, the final modeling datasets comprised 888 points (293 presence, 595 absence) for both HSM_{STD} and HSM_{HR} models.

Predictor Extraction

Depth, slope, and rugosity were extracted from the 1m DEM and its derivatives at each training point using bilinear interpolation. Substrate stability class and effective fetch were assigned to training points by nearest-feature join from the respective shoreline segment datasets. Oceanographic predictor values from SSM and SSM_{HR} were extracted from the validated 5m interpolated raster surfaces.

Predictor Selection and Multicollinearity Assessment

Prior to model fitting, final predictors were assessed again for collinearity using Pearson correlation coefficients computed from pairwise complete observations (Appendix 2). Variance inflation factors were assessed using generalized VIF (Fox & Monette, 1992), computed from a binomial GLM approximating the sdmTMB fixed-effect structure, to appropriately handle the

multi-level substrate factor. Rugosity was removed entirely due to near-perfect collinearity with slope ($r = 0.99$). Salinity was retained despite moderate collinearity with nitrate concentration in the HSM_{STD} predictor set ($r = 0.80$) and sea surface temperature ($r = -0.57$) because of its known influence as a variable affecting kelp growth (Farrugia Drakard et al., 2025). Surface and bottom current velocity were more strongly correlated in the HSM_{HR} predictor set ($r = 0.85$) than in HSM_{STD} ($r = 0.62$), but both were retained as they represent distinct seasonal and depth-stratified aspects of the hydrodynamic regime, with post-selection $GVIF^{(1/(2 \cdot Df))}$ below 2.4 for both variables in both models. All other predictors in the final model had $GVIF^{(1/(2 \cdot Df))}$ values below 2.0, indicating acceptable collinearity levels for model fitting.

Model fit was compared across two predictor configurations and evaluated using spatial random field standard deviation (σ_O) and marginal likelihood criterion (ML). In this initial comparison, bottom current velocity and bed shear stress (TAUC) were used as candidate variables in different predictor configurations to capture hydrodynamic energy influences on the gametophyte life stage of *Nereocystis*. Log-transformed bed shear stress was tested in the initial set of predictors but was ultimately replaced by winter mean bottom current velocity, which provides a more mechanistically direct measure of forces that may dislodge the holdfast from the substrate. The predictor configuration in which bed shear stress was replaced by bottom current (Table 2) was selected as the final model based on a decrease in the spatial random field standard deviation and lower marginal likelihood criterion, indicating improved fixed-effect explanation of spatial structure and overall model fit (Appendix 3).

Table 2. Final predictor variables included in both HSM_{STD} and HSM_{HR} models, with seasonal windows and summary metrics defined by ecological relevance to *N. luetkeana* life history.

Variable	Daily Metric	Seasonal Average	Months
Surface Temperature	Maximum	Summer	April 2024 - September 2024
Surface Current Speed	Maximum	Summer	April 2024 - September 2024
Surface Nitrate concentration	Mean	Summer	April 2024 - September 2024
Bottom Current Speed	Maximum	Winter	October 2023 - February 2024
Bottom Temperature	Mean	Winter	October 2023 - February 2024
Surface Salinity	Minimum	Spring	March 2024 - June 2024
Surface TSS	Mean	Spring	March 2024 - June 2024

Model Fitting, Variable Response, and Suitability Predictions

Two binomial sdmTMB models were fitted to predict *N. luetkeana* habitat suitability across the nearshore waters of San Juan County, Washington (0-12m MLLW). The first model (HSM_{STD}) used hydrodynamic predictors interpolated from the Salish Sea Model at approximately 1km resolution; the second model (HSM_{HR}) used hydrodynamic predictors from the high-resolution Salish Sea Model at approximately 180m resolution. Both models used identical fixed-effect structures for direct comparability and were fitted using the sdmTMB R package (version 1.0.0) in R (Anderson et al., 2025).

Both models incorporated a Matérn spatial random field with a barrier mesh to restrict spatial correlation across land boundaries. The barrier mesh was constructed from a land polygon derived from bathymetry and simplified to 500m tolerance, with a minimum triangle edge length of 0.5km and a range fraction of 0.1 across land barriers, setting spatial correlation across land to 10% of open-water range.

The final model formula included depth, slope, effective fetch, substrate stability class, sea surface temperature (summer mean of daily maximum), bottom temperature (winter mean of

daily mean), spring surface salinity (mean of daily minimum), summer surface current (mean of daily maximum), winter bottom current (mean of daily maximum), summer nitrate concentration (mean of daily mean), and spring total suspended solids (mean of daily mean). All continuous predictors were standardized to mean zero and unit standard deviation using parameters derived from the HSM_{STD} training dataset and applied consistently to both models. A standard intercept parameterization was used with Stable substrate as the reference level.

Partial dependence response curves were generated for all clearly defined predictors whose confidence intervals did not cross zero in the HSM_{HR} final model, with HSM_{STD} shown for comparison. Curves were computed by varying each predictor across its observed range while holding all other continuous predictors at their training dataset means and fixing substrate at the Stable reference level. The spatial random field was excluded from predictions to isolate fixed-effect relationships. Confidence intervals were estimated using simulation-based standard errors (n=300 draws) converted to the probability scale.

Habitat suitability predictions were generated across a 30m prediction grid covering the 0-12m MLLW depth domain and resampled to 5m resolution for final mapping. Suitable habitat was defined as cells where predicted suitability exceeded a probability threshold of 0.5. Total suitable habitat area was calculated as the sum of 5m grid cells exceeding the 0.5 probability threshold, converted to hectares. Prediction uncertainty was estimated by generating 500 simulations from the model and converting the resulting standard errors to the probability scale.

Model Validation

Spatial Block Cross-Validation

Spatial block cross-validation was conducted using island-based folds to assess model transferability across spatially independent regions of the study area. Three folds were defined

corresponding to San Juan Island, Lopez Island, and Orcas Island, the three islands with sufficient training point density for meaningful fold-level area under the curve (AUC) computation. Smaller islands throughout the study extent lacked the minimum point density required for reliable AUC estimation and were excluded from cross-validation folds. Held-out fold sizes ranged from 124 to 198 points, with presence points ranging from 29 to 80 and absence points ranging from 92 to 143 per fold. Island separation of 2-10km exceeds the estimated Matérn spatial autocorrelation range (~2km) of the final models, consistent with the spatial independence assumption underlying blocked cross-validation (Roberts et al. 2017).

In each fold, the held-out island was withheld entirely from model fitting. The barrier mesh was rebuilt from the training fold points and scale parameters were recomputed from the training fold data for each iteration. Both HSM_{STD} and HSM_{HR} models were refit per fold. AUC was computed for each held-out fold and averaged across folds to produce mean cross-validation AUC.

Field Observations for Independent Validation (2025 Snorkel Surveys)

Independent presence/absence observations were collected during snorkel surveys conducted across San Juan Island in July and August of 2025 to evaluate model predictions against data temporally and methodologically independent of the 2024 training dataset. Surveys were conducted at 19 sites representing a range of oceanographic and abiotic conditions, selected to be as evenly distributed as possible across publicly accessible areas and sites accessed with private permission. Sites included areas with known or historical kelp occurrence as well as rocky areas without recent kelp records.

At each site, a minimum of 12 quadrats (0.5 m²) were sampled at intervals of at least 1m to ensure spatial independence. Quadrats were tossed haphazardly within the survey area.

Presence or absence of *Nereocystis luetkeana* was recorded for each quadrat, along with bulb and stipe counts where kelp was present. Geographic coordinates were recorded using a Garmin eTrex 10 GPS receiver with WAAS enabled (estimated horizontal accuracy 3-5m).

Survey waypoints were reviewed prior to analysis and filtered to retain only observations with complete and accurate coordinates, minimum point spacing of 1m, and a successfully recorded presence/absence determination. Of 369 waypoints recorded across all sites, 353 passed these filters, of which 345 fell within the 0-12m MLLW modeling domain with complete oceanographic predictor coverage. Substrate and effective fetch at each validation point were assigned via nearest-feature join, while all other predictors were extracted from continuous rasters at each validation point, consistent with the treatment of training data. Predictor values at validation locations were extracted from the 2024 interpolated raster surfaces used for model fitting, as 2025 covariate data were unavailable. The independent validation therefore primarily assesses spatial transferability of model predictions to a geographically restricted subset of the study area rather than strict temporal prediction.

Independent Validation

Discrimination performance of the models against the 2025 snorkel survey data was assessed using the area under the receiver operating characteristic curve (AUC), computed using the Wilcoxon rank-sum statistic. AUC was calculated for all depths combined and stratified by depth (shallow $\leq 5\text{m}$, deep $> 5\text{m}$) to characterize model performance across the depth distribution of the validation dataset. Threshold optimization via maximum True Skill Statistic (TSS) was conducted on in-sample training data. Given that TSS threshold optimization was conducted on in-sample training predictions including the spatial random field, the resulting high TSS values (HSM_{STD} 0.985, HSM_{HR} 0.980) reflect fitted rather than predictive performance and are not

appropriate as a basis for habitat area estimation. A fixed probability threshold of 0.5 was therefore used for all habitat area estimates rather than the optimized thresholds. Receiver operating characteristic (ROC) curves were generated for both models by computing sensitivity and specificity across 200 evenly spaced probability thresholds and are presented alongside AUC values to characterize model discrimination across the full range of suitability classification thresholds.

Results

Model Fit and Fixed Effects

The model intercepts were -9.78 (95% Confidence Interval (CI): -16.55 to -3.01) for HSM_{STD} and -7.14 (95% CI: -13.23 to -1.05) for HSM_{HR} , corresponding to predicted occurrence probabilities below 0.1% when all predictor values are held at their mean on stable substrate, with the spatial random field excluded. These strongly negative intercepts reflect the low base rate of kelp occurrence on the logarithmic-odds scale. The spatial random field provides substantial additional spatial structure beyond the fixed effects, helping to identify locations where environmental conditions yield higher probabilities of kelp occurrence.

Regression coefficients for all predictors with 95% confidence intervals are summarized in Table 3 and shown in Figure 1. Variables with confidence intervals that did not span zero were considered to have a clearly defined association with *N. luetkeana*. Slope (HSM_{STD} : $\beta = 2.82$, CI: 1.27 to 4.37; HSM_{HR} : $\beta = 1.76$, CI: 0.33 to 3.19) and effective fetch (HSM_{STD} : $\beta = 6.88$, CI: 2.82-10.95; HSM_{HR} : $\beta = 4.19$, CI: 1.26 to 7.12) were positive predictors in both models. Mobile substrate (sand, mud, and organic material) was a negative predictor in both models (HSM_{STD} : $\beta = -4.72$, CI: -8.18 to -1.25; HSM_{HR} : $\beta = -4.18$, CI: -7.26 to -1.10), while semi-stable substrate (gravel and cobble) had CIs spanning zero in both models. SST, bottom temperature, salinity,

and TSS had confidence intervals spanning zero in both models, indicating no clearly defined association with *N. luetkeana* at this scale.

Three predictors had CIs excluding zero in HSM_{HR} but not HSM_{STD}: surface current ($\beta = 1.75$, CI: 0.12 to 3.37), bottom current ($\beta = 1.65$, CI: 0.10 to 3.21), and nitrate concentration ($\beta = 3.60$, CI: 0.50 to 6.69). Depth had a confidence interval excluding zero in HSM_{HR} ($\beta = -1.11$, CI: -1.94 to -0.28) but not in HSM_{STD} ($\beta = -0.82$, CI: -1.68 to 0.04), where the confidence interval marginally included zero.

The spatial random field estimated for HSM_{HR} had a marginal standard deviation (σ_O) of 9.10 and a Matérn range of 2.82 km, compared to σ_O of 13.97 and a range of 2.40 km for HSM_{STD} (Table 4). The lower σ_O in HSM_{HR} indicates that the fixed effects in that model explained a greater proportion of spatial variation, consistent with the expectation that finer-resolution predictors capture ecologically relevant spatial structure that 1 km resolution cannot resolve. The marginal log-likelihood criterion was 177.97 for HSM_{HR} and 198.97 for HSM_{STD}. Lower ML values indicate better fit, and the difference of approximately 21 units is consistent with improved fit in HSM_{HR}. The longer Matérn range in HSM_{HR} (2.82 vs 2.40 km) is consistent with this interpretation: when fine-scale spatial heterogeneity is absorbed by the fixed effects, the residual random field operates at a broader spatial scale rather than compensating for unresolved predictor variation at fine scales.

Table 3. Fixed-effect regression coefficients for all predictors in the HSM_{STD} and HSM_{HR} final models. Estimates are on the log-odds scale; all continuous predictors were standardized to a mean of zero and a unit standard deviation prior to model fitting. Confidence intervals were derived from Hessian-based standard errors. Predictors whose 95% confidence intervals exclude zero are considered to have a clearly defined association with *N.luetkeana*. Substrate contrasts are relative to the stable substrate reference level (bedrock and consolidated rock). The intercept represents the log-odds of *N.luetkeana* presence at mean predictor values on stable substrate with the spatial random field excluded.

Model	Term	Estimate	std.error	conf.low	conf.high
HSM _{STD}	Intercept (Stable substrate baseline)	-9.7795	3.4538	-16.5488	-3.0102
HSM _{STD}	Depth	-0.8195	0.4392	-1.6803	0.0412
HSM _{STD}	Slope	2.8194	0.7897	1.2716	4.3672
HSM _{STD}	Fetch	6.8813	2.0744	2.8155	10.9471
HSM _{STD}	Substrate: Semi-stable (gravel/cobble)	-3.4351	1.9124	-7.1833	0.3131
HSM _{STD}	Substrate: Mobile (sand/mud/organic)	-4.7151	1.7672	-8.1788	-1.2514
HSM _{STD}	SST summer (mean of daily max)	-6.1215	5.6651	-17.2248	4.9819
HSM _{STD}	Bottom temp winter (mean of daily mean)	7.2373	6.297	-5.1046	19.5793
HSM _{STD}	Salinity spring (mean of daily min)	-6.8773	4.3721	-15.4465	1.6919
HSM _{STD}	Surface current summer (mean of daily max)	-0.5905	2.1486	-4.8017	3.6207
HSM _{STD}	Bottom current winter (mean of daily max)	3.1318	2.8241	-2.4034	8.667
HSM _{STD}	Nitrate concentration summer (mean of daily mean)	3.962	3.6892	-3.2687	11.1926
HSM _{STD}	TSS spring (mean of daily mean)	-0.6525	4.097	-8.6825	7.3775
HSM _{HR}	Intercept (Stable substrate baseline)	-7.1432	3.1066	-13.2321	-1.0544
HSM _{HR}	Depth	-1.1121	0.4222	-1.9397	-0.2845
HSM _{HR}	Slope	1.758	0.7301	0.3269	3.189
HSM _{HR}	Fetch	4.1906	1.4962	1.2581	7.1231
HSM _{HR}	Substrate: Semi-stable (gravel/cobble)	-1.9087	1.511	-4.8701	1.0528
HSM _{HR}	Substrate: Mobile (sand/mud/organic)	-4.1791	1.5696	-7.2554	-1.1027

HSM _{HR}	SST summer (mean of daily max)	-0.5058	1.2946	-3.0432	2.0317
HSM _{HR}	Bottom temp winter (mean of daily mean)	0.5982	1.0764	-1.5115	2.708
HSM _{HR}	Salinity spring (mean of daily min)	-0.454	1.5467	-3.4854	2.5774
HSM _{HR}	Surface current summer (mean of daily max)	1.7464	0.8281	0.1234	3.3694
HSM _{HR}	Bottom current winter (mean of daily max)	1.6529	0.7935	0.0977	3.208
HSM _{HR}	Nitrate concentration summer (mean of daily mean)	3.598	1.5787	0.5039	6.6921
HSM _{HR}	TSS spring (mean of daily mean)	1.3174	1.2414	-1.1157	3.7506

Table 4. Summary of spatial random field parameters, model fit criteria, and suitable habitat area estimates for HSM_{STD} and HSM_{HR} final models. Spatial standard deviation (σ_O) and Matérn range were estimated from the model posterior. Lower marginal log-likelihood (ML) criterion values indicate better model fit. Suitable habitat area was calculated as the sum of 5m grid cells with predicted suitability exceeding 0.5, converted to hectares. TSS-optimized thresholds were derived from in-sample training predictions including the spatial random field and are reported for reference only; the fixed 0.5 threshold was used for all habitat area estimates.

	HSM _{STD}	HSM _{HR}
σ_O (spatial SD)	13.97	9.10
Matérn range (km)	2.40	2.82
ML criterion	198.97	177.97
Suitable habitat at 0.5 threshold (ha)	2,241	1,705
TSS-optimized threshold	0.44	0.40
In-sample TSS	0.985	0.980

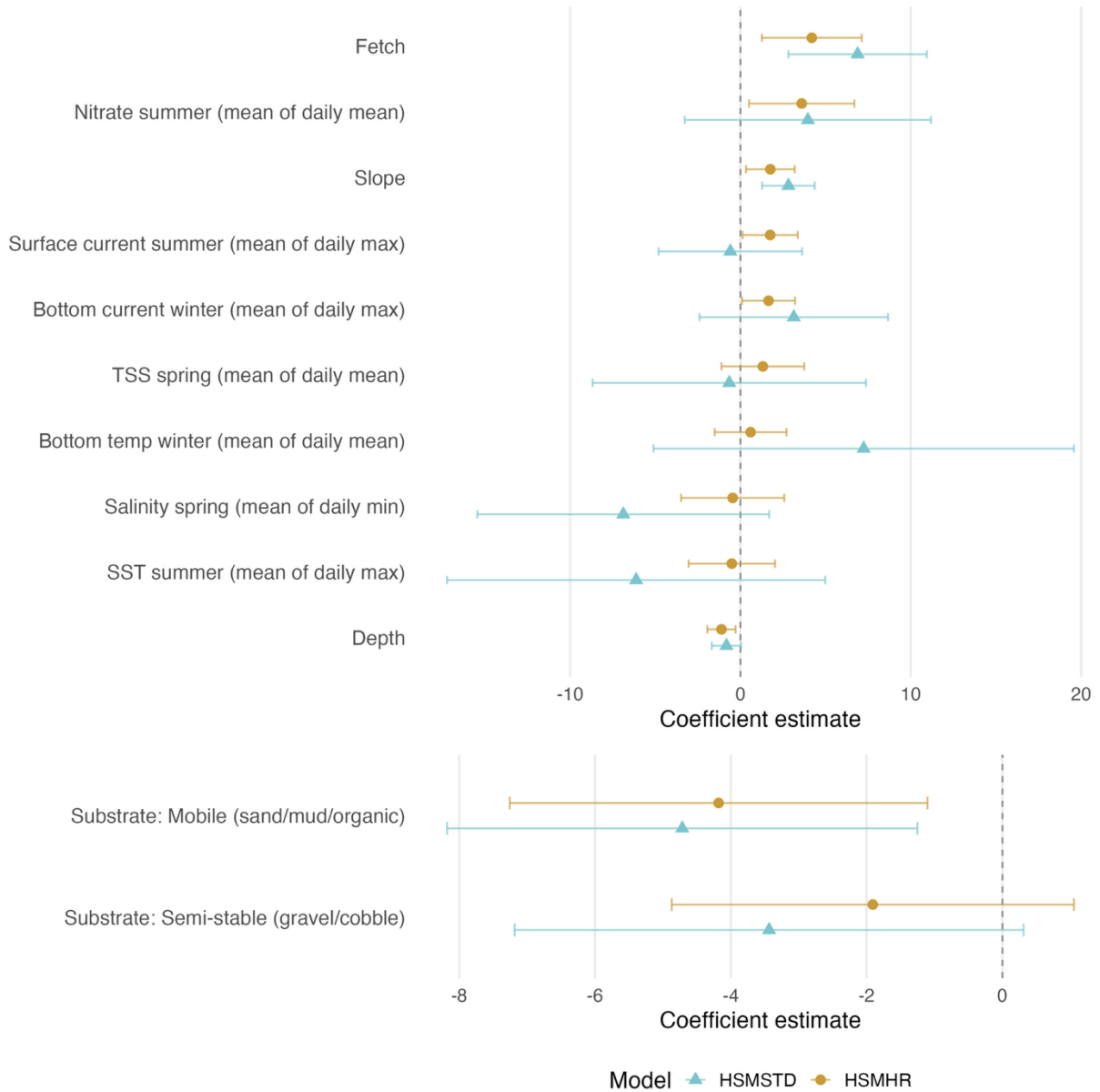


Figure 1. Regression coefficients and 95% confidence intervals for all predictors in the HSM_{STD} and HSM_{HR} final models, excluding the intercept (stable substrate). Panel A shows continuous predictors ordered by HSM_{HR} coefficient magnitude. Panel B shows substrate contrasts relative to the stable substrate reference level (bedrock and consolidated rock). Points and error bars represent coefficient estimates and 95% confidence intervals. Estimates whose intervals exclude zero indicate a clearly defined association with *N.luetkeana* presence. Predictors are shown for both models to facilitate direct comparison of models constructed with different hydrodynamic oceanographic predictor resolutions.

Predictor Variable Response Curves

Physical habitat predictors

Slope and effective fetch showed the clearest and most consistent relationships with kelp suitability across both models (Figure 2). Both predictors produced S-shaped response curves, with suitability remaining near zero at low values before increasing steeply through a transitional range and approaching 1.0 at high values. For slope, the 0.5 suitability threshold occurred at approximately 28° in HSM_{STD} and 32° in HSM_{HR}. For fetch, thresholds occurred at approximately 21 km and 24 km respectively. In both cases HSM_{HR} required slightly higher predictor values to reach equivalent suitability, indicating a more conservative response at intermediate values. Curve shapes were otherwise consistent between models, and both predictors reached maximum predicted suitability within the observed training range.

Depth showed no interpretable marginal relationship with kelp suitability in either model across the 0-12m domain. Predicted suitability remained near zero throughout (maximum 0.004 in HSM_{HR}, 0.0002 in HSM_{STD}), and confidence intervals were wide across the full range in both models. Although depth had a confidence interval excluding zero in HSM_{HR} ($\beta = -1.11$, CI: -1.94 to -0.28), the response curve provides no evidence of a meaningful fixed-effect relationship within this depth range.

Hydrodynamic predictors

Surface current and bottom current only have clear positive associations in HSM_{HR}, still both models are shown for comparison. Surface current showed a weak positive relationship in HSM_{HR}, with predicted suitability reaching only 0.18 at the maximum observed value (~1.46 m/s) and remaining near zero across most of the observed range. The HSM_{STD} curve for surface

current was essentially flat throughout. Given the limited range of the response, the contribution of surface current to predicted suitability is understood to be modest.

Compared with other predictor variables considered in this analysis, bottom current showed the largest divergence between high and low model resolutions. The HSM_{STD} curve reached 0.90 predicted suitability at the upper end of the observed range (~ 0.64 m/s), crossing the 0.5 threshold at approximately 0.56 m/s. The HSM_{HR} bottom current curve, despite this variable having a positive association with *N.luetkeana*, reached a maximum predicted suitability of only 0.31 and did not cross the 0.5 threshold within the observed range. Both curves increased steadily but never flattened, indicating that the relationship extends beyond the range of training data.

Water Quality Predictor

Nitrate concentration had a consistently positive relationship with *N.luetkeana*, with coefficient confidence intervals excluding zero only in HSM_{HR} . The HSM_{HR} response curve remained near zero across most of the observed range (0.025-0.105 gN/m³), rising sharply only above approximately 0.085 gN/m³ and crossing the 0.5 suitability threshold at ~ 0.100 gN/m³, within the uppermost 5% of observed values. The HSM_{STD} curve did not exceed a predicted suitability value of 0.28 across the full range. The threshold behavior at the data boundary means this relationship is poorly characterized: the curve reflects a steep rise at nitrate concentration values where training data are sparse, meaning that the shape of the relationship at high nitrate concentration is not captured by the available data. Whether this reflects a genuine nitrate concentration threshold for suitability or is related to spatial correlation with co-occurring environmental conditions (particularly salinity, which is moderately correlated with nitrate concentration in both predictor sets) is not resolved here and warrants further investigation.

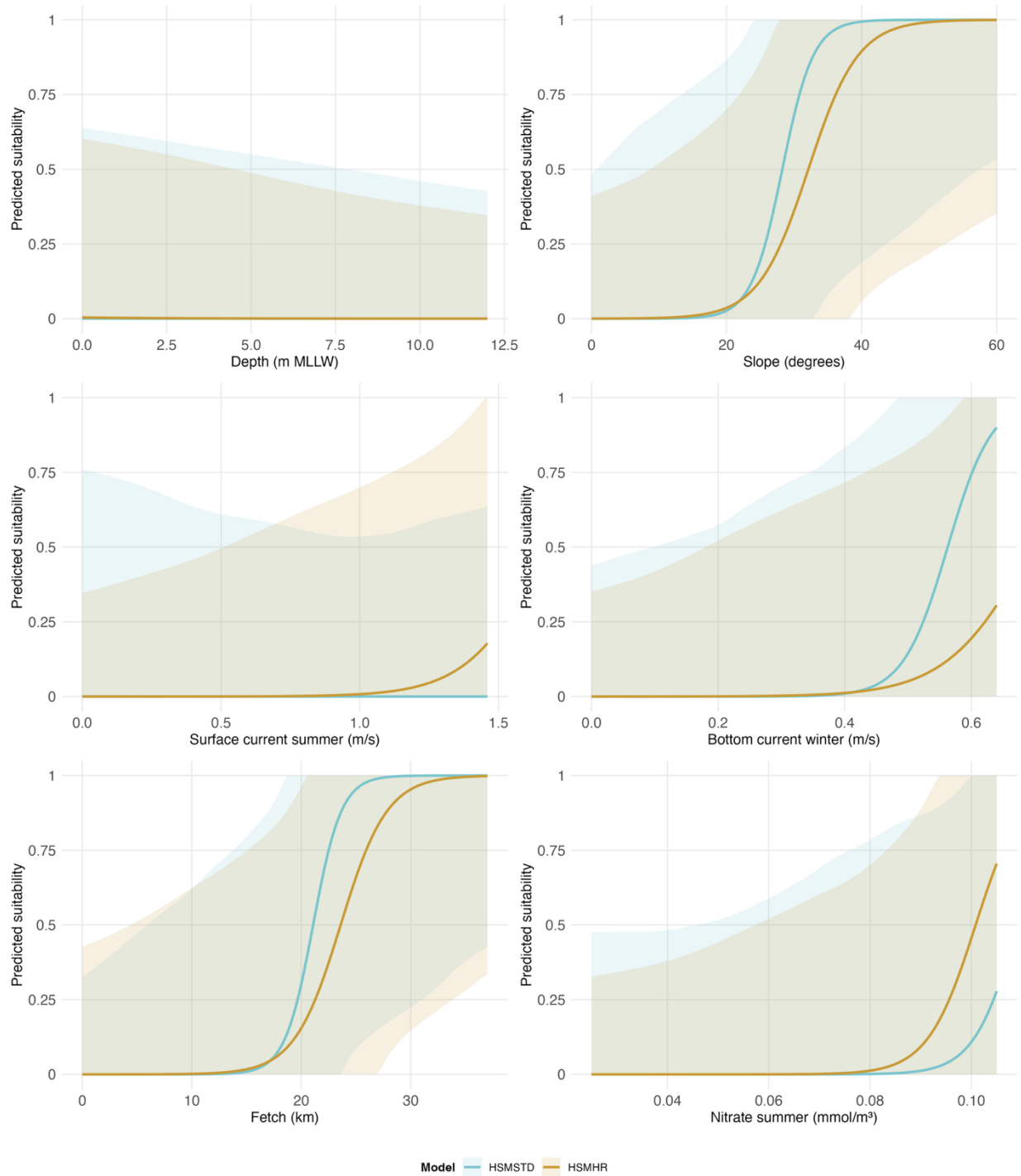


Figure 2. Partial dependence response curves for predictors with confidence intervals excluding zero in at least one HSM. Curves show predicted kelp suitability as a function of each predictor while holding all other continuous predictors at training dataset means, with substrate fixed at the stable reference level and the spatial random field excluded. Shaded bands represent 95% confidence intervals (n=300 draws).

Predicted Habitat Distribution and Area Estimates

Habitat suitability predictions in both models are concentrated and generally continuous along exposed rocky shorelines throughout the study area, with highest suitability values consistently associated with the shorelines facing outer channels and straits around San Juan, Orcas, and Lopez Islands and smaller exposed islands throughout the archipelago (Figure 3). Interior channels and bays with low fetch and finer substrate showed low predicted suitability in both models. Predicted suitability broadly overlaps with the Samish Indian Nation's 2024 kelp canopy polygons, although this alignment reflects the training points derived from the polygons and the environmental predictors shared during training and prediction.

At a probability threshold of 0.5, HSM_{STD} predicted 2,241 ha of suitable habitat and HSM_{HR} predicted 1,705 ha, approximately 24% less than HSM_{STD} (Figure 4). The lower HSM_{HR} area estimate is consistent with the more conservative fixed-effect responses, possibly related to observed patterns in the slope and fetch response curves, where HSM_{HR} required higher predictor values to reach equivalent suitability. The 0.5 threshold was applied as an ecologically interpretable cutoff representing equal probability of *N. luetkeana* presence and absence. TSS-optimized thresholds derived from in-sample training predictions (HSM_{STD} 0.44, HSM_{HR} 0.40) are reported in Table 4 for reference but were not used as the primary basis for area estimates given their in-sample derivation.

Prediction uncertainty was generally higher in HSM_{STD} than HSM_{HR} across the study area, consistent with HSM_{STD} 's larger spatial random field standard deviation ($\sigma_O = 13.97$ vs 9.10), indicating that HSM_{STD} relies more heavily on the random field to explain spatial variation not captured by fixed effects (Figure 5).

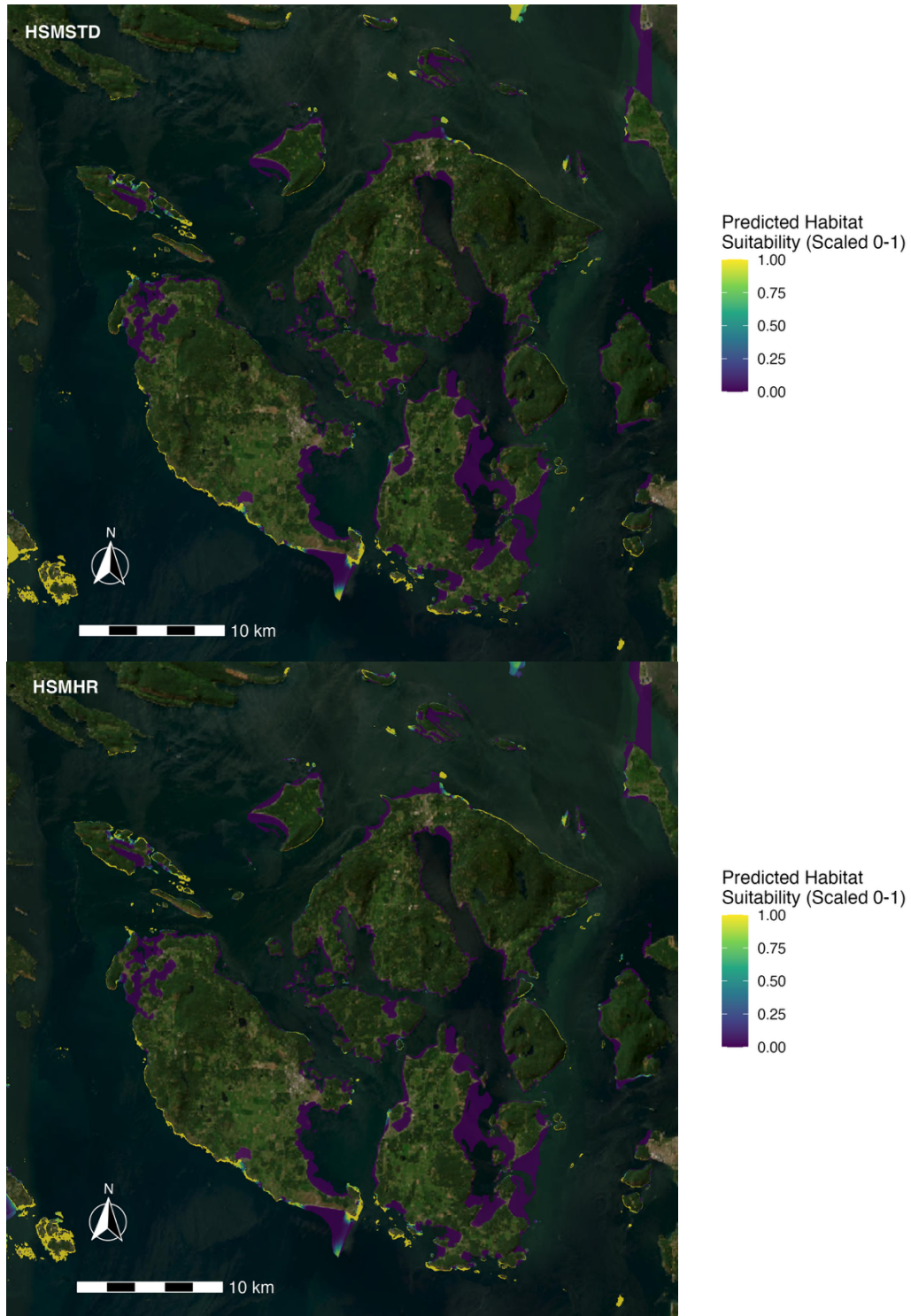


Figure 3. Predicted habitat suitability for *N.luetkeana* across the San Juan Islands, Washington for HSM_{STD} (~1 km resolution, top) and HSM_{HR} (~180 m resolution, bottom). Suitability values represent the probability of *N.luetkeana* occurrence predicted by each model across the 0-12m MLLW depth domain. Predictions are based on 2023-2024 oceanographic conditions.

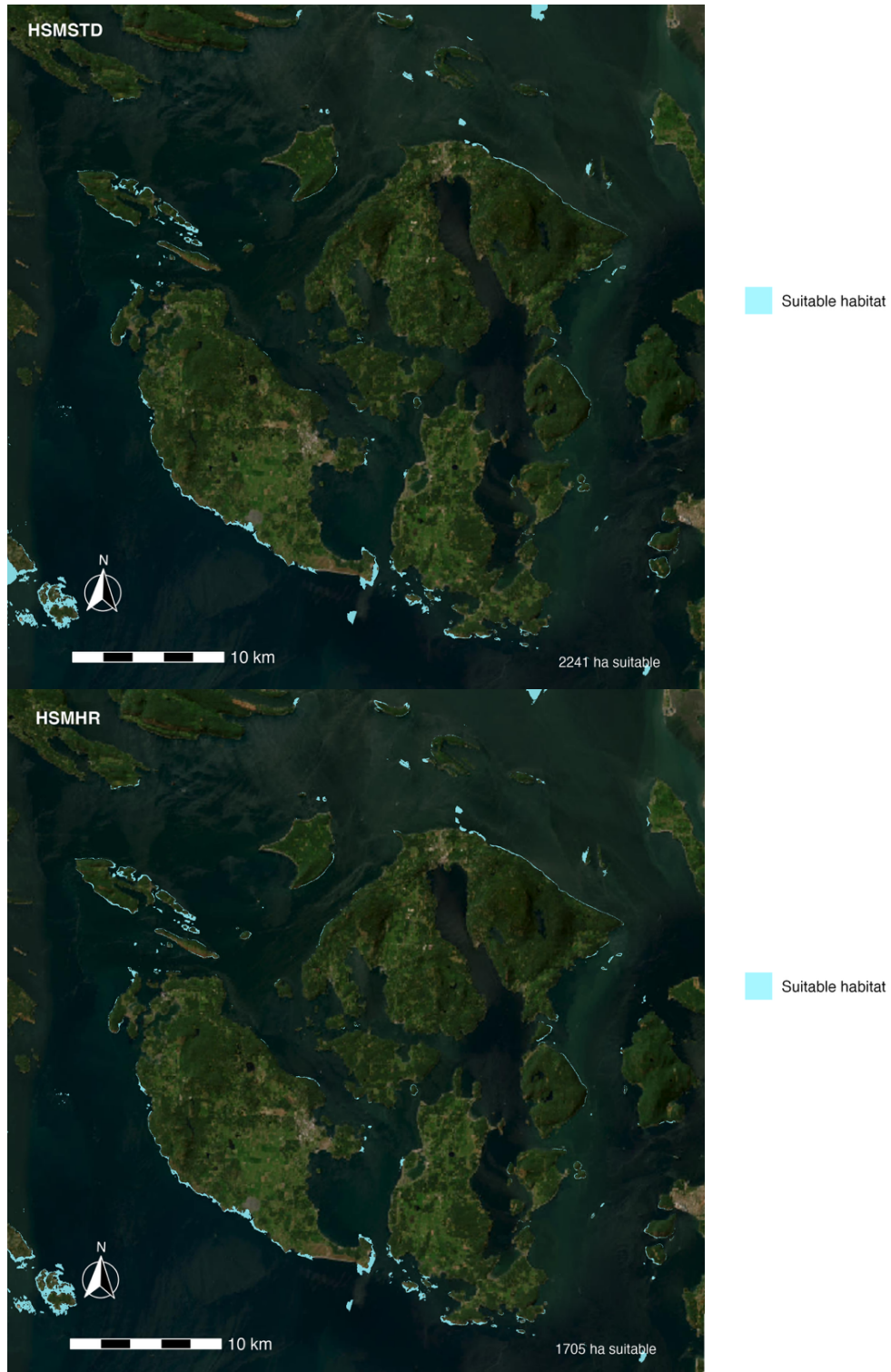


Figure 4. Binary suitable habitat area available for *N. luetkeana* in the San Juan Islands, Washington for HSM_{STD} (top, 2,241 ha) and HSM_{HR} (bottom, 1,705 ha), defined as grid cells with predicted suitability exceeding 0.5. Suitable habitat area was calculated from 5m resolution prediction grids across the 0-12m MLLW depth domain.

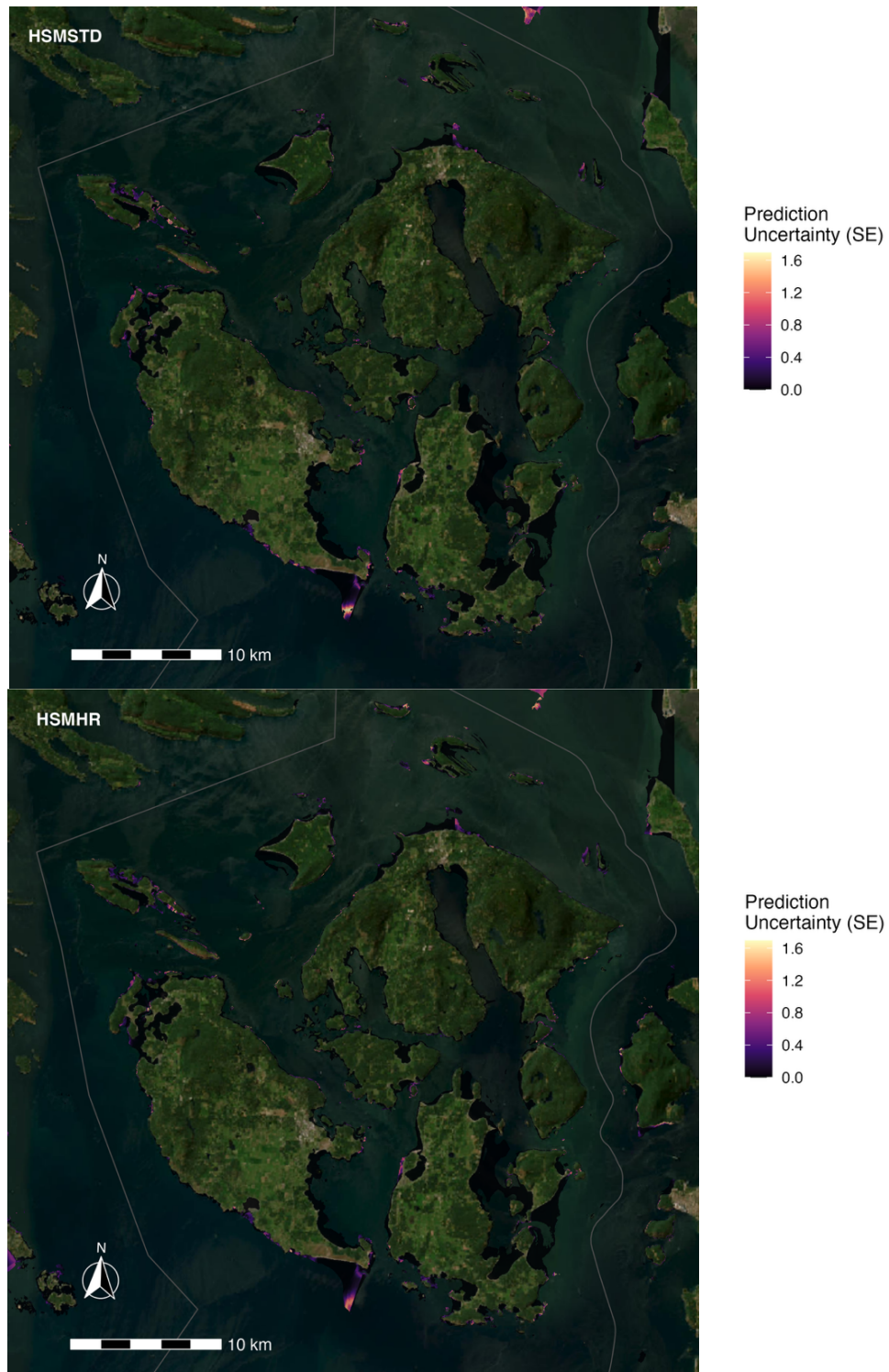


Figure 5. Prediction uncertainty for HSM_{STD} (top) and HSM_{HR} (bottom) habitat suitability models, expressed as standard error on the probability scale estimated from 500 simulation draws. Darker colors indicate greater prediction confidence; lighter colors indicate greater prediction uncertainty.

Model Validation

Spatial Block Cross-Validation

Mean cross-validation AUC was 0.895 for HSM_{STD} and 0.918 for HSM_{HR}, indicating good discriminative performance for both models across spatially independent folds (Figure 6). HSM_{HR} outperformed HSM_{STD} on San Juan Island (0.939 vs 0.801) and Lopez Island (0.990 vs 0.971), while HSM_{STD} outperformed HSM_{HR} on Orcas Island (0.912 vs 0.826).

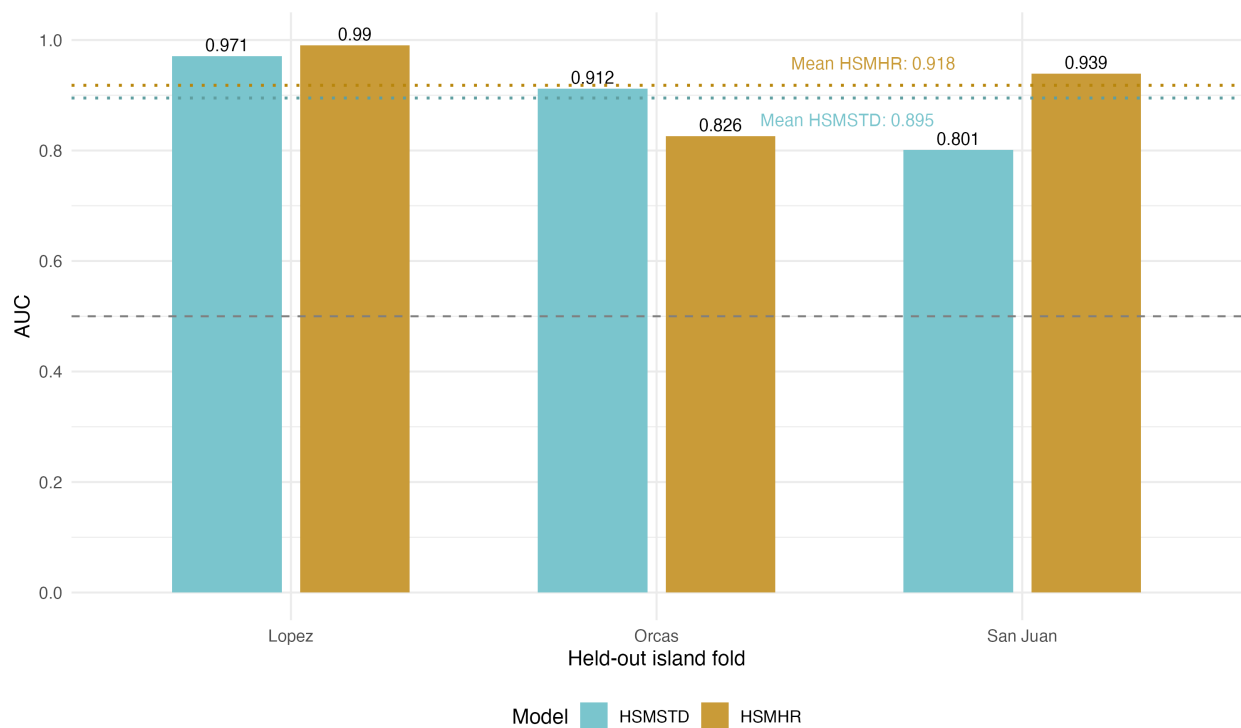


Figure 6. Spatial block cross-validation AUC for HSM_{STD} and HSM_{HR} final models across three island-based folds (San Juan, Lopez, and Orcas Islands). Dotted lines indicate mean cross-validation AUC for each model. Dashed line indicates no-skill discrimination (AUC=0.5). Barrier mesh and scale parameters were recomputed per fold.

Independent Validation

Overall AUC against the 2025 snorkel survey data was 0.722 for HSM_{STD} and 0.698 for HSM_{HR} (Figure 7), notably lower than cross-validation AUC for both models (Table 5). When

stratified by depth, validation showed contrasting patterns between models. At shallow sites ($\leq 5\text{m}$, $n=302$), HSM_{STD} achieved an AUC of 0.756 and HSM_{HR} achieved 0.718. At deeper sites ($>5\text{m}$, $n=43$), the AUC for HSM_{STD} was 0.485 (essentially no-skill discrimination) while HSM_{HR} achieved 0.559 (Figure 7). The shallow zone dominates the overall independent validation AUC for both models given that 302 of 345 validation points fall at depths of 5m MLLW or less, which partially explains this overall divergence from cross-validation performance, as the models were evaluated primarily on shallow environments (Figure 8).

ROC curves for both models against the independent validation dataset are shown in Figure 9. HSM_{STD} showed consistently higher sensitivity than HSM_{HR} across most of the low false positive rate range ($\text{FPR} < 0.40$), indicating better discrimination of true presences while limiting false positive predictions. The two curves converged above 0.40, with neither model showing a consistent advantage at higher false positive rates. The overall AUC difference (HSM_{STD} 0.722 vs HSM_{HR} 0.698) therefore primarily reflects the advantage of HSM_{STD} at low false positive rates rather than an advantage across all suitability classification thresholds.

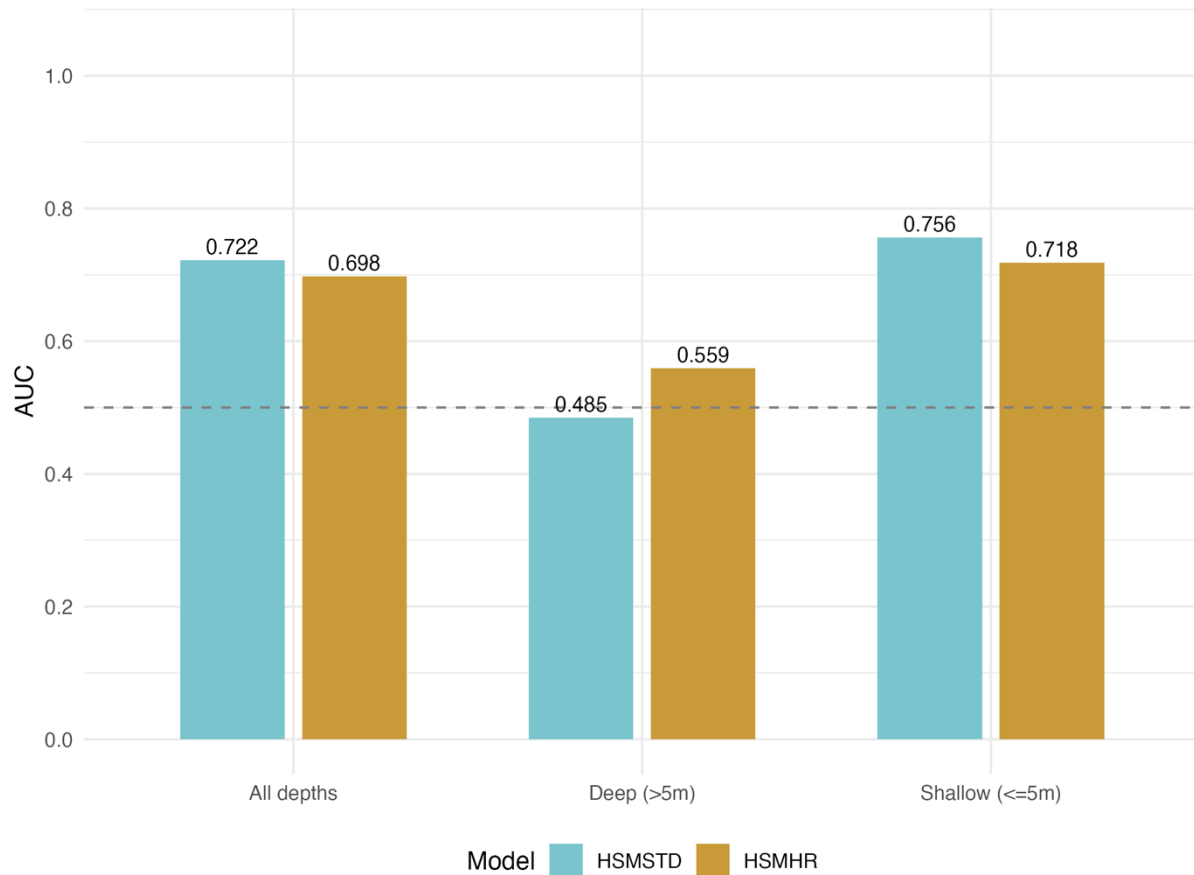


Figure 7. Independent validation discrimination performance for HSM_{STD} and HSM_{HR} models against 2025 snorkel survey data (n=345). Left: overall AUC for both models at all depths combined. Center: depth-stratified AUC for shallow zone ($\leq 5\text{m}$, n=302). Right: depth-stratified AUC for deep zone ($> 5\text{m}$, n=43). Dashed line indicates no-skill discrimination (AUC = 0.5).

Table 5. Summary of model discrimination performance for HSM_{STD} and HSM_{HR} final models. Cross-validation (CV) AUC values were computed using three-fold island-based spatial block cross-validation (San Juan, Lopez, and Orcas Islands), with barrier mesh and scale parameters recomputed per fold. Independent validation AUC values were computed against 2025 snorkel survey data (n=345) withheld entirely from model fitting, stratified by depth to characterize performance across the validation depth distribution.

Metric	Value
CV mean AUC (HSM _{STD})	0.895
CV mean AUC (HSM _{HR})	0.918
CV AUC San Juan (HSM _{STD})	0.801
CV AUC San Juan (HSM _{HR})	0.939

CV AUC Lopez (HSM _{STD})	0.971
CV AUC Lopez (HSM _{HR})	0.99
CV AUC Orcas (HSM _{STD})	0.912
CV AUC Orcas (HSM _{HR})	0.826
Independent AUC all depths (HSM _{STD})	0.722
Independent AUC all depths (HSM _{HR})	0.698
Independent AUC shallow ≤5m (HSM _{STD})	0.756
Independent AUC shallow ≤5m (HSM _{HR})	0.718
Independent AUC deep >5m (HSM _{STD})	0.485
Independent AUC deep >5m (HSM _{HR})	0.559

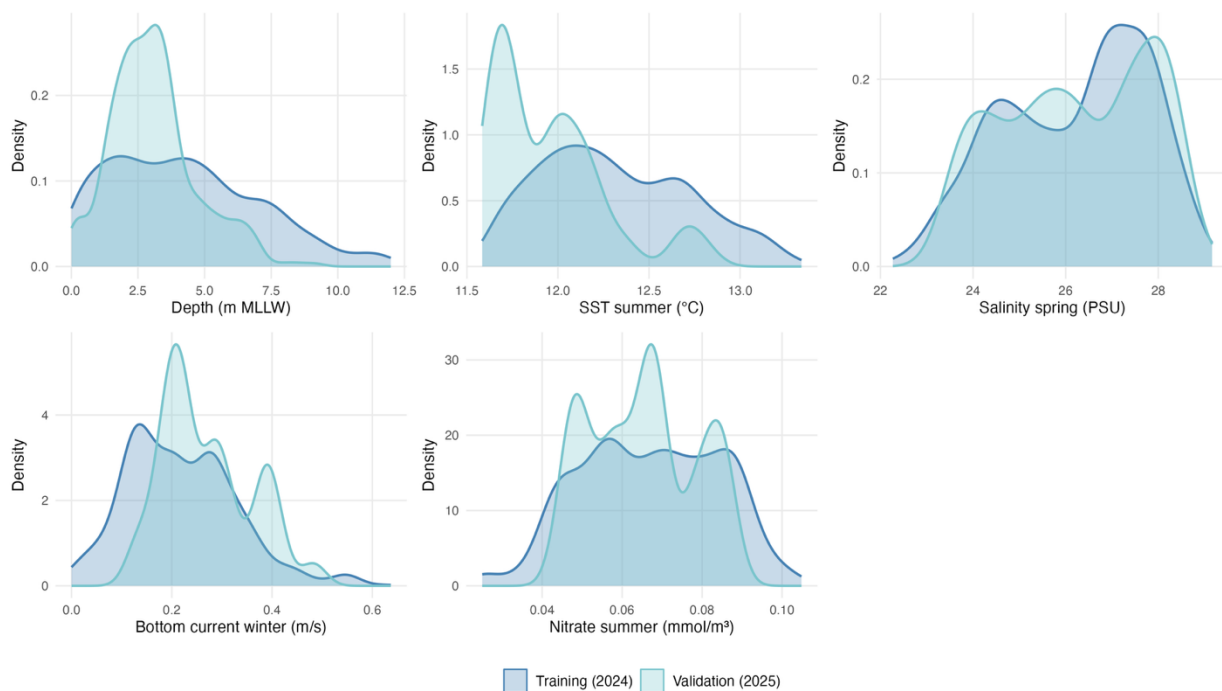


Figure 8. Comparison of predictor variable distributions between HSM_{STD} model training data (2024 presence and absence points, n=888) and independent validation data (2025 snorkel surveys, n=345). Density curves show distributions of depth, SST, salinity, bottom current, and nitrate concentration at training and validation locations. These covariates were selected for visualization because they represent the primary variables expected to differ between training and validation datasets based on the geographic restriction of validation surveys to San Juan Island.

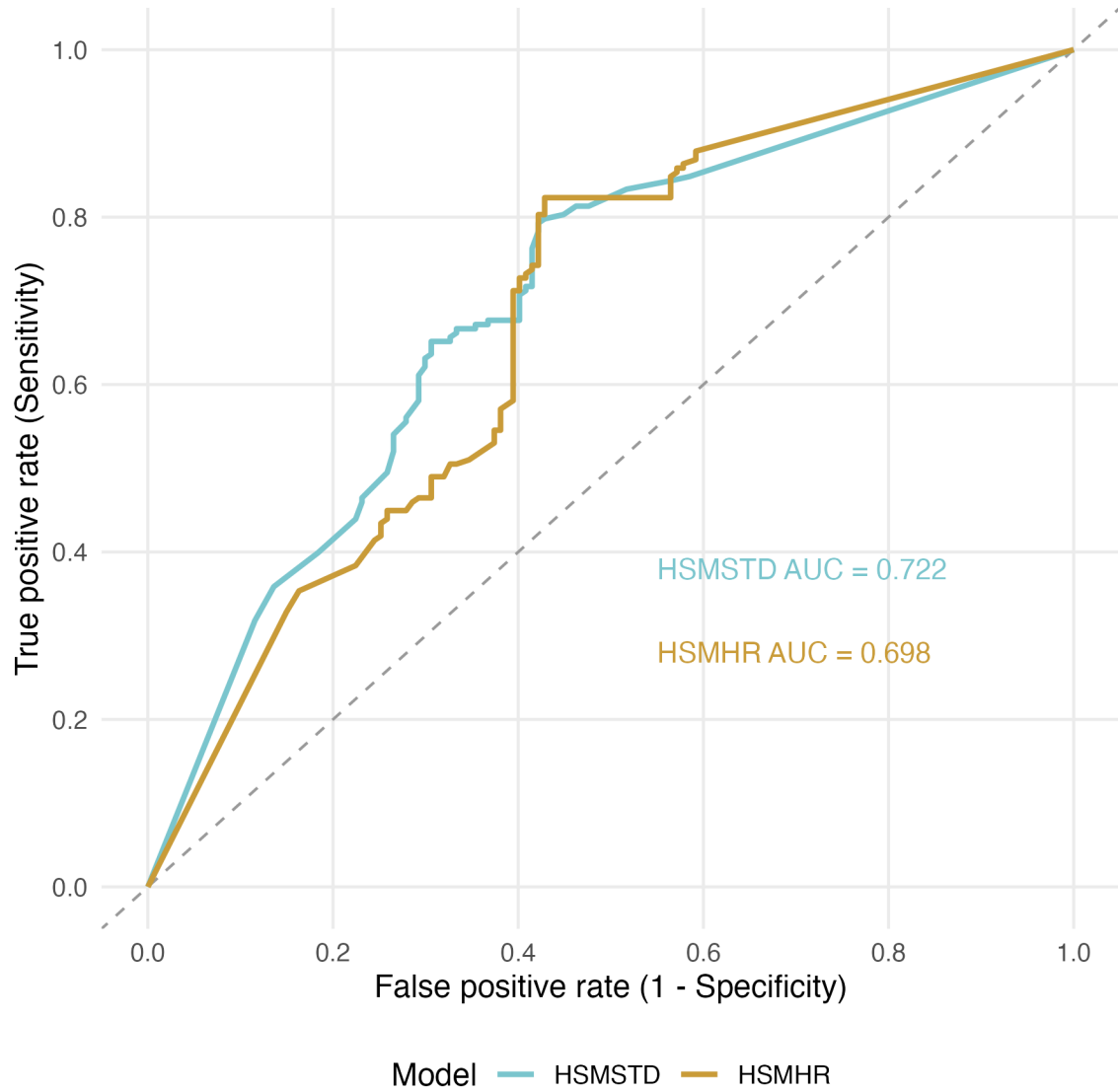


Figure 9. Receiver operating characteristic (ROC) curves for HSM_{STD} and HSM_{HR} models evaluated against independent validation data (2025 snorkel surveys, $n=345$). Each curve plots true positive rate (sensitivity) against false positive rate (1 - specificity) across 200 evenly spaced probability thresholds. The dashed diagonal line indicates no-skill discrimination. HSM_{STD} achieves higher sensitivity than HSM_{HR} across most of the low false positive rates (< 0.40), with curves converging above this threshold. AUC values are shown for each model.

Discussion

Ecological Interpretation of Significant Predictors

Effective fetch and slope emerged as clearly defined predictors in both habitat suitability models, with both demonstrating positive associations with *N. luetkeana* occurrence. Fetch is the dominant predictor, having the largest coefficient and demonstrating a consistent positive association in both models. HSM_{HR} is more conservative about labeling sheltered sites suitable, indicated by the HSM_{HR} fetch response curve threshold being slightly higher at a 0.5 suitability (24km vs 21km). Slope demonstrated a pattern similar to fetch, with a higher threshold for suitability in the HSM_{HR} model (32°) vs the HSM_{STD} model (28°). Fetch and slope likely reflect hard substrate availability as steep and wave-swept environments are often characterized by bedrock and consolidated rock (Bekkby et al., 2009). The rocky, exposed nature of a large portion of shoreline in the San Juan Islands consists of stable substrate that supports *N. luetkeana* recruitment and holdfast attachment (Washington State Department of Natural Resources, 2024). Mobile substrate showed the largest negative coefficient of any predictor in both models, reflecting the inability of *N. luetkeana* to successfully establish holdfasts on unstable sediments. Soft, mobile sediment is more characteristic of gradually sloping and less exposed shorelines, reinforcing the interpretation that fetch, slope, and substrate together capture similar physical habitat conditions relevant to *N. luetkeana* establishment in the San Juan Islands.

In HSM_{HR}, three additional predictors showed clearly defined associations not resolved in HSM_{STD}: surface current, bottom current, and nitrate concentration. However, the practical contribution of these variables to predicted suitability requires careful interpretation given the threshold behavior of their response curves. Hydrodynamic energy regimes are important for kelp in this region, and generally, increased current velocity shows a positive association with *N.*

luetkeana occurrence. Specifically, increased bottom current speed in the winter shows a positive association with *N. luetkeana*, presumably reflecting conditions that are favorable for spore dispersal, sexual reproduction, and early development of sporophytes in winter (Burnett et al., 2024). The divergence between models in the bottom current response curve, with HSM_{STD} suitability approaching 0.90 at the upper end of observed values while HSM_{HR} reached only 0.31, suggests that coarser resolution predictors may conflate localized high-current conditions with broader spatial patterns, increasing the effect of current velocity in the standard resolution HSM_{STD} model. Surface current showed a weak but clearly defined positive association in HSM_{HR} , although maximum predicted suitability remained below 0.20 across the observed range, limiting the value of this variable in identifying suitable habitat.

Nitrate concentration is recognized as a limiting nutrient for kelp growth, particularly during the summer growing season when photosynthetic demand is highest (Dayton, 1985). In the San Juan Islands, nitrate in nearshore environments is predominantly marine-derived, supplied through tidal mixing and upwelling of nutrient-rich waters of the Pacific rather than terrestrial runoff. Over summer, nitrate concentration is progressively depleted through biological uptake (Khangaonkar et al., 2012), and sites with sustained exposure to deeper, nutrient-rich water through strong tidal mixing or upwelling may support kelp growth and persistence into late summer. The positive association between nitrate concentration and *N. luetkeana* occurrence in HSM_{HR} is consistent with this mechanism, although the response curve rose only within the uppermost 5% of observed nitrate concentration values and the relationship may partially reflect spatial co-occurrence with broader oceanographic conditions associated with strong tidal mixing and marine water influence that simultaneously elevate both nitrate concentration and salinity in this system. The correlation between nitrate concentration and

salinity differs substantially between models: $r = 0.80$ in HSM_{STD} versus $r = 0.44$ in HSM_{HR} (Appendix 2), which suggests that finer resolution of salinity partially decouples these variables by capturing fine-scale spatial heterogeneity that coarser resolution interpolation smooths over. The response curve behavior at the boundary of the data means this relationship remains poorly characterized regardless of resolution.

Although depth showed a clearly defined negative association with *N. luetkeana* in HSM_{HR} only ($\beta = -1.11$, CI: -1.94 to -0.28), the response curve indicates that there is no evidence of a meaningful fixed-effect relationship in either model across the 0-12m domain. This discrepancy likely reflects the spatial random field absorbing depth-correlated spatial structure rather than a direct effect of depth on kelp suitability. The discriminatory power of depth as a predictor variable is further limited by the restriction of the modeling domain to 0-12m, the range where kelp was observed to occur, leaving insufficient depth variation to distinguish *N. luetkeana* presence from absence.

TSS was included in both models to capture water clarity as a potential control on light availability for photosynthesis, although this variable showed an unresolved association with *N. luetkeana* in both models. The narrow range of TSS values across the study area (1.59-2.11 mg/L) likely limited the model's ability to detect meaningful differences related to this variable, suggesting that water clarity does not vary sufficiently across the nearshore environments of the San Juan Islands to function as a clear predictor of kelp habitat at this resolution (Appendix 5).

Temperature (both summer SST and winter seafloor temperature) showed unresolved associations in both models. Although temperature variables demonstrate directionally expected relationships within the moderate temperature range observed at training points— higher maximum summer SST is associated with absence while higher mean winter bottom temperature

is associated with presence— temperature effects are not clearly resolved in either model. Summer SST at training points ranged from 11.3 to 14.5°C in HSM_{HR} and 11.6 to 13.3°C in HSM_{STD}. While the upper end of this range approaches temperatures associated with reduced sporophyte production, most values fall within a favorable range for *N. luetkeana* sporophytes in this region (Weigel et al., 2023). Winter bottom temperatures ranged from 7.3 to 9.5°C in HSM_{HR} and 7.7 to 8.8°C in HSM_{STD}, falling below the thermal range where gametophyte densities peak (Weigel et al., 2023), though gametophyte maturation has been demonstrated at temperatures as low as 5°C in *N. luetkeana* (Vadas, 1972). However, this narrow range of bottom temperatures across training points provides insufficient contrast to discriminate presence from absence.

Similarly, salinity demonstrates an unresolved effect in both models. HSM_{HR} salinity shows a negative coefficient near zero, with confidence intervals spanning zero, while in HSM_{STD}, salinity has a stronger negative coefficient though it has much wider 95% confidence intervals that also span zero. The negative coefficient is unexpected given that higher salinity is related to marine water influence and co-occurs with higher nitrate concentrations. When these variables are included jointly, salinity's independent contribution becomes difficult to resolve. Its unresolved association may reflect residual correlation with nitrate concentration, which better captures marine influence in this system rather than representing a direct physiological effect of salinity on *N. luetkeana*. Additionally, for HSM_{HR}, the distribution of spring salinity values at presence and absence points showed substantial overlap, with both groups concentrated near 28-29 PSU (Appendix 5), consistent with the model's inability to resolve a clear directional association.

The differences in predictor resolution between HSM_{STD} and HSM_{HR} can be summarized by the number of clearly defined predictors and the magnitude of fixed-effect uncertainty. HSM_{HR} was able to resolve additional predictors that HSM_{STD} could not (depth, surface current, bottom current, nitrate concentration), and did so with a lower spatial random field standard deviation ($\sigma_O = 9.10$ vs 13.97), indicating that high-resolution hydrodynamic variables used in the HSM_{HR} model provided more fine detail to support clearer separation of the effects from other fixed predictors. This reduction in residual spatial variance supports the interpretation that higher hydrodynamic resolution better characterizes environmental gradients relevant to *N.luetkeana* habitat in the complex marine environment of the San Juan Islands.

Model Performance Comparison

Both models demonstrated strong cross-validation performance, with mean AUC of 0.895 for HSM_{STD} and 0.918 for HSM_{HR} across three spatially independent island folds. These values indicate good discriminative ability within the environmental and geographic conditions represented by the 2024 training dataset. However, cross-validation and independent validation tell different stories about model reliability. Cross-validation with island-based blocking assesses how well the model generalizes across spatially independent regions within the same year of training data, testing spatial transferability but not temporal or methodological independence. Independent validation against 2025 snorkel surveys tests whether the model makes useful predictions when evaluated against data collected by a different method, in a different year, at locations that were not used to fit the model. The substantially lower independent validation AUC values (HSM_{STD} 0.722, HSM_{HR} 0.698) reflect the more rigorous nature of this test and should be interpreted as the more conservative and realistic estimate of model predictive performance.

The divergence between cross-validation and independent validation AUC values can be attributed to several sources of mismatch between validation and training datasets. The most notable is the depth distribution of independent validation points: 302 of 345 validation points fall at depths of 5m or less, while training data extends more evenly across the full 0-12m domain (Figure 8). This means that independent validation disproportionately evaluates model performance in shallow environments and underrepresents the deeper portions of the modeling domain. SST at validation locations was also more narrowly distributed around 12°C compared to the broader thermal range in training data, suggesting validation points were restricted to a slightly cooler subset of the environment the models were trained on. Bottom current values at validation locations were similarly concentrated at lower velocities than the training distribution, so the model was evaluated primarily in calmer nearshore environments. All of these environmental differences reflect the geographic restriction of validation surveys to San Juan Island, and the shallower habitats surveyed via snorkeling. Additionally, predictor values at validation locations were extracted from 2024 raster surfaces rather than 2025 conditions, and kelp distribution is inherently dynamic between years. Beds that were present in 2024 may have shifted, contracted, or expanded by 2025, introducing biological mismatch that environmental predictors alone cannot account for. While these differences between datasets account for much of the divergence, the low HSM_{STD} AUC at depths greater than 5m suggests that some of the divergence reflects genuine limitations in the models rather than error introduced by validation design alone.

Depth-stratified independent validation revealed a particularly important limitation for HSM_{STD} . At shallow sites ($\leq 5m$), both models performed reasonably well (HSM_{STD} 0.756, HSM_{HR} 0.718), consistent with the concentration of training data and validation points in this

depth range. At deeper sites ($>5\text{m}$, $n=43$), the HSM_{STD} AUC of 0.485 indicates essentially no-skill discrimination, whereas the HSM_{HR} achieved an AUC of 0.559. This suggests that the HSM_{STD} model's performance at depth is limited for applications requiring predictions across the full 0-12m domain and is consistent with coarser resolution predictors being insufficient to capture the oceanographic conditions that structure kelp habitat in deeper nearshore environments.

The cross-validation results reveal differences between Orcas Island, where HSM_{STD} outperformed HSM_{HR} (0.912 HSM_{STD} vs 0.826 HSM_{HR}), and San Juan Island (0.801 HSM_{STD} vs 0.939 HSM_{HR}) and Lopez Island (0.971 HSM_{STD} vs 0.990 HSM_{HR}) where HSM_{HR} showed stronger performance. Diagnostic analysis of training point distributions revealed that Orcas presence points are spatially concentrated along the northern shoreline and are associated with notably lower salinity ($\sim 23\text{-}24$ PSU) and lower effective fetch than presence points on the other two islands, which peak near 27-28 PSU (Appendix 4). When the model is trained without the Orcas fold, predictor-response relationships are learned primarily from higher-salinity, more exposed environments on San Juan and Lopez Islands. Transferring these relationships to the Orcas fold where kelp primarily occupies a distinct portion of the island across a much broader area reduces predictive accuracy in both models. This pattern highlights a limitation of island-based spatial blocking approach when environmental conditions differ systematically across folds rather than varying continuously across the study area. Block-level AUC estimates are sensitive to the environmental representativeness of each fold, not just its spatial independence.

The validation results support a nuanced interpretation of model performance when considered all together. Cross-validation AUC values in the range of 0.895 to 0.918 indicate that both models generalize well spatially within the conditions and year represented by training data,

and HSM_{HR}'s advantage on San Juan and Lopez folds suggests meaningful improvement from higher predictor resolution. Independent validation AUC values of 0.722 and 0.698 are modest but remain above no-skill discrimination, and the multiple sources of predictor mismatch between training and validation datasets (Figure 8) mean these values likely underestimate true predictive performance under more closely aligned environmental conditions. In summary, cross-validation performance is the more reliable indicator of the model's ability to generalize spatially within the training year, while independent validation reveals the model's limits of transferability to novel conditions. These limitations are partially methodological and partially reflect genuine uncertainty in kelp distribution between years.

Prediction Output Comparison and Uncertainty Estimates

Both models produced broadly similar spatial predictions of *N. luetkeana* habitat suitability, with high suitability concentrated along exposed outer coastlines and lower suitability in interior channels, bays, and sheltered inlets (Figures 3, 4). The general agreement between models in the location of high-suitability habitat is consistent with the shared predictor set and similar coefficient directions for the strongest predictors (fetch, slope, and mobile substrate) which drive the spatial pattern of predictions in both models. Where models diverge, HSM_{HR} identifies suitable habitat on the shores of smaller, more isolated rocky outcroppings that HSM_{STD} does not resolve as suitable, and HSM_{HR} predictions in continuously suitable areas appear somewhat more spatially coherent and tightly bounded, reflected in the lower total suitable area estimate (1,705 ha vs 2,241 ha). This pattern is consistent with HSM_{HR}'s finer resolution capturing environmental gradients at small island and reef edges that coarser interpolation smooths over, producing more spatially precise boundaries between suitable and unsuitable habitat rather than simply predicting less suitable area overall.

Prediction uncertainty was generally higher in HSM_{STD} than HSM_{HR} across the study area, consistent with HSM_{STD} 's larger spatial random field standard deviation (Figure 5). Where predicted suitability is high and uncertainty is low model predictions can be used with greater confidence, and where uncertainty is elevated, suitability predictions should be evaluated through supplementary field observations and experiments.

Areas where both models predict high suitability and have high certainty represent strong candidates for kelp habitat assessment and restoration planning, as agreement between the independent models provides stronger evidence of genuine environmental suitability than either model alone. Areas where models diverge, particularly smaller rocky outcroppings and transitional zones identified as suitable by HSM_{HR} but not HSM_{STD} , should be verified in the field before restoration investment, as these predictions depend on detailed hydrodynamic gradients that HSM_{HR} resolves but that remain poorly understood within the range of the training data. An interactive map of HSM_{HR} suitability predictions and uncertainty estimates is available as a supplementary resource online and may support site-level exploration by resource managers and restoration practitioners (<https://ebews.github.io/bull-kelp-habitat-suitability-final/>).

With these results in mind, the two models tested here serve distinct purposes for research and management planning. HSM_{STD} predictions are well-suited for regional habitat assessments, priority-setting across the full archipelago, and identifying broad areas warranting further survey effort. HSM_{HR} predictions are more appropriate for more specific site selection, distinguishing suitable from unsuitable habitat at a site-level scale, and identifying specific locations for targeted field verification prior to experimental restoration site prioritization.

Limitations

The following limitations should be considered in interpretation of these model predictions and results. As previously discussed, independent validation AUC values likely underestimate true predictive performance given the environmental and temporal mismatch between training and validation datasets, though the degree of underestimation cannot be quantified without a more closely matched validation dataset.

The models reflect environmental suitability based on physical and oceanographic conditions and do not account for all factors relevant to kelp establishment and persistence. Biotic factors including grazing pressure from herbivores and competition from other algae, which are known to influence kelp distribution in this region, were not included as predictors and may reduce actual *N. luetkeana* occurrence in areas the models identify as environmentally suitable. Site disturbance history and actual substrate patterns at scales finer than the prediction grid may also limit kelp presence at sites that appear to be highly suitable in the models.

These predictions should be interpreted as characterizations of habitat suitability under the specific oceanographic conditions of 2023-2024 rather than long-term average suitability. *N. luetkeana* is an annual species with high interannual variability in recruitment, distribution, and canopy extent. A single year of oceanographic predictors is unlikely to capture the full range of conditions that drive kelp persistence across years, and predictions may not be generalizable to years with substantially different temperature, current, or nutrient regimes. Incorporating multiple years of hindcast data from the Salish Sea Model and multiple years of kelp presence/absence data into future modeling efforts would improve the robustness of habitat suitability predictions across variable oceanographic conditions.

Conclusion

This study presents the first habitat suitability models for *N. luetkeana* in the San Juan Islands, using oceanographic data derived from standard and high-resolution configurations of the Salish Sea Model. The results directly address the three motivating research questions: what is the distribution of suitable habitat for *N. luetkeana* in this region, which environmental factors are associated with *N. luetkeana* occurrence, and whether finer resolution of hydrodynamic predictor data improves model performance in terms of spatial detail, estimated suitable habitat area, and predictive accuracy.

Suitable habitat for *N. luetkeana* in the San Juan Islands is concentrated along exposed rocky outer coastlines throughout the study area, with both models identifying consistent spatial patterns driven primarily by wave exposure and substrate stability. Effective fetch and slope were the dominant positive predictors across both models, with mobile substrate as a strong negative predictor, together indicating that wave exposure and hard substrate availability are the primary physical requirements for *N. luetkeana* canopy establishment and persistence in this system.

Finer resolution hydrodynamic predictor data improved habitat suitability predictions in several ways. HSM_{HR} resolved four additional predictor associations not clearly defined in HSM_{STD} (depth, surface current, bottom current, and nitrate concentration) and produced a substantially lower spatial random field standard deviation ($\sigma_O = 9.10$ vs 13.97), indicating that finer resolution data helps capture ecologically relevant spatial structure that standard resolution oceanographic data cannot resolve. HSM_{HR} outperformed HSM_{STD} in cross-validation with AUC values of HSM_{HR} 0.918 vs HSM_{STD} 0.895, while independent validation against 2025 snorkel survey data yielded more modest AUC values for both models (HSM_{STD} 0.722, HSM_{HR} 0.698),

although these values likely underestimate true predictive performance given spatial and temporal mismatch between training and validation datasets. Taken together, these results support the conclusion that higher-resolution hydrodynamic predictor data meaningfully improves habitat suitability modeling in the complex tidal environment of the San Juan Islands, particularly in improving spatial precision at a site-specific level.

These findings have direct implications for *N. luetkeana* monitoring and restoration planning in the San Juan Islands. The suitability maps provide a spatial framework for prioritizing field surveys and identifying candidate restoration sites, with areas of high suitability and high model agreement representing the strongest candidates for restoration consideration and areas of model divergence warranting field verification before investment. The direct comparison of HSM_{STD} and HSM_{HR} predictions demonstrates that oceanographic predictor resolution has practical consequences for habitat area estimates, predictor interpretation, and site-level spatial precision.

The primary limitations of this work are the restriction to a single year of oceanographic conditions and the absence of biotic predictors known to influence kelp distribution, such as grazing pressure and algal competition. Multi-year spatiotemporal modeling using repeated Salish Sea Model hindcasts and Samish kelp canopy data should be prioritized in future extensions of this work within the sdmTMB framework. Spatiotemporal analysis was not completed within the scope of this study as high resolution SSM hindcast data was only available for 2023 and 2024; as more high resolution SSM hindcast data becomes available the application of this framework across years would enable estimation of interannual variability in suitable habitat and provide a basis for understanding changes in *N. luetkeana* distribution over time. Additionally, expanded validation surveys covering all three major islands and deeper portions

of the modeling domain would substantially improve assessment of model transferability. The incorporation of hydrodynamic and biogeochemical predictor data from SSM2.0, the next-generation configuration of the Salish Sea Model that eliminates bathymetric smoothing, represents an important extension of this work. Khangaonkar et al. (2026) demonstrated that bathymetric smoothing in SSM_{HR} led to underprediction of currents and temperature in nearshore intertidal regions, with SSM2.0 showing substantially improved skill. These are the environments where *N. luetkeana* habitat predictions are most consequential, suggesting that SSM2.0-derived predictors could improve model performance in nearshore areas where SSM_{HR} resolution limitations are most pronounced. Additionally, high-resolution nitrate concentration data, when available from ongoing water quality model development, would directly address the most significant predictor resolution gap identified in this study. Repeated application of this modeling framework to future Samish kelp canopy data would enable tracking of habitat suitability shifts over time, connecting observed changes in kelp canopy extent to the environmental conditions this study identifies as most relevant to *N. luetkeana* persistence in the San Juan Islands.

Appendix 1. Shoreline barrier feature repair locations

Table A1. Shoreline barrier feature repair locations for spatial interpolation of SSM (HYD and WQM) and SSM_{HR} (HYD) variables. Coordinates provided mark the general area where manual edits were used to reconnect or reestablish land areas in the shoreline barrier feature used for Spline with Barriers interpolation.

Island	Location and Description	Coordinates
Lopez Island	John's Point near Barlow Bay (south side of Lopez), shoreline reestablished along Barlow Bay Rd.	48.4325311, -122.8766745
Lopez Island	Peninsula northeast of Islandale (southeast side of Lopez), shoreline reestablished along Sperry Rd.	48.4614854, -122.8198040
Lopez Island	Spencer Spit State Park (northeast side of Lopez) shoreline extended slightly to capture this point of the island	48.5362638, -122.8559771
Lopez Island	Fisherman Bay Preserve (west side of Lopez), shoreline reestablished along Bayshore Rd.	48.50540371, -122.93414255
Lopez Island	Humphrey Head (north side of Lopez), shoreline reestablished along Shoreland Dr.	48.553046, -122.871430
San Juan Island	White Point (northwest side of San Juan), shoreline extended along White Point Rd.	48.595896, -123.165690
San Juan Island	Davidson Head (north side of San Juan), shoreline extended slightly to the north and east to reestablish the land feature	48.623042, -123.148684
San Juan Island	Yacht Haven (northwest side of San Juan), shoreline reestablished along Yacht Haven Rd.	48.577630, -123.150273
San Juan Island	Yacht Haven (northwest side of San Juan), shoreline extended along De Haro Rd.	48.5729819, -123.1620100
San Juan Island	Yacht Haven (northwest side of San Juan), shoreline extended along Shorett Dr. and Inskeep Ln.	48.587681, -123.165102
San Juan Island	Cape San Juan near Cattle Point (south side of San Juan), shoreline reestablished along Cape Dr.	48.460072, -122.964384
Henry Island	Henry Island Preserve, shoreline reestablished at central isthmus of the island	48.596541, -123.183406
Shaw Island	University of Washington Biological Preserve (southwest side of Shaw), shoreline reestablished out to Point George	48.557980, -122.982480
Shaw Island	University of Washington Biological Preserve (south side of Shaw), shoreline extended along Lyckebo Ln.	48.552093, -122.965134
Shaw Island	Shaw County Park (south side of Shaw), shoreline extended at Squaw Bay Rd.	48.560831, -122.942006
Shaw Island	Shaw County Park (south side of Shaw), shoreline extended along Indian Cove Rd.	48.5642919, -122.9261761

Shaw Island	Neck Point (northwest side of Shaw), shoreline extended along Neck Point Rd.	48.582698, -123.004968
Decatur Island	Kimball Preserve (south side of Decatur), shoreline extended along Reed Shipyard Rd. to Kimball Preserve	48.48681, -122.818547
Decatur Island	Sylvan Cove (west side of Decatur), shoreline extended along Peninsula Rd.	48.5115999, -122.8363267
Decatur Island	Decatur Head (east side of Decatur), shoreline extended along Decatur Head Dr.	48.5111603, -122.7851249
Spieden Island	West side of island, shoreline extended to restore continuity	48.645450, -123.151727
Sucia Island	Shoreline reestablished along full extent of island to restore continuity	48.7628002, -122.9128839

Appendix 2. Correlation matrix for HSM_{STD} and HSM_{HR} full predictor candidate sets

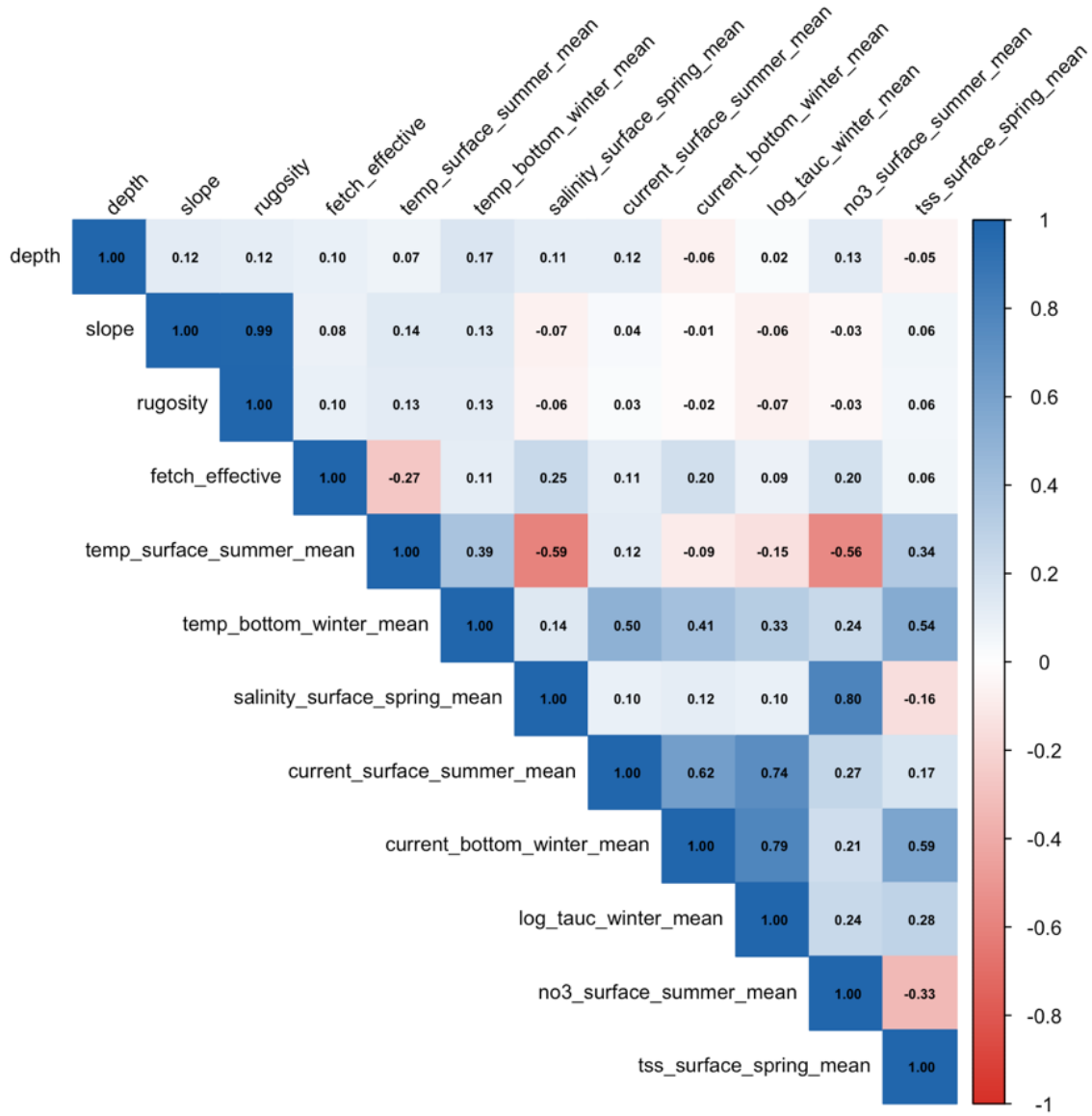


Figure A2.1. Correlation matrix for HSM_{STD} full predictor candidate set. Pearson correlation coefficients were computed for all candidate predictor variables prior to final predictor selection to identify collinear variables. Color intensity indicates correlation strength. Variable pairs with $|r| > 0.70$ were flagged as potentially collinear. Notable correlation is observed with rugosity and slope ($r = 0.99$), nitrate and salinity ($r = 0.80$), current velocity and bed shear stress variables ($r > 0.74$), and TSS, turbidity, and other light extinction variables ($r > 0.93$)

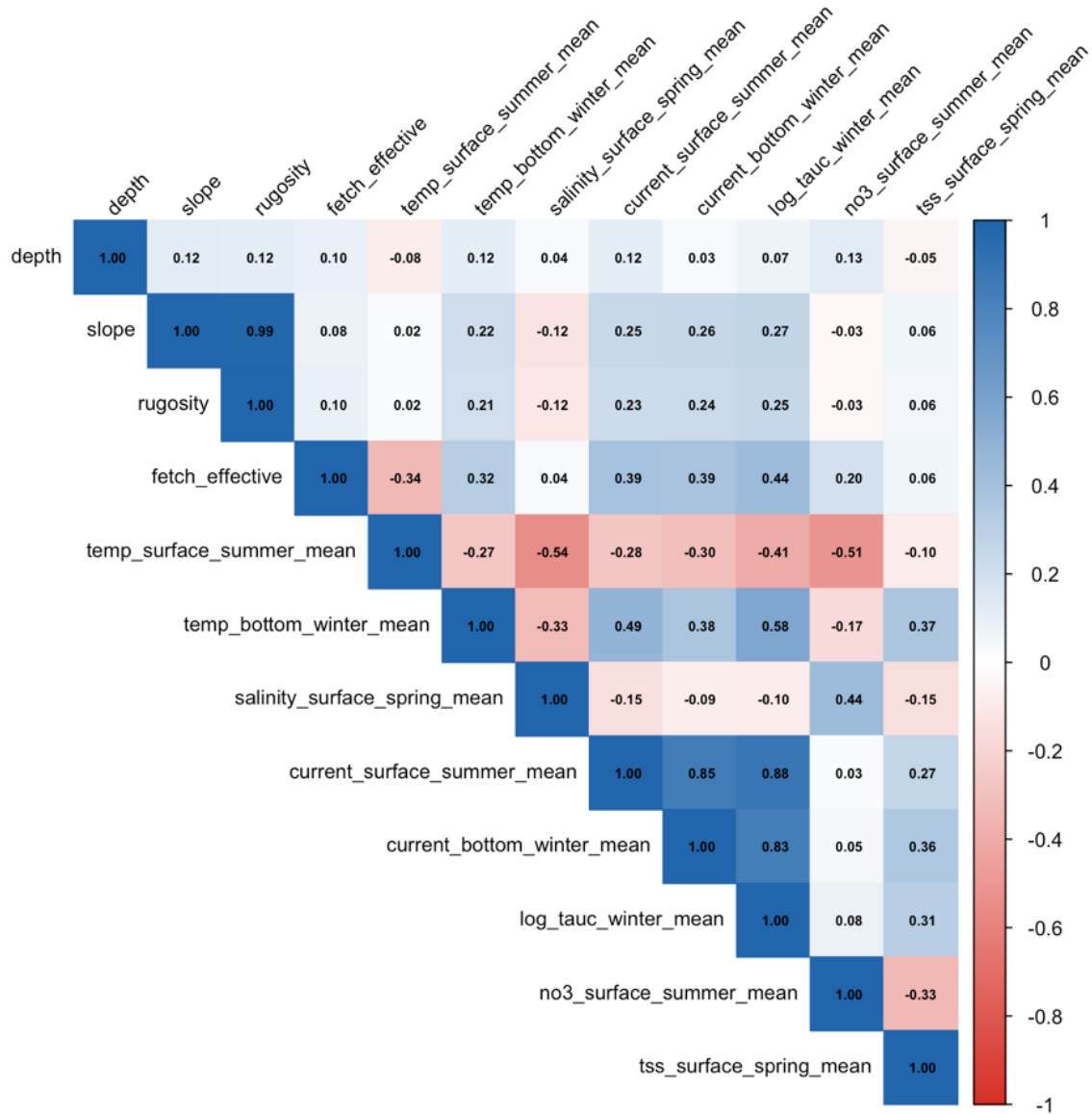


Figure A2.2. Correlation matrix for HSM_{HR} full predictor candidate set. Pearson correlation coefficients were computed for all candidate predictor variables prior to final predictor selection to identify collinear variables. Color intensity indicates correlation strength. Variable pairs with $|r| > 0.70$ were flagged as potentially collinear. Notable correlation is observed with rugosity and slope ($r = 0.99$), current velocity variables and bed shear stress variables ($r > 0.83$), and TSS, turbidity, and other light extinction variables ($r > 0.93$)

Appendix 3. Model comparisons under different predictor configurations.

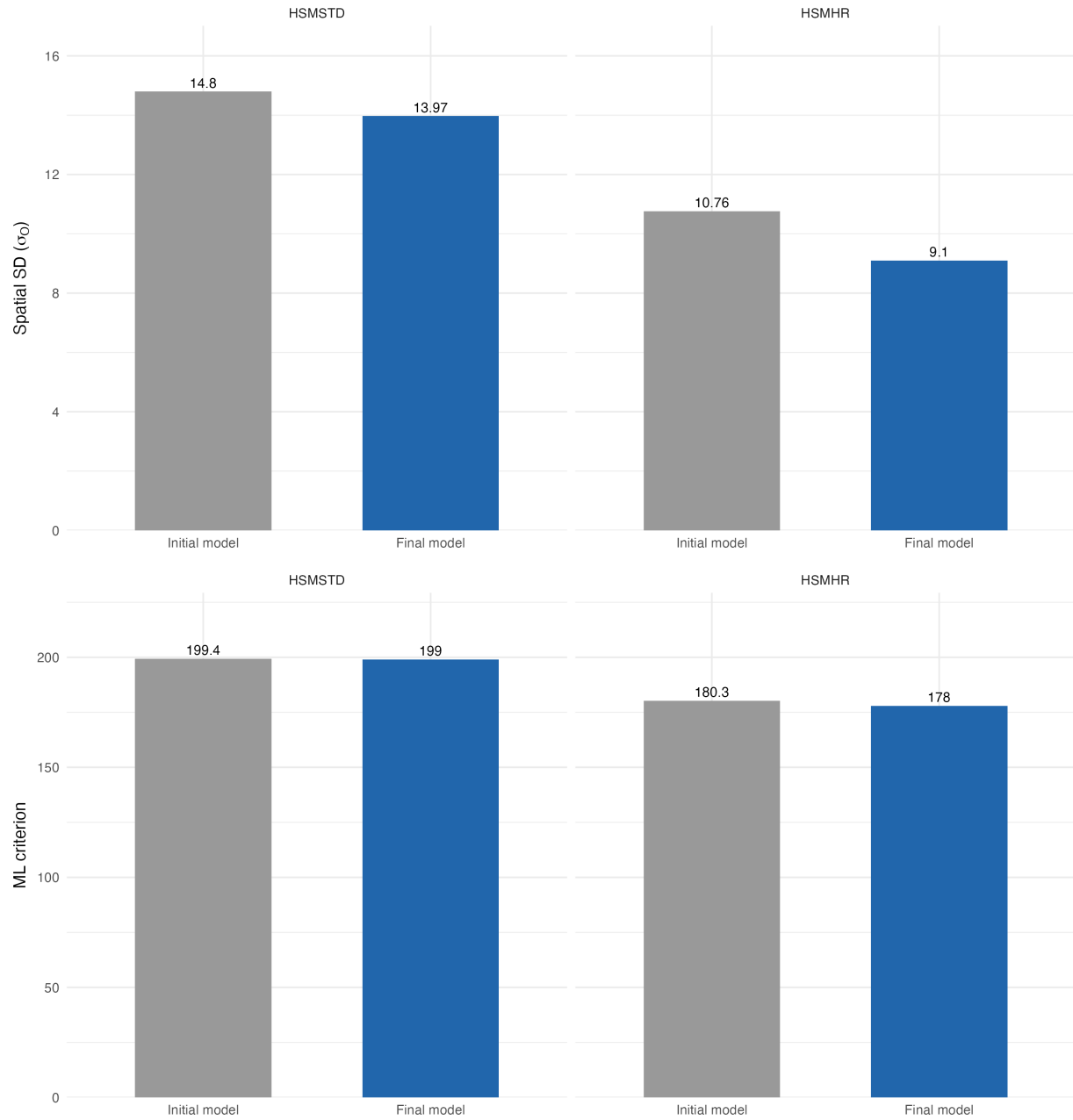


Figure A3. Comparison of spatial random field standard deviation (σ_0) and marginal likelihood (ML) criterion between HSM_{STD} and HSM_{HR} models under two configurations. The initial predictor configuration included winter bottom bed shear stress, while the final predictor configuration fit the model with winter bottom current instead as a mechanistic predictor for hydrodynamic conditions affecting *N. luetkeana* gametophytes over winter. The final predictor configuration was selected based on a decrease in both σ_0 and ML criterion, indicating improved fixed-effect explanation of spatial structure and overall model fit.

Appendix 4. Orcas island cross-validation diagnostics

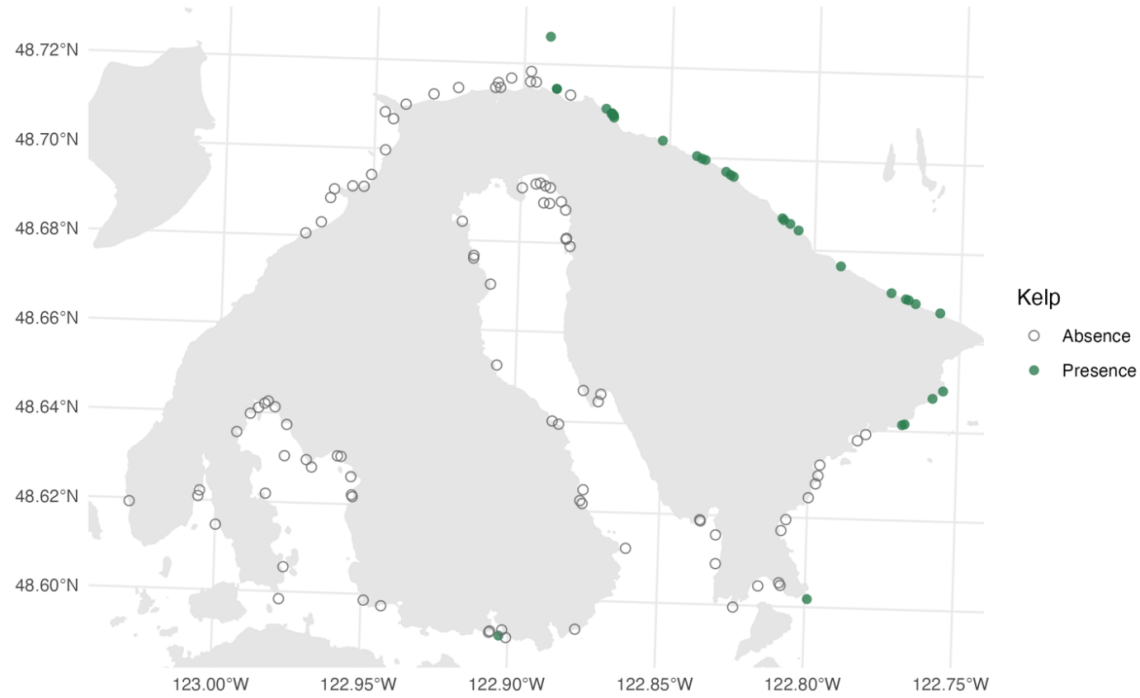


Figure A4.1. Spatial distribution of training presence and absence points on Orcas Island.

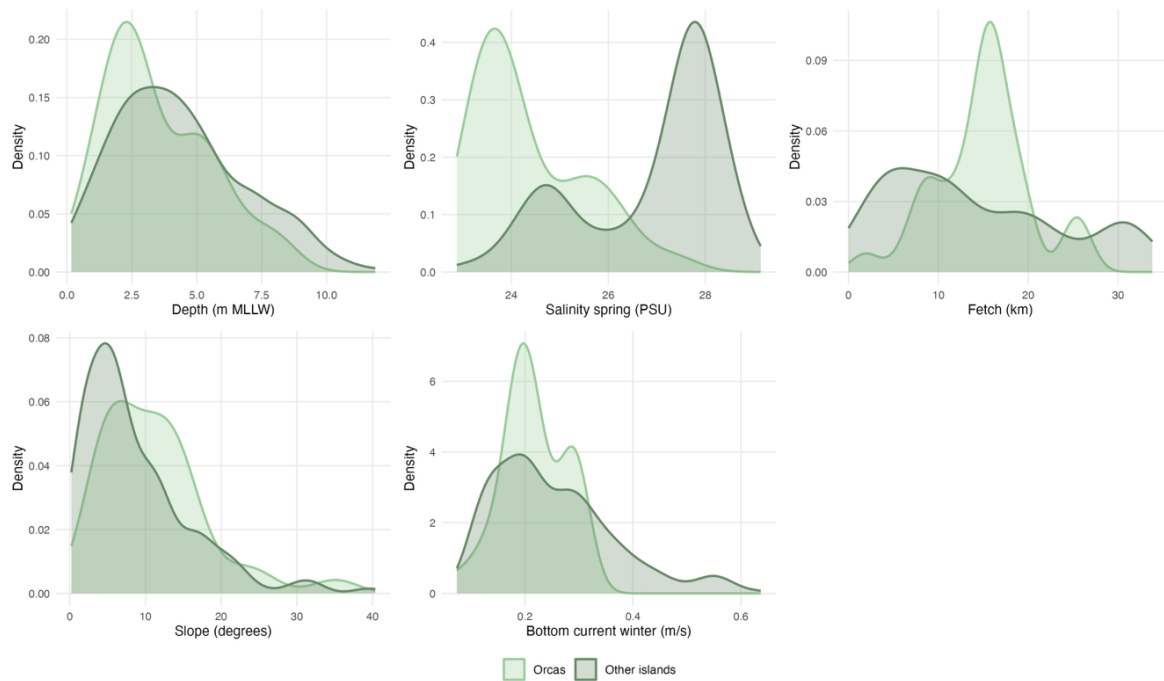


Figure A4.2. Density plots comparing distribution of environmental factors at Orcas Island kelp presence points compared to presence points on San Juan and Lopez Islands.

Appendix 5. Distribution of predictor values at *N. luetkeana* presence and absence points

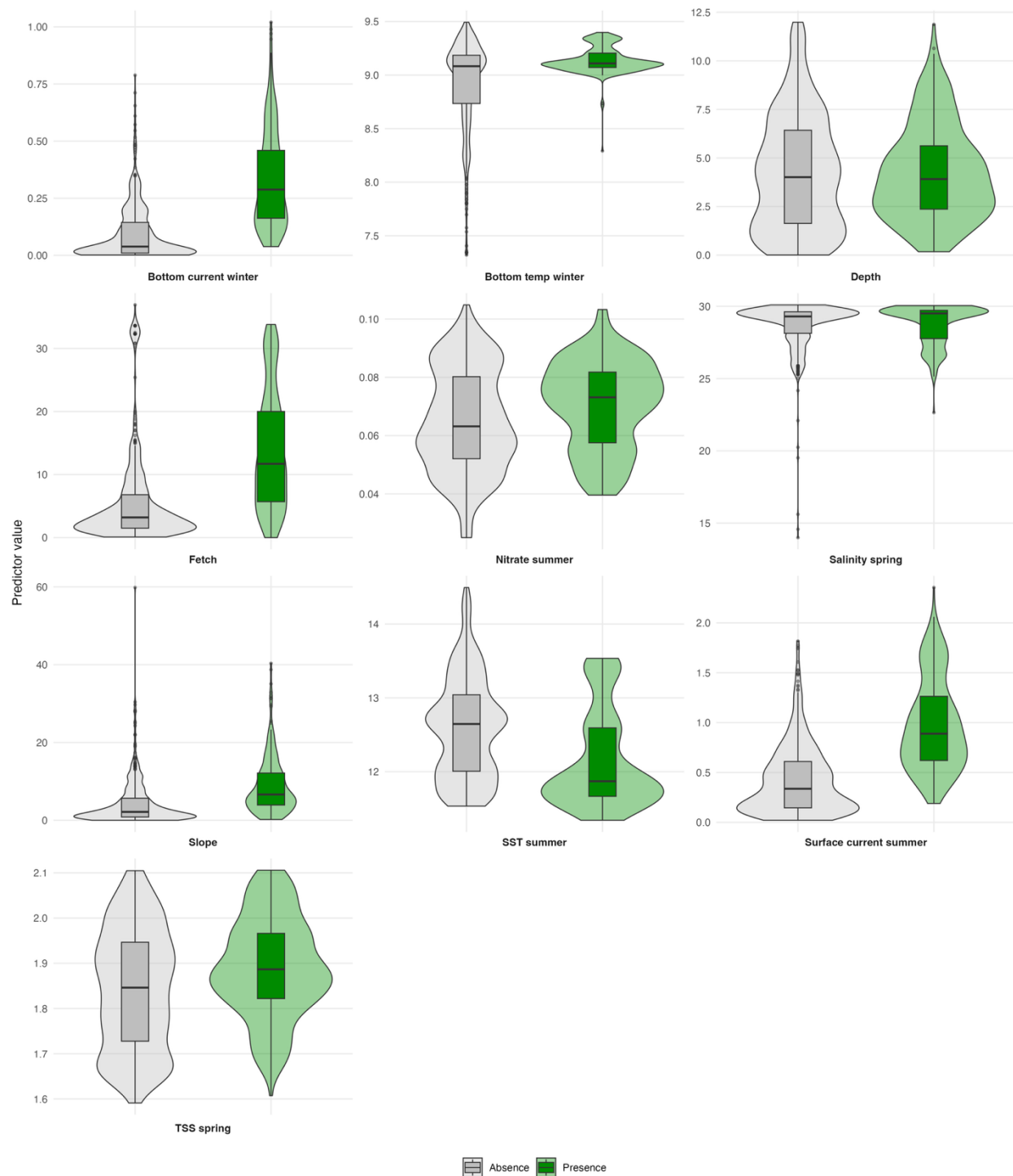


Figure A5.1. Distribution of predictor values at *N. luetkeana* presence and absence points in the HSM_{HR} training dataset (n=888). Violin plots show the full distribution, boxplots within show median and interquartile range.

References

- Anderson, S. C., Ward, E. J., English, P. A., Barnett, L. A. K., & Thorson, J. T. (2025). sdmTMB: An R Package for Fast, Flexible, and User-Friendly Generalized Linear Mixed Effects Models with Spatial and Spatiotemporal Random Fields. *Journal of Statistical Software*, 115, 1–46. <https://doi.org/10.18637/jss.v115.i02>
- Bekkby, T., & Moy, F. E. (2011). Developing spatial models of sugar kelp (*Saccharina latissima*) potential distribution under natural conditions and areas of its disappearance in Skagerrak. *Estuarine, Coastal and Shelf Science*, 95(4), 477–483. <https://doi.org/10.1016/j.ecss.2011.10.029>
- Bekkby, T., Moy, F. E., Kroglund, T., Gitmark, J. K., Walday, M., Rinde, E., & Norderhaug, K. M. (2009). Identifying Rocky Seabed Using GIS-Modeled Predictor Variables. *Marine Geodesy*, 32(4), 379–390. <https://doi.org/10.1080/01490410903297816>
- Berry, H. D., Mumford, T. F., Christiaen, B., Dowty, P., Calloway, M., Ferrier, L., Grossman, E. E., & VanArendonk, N. R. (2021). Long-term changes in kelp forests in an inner basin of the Salish Sea. *PLOS ONE*, 16(2), e0229703. <https://doi.org/10.1371/journal.pone.0229703>
- Burnett, N. P., Ricart, A. M., Winquist, T., Saley, A. M., Edwards, M. S., Hughes, B., Hodin, J., Baskett, M. L., & Gaylord, B. (2024). Bimodal spore release heights in the water column enhance local retention and population connectivity of bull kelp, *Nereocystis luetkeana*. *Ecology and Evolution*, 14(8), e70177. <https://doi.org/10.1002/ece3.70177>
- Calloway, M., Oster, D., Berry, H., Mumford, T., Naar, N., Peabody, B., Hart, L., Tonnes, D., Copps, S., Selleck, J., Allen, B., & Toft, J. (2020). *Puget Sound kelp conservation and recovery plan*. Northwest Straits Commission. <https://nwstraits.org/sites/default/files/2025-10/pugetsoundkelpconservationandrecoveryplan.pdf>
- Christiaen, B., Berry, H., Ferrier, L., Dowty, P., Gaeckle, J., Johnson, L., Sanchez, M., & Cowdrey, T. (2023). *Marine vegetation in the southern Salish Sea: Final report to the SeaDoc Society* (Nos. A17-0568-S012). https://dnr.wa.gov/sites/default/files/2025-03/aqr_nrsh_final_report_a17_0568_s012_f20ap12280.pdf
- Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado, Boulder. (2014). *Continuously updated digital elevation model (CUDEM) – 1/9 arc-second resolution bathymetric-topographic tiles: San Juan Island, Washington subset* [Data set]. NOAA National Centers for Environmental Information. <https://doi.org/10.25921/ds9v-ky35>
- Dayton, P. K. (1985). Ecology of Kelp Communities. *Annual Review of Ecology and Systematics*, 16(1), 215–245. <https://doi.org/10.1146/annurev.es.16.110185.001243>
- Dethier, M. (1990). *A Marine and Estuarine Habitat Classification System for Washington State*. Dept. of Natural Resources, Washington Natural Heritage Program. https://dnr.wa.gov/sites/default/files/2025-03/amp_nh_marine_class.pdf

- Dethier, M. (2014). *Shoreline habitat classifications*. Puget Sound Institute.
<https://www.eopugetsound.org/habitats/shore-types>
- Eckman, J., Duggins, D., & Siddon, C. (2003). Current and wave dynamics in the shallow subtidal: Implications to the ecology of understory and surface-canopy kelps. *Marine Ecology Progress Series*, 265, 45–56. <https://doi.org/10.3354/meps265045>
- Eger, A. M., Marzinelli, E. M., Beas-Luna, R., Blain, C. O., Blamey, L. K., Byrnes, J. E. K., Carnell, P. E., Choi, C. G., Hessing-Lewis, M., Kim, K. Y., Kumagai, N. H., Lorda, J., Moore, P., Nakamura, Y., Pérez-Matus, A., Pontier, O., Smale, D., Steinberg, P. D., & Vergés, A. (2023). The value of ecosystem services in global marine kelp forests. *Nature Communications*, 14(1), 1894. <https://doi.org/10.1038/s41467-023-37385-0>
- Fales, R. J., Weigel, B. L., Carrington, E., Berry, H. D., & Dethier, M. N. (2023). Interactive effects of temperature and nitrogen on the physiology of kelps (*Nereocystis luetkeana* and *Saccharina latissima*). *Frontiers in Marine Science*, 10, 1281104.
<https://doi.org/10.3389/fmars.2023.1281104>
- Farrugia Drakard, V., Hollarsmith, J. A., & Stekoll, M. S. (2025). Hyposaline conditions impact the early life-stages of commercially important high-latitude kelp species. *Journal of Phycology*, 61(2), 317–329. <https://doi.org/10.1111/jpy.70003>
- Filbee-Dexter, K., Wernberg, T., Grace, S. P., Thormar, J., Fredriksen, S., Narvaez, C. N., Feehan, C. J., & Norderhaug, K. M. (2020). Marine heatwaves and the collapse of marginal North Atlantic kelp forests. *Scientific Reports*, 10(1), 13388.
<https://doi.org/10.1038/s41598-020-70273-x>
- Fox, J., & Monette, G. (1992). Generalized Collinearity Diagnostics. *Journal of the American Statistical Association*, 87(417), 178–183. <https://doi.org/10.2307/2290467>
- García-Reyes, M., Thompson, S. A., Rogers-Bennett, L., & Sydeman, W. J. (2022). Winter oceanographic conditions predict summer bull kelp canopy cover in northern California. *PLOS ONE*, 17(5), Article 5. <https://doi.org/10.1371/journal.pone.0267737>
- GeoTerra, Inc. (2024). *WADNR Kelp and Seagrass Imagery Acquisition Report* (GT#240126). Prepared for NV5 GeoSpatial.
- Gregr, E. J., Palacios, D. M., Thompson, A., & Chan, K. M. A. (2019). Why less complexity produces better forecasts: An independent data evaluation of kelp habitat models. *Ecography*, 42(3), 428–443. <https://doi.org/10.1111/ecog.03470>
- Hauser, A., Attrill, M., & Cotton, P. (2006). Effects of habitat complexity on the diversity and abundance of macrofauna colonising artificial kelp holdfasts. *Marine Ecology Progress Series*, 325, 93–100. <https://doi.org/10.3354/meps325093>
- Hollarsmith, J. A., Andrews, K., Naar, N., Starko, S., Calloway, M., Obaza, A., Buckner, E., Tonnes, D., Selleck, J., & Therriault, T. W. (2022). Toward a conceptual framework for managing and conserving marine habitats: A case study of kelp forests in the Salish Sea. *Ecology and Evolution*, 12(1), e8510. <https://doi.org/10.1002/ece3.8510>
- Khangaonkar, T., Nugraha, A., Xu, W., Long, W., Bianucci, L., Ahmed, A., Mohamedali, T., & Pelletier, G. (2018). Analysis of Hypoxia and Sensitivity to Nutrient Pollution in Salish

- Sea. *Journal of Geophysical Research: Oceans*, 123(7), 4735–4761.
<https://doi.org/10.1029/2017JC013650>
- Khangaonkar, T., Sackmann, B., Long, W., Mohamedali, T., & Roberts, M. (2012). Simulation of annual biogeochemical cycles of nutrient balance, phytoplankton bloom(s), and DO in Puget Sound using an unstructured grid model. *Ocean Dynamics*, 62(9), 1353–1379.
<https://doi.org/10.1007/s10236-012-0562-4>
- Khangaonkar, T., Wang, T., Yun, S. K., Ni, W., Nugraha, A., Mitra, A., & Zhang, Y. J. (2026). Simultaneous resolution of near-shore intertidal and deep estuarine circulation by eliminating bathymetric smoothing in modeling—The Salish Sea. *Ocean Dynamics*, 76(4), 42. <https://doi.org/10.1007/s10236-026-01794-8>
- Krumhansl, K., & Scheibling, R. (2012). Production and fate of kelp detritus. *Marine Ecology Progress Series*, 467, 281–302. <https://doi.org/10.3354/meps09940>
- McKenna, G. E., & Claar, D. C. (2026). Synthesizing disparate data for a comprehensive view of floating kelp distribution in Washington State, USA. *Journal of Applied Phycology*, 38(2), 1737–1751. <https://doi.org/10.1007/s10811-026-03829-5>
- McKenna, G., Fox, V., Wolf, A., & Ainsworth, C. (2022). *Bull Kelp Restoration Site Selection Tool*. University of Washington.
<https://experience.arcgis.com/experience/b11daaa83ff045f1a9d88b2b926e1f75/page/Suitability-Map>
- Naar, N. (2020). APPENDIX B: The Cultural Importance of Kelp for Pacific Northwest Tribes. In M. Calloway, D. Oster, H. Berry, T. Mumford, N. Naar, B. Peabody, L. Hart, D. Tonnes, S. Copps, J. Selleck, B. Allen, & J. Toft, *Puget Sound Kelp Conservation and Recovery Plan* (p. B-1-B-9). Northwest Straits Commission.
<https://nwstraits.org/sites/default/files/2025-10/pugetsoundkelpconservationandrecoveryplan.pdf>
- Norderhaug, K., Filbee-Dexter, K., Freitas, C., Birkely, S., Christensen, L., Møllerud, I., Thormar, J., Van Son, T., Moy, F., Vázquez Alonso, M., & Steen, H. (2020). Ecosystem-level effects of large-scale disturbance in kelp forests. *Marine Ecology Progress Series*, 656, 163–180. <https://doi.org/10.3354/meps13426>
- Pérez-Matus, A., Micheli, F., Konar, B., Shears, N., Low, N. H. N., Okamoto, D. K., Wernberg, T., Krumhansl, K. A., Ling, S. D., Kingsford, M., Navarrete-Fernandez, T., Ruz, C. S., & Byrnes, J. E. K. (2025). Kelp forests as nursery and foundational habitat for reef fishes. *Ecology*, 106(2), e70007. <https://doi.org/10.1002/ecy.70007>
- Premathilake, L., & Khangaonkar, T. (2022). Explicit quantification of residence and flushing times in the Salish Sea using a sub-basin scale shoreline resolving model. *Estuarine, Coastal and Shelf Science*, 276, 108022. <https://doi.org/10.1016/j.ecss.2022.108022>
- Questad, E. J., Kellner, J. R., Kinney, K., Cordell, S., Asner, G. P., Thaxton, J., Diep, J., Uowolo, A., Brooks, S., Inman-Narahari, N., Evans, S. A., & Tucker, B. (2014). Mapping habitat suitability for at-risk plant species and its implications for restoration and reintroduction.

- Ecological Applications: A Publication of the Ecological Society of America*, 24(2), 385–395. <https://doi.org/10.1890/13-0775.1>
- Samish Indian Nation. (2024, December 6). *A Decade of Disappearance: Bull Kelp in the San Juan Islands*. ArcGIS StoryMaps. <https://storymaps.arcgis.com/stories/b9f979a547004c32a616b5319a6410c0>
- Samish Indian Nation GIS. (2024). *2024 Digitized Kelp Canopy Polygons* [Dataset].
- Smale, D. A. (2020). Impacts of ocean warming on kelp forest ecosystems. *New Phytologist*, 225(4), 1447–1454. <https://doi.org/10.1111/nph.16107>
- Springer, Y., Hays, C., Carr, M., & Mackey, M. (2007). *Ecology and management of the bull kelp, Nereocystis luetkeana: A synthesis with recommendations for future research*. Lenfest Ocean Program, The Pew Charitable Trusts. https://www.lenfestocean.org/-/media/legacy/lenfest/pdfs/springer_underlying_report_0.pdf
- Steneck, R. S., Graham, M. H., Bourque, B. J., Corbett, D., Erlandson, J. M., Estes, J. A., & Tegner, M. J. (2002). Kelp forest ecosystems: Biodiversity, stability, resilience and future. *Environmental Conservation*, 29(4), 436–459. <https://doi.org/10.1017/S0376892902000322>
- Straub, S. C., Wernberg, T., Thomsen, M. S., Moore, P. J., Burrows, M. T., Harvey, B. P., & Smale, D. A. (2019). Resistance, Extinction, and Everything in Between – The Diverse Responses of Seaweeds to Marine Heatwaves. *Frontiers in Marine Science*, 6. <https://doi.org/10.3389/fmars.2019.00763>
- Teagle, H., Hawkins, S. J., Moore, P. J., & Smale, D. A. (2017). The role of kelp species as biogenic habitat formers in coastal marine ecosystems. *Journal of Experimental Marine Biology and Ecology*, 492, 81–98. <https://doi.org/10.1016/j.jembe.2017.01.017>
- Vadas, R. L. (1972). Ecological implications of culture studies on *Nereocystis luetkeana*. *Journal of Phycology*, 8(2), Article 2. <https://doi.org/10.1111/j.1529-8817.1972.tb04025.x>
- Washington State Department of Natural Resources. (2024). *ShoreZone Inventory* [Dataset]. <https://gis.dnr.wa.gov/portal/home/item.html?id=120c9827842b4c279475a54f30ac670f>
- Washington State Department of Natural Resources. (2025, October 6). *WA Bathymetry CoNED MLLW*. <https://geo.wa.gov/datasets/00dead70090640dd83d36d5880c97b7e/about>
- Weigel, B. L., Small, S. L., Berry, H. D., & Dethier, M. N. (2023). Effects of temperature and nutrients on microscopic stages of the bull kelp (*Nereocystis luetkeana*, Phaeophyceae). *Journal of Phycology*, 59(5), 893–907. <https://doi.org/10.1111/jpy.13366>
- Williams, M., Van Damme, C., & Flury, B. (2024). *Expanding a bull kelp habitat suitability model with ROV derived data* [Unpublished report]. Seattle Aquarium. <https://experience.arcgis.com/experience/17d4311a53454108a1196b90bd236547/page/M>
ap