

Methods note · Wildlife and ecology

Reading species models beyond sampled ocean conditions

Amina Okafor · 14 October 2026

Fictional author and institute · Editorial demonstration, not peer reviewed

ABSTRACT

A complete species-distribution map can conceal large gaps in environmental support. This methods note uses published environmental-novelty, transferability, and structured-validation research to distinguish interpolation from extrapolation in ocean applications. A hypothetical temperature-salinity example separates a new predictor range from a novel combination that lies within every individual range. Further decision cases examine structural disagreement beyond training conditions, clamping, observation bias, and validation splits matched to intended deployment. No real species model is fitted or range shift predicted, and terrestrial or freshwater case-study results are not imported as marine performance rankings. The practical output is a prediction accompanied by its evidence boundary, uncertainty, and sensitivity, making explicit where additional surveys would be more informative than an unqualified probability surface.

Keywords: species distribution · extrapolation · environmental novelty · model transferability · cross-validation

A source-based editorial text, not a report of original Still North research. Examples and imagery are illustrative.

A species-distribution model can fill an unsampled ocean sector with precise-looking probabilities. The map's visual completeness does not establish that the species has been observed there, that its response to the new environment is known, or that the model has been validated under those conditions. The most useful companion to a prediction map is often a map of where its evidence stops.

This methods note asks how to recognise and report extrapolation when an ecological model is projected beyond sampled ocean conditions. It draws on published methods for environmental novelty, transferability, and structured validation. All species, predictor ranges, and model outputs used in the worked comparison are hypothetical. The note does not fit a model, predict a real range shift, or claim that an environmental association identifies a physiological limit.

What exactly is being transferred?

A correlative distribution model learns a relationship between observations of occurrence or abundance and selected predictors. Its target may be presence probability, expected count, encounter rate, or relative suitability, depending on the data and model. These outputs are not interchangeable. Presence-only and background data, for example, require assumptions that differ from a designed detection-and-non-detection survey. Identify the response and observation process before interpreting any number between zero and one as occupancy probability.

Projection can transfer across geography, season, year, depth, or environmental conditions. A nearby location may be environmentally novel, while a distant location may resemble the training data in the selected predictors. Neither resemblance nor geographical proximity guarantees transferability if unmeasured processes differ. Prey, competitors, access, movement constraints, and observa-

tion methods can change while the plotted temperature and ice values remain familiar.

A worked distinction between two kinds of novelty

Imagine a fictional shelf-species model trained where temperatures range from 0°C to 4°C and salinities from 30 to 34 on the Practical Salinity Scale. Suppose the sampled environment strongly couples cold water with fresher conditions and warmer water with saltier conditions. A projection cell at 6°C and salinity 32 exceeds the sampled temperature range. This is straightforward univariate novelty, visible from the training minima and maxima.

A second cell at 3°C and salinity 30.2 is inside each variable's separate range, yet may lie far from the combinations actually sampled. A check that asks only whether every predictor falls between its minimum and maximum would miss it. Mesgaran, Cousens, and Webber developed an environmental-novelty method that distinguishes new ranges from new combinations by accounting for covariate relationships. Their case study is terrestrial; the statistical distinction applies equally to an explicitly defined marine predictor space.

The labels depend on the actual joint distribution, not just the endpoints stated here. A covariance-based distance can summarise novelty, but its assumptions and scaling matter, and a single ellipse does not describe every complex occupied region. The novelty calculation identifies a problem in the input space. It does not quantify the true prediction error or prove that the species cannot use a novel combination.

Hypothetical projection cell	Temperature	SP	Evidence question
Familiar sampled combination	1°C	31	Interpolation, subject to coverage
Beyond one predictor range	6°C	32	New temperature response
Novel joint combination	3°C	30.2	Unobserved temperature-salinity pairing

Model agreement inside the data may hide disagreement outside

Consider two hypothetical fitted models that predict similar occurrence within the sampled environment. One extends a positive temperature relationship indefinitely; another imposes a smooth peak and decline beyond the warmest observations. They could return 0.8 and 0.2 at the same unsampled warm cell while performing similarly during ordinary interpolation. The training data cannot directly adjudicate the unobserved portion of those curves. A tight interval from one fitted model may omit this structural uncertainty.

Clamping predictors to their training range avoids some extreme numerical behaviour but does not create observations. A clamped 6°C cell may receive the prediction associated with 4°C, despite the species' response above 4°C being unknown. If clamping is used, map it and explain its effect. Do not present the visually smooth boundary as evidence of biological resilience or a demonstrated habitat threshold. Refusing a prediction in unsupported conditions can be more informative than a complete colour surface.

Validation should resemble the intended journey

Randomly splitting nearby observations can place almost identical environmental and spatial conditions in both training and test sets. The resulting score may describe interpolation among familiar samples rather than transfer to a new sea, year, or season. Wenger and Olden show how strong within-dataset performance can coexist with weak transferability, and propose testing on deliberately distinct groups. Their fish examples concern freshwater systems, so their model rankings are not claims about the best Arctic marine algorithm.

Roberts and colleagues examine cross-validation for structured ecological data and show why spatial, temporal, hierarchical, and other dependencies affect error estimates. They also identify a necessary caution: blocking can itself create extrapolation by changing the predictor ranges available for training. That may be appropriate when extrapolation is the intended task, but it can overstate error if the goal is ordinary interpolation. Validation is a model of deployment, not a ritual that becomes correct merely by using larger blocks.

For an intended future-season projection, an appropriate exercise might withhold entire years or environmentally distinct periods, then report performance separately in familiar and novel conditions. For projection into a new sector, hold out meaningful spatial units rather than individual neighbouring records. The design must still leave enough independent groups to assess variability. No clever split can manufacture a test environment that has never been observed anywhere in the record.

Separate occurrence from observation

An apparent distribution boundary may reflect the survey. Ships avoid difficult ice, tags sample selected individuals, and acoustic instruments detect only audible and recognisable signals within their effective range. If those observation conditions covary with temperature or ice concentration, a model may learn survey access or detectability as part of the species relationship. Transferring it to a new recording system or sampling season can then change predictions without any biological range shift.

Document effort, non-detection meaning, gear, detector settings, and sampled demographic groups. Where supported, model the observation process or restrict comparisons to compatible data. Adding every available predictor does not automatically solve confounding; some may encode effort indirectly. Biological knowledge should help specify plausible mechanisms and transfer limits, but should not be used to erase contradictory data or declare an unsupported response curve correct.

Use the projection as a decision aid, not an observation

Suppose a hypothetical protected-area assessment encounters a high predicted value in a novel warm-fresh combination. Three defensible outputs are the conditional prediction, the novelty flag, and the sensitivity across plausible model responses. Together they can prioritise additional surveys or identify a decision vulnerable to missing information. Reporting only the high value would imply more support than exists; reporting only uncertainty would discard potentially useful conditional information.

Novelty thresholds should be chosen transparently, not tuned until inconvenient cells disappear. Preserve training-data coverage, predictor definitions, spatial and temporal resolution, model version, extrapolation settings, and validation design. A scenario-driven future projection must also identify the environmental scenario rather than presenting it as a forecast with known probability. Distinguish uncertainty in future conditions from uncertainty in the species response to those conditions.

Limitations and the meaningful map

This note addresses correlative projection, not a complete mechanistic habitat model. Environmental similarity cannot establish dispersal, prey access, population persistence, or reproductive success. The hypothetical cases contain no sampling error calculation, and the cited validation methods do not guarantee transfer to unprecedented oceans. Their contribution is to make different kinds of unsupported inference visible before a map is used.

A useful model map can therefore contain a deliberate boundary: here the data support interpolation; here the model extrapolates; here no defensible probability is supplied. That boundary is not a failure of ambition. It is

the information needed to decide where a prediction can guide action, where caution must dominate, and which next observation could most improve the question.

References

1. Mesgaran, M. B., Cousens, R. D., and Webber, B. L. (2014). Here be dragons: a tool for quantifying novelty due to covariate range and correlation change when projecting species distribution models. *Diversity and Distributions*, 20, 1147–1159.
2. Wenger, S. J., and Olden, J. D. (2012). Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods in Ecology and Evolution*, 3, 260–267.
3. Roberts, D. R., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40, 913–929.