Michael Mahoney [1]
mike.mahoney.218@gmail.com
https://mm218.dev

[1] Graduate Program in Environmental Science, SUNY College of Environmental Science and Forestry

Assessing spatial models is hard

Assessing spatial models can be difficult. Model errors may exhibit spatial autocorrelation, model predictions are often aggregated to multiple spatial scales by users, and models are often used to extrapolate outside the boundaries of their training data. To adjust for these considerations, modelers must choose from a dizzying array of assessment protocols and metrics which may give different or even contradictory results. Researchers are also often responsible for implementing these techniques themselves, or otherwise welding together multiple packages with incompatible interfaces in order to properly assess their models.

New tools for R's tidymodels modeling framework aim to reduce this complexity, implementing common model assessment tasks in a straightforward, computationally efficient, and easy-to-learn manner.

spatialsample

The `spatialsample` package (Mahoney et al. 2023; Mahoney and Silge 2023) provides methods for spatial cross-validation. These methods provide more accurate estimates of model performance when extrapolating into new areas, and are particularly useful when there is no independent probability sample available for assessing predictions.

Functions in `spatialsample` provide a standardized interface to several of the most popular spatial CV approaches, and return objects which extend classes from the `rsample` (Frick et al. 2022) package. As a result, these objects can seamlessly be used with any functions within the tidymodels modeling framework.

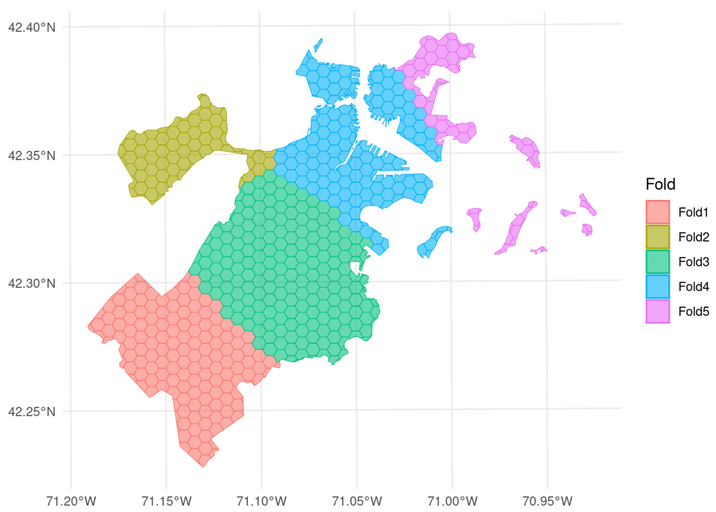

Figure 1: Spatial cross-validation methods, such as this instance of spatially clustered CV, assign observations to different folds of k-fold CV based on their geographic position.

Assessing spatial models in scientifically sound ways **is hard. New tools** for the tidymodels modeling framework can help **make it easier.**

waywiser

The `waywiser` package (Mahoney 2023a, 2023b) is an ergonomic toolbox providing a consistent user interface for assessing spatial models. To that end, `waywiser` makes it easy to compute model assessment metrics, evaluate predictions aggregated to multiple scales, and identify areas where your model can't safely make predictions. The multi-scale assessment protocol follows Riemann et al. (2010), and is designed to help modelers identify minimum mapping units and regions where models underperform, as well as to estimate how well models perform at relevant levels of aggregation.

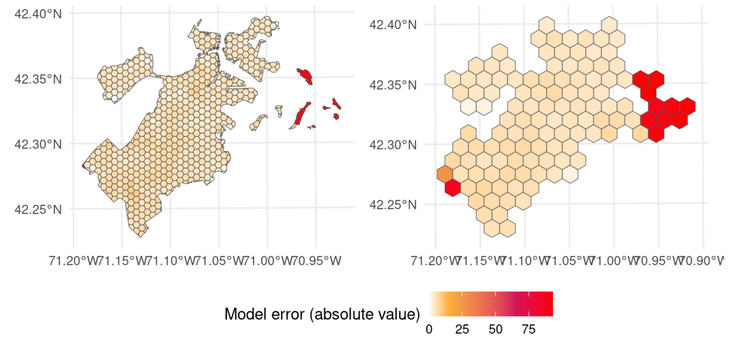

Figure 2: Assessing model predictions at multiple spatial scales can help you identify where your model underperforms and how well your model performs at the scales you and your users care about.

Additional functions implement the "area of applicability" approach from Meyer and Pebesma (2021). This method helps modelers estimate whether or not a novel observation falls within the feature space a model was trained on, and accordingly whether or not the model can be reliably used to generate predictions for that observation. These approaches are particularly useful for spatial models when extrapolating into "unknown" space.

As with `spatialsample`, `waywiser` is built to extend existing tidymodels packages, namely `yardstick` (Kuhn, Vaughan, and Hvitfeldt 2023) and `applicable` (Gotti and Kuhn 2022). Functions from `waywiser` share a standardized interface and are size- and type-stable in their outputs. Metric functions can be used as a drop-in replacement for `yardstick` functions, AOA objects behave identically to objects from `applicable`, and results are returned as either `tibble` (Müller and Wickham 2023) or `sf` (Pebesma 2018) objects as appropriate.

References

Frick, Hannah, Fanny Chow, Max Kuhn, Michael J Mahoney, Julia Silge, and Hadley Wickham. 2022. *Rsample: General Resampling Infrastructure*. https://CRAN.R-project.org/package=rsample.
Gotti, Marly, and Max Kuhn. 2022. *Applicable: A Compilation of Applicability Domain Methods*. https://CRAN.R-project.org/package=applicable.
Kuhn, Max, Davis Vaughan, and Emil Hvitfeldt. 2023. *Yardstick: Tidy Characterizations of Model Performance*. https://CRAN.R-project.org/package=yardstick.
Mahoney, Michael J. 2023b. *waywiser: Ergonomic Methods for Assessing Spatial Models*. https://CRAN.R-project.org/package=waywiser.
———. 2023a. "waywiser: Ergonomic Methods for Assessing Spatial Models." https://doi.org/10.48550/arXiv.2303.11312.
Mahoney, Michael J, Lucas K Johnson, Julia Silge, Hannah Frick, Max Kuhn, and Colin M Beier. 2023. "Assessing the Performance of Spatial Cross-Validation Approaches for Models of Spatially Structured Data." https://doi.org/10.48550/arXiv.2303.07334.
Mahoney, Michael J, and Julia Silge. 2023. *spatialsample: Spatial Resampling Infrastructure*. https://CRAN.R-project.org/package=spatialsample.
Meyer, Hanna, and Edzer Pebesma. 2021. "Predicting into Unknown Space? Estimating the Area of Applicability of Spatial Prediction Models." *Methods in Ecology and Evolution* 12 (9): 1620–33. https://doi.org/https://doi.org/10.1111/2041-210X.13650.
Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. https://CRAN.R-project.org/package=tibble.
Pebesma, Edzer. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10 (1): 439–46. https://doi.org/10.32614/RJ-2018-009.
Riemann, Rachel, Barry Tyler Wilson, Andrew Lister, and Sarah Parks. 2010. "An Effective Assessment Protocol for Continuous Geospatial Datasets of Forest Characteristics Using USFS Forest Inventory and Analysis (FIA) Data." *Remote Sensing of Environment* 114 (10): 2337–52. https://doi.org/https://doi.org/10.1016/j.rse.2010.05.010.